

# **Regulation of UDP- glucuronosyltransferase 1A genes in the intestine**

by

**Siti Nurul Mubarakah**

*Thesis  
Submitted to Flinders University  
for the degree of*

**Doctor of Philosophy**

College of Medicine and Public Health  
18 October 2018

---

# TABLE OF CONTENTS

List of Figures .....	vi
List of Tables .....	x
Summary .....	xi
Declaration .....	xiii
Acknowledgements.....	xiv
Publication .....	xvi
Conference Proceedings.....	xvi
Scholarships and Awards in Support of This Thesis.....	xviii
Abbreviations .....	xix
CHAPTER 1. INTRODUCTION.....	1
1.1 Glucuronidation and human UGTs.....	1
1.1.1 Glucuronidation .....	1
1.1.2 Human UDP-Glucuronosyltransferase (UGT) enzymes .....	3
1.1.3 The UGT gene super family.....	5
1.2 Expression of UGT genes: general aspects .....	10
1.2.1 Tissue specific distribution of human UGTs.....	10
1.2.2 Constitutive and inducible regulation of UGTs .....	12
1.3 Intestinal UGT enzymes.....	15
1.3.1 Intestinal UGTs in phenolic compound glucuronidation .....	15
1.3.2 Intestinal microbiota and UGT activities.....	18
1.3.3 Expression of intestinal UGTs .....	19
1.4 Intestinal development and homeostasis .....	20
1.4.1 Intestinal development is critically regulated by homeobox factors including the master regulator CDX2.....	20
1.4.2 The intestine is maintained by continual epithelial turnover that is controlled by developmental regulators. ....	22
1.4.3 Constitutive regulation of intestinal UGTs by CDX2 and its partners.....	27
1.5 <i>In vitro</i> models for intestinal UGT studies.....	31
1.5.1 Caco-2 as a model for enterocytes .....	31
1.5.2 The intestinal organoid model.....	32
1.5.3 'Humanised' UGT transgenic mouse models.....	34
1.6 UGT1A8 is the prototypical intestinal-specific UGT gene .....	36

1.6.1	UGT1A8 expression, activities and substrates.....	36
1.6.2	UGT1A8 polymorphisms in drug metabolism and cancer risk.....	38
1.6.3	UGT1A8 gene regulation.....	40
1.7	Project overview and experimental aims.....	44
CHAPTER 2. MATERIALS AND GENERAL METHODS.....		46
2.1	Materials.....	46
2.1.1	Chemicals, reagents and buffers.....	46
2.1.2	Mammalian cell lines.....	46
2.1.3	Expression vectors and reporters.....	46
2.1.4	Oligonucleotides.....	47
2.2	Methods.....	47
2.2.1	Mammalian cell culture, maintenance and stocks.....	47
2.2.2	Bacterial culture and stocks.....	47
2.2.3	Cryopreservation of cells (freezing stocks).....	47
2.2.4	Polymerase chain reaction (PCR) amplification.....	48
2.2.5	Total RNA extraction.....	49
2.2.6	cDNA synthesis.....	50
2.2.7	Quantification of nucleic acid (DNA or RNA).....	50
2.2.8	Molecular cloning using restriction enzymes.....	51
2.2.9	Competent cell preparation.....	51
2.2.10	Plasmid DNA extraction.....	52
2.2.11	Agarose gel electrophoresis.....	52
2.2.12	DNA Sequencing.....	53
2.2.13	Plasmid-DNA transfection.....	53
2.2.14	Luciferase reporter assays.....	54
2.2.15	Chromatin Immunoprecipitation (ChIP) assays.....	55
2.2.16	Electrophoretic mobility shift assay (EMSA).....	56
2.2.17	Western blotting.....	57
2.2.18	Cell imaging (EVOS® FL).....	58
2.2.19	Statistics.....	58
CHAPTER 3. REGULATION OF INTESTINAL UGTs BY A COMBINATORIAL MECHANISM OF CDX2 AND HNF4 $\alpha$ .....		59
3.1	Introduction.....	59
3.2	Methods.....	63
3.2.1	Plasmids.....	63

3.2.2	Mutagenesis of binding sites in <i>UGT1A8</i> , -1A9 and -1A10 promoter constructs .....	63
3.2.3	Transfections of CDX2 and HNF4 $\alpha$ expression vectors.....	64
3.2.4	Chromatin Immunoprecipitation (ChIP) assays in Caco-2 cells.....	64
3.2.5	Non-radioactive EMSA assays in Caco-2 cells .....	64
3.2.6	CDX2 and HNF4 $\alpha$ siRNA design and transfection .....	64
3.2.7	Transcriptomic data profiling of Colon Adenocarcinoma (COAD) from The Cancer Genome Atlas (TCGA) .....	65
3.2.8	Oligonucleotides.....	65
3.3	Results and Discussion.....	67
3.3.1	CDX2 synergistically interacts with HNF4 $\alpha$ to further induce <i>UGT1A8</i> promoter activity .....	67
3.3.2	Co-regulation of CDX2/HNF4 $\alpha$ is mediated through a novel synergistic composite element .....	69
3.3.3	The CDX2/HNF4 $\alpha$ composite element is conserved in the <i>UGT1A8</i> , -1A9 and -1A10 promoters .....	77
3.3.4	Dominant negative HNF4 $\alpha$ does not inhibit the ability of wild-type HNF4 $\alpha$ to synergise with CDX2 in regulation of the <i>UGT1A8</i> and -1A9 promoters .....	81
3.3.5	The intestinal homeobox factor Barx2 does not synergise with HNF4 $\alpha$ .....	83
3.3.6	Regulation of endogenous <i>UGT1A8-1A10</i> by CDX2 and HNF4 $\alpha$ in Caco-2 cells .....	84
3.3.7	A regulatory feedback loop involving CDX2 and HNF4 $\alpha$ is confirmed at the mRNA level.....	86
3.3.8	Differential CDX2 and HNF4 $\alpha$ functions in the regulation of <i>UGT1A8</i> and -1A9 in hepatic vs intestinal contexts.....	88
3.3.9	Data analysis from TCGA database supports positive correlation of <i>UGT1A8</i> and -1A10 expression with CDX2 and HNF4 $\alpha$ levels in human colon samples.....	91
3.4	Conclusions .....	94

CHAPTER 4.	<i>UGT1A8</i> GENE REGULATION IN RESPONSE TO FLAVONOIDS, BUTYRATE INVOLVEMENT AND MECHANISM ANALYSIS .....	99
4.1	Introduction.....	99
4.2	Methods.....	102
4.2.1	Caco-2 cells and culture .....	102
4.2.2	Chemical treatment of Caco-2 cells .....	103
4.2.3	Luciferase assays.....	103
4.2.4	Bradford assay for the determination of protein concentration .....	104
4.2.5	Plasmids.....	104

4.2.6	Mutagenesis of PPAR response element (PPAR-RE) in the <i>UGT1A8</i> promoter .....	106
4.2.7	Fluorescence activated cell sorting (FACS) of Vp 16-PPAR $\alpha$ / GFP co-transfection	106
4.2.8	Cell Electroporation .....	107
4.2.9	Lentiviral packaging and transduction.....	107
4.2.10	Oligonucleotides.....	108
4.3	Results and Discussion.....	110
4.3.1	Screening of <i>UGT1A8</i> promoter response to flavonoids identifies genistein as a potent inducer of promoter activity in NaB-induced differentiated Caco-2 cells .....	110
4.3.2	Genistein and butyrate synergistically induce <i>UGT1A8</i> promoter activity and <i>UGT1A8</i> mRNA expression .....	114
4.3.3	Mechanistic analysis of <i>UGT1A8</i> regulation by genistein.....	116
4.3.4	Analysis of <i>UGT1A8-1A10</i> gene regulation by PPAR $\gamma$ .....	122
4.3.5	Potential role of PPAR $\alpha$ in regulating <i>UGT1A8-1A10</i> gene expression .....	126
4.3.6	PPAR $\alpha$ and PPAR $\gamma$ attenuate HNF4 $\alpha$ -mediated <i>UGT1A9</i> activation .....	129
4.3.7	The putative DR1 PPAR response element in the <i>UGT1A8</i> promoter is important for transcriptional activation .....	131
4.3.8	Neither NF-YA, nor C/EBP mediates genistein action in <i>UGT1A8</i> regulation.....	132
4.3.9	Possible roles for CDX2 and SMAD2 in <i>UGT1A8</i> regulation.....	134
4.3.10	Genistein action involves altered histone modifications in the <i>UGT1A8</i> proximal promoter region .....	136
4.3.11	Comparison of gene expression in human colonic cancer cell lines.....	138
4.3.12	Optimizing Caco-2 transfection methods .....	143
4.4	Conclusions .....	145
CHAPTER 5. EXPLORATION OF ENTEROID AND ORGANOID CULTURE FROM <i>UGT1A8</i> /LUCIFERASE-TRANSGENIC MICE .....		151
5.1	Introduction.....	151
5.1.1	Intestinal stem cells and identification of markers .....	153
5.1.2	Organoid culture of intestinal epithelium .....	154
5.1.3	Aims .....	157
5.2	Methods.....	157
5.2.1	Development and maintenance of the <i>UGT1A8</i> /Luc transgenic mice .....	157
5.2.2	Preparation of Matrigel, culture plates, and media .....	158
5.2.3	Isolation of mouse intestinal crypts .....	159
5.2.4	Enteroid/organoid culture from isolated crypts .....	160
5.2.5	Passaging of the organoid culture .....	160
5.2.6	Freezing organoids.....	161

5.2.7	Enteroid/organoid transfection.....	161
5.2.8	Organoid transduction.....	162
5.2.9	Mouse tissue and enteroid/organoid RNA extraction, reverse transcription and qRT-PCR.....	163
5.2.10	Whole mount immunofluorescence (IF) of organoids.....	164
5.2.11	Oligonucleotides.....	165
5.3	Results and discussion.....	165
5.3.1	Generation of enteroids and colonoids from isolated crypts.....	165
5.3.2	Observation of structures during organoid culture.....	167
5.3.3	Organoid formation in the presence or absence of intestinal fibroblasts/ISEMFs....	169
5.3.4	Exploration of gene delivery methods in <i>UGT1A8</i> /Luc-mouse intestinal enteroids..	171
5.3.5	Immunofluorescent staining of organoid.....	174
5.3.6	Challenges and opportunities in organoid culture.....	175
5.3.7	Detection of <i>UGT1A8</i> /luciferase expression in <i>UGT1A8</i> /Luc tissues and enteroids.	178
5.3.8	Expression of mouse- <i>Cdx2</i> , <i>Hnf4<math>\alpha</math></i> , <i>Ppar<math>\gamma</math></i> , and <i>Ugt1a7c</i> in the <i>UGT1A8</i> /Luc enteroids.....	179
5.4	Conclusion.....	180
CHAPTER 6. GENERAL DISCUSSION AND CONCLUSION.....		183
APPENDICES.....		191
	Appendix 1.....	191
	Appendix 2.....	196
	Appendix 3.....	198
REFERENCES.....		214

## LIST OF FIGURES

Figure 1.1.	The glucuronidation mechanism. ....	3
Figure 1.2.	Model of UGT enzyme topology within the endoplasmic reticulum. ....	5
Figure 1.3.	The phylogenetic tree of UGT1, UGT2, UGT3 and UGT8. ....	6
Figure 1.4.	The human UGT1 family complex locus. ....	7
Figure 1.5.	Illustration of trophoctoderm developmental fate, as regulated by CDX2. ....	21
Figure 1.6.	Intestinal epithelial self-renewal. ....	23
Figure 1.7.	Graphical overview of major transcription factors involved in intestinal epithelial regulatory networks. ....	25
Figure 1.8.	Detection of UGT1A protein in the human intestine. ....	28
Figure 1.9.	Immunofluorescence staining showing UGT localization in mouse blastocyst. ....	29
Figure 1.10.	Illustration of enteroid culture. ....	34
Figure 1.11.	Representative images of human <i>UGT1A8</i> promoter activity identification in transgenic mouse by bioluminescence assay. ....	35
Figure 1.12.	Raloxifene glucuronidation by UGT enzymes. ....	39
Figure 1.13.	Identified transcription factors (TFs) and their cognate cis-regulatory elements (CREs) in the proximal promoter of <i>UGT1A8</i> . ....	42
Figure 1.14.	Schematic of the project. ....	45
Figure 3.1.	Model of intestinal regulatory networking in differentiated IECs involving CDX2, HNF1 $\alpha$ and HNF4 $\alpha$ . ....	60
Figure 3.2.	Phylogenetic analysis of ~1k-proximal promoter regions of <i>UGT1A genes</i> . ....	61
Figure 3.3.	Synergistic activation of the <i>UGT1A8</i> promoter by CDX2 and HNF4 $\alpha$ . ....	68
Figure 3.4.	The HNF4 $\alpha$ RE at -44bp in the <i>UGT1A8</i> proximal promoter is required for synergistic induction by CDX2 and HNF4 $\alpha$ . ....	69
Figure 3.5.	Synergistic regulation of the <i>UGT1A8</i> promoter is independent of the consensus -70CDX2RE and the Sp1/Inr element. ....	71
Figure 3.6.	EMSA analysis of HNF4 $\alpha$ and CDX2 binding to the novel composite element of the <i>UGT1A8</i> promoter region. ....	73
Figure 3.7.	EMSA mutational analysis of CDX2 and HNF4 $\alpha$ binding to the synergistic composite element designated at -44HNF4 $\alpha$ RE. ....	75
Figure 3.8.	The CDX2 and HNF4 $\alpha$ motifs in the composite element are both required for synergistic promoter activation. ....	77
Figure 3.9.	Alignment of the human UGT1A proximal promoters. ....	78
Figure 3.10.	The CDX2/HNF4 $\alpha$ composite element is functionally conserved in the <i>UGT1A8-1A10</i> promoters. ....	80

Figure 3.11. The CDX2/HNF4 $\alpha$ composite element does not mediate dominant negative (DN) HNF4 $\alpha$ action. ....	82
Figure 3.12. The combination of Barx2 and HNF4 $\alpha$ does not induce the UGT1A8 promoter. ....	84
Figure 3.13. Inhibition of endogenous UGT1A8-1A10 expression by inhibition of HNF4 $\alpha$ or CDX2 expression or function. ....	85
Figure 3.14. CDX2 and HNF4 $\alpha$ regulate each other's expression in Caco-2 and HT-29 colon cancer cell lines. ....	87
Figure 3.15. UGT1A8 and -1A9 are differently regulated by CDX2 and HNF4 $\alpha$ in hepatic and intestinal cell models. ....	89
Figure 3.16. HNF4 $\alpha$ activation via predicted and functional motifs within 1kb of the <i>UGT1A9</i> promoter region in the intestinal Caco-2 cell line. ....	91
Figure 3.17. Correlation analysis of UGT1A8-1A10, CDX2 and HNF4 $\alpha$ levels in normal colon samples (n=41) from the Colon Adenocarcinoma (COAD) dataset of The Cancer Genome Atlas (TCGA). ....	92
Figure 3.18. Analysis of UGT1A8-1A10, CDX2 and HNF4 $\alpha$ levels in colon cancer samples (n=480) from the TCGA dataset. ....	94
Figure 3.19. A model showing differential regulation of specific intestinal UGT1A8 and hepatic/intestinal UGT1A9 by CDX2 and HNF4 $\alpha$ . ....	96
Figure 4.1. Screening for activation of <i>UGT1A8</i> promoter by flavonoid and steroid in differentiated Caco-2 cells. ....	111
Figure 4.2. Differentiation markers increase following sodium butyrate-induced differentiation. ....	113
Figure 4.3. Identification of genistein as an inducer of <i>UGT1A8</i> promoter activity and gene expression. ....	115
Figure 4.4. Schematic model for the synergistic regulation of the <i>UGT1A8</i> promoter by NaB and genistein. ....	116
Figure 4.5. Potential mechanism of UGT1A8 mRNA induction by genistein via PPAR $\gamma$ and ER $\beta$ . ....	118
Figure 4.6. Mechanistic analysis of UGT1A8 regulation by genistein at the promoter level. ....	119
Figure 4.7. A putative PPAR binding site within <i>UGT1A8</i> , -1A9 and -1A10 1kb promoters. ....	120
Figure 4.8. Genistein activates the -1kb <i>UGT1A8</i> , -1A9 and -1A10 promoters, likely via PPAR $\gamma$ . ....	121
Figure 4.9. Analysis of UGT1A9 and -1A10 mRNA induction by genistein. ....	122
Figure 4.10. PPAR $\gamma$ -RXR $\alpha$ overexpression enhances the ability of a PPAR $\gamma$ agonist to induce UGT1A8 mRNA expression in Caco-2 cells. ....	123
Figure 4.11. PPAR $\gamma$ induces <i>UGT1A8</i> promoter activity via a PPRE at location -818 to -805nt in COS-7 cells. ....	124

Figure 4.12. PPAR $\gamma$ may require the co-activator PGC-1 to induce UGT1A8 promoter activity in Caco-2 cells.....	125
Figure 4.13. VP16-PPAR $\alpha$ induces UGT1A8, UGT1A9, UGT1A10 and L-FABP mRNA expression. ....	127
Figure 4.14. <i>UGT1A8</i> and <i>-1A9</i> -1kb promoters do not respond to ligand-activated PPAR $\alpha$ in Caco-2 cells.....	128
Figure 4.15. The presence of PPAR $\alpha$ or $\gamma$ repress HNF4 $\alpha$ -mediated <i>UGT1A9</i> promoter activation. ....	130
Figure 4.16. A DR1 PPAR response element is important for <i>UGT1A8</i> basal promoter activity..	132
Figure 4.17. NF-YA does not mediate induction of the <i>UGT1A8</i> promoter by genistein.....	133
Figure 4.18. C/EBP $\alpha$ over-expression in Caco-2 cells does not induce <i>UGT1A8</i> promoter activation. ....	134
Figure 4.19. UGT1A8 promoter response to genistein treatment in combination with SMAD2 and CDX2 overexpression.....	136
Figure 4.20. Genistein induces H3K4me3 and H3Kac enrichment at the <i>UGT1A8</i> promoter.....	137
Figure 4.21. Growth and morphological variation of colon cancer cell lines. ....	139
Figure 4.22. Basal mRNA expression of UGT1A8, -1A9, -1A10, CDX2 and HNF4 $\alpha$ in human colonic cell lines. ....	142
Figure 4.23. Optimization of transfection methods for Caco-2 cells. ....	145
Figure 4.24. Schematic representation of the mode of action of genistein and butyrate in regulating UGT1A8. ....	149
Figure 5.1. Illustration of a Matrigel dome of enteroid/organoid culture, in a single well of 48 well-plate. ....	160
Figure 5.2. Successful enteroid formation from crypts of <i>UGT1A8</i> /Luc-mice. ....	166
Figure 5.3. Representative images of structures observed during <i>UGT1A8</i> /Luc-mice organoid expansion.....	168
Figure 5.4. Comparison of organoid culture in the presence and absence of fibroblasts/ISEMFs. ....	170
Figure 5.5. Organoid growth in the presence of fibroblasts/ISEMFs cannot be maintained without exogenous growth factors.....	171
Figure 5.6. GFP transfection into enteroids by lipofection and lentiviral transduction.....	172
Figure 5.7. Transfection of fluorescent labelled, fluorescein amidite (FAM) siRNA into organoids. ....	174
Figure 5.8. Immunofluorescence of organoids and intestinal fibroblast.....	175
Figure 5.9. Uneven organoid dissociation into single cells results in unequal organoid growth.	176
Figure 5.10. Comparison of two stages of organoid viability. ....	177

Figure 5.11. Persistent enterosphere structure after 4 weeks culture in growth factor deficient media. ....	177
Figure 5.12. Expression of <i>UGT1A8</i> /Luc reporter and <i>Ugt8</i> mRNA in tissues and enteroids from <i>UGT1A8</i> /Luc-mice. ....	179
Figure 5.13. Expression of mouse mPPAR $\gamma$ , mHNF4 $\alpha$ , mCdx2 and mouse <i>Ugt1a7c</i> in <i>UGT1A8</i> /Luc organoids. ....	180

## LIST OF TABLES

Table 1.1. mRNA expression of UGT1As and 2Bs in human tissues.....	12
Table 1.2. Nuclear receptors, ligands, and their target UGTs.....	14
Table 1.3. Phenolic compounds metabolised by intestinal UGTs. ....	17
Table 3.1. List of oligonucleotides used in studies described in Chapter 3. ....	65
Table 4.1. Oligonucleotides used in studies described in Chapter 4.....	108
Table 5.1. Growth factors and reagents for enteroid culture.....	155
Table 5.2. Oligonucleotides used for qRT-PCR analysis of organoids and mouse tissue.....	165

## SUMMARY

UDP-glucuronosyltransferases are a superfamily of enzymes involved in Phase II metabolism of small lipophilic chemicals; by conjugating these chemicals with sugars, UGTs render them more water soluble and readily eliminated. Human intestine is constantly exposed to lipophilic chemicals ingested as part of the diet, or as supplements or drugs. These chemicals can have positive bioactive effects or can potentially be toxic or carcinogenic. The activity of intestinal UGT enzymes is crucial as the first line of defence rendering these substances more water soluble, thus, facilitating their inactivation and excretion. The constitutive expression of different *UGT* genes is regulated in a defined tissue-specific manner. In addition, their expression within specific tissues can be induced by small molecules that include UGT substrates, allowing a feedback response. A cluster of *UGT* genes, *UGT1A8*, *-1A9*, and *-1A10*, are highly expressed in the intestine. However; while *UGT1A8* and *UGT1A10* are exclusively extrahepatic, *UGT1A9* expressed in both liver and intestine. With the evidence linking the risk of colorectal cancer with the level of intestinal UGT activity, it is of particular importance to determine how the spatiotemporal regulation of intestinal UGTs is mediated, both in terms of constitutive and inducible expression. CDX2 is an intestinal master transcription factor which works in partnership with HNF4 $\alpha$  to control gene expression during intestinal development and intestinal epithelial renewal. In this study, we provided a defined novel mechanism by which CDX2 and HNF4 $\alpha$  control *UGT1A8*, *-1A9* and *-1A10* expression at the promoter level. Using a variety of molecular techniques, we showed that CDX2 and HNF4 $\alpha$  synergistically induce *UGT1A8-1A10* intestinal expression via a conserved composite element of 12 nt located at the proximal promoter of the three genes. Moreover, this work identified the first known functional CDX2 binding motif in *UGT1A9* helping to explain its intestinal expression. We also examined how HNF4 $\alpha$  controls hepatic expression of *UGT1A9* leading to a model in which HNF4 $\alpha$  acts via separate intestinal and hepatic regulatory modules in this gene. Overall our study showed that the CDX2 /HNF4 $\alpha$  nexus that is critical in developmental patterning and maintenance of intestine also defines *UGT1A8*, *-1A9* and *-1A10* expression.

Inducible regulation of UGTs in the intestine involves dietary constituents and products of

microbiota activity; reports show that the activation of UGTs by chemicals that are also UGT substrates constitutes a feedback-regulatory mechanism. Using Caco-2 cells carrying an integrated *UGT1A8* promoter-reporter construct, we screened for chemicals that could induce promoter activity. We identified genistein as a flavonoid that most potently induced the *UGT1A8* promoter. The effect of genistein on *UGT1A8* (as well as *-1A9* and *-1A10*) expression was enhanced synergistically by butyrate. Butyrate is a fermentation product of gut microbiota and is well known as HDAC inhibitor. Via inducing chromatin remodelling, butyrate can promote access of ligand-induced transcription factors to target genes. We assessed whether genistein might function as ligand for various ligand-dependent transcription factors including PPAR. Our findings using antagonist assays supported involvement of PPAR $\gamma$  at the promoter and mRNA level. It is also possible that PPAR $\gamma$  activity is regulated post-transcriptionally via genistein; moreover, we found evidence that genistein can alter chromatin accessibility at the *UGT1A8* promoter. These studies define at least one pathway for inducible regulation of intestinal UGTs by flavonoids and butyrate; given that *UGT1A8* conjugates many carcinogenic compounds, this induction may be involved in the protective effects of flavonoid and fibre rich diets on cancer risk.

To allow a better understanding of physiological aspects of constitutive and ligand-activated regulation of intestinal *UGT* genes, we established an intestinal organoid culture system using *UGT1A8* promoter-reporter transgenic mice. This model more closely represents the heterogeneity and structural features of normal intestine than cell line models but maintains benefits of cell lines such as amenability to genetic manipulation. In Matrigel based culture containing appropriate growth factors, spontaneous growth of enteroids or organoids from isolated intestinal crypts were observed, and these could be maintained throughout several passages as well as frozen for long-term storage. Although requiring further optimization, we found that genetic manipulation in organoids is possible using DNA and RNA transfection. We also showed that expression of *UGT1A8* and its main developmental and inducible regulators (CDX2, HNF4 $\alpha$ , PPAR) in intestinal organoids is comparable to that in intact adult intestine. This preliminary exploration provides new scope for studies using humanized mice and organoid technology.

## DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.



Siti Nurul Mubarokah

## ACKNOWLEDGEMENTS

In the name of God, the Most Gracious and the Most Merciful. I have placed my trust in You, there is no might and no power except by You. My praises all to You for Your countless blessings and strengths throughout my life.

First and foremost, I would like to express my sincere gratitude to my supervisor Dr. Robyn Meech for her continuous support and guidance, motivation and immense knowledge. I always see you as a role model and inspiration of a scientist with integrity, kindness and compassionate heart.

My sincere appreciation and respect also go to my co-supervisor Prof. Ross McKinnon and Dr. Michael Ward, and to Dr. Dong Gui Hu and Prof. Peter MacKenzie, thank you for your enormous support, advices and expertise throughout this thesis project, which could not be accomplished without it.

I would like to thank my fellow labmates and friends. To Alex Haines, thank you for your involvement and enormous contribution. Thanks also to Julie-Ann Hulin, Apichaya Chanawong (Bo), Dhilushi Wijayakumara, Shuang Cui, Lu Lu, Anne Rogers, Tran Nguyen and many other colleagues in the Department of Clinical Pharmacology for the brilliant ideas that shared, motivation and tremendous support throughout this project. It has always been a pleasure to work with you.

I would also like to acknowledge HPEQ PhD scholarship from the Indonesian Government and Flinders Tuition Fee Sponsorship for the funding throughout the period of my PhD candidature.

To my mentor Prof. M Aris Widodo MS SpFK PhD, whom I am truly grateful for his continuous support and encouragement to start off career on the right path. To my fellow PhD Noer Aini, Rahma Triliana, and Rio Risandiansyah, thank you for supporting each other with motivation and prayer. My gratitude also goes to all my colleagues and friends in the Medical Faculty, University of Islam Malang and to all my fellow Indonesian students in Adelaide.

To my mother and father, I owe you everything, I cannot thank you enough for your endless love, wise counsel and all you have done to support me during my PhD, including caring my baby son back in Indonesia. Thanks also to my siblings Azis, Ema and Ela, my parents, brothers and sisters of in-laws, also my grandmother, thank you for your constant prayers, comfort and support that I always need.

Finally, I am especially grateful for my wonderful husband Eko Susanto for your unconditional love, sympathetic ear and countless sacrifices that enable me to get to this point. For my beautiful and joyful children, Zach Al-Farabi, Zia Az-Zahra and Zen Al-Ghazali whom I am so proud of and dedicate this thesis for, I love you eternally.

## PUBLICATION

MUBAROKAH, N., HULIN, J. A., MACKENZIE, P. I., MCKINNON, R. A., HAINES, A. Z., HU, D. G. & MEECH, R. 2018. Cooperative Regulation of Intestinal UDP-Glucuronosyltransferases 1A8, -1A9, and 1A10 by CDX2 and HNF4 $\alpha$  is Mediated by a Novel Composite Regulatory Element. *Mol Pharmacol*, 93, 541-552.

## CONFERENCE PROCEEDINGS

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Combinatorial regulation of UGT1A8 gene by intestinal factors Cdx2 and HNF4 $\alpha$ . Oral and poster presentation, Consortium of Biological Sciences 2017, Kobe, Japan, December 2017.

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Combinatorial regulation of UGT1A8 gene by Cdx2 and HNF4 $\alpha$ . Poster and teaser presentation, ComBio2017, Adelaide, SA, Australia, October 2017.

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Study of a novel regulatory mechanism that controls UGT1A8 synergistic activation by Cdx2 and HNF4 $\alpha$ . Poster Presentation, The Australian Society of Clinical and Experimental Pharmacologists and Toxicologists - The Molecular Pharmacology of G protein coupled receptors (ASCEPT-MPGCR) Joint Scientific Meeting, Melbourne, VIC, Australia, December 2014.

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Hepatocyte nuclear factor 4 $\alpha$  (HNF4 $\alpha$ ) involves caudal-related homeodomain protein 2 (Cdx2) to regulate the intestinal-specific expression of UGT1A8 and -1A10 genes, but independently regulate UGT1A9 gene. Poster presentation, The 20<sup>th</sup> International Symposium on Microsomes and Drug Oxidations (MDO2014), Stuttgart, Baden-Württemberg, Germany, May 2014.

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Cdx2 and HNF4 $\alpha$  synergize to regulate UDP-Glucuronosyltransferase (UGT) 1A8 gene expression. Poster Presentation, 53rd Annual Meeting of the Society of Toxicology (SOT), Phoenix, Arizona, USA, March 2014.

Mubarokah, S. N., MacKenzie, P. I., Meech, R. Synergistic regulation of UDP-Glucuronosyltransferase (UGT) 1A8, -1A9 and -1A10 gene expression by Cdx2 and HNF4 $\alpha$ . 2013 ASCEPT Annual Scientific Meeting, Melbourne, VIC, Australia, December 2013.

Mubarokah, S. N., MacKenzie, P. I., McKinnon, R. A., Ward, M., Meech, R. The role of PPAR $\gamma$  in the synergistic effect of genistein and sodium butyrate on UGT1A8 gene expression. Poster presentation, Annual scientific meeting of the Australian Society for Medical Research (ASMR), Adelaide, SA, Australia, June 2013.

Mubarokah, S. N., MacKenzie, P. I., McKinnon, R. A., Ward, M., Meech, R. Synergistic effect of genistein and sodium butyrate on UGT1A8 gene expression, and the PPAR $\gamma$  possible role. Poster Presentation, The Australian Health and Medical Research (AHMR) Congress, Adelaide, SA, Australia, November 2012.

Mubarokah, S. N., MacKenzie, P. I., McKinnon, R. A., Ward, M., Meech, R. Peroxisome proliferator-activated receptor gamma (Ppar $\gamma$ ) may be involved in the regulation of UGT1A8 expression by genistein. Oral Presentation. Combio 2012, Combined ASBMB, ASPS, ANZCDB, NZSBMB, NZSPB, AMATA and APPS/MPPP Annual Meeting, Adelaide, SA, Australia, September 2012.

Mubarokah, S. N., MacKenzie, P. I., McKinnon, R. A., Ward, M., Meech, R. UDP-Glucuronosyltransferase 1A8 expression in sodium butyrate-treated caco-2 cells increased by genistein treatment (potential role in butyrate presence of colon carcinogens inactivation). Oral Presentation, International Conference of Life Science 2012, Brawijaya University, Batu – Malang, East Java, Indonesia, July 2012.

Mubarokah, S. N., Meech, R. Potential role of isoflavone consumption in anticancer effect. Oral Presentation, The International Conference of TCAM 2012 (Traditional and Complementary Alternative Medicine), Surakarta, Central Java, Indonesia, June 2012.

Mubarokah, S. N., MacKenzie, P. I., McKinnon, R. A., Ward, M., Meech, R. Isoflavone genistein increases UDP-Glucuronosyltransferase 1A8 expression in sodium butyrate-induced differentiated Caco-2 cells. Poster presentation, The Australian Health and Medical Research (AHMR) Congress, Adelaide, SA, Australia, June 2012.

## **SCHOLARSHIPS AND AWARDS IN SUPPORT OF THIS THESIS**

Health Professional Education Quality (HPEQ) program of Directorate General of Higher Education of Indonesian Ministry of National Education for Post Graduate Scholarship, 2011 – 2014.

Flinders University Tuition Fee Sponsorship, 2014-2015, 2017.

Flinders University College of Medicine, Nursing and Health Sciences Conference grant to attend and present at the Consortium of Biological Sciences 2017, Kobe, Japan, December 2017.

ASCEPT committee student educational conference travel award to attend and present at the ASCEPT-MPGCR Joint Scientific Meeting, Melbourne, VIC, Australia, December 2014.

Flinders University Faculty of Health Sciences Conference grant to attend and present at the ASCEPT-MPGCR Joint Scientific Meeting, Melbourne, VIC, Australia, December 2014.

Flinders University Research student international conference travel award to attend and present at the 20<sup>th</sup> International Symposium on Microsomes and Drug Oxidations (MDO2014), Stuttgart, Baden-Württemberg, Germany, May 2014.

ASCEPT committee student educational conference travel award to attend and present at the Annual ASCEPT Scientific Meeting, Melbourne, VIC, Australia, December 2013.

Flinders University Faculty of Health Sciences Conference grant to attend and present at the 2013 ASCEPT Annual Scientific Meeting, Melbourne, VIC, Australia, December 2013.

Microsomes and Drug Oxidations committee travel award for international student of research higher degree program, 2013.

2014 SOT/AstraZeneca/SOT Endowment Fund/IUTOX Travel Award to attend and present at the SOT's 53rd Annual Meeting at the Phoenix Convention Center in Phoenix, Arizona, USA, 2013.

Flinders University Faculty of Health Sciences Conference grant to attend and present at the Combio 2012, Combined ASBMB, ASPS, ANZCDB, NZSBMB, NZSPB, AMATA and APPS/MPPP Annual Meeting, Adelaide, SA, Australia, September 2012.

Best Oral Presenter at the 5<sup>th</sup> International Conference of Life Science (ICSL), East Java, Indonesia, 2012.

## ABBREVIATIONS

ABCB1	ATP-binding cassette sub-family B member 1
ABCG2	ATP-binding cassette sub-family G member 2
AMPK	Adenosine monophosphate-activated protein kinase
AR	Androgen receptor
ARE	Antioxidant response element
AhR	Aryl hydrocarbon receptor
AT	Anion transporter
BE	Barrett's esophagus
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
BSP	Bone sialoprotein
bp	Base pair
C/EBP $\alpha$	CCAAT-enhancer-binding protein-alpha
CA	Clofibric acid
CAR	Constitutive androstane receptor
CBC	Crypt base columnar
cDNA	Complementary DNA
Cdx2	Caudal-related homeobox 2
ChIP	Chromatin immunoprecipitation
CLDN2	Claudin-2
COAD	Colon adenocarcinoma
CRC	Colorectal cancer
CRISPR	Clustered regularly interspaced short palindromic repeats
CSC	Cancer stem cell
Ct	Cycle threshold
CYP450	Cytochrome P450
DH5 $\alpha$	Doug-Hanahan-5 alpha
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulfoxide
DN	Dominant negative
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferases
dNTP	Deoxynucleotide triphosphate
DR1	Direct repeat-1
DTT	Dithiothreitol
EC	Esophageal cancer
<i>E.coli</i>	<i>Escherichia coli</i>
E1	Estrone
E2	Estradiol
E3	Estriol
ECM	Extracellular matrix
EGCG	Epigallocatechin gallate
EGF	Epidermal growth factor
EMSA	Electrophoretic mobility shift assay
EPI	Epirubicin
ER	Estrogen receptor
ERK	Extracellular Signal-regulated Kinase
ESC	Embryonic stem cell
ESCC	Esophageal squamous cell carcinoma
EtBr	Ethidium bromide
FA	Fatty acid
FACS	Fluorescence-activated cell sorting

FAM	Fluorescein amidite
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
FXR	Farnesoid X receptor
GFP	Green fluorescent protein
GIT	Gastrointestinal tract
GPR49	G-protein coupled receptor 49
GR	Glucocorticoid receptor
GSK	Glycogen synthase kinase
GST	Glutathione S-transferase
H3K4me3	Histone 3 lysine 4 trimethylation
H3K27me3	Histone 3 lysine 27 trimethylation
H3Kac	Histone 3 lysine acetylation
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDACi	HDAC inhibitor
Hhex	Hematopoietically-expressed homeobox
HNF1 $\alpha$	Hepatocyte nuclear factor 1-alpha
HNF4 $\alpha$	Hepatocyte nuclear factor 4-alpha
HT	Hesperitin
IEB	Isoosmolar electroporation buffer
IEC	Intestinal epithelial cell
IESC	Intestinal epithelial stem cell
IF	Immunofluorescence
IFD	Induced fit docking
ISEMF	Intestinal sub-epithelial myofibroblast
iPSC	Inducible pluripotent stem cell
kb	Kilo bp
kDa	Kilo Dalton
LB	Luria-Bertani
LBD	Ligand binding domain
LETF	Liver-enriched transcription factors
L-FABP	Liver-type fatty acid-binding protein
LGR5	Leucine-rich repeat-containing G-protein coupled receptor 5
LPH	Lactase-phlorizin hydrolase
LRC	Label-retaining cell
LXR	Liver X receptor
MC	Methylcholanthrene
MCT	Monocarboxylate transporter
MMF	Mycophenolate mofetil
MMP	Matrix metalloproteinase
MPA	Mycophenolic acid
MPAG	MPA-glucuronide
mRNA	Messenger RNA
NaB	Sodium butyrate
NCBI	National Center for Biotechnology Information
NEAA	Non-essential amino acid
NF- $\kappa$ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NF-YA	Nuclear transcription factor Y-alpha
NR	Nuclear receptor
Nrf2	Nuclear factor erythroid-related factor 2
NST	Nucleotide sugar transporter
MAPK	Mitogen-activated protein kinase
OD	Optical Density
OHCE	Hydroxy-catecholestrogens

Oct4	Octamer-binding transcription factor 4
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PGC-1	PPAR $\gamma$ coactivator-1
PK	Protein kinase
PLB	Passive lysis buffer
PMA	Phorbol 12-myristate 13-acetate
PPAR	Peroxisome proliferator-activated receptor
PPRE	Peroxisome proliferator response element
pSmad	Phosphorylated Smad
PXR	Pregnane X receptor
qISC	Quiescent intestinal cell
qRT-PCR	Quantitative real time-PCR
RE	Response element
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
ROCK	Rho-associated protein kinase
R-spondin/ RSPO	Roof plate-specific spondin
RT	Room temperature
RXR	Retinoid X receptor
SBE	Smad binding element
SCFA	Short chain fatty acid
SDM	Site-directed mutagenesis
SDS	Sodium dodecyl sulfate
SERM	Selective estrogen receptor modulator
SI	Sucrose-isomaltase
siRNA	Small interfering RNA
SOC	Super optimal with catabolite repression
Sox2	Sex determining region Y-box 2
Sp1	Specificity protein 1
STK31	Serine/threonine kinase 31
TA	Transit amplifying
TAD	Transactivation domain
TAE	Tris-acetate buffer
TAM	Tamoxifen
TBE	Tris-borate-EDTA buffer
tBHQ	tert-Butylhydroquinone
TCDD	Tetrachlorodibenzo-p-dioxin
TCGA	The cancer genome atlas
TE	Tris-EDTA buffer
TF	Transcription factor
TGF- $\beta$	Transforming growth factor- $\beta$
TSS	Transcription start site
UDCA	Ursodeoxycholic acid
UDP	Uridine diphosphate
UDPGA	UDP-glucuronic acid
UDPGlcNAC	UDP-N-acetyl glucosamine
UGT	UDP-glucuronosyltransferase
UV	Ultra violet
VDR	Vitamin D receptor
XRE	Xenobiotic response element

# CHAPTER 1. INTRODUCTION

## 1.1 Glucuronidation and human UGTs

### 1.1.1 Glucuronidation

On a daily basis, humans are in contact with large numbers of foreign substances/xenobiotics, such as environmental contaminants, ingested dietary constituents, and drugs. These xenobiotics may be pharmacologically active or potentially toxic. Many are in the form of lipophilic compounds. In addition, many endogenous lipophilic molecules such as bile acids, bilirubin and steroid hormones are generated and used within the body as part of normal physiology, while others are generated as by-products of metabolic processes. These lipophilic molecules cannot be easily excreted from the body since they are not water soluble, and may hence accumulate in human tissues and increase to toxic levels over time.

To facilitate their excretion, lipophilic molecules undergo biotransformation processes that enhance their hydrophilicity. This process is mediated by drug/xenobiotic metabolising enzymes.

Traditionally, metabolism is classified into 2 major pathways: Phase I and Phase II. This concept was first introduced by (Williams, 1959). Phase I metabolic enzymes catalyse oxidation, reduction, and hydrolysis, with the cytochrome P450s (CYP450s) being the major family of Phase I enzymes. Metabolites of Phase I are substrates that are susceptible to be further metabolised by Phase II enzymes through conjugation reactions. Conjugation converts lipophilic compounds to more polar molecules, enabling transport to a more aqueous compartment of the cell and enabling cellular efflux and eventually elimination via urine or bile/faeces (Dutton, 1980, Mackenzie *et al.*, 1997). Phase II enzymes include UDP-Glucuronosyltransferases (UGTs), sulfotransferases, and glutathione S-transferases. UGTs have been reported to be the most effective enzymes in xenobiotic metabolism (Markey, 2002). They are responsible for catalysing 40-70% of all clinically used drugs metabolised in Phase II (Evans and Relling, 1999, Smith, 2013).

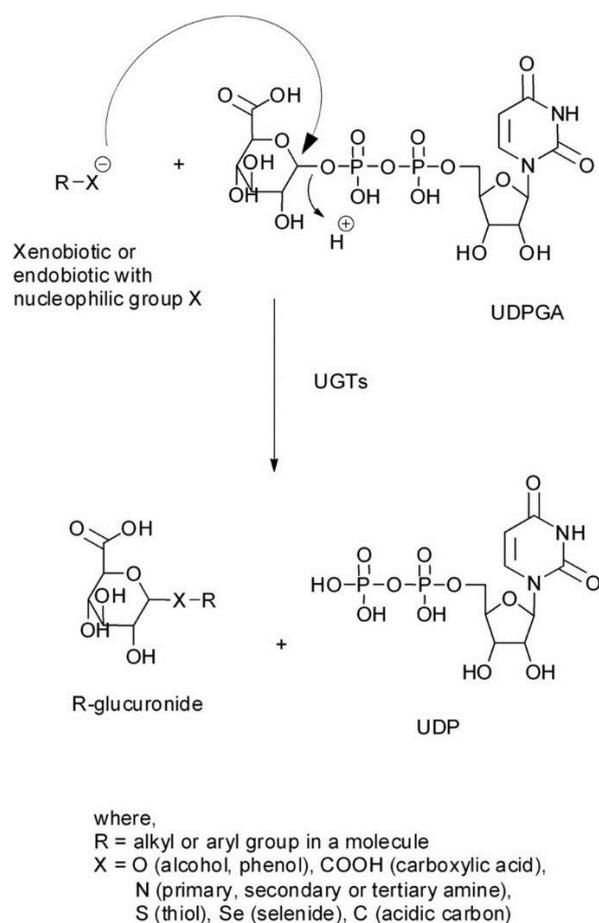
For some compounds, conjugation is the initial step in biotransformation. Hence conjugation should not always be considered as a 'Phase II' metabolic process. For example, morphine is

directly glucuronidated; paracetamol is predominantly glucuronidated or conjugated with sulfate (Anzenbacher and Zanger, 2012); and phenols, naphthols, alcohols, amines, and carboxylic acids can be directly metabolised without first being subject to Phase I reactions (Sanchez and Kauffman, 2010, Tephly and Burchell, 1990).

The majority of drug metabolites in bile and urine are glucuronides (Markey, 2002, Dutton, 1980). This implies that glucuronidation is the principal pathway of Phase II metabolism. Glucuronidation generally renders drugs and other molecules biologically inactive as well as promoting their excretion from the body (Mackenzie *et al.*, 1997).

The glucuronidation reaction is catalysed by a large superfamily of UGT enzymes. This process involves covalent addition of a hydrophilic sugar moiety (glucuronide) from the co-factor uridine diphosphate glucuronic acid (UDPGA) to a nucleophilic group of the substrate (Mackenzie *et al.*, 2010a, Guillemette *et al.*, 2010, Dutton, 1980, Rowland *et al.*, 2013). UGT enzymes possess the ability to catalyse the transfer of the glucuronic-acid moiety at many functional sites, such as carbonyl, carboxyl, sulfonyl, hydroxyl (alcoholic, phenolic), and amine groups (primary, secondary, or tertiary). In general each UGT has a broad substrate specificity, this together with the large number of structurally divergent UGTs in the superfamily allows an extraordinarily large number of molecules to be the target of glucuronidation (Radomska-Pandya *et al.*, 1999, Tukey and Strassburg, 2001, Dutton, 1980, Mackenzie *et al.*, 1997). The glucuronidation reaction is depicted in Figure 1.1.

Glucuronidation plays a significant role in the clearance of chemicals and biologically active toxins from the body, and considering many drugs and their metabolites are conjugated, it also contributes towards drug bioavailability (Miners *et al.*, 2004). It is therefore very important to improve our understanding and knowledge of the enzymes responsible for glucuronidation, including the molecular mechanisms that control UGT expression.



**Figure 1.1. The glucuronidation mechanism.**

Glucuronidation is a nucleophilic substitution facilitated by UDP-Glucuronosyltransferases. The nucleophilic group of the aglycone (X) attacks the anomeric carbon atom of the glucuronic acid moiety of the co-substrate UDPGA to form a glucuronide. This liberates the by-product, uridine diphosphate (UDP). Image is reproduced from (Argikar, 2012), with permission from the American Society for Pharmacology and Experimental Therapeutics (ASPET).

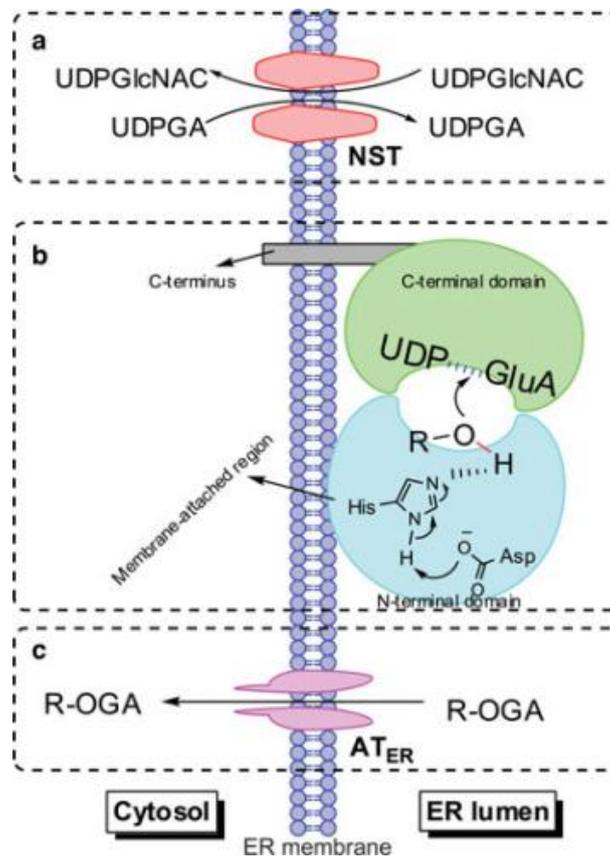
### 1.1.2 Human UDP-Glucuronosyltransferase (UGT) enzymes

The human UDP-Glucuronosyltransferases are a superfamily of enzymes comprising 22 members.

These enzymes are expressed in both hepatic and extrahepatic tissues (Guillemette *et al.*, 2010).

Recent studies by Court *et al* identified at least 29 human tissues (from 26 adults and 3 foetal samples) expressing UGTs at different mRNA levels. These were adult liver, foetal liver, salivary gland, stomach, small intestine, pancreas, colon, trachea, lung, heart, bone marrow, spleen, thymus, adrenal gland, thyroid, ovary, uterus, placenta, breast, testis, prostate, adipose, skeletal muscle, foetal brain, cerebellum, adult nasal mucosa and foetal nasal mucosa. Consistent with where most glucuronidation takes place, the liver is the tissue which expresses the highest overall level of UGTs (Court *et al.*, 2012).

As reviewed by (Wu *et al.*, 2011), the UGT enzyme is localised in the membrane of the endoplasmic reticulum (ER), with the active site facing the luminal side where the conjugation reaction occurs. Figure 1.2 shows the topology of a UGT enzyme. A single UGT consists of two domains- the N-terminal domain and the C-terminal domain. The N-terminal domain binds the aglycone (lipophilic substrate) and the C-terminal domain binds the co-substrate UDPGA (Figure 1.2b). Several studies show that UGTs are able to form homo- or hetero-dimers or oligomers (Finel and Kurkela, 2008, Radomska-Pandya *et al.*, 2005, Bock and Köhle, 2009). However, the nature of their dimeric or oligomeric structures, as well as their functional implication, remain to be fully understood. One of the key steps in glucuronidation is transporting UDPGA across the ER membrane, from the cytosol to the lumen, to allow UDPGA to access the active site of the enzyme. UDPGA is synthesized in the cytoplasm and transported into the ER by nucleotide sugar transporters (NSTs). Several NSTs that can transport UDPGA have been identified and characterised, with SLC35D1 (UGTrel7) displaying the highest activity (Kobayashi *et al.*, 2006, Muraoka *et al.*, 2001). The function of NSTs require counter transport by exporting another nucleotide sugar, UDP-*N*-acetyl glucosamine (UDPGlcNAC) out to the cytosol (Figure 1.2a). Another important step in glucuronidation is transporting the glucuronides from the lumen of the ER out to the cytosol, which is mediated by ER-localised organic anion transporters (Kardassis *et al.*) (Csala *et al.*, 2004) (Figure 1.2c).



**Figure 1.2. Model of UGT enzyme topology within the endoplasmic reticulum.**

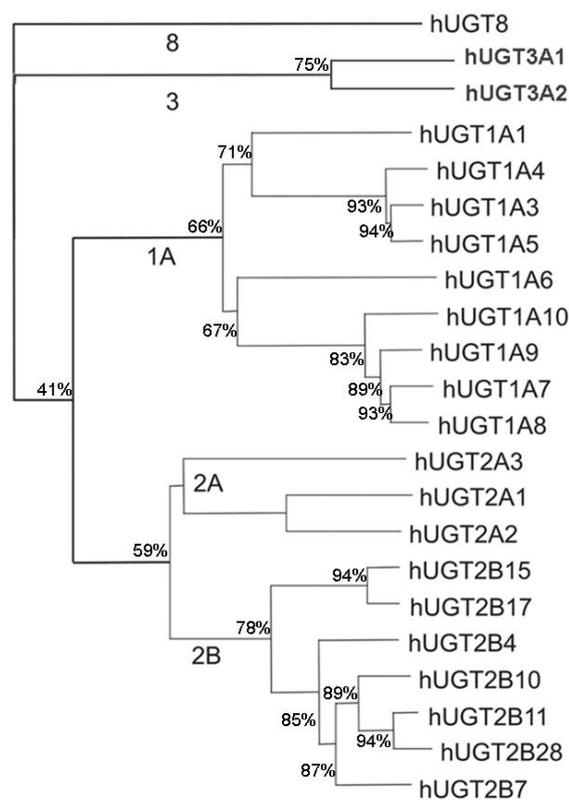
a) Transport of UDPGA by NSTs. b) Functional UGT with N- and C- terminal domains. c) Transport of glucuronides by  $AT_{ER}$  via diffusion. R-OH, phenolics; R-OGA, glucuronide. Image is reproduced from (Wu *et al.*, 2011), with permission from Elsevier.

The subcellular location of UGTs gives direct access to products of Phase I reactions, bringing advantages to their function as Phase II metabolising enzymes (Kiang *et al.*, 2005). However, the fact that UGTs are located within the ER contributes to the difficulties in identifying *in vivo* effects of UGT enzymes, and is responsible for the phenomenon of "latent" enzyme activity (Kiang *et al.*, 2005). Dutton and co-workers (Dutton, 1980) were the first to reveal that little or no enzyme activity could be observed unless membrane disrupting agents were added to the preparation of microsomes, with an increase in UGT activity of around 20 fold seen when adding detergent (Dutton, 1980).

### 1.1.3 The UGT gene super family

The UGT superfamily consists of four families, separated based on amino acid identity: UGT1, UGT2, UGT3 and UGT8 (Mackenzie *et al.*, 2005a). The UGT1 and UGT2 families mainly utilise

UDP-glucuronic acid as the glycosyl (sugar) donor, but may also use other UDP-sugars such as UDP-glucose and UDP-xylose in some contexts. These families represent the earliest and most extensively characterised UGT enzymes (Mackenzie *et al.*, 2005a, Meech and Mackenzie, 2010). Members of the UGT1 and UGT2 families were defined initially by their enzymatic activities and subsequently purified to homogeneity. Later, the mRNAs encoding these UGTs were identified by sophisticated molecular cloning methods (Mackenzie *et al.*, 1984). Once UGT mRNA sequences were available, the families were expanded by use of molecular cloning tools to identify new related genes.

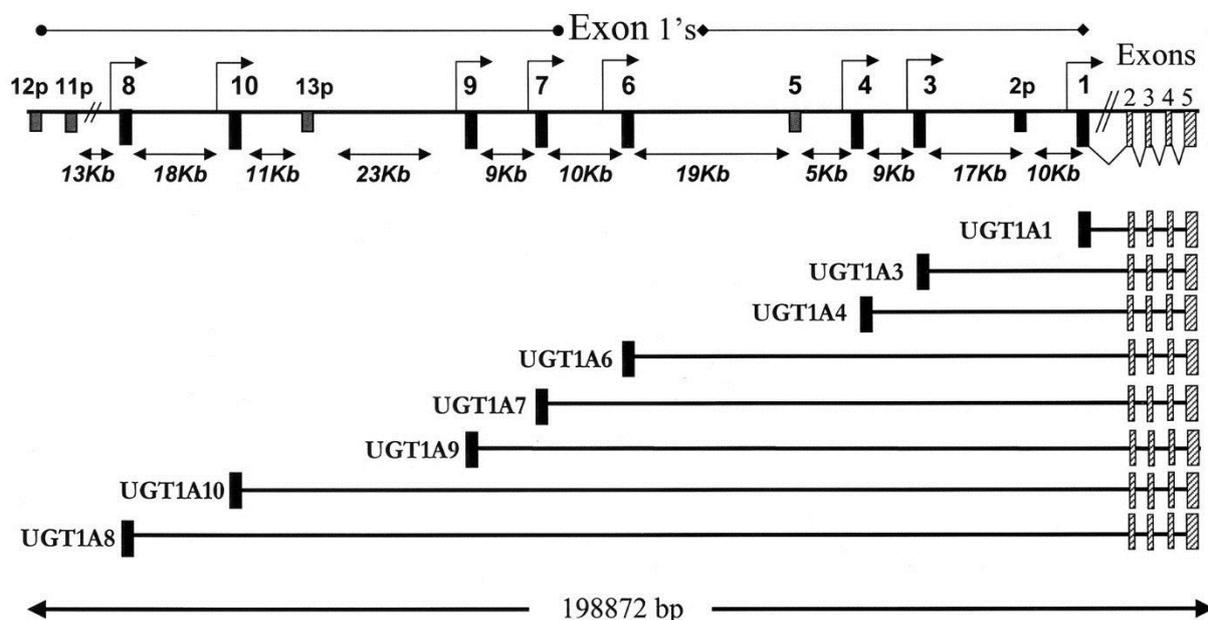


**Figure 1.3. The phylogenetic tree of UGT1, UGT2, UGT3 and UGT8.**

The amino acid sequence homology between enzymes is represented by percentage similarity values. Information for the figure was collected from (Meech and Mackenzie, 2010, Gregory *et al.*, 2004b, Guillemette, 2003, Gregory, 2004).

A phylogenetic tree of the human UGTs is depicted in Figure 1.3. UGT1A and UGT2 families share a similarity of 41% (Guillemette, 2003). UGT2A and UGT2B families share a 59% sequence homology.

The UGT3 and UGT8 families form a separate branch from the UGT1 and UGT2 families. A single UGT1 locus on chromosome 2q37 encodes all UGT1A enzymes. There are nine functional isoforms of UGT1A (UGT1A1, UGT1A3-UGT1A10), while UGT1A2, UGT1A11, UGT1A12 and UGT1A13 are considered pseudogenes. The UGT1A locus is comprised of multiple variants of exon 1 located upstream of four common exons (termed exons 2-5). The different UGT1A isoforms are formed through alternative splicing of a single exon 1 to the common exons (Mackenzie *et al.*, 2005a). Organisation of the UGT1A complex locus is depicted in Figure 1.4.



**Figure 1.4. The human UGT1 family complex locus.**

The UGT1A family consists of 12 isoforms; 9 functional isoforms indicated with transcription arrows, and 4 pseudogenes labelled with p (2p, 11p, 12p, 13p). The positions and distances relative to exon 2-5 (striped rectangle) and within exon 1 (black rectangle) is indicated. Exons 2-5 join exon 1 in the mature transcript. The human UGT1A locus extends for 198,872 bp as calculated from the DNA sequence (Tukey and Strassburg, 2001). Image is derived from (Tukey and Strassburg, 2001) with permission from The American Society for Pharmacology and Experimental Therapeutics (ASPET).

Six UGT1A members are expressed in the liver, the main organ where glucuronidation occurs. These are UGT1A1, UGT1A3, UGT1A4, UGT1A5, UGT1A6 and UGT1A9 (Tukey and Strassburg, 2000). Furthermore, a high degree of sequence homology (>80%) is observed within the UGT1A3-1A5 cluster and within the UGT1A7-UGT1A10 cluster. UGT1A1, the UGT1A3-1A5 cluster, and UGT1A6 are mainly hepatically expressed. Identification of UGT enzymes in non-liver tissues by Strassburg *et al* revealed that UGT1A7 and UGT1A10 are not hepatically expressed, and both

enzymes are found specifically in the gastrointestinal tract (Strassburg *et al.*, 1997). Later, the same research group showed that UGT1A8 is also extrahepatic, being expressed mainly in colon, which brought attention to the importance of glucuronidation at the distal end of the digestive tract. UGT1A9 is expressed in the gastrointestinal tract as well as in liver, thus UGT1A7-UGT1A10 is commonly described as the gastrointestinal UGT cluster (Strassburg *et al.*, 1998).

Together the UGT1A family members conjugate a very large number of small molecules. The divergent substrate specificity of multiple UGT1A isoforms is determined by a substrate-binding domain in the amino-terminal half of the protein that is encoded by exon 1. This substrate-binding site is considered 'loose,' which allows multiple substrates to bind to one isoform (Meech and Mackenzie, 1997). The cofactor UDPGA on the other hand, binds to a region in the carboxyl half of the UGT protein, which is identical in all UGT1A isoforms and is highly conserved (Tukey and Strassburg, 2001). Both aglycone and UDPGA binding regions are presumed to cooperate for transferring glucuronic acid to the substrate. The UDPGA-binding C-terminal regions of the UGT1 and UGT2 families are relatively more closely related than the N-terminal aglycone-binding regions (Tukey and Strassburg, 2000).

The UGT2A and 2B subfamilies form a very large gene cluster located on chromosome 4q13.2. The seven functional enzymes of the UGT2B family (2B4, 2B7, 2B10, 2B11, 2B15, 2B17 and 2B28) are encoded by individual unique genes, which consist of six exons each (Hu *et al.*, 2014a). All UGT2B enzymes are expressed in the liver (Tukey and Strassburg, 2000). The UGT2A1 and UGT2A2 enzymes are generated by alternative splicing of a gene that has two variants of exon1 located upstream of five common exons (termed exons 2-6). Thus, they each have a unique first exon but share the common exons. In contrast, the UGT2A3 enzyme is encoded by a single gene containing six unique exons. UGT2A enzymes are expressed in the liver, small intestine, brain and foetal lung (Guillemette, 2003).

There is a degree of overlap in the aglycone specificities of UGT1A and UGT2B enzymes. Indeed, studies in UGT1A-deficient rats have shown that compensatory induction of UGT2B genes can occur in the intestinal tract (Wang *et al.*, 2009).

The UGT8 gene family only has one member, UGT8A1 (usually abbreviated as UGT8), which is encoded by a gene of five exons on chromosome 4q26 (Mackenzie *et al.*, 2005a, Stahl *et al.*, 1994). The activity of the UGT8 enzyme was originally described in 1968 in embryonic chicken brain; based on this activity the enzyme was named ceramide galactosyltransferase (CGT) (Basu *et al.*, 1968). The first human cDNA encoding CGT/UGT8 was isolated by (Ichikawa *et al.*, 1996). The UGT8 name was assigned by the Human UGT nomenclature committee (Mackenzie *et al.*, 2005a). The UGT8 enzyme uses UDP-galactose as the sugar donor and performs almost exclusively conjugation of galactose to ceramide (Ichikawa *et al.*, 1996). UGT8 is essential for biosynthesis of glycosphingolipids, cerebrosides and sulfatides in the nervous system. Consistent to its function, this enzyme is highly expressed in the nervous system, but is also found in considerable levels in the gastrointestinal tract, kidney, thymus, and bone marrow (Meech *et al.*, 2015). Recently, its novel function as a regulator in homeostasis and signalling of bile acid was discovered (Meech *et al.*, 2015).

The UGT3 family was the last UGT family to be identified in approximately the year 2000 based on homology searches of the newly available complete human genome sequence. The UGT3 genes were sufficiently dissimilar to the other known UGTs to be defined as a new family (< 50% sequence homology) (Tukey and Strassburg, 2000). The UGT3 family is divided into two members, namely UGT3A1 and UGT3A2, with the two genes located on chromosome 5p13.2. The functions of the UGT3 enzymes remained elusive for some time as no activities could be identified using a range of common UGT substrates and UDPGA as the sugar donor. Ultimately it was discovered that neither UGT3 enzyme use UDPGA as co-substrate. UGT3A1 utilises UDP-N-acetylglucosamine as the sugar donor (Meech and Mackenzie, 2010) with a variety of aglycone substrates including ursodeoxycholic acid (UDCA), a molecule used for the treatment of cholestasis. UGT3A2 utilises UDP-glucose and UDP-xylose as its sugar donors, with limited activity with UDP-galactose (Meech and Mackenzie, 2010). The major UGT3A2 substrate identified to date is 4-methylumbelliferone; it is also active with 1-hydroxypyrene, 7-hydroxycoumarin, naringenin and genistein (MacKenzie *et al.*, 2011). Gene expression analysis by reverse-transcriptase-PCR (RT-PCR) showed that UGT3A1 is mainly expressed in liver and kidney, and

also at a low level in stomach, duodenum, colon, and testes (Mackenzie *et al.*, 2008). UGT3A2 mRNA expression was detected in thymus, testes and kidney but not liver (MacKenzie *et al.*, 2011).

## **1.2 Expression of UGT genes: general aspects**

### **1.2.1 Tissue specific distribution of human UGTs**

In humans, among the UGT families, the UGT1A and UGT2B families are thought to contribute most significantly to Phase II xenobiotic and endobiotic metabolism with UGT2A, UGT3 and UGT8 family enzymes playing more minor roles (Meech *et al.*, 2012). Understanding the distribution of UGT enzymes can help predict potential drug-drug interactions mediated by these enzymes (Kiang *et al.*, 2005). Thus, extensive information has been gathered detailing the distribution of the UGT1A and UGT2Bs in human tissues.

Each UGT isoform is expressed in a tissue-specific manner, although their patterns are highly overlapping. The liver is certainly the most important organ expressing the largest variety of UGT enzymes; hence the greatest number of studies focus on UGTs in this organ. Studies of UGT mRNA expression show that all UGT1A genes (except UGT1A8) and all UGT2B genes are expressed in the liver. However, the mRNA levels of UGT1A5, UGT1A7, UGT1A10 and UGT2B11 that were detected were extremely low, with disagreement between studies (Ohno and Nakajin, 2009, Court *et al.*, 2012). Among the UGT1A family genes, UGT1A1 and UGT1A9 are the most abundantly expressed in the liver. However, UGT2B genes are more abundant in the liver overall, with UGT2B4 being reported as the highest hepatically expressed UGT (Wu *et al.*, 2011).

Non-hepatic expression of UGTs was first addressed around two decades ago (Strassburg *et al.*, 1997, Strassburg *et al.*, 1998), with evidence that small intestine and colon are major sites of UGT1A and UGT2B glucuronidation activities (Ohno and Nakajin, 2009). Metabolism at the site of absorption directly affects oral drug efficacy by limiting the level of active drug entering the systemic circulation. In addition, UGT activities in the gastrointestinal tract affect the absorption of dietary chemicals and may provide protection against ingested toxins and carcinogens. Thus, gastrointestinal UGTs have been the subject of extensive studies acknowledging these significant

functions. Section 1.3 details on our current understanding of the expression and activity of UGTs in the gastrointestinal tract. Briefly however, the most important UGT isoforms in the small intestine and colon are UGT1A1, UGT1A8, UGT1A9 and UGT1A10 (Wu *et al.*, 2011). Among the UGT2B genes, UGT2B7, UGT2B17 and UGT2B15 are the three most abundantly expressed in the small intestine and colon (Wu *et al.*, 2011).

Table 1.1 shows UGT mRNA expression in various human tissue samples, as assayed using real-time reverse transcriptase-polymerase chain reaction (RT-PCR) (Court *et al.*, 2012, Ohno and Nakajin, 2009, Nakamura *et al.*, 2008), as well as compiled data from (Gregory, 2004) and (Guillemette *et al.*, 2014) showing UGT expression in major glucuronidation tissues such as liver, kidney, and gastrointestinal tract (oesophagus, stomach, small intestine, colon), as well as at other extra-hepatic sites that will not be discussed here. In comparing these reports, differences are observed regarding UGT expression in some tissues, which may be caused by large interindividual variability, influenced by factors such as genetic polymorphism, ethnicity, age or gender. Moreover, care must also be taken when correlating UGT mRNA levels to protein levels. A study (Izukawa *et al.*, 2009) in human liver found that not all hepatic UGT mRNA and protein levels are well correlated; in general, further studies of protein levels are required for conclusive information.

**Table 1.1. mRNA expression of UGT1As and 2Bs in human tissues.**

mRNA	Tissue																					
	Liver	Esophagus	Stomach	Small Intestine	Colon	Brain	Cerebellum	Thyroid	Thymus	Breast	Testis	Ovary	Placenta	Cervix	Heart	Lung	Trachea	Kidney	Bladder	Spleen	Uterus	Pancreas
1A1	+	-	+	+	+	-	-	+/-	+/-	+	+	-	+/-	-	-	-	+/-	+	-	-	+	-
1A3	+	-	+	+	+	-	-	+/-	+/-	-	+/-	-	-	-	-	+/-	+/-	+	-	-	+	-
1A4	+	-	+	+	+	-	-	+/-	+/-	-	+/-	-	+/-	-	-	-	+/-	+	-	-	+	+
1A5	-	+	+	+	+	+/-	-	-	+	-	+	-	+	+	-	-	+	+	+	-	+	-
1A6	+	-	+	+	+	+	-	+/-	+/-	-	+	-	+/-	-	-	+/-	+	+	+	+/-	+	-
1A7	-	+	+	+	+	-	-	+/-	+/-	-	+/-	-	+/-	+	-	-	+	+	-	-	-	-
1A8	-	+	-	+	+	-	-	-	+/-	+/-	+/-	-	-	-	-	-	+	+	+	-	-	-
1A9	+	+	-	+	+	-	-	+/-	+/-	+	+	-	+/-	-	-	-	+/-	+	+	-	-	-
1A10	-	+	+	+	+	-	-	-	+/-	-	+/-	-	+/-	-	-	-	+	+	-	-	-	-
2B4	+	+	-	+	+	-	-	-	+	+	+	-	-	-	+/-	+	+/-	+	-	-	-	-
2B7	+	+	+	+	+	+	-	-	+/-	+	+/-	-	+/-	-	-	+	-	+	-	-	+	+
2B10	+	+	-	-	-	-	-	-	-	+	+	-	-	-	-	+	-	+	-	-	-	-
2B11	+	-	-	+	-	+/-	+/-	+/-	-	+	+/-	-	-	-	-	+	-	+	-	-	-	+
2B15	+	+	+	+	+	-	-	+/-	-	+	+	-	-	-	-	+	+	+	-	-	-	+
2B17	+	-	+	+	+	+/-	+/-	+/-	+	+	+	+	+/-	+	-	+	+	+	-	+	+	+
2B28	+	-	-	-	-	-	-	-	-	+	-	-	-	NT	NT	-	-	-	+	-	-	-

(+) indicates that UGT mRNA transcript has been detected in the tissue, (-) indicates the absence of mRNA transcript, (+/-) indicates discrepancies between studies, and NT indicates that the presence of mRNA transcript has not been tested.

### 1.2.2 Constitutive and inducible regulation of UGTs

Regulation of UGTs at the transcriptional level has been the subject of many studies over the last two decades, and we now have a relatively good, although still incomplete, understanding of transcription factors that regulate UGT expression. There are two major aspects of UGT regulation that are of interest: constitutive and inducible expression. These are discussed here in general terms, and more specifically with respect to intestinal UGTs in section 1.4.

Studies of constitutive UGT expression generally examine the factors that determine tissue specificity (for example liver vs intestine) and regional patterning (i.e. along the proxomodistal axis of the gastrointestinal tract). They also examine mechanisms that control differential expression

between progenitor cells and various differentiated cell types of the organ. These tissue-, region-, and cell-type specific patterns are typically controlled by the same transcription factors that underlie the development, patterning and regeneration of these tissues.

Studies into the regulation of the hepatically expressed rat *Ugt2b1* gene were the first to reveal tissue-specific transcription factors controlling UGT gene expression. Using functional and DNA binding assays in HepG2 cells, two liver-enriched transcription factors namely hepatocyte nuclear factor-1 alpha (HNF1 $\alpha$ ) and CAAT/enhancer binding protein-alpha (C/EBP $\alpha$ ), were shown to bind to the *Ugt2b1* promoter region and induce its activity (Hansen *et al.*, 1997, Hansen *et al.*, 1998). Many hepatic UGTs are now known to be controlled by HNF1 $\alpha$  as well as another liver enriched factor HNF4 $\alpha$ , through direct binding of these factors to their gene promoters (Hu *et al.*, 2014a).

Although HNF1 $\alpha$  was initially considered to be the main transcriptional activator for liver specific genes including UGTs, its identification in non-hepatic tissue, particularly gastrointestinal tract in mice (Kuo *et al.*, 1990) and humans (Uhlén *et al.*, 2015), suggested that it may also regulate the expression of intestinal UGTs. This theory was confirmed by a study from our laboratory on the intestinally expressed UGT1A8, -1A9 and -1A10 genes as described in detail in section 1.4. Our studies also identified a role for the intestine-specific transcription factor caudal-related homeodomain protein 2 (CDX2) in regulation of these intestinal UGTs (Gregory *et al.*, 2004a). In many cases, there is a requirement for transcription factors to work in combination to define cell-type specific expression (Verzi *et al.*, 2013, Biggin, 2011, Davidson and Levine, 2008) and as discussed further in section 1.4, this paradigm is applicable to intestinal UGT regulation.

Studies of the inducible expression of UGTs focus on their response to chemical inducers, which may be endobiotics (i.e. signalling molecules such as steroids or bile acids) or xenobiotics such as toxins, carcinogens, drugs, and other bioactive dietary and environmental chemicals. These pathways are regulated by a number of transcription factor families but most prominent among them are the nuclear receptor (NR) family. Nuclear receptors bind small molecules as ligands and this induces their activation and often their translocation to the nucleus where they can regulate target genes. This provides a critical feedback mechanism that senses levels of small molecules

and enables the cells to alter expression of genes that control the levels of these small molecules (such as UGTs and other drug metabolism pathway genes) that help clear the small molecule. Such feedback loops in regulatory circuits involving UGTs and their own substrates have recently been reviewed (Bock, 2012). It should be noted that regulatory feedback can also involve transcription factors other than nuclear receptors. For example, the cytotoxic anticancer drug epirubicin, a substrate for UGT2B7, upregulates UGT2B7 via the p53 transcription factor (Hu *et al.*, 2014b).

Nuclear receptors can also contribute to the tissue specific expression pattern of UGTs, particularly when they respond to endogenous signalling molecules that are concentrated in specific tissues, such as steroids (Hu *et al.*, 2014a). Finally, polymorphisms in *UGT* gene promoters and other regulatory regions may cause variations in UGT expression between individuals and this is applicable to both constitutive and inducible expression paradigms (Hu *et al.*, 2014a, Mackenzie *et al.*, 2010b, Gregory *et al.*, 2004b, Gregory *et al.*, 2004a). Table 1.2 summarises nuclear receptors and their target UGTs. The roles of these factors in control of intestinal UGTs are described in more detail in section 1.4.

**Table 1.2. Nuclear receptors, ligands, and their target UGTs**

<b>Receptors</b>	<b>Ligands</b>	<b>Target UGTs</b>
ER $\alpha$	17-beta estradiol, Tamoxifen	1A4, 2B15, 2B17
AR	Testosterone, Dihydrotestosterone, R1881, Flutamide	1A1, 1A3, 2B10, 2B11, 2B15, 2B17, 2B28
GR	Cortisone, Dexamethasone	1A1
CAR	1,4-bis[2-(3,5-dichloropyridyloxy)] benzene (TCPOBOP), 3a,5a-androstenol, 3a,5a-androstanol, Artemisinin	1A1, 2B7
FXR	Chenodeoxycholic acid, Cholic acid, Deoxycholic acid, Z-Guggulsterone	1A3, 2B4, 2B7
LXR	24(S)-Hydroxycholesterol, 24(S),25-Epoxycholesterol	1A3

PPAR $\alpha$	Clofibric acid, Gemfobrozil, Bezafirate, GW 7647, 20-hydroxyeicosatetraenoic acid(Siess <i>et al.</i> ), 11,12-epoxyeicosatrienoic acid (EET)	1A3, 1A6, 1A9, 2B4
PPAR $\gamma$	Ciglitazone, Troglitazone, 15-Deoxy-D12, 14-prostaglandin J2, GW 9662	1A9
PXR	Rifampicin, Lithocholic acid, Hyperforin	1A1
VDR	25-Hydroxyvitamin D3, 1,25-Dihydroxyvitamin D3	2B15, 2B17

Table is partly taken from review by (Hu *et al.*, 2014a), available for thesis purposes from Taylor & Francis.

## 1.3 Intestinal UGT enzymes

### 1.3.1 Intestinal UGTs in phenolic compound glucuronidation

As mentioned, glucuronidation in the human gastrointestinal tract plays a critical role as the first line of defence against potentially toxic and carcinogenic substances that are ingested. It also controls the activity of some dietary nutrients and bioactive compounds, and controls efficacy and potential toxicity of oral drugs, and intravenous drugs that reach the intestine following enterohepatic circulation (Gregory *et al.*, 2004b, Wu *et al.*, 2011, Radomska-Pandya *et al.*, 1998).

For numerous compounds, intestinal glucuronidation often serves as the major metabolic pathway and most efficient elimination process. One of the most well studied groups is phenolic compounds. These compounds are known to be efficiently glucuronidated along the gastrointestinal tract, leading to a very poor oral bioavailability (Wu *et al.*, 2011, Gao and Hu, 2010), with only 5% of dietary phenols reaching the plasma unchanged (Clifford, 2004). Dietary polyphenols are naturally occurring and widely found in vegetables, fruits and herbs, and thus are consumed regularly by humans; they include flavonoids (flavones, flavonols, flavanones, isoflavones), stilbenes, quinones, and many more. Synthetic phenolic compounds are also found in drugs, such as acetaminophen, raloxifene, tamoxifen, and mycophenolic acid. Table 1.3 shows examples of phenolic compounds predominantly metabolised by UGTs expressed in colon. Some of the compounds have been reviewed in (Wu *et al.*, 2011).

In addition to being extensively metabolised by intestinal UGTs, flavonoids can induce endogenous UGT gene expression *in vivo* and *in vitro* (Petri *et al.*, 2003, Galijatovic *et al.*, 2001). Petri *et al.* reported that in humans, acute ingestion of quercetin, a flavonoid abundant in capers and coriander (Bhagwat *et al.*, 2014), increased UGT1A1 mRNA by 2.4 fold in enterocytes. An increase in UGT1A1 expression was also seen in an *in vitro* experiment using the colon cancer cell line Caco-2: chrysin, (a flavonoid found in honey and propolis (Siess *et al.*, 1996)), at a concentration of 25  $\mu$ M, significantly induced UGT1A1 without affecting UGT1A6, -1A9 and -2B7 levels. The result also showed that an increase in UGT1A1 expression increased glucuronidation of N-hydroxy-PhIP, a mutagen, by 10 fold (Galijatovic *et al.*, 2001). These findings suggest that modification of intestinal UGT expression by regular dietary constituents can help prevent carcinogenesis by enhancing the intestinal capability to metabolise carcinogens.

The role of UGTs in cancer of the gastrointestinal tract is likely to be complex. In humans most cancers occur in colon (rarely in small intestine). In normal colonic mucosa UGT1A enzymes are continuously expressed at a high level; this expression is reduced significantly in adenocarcinoma tissues (Wang *et al.*, 2012). This is consistent with the relatively less differentiated state of carcinoma cells but also suggests that UGT1A expression may play a role in the progression of colon carcinogenesis. In support of the latter idea, a recent study (Liu *et al.*, 2016) demonstrated that deletion of the UGT1 locus in the intestinal crypt in mice causes a reduction in active p53 and leads to repression of apoptosis. Moreover, when colon cancer was induced in the UGT1A null mice, the number and size of tumours was greater than in wildtype mice (Liu *et al.*, 2016). Exactly how loss of UGT1A1 promotes colon cancer progression is unclear.

**Table 1.3. Phenolic compounds metabolised by intestinal UGTs.**

Phenolic compounds	Metabolised by UGT	Reference
<b>Flavones</b>		
Apigenin Chrysin	UGT1A6	(Liu <i>et al.</i> , 2007)
Baicalein	UGT1A8, 1A9	(Zhang <i>et al.</i> , 2007)
Wogonin Oroxylin A	UGT1A7, 1A8, 1A9, 1A10	(Zhou <i>et al.</i> , 2010)
<b>Isoflavones</b>		
Genistein Daidzein Glycitein Formonentin Biochanin A	UGT1A1, 1A8, 1A9, 1A10	(Tang <i>et al.</i> , 2009)
<b>Flavonols</b>		
Kaempferol Quercetin	UGT1A9	(Oliveira and Watson, 2000)
<b>Chalcones</b>		
Xanthohumol	UGT1A8,1A9, 1A10, 1A1, 1A7, 2B7	(Ruefer <i>et al.</i> , 2005)
<b>Stilbenes</b>		
Resveratrol	UGT1A1, 1A8, 1A9, 1A10	(Brill <i>et al.</i> , 2006, Miksits <i>et al.</i> , 2010, Iwuchukwu <i>et al.</i> , 2011)
Combretastatin	UGT1A9	(Aprile <i>et al.</i> , 2010)
<b>Coumarins</b>		
4-methylumbelliferone	UGT1A6, 1A7, 1A10	(Uchaipichat <i>et al.</i> , 2004, Antonio <i>et al.</i> , 2003)
Daphnetin	UGT1A9, 1A6	(Liang <i>et al.</i> , 2010)
<b>Phenolic drugs</b>		
Entacapone	UGT1A7, 1A10	(Luukkanen <i>et al.</i> , 2005)
Mycophenolic acid	UGT1A8, 1A10	(Mackenzie, 2000)
Tamoxifene	UGT1A8, 1A10, 2B7	(Sun <i>et al.</i> , 2007)
Raloxifene	UGT1A1, 1A8	(Mizuma, 2009)
Propofol	UGT1A7, 1A8, 1A10, 1A9	(Court <i>et al.</i> , 2012)
Irinotecan	UGT1A1	(O'dwyer and Catalano, 2006)

### 1.3.2 Intestinal microbiota and UGT activities

The intestine is home to most of the human microbiota; interaction of Phase I and Phase II metabolic enzymes with microbiota activity is likely to play an important role in maintaining intestinal homeostasis. Studies to understand the mechanisms by which microbiota regulate drug and xenobiotic metabolism are relatively recent (Swanson, 2015). Claus *et al* demonstrated that the “xenobiotic sensor” transcription factors constitutive androstane receptor (CAR) and PXR (regulators of several CYP and UGT genes) were significantly increased in the gut of germ-free mice. The endogenous levels of CYP enzymes were also altered in the germ-free mice (Claus *et al.*, 2011). Another study reported upregulation of the Phase II biotransformation enzymes glutathione S-transferases (GSTs) by butyrate in primary human colon tissue and colon cancer cell lines (HT29 and LT97) (Pool-Zobel *et al.*, 2005). Butyrate is generated by gut microbiota as discussed further below. Overall, current data suggests a potential function of microbiota metabolic products in the regulation of intestinal metabolic enzymes.

Butyrate is a short chain fatty acid (SCFA) produced from the fermentation of dietary fibre by the intestinal microbiota (McIntyre *et al.*, 1993, VanHook, 2015). High levels of dietary fibre are usually found in plant-based food, which also contain high level of polyphenols that can also give rise to butyrate in the intestine (Blaut *et al.*, 2003). Butyrate provides a main energy source for intestinal epithelial cells and plays an important role in the regulation of many intestinal functions. It is associated with intestinal barrier protection and integrity (Kelly *et al.*, 2015), is a modulator of intestinal inflammatory processes related to metabolic diseases (Chang *et al.*, 2014, Puddu *et al.*, 2014, Brahe *et al.*, 2013), and may even be neuroprotective due to the gut-brain axis (Bourassa *et al.*, 2016).

Many studies have implicated high fibre diets with polyphenol rich constituents in lowering the risk of cancer. For example, dark green vegetable consumption was reported to be related to a lower risk of Barrett's oesophagus (Jiao *et al.*, 2013), while colorectal cancer is associated with a low intake of fruit and vegetables (Terry *et al.*, 2001, Pericleous *et al.*, 2013). Since the risk of carcinogenesis is also associated with the level of UGT activity (Wallig, 2004), it is important to

gain an understanding of the interplay between intestinal microbiota-producing butyrate, phenolic compound intake and regulatory mechanisms of intestinal UGT expression. One of the goals of this project was to better understand regulatory mechanisms of intestinal UGTs, including the roles of dietary bioflavones and butyrate (see section 1.2).

### **1.3.3 Expression of intestinal UGTs**

As the main organ system in drug metabolism after liver, the gastrointestinal tract (oesophagus, stomach, small intestine, and colon) expresses UGTs at high levels, although the diversity of isoforms expressed is much lower than in liver. As discussed in detail in section 1.4, the innermost layer of the intestinal tract is the mucosal layer in which absorptive and secretory functions are carried out. Very early work indicated that the UGT1 subfamily plays a greater role in intestinal glucuronidation than UGT2, with its activity around 6 fold higher overall (Dubey and Singh, 1988). Furthermore, UGT1A expression was shown to be localised in the epithelial cells of the mucosa, and not in deeper submucosa or muscularis layers of the intestine. This is in accordance with UGT function, as expression in the mucosal epithelial cells allows direct contact with xenobiotics.

In the last two decades, the differential patterns of UGT1A isoforms along the gastrointestinal tract from the upper (proximal) to the lower (distal) end has been defined (Gregory *et al.*, 2004b).

Strassburg *et al* examined a set of human tissues and concluded that UGT1A7 is expressed in the upper GI tract, mainly in the stomach, and UGT1A8 is most highly expressed in the lower part of GI tract, e.g. colon (Strassburg *et al.*, 1998). UGT1A10 appears to be expressed throughout small and large intestine (Gregory *et al.*, 2004b). Expression of UGT1A8 in the colon may be important for metabolism of substances which are not glucuronidated in the upper regions of the gastrointestinal tract, and the prolonged exposure time of colon mucosa to the substances may lead to potential interplay of the conjugating activities with intestinal microbiota (Strassburg *et al.*, 1998).

Zhou *et al* and Tang *et al* found that the correlation between glucuronidation rates (in heterologous expression systems and human tissue microsomes) and UGT expression levels in tissues is able to predict tissue-specific UGT activity profiles. In particular, tissue-specific glucuronidation of phenolic compounds (i.e. bioflavones) was predicted using microsomes from human liver and

intestine, suggesting that in liver UGT1A1 and UGT1A9 are likely to be most important for this metabolism while UGT1A8 and UGT1A10 play the major role in intestine (Zhou *et al.*, 2010, Tang *et al.*, 2010).

The mechanisms that regulate UGT expression in the intestine are complex and involve both constitutive and inducible factors as introduced in section 1.2.2. In order to explain the mechanisms of constitutive UGT regulation in intestine, it is critical to first describe the structure and development of the intestine, and the key factors that are involved in its development and turnover. This discussion is provided in the next section.

## **1.4 Intestinal development and homeostasis**

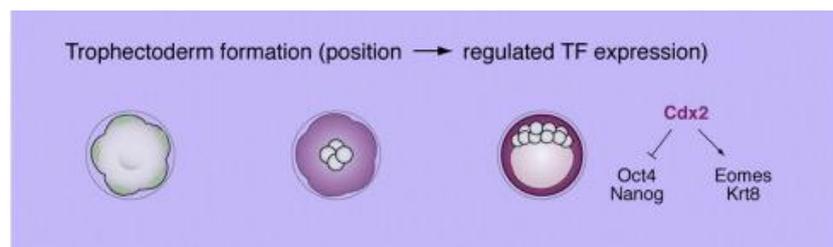
### **1.4.1 Intestinal development is critically regulated by homeobox factors including the master regulator CDX2**

During embryogenesis, the intestine is developed from the endoderm layer that formed at the end of gastrulation. The endodermal sheets have anterior (anterior intestinal portal) and posterior (caudal intestinal portal) ends. These endodermal formations are determined and patterned by the expression of developmental regulators. For example, studies in mice show that the anterior endoderm is mainly patterned by the expression of Sox2 and the homeobox protein Hhex, and gives rise to foregut structures, e.g. stomach. The posterior endoderm will give rise to the small and large intestines, and activity of the homeobox factor Cdx2 is essential in their formation (Noah *et al.*, 2011).

Cdx2 is a member of the caudal-related homeobox family. Homeobox genes are key factors in controlling normal embryonic development particularly anterior-posterior patterning (McGinnis and Krumlauf, 1992). In the mouse blastocyst, Cdx2 is initially expressed specifically in the trophectoderm that gives rise to extraembryonic membrane (Beck *et al.*, 1995). Ralston and Rossant showed that Cdx2 drives trophectoderm commitment into maturation by negatively regulating pluripotency factors octamer-binding transcription factor 4 (Oct4) and Nanog genes in the trophectoderm. They also positively regulate trophectoderm lineage markers, like eomesodermin (Eomes) and Keratin 8 (Krt8) (Ralston and Rossant, 2008). Mouse embryos with

genetic deletion of Cdx2 (Cdx2 null) show disruption of trophoectoderm formation leading to pre-implantation death. Figure 1.5 illustrates the trophoectoderm maturation and the role of Cdx2 in the lineage-specific regulatory mechanism (Ralston and Rossant, 2008).

The role of Cdx2 in development of the inner cell mass can be assessed by rescuing the trophoectoderm (and hence implantation) defect by tetraploid fusion. In these rescued embryos, it is evident that Cdx2 is necessary for the normal progression of the blastocyst through the gastrulation and tail bud elongation stages. Cdx2 expression becomes restricted to posterior gut endoderm at about day 12.5 post coitum in mice (around halfway through development). Consistent with this expression pattern, Cdx2 null mice have posterior truncations of embryonic structures showing that Cdx2 controls anterior-posterior patterning (Chawengsaksophak *et al.*, 2004). Moreover, the intestinal tissue is transformed to a foregut phenotype.



**Figure 1.5. Illustration of trophoectoderm developmental fate, as regulated by CDX2.**

The image is reproduced from (Ralston and Rossant, 2008), with permission from © Elsevier.

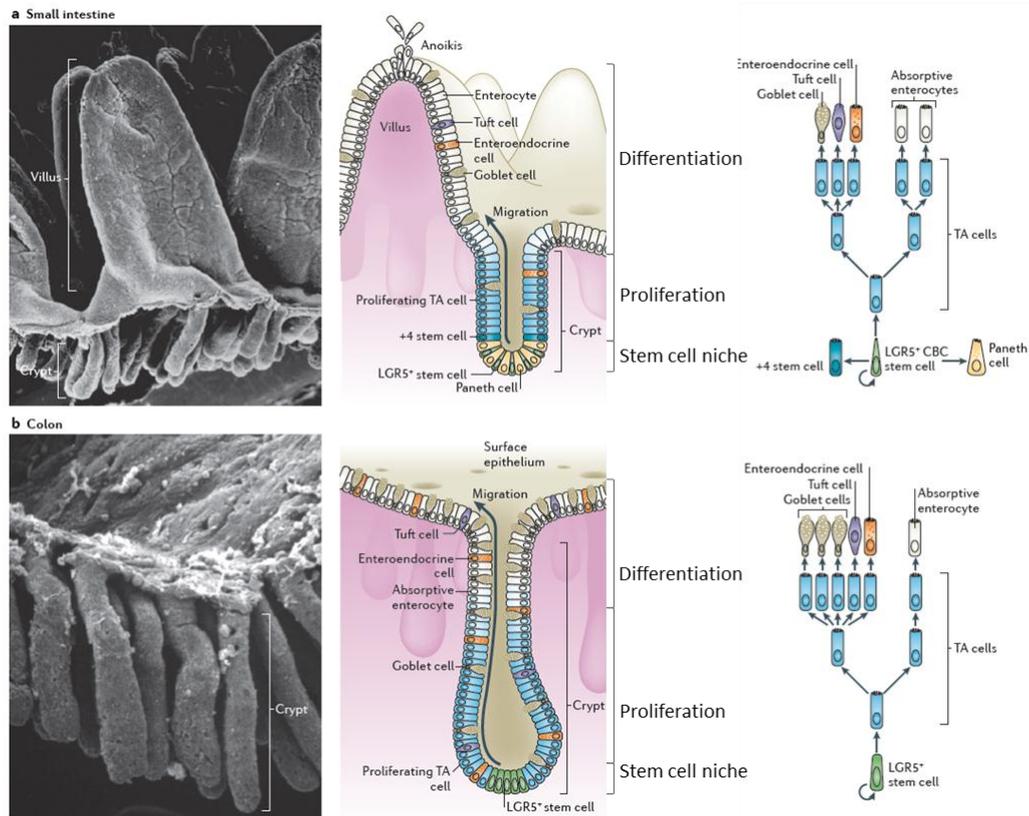
As well as being important for early specification of gut structures, Cdx2 is also important for the differentiation of intestinal cell types. In mouse, the level of Cdx2 increases 6 to 7 fold during development of the intestinal epithelium (James *et al.*, 1994). Cdx2 has also been shown to continue to express at a high level in the adult stage, and shown to remain exclusively expressed in the intestinal tract, with no detectable Cdx2 transcripts in other tissues, such as stomach, liver, brain, lung, pancreas, testis, skeletal muscle, heart, spleen and kidney (Suh *et al.*, 1994). The germline Cdx2 null mouse model shows embryonic lethality, therefore, conditional transgenic Cdx2 null mice have been generated to examine the importance of Cdx2 in controlling intestinal identity at early developmental stages and also in adults. The study of conditional Cdx2 null mice showed

failure in colon formation and severely abnormal intestinal cell differentiation at every stage (Gao *et al.*, 2009), thus highlighting Cdx2 as the key factor in regulating intestinal homeostasis, both during development and throughout life. Additional studies confirm that in adults, Cdx2 is the key player in the maintenance of intestinal epithelial continuous renewal, by regulating cellular proliferation and differentiation as discussed further below (Suh and Traber, 1996, Silberg *et al.*, 2000, Beck *et al.*, 1999).

#### **1.4.2 The intestine is maintained by continual epithelial turnover that is controlled by developmental regulators.**

The mammalian intestinal epithelium continuously undergoes a dynamic process of cell renewal, and these cells are counted as the most rapidly turned over in the human body. This turnover involves the continual activity of an anatomically defined stem cell compartment as discussed more in this section.

The small intestine epithelium is comprised of columnar epithelial cells (enterocytes) arranged into a series of villi. These villi increase the area for absorption, each enterocyte also has microvilli and this further increases the absorptive surface area. Other rarer cell types found in the epithelial layer have secretory functions such as goblet cells or enteroendocrine functions. In the colon, the epithelium is also simple columnar but lacks villi. In both small intestine and colon, the epithelium forms invaginations called crypts which contain the stem cell niche compartment.



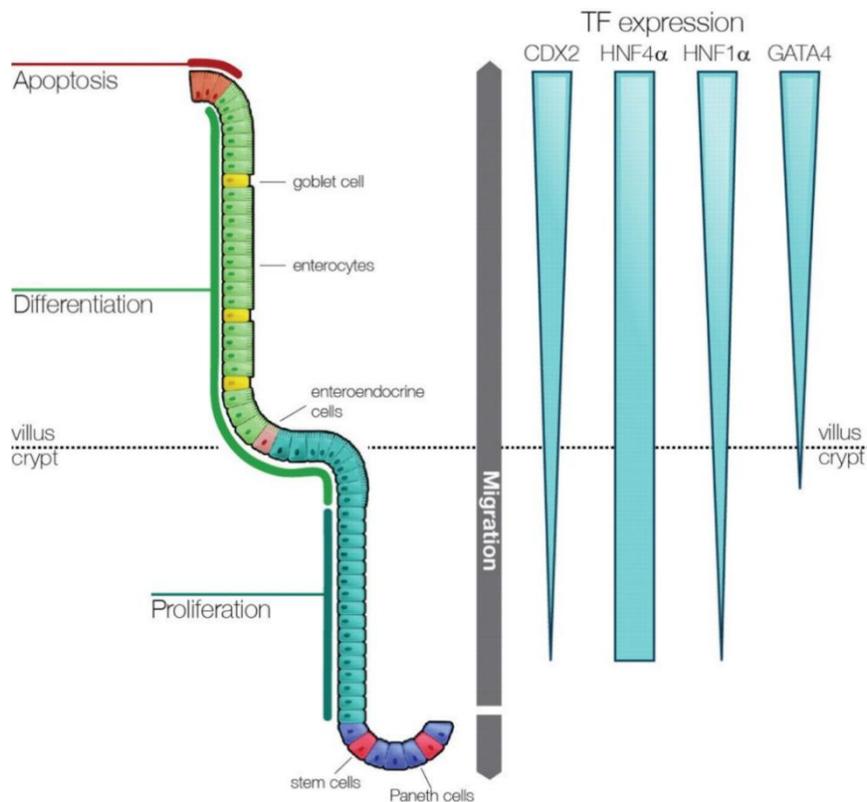
**Figure 1.6. Intestinal epithelial self-renewal.**

The crypt base is rich with highly proliferating progenitor of transit-amplifying (TA) cells, derived from Leu-rich repeat-containing G protein-coupled receptor 5-expressing (LGR5+) cells at the stem cell niche. These then migrate and differentiate into various mature cells. a) In the small intestine, the differentiated cells populate the intestinal villi, with the differentiated cell compartment making up more than half of the crypt. Paneth cells in the intestinal crypt are supplied every 3-6 weeks from the TA cells. b) In the colon the composition of differentiated cells occupy only one third, whereas the remaining two thirds are the crypt cells. The differentiation hierarchy and variants of lineage cells is described in the right panel. Images a and b are reproduced from (Barker, 2014), with permission from © Nature Publishing Group.

During turnover of the epithelium, stem cells in the crypt give rise to proliferative transit-amplifying (TA) cells that migrate up to the surface and subsequently differentiate into various functional mature cells (Barker, 2014). The turnover time for ileum epithelial cells of the small intestine is 3 to 5 days, whereas for colon cells it is 1 to 2 days. The shorter time for colon cell turnover might be related to the shorter distance of the colon crypt to migrate and differentiate on the luminal surface (Creamer *et al.*, 1961). Figure 1.6 shows the structural compartments of the small intestine and colon epithelium; the TA cells can be seen to span the length of the proliferation compartment of the crypt while stem cells defined by the marker Leu-rich repeat-containing G protein-coupled receptor 5 (LGR5) are located at the base of the crypt. In the small intestine, Paneth cells are

located adjacent to the stem cells in the crypt base; moreover a second stem cell type called the plus-four (+4) cell based on its location is also described (reviewed in (Olsen *et al.*, 2012)).

The developmental regulator Cdx2 plays a critical role in adult intestinal turnover. Since Cdx2 was first shown to be responsible for transcriptional activation of sucrose-isomaltase (SI), which is exclusively expressed in mature intestinal epithelial cells (Suh *et al.*, 1994), many other studies have also described Cdx2 as having a role in transcriptional control of multiple intestinal specific genes in adult (reviewed in (Olsen *et al.*, 2012)). The networks that it controls can be partly inferred from its distribution in the intestinal epithelium. In the mature intestine, Cdx2 and HNF1 $\alpha$  are highly expressed at the villus tips and decrease towards the crypt, whereas GATA-4 is only observed in the villus and not expressed in the crypt, and HNF4 $\alpha$  expression is evenly distributed along the crypt-villus axis (reviewed in (Olsen *et al.*, 2012)). The increased expression of Cdx2 in the villus cells (relative to the crypt cells) suggests a greater role in differentiation than in proliferation as will be discussed further below in the context of its target gene networks. Figure 1.7 shows the distribution of the major transcription factors controlling proliferation and differentiation of the intestinal epithelium.



**Figure 1.7. Graphical overview of major transcription factors involved in intestinal epithelial regulatory networks.**

Along the villus-crypt axis, the cell types and location are indicated, and the expression levels of transcription factors Cdx2, HNF4 $\alpha$ , HNF1 $\alpha$  and GATA-4 are depicted. Image is reproduced partly from (Olsen *et al.*, 2012), which is available for reproduction in the thesis without permission.

HNF4 $\alpha$  is well known as a master regulator of hepatic gene expression, and it is essential in hepatocyte development and differentiation (Parviz *et al.*, 2003). However, HNF4 $\alpha$  is also expressed in the intestine (Duncan *et al.*, 1994), kidney, and pancreas from early embryonic stages (Taraviras *et al.*, 1994). HNF4 $\alpha$ -null mice show embryonic lethality with abnormal gastrulation (Chen *et al.*, 1994). HNF4 $\alpha$  is reported to regulate the development visceral endoderm during embryonic development and thus the formation of the intestinal tract (Duncan *et al.*, 1997).

HNF4 $\alpha$  is expressed in the small and large intestine, from the crypt to villus. A study in a foetal mouse model, modified using Cre-*loxP* technology to remove HNF4 $\alpha$  from the colon epithelium, demonstrated that HNF4 $\alpha$  plays a significant role in controlling normal colon development, function and differentiation. In the HNF4 $\alpha$ -null colon, colon crypt development failed to occur, goblet-cell

maturation was disturbed, and this consequentially disrupted expression of multiple HNF4 $\alpha$ -site containing genes (Garrison *et al.*, 2006). Using a similar Cre-**loxP** technique to generate HNF4 $\alpha$  loss in the small intestinal epithelium, Cattin *et al* reported that HNF4 $\alpha$  controls normal function, stability and homeostasis between proliferation and differentiation of intestinal epithelial cells. HNF4 $\alpha$  loss impaired maturation of enterocyte and endocrine cells, and increased the number of crypt and goblet cells (Cattin *et al.*, 2009).

Large scale transcriptomic and epigenomic studies have been performed to understand the transcriptional networks controlled by factors such as Cdx2 and HNF4 in mouse intestinal development and homeostasis, and in some cases also in human-derived intestinal models. Epigenomic analysis has examined genome-wide DNA-protein binding sites for these factors using techniques such as ChIP-chip analysis, (a coupled protocol of chromatin immunoprecipitation (ChIP) and a hybridisation genome microarray (Ren *et al.*, 2000)), or ChIP-seq (ChIP coupled with massively multiparallel sequencing).

Transcriptome data from mouse intestinal cells revealed that genes containing a HNF4 $\alpha$  binding site in their promoter are upregulated during development of the intestine, and also in intestinal cells as they migrate from crypt to villus and differentiate (Stegmann *et al.*, 2006). Subsequently, Boyd *et al* used ChIP-chip to show the direct binding of the HNF4 $\alpha$  protein to the promoters of target genes involved in intestinal development and differentiation. These targets included the Cdx2 promoter indicating that HNF4 $\alpha$  plays a role in regulation of Cdx2. Interestingly, binding of Cdx2 protein to the Cdx2 promoter was also found (Boyd *et al.*, 2009), and another study later confirmed a Cdx2 positive autoregulation mechanism in the intestine (Barros *et al.*, 2011).

Two more recent epigenomic studies revealed that Cdx2 dynamically co-occupies target sites with different transcription factors in progenitor and mature intestinal cells. In progenitor cells, Cdx2 partners with GATA-6 and GATA-4, whereas during differentiation, Cdx2 relocates and partners with HNF4 $\alpha$ . The interaction of HNF4 $\alpha$  and Cdx2 is considered critical in regulating the genes expressed during cell differentiation (Verzi *et al.*, 2010, San Roman *et al.*, 2015).

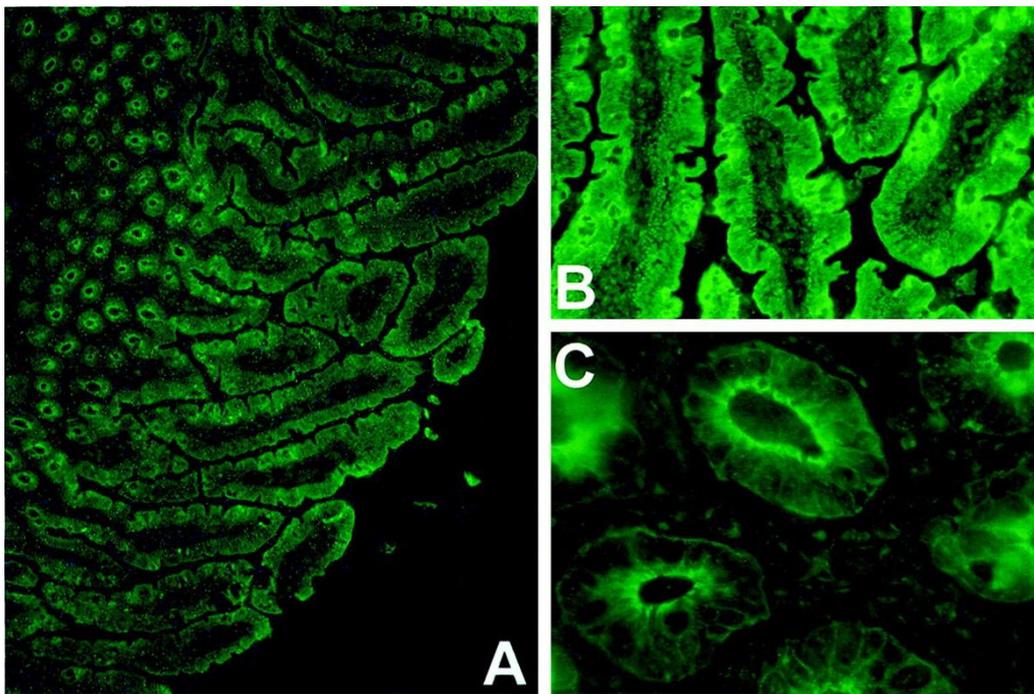
Using a mouse model, Verzi *et al* investigated the functional interaction of Cdx2 and HNF4 $\alpha$  and how their ability to occupy binding elements leads to regulation of specific gene expression. They observed joint co-occupancy of Cdx2 and HNF4 $\alpha$  at specific regions in the intestinal cell genome. The co-occupancies were found within 300 bp of a *cis*-element or most often closer, demonstrating that the joint occupancies of Cdx2 and HNF4 $\alpha$  potentially activate genes through functional enhancers. Using conditional double mutant Cdx2 and HNF4 $\alpha$  knock out mice, the cooperation of Cdx2 and HNF4 $\alpha$  in stimulating transcription was confirmed. Moreover, Verzi *et al* elucidated that the interaction between Cdx2 and HNF4 $\alpha$  is controlled epigenetically, with the chromatin accessibility maintained by Cdx2. They found that the loss of Cdx2 compromises chromatin configuration and severely affects HNF4 $\alpha$  binding, as well as disrupting proximal gene expression, whereas HNF4 $\alpha$  loss does not affect Cdx2 occupancy (Verzi *et al.*, 2013).

### **1.4.3 Constitutive regulation of intestinal UGTs by CDX2 and its partners**

As early as 1988, both Phase I and Phase II metabolic enzymes were found to be distributed in the same fashion from the crypt to villus surface of the small intestine, where the mature villus expresses higher level of metabolic enzymes, whilst this declines toward crypt cells. Interestingly expression of Phase II enzymes is higher than Phase I enzymes suggesting that, unlike in liver, Phase II activity is dominant in intestine.

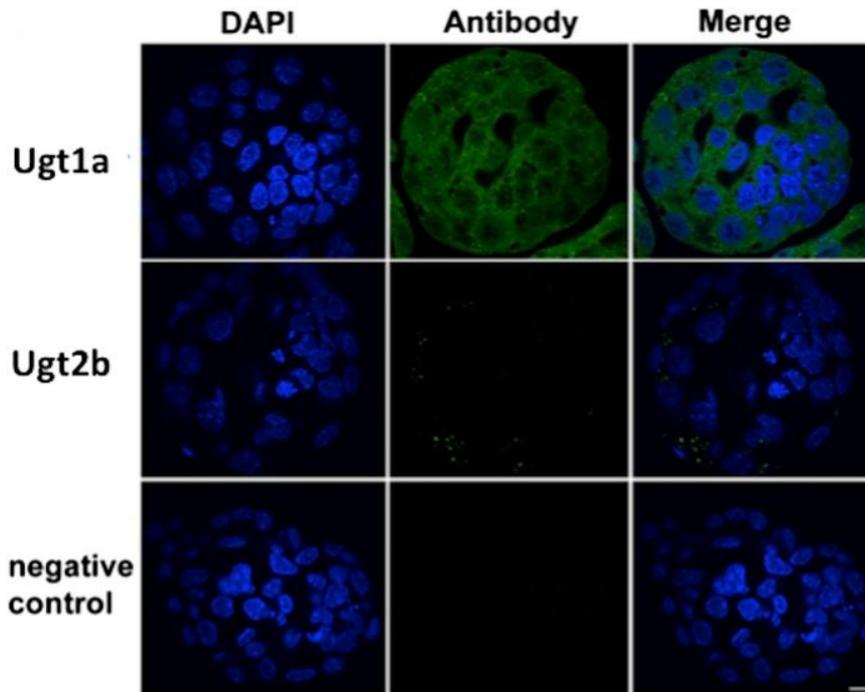
In 1988, Dubey and Singh showed the presence of UGT1 and UGT2 in the crypt as well as the villi, with UGT1 expression greater than UGT2. UGT expression in undifferentiated crypt cells is about 4 fold lower than in the differentiated villus cells, which is consistent with the high need for detoxifying metabolic functions in active absorptive cells. Moreover, cells in the crypt become progressively more differentiated as they migrate up towards the lumen/villi and immunofluorescence staining of UGT1A protein in the small intestine crypts has shown that it is concentrated in the apical region towards the lumen (Strassburg *et al.*, 2000) (see Figure 1.8). Hence overall it appears that UGT expression is associated with a differentiated phenotype. However, it seems likely that that cells at every stage of intestinal lineage progression require at least some degree of protection from toxic insults including undifferentiated dividing crypt cells, and

the presence of UGTs at moderate levels in TA cells suggests a role in this protection against harmful chemicals. In support of this idea, UGT activity in the crypt was shown to provide protection against genotoxic compounds during regeneration (Patel *et al.*, 1997). Moreover, although the level of UGT expression is higher in cells of the luminal surface/villi than crypt cells, it may be highly responsive to induction in crypt cells. For example, rat UGT1 activity in the replicating (TA) crypt cells was found to be 4 fold more responsive to 3-methylcholanthrene (3MC) than that in the differentiated villus cells (Dubey and Singh, 1988).



**Figure 1.8. Detection of UGT1A protein in the human intestine.**

A) Localisation of UGT1A protein in the epithelial cell layer and crypt of ileum at 40x magnification showing homogenous staining of villi and ring pattern of crypt. B) Higher magnification of UGT1A protein in villi (400x). The protein is confirmed as not being expressed in submucosa. C) UGT1A expression in the cross-section of crypt is shown concentrated in crypt enterocyte towards the lumen (400x). The figure is reproduced from (Strassburg *et al.*, 2000), with permission from The American Society for Biochemistry and Molecular Biology, Inc.



**Figure 1.9. Immunofluorescence staining showing UGT localization in mouse blastocyst.**

Pan-specific antibody against UGT1 and UGT2B were used for staining, with DAPI for nuclear staining. Only UGT1 signal is shown in high intensity throughout blastocyst, with negative from control. Image is reproduced from (Collier *et al.*, 2014), with permission from © The American Society for Pharmacology and Experimental Therapeutics.

The developmental regulation of UGT expression is a poorly studied topic. This is partly because phase II detoxification activities are considered to be generally low in embryonic/fetal stages. Indeed Court *et al* reported that the total UGT mRNA level in human fetal liver is approximately three times less than in the adult (Court *et al.*, 2012). Moreover, the lack of conservation between human and rodent UGT genes renders the mouse, which is widely used to study embryonic development, a poor model to understand human UGT regulation. Regardless, UGT expression has been demonstrated in embryonic and foetal development (Strassburg *et al.*, 2002) and some UGTs may be expressed even in very early embryos. Indeed the mouse pre-implantation blastocyst shows expression of Ugt1, but not Ugt2 enzymes (Figure 1.9), although their functions in this context are unknown (Collier *et al.*, 2014).

While little is known about how UGT expression is regulated at early stages of intestinal development, it might be inferred that factors that control UGT expression during adult intestinal

turnover are also involved in their expression during intestinal development. In this regard it is relevant that previous work from our laboratory showed that human CDX2 binds and regulates the *UGT1A8*, and *-1A10* promoters in the Caco-2 cell line (used as a model of intestinal epithelial cells) (Gregory *et al.*, 2004a). Interestingly, these studies suggested that CDX2 may not be able to bind to an equivalent site in the *UGT1A9* promoter. In addition, our group has shown that HNF1 $\alpha$  synergises with CDX2 to regulate *UGT1A8* and *-1A10* promoters in these cells (Gregory *et al.*, 2004a). The details of this regulation are provided in section 1.1.6 which focusses on the functions and regulation of the prototypical intestinal UGT isoform UGT1A8. However, briefly, these studies were interpreted as indicating that CDX2 plays a key role in restricting the expression of these two extrahepatic genes (*UGT1A8*, *-1A10*) to the intestine (Gregory *et al.*, 2004a).

Despite the central role that HNF4 $\alpha$  plays in controlling intestinal gene expression in cooperation with Cdx2 in mice, to date there have been no studies investigating the role of HNF4 $\alpha$  transcriptional regulation of UGTs in an intestinal context. However, consistent with its role as constitutive regulator of hepatic genes (Gonzalez, 2008), HNF4 $\alpha$  was shown to regulate hepatic UGT1A9 expression (Barbier *et al.*, 2005). As discussed previously, UGT1A9 is expressed in the liver and intestine, whereas *UGT1A7*, *-1A8*, and *-1A10* are extrahepatic. The main HNF4 $\alpha$  response element (HNF4 $\alpha$ RE) identified in the promoter of UGT1A9 by Barbier *et al* was shown to be only partially conserved in the *UGT1A7*, *-1A8*, and *-1A10* genes (Barbier *et al.*, 2005). They concluded that the sequences differences in the HNF4 $\alpha$  motifs in the extrahepatic *UGT* genes prevented their regulation by HNF4 $\alpha$  and thus contributed to their lack of hepatic expression, This work was complemented by findings that HNF1 $\alpha$  and HNF4 $\alpha$  are both required for hepatic regulation of UGT1A9 expression (Gardner-Stephen and Mackenzie, 2007).

Unresolved by these previous studies were two questions: firstly, whether HNF4 $\alpha$  plays any role in the intestinal expression of the extrahepatic UGTs and if so through what elements; secondly what directs UGT1A9 expression to the intestine given that no functional CDX2 binding site has been identified in this gene. Gaining a better understanding of UGT regulation in the intestine in general, and answering these two questions in particular, were some of the goals of this project.

## 1.5 *In vitro* models for intestinal UGT studies

### 1.5.1 Caco-2 as a model for enterocytes

The Caco-2 (Caco-2) cell line originally developed from a human colorectal adenocarcinoma by Jorgen Fogh (Fogh *et al.*, 1977) (Memorial Sloan-Kettering Cancer Center, New York), has been utilised for many years as a model to study intestinal epithelial cell functions, including studies of UGT function and regulation. Many studies investigating how flavonoids control UGT expression have used Caco-2 as a model, such as those showing the induction of UGT1A1 by quercetin (Petri *et al.*, 2003), chrysin, resveratrol, and curcumin (Galijatovic *et al.*, 2001, Iwuchukwu *et al.*, 2011). Studies into the transcriptional regulation of *UGT1A8-1A10* genes by Gregory *et al.* were also conducted in Caco-2 (Gregory *et al.*, 2004a, Gregory *et al.*, 2004b), and CDX2 has been the most investigated differentiation-related transcription factor studied in this context (reviewed in (Olsen *et al.*, 2012)).

When properly cultured, the Caco-2 cell line is able to provide a reasonable model of the biological and biochemical characteristics of the intestinal epithelium. Caco-2 is the only intestinal cell line that possesses the ability to spontaneously differentiate *in vitro* into polarised enterocyte monolayers under normal growing conditions without any inducers. This is a condition that cannot be achieved by other intestinal cell line models, for example HT-29, which requires specific inducers for differentiation (Rousset, 1986). Relative to non-immortal human primary enterocytes that can only be maintained for a short period in culture, it has the advantages of simplicity and reproducibility.

Full differentiation of Caco-2 cells is well characterised, and structurally and functionally resembles formation of an intestinal epithelium. The development of microvilli, tight junctions between adjacent cells and dome formation indicates structural characteristics of differentiated and polarised cells, whereas expression of brush border associated enzymes and other intestinal specific enzymes such as alkaline phosphatase, sucrase-isomaltase, lactase and aminopeptidase, characterise the typical intestinal function (Pinto *et al.*, 1983). Characteristics of colon cells (colonocytes) are progressively lost during Caco-2 cell transformation into enterocytes. Engle *et al.*

reported that characteristics of both colonocytes and enterocytes are present in Caco-2 cells immediately once the cells are confluent. Over subsequent weeks of post-confluent culture, there is spontaneous transformation into the enterocyte phenotype in which colonocyte specific-proteins decrease and those of enterocytes increase, transforming Caco-2 into a model of the small intestine that is particularly useful for applications such as drug screening (Engle *et al.*, 1998).

Despite the great potential provided by Caco-2 cells to model the intestinal epithelium, the culture protocol to differentiate Caco-2 cells is considered time consuming (requiring 21 days), and labour intensive. To overcome these concerns, Yamashita *et al* developed a short 5 day culture protocol involving butyrate supplementation in medium containing 10% serum. Although the monolayer integrity is less stable and decreases after 5-6 days, this short protocol provides broadly equivalent intestinal properties to the traditional methods of the 21 day protocol (Yamashita *et al.*, 2002).

While differentiated Caco-2 cells have a gene expression profile comparable to what is found in differentiated cells of the normal intestinal epithelium, proliferating Caco-2 cells have only minimal similarity to the normal intestinal proliferating crypt cells (Tremblay *et al.*, 2006). Engle *et al* has reported that Caco-2 cells share more features of the differentiated foetal intestine than with the adult crypt phenotype (Engle *et al.*, 1998). A major limitation of the Caco-2 cell system is that it does not reproduce the heterogeneity of cell types seen in the normal intestine. These limitations of Caco-2 cells compared to normal intestinal epithelium indicate that caution should be taken when extrapolating from *in vitro* data to the *in vivo* state.

### **1.5.2 The intestinal organoid model**

Cell line and *in vivo* (usually mouse) models have advantages as well as limitations in studies of the intestine, including those related to drug and xenobiotic metabolism. Alternative models that marry the convenience and long-term culture of immortal cell lines with the physiological relevance of animal models would be a powerful tool for intestinal research. Within the last few years, such an *ex vivo* model has been developed for the 3 dimensional (3D) culture of isolated primary intestinal tissue. This model called the 'organoid' or sometimes 'mini-gut', was developed largely by Sato *et al.* (Sato *et al.*, 2009, Sato *et al.*, 2011). Both small intestinal and colonic organoids can be

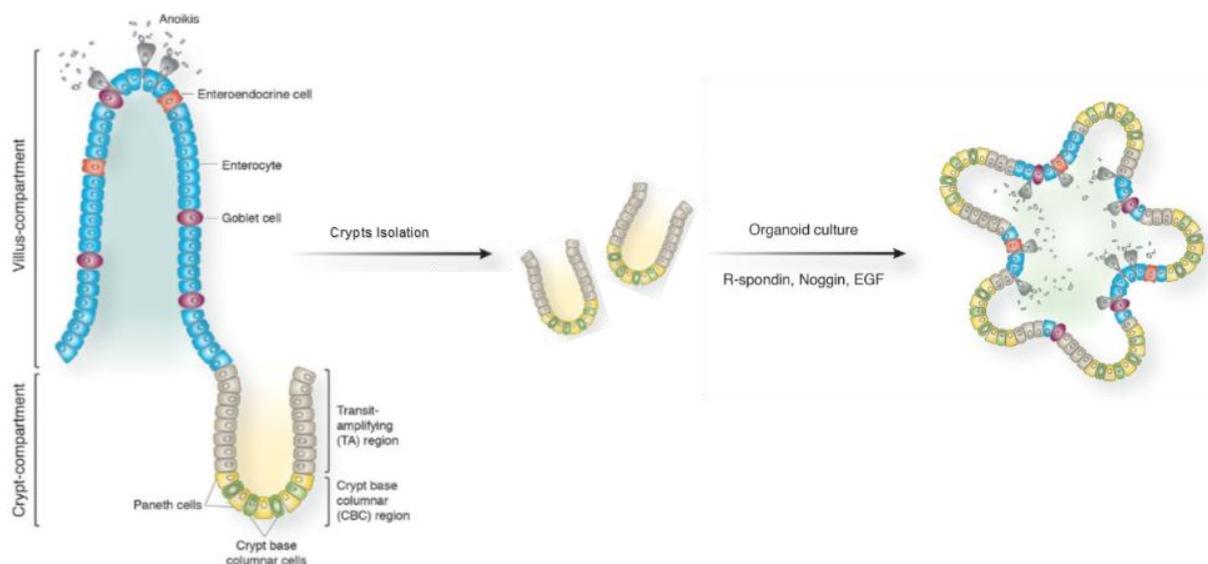
made by appropriate culture of intestinal stem cells present at the bottom of the intestinal crypt. These stem cells can give rise to a self-organising structure with similar crypt-villus architecture, diversity of cell types, and physiology to normal intestine. Importantly this organoid culture can also be expanded continuously (Sato *et al.*, 2009).

A nomenclature for organoid terminology has been proposed as follows (Stelzner *et al.*, 2012). The term organoid or sometimes 'reconstituted organoid' is suggested to be applied when the culture contains a combination of the epithelial cells (colon or small intestinal derived) with another cell type, e.g. fibroblasts, whereas a culture system which is derived from crypt or a single stem cell without a mesenchymal component is termed an enteroid (if derived from intestine) or colonoid (if derived from colon). However, in practice the generic term organoid is often applied to all of these different culture types.

As previously discussed, the intestinal crypt contains cycling LGR5+ stem cells that can give rise to all epithelial lineages (Barker *et al.*, 2007). Sato *et al.* (Sato *et al.*, 2009) found that a single crypt containing LGR5+ cells, or even a single sorted LGR5+ cell could give rise to an enteroid culture. The culture method of Sato *et al.* places isolated intestinal crypts or sorted LGR5+ cells into an extracellular matrix mixture called Matrigel and cultures them in media supplemented with growth factors which reproduce the niche signals produced by the bottom of the crypt. These factors are roof plate-specific spondin 1 (R-spondin1), epidermal growth factor (EGF), the bone morphogenetic protein inhibitor Noggin, and a Notch ligand. In addition, Sato *et al.* found that to increase survival of single cells, a ROCK inhibitor reagent is required to prevent anoikis (e.g Y-27632) (Sato *et al.*, 2009).

Colonoid culture has also been successfully developed by adapting the method for culturing enteroids (Sato *et al.*, 2011). However, to grow colonoids, Sato *et al.* found that Wnt3A is required to be added to the combination of growth factors, suggesting differences between the signals produced by the small intestine and colon stem cell niche. Moreover, long-term enteroid and colonoid culture has been found to require the addition of nicotinamide, as well as a small molecule inhibitor of Alk, and p38 inhibitor (Sato *et al.*, 2011). Figure 1.10 illustrates the method of enteroid culture system from isolated crypts.

Since enteroids/organoids resemble intestinal epithelium much better than cell line cultures, organoids are considered an excellent model for pharmacologic, immunologic, pathologic and genetic studies of the intestinal tract (Sato *et al.*, 2009). In the last few years, organoids have been widely used for both basic and clinical applications. These include investigations into the molecular regulation of intestinal stem cell self-renewal, proliferation, and differentiation (Sato and Clevers, 2013, Barker, 2014, Sato *et al.*, 2011), pathological mechanisms of intestinal epithelial dysfunction (Günther *et al.*, 2011), development of gene therapy, drug development and toxicity (Grabinger *et al.*, 2014, Ranga *et al.*, 2014), and many more (reviewed in (Meneses *et al.*, 2016)). However, this new technology is only just now being applied to studies of drug and xenobiotic metabolic enzymes. This year, Lu *et al* reported the first use of the organoid model to study the function of UGTs in the intestine (Lu *et al.*, 2017).



**Figure 1.10. Illustration of enteroid culture.**

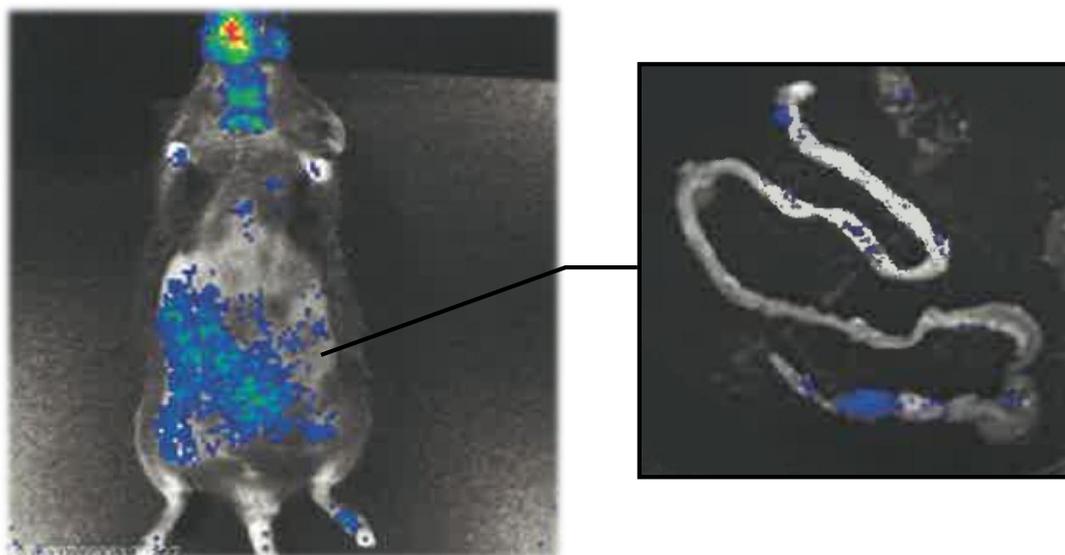
The diagram shows the isolation of crypts from the luminal surface of the intestinal wall. Upon culture these crypts grow through the self-renewal of stem cells and expansion of the TA population that gives rise to differentiated cell types. The structure usually forms a closed sphere with buds that represent the newly formed crypt-villus structures. The image is adapted from (Leushacke and Barker, 2014), with permission from BMJ Publishing Group Ltd.

### 1.5.3 'Humanised' UGT transgenic mouse models

As mentioned previously, studies of UGT function and regulation in mice are difficult because the *UGT* gene loci are not well conserved between humans and rodents. Likely mouse homologs can

only be identified for a few of the human *UGT* genes, and there are no clear homologues of the intestinal *UGT1A8-1A10* gene cluster (mice do have intestinal *Ugts*, but their sequence and function have diverged greatly during evolution). This limitation and the desire to study UGTs in a whole animal system has led to the development of so called 'humanised' mice that carry the human *UGT* genes. In particular, a mouse has been developed in which the mouse *Ugt1* locus has been deleted and the entire human *UGT1* gene locus has been inserted (Chen *et al.*, 2005). These mice allow the expression and functions of human UGT genes to be studied *in vivo* (Cai *et al.*, 2010).

Another approach to study *UGT* regulation *in vivo* is to generate transgenic mice carrying just the regulatory regions of relevant *UGT* genes. In our laboratory we previously developed a mouse that carries a 13kb segment of the *UGT1A8* promoter linked a luciferase reporter gene, thus enabling bioluminescence detection of promoter activity *in vivo*. This *UGT1A8* humanised mouse displays a distinct pattern of promoter activity in the mouse gastrointestinal tract, as shown in Figure 1.11.



**Figure 1.11. Representative images of human *UGT1A8* promoter activity identification in transgenic mouse by bioluminescence assay.**

The 'humanised' transgenic mice carrying 13k-*UGT1A8* promoter linked to luciferase reporter was developed previously. Real-time imaging of bioluminescence assay at exposure time of 15 minutes shows the presence of human *UGT1A8* promoter activity localised in the mouse abdomen. Isolation of intestinal tract showed expression of human-*UGT1A8* in the intestine with activation shown in the colon.

In this thesis project we originally proposed to use this *UGT1A8*-luciferase transgenic mouse model in a number of aspects of studies; however as described in Chapter 5, due to time constraints we reduced the scope of these plans. Ultimately, these mice were used to develop the enteroid/organoid model to provide a novel tool for the study of intestinal-specific UGT expression with specific focus on *UGT1A8* as described in detail in Chapter 5.

## **1.6 UGT1A8 is the prototypical intestinal-specific UGT gene**

### **1.6.1 UGT1A8 expression, activities and substrates**

As previously discussed, several *UGT* genes are expressed in the intestine, however of those only *UGT1A8* and *UGT1A10* show exclusively extrahepatic expression. The *UGT1A8* gene has become the model of choice to study intestinal-specific regulatory mechanisms of UGTs. Strassburg *et al*, initially isolated the *UGT1A8* cDNA from human colon tissue (Strassburg *et al.*, 1998). *UGT1A8* shares 94% sequence homology to *UGT1A10* and 90% to *UGT1A7* (Mojarrabi and Mackenzie, 1998). The expression of these UGTs vary from the proximal to the distal ends of the GI tract; *UGT1A10* is expressed throughout the GI tract and in the bile ducts; *UGT1A7* is expressed proximally in oesophagus and stomach, ; and *UGT1A8* is mainly expressed distally in the duodenum and colon (reviewed in (Gregory *et al.*, 2004b)).

The initial studies of Strassburg *et al* did not detect any activities for *UGT1A8*; However later studies, including from our laboratory, showed that *UGT1A8* has high catalytic activity with multiple substrates, mainly flavonoids and other phenolic substrates. *UGT1A8* substrates reported by Cheng *et al* are dietary phenols such as 7-hydroxyflavone, chrysin, apigenin, fisetin, quercetin, naringenin, the isoflavone genistein, coumarin rings of scopoletin, 4-hydroxycoumarin, umbelliferone, 4-methylumbelliferone, esculetin, phenolic drugs and xenobiotics such as *p*-nitrophenol, eugenol, 1-naphthol, propofol, carvacrol, 4-hydroxybiphenil, 2-hydroxybiphenyl and mycophenolic acid (MPA) (Cheng *et al.*, 1999).

Flavonoids, which are the largest subclass of polyphenols, are widely consumed by humans, with an estimated 428±49 mg of flavonoids consumed by European adults daily (Vogiatzoglou *et al.*,

2015). Many flavonoids have anti-cancer properties and are considered beneficial components of plant-rich diets. The high capacity of UGT1A8 in metabolising dietary phenols, especially flavonoids, contributes significantly to their low (around 5%) bioavailability in the body (Clifford, 2004). While UGT1A8 is responsible for eliminating flavonoids that may have anticancer activities, UGT1A8 may also play a role in cancer prevention by metabolising carcinogens as described below.

UGT1A8 conjugates carcinogenic metabolites of benzo(a)pyrene, 2-acetylaminofluorene (Mojarrabi and Mackenzie, 1998), primary amines of 2-aminobiphenyl, 4-aminobiphenyl, and the secondary amine of diphenylamine (Cheng *et al.*, 1998), a heterocyclic aromatic amine contained in cooked meats, 2-hydroxyamino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine (*N*-OH-PhIP) (Nowell *et al.*, 1999) and a toxic metabolite of antineoplastic agent irinotecan, 7-ethyl-10-hydroxycamptothecin (SN-38) (Gagné *et al.*, 2002). UGT1A8 is also active towards hydroxylated estrogens (including estrone, 2-hydroxyestrone, 4-hydroxyestrone, 2-hydroxyestradiol, 4-hydroxyestradiol, diethylstilbestrol, 17 $\alpha$ -ethynyl estradiol) and androgens (testosterone, dihydrotestosterone, 5 $\alpha$ -androstane-3 $\alpha$ , 17 $\beta$ -diol, epitestosterone) (Cheng *et al.*, 1998, Cheng *et al.*, 1999). Consistent to the pattern of UGT1A8 expression, glucuronidation of steroids is reported to occur significantly along the intestinal tract, with the highest activity at the end of the ileum and colon (Radomska-Pandya *et al.*, 1998). Some of the steroids conjugated by UGT1A8 also have carcinogenic properties such as 4-hydroxycatecholesterol (Lépine *et al.*, 2004), which is associated with breast and uterine carcinogenesis (Jefcoate *et al.*, 2000, Zhu and Conney, 1998)

UGT1A8 also conjugates a large number of clinically used drugs. Cheng *et al.* found that UGT1A8, and not UGT1A10, is capable of metabolising opioids (buprenorphine, morphine, naltrexone, naloxone, and nalorphine), bile acids, fatty acids, retinoids and drugs including ciprofibrate, furosemide, diflunisal, 4-hydroxytamoxifen, phenolphthalein, and hexafluoro-2-propanol (Cheng *et al.*, 1999). Triglitazone, an antidiabetic agent, is also metabolised by UGT1A8 (Watanabe *et al.*, 2002), and recently UGT1A8 was shown to play the major role in intestinal glucuronidation of a new antineoplastic agent in clinical trials OST167 (Ramírez *et al.*, 2015). Overall, these data

suggest that UGT1A8 has considerable relevance for metabolism of a subset of drugs in the intestine (Cheng *et al.*, 1999). Where the clearance effect is significant (Williams *et al.*, 2004), UGT1A8 polymorphisms might be clinically important in the modulation of therapeutic outcome.

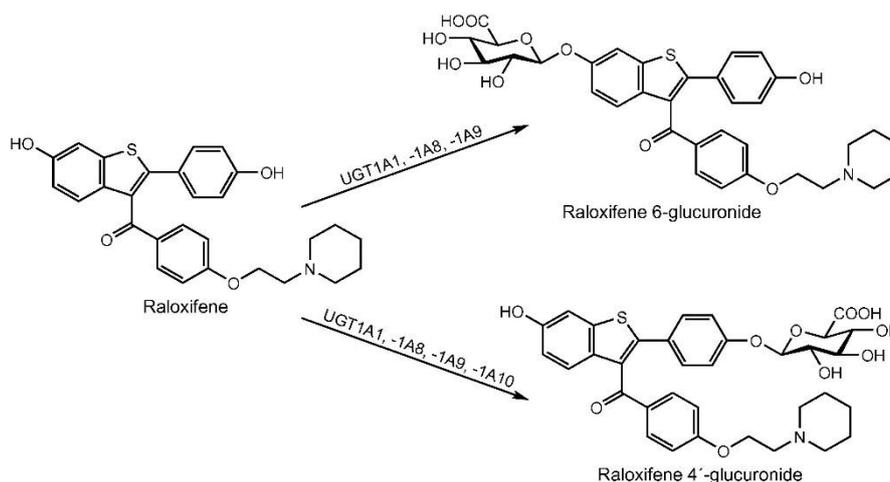
### 1.6.2 UGT1A8 polymorphisms in drug metabolism and cancer risk

Naturally occurring UGT1A8 genetic polymorphisms have been reported. Huang *et al.* (Huang *et al.*, 2002) identified four variants in Caucasian, African-American and Asian populations.

UGT1A8\*1 is the wild type allele. The UGT1A8\*2 and UGT1A8\*3 alleles create amino acid substitutions, and thus generate variant UGT1A8 proteins. The nucleotide change for UGT1A8\*2 is at position 518 (518 C>G), and creates a protein change at the position 173 (A<sup>173</sup>G), whereas the nucleotide change for UGT1A8\*3 is at position 830 (830 G>A), and creates a protein change at the position 277 (C<sup>277</sup>Y). The allelic frequencies showed that UGT1A8\*1a > UGT1A8\*2 > UGT1A8\*3. Indeed UGT1A8\*3 is the rarest variant with only 2.2% of the total population displaying this polymorphism in studies by (Huang *et al.*, 2002) and (Wang *et al.*, 2013).

An *in vitro* study found that UGT1A8\*3 has significantly lower activity than UGT1A8\*1 towards various substrates, namely 4-methylumbelliferone, octylgallate, hydroxy-2-acetylaminofluorenes and hydroxylbenzo[*a*]pyrenes. In contrast, UGT1A8\*2 showed activity comparable to UGT1A8\*1 (Huang *et al.*, 2002). The activity of the UGT1A8 allelic variants towards mycophenolic acid (MPA) has also been extensively studied (Bernard *et al.*, 2006, Bernard and Guillemette, 2004, Xie *et al.*, 2015, L'Aurette *et al.*, 2008, Levesque *et al.*, 2007). MPA is an active substance of mycophenolate mofetil (MMF), an approved immunosuppressant used to prevent rejection after organ transplantation. In the body, MPA is rapidly transformed by glucuronidation to inactive and readily excreted 7-O-MPA-glucuronide (MPAG) (Shaw and Nowak, 1995). Extrahepatic metabolism of MPA to MPAG is performed by UGT1A8 (Mackenzie, 2000, Cheng *et al.*, 1999). An *in vitro* study showed that UGT1A8\*3 has substantially decreased MPAG formation activity along with later discovered variants \*5, \*7, \*8, and \*9. Those variants occurred in 2.8% of Caucasians and 4.8% of African Americans. Modest reduction in MPA glucuronidation was also found to be caused by the UGT1A8\*2 variant (Bernard and Guillemette, 2004, Bernard *et al.*, 2006) and this may be relevant in renal transplant patients (Xie *et al.*, 2015).

The second generation selective estrogen receptor modulator (SERM), raloxifene, has low bioavailability due to extensive metabolism in the intestine by UGT1A1, UGT1A8, UGT1A9 and UGT1A10 (Figure 1.12) (Kemp *et al.*, 2002). The role of UGT1A8 in raloxifene glucuronidation is more prominent than that of the intestinal specific UGT1A10 (Mizuma, 2009), even though UGT1A8 is less abundant than UGT1A10 in the intestine (Ohno and Nakajin, 2009). An *in vitro* study showed that, compared to wild-type UGT1A8, raloxifene glucuronidation by UGT1A8\*3 to both its metabolites (6- and 4'-glucuronide) is significantly lower. UGT1A8\*2 also showed lower activity in raloxifene 4'-glucuronidation, (Kokawa *et al.*, 2013). These findings have been confirmed by another study (Sun *et al.*, 2013).



**Figure 1.12 Raloxifene glucuronidation by UGT enzymes.**

Image is reproduced from (Kokawa *et al.*, 2013) with permission from Elsevier.

As well as drug disposition, UGT1A8 polymorphisms are involved in cancer risk. Thibaudeau *et al* reported that UGT1A8\*3 is associated with a 12 fold reduction in carcinogenic 4-hydroxy-catecholestrogen metabolism (Thibaudeau *et al.*, 2006). Their findings suggest that the presence of UGT1A8 variants may result in a differential response to carcinogenic exposure, and may modulate the risk of cancer. Consistent with this idea, Wang *et al* found a high incidence of UGT1A8 polymorphism in colorectal cancer patients and established that the UGT1A8\*3 allele is a risk factor for colorectal cancer (Wang *et al.*, 2013). In addition, in a case-controlled study, high

activity UGT1A8 variants were found less frequently in patients with oesophageal squamous cell carcinoma (Dura *et al.*, 2012). Overall, these studies suggest that UGT1A8 plays a protective role against carcinogenesis.

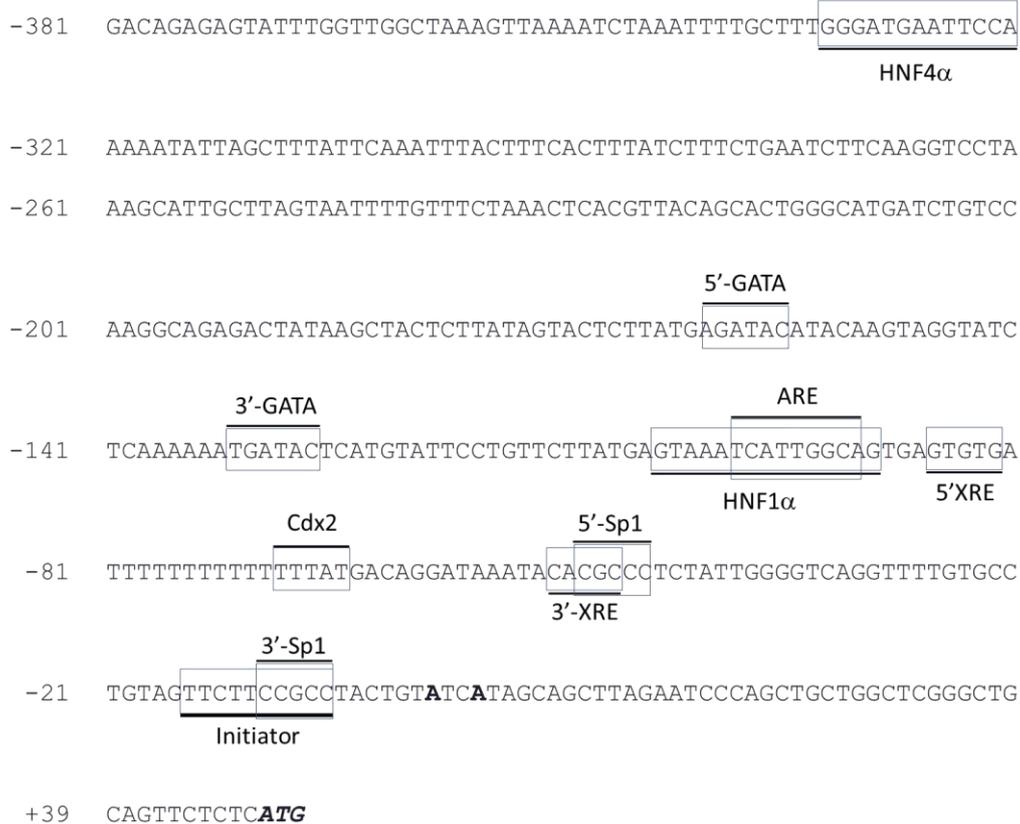
### 1.6.3 UGT1A8 gene regulation

A number of studies have examined the constitutive and inducible regulation of UGT1A8 expression by tissue-specific and ligand activated transcription factors. Kalthoff *et al.* showed that coffee, a flavonoid rich beverage, is able to induce endogenous UGT1A8 expression significantly in Caco-2 and KYSE70 cells (Kalthoff *et al.*, 2010a). The mechanism of action appeared to involve the ligand activated transcription factors AhR (aryl hydrocarbon receptor) and Nrf2 (nuclear factor erythroid-related factor 2) acting via Xenobiotic (XRE) and antioxidant (ARE) response elements in the proximal *UGT1A8* promoter (Kalthoff *et al.*, 2010a) (Figure 1.13). Another study by the same group showed synergistic transcriptional regulation of UGT1A8 by AhR and Nrf2 (Kalthoff *et al.*, 2010b) in response to 2,3,7,8-tetrachlordibenzo-p-dioxin (TCDD) (a highly toxic organic pollutant) (Pelclová *et al.*, 2006), and *tert*-butylhydroquinone (tBHQ) (a potential carcinogen of synthetic antioxidant) (Gharavi and El-Kadi, 2005). These studies suggest that UGT1A8 may be induced by its own substrates as part of a feedback response.

Studies of the constitutive regulation of UGT1A8 have been performed mainly by our laboratory. Our group has characterized the proximal promoter structure, as well as distal regulatory regions/enhancers (reviewed in (Hu *et al.*, 2014a)). The transcription start site (Uhlén *et al.*) in the *UGT1A8* promoter was initially thought to occur in a series of 14 consecutive thymidines (T-box) located at approximately 120 bp from the initiator codon, which act as a TATA Box (Gong *et al.*, 2001). However, promoter analysis by Gregory *et al* reported that *UGT1A8* promoter activity is driven not by the T-box, but instead by an overlapping Sp1/ initiator-like region which binds Specificity protein 1 (Sp1) and initiator-like protein and transcription starts at -48 bp from the ATG translation initiation codon. In Caco-2 cells, where the *UGT1A8* promoter is highly active, mutating the Sp1 site decreased the promoter activity to near background levels (Gregory *et al.*, 2004b). Mutation of the T-box showed that it does not function as TSS, and a similar result was found the

T-box was mutated in the highly conserved *UGT1A9* and *UGT1A10* promoters (Mackenzie *et al.*, 2005b).

Studies by Gregory *et al* (Gregory *et al.*, 2004a) identified adjacent binding sites for CDX2 and HNF1 $\alpha$  within the proximal *UGT1A8* promoter. They used promoter-reporter analysis to demonstrate that constitutive activity of the *UGT1A8* promoter is regulated by CDX2 in Caco-2 cells. Moreover, they showed that CDX2 also controls expression of other intestinal UGTs, such as *UGT1A10* and *UGT2B7* (Gregory *et al.*, 2004b, Gregory *et al.*, 2004a). Consistent with their adjacent binding sites, CDX2 and HNF1 $\alpha$  were found to co-regulate the *UGT1A8* promoter. In particular, when both CDX2 and HNF1 $\alpha$  binding sites were mutated the promoter activity was decreased in Caco-2 cells, whereas mutation of only one site was able to be compensated for by the other. Overexpression of CDX2 or HNF1 $\alpha$  alone induced *UGT1A8* promoter activity, and when both factors were simultaneously overexpressed, a greater than additive effect was observed. This synergistic action required the presence of the HNF1 $\alpha$  binding site but not the CDX2 site, suggesting that HNF1 $\alpha$  might bind and recruit CDX2 to the *UGT1A8* promoter (Gregory *et al.*, 2004a). Furthermore, Mackenzie *et al* (Mackenzie *et al.*, 2005b) reported that two polymorphisms in the CDX2 binding element in the *UGT1A8* promoter that were detected in colon tumour tissue (Wicking *et al.*, 1998) did not alter the ability of CDX2 to regulate the *UGT1A8* promoter (Mackenzie *et al.*, 2005b). The physical interaction between Cdx2 and HNF1 $\alpha$  has been previously reported (MITCHELMORE *et al.*, 2000) supporting a co-recruitment model to explain the apparent redundancy of the CDX2 binding site for regulation of *UGT1A8* by CDX2.



**Figure 1.13. Identified transcription factors (TFs) and their cognate cis-regulatory elements (CREs) in the proximal promoter of *UGT1A8*.**

The translation initiation ATG codon is 48 base pairs downstream of the transcription start site. ARE, antioxidant response element; Cdx2, caudal-related homeodomain protein 2; HNF1 $\alpha$ , hepatocyte nuclear factor 1 alpha; HNF4 $\alpha$ , hepatocyte nuclear factor 4 alpha; Sp1, the specificity protein 1; XRE, xenobiotic response element.

The physical interaction of CDX2 and HNF1 $\alpha$ , which also involves interaction with GATA factors, has been reported as playing a role in regulating other intestinal expressed genes, such as claudin-2 (CLDN2), sucrase-isomaltase, and lactase-phlorizin hydrolase (LPH) (Boudreau *et al.*, 2002, MITCHELMORE *et al.*, 2000). Furthermore, transcriptional regulation by a combinatorial mechanism of Cdx2, GATA and HNF4 $\alpha$  is reported to maintain intestinal homeostasis in mouse (San Roman *et al.*, 2015). This group of intestinal expressed transcription factors potentially assemble at tissue-specific cis-regulatory sites where they may physically interact to regulate intestinal-specific gene expression.

As shown in Figure 1.13, the *UGT1A8* proximal promoter contains GATA binding sites adjacent to Cdx2 and HNF1 $\alpha$  binding sites (Gregory, 2004). In the study by (Gregory, 2004), the GATA factors, namely GATA-4, -5 and -6, were reported as being able to bind to their cognate sites in the *UGT1A8* promoter. However, none of the GATA factors tested in their study were able to activate the promoter, either alone or in combination with Cdx2 or HNF1 $\alpha$ . This study suggested that the presence of GATA sites may not be required for *UGT1A8* promoter activation, nor do they facilitate interaction with HNF1 $\alpha$  or Cdx2 under the examined conditions.

As discussed previously, HNF4 $\alpha$  was shown to physically interact with Cdx2 in mouse intestine *in vivo* (Verzi *et al.*, 2013), and combinatorial activities of HNF4 $\alpha$  and Cdx2 were shown to regulate cell maturation and a cohort of functional enterocyte genes (San Roman *et al.*, 2015). To date the role of HNF4 $\alpha$  in regulating intestinal UGT expression has not been studied, although it has been studied in hepatic expression. Using promoter analysis in HepG2 cells, Barbier *et al* (Barbier *et al.*, 2005) demonstrated that HNF4 $\alpha$  plays a key role in regulating *UGT1A9* expression in the liver. Moreover, despite high sequence similarity of the *UGT1A7-UGT1A10* promoters, minor base-pair differences in the analogous HNF4 $\alpha$  binding sites in the *UGT1A7*, -1A8 and -1A10 promoters renders them non-functional, which may provide an explanation for the absence of expression of these genes in the liver. It is notable that the HNF4 $\alpha$  binding site(s) that are involved in hepatic expression of *UGT1A9*, and that were shown to be non-functional in the *UGT1A7*, -1A8 and -1A10 promoters are located at a considerable distance from the CDX2 and HNF1 $\alpha$  binding elements in the proximal promoter (Figure 1.13). It is possible that more proximal HNF4 $\alpha$  binding elements that function in the intestine are yet to be discovered and this possibility was addressed in this thesis project.

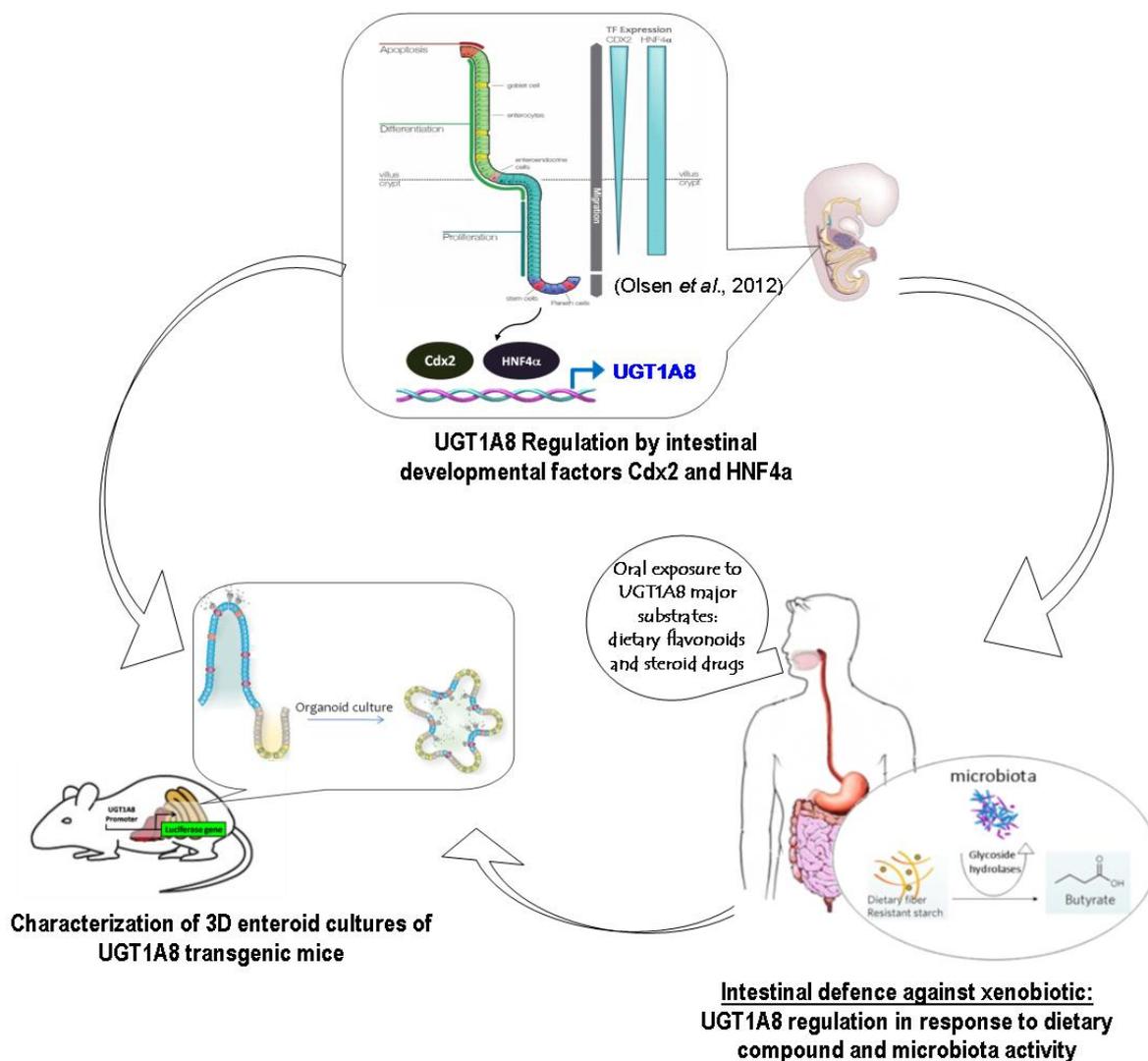
## 1.7 Project overview and experimental aims

Current knowledge suggests that CDX2 is involved in controlling intestinal *UGT* expression consistent with its role in intestinal development and the maintenance of intestinal adult epithelium particularly through control of cell differentiation. Currently however, we do not know whether HNF4 $\alpha$  (and the coordinated activities of Cdx2 with HNF4 $\alpha$ ) direct intestinal UGT expression as they do the expression of other intestinal genes. In addition, it remains unclear how *UGT1A9* expression is directed to the intestine in the absence of any known functional CDX2 binding site in its promoter.

*UGT1A8* is possibly the most important UGT for the metabolism of endo- and xenobiotic compounds in the GI tract, particularly in the lower end of the intestinal tract, and a high level of *UGT1A8* may prevent colorectal carcinogenesis. The induction of *UGT1A8* expression may be controlled by its own substrates such as dietary flavonoids and steroid compounds; however, the range of substrates that can induce *UGT1A8*, and the mechanisms of their action are not well characterised. In addition, it is unknown how the activity of intestinal microbiota might influence *UGT1A8* expression.

Studies of *UGT* gene regulation and function would be advanced by the application of models that represent the intestine better than cancer cell lines such as Caco-2. Models such as intestinal 3-D organoid culture and 'humanized' transgenic mice provide such an opportunity but have not been well developed in terms of intestinal UGTs.

These considerations have led to our aims presented below and illustrated schematically in Figure 1.14.



**Figure 1.14. Schematic of the project.**

The project integrates studies of constitutive/developmental regulation of intestinal UGTs with studies of their inducible regulation by dietary chemicals focussed on *UGT1A8*. To facilitate future extension of these studies, an organoid culture model is developed using *UGT1A8* transgenic mice.

Project aims:

1. To investigate the potential interplay of intestinal developmental factors CDX2 and HNF4α in the transcriptional regulation of the intestinally expressed *UGT1A8* and *-1A10* genes and the intestinal/hepatic expressed *UGT1A9* gene.
2. To identify flavonoids and steroids that can induce *UGT1A8* expression and determine their mechanism of regulation, including their interaction with the microbial metabolite butyrate.
3. To establish an organoid model from the intestines of *UGT1A8* transgenic mice and develop protocols to manipulate gene expression in this model.

## CHAPTER 2. MATERIALS AND GENERAL METHODS

This chapter presents general methods and materials that were used in all experimental chapters of this thesis (Chapter 3, 4 and 5). Each experimental chapter also includes a sub-section of Methods that were specific to the studies therein; these specific methods subsections are 3.2, 4.2 and 5.2.

### 2.1 Materials

#### 2.1.1 Chemicals, reagents and buffers

The standard purity grade of chemicals used in experimental procedures of this research project was analytical grade. The list of chemicals and reagents and suppliers are categorised and presented in Appendix 1. Preparations of buffers used during this study are listed in Appendix 2.

#### 2.1.2 Mammalian cell lines

The Caco-2 (ATCC HTB-37), HepG2 and COS-7 cell lines were obtained from The American Type Culture Collection (Manassas, VA). HCT-116, SW-480 cell lines were obtained from Dr. Dong Gui Hu (Department of Clinical Pharmacology, Flinders University), and HT-29 was obtained from Dr. Michael Michael (Flinders Centre for Innovation in Cancer, South Australia).

#### 2.1.3 Expression vectors and reporters

The details of expression vectors used in this study, including generation of specific plasmid constructs, are described in each Chapter of this thesis. Control vector pCMX was purchased from Addgene and pcDNA3 from Invitrogen.

The promoter-reporter constructs containing *UGT1A8*, *-1A9* and *-1A10* promoter regions linked to the firefly luciferase reporter gene in pGL3-Basic were made previously in our laboratory by Dr. Philip Gregory. The control reporter vectors pGL3-Basic and pRL-null (expressing *Renilla* luciferase) were purchased from Promega.

#### **2.1.4 Oligonucleotides**

Oligonucleotides used in this study are listed in each Chapter of this thesis. Oligonucleotide primers and siRNAs were from the following suppliers: Geneworks (Hindmarsh, SA, Australia), Integrated DNA Technology (IDT, Coralville, Iowa, USA), Sigma-Aldrich Australia (Merck KGaA, Darmstadt, Germany), GenePharma (Shanghai, China) and Macrogen (Seoul, Korea). All primers were of standard purification quality (desalted).

## **2.2 Methods**

### **2.2.1 Mammalian cell culture, maintenance and stocks**

The cell lines used throughout this study were cultured in complete high-glucose Dulbecco's modified Eagle's medium (DMEM); complete DMEM media was prepared by adding 1  $\mu$ M sodium pyruvate, and supplemented with 1% non-essential amino acids and 10% foetal bovine serum (FBS). The cell culture was maintained in a humidified incubator at 5% CO<sub>2</sub>, at 37°C. Carryover cells were cultured in T75 flasks; media were replaced every two days between subculturing. When reaching ~80% confluency, the cells were ready for either subculturing or harvesting for experiments.

### **2.2.2 Bacterial culture and stocks**

The bacterial strain used for cloning purposes was *E. coli* DH5 $\alpha$ <sup>TM</sup>. Bacteria were either cultured on Luria-Bertani (LB) agar or in LB broth containing the appropriate antibiotic for selection of bacteria carrying transformed plasmids (100  $\mu$ g/mL ampicillin or 30  $\mu$ g/mL kanamycin). The agar cultures were incubated for 16-20 hours in a Shimaden (Scientific Equipment Manufacturer) incubator at 37°C. The broth cultures were grown in 50-100 mL volume in an Innova 4330 incubator shaker at 37°C and 220 rpm.

### **2.2.3 Cryopreservation of cells (freezing stocks)**

#### **2.2.3.1 Freezing of mammalian cell stocks**

Mammalian cell line stocks were made to maintain the availability of early passage cell lines for later use. To prepare the stocks, cells were harvested by trypsinisation and resuspended in complete DMEM media. Cells were then pelleted by centrifugation for 5 minutes at 1000 rpm at

room temperature (RT) followed by resuspension in FBS containing 10% dimethylsulfoxide (DMSO), transferred into Nunc cryotube vials and stored at -80°C for temporary storage, or in liquid nitrogen for long-term storage. To recover cells, vials were thawed rapidly in a 37°C water bath for less than 1 minute, to avoid a toxic effect of DMSO on the cells. Prewarmed complete DMEM media was used to dilute the thawed cells and they were then transferred to T25 or T75 flasks. After 24 hours, the media was replaced to remove residual DMSO.

### **2.2.3.2 Bacterial glycerol stocks**

Bacterial stocks were prepared by combining 250 µL of bacterial broth culture with 250 µL of 50% glycerol and stored at -80°C. To ensure the viability of stocks was maintained; upon use the frozen tubes were kept on ice to minimize thawing.

### **2.2.4 Polymerase chain reaction (PCR) amplification**

The PCR products for qualitative analytical or preparative purposes were generated by amplification using a BioRad iCycler™ thermal cycler (Waukegan, IL, USA), whereas the quantitative real-time PCR (qRT-PCR) reactions were performed using a Rotorgene™ 3000 (Corbett Life Science, Mortlake, NSW, Australia).

#### **2.2.4.1 PCR for plasmid DNA construction and screening**

Preparative polymerase chain reaction (PCR) was performed to generate DNA fragments that were ligated into plasmid vectors. These reactions were performed using Phusion Hot Start II High-Fidelity PCR Master Mix (Thermo Fisher Scientific) according to instructions in the manual. Analytical PCR was performed to screen bacterial colonies for the presence of correct plasmid transformants using Phire Hot Start enzyme according to the manufacturer's protocols (Thermo Fisher Scientific). Optimal PCR annealing temperature was calculated based on the primer melting temperature ( $T_m$ ).

#### **2.2.4.2 Site-directed mutagenesis**

Mutations were introduced into target regulatory elements within the various promoter-reporter constructs via the *in vitro* site-directed mutagenesis (SDM) method using the QuikChange® Site-Directed Mutagenesis System with *PfuTurbo*® polymerase (Stratagene, La Jolla, CA) and primers

containing the desired mutation. Primers used for site-directed mutagenesis are listed in the relevant Chapter of this thesis. All primers for SDM were purified by polyacrylamide gel electrophoresis (PAGE) prior to use to increase mutation efficiency. The SDM amplification reaction was performed using the following cycling parameters: 1 cycle of 30 seconds at 95°C for initial denaturation; 18 cycles of 30 second denaturation at 95°C, 1 minute annealing at 55°C and 7 minutes extension at 68°C. The PCR products were then incubated with DpnI restriction enzyme for 2 hours to digest the plasmid template and transformed into DH5 $\alpha$  *E. coli* competent cells. Colonies were selected and cultured overnight in LB media before preparation of plasmid DNA as in section 2.2.10. Mutant plasmids were identified by Sanger sequencing as in section 2.2.12.

### **2.2.4.3 Quantitative reverse transcriptase PCR (qRT-PCR)**

The qRT-PCR analysis was performed to assess the level of mRNA expression. The first step in qRT-PCR was cDNA synthesis by reverse transcription (RT) as described in section 2.2.6. For the PCR step, 2  $\mu$ L of cDNA was used in a 20 $\mu$ L reaction containing 1 x GoTaq PCR Master Mix (Promega) and 0.5  $\mu$ M each of the forward and reverse primers for relevant target gene. The qRT-PCR was performed using a RotorGene 3000 instrument (Corbett Research, NSW, Australia) under the following cycling conditions: 95°C for 15 min, followed by 40 cycles of 10 s at 95°C, 15 s at 56°C to 60°C, and 20 s at 72°C. qRT-PCR primers that were used in this study are listed in the corresponding Chapters. Each sample for qRT-PCR analysis was amplified in duplicate. The relative abundance of the target mRNA (i.e. UGT1A8, -1A9, -1A10, CDX2 and HNF4 $\alpha$ ) was calculated using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001) and normalised to the reference gene 18s rRNA. For induction or inhibition studies, the expression data was generally presented relative to the control (mock treated) condition with the Student's t-test used to assess significance.

## **2.2.5 Total RNA extraction**

### **2.2.5.1 RNA extraction from monolayer cells**

Total RNA was extracted from cells using TRIzol® Reagent. One mL of TRIzol® Reagent was added directly to the cell culture well containing  $1 \times 10^5$  cells and incubated at room temperature for 5 minutes to allow complete cell dissociation (lysis). The cell lysates were transferred into

microtubes and 200  $\mu\text{L}$  chloroform was added. RNA in the reaction mixture was partitioned by centrifugation at 12,000 x g for 15 minutes at 4°C. The upper aqueous phase produced after centrifugation was transferred into a new tube and the RNA precipitated by adding 500  $\mu\text{L}$  isopropanol, and incubated for 10 minutes at room temperature. To pellet the RNA, the tubes were centrifuged at 12,000 x g at 4°C. RNA pellets were air-dried and resuspended in 20  $\mu\text{L}$  nuclease-free water. All RNA samples were kept at -80°C for long-term storage.

### **2.2.5.2 RNA extraction from tissues**

Total RNA from UGT1A8-transgenic mouse tissues were harvested using TRIzol® Reagent with the same methods as described previously; however, mouse tissue samples were prepared by homogenisation using a micro-pestle prior to lysis with TRIzol® Reagent.

### **2.2.6 cDNA synthesis**

To generate cDNA, 2  $\mu\text{g}$  of RNA was treated with DNaseI and incubated at 37°C in a thermal block. Following the reaction, EDTA at a final concentration of 2.5 mM was added and the reaction was heated to 75°C for 5 minutes to inactivate the DNaseI. An 8  $\mu\text{L}$  aliquot of the DNaseI-reaction (containing 1  $\mu\text{g}$  of RNA) was combined with 1  $\mu\text{L}$  of 10 mM dNTPs and 1  $\mu\text{L}$  of 53 ng/ $\mu\text{L}$  random primers (New England Biolabs) and incubated at 65°C for 5 minutes, then placed on ice for 2 minutes. The cDNA was synthesised by adding 2  $\mu\text{L}$  10 x reverse transcriptase buffer, 50 units of NxGen M-MuLV RT (Lucigen) and RNase inhibitor (Lucigen) and nuclease free water up to 20  $\mu\text{L}$ . The reaction was incubated at 42°C for 60 minutes and then at 90°C for 10 minutes. The resulting cDNA was diluted 1:5 in nuclease free water (Promega).

### **2.2.7 Quantification of nucleic acid (DNA or RNA)**

The concentration of DNA or RNA was measured using a NanoDrop™ 2000 (Thermo Fisher Scientific) spectrophotometer using the manufacturer's protocol. One  $\mu\text{L}$  of DNA or RNA sample was placed onto the pedestal and absorbance measured at 260 nm. Concentration was calculated using the formula: dsDNA concentration = 50  $\mu\text{g}/\text{mL}$   $\times$  OD260  $\times$  dilution factor; RNA concentration = 40  $\mu\text{g}/\text{mL}$   $\times$  OD260  $\times$  dilution factor. Purity was assessed via the 260/280 ratio (nucleic acid/protein) with 1.8 – 2.0 considered to represent good quality.

## **2.2.8 Molecular cloning using restriction enzymes**

### **2.2.8.1 Restriction digestion**

Restriction enzymes used for cloning in this study were all purchased from New England Biolabs (Mackenzie *et al.*). Restriction digestion of vector or insert DNA (including PCR products – see section 2.2.4.1) was performed in a 30-50  $\mu\text{L}$  digestion reaction containing 1-3  $\mu\text{g}$  DNA, 1  $\mu\text{L}$  of each restriction enzyme, 1x appropriate enzyme buffer in nuclease-free water. The reaction was incubated at 37°C for 1 hour to overnight. Subsequently the digested DNA fragments were purified either using the QIAquick PCR purification kit or by gel extraction protocol (QIAGEN).

### **2.2.8.2 Ligation**

Cloning of insert DNA into the vector was performed using the Quick Ligation™ Kit (Mackenzie *et al.*) following the manufacturer's protocol. Briefly, 20-100 ng of the purified insert and vector DNA were combined with 1  $\mu\text{L}$  of Quick Ligase and 5  $\mu\text{L}$  of 2x Quick Ligase Reaction Buffer and nuclease free water to 10  $\mu\text{L}$  final volume. The reaction was incubated for 5 minutes at room temperature (25°C), and chilled on ice for 2-5 minutes. The ligation products (2  $\mu\text{L}$ ) were then transformed into DH5 $\alpha$  *E. coli* competent cells or stored at -20°C until ready for transformation.

### **2.2.8.3 Transformation**

The DH5 $\alpha$  *E. coli* competent cells were used to transform plasmids including molecular cloning ligation products. Fifty  $\mu\text{L}$  of competent cells were thawed and kept on ice, and 0.1-5 ng of plasmid (or 2  $\mu\text{L}$  of ligation mixture) was added to the competent cells and mixed gently by flicking the tube, and then incubated on ice for 30 minutes. The cells were heat-shocked at 42°C for 30 seconds and returned to ice for 2 minutes. LB media containing no antibiotic, or super optimal (SOC) media (950  $\mu\text{L}$ ) was added to the transformation mixture and incubated in a shaker incubator at 37°C for 1 hour. Fifty to 200  $\mu\text{L}$  of the transformation mixture was spread onto a pre-warmed LB-agar plate with appropriate antibiotic. The plate was incubated for 16 to 20 hours in a 37°C incubator to allow the growth of bacterial colonies.

## **2.2.9 Competent cell preparation**

The protocol to prepare DH5 $\alpha$  *E. coli* competent cells was modified from (Hanahan *et al.*, 1991).

Briefly, 1 mL of overnight DH5 $\alpha$  culture was inoculated into 100 mL LB media without any antibiotic, and incubated in a shaker incubator at 37°C until the culture density reached an OD<sub>600</sub> of 0.25 – 0.3. The culture was transferred into falcon tubes and centrifuged at 3,000 x g at 4°C for 10 minutes and the bacterial pellet resuspended in 32 mL of cold CCMB80 buffer (80 mL CaCl<sub>2</sub>, 20 mL MnCl<sub>2</sub>, 10 mM MgCl<sub>2</sub>, 10 mM KOAc pH 7.0, 10% glycerol, pH was adjusted to 6.4). The tubes were centrifuged for 10 minutes and the pellet resuspended in 4 mL of cold CCMB80 buffer. In the cold room with microtubes kept on ice, the 4 mL DH5 $\alpha$  in CCMB80 was divided into 50  $\mu$ L aliquots and stored at -80°C until required.

### **2.2.10 Plasmid DNA extraction**

The plasmid DNA from bacterial cultures was isolated using a commercial extraction kit from QIAGEN. For small scale bacterial culture of 1-5 mL, QIAprep Spin Miniprep Kit was used, whereas QIAGEN® Plasmid midi kit was used for a large scale bacterial cultures of up to 100 mL. The protocols were carried out as per the manufacturer's instruction manual. The concentration of isolated plasmids was measured using NanoDrop™ 2000 spectrophotometer.

### **2.2.11 Agarose gel electrophoresis**

Agarose gel electrophoresis was performed to analyse products of PCR amplification, restriction digest or plasmid preparation. Agarose gels (at 1-2%) were made in TAE buffer, and ethidium bromide as added to a final concentration of 0.2 – 0.5  $\mu$ g/mL (2-3  $\mu$ L of the lab stock) to allow visualisation of DNA under ultra violet (UV) light. The DNA samples were mixed with glycerol or dye containing glycerol prior to loading them on to the gel. Standard DNA ladders (Mackenzie *et al.*) in the 1kilo base pair (bp) or 100 bp range were used to track the size of the DNA samples. Electrophoresis was performed at 80-120 V, at constant ampere in a Bio-Rad Mini-Sub Gel GT electrophoresis system. The DNA in the gel was visualized and documented in a GeneGenius bio-imaging system apparatus (Syngene, Cambridge, England) using GeneSnap version 6.04 software. For isolating DNA from the gel, QIAquick Gel Extraction kit was used according to the manufacturer's instructions.

## **2.2.12 DNA Sequencing**

DNA sequencing was performed to confirm the sequence of inserted DNA following plasmid construction or to confirm the occurrence of mutations following SDM. Sequencing procedures were conducted by DNA Sequencing Service Facility (SA Pathology, Flinders Medical Centre, South Australia) using an ABI 3130xl Genetic Analyser Sequencer and Big Dye Terminator Cycle Sequencing Version 3.1 Chemistry (Applied Biosystem, Foster City, CA). The National Center for Biotechnology Information (NCBI) sequence nucleotide BLAST was used to analyse the sequencing data.

## **2.2.13 Plasmid-DNA transfection**

### **2.2.13.1 *Lipofectamine® 2000 protocol***

Transient transfections were performed in various cell lines including Caco-2, HT-29 and HepG2 cells to study UGT gene expression and regulation. Caco-2 cells were used most extensively. Transfections of luciferase promoter-reporter vectors together with various expression vectors (effectors) were performed in a 48-well plate format. For each well,  $7.5 \times 10^4$  cells were co-transfected with 250 ng promoter-reporter constructs (pGL3-Basic, wild type or mutant UGT promoter: UGT1A8, -1A9 or -1A10), 5 ng of internal control pRL-null vector and 250 ng of each effector construct (e.g. empty vector pCMX, CDX2, HNF4 $\alpha$ ). Plasmids were diluted in serum free media and incubated with 1.25  $\mu$ L Lipofectamine® 2000 reagent (ratio of DNA: Lipofectamine was 1:2.5). The total amount of DNA was kept constant at 0.5  $\mu$ g per well using additional pCMX plasmid as 'filler'. All transfections were carried out using the reverse transfection method by mixing the cell and pre-incubated DNA-Lipofectamine® 2000 complex in the tube before plating. Incubation time and reaction volume were as per the manufacturer's protocol. The cells were harvested 48 hours post-transfection in 1x Passive Lysis Buffer (Promega) for subsequent luciferase assays.

Transfection of siRNAs was performed in 48-well plates using the same number of cells. Five pmol of siRNA (CDX2-siRNA, HNF4 $\alpha$ -siRNA or negative control siRNA) diluted in serum free media was pre-incubated with 1.25  $\mu$ L Lipofectamine® 2000 and combined with the media containing cells

and plated into 48-well plates. After 24 hours, transfection was repeated by removing media and replacing with fresh media containing siRNA-Lipofectamine® 2000 complex. Forty-eight hours after the first transfection, cells were harvested in TRIzol® Reagent for preparation of total RNA and subsequent qRT-PCR analysis of target gene expression. All experiments were performed in duplicate or triplicate and repeated independently three to nine times.

#### **2.2.13.2 Lipofectamine® LTX protocol**

Lipofectamine® LTX was also used for transient transfection of cells with expression plasmids, particularly when analysis of endogenous target gene expression was to be performed, because of its better efficiency relative to Lipofectamine® 2000. The cell number, DNA amount and well plate format were as described above for the Lipofectamine® 2000 protocol. Transfection was carried out as per manufacturer's instructions. Briefly, a total of 1 µg DNA was diluted into 100 µL of serum free media and 1 µL PLUS™ reagent (1:1 ratio DNA to PLUS) was added directly to the DNA dilution, and the mixture was incubated for 15 minutes at room temperature. Following incubation, 2.5 µL Lipofectamine® LTX was added to the diluted DNA mixture and incubated for an additional 30 minutes to allow formation of DNA-Lipofectamine® LTX complexes. The mixture was added to 400 µL cell suspension and mixed by inverting the tube gently several times, before plating into the well of a 48-well plate. Cell culture was incubated in a CO<sub>2</sub> incubator at 37°C for 48 hours without changing the media before harvest for gene expression analysis.

#### **2.2.14 Luciferase reporter assays**

Transfected cells in 48-well plates were washed once with phosphate-buffered saline (containing 137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, in a pH of 7.4) and lysed by adding 75 µL of 1 x Passive Lysis Buffer (Promega), followed by gentle rocking for 1 hour at room temperature. The lysate was then either be stored at -20°C or directly assayed. To measure the firefly and Renilla luciferase activities, 20 µL of lysate was assayed using the Dual-Luciferase Reporter Assay System and a Packard TopCount luminescence and scintillation counter (PerkinElmer Life and Analytical Sciences, Waltham, MA) as per manufacturers' instructions. The activities of promoter-reporter vectors were normalised to that of empty pGL3-Basic vector. Induction or inhibition of

promoter activities as a result of co-transfection with expression vectors was represented as fold-change relative to co-transfection with empty expression vector (i.e. pCMX). A two-tailed Student's t-test used to determine statistical significance

### **2.2.15 Chromatin Immunoprecipitation (ChIP) assays**

ChIP was used to detect HNF4 $\alpha$  binding to the *UGT1A8-1A10* proximal promoters. The protocol was adapted from the  $\mu$ ChIP method by (Dahl and Collas, 2008). Briefly,  $7.5 \times 10^6$  Caco-2 cells were reverse transfected with 20  $\mu$ g HNF4 $\alpha$  and pCMX control using 100  $\mu$ L Lipofectamine<sup>®</sup> 2000 reagent and plated in T75 flasks. Forty-eight hours post-transfection, formaldehyde was added to media to 1% final concentration and incubated for 30 minutes at room temperature to crosslink interacting proteins and DNA in the cells. Subsequently 125 mM glycine was added to stop the reaction. Cells were collected by centrifugation, washed and resuspended in ChIP lysis buffer and the chromatin was fragmented by sonication (Sonics Vibracell VCX130, John Morris Scientific) for 20 seconds ON and 30 seconds OFF pulse at 25% amplitude, for a total of 12-18 bursts to generate DNA fragments of around 0.2-1 kb in size. The fragmented chromatin was diluted with ChIP dilution buffer and pre-cleared using Protein G ChIP-grade magnetic beads (Cell Signaling Technology). Chromatin was divided into three aliquots: input (total chromatin), IgG (normal rabbit IgG added, sc-2027, Santa Cruz Biotechnology) and HNF4 $\alpha$  (rabbit HNF4 $\alpha$  antibody added, sc-6556, Santa Cruz Biotechnology). Two micrograms of antibody was used per aliquot and incubated overnight at 4°C on a gentle rotator. Immunocomplexes were captured using Protein-G magnetic beads for two hours followed by extensive washing as in the original protocol. After the final wash of the beads, reversal of DNA/protein complex crosslinking was achieved by heating to 75 °C for 4 hours and then incubating with proteinase-K for 1 hour. Released genomic DNA was purified using PCR-purification kits (QIAGEN) and analysed by quantitative genomic PCR with primers designed for target (*UGT1A8*, *-1A9* and *1A10* proximal promoters) and control loci, and quantified by the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001). Primers used for ChIP assay are listed in the corresponding Chapters.

## **2.2.16 Electrophoretic mobility shift assay (EMSA)**

### **2.2.16.1 *Preparation of nuclear extract***

Nuclear extract was prepared freshly from transfected Caco-2 cells. Transfection was performed in T75 flasks in the same manner as for ChIP. Cells were harvested 48 hours after transfection. Cells were scraped and lysed in hypotonic lysis buffer containing 20 mM Tris pH 7.5, 10 mM MgCl<sub>2</sub>, 2 mM EDTA, 1% Triton-X, 10% glycerol, 10 mM KCl, 1 mM dTT (dithiothreitol). The nuclear pellet was collected from centrifugation for 10 minutes at 14,000 rpm, at 4°C and resuspended in nuclear buffer containing 50 mM Tris-HCl, pH 7.9, 500 mM KCl, 2 mM DTT, 5 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 10% sucrose, 20% glycerol and 1x complete protease inhibitor cocktail.

### **2.2.16.2 *Non-radioactive EMSA assays***

The nuclear extract was used immediately for EMSA or stored at -80°C for use within 1 week. The EMSA protocol was modified from the LUEGO method (Jullien and Herman, 2011) that generates tripartite fluorescently labelled DNA probes. The LUEGO Tag oligonucleotide (5'-gtgccctggctctgg-3') was labelled with fluorophore Cy5. Oligonucleotide probes corresponding to the target element were designed such that one included a terminal segment complementary to the LUEGO Tag sequence. EMSA probe oligonucleotides and the labelled LUEGO Tag oligonucleotide were purchased from Integrated DNA Technologies. Oligonucleotide probes were annealed using Thermocycler (Bio-Rad) with the following program: 94°C for 2 minutes, cooling down at 2°C/second to 70°C and cooling down at 0.1°C/second to 18°C. Annealing reaction was composed of 2.5 µL LUEGO Tag oligonucleotide, 2.5 µL top strand oligonucleotide (100 µM), 1 µL bottom strand oligonucleotide (100 µM), in a total volume of 100 µL in annealing buffer with 50 mM NaCl.

EMSA reactions for DNA-protein complexes consisted of 15 µg nuclear extract, 2 µL EMSA binding buffer (25 mM Tris-HCl, pH 7.6, 100 mM KCl, 5 mM MgCl<sub>2</sub>, 0.5 mM dithiothreitol, 0.5 mM EDTA, and 10% glycerol), 2 µL of probe (1 µM), 2 µL poly (deoxyinosinic-deoxycytidylic) (0.1 µg/µL) and water to a total volume of 10 µL per reaction. Incubation for EMSA reaction was 20 minutes at room temperature immediately after adding the probes. Following incubation, for supershift assays, 0.4 µg antibody was added to the reaction and incubated for a further 20

minutes. For competition assays, unlabelled wild-type competitor oligonucleotide (100-fold excess concentration) were added prior addition of probes. After incubation, the complexes were loaded into a 4% non-denaturing polyacrylamide gel prepared with 0.5x TBE (Tris borate-EDTA) that had been pre-electrophoresed (1 hour at 50V, 4°C in 0.5x TBE). The gel was electrophoresed for an additional 1 hour 45 minutes at 50 V and 4°C. Visualisation of the gel was performed using a Typhoon 9400 Scanner (GE Healthcare Life Science).

### **2.2.17 Western blotting**

Western blotting was performed to assess the CDX2 protein expression level in Caco-2 cells following drug (BI6015) treatment. Treated cells in 6-well plates were harvested by scraping cells after 3 washes with cold PBS, and cells were gently pelleted at 1,500 rpm for 5 minutes. One mL radioimmunoprecipitation assay (Chen *et al.*) buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulphate (SDS)) containing 1x Protease inhibitors was added to lyse the cell pellet; to aid lysis, cells in RIPA buffer were passed through a syringe. The lysate was centrifuged at 13,000 rpm for 10 minutes at 4°C and the supernatant was transferred to a new tube and stored at -20°C.

The protein concentration of lysate was determined by spectrophotometry, before. For each sample, 20 µg of protein was diluted in a 1:1 ratio with 2 x Laemmli sample loading buffer and heated at 95°C for 5 minutes. Samples were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) along with molecular weight markers (Precision Plus Protein™ WesternC™ Standards, Bio-Rad). Electrophoresis conditions were 70 V through the stacking gel and 120 V through the resolving gel for 1-hour. Proteins were transferred onto a nitrocellulose membrane (Trans-blot nitrocellulose, Bio-Rad) by assembling a transfer sandwich of gel and membrane and placing in the cassette in a cooled Mini Trans-Blot cell (Bio-Rad) tank. Transfer occurred at 25 Volts, constant current of 10 mAmps, overnight at 4°C. Following transfer, the membrane was rinsed with TBST (Tris-buffered saline with 0.2% Tween-20), and blocked in 3% (w/v) of non-fat milk in TBST (blocking buffer) for 90 minutes at room temperature. The membrane was washed 3 times with TBST following milk blocking for further blocking in 1% blocking buffer solution

containing primary antibody (CDX2) at 1:5000 dilution overnight at 4°C on a gentle rocking platform. Following primary antibody blocking, the membrane was washed 3 times with TBST and incubated in the HRP-conjugated secondary antibody solution of 1% blocking buffer (1:2000) for 1-hour at RT. Finally, the membrane was rinsed 3 times with TBST and detection performed by adding chemiluminescence substrate (enhanced SuperSignal West Pico chemiluminescent (ECL) HRP substrate) according to the manufacturer's instructions and imaging on an ImageQuant LAS 4000 (GE Healthcare Life Science).

### **2.2.18 Cell imaging (EVOS® FL)**

To analyse transfection efficiency, and capture cell images or organoid images for morphological analysis, imaging was performed using an EVOS® FL Cell Imaging System microscope (Thermo Fisher Scientific, USA).

### **2.2.19 Statistics**

All graphs in this study were generated using Microsoft Excel 2010 software, data significances were analysed using Student's *t* test or one-way ANOVA using Tukey's post-hoc test.

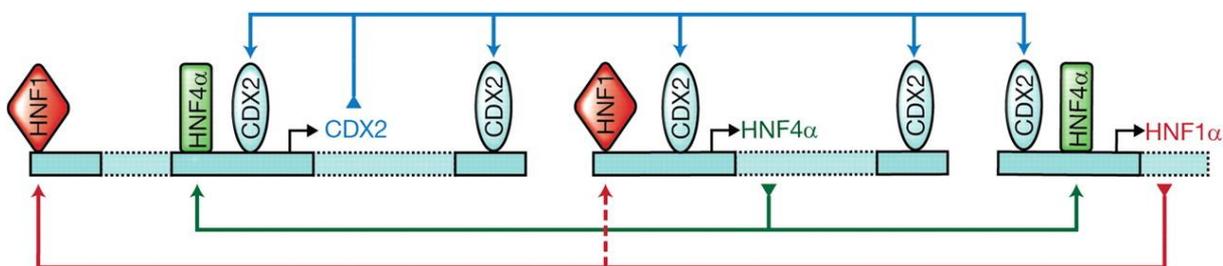
# CHAPTER 3. REGULATION OF INTESTINAL UGTs BY A COMBINATORIAL MECHANISM OF CDX2 AND HNF4 $\alpha$

## 3.1 Introduction

Identification and characterisation of the UGT1A family shows that expression and distribution of these enzymes are spatiotemporally and tissue-specifically regulated beginning in the early embryo and continuing in adult (Collier *et al.*, 2014, Dubey and Singh, 1988, Strassburg *et al.*, 1997, Strassburg *et al.*, 1998). Although glucuronidation activity is considered to mainly occur in the liver, the expression of UGT1A enzymes in the gastrointestinal tract (GIT) reflects the importance of this organ in metabolising a diverse range of ingested substances, which includes nutrients, as well as potentially toxic and carcinogenic chemicals. The UGT1A7-1A10 gene cluster, which share high similarity in their first exon sequences (>70%) are considered to be (with the exception of UGT1A9) extrahepatic, with expression almost exclusively in GIT. UGT1A7 is mainly expressed in the upper GIT (stomach) and is absent in the intestine and colon, UGT1A10 expressed along the GIT from oesophagus to colon, whereas UGT1A8 and UGT1A9 are expressed in the intestine and colon, but absent in the stomach (Gregory *et al.*, 2004b, Strassburg *et al.*, 1997, Strassburg *et al.*, 1998).

In considering what regulatory factors might restrict expression of these genes to the GIT, early work focussed in the intestinal specific regulator CDX2. CDX2 is considered a key player in intestinal development and in maintenance of adult intestinal homeostasis by modulating intestinally expressed genes associated with both the proliferation and differentiation of intestinal cells (Suh and Traber, 1996, Gao *et al.*, 2009). Recent work has expanded our understanding of the regulatory networks important in initiating and maintaining intestinal epithelial differentiation. HNF1 $\alpha$  and HNF4 $\alpha$  are two transcription factors that, although not intestinal-specific, are critically involved with CDX2 in the regulatory network in differentiated intestinal epithelial cells (IECs). As previously discussed (Chapter 1) CDX2 expression levels increase in IECs as they migrate toward the villus tips and mature, HNF1 $\alpha$  expression is also highest in the villus, while the level of HNF4 $\alpha$

is similar from crypt to villus (Olsen *et al.*, 2012). These factors function together in transcriptional complexes, and also regulate one another's expression. In particular, CDX2 regulates its own promoter via autoregulation, and also binds to the promoters of *HNF1 $\alpha$*  and *HNF4 $\alpha$*  to activate these genes. HNF4 $\alpha$  is also able to activate CDX2 and HNF1 $\alpha$  expression. Furthermore, in liver cells HNF1 $\alpha$  binds to the promoter and regulates HNF4 $\alpha$  (Olsen *et al.*, 2012). Figure 3.1 shows a model of regulatory network interactions between CDX2, HNF1 $\alpha$  and HNF4 $\alpha$  in the intestine based on ChIP-seq and ChIP-chip analysis (Olsen *et al.*, 2012).



**Figure 3.1. Model of intestinal regulatory networking in differentiated IECs involving CDX2, HNF1 $\alpha$  and HNF4 $\alpha$ .**

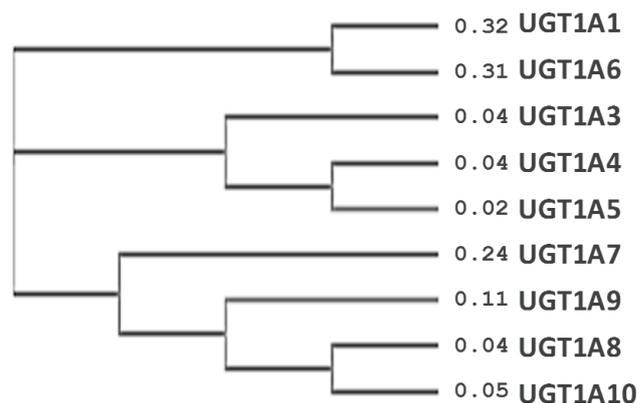
Image is reused from (Olsen *et al.*, 2012). The publisher, the American Physiological Society, allows content to be reused in a thesis without formal permission.

During differentiation of IEC it is proposed that HNF4 $\alpha$  and/or HNF1 $\alpha$  partner with CDX2 to regulate target genes (Verzi *et al.*, 2010, San Roman *et al.*, 2015, Boudreau *et al.*, 2002, Benoit *et al.*, 2010). In addition to CDX2 selective co-regulation, CDX2 would pair with different partners during intestinal progenitor cell proliferation. GATA factors such as GATA-6 are reported to co-occupy with CDX2 in these proliferating cells, and expression of genes associated with differentiation/maturation are not activated in these cells (Verzi *et al.*, 2010).

Intestinal UGT1As are known to be most highly expressed in the differentiated epithelial cells of the lumen where contact with substrates occurs (Strassburg *et al.*, 2000, Dubey and Singh, 1988, Mariadason *et al.*, 2002). Previously in our laboratory, Gregory *et al.* reported that CDX2 and HNF1 $\alpha$  bind to cognate response element within the proximal promoters of *UGT1A8* and *-1A10*

and synergistically induce their activity in Caco-2 cells (Gregory *et al.*, 2004a). Moreover, previous investigation by Gregory (Gregory, 2004) identified at least two GATA motifs within the proximal ~200 bp of the *UGT1A8* promoter (Figure 1.13 in Chapter 1). However, binding of these GATA factors (GATA-4, -5 and -6) to their cognate response elements was unable to induce *UGT1A8* promoter activity, and the same result was shown when GATA factors were in combination with CDX2 or HNF1 $\alpha$ . Overall, these findings are consistent with the reports that CDX2 - HNF1 $\alpha$  cooperation is involved in the up-regulation of genes associated with the differentiated phenotype (such as UGTs) and that CDX2 - GATA partnership is more important for regulation of proliferation-associated genes. Although present in relatively low levels, UGT1A expression in the crypt compartment might also play a role in protection of progenitor and stem cells against genotoxic insults (Patel *et al.*, 1997); however, it is unclear how this may be regulated at present.

The promoter regions spanning up to -1kb from the transcription start sites of *UGT1A8-1A10*, are highly homologous at >75% identity (Gregory *et al.*, 2003) but they show much lower homology with other UGT1A genes. This suggests the importance of these promoter regions in controlling *UGT1A8-10* expression in the GI tract. A phylogenetic tree generated using the -1k-promoter sequences of all UGT1A genes is shown in Figure 3.2.



**Figure 3.2. Phylogenetic analysis of ~1k-proximal promoter regions of *UGT1A* genes**

The number at each terminal branch represent the distance/ between each UGT and its nearest neighbour. The Clustal Omega multiple sequence alignment program was utilized to generate this phylogenetic tree.

Although the proximal promoter regions of *UGT1A8-1A10* are very similar, there are small sequence differences in the predicted binding element for CDX2 in the *UGT1A9* promoter relative to the *UGT1A8* and *UGT1A10* promoters. These differences were proposed to explain the observations from mutational analyses that the putative CDX2 regulatory element is not functional in the *UGT1A9* promoter (Gregory *et al.*, 2004a).

This study aimed to enhance our understanding of intestinal UGT (*UGT1A8*, *1A9* and *-1A10*) regulation by CDX2 and other intestinal factors. Recent work in mouse using genome-wide ChIP-Seq profiling has shown that Cdx2 and HNF4 $\alpha$  co-occupy adjacent sequence elements of promoter regulatory regions and activate the genes responsible for epithelial maturity, survival and intestinal function such as brush border formation, and absorption (Verzi *et al.*, 2013). Although several HNF4 $\alpha$  motifs have been identified within the *UGT1A8*, *1A9* and *-1A10* promoters, the role of HNF4 $\alpha$  in regulation of these UGTs in an intestinal cell context had never been studied. HNF4 $\alpha$  is expressed in liver as well as intestine, and previous findings in our laboratory showed that HNF4 $\alpha$  could activate the *UGT1A9* promoter, but not the *UGT1A8* and *-1A10* promoters in liver cells (Gardner-Stephen and Mackenzie, 2008) and this has been proposed as at least partial explanation for the presence of *UGT1A9*, and absence of *UGT1A8* and *-1A10*, in liver. This study hoped to better explain mechanisms underlying the intestinal specific *UGT1A8* and *-1A10*, and hepatic/intestinal expression patterns of *UGT1A9*.

Using Caco-2 cells, the specific aims of this study are:

1. To investigate potential interaction of CDX2 and HNF4 $\alpha$  in the regulation of *UGT1A8* at the promoter and mRNA level.
2. To identify the binding motif that mediates regulation of *UGT1A8* by CDX2 and HNF4 $\alpha$ .
3. To identify role of the CDX2 and HNF4 $\alpha$  partnership in the regulation of *UGT1A9* and *-1A10*.
4. To understand the transcriptional mechanism that defines the hepatic/extrahepatic expression pattern of *UGT1A8*, *-1A9* and *-1A10* involving specific regulatory modules.

Note that data and text in this Chapter is published in part in “Cooperative regulation of intestinal UDP-glucuronosyltransferases by CDX2 and HNF4 $\alpha$  is mediated by a novel composite regulatory element” Authors: Siti Nurul Mubarakah, JulieAnn Hulin, Peter I Mackenzie, Ross A McKinnon, Alex Z Haines, Dong Gui Hu, and Robyn Meech; Molecular Pharmacology (see Appendix 3).

## **3.2 Methods**

### **3.2.1 Plasmids**

#### **3.2.1.1 Promoter luciferase constructs**

All of the wild-type *UGT1A8*, -1A9 and -1A10 promoter-reporter constructs in pGL3-Basic were previously made in our laboratory by Dr. Philip Gregory as described in (Gregory *et al.*, 2003).

#### **3.2.1.2 Expression vectors**

The HNF4 $\alpha$ -pCMX expression plasmid was previously cloned by Dr. Dione A. Gardner-Stephen in our laboratory, and the CDX2 expression plasmid was a kind gift from Dr. Cathy Mitchelmore (University of Copenhagen, Copenhagen, Denmark). The Barx2-pcDNA3 expression plasmid was cloned by Mr. Lizhe Zhuang in our laboratory.

### **3.2.2 Mutagenesis of binding sites in *UGT1A8*, -1A9 and -1A10 promoter constructs**

Mutation of desired nucleotides within the response elements of the UGT promoter-reporter constructs was achieved by *in vitro* site-directed mutagenesis (SDM) method using the QuikChange<sup>®</sup> Site-Directed Mutagenesis System with *PfuTurbo*<sup>®</sup> polymerase (Stratagene, La Jolla, CA), and primers containing the desired mutation (Integrated DNA Technologies, CA) as described in Chapter 2. All oligonucleotides are listed in Table 3.1. Promoter-reporter constructs containing the correct mutations were identified by sequencing (DNA Sequencing Service, SA Pathology, South Australia). The mutations that were introduced are as follows: 1) in the -1kb-*UGT1A8* promoter region the HNF4 $\alpha$  response element (RE) at -811bp to -798bp was mutated from "TGACCTCAGGGAG" to "TGATTTTCAGGGAG"; 2) in the *UGT1A8*, -1A9,-1A10 promoter regions (-1kb- and -0.19kb) the HNF4 $\alpha$ RE at -44bp to -32bp was mutated from "TCTATTGGGGTCA" to "TCTATTGGAATCA"; 3) in the 0.19kb *UGT1A8* promoter region the HNF4 $\alpha$ RE at -44bp to -32bp was mutated from "TCTATTGGGGTCA" to "TCTATCCGGGTCA" (i.e.

a different mutation). In addition, we used a variant of the UGT1A8 promoter construct carrying a mutation in the CDX2RE at location -70bp to -66bp, and a variant carrying a mutation of the Sp1/Initiator-likeRE at location -15bp to -5bp, that were made previously by Dr. Philip Gregory (Gregory *et al.*, 2003).

### **3.2.3 Transfections of CDX2 and HNF4 $\alpha$ expression vectors**

Expression vectors for CDX2 and HNF4 $\alpha$  were transfected into Caco-2 cells using Lipofectamine® LTX as described in Chapter 2, Section 2.2.13.2. Endogenous UGT mRNA levels were analysed by qRT-PCR. All primers are listed in Table 3.1.

### **3.2.4 Chromatin Immunoprecipitation (ChIP) assays in Caco-2 cells**

ChIP assay was performed as described in Chapter 2, Section 2.2.15. qPCR analysis was performed following ChIP protocols to detect HNF4 $\alpha$  binding to DNA sequences in the proximal promoters of the *UGT1A8*, *-1A9* and *-1A10* genes, and at a control non-target locus, using primers listed in Table 3.1.

### **3.2.5 Non-radioactive EMSA assays in Caco-2 cells**

Nuclear extract was prepared from Caco-2 cells transfected with HNF4 $\alpha$  and CDX2 expression plasmids and EMSA assay was performed as described in Chapter 2, Section 2.2.16. For supershift analysis to confirm complex specificity, 1  $\mu$ g of HNF4 $\alpha$  (sc-6556; Santa Cruz Biotechnology) or Cdx2 (Biogenex, San Ramon, CA) antibody was included per binding reaction. Probes for detecting protein-DNA complexes in the reaction are shown in Table 3.1.

### **3.2.6 CDX2 and HNF4 $\alpha$ siRNA design and transfection**

CDX2 and HNF4 $\alpha$  siRNAs (GenePharma, Shanghai, China) were reverse transfected in triplicate wells of a 48-well plate using Lipofectamine® 2000 in 48-well plates as described in Chapter 2. The transfection was repeated after 24 hours to enhance target gene knockdown efficacy. Forty-eight hours after the first transfection, cells were harvested for qRT-PCR analysis as in Chapter 2. Primers used for qRT-PCR assay are listed in Table 3.1.

### 3.2.7 Transcriptomic data profiling of Colon Adenocarcinoma (COAD) from The Cancer Genome Atlas (TCGA)

Analysis of RNA sequencing (RNAseq) transcriptomic data from Colon Adenocarcinoma (COAD) was conducted by Dr. Dong Gui Hu and Dr. Robyn Meech; data were downloaded from The Cancer Genome Atlas (TCGA) (<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 normalised using the upper quantile normalisation method. Spearman's correlation analyses between the expression levels of two UGT genes (e.g. UGT1A8, -1A10) and two transcription factors (CDX2, HNF4a) 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.

### 3.2.8 Oligonucleotides

Table 3.1 shows oligonucleotides used in this Chapter, the mutated bases are indicated by bold letters.

**Table 3.1. List of oligonucleotides used in studies described in Chapter 3.**

Name	Sequence (5' → 3')
Quantitative real time PCR	
18srRNA-Forward	CGATGCTCTTAGCTGAGTGT
18srRNA-Reverse	GGTCCAAGAATTTACCTCT
UGT1A8-Forward	CTGCTGACCTGTGGCTTTGCT
UGT1A8-Reverse	CCATTGAGCATCGGCGAAAT
UGT1A9-Forward	GAGGAACATTTATTATGCCACCG
UGT1A9-Reverse	GCAACAACCAAATTGATGTGT
UGT1A10-Forward	CCTCTTTCCTATGTCCCAATGA
UGT1A10-Reverse	GCAACAACCAAATTGATGTGTG
CDX2-Forward	ATCACCATCCGGAGGAAAG
CDX2-Reverse	TGCGGTTCTGAAACCAGATT
HNF4a-Forward	CAGCACTCGAAGGTCAAGCTA
HNF4a-Reverse	ACGGGGGAGGTGATCTGT

Site directed mutagenesis (SDM) of promoter reporter	
UGT1A8 -811HNF4amutFor	CATC <b>ATT</b> ACTG <b>ATTT</b> CAGGGAGTGCCCAG
UGT1A8 -811HNF4amut Rev	CCCTGAA <b>AAT</b> CAGT <b>AAT</b> GATGTCATCTTTGTGT
UGT1A8 -44HNF4amutFor	ATTGGA <b>AAT</b> CAGGTTTTGTGCCTGTAGTTC
UGT1A8 -44HNF4amut Rev	CCTG <b>ATT</b> CCAATAGAGGGCGTGATTTATCCTG
UGT1A8 -44CrypticCDX2-composite mutFor	CCTCTAT <b>CCGGG</b> TCAGGTTTTGTGCCTGTAGTT
UGT1A8 -44CrypticCDX2-composite mut Rev	GACCC <b>GG</b> ATAGAGGGCGTGATTTATCCTGTCA
EMSA	
LUEGO-Tag	GTGCCCTGGTCTGG
Consensus HNF4 from apoCIIIAPF1	GTGCCCTGGTCTGGCGCTGGGCAAAGGTCACCTGC
UGT1A8Cdx2 -70 wt Top	GTGCCCTGGTCTGGTTTTTTTTTATGACAGGATA
UGT1A8/-1A9/-1A10HNF4a half cons/ Cryptic CDX2-composite wt Top	GTGCCCTGGTCTGGCGCCCTCTATT <b>GGGG</b> TCAGGTTTT
UGT1A8/-1A9/-1A10HNF4a half cons-composite m Top	GTGCCCTGGTCTGGCGCCCTCTATT <b>GAA</b> GTCAAGGTTTT
UGT1A8/-1A9/-1A10 Cryptic CDX2-composite m I Top	GTGCCCTGGTCTGGCGCCCTCTAT <b>CCGGG</b> TCAGGTTTT
UGT1A8/-1A9/-1A10 Cryptic CDX2-composite m II Top	GTGCCCTGGTCTGGCGCCCTCTAT <b>CCCGG</b> TCAGGTTTT
ChIP-qPCR	
ChIP negative control locus For	ACATACTCAGATGGAAATGAGAA
ChIP negative control locus Rev	AGCTCAACATTCTGCTGAAC
UGT1A8 -44HNF4a locus For	TTTTGGTACCTCAAAAAATGATACTC
UGT1A8 -44HNF4a locus Rev	AGCCACGCGTGAAGTGCAGCCCGAGC
UGT1A9 -57HNF4a locus For	TATGGATGGGGGCAGTC
UGT1A9 -57HNF4a locus Rev	CTCCTATGATACAGTAGGTGGG
UGT1A10 -47HNF4a locus For	AGTAGGTACCTCAGCAAATGATACTC
UGT1A10 -47HNF4a locus Rev	CCACCCCGGGCGAGCCATGAGAGAACTG
siRNA	
hHNF4 $\alpha$ siRNA	CGUCAAGGAUGCGUAUGGACACCCGGC
hCDX2 siRNA	AACCAGGACGAAAGACAAUA
human neg control siRNA	UUCUCCGAACGUGUCACGU

### 3.3 Results and Discussion

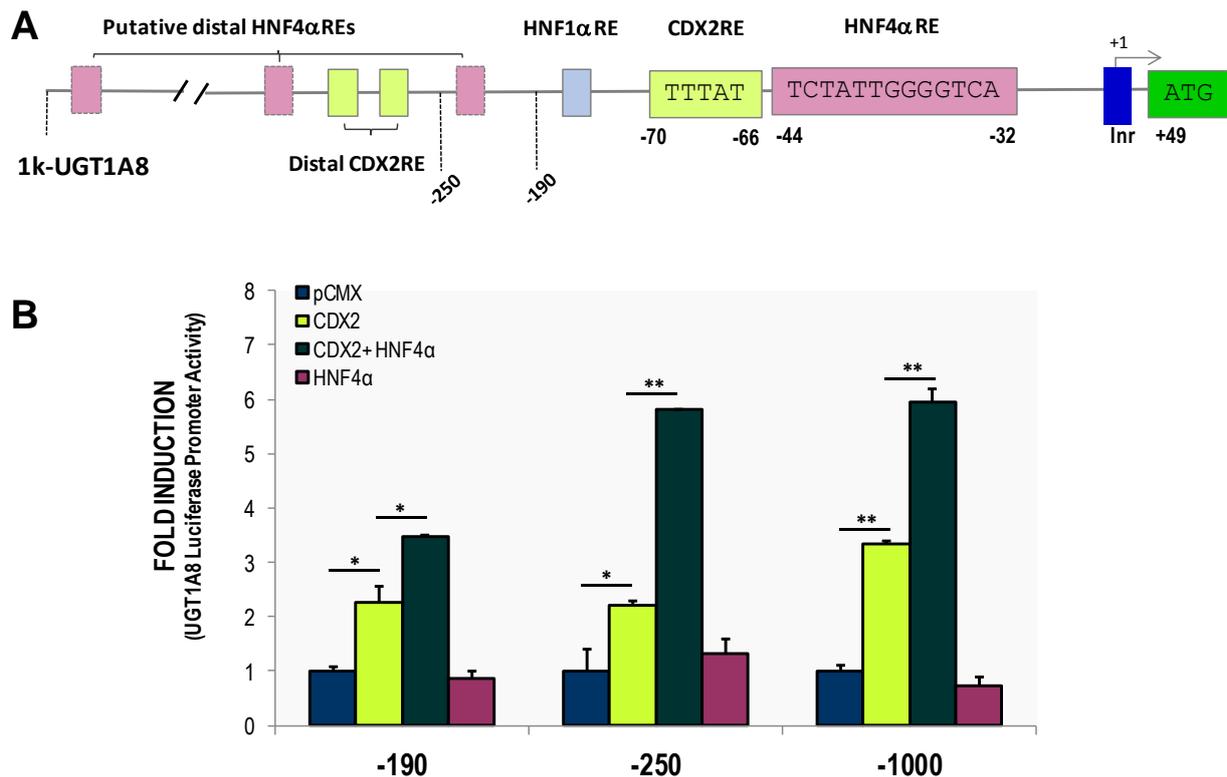
#### 3.3.1 CDX2 synergistically interacts with HNF4 $\alpha$ to further induce *UGT1A8* promoter activity

As described in the introduction, a functional CDX2 motif was previously identified in the *UGT1A8* and *-1A10* promoters, but the equivalent motif was non-functional in the *UGT1A9* promoter (Gregory *et al.*, 2004a), leaving unresolved the question of how CDX2 may control the intestinal expression of UGT1A9. CDX2 and HNF4 $\alpha$  are known to act in combinatorial manner during intestinal development and differentiation (Verzi *et al.*, 2010, San Roman *et al.*, 2015). Numerous HNF4 $\alpha$  binding motifs are predicted within the proximal promoters of *UGT1A8-1A10* based on bioinformatic analysis. In addition, previous studies of the *UGT1A9* promoter identified HNF4 $\alpha$  motifs that function in a liver cell context (Gardner-Stephen and Mackenzie, 2007). However, HNF4 $\alpha$  has not been previously examined as a regulator of any UGT in an intestinal context.

To investigate whether CDX2 and HNF4 $\alpha$  may regulate intestinal UGTs in a combinatorial manner, we began by examining *UGT1A8*, the prototypical intestinal-specific UGT. Within the proximal 190 bp length of the *UGT1A8* promoter, we predicted an HNF4 $\alpha$  binding motif at -44bp which is very close to the previously identified functional CDX2 and HNF1 $\alpha$  binding sites, located at -70bp and -148bp respectively. Three other putative HNF4 $\alpha$  motifs (HNF4 $\alpha$ RE) were also predicted within the -1k promoter region, including a HNF4 $\alpha$ RE at -290bp previously identified in our laboratory (Gardner-Stephen and Mackenzie, 2007), a HNF4 $\alpha$ RE at -360bp homologous to that previously identified in *UGT1A9* by Barbier *et al.* (Barbier *et al.*, 2005) and a HNF4 $\alpha$ RE at -811bp (Figure 3.3A). The last HNF4 $\alpha$ RE overlapped with a functional PPAR-RE previously identified in *UGT1A9* (nt -719 to -706 bp from ATG) (Barbier *et al.*, 2003). All three distal HNF4 $\alpha$ RE were previously shown to be non-functional in *UGT1A8* promoter activation in liver cell model, HepG2 (Gardner-Stephen and Mackenzie, 2007). These motifs however, had never been investigated in the intestinal cell context.

To examine the roles of the various predicted HNF4 $\alpha$ RE in the regulation of *UGT1A8*, we tested the response of promoter-reporter vectors carrying several lengths of the proximal -1kb *UGT1A8*

promoter to the expression of HNF4 $\alpha$ , CDX2, and combination of CDX2 and HNF4 $\alpha$ . As shown in Figure 3.3B, HNF4 $\alpha$  alone could not induce promoter activity in the -0.19kbb-, -0.25k-, or -1kb- UGT1A8 promoter constructs. In contrast, CDX2-alone induced all three UGT1A8 promoter constructs. Interestingly, the combination of HNF4 $\alpha$  and CDX2 induced UGT1A8 promoter activity by approximately 2-fold more than CDX2 alone (Figure 3.3B). These results suggest that a CDX2 and HNF4 $\alpha$  synergistically regulate the UGT1A8 promoter.

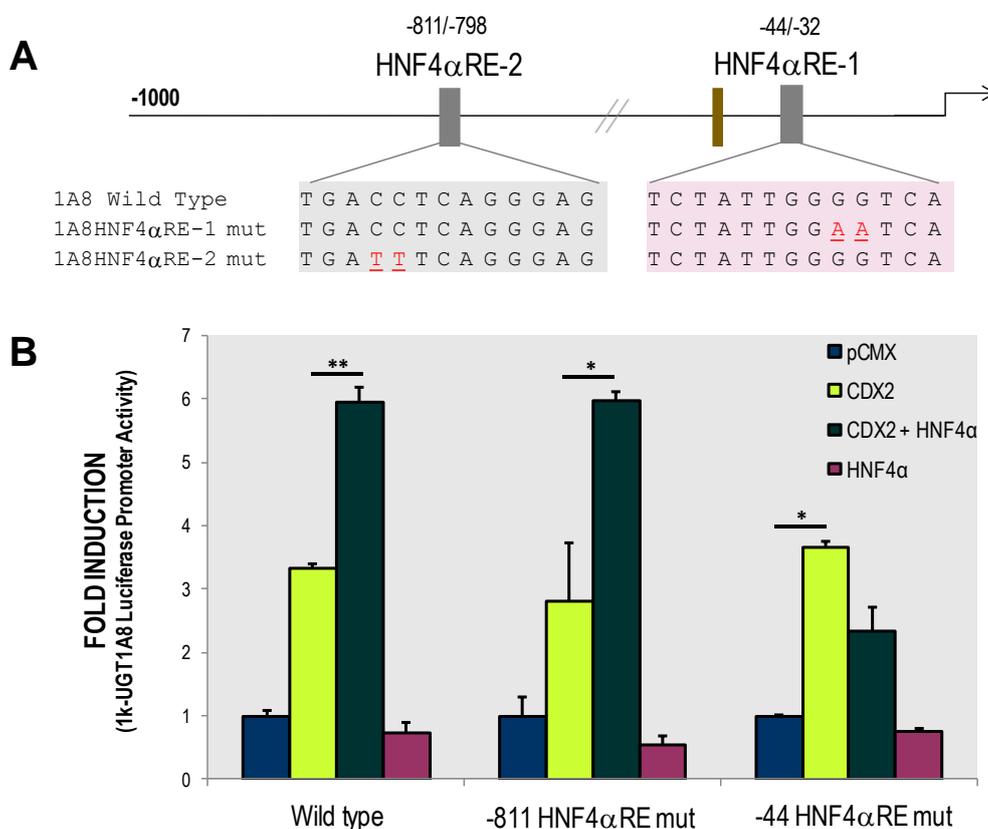


**Figure 3.3. Synergistic activation of the UGT1A8 promoter by CDX2 and HNF4 $\alpha$ .**

A. Schematic representation of the -1kb UGT1A8 promoter containing predicted sites of four HNF4 $\alpha$ , three CDX2, one HNF1 $\alpha$  and Sp1/Inr element, as indicated. Transcription start site is shown as +1. B. Caco-2 cells were co-transfected with three different lengths of UGT1A8 promoter constructs (-1kb, -0.25kb- and -0.19kb-) and expression plasmids for CDX2, HNF4 $\alpha$  or a combination of both plasmids. pRL-null was used as internal control. Cells were assayed 48h post-transfection for luciferase activities using promoter-less pGL3 basic to normalise the activity, and is expressed as fold induction over empty vector pCMX. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### 3.3.2 Co-regulation of CDX2/HNF4 $\alpha$ is mediated through a novel synergistic composite element

The CDX2/HNF4 $\alpha$  synergistic induction was observed with all promoter constructs including the shortest -0.19kb fragment (Figure 3.4B), suggesting that the transcriptional mechanism is mediated by the newly predicted HNF4 $\alpha$ RE at -44 bp (called hereafter -44HNF4 $\alpha$ RE) and not the more distal elements. To test this idea, we mutated the proximal -44HNF4 $\alpha$ RE and as a control we also mutated the distal HNF4 $\alpha$  site at -811bp in the -1kb-UGT1A8 promoter. Mutation of -44HNF4 $\alpha$ RE ablated the CDX2/HNF4 $\alpha$  synergistic effect, whereas mutation of the -811HNF4 $\alpha$ RE had no effect, confirming that -44HNF4 $\alpha$ RE is required for the CDX2/HNF4 $\alpha$  synergy.



**Figure 3.4. The HNF4 $\alpha$ RE at -44bp in the *UGT1A8* proximal promoter is required for synergistic induction by CDX2 and HNF4 $\alpha$ .**

A. Schematic representation of 1k-UGT1A8 promoter constructs containing mutations at two different HNF4 $\alpha$  sites- the proximal -44 bp (HNF4 $\alpha$ RE-1) and distal site at -811bp (HNF4 $\alpha$ RE-2), mutated bases are indicated by red letters. B. CDX2 and HNF4 $\alpha$  synergistic activation of the *UGT1A8* promoter was blocked by mutation of -44HNF4 $\alpha$ RE but not -811HNF4 $\alpha$ RE. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

The presence of the previously identified functional CDX2 binding site at -70bp (called hereafter -70CDX2RE) in proximity to the -44HNF4 $\alpha$ RE (Figure 3.5A), suggested that CDX2 binding to this motif may mediate CDX2 cooperation with HNF4 $\alpha$ . To test this idea, mutation of -70CDX2RE was introduced into the 0.19k-UGT1A8 promoter and co-transfected with CDX2, HNF4 $\alpha$  and a combination of CDx2 and HNF4 $\alpha$ . Unexpectedly, whilst the mutation blocked *UGT1A8* promoter activation by CDX2 alone, the combination of CXD2 and HNF4 $\alpha$  was still observed to activate the mutant promoter (Figure 3.5B). This suggests recruitment of CDX2 to -70CDX2RE was not required for the synergistic induction by CDX2 and HNF4 $\alpha$ .

Another element in the *UGT1A8* promoter in close proximity to the -44HNF4 $\alpha$ RE is an Sp1/Initiator(Du *et al.*)-like motif, as indicated in Figure 3.5C. The Sp1/Inr-like site is believed to mediate recruitment of the basal transcription machinery. It was previously speculated that HNF1 $\alpha$  might promote interaction of CDX2 with the basal transcription machinery and this might play a role in CDX2/HNF1 $\alpha$  synergy (Gregory *et al.*, 2004a). Therefore, we sought to examine whether activation by CDX2 alone or the synergistic activation by CDX2 and HNF4 $\alpha$  involved the Sp1/Inr-like motif. We co-transfected the -0.19k-UGT1A8 promoter wild-type and mutant (Sp1/Inr-like mut) constructs with CDX2, HNF4 $\alpha$  or combination of both expression plasmids. The promoter construct with the mutated Sp1/Inr-like motif could not be induced by CDX2 alone; surprisingly however it was still induced by the combination of CDX2 and HNF4 $\alpha$ .



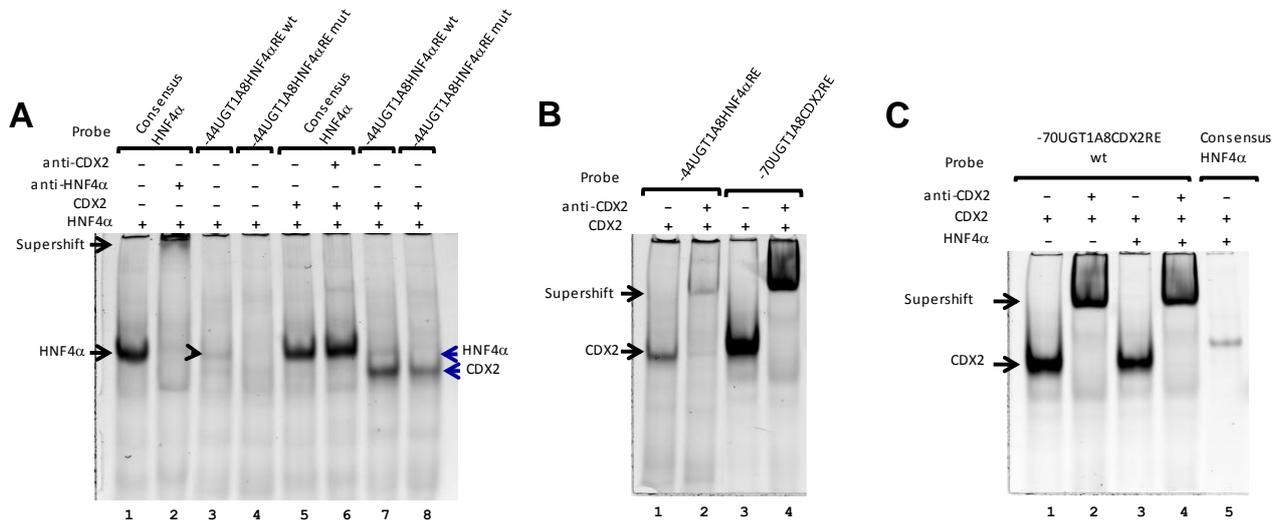
As shown in Figure 3.6A, extract expressing HNF4 $\alpha$  formed a strong complex on the consensus HNF4 $\alpha$ RE that was supershifted by HNF4 $\alpha$  antibody (lanes 1, 2). The HNF4 $\alpha$  extract formed a comparatively weaker complex on the -44bp HNF4 $\alpha$ RE probe (lane 3) but mutation of the HNF4 $\alpha$  core recognition motif prevented this complex from forming (lane 4) demonstrating specificity. Of note, previous studies showed that binding of HNF4 $\alpha$  to the functional upstream HNF4 $\alpha$ REs in *UGT1A9* was also much weaker than to a consensus HNF4 $\alpha$  probe (Gardner-Stephen and Mackenzie, 2007).

When nuclear extracts containing both CDX2 and HNF4 $\alpha$  were assayed for binding to the -44HNF4 $\alpha$ RE probe, an additional faster migrating complex was formed (lanes 7, 8). This complex was not formed on the consensus HNF4 $\alpha$ RE probe (lanes 5, 6). Moreover, this additional complex was not ablated by mutation of the core HNF4 $\alpha$  recognition motif in the probe (lane 8), suggesting the complex is not mediated by HNF4 $\alpha$  binding.

To test whether the faster migrating complex formed on the -44HNF4 $\alpha$ RE probe when incubated with CDX2-containing extract actually contained CDX2 we used EMSA/supershift analysis. The -70bp CDX2RE probe was used as a positive (consensus) control for CDX2 binding. Figure 3.6B shows that nuclear extracts containing CDX2 formed a robust complex with the -70CDX2RE probe (lane 3), which migrated similarly to the (comparatively weaker) complex formed on the -44HNF4 $\alpha$ RE probe (lane 1). Supershift analysis showed that the complexes formed on both the -44HNF4 $\alpha$ RE and -70CDX2RE probes were shifted by addition of anti-CDX2 antibody (lane 2 and 4), confirming that these complexes contained CDX2. This CDX2 supershift data, together with the observation that the extract containing CDX2 still forms a complex on the 44HNF4 $\alpha$ RE probe when the core HNF4 $\alpha$  motif is mutated suggests that CDX2 might bind to the -44bp HNF4 $\alpha$  RE probe independently of HNF4 $\alpha$ .

We also examined whether HNF4 $\alpha$  might bind to the -70bp CDX2RE (Figure 3.6C). CDX2-containing extract formed a robust complex with this probe that was shifted by CDX2 antibody (lanes 2 and 4). However, there were no additional complexes formed by extracts that contain both

CDX2 and HNF4 $\alpha$  (compare lanes 1 and 3). The consensus HNF4 $\alpha$  probe was used as a positive control of HNF4 $\alpha$  complex migration (lane 5). Thus HNF4 $\alpha$  does not bind to the -70bp CDX2RE; a finding is consistent with the redundancy of the -70bp CDX2RE for CDX2-HNF4 $\alpha$  synergy (Figure 3.5B).



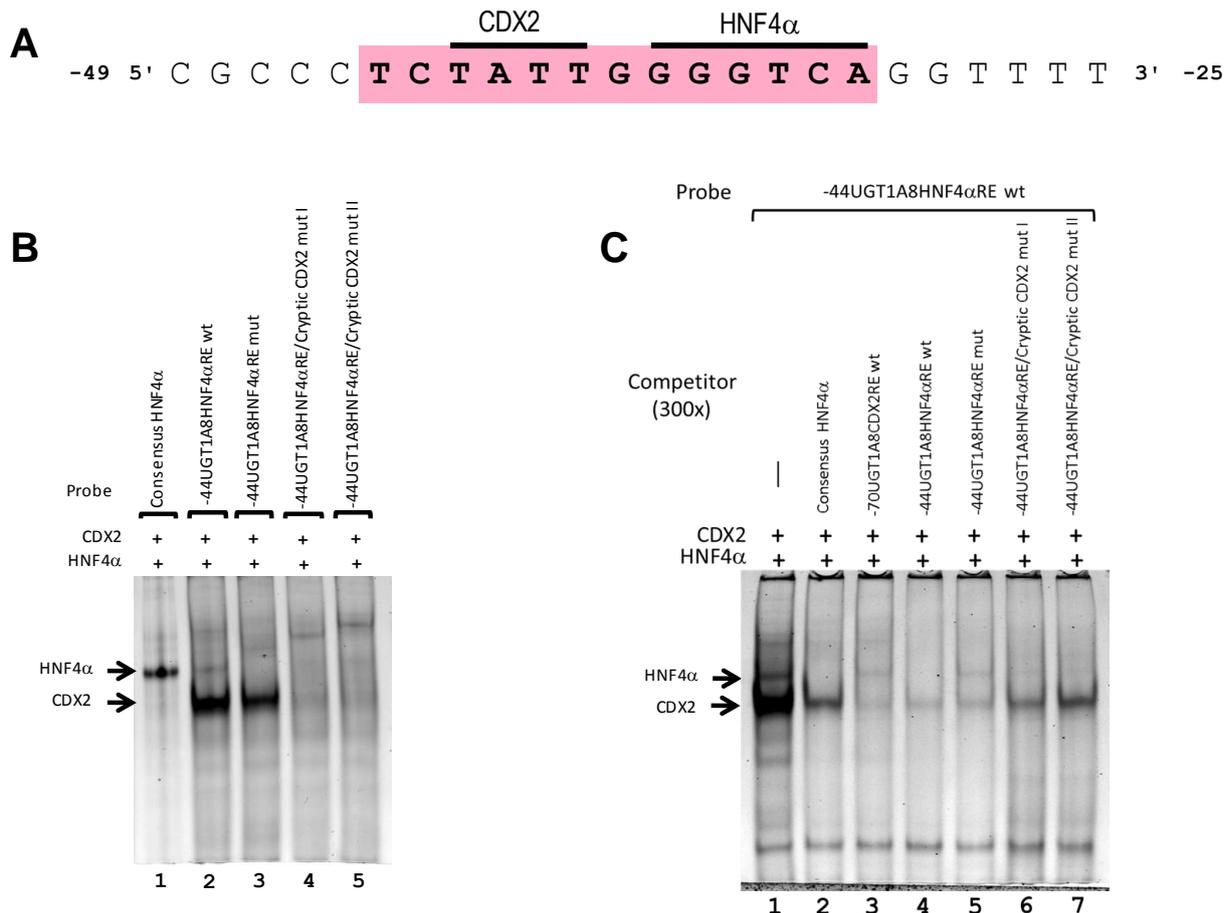
**Figure 3.6. EMSA analysis of HNF4 $\alpha$  and CDX2 binding to the novel composite element of the UGT1A8 promoter region.**

A. Lanes 1, 2, 5, 6 of EMSA assays shows single complex formed by HNF4 $\alpha$  consensus probe with Caco-2 nuclear extracts containing either HNF4 $\alpha$  or CDX2 and HNF4 $\alpha$ , anti-HNF4 $\alpha$  antibody prevents complex formation, anti-CDX2 antibody does not inhibit complex formation on the HNF4 $\alpha$  consensus probe. Lane 3, a weaker complex that migrates at the same position as the complex formed on the HNF4 $\alpha$  consensus probe is formed on -44UGT1A8HNF4 $\alpha$  RE probe by extracts containing HNF4 $\alpha$ ; Lane 4, this complex is ablated by mutation of the core of HNF4 $\alpha$  motif. Lanes 7, 8, the -44UGT1A8HNF4 $\alpha$ RE probe forms an additional complex with nuclear extracts containing both CDX2 and HNF4 $\alpha$  (blue arrow). B. The -44HNF4 $\alpha$ RE probe forms a complex with nuclear extracts containing CDX2 (lane 1) which migrates at the same position as a complex formed by consensus --70CDX2RE probe with nuclear extracts containing CDX2 (lane 3). Lanes 2 and 4, addition of anti-CDX2 antibody supershifts the complex. C. The -70CDX2RE probe forms a single complex with nuclear extracts containing CDX2 only (lane 1) or CDX2 and HNF4 $\alpha$  (lane 3). No HNF4 $\alpha$  complex is evident (see lane 5 for comparison). The CDX2 complex formation is supershifted by addition of anti-CDX2 antibody (lanes 2 and 4).

Overall, these data suggest that the -44bp HNF4 $\alpha$ RE, which we have identified as mediating a novel synergistic response to HNF4 $\alpha$  and CDX2, binds to both HNF4 $\alpha$  and CDX2; moreover, the HNF4 $\alpha$  core recognition motif is not required for CDX2 binding. Further analysis of the sequence of this element showed that it contains a cryptic CDX2-like binding motif with the sequence TATT immediately adjacent to the HNF4 $\alpha$  core recognition motif (Figure 3.7A). To test whether this TATT

motif might mediate binding to CDX2, we mutated the motif (two different mutations, see Table 3.1 for details) in the -44bp HNF4 $\alpha$ RE probe and performed EMSA with extracts containing both HNF4 $\alpha$  and CDX2. As shown in Figure 3.7B, mutation of the CDX2-like binding motif in the -44HNF4 $\alpha$ RE probe abolished the formation of the CDX2 complex (lane 4 and 5). In contrast, mutation of the HNF4 $\alpha$  core motif abolished formation of the HNF4 $\alpha$  complex (lane 2) but not the CDX2 complex (lane 3).

We further used unlabelled oligonucleotide competition to confirm the role of these two motifs in binding to CDX2 and HNF4 $\alpha$  respectively (Figure 3.7C). The -44HNF4 $\alpha$ RE probe formed both the CDX2 and HNF4 $\alpha$  complexes (lane 1); a consensus HNF4 $\alpha$ RE competitor blocked formation of the HNF4 $\alpha$  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 HNF4 $\alpha$  complex (lane 3), whereas the -44HNF4 $\alpha$ RE (self) competitor blocked both complexes (lane 4). A -44HNF4 $\alpha$ RE competitor with a mutated HNF4 $\alpha$  motif did not block the HNF4 $\alpha$  complex but reduced the CDX2 complex (lane 5); in contrast, a -44HNF4 $\alpha$ RE competitor with a mutated CDX2 motif had little effect on the CDX2 complex but blocked the HNF4 $\alpha$  complex (lanes 6, 7). These data further confirm that the -44HNF4 $\alpha$ RE is a composite of two motifs that likely mediate adjacent binding of HNF4 $\alpha$  and CDX2.



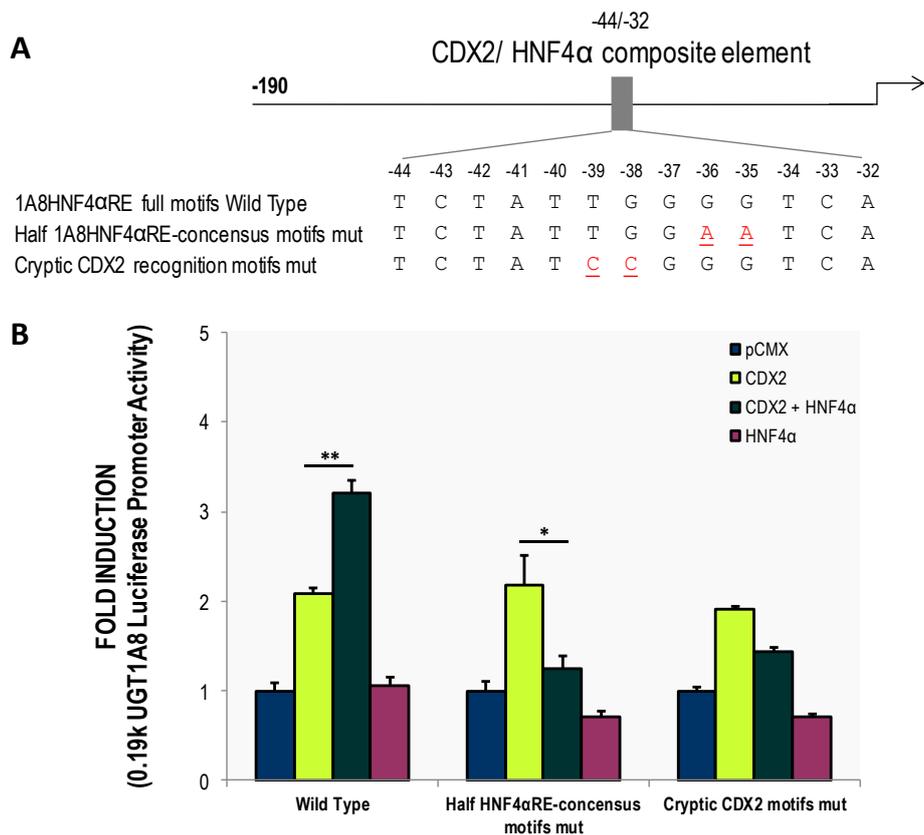
**Figure 3.7. EMSA mutational analysis of CDX2 and HNF4 $\alpha$  binding to the synergistic composite element designated at -44HNF4 $\alpha$ RE.**

A. Sequence of the novel composite element, identified as -44HNF4 $\alpha$ RE, shows predicted recognition motifs bound by CDX2 (TATT) and HNF4 $\alpha$  (GGGTCA). B. -44HNF4 $\alpha$ RE probe containing mutation of CDX2 recognition motif inhibited CDX2 complex formation. Using nuclear extracts containing both CDX2 and HNF4 $\alpha$ , lane 1 shows HNF4 $\alpha$  complex formed by consensus HNF4 $\alpha$  probe but not CDX2. -44HNF4 $\alpha$ RE probe shows the same HNF4 $\alpha$  migration complex formation with additional formation of faster CDX2 complex (lane 2). HNF4 $\alpha$  complex formation is lost by mutation of the core of the -44HNF4 $\alpha$ RE probe (GGGTCA) (lane 3). When CDX2 motif within -44HNF4 $\alpha$ RE is mutated, both faster (CDX2) and slower (HNF4 $\alpha$ ) complexes are ablated (lane 4 and 5). C. EMSA competitor analysis. Both CDX2 and HNF4 $\alpha$  complex formed on -44HNF4 $\alpha$  probe with nuclear extracts containing CDX2 and HNF4 $\alpha$  (lane 1). Competitor of consensus HNF4 $\alpha$  ablates HNF4 $\alpha$  complex formation (lane 2). Competitor of consensus -70CDX2RE modestly reduces the complex formation (lane 3). Competitor of wildtype -44HNF4 $\alpha$ RE ablates formation of HNF4 $\alpha$  and CDX2 complex (lane 4). Competitor of -44HNF4 $\alpha$ RE containing mutation of HNF4 $\alpha$  core motif ablates CDX2 complex but not HNF4 $\alpha$  complex (lane 5). Competitor of -44HNF4 $\alpha$ RE containing mutation of CDX2 motif ablates HNF4 $\alpha$  complex but not CDX2 complex (lane 6 and 7).

Based on the data shown above, we designate the newly identified -44HNF4 $\alpha$ RE in the UGT1A8 promoter a composite element which can bind to both CDX2 and HNF4 $\alpha$  and mediate synergistic regulation by these two factors. The discovery and characterization of this novel element is

considered a major finding in this study. One curious aspect of our EMSA data is that interaction of the composite element probe (-44HNF4 $\alpha$ RE) with extracts containing both CDX2 and HNF4 $\alpha$  produced two distinct complexes that migrated equivalently to the complexes formed when separate CDX2 and HNF4 $\alpha$  extracts were used. If there was simultaneous binding of both factors to the probe, a different slower migrating (i.e. larger) complex might be expected to be formed. However, the absence of such a larger complex might be an artefact of the technique; in particular, steric hindrance might prevent both factors binding simultaneously to the probe. It is conceivable that simultaneous binding to the native element within genomic DNA may involve conformational changes that prevent steric hindrance (Ismail *et al.*, 2010), and the short EMSA probe may not be able to reproduce such a conformation. Regardless, simultaneous binding of both factors is the best explanation for our observation that their synergistic activity is lost upon mutation of either motif; in future work this might be further supported by analyses such as sequential ChIP with antibodies to both factors.

To assess the functional significance of the cryptic CDX2 motif in the -44HNF4 $\alpha$ RE, mutation of the CDX2 motif was introduced into the -0.19k-UGT1A8 promoter construct (Figure 3.8A) and synergistic induction by CDX2 and HNF4 $\alpha$  was assessed. As shown in Figure 3.8B, mutation of the cryptic CDX2 motif inhibited the synergistic activation of the promoter by CDX2 and HNF4 $\alpha$  as effectively as mutating the HNF4 $\alpha$  motif, showing that both motifs are required for synergy and that the loss of one motif cannot be compensated by the other.



**Figure 3.8. The CDX2 and HNF4 $\alpha$  motifs in the composite element are both required for synergistic promoter activation.**

A. Schematic representation of mutation generated within the CDX2/HNF4 $\alpha$  composite element in the -0.19k-UGT1A8 construct. Mutated nucleotides are indicated in red. B. Mutation of either CDX2 or HNF4 $\alpha$  motif blocks synergistic induction of the -0.19k-UGT1A8 promoter by CDX2 and HNF4 $\alpha$ . For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### 3.3.3 The CDX2/HNF4 $\alpha$ composite element is conserved in the *UGT1A8*, -1A9 and -1A10 promoters

As mentioned previously, the proximal promoters of *UGT1A8*, -1A9 and -1A10 share a high degree of sequence homology, and the newly identified -44HNF4 $\alpha$ RE in *UGT1A8* is fully conserved to the corresponding HNF4 $\alpha$ RE in -57bp (from transcription start site [TSS]) of the *UGT1A9* promoter (designated -57HNF4 $\alpha$ RE) and -47bp (from TSS) of the *UGT1A10* promoter (designated -47HNF4 $\alpha$ RE) (Figure 3.10A). Consistent with our experimental data showing the importance of this CDX2/HNF4 $\alpha$  composite element in regulation of *UGT1A8-1A10* by intestinal factors, Figure 3.9 shows that among all *UGT1A* genes, this site is only conserved in the promoter regions of the extrahepatic *UGT1A8-1A10* genes.



slightly, which is likely to be mediated by the -47HNF4 $\alpha$ RE as this small induction was consistently lost when the motif was mutated in both the -1kb- and -0.19kb-promoters. This finding suggests that CDX2 and HNF4 $\alpha$  synergistically regulate the *UGT1A10* promoter in a similar manner to *UGT1A8*; moreover, HNF4 $\alpha$  alone may also be capable of regulating *UGT1A10*.

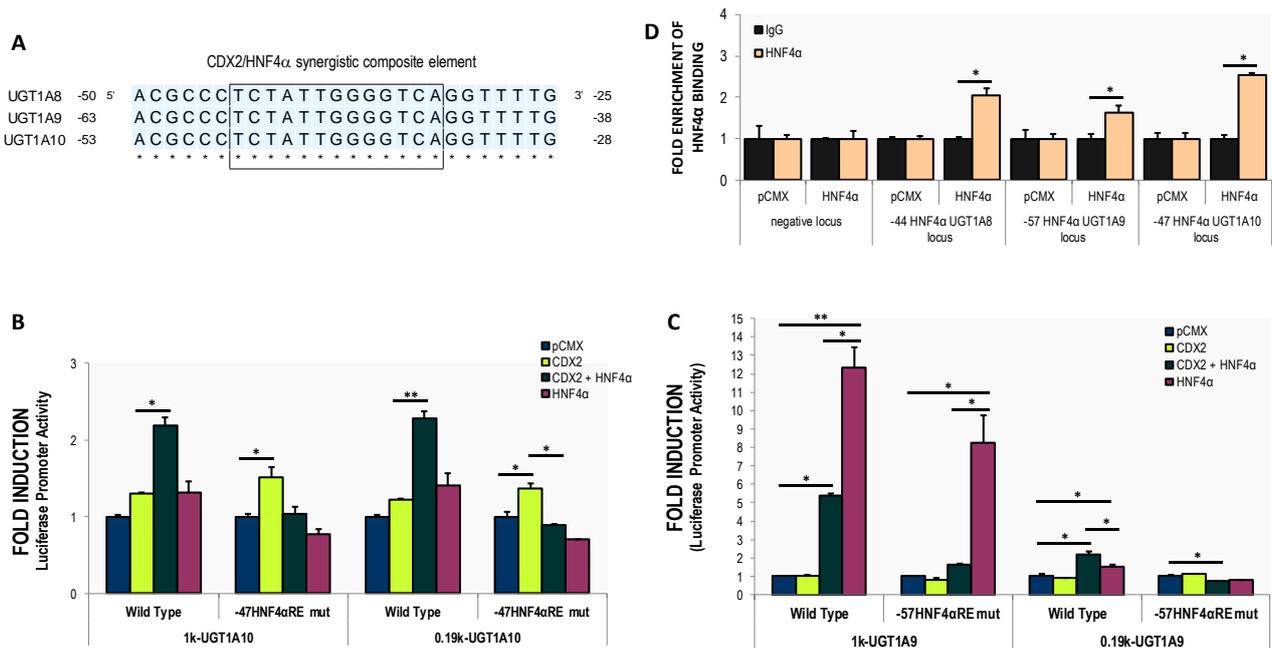
Next, promoter mutation analysis was performed to assess the function of the CDX2/HNF4 $\alpha$  composite motif in the *UGT1A9* promoter (-57HNF4 $\alpha$ RE). Mutations of the -57HNF4 $\alpha$ RE were introduced in the -1kb- and -0.19kb-*UGT1A9* promoter constructs. The wildtype and mutant *UGT1A9* promoter constructs were then co-transfected into Caco-2 cells with CDX2, HNF4 $\alpha$  or the combination of both CDX2 and HNF4 $\alpha$  expression plasmids. As shown in Figure 3.10C, the -0.19kb-*UGT1A9* promoter construct was not induced by CDX2 alone, which is consistent with the reported non-functional -70CDX2RE motif in *UGT1A9* (Gregory *et al.*, 2004a). HNF4 $\alpha$  alone induced the promoter slightly, apparently via the -57HNF4 $\alpha$ RE since mutation of this site appeared to ablate the activation; this result is similar to the activation of the *UGT1A10* proximal promoter by HNF4 $\alpha$  alone. Importantly, the -0.19kb-*UGT1A9* construct showed synergistic induction by the combination of CDX2 and HNF4 $\alpha$  (Figure 3.10C), which was ablated by mutation of the -57HNF4 $\alpha$ RE. This data indicates that the novel CDX2/HNF4 $\alpha$  composite element that is conserved in intestinally expressed *UGT1A8*, -1A9 and -1A10, mediates synergistic induction of all three promoters by CDX2 and HNF4 $\alpha$ . We propose that this CDX2/HNF4 $\alpha$  composite element is a central regulatory module for intestinal UGT expression.

When the -1kb-*UGT1A9* promoter construct was co-transfected with CDX2 and HNF4 $\alpha$  in Caco-2 cells, we observed that HNF4 $\alpha$  alone greatly induced the promoter (Figure 3.10C). This result was consistent with the previous report that the -1kb-*UGT1A9* promoter is regulated by HNF4 $\alpha$  in liver cells (Barbier *et al.*, 2005). HNF4 $\alpha$ -mediated induction of the -1kb-*UGT1A9* promoter was reduced by around 40% upon mutation of the -57HNF4 $\alpha$ RE site (Figure 3.10C), suggesting that this newly identified motif only partially mediates HNF4 $\alpha$  action in Caco-2 cells; the multiple upstream HNF4 $\alpha$ REs in this construct likely contribute the remainder of the induction. Interestingly, co-

expression of CDX2 and HNF4 $\alpha$  reduced activation of the -1kb-*UGT1A9* promoter relative to HNF4 $\alpha$  alone. This latter result may indicate competition between binding of HNF4 $\alpha$  to the upstream HNF4 $\alpha$ REs and the proximal -57HNF4 $\alpha$ RE, as discussed later (section 3.3.8).

Overall, the data presented here indicates that the new HNF4 $\alpha$ /CDX2 composite binding element (designated as -44HNF4 $\alpha$ RE, -57HNF4 $\alpha$ RE and -47HNF $\alpha$ 4RE in the *UGT1A8*, -1A9 and -1A10 proximal promoters respectively) can mediate HNF4 $\alpha$ /CDX2 synergy on all three promoters.

Consistent with this idea, we also showed that all three promoters recruit HNF4 $\alpha$  to this region using chromatin immunoprecipitation (ChIP) assays in Caco-2 cells (Figure 3.10D). The discovery of this new composite element suggests a mechanism by which CDX2 can induce *UGT1A9* promoter activity in intestinal cells, given that the previously identified 'canonical' -70bp CDX2 RE was found to be non-functional (Gregory *et al.*, 2004b, Gregory *et al.*, 2004a, Gregory, 2004).



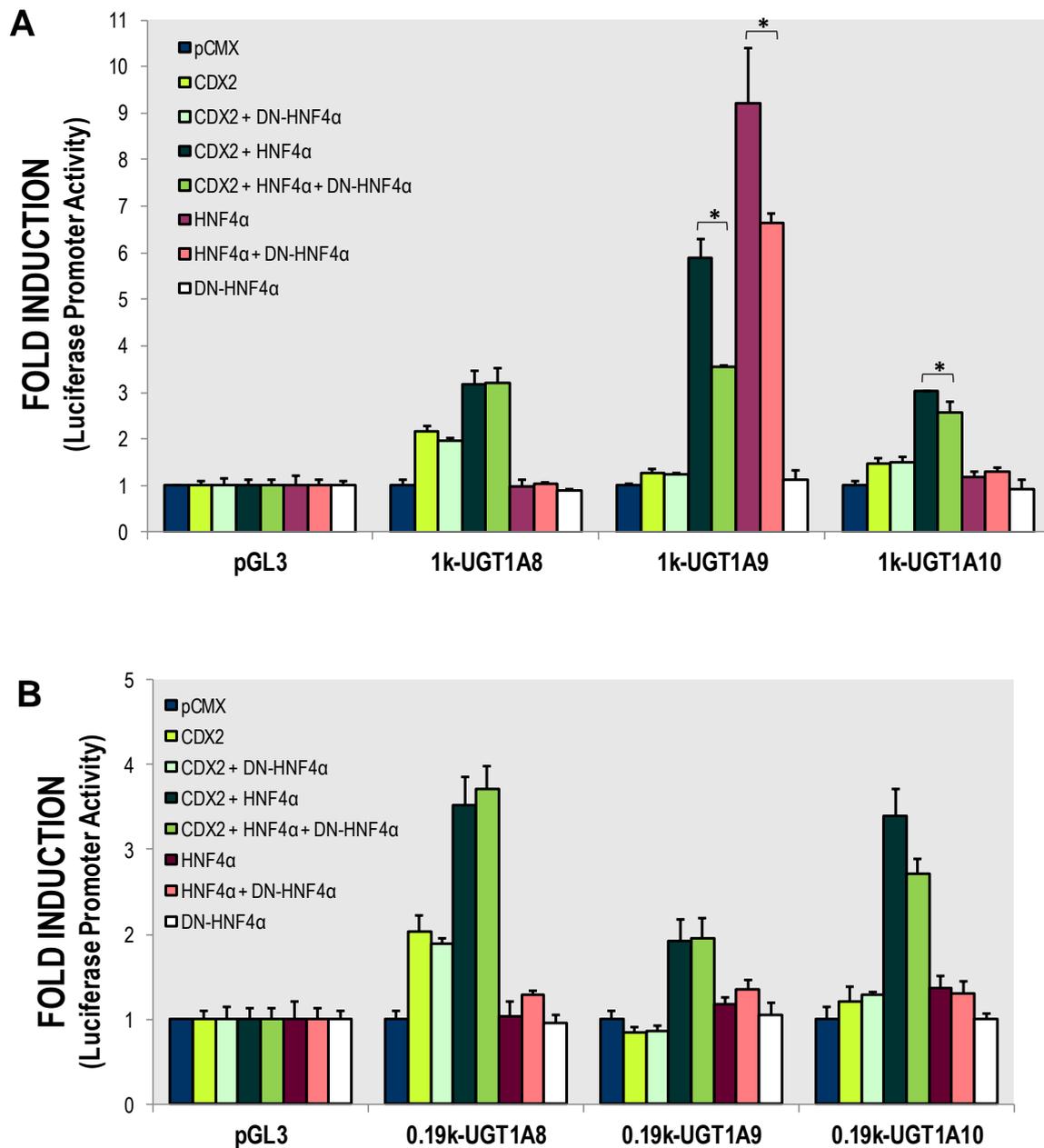
**Figure 3.10. The CDX2/HNF4 $\alpha$  composite element is functionally conserved in the *UGT1A8-1A10* promoters.**

A. Alignment of sequences of the *UGT1A8*, -1A9 and -1A10 promoters shows complete conservation of the CDX2/HNF4 $\alpha$  composite element. B. Synergistic induction UGT1A10 -1kb- and -0.19kb promoters by CDX2 and HNF4 $\alpha$ . Mutation of the -47HNF4 $\alpha$ RE (i.e. CDX2/HNF4 $\alpha$  composite element) prevents the promoter induction. C. HNF4 $\alpha$  alone but not CDX2 alone can activate the *UGT1A9* promoters. Synergistic induction of UGT1A9 promoter activity by CDX2 and HNF4 $\alpha$  is only observed in the short -0.19kb promoter; mutation the HNF4 $\alpha$  binding motif ablates the synergy. Mutation of the -57HNF4 $\alpha$ RE (i.e. CDX2/HNF4 $\alpha$  composite element) reduces induction of -1kb-UGT1A9 promoter by HNF4 $\alpha$ , and ablates synergistic activation of the -

0.19kb-UGT1A9 promoter. D. CHIP-qPCR analysis shows HNF4 $\alpha$  binding to the region spanning the conserved CDX2/HNF4 $\alpha$  synergistic composite element in the promoter region of the *UGT1A8*, *-1A9* and *-1A10* genes. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### **3.3.4 Dominant negative HNF4 $\alpha$ does not inhibit the ability of wild-type HNF4 $\alpha$ to synergise with CDX2 in regulation of the *UGT1A8* and *-1A9* promoters**

Previous work showed that a dominant negative (DN) form of HNF4 $\alpha$  could inhibit the activity of wildtype HNF4 $\alpha$  in regulation of target genes (unpublished data from our laboratory). We co-transfected the -1kb and -0.19kb *UGT1A8*, *-1A9* and *-1A10* promoters constructs with CDX2, HNF4 $\alpha$  or the combination of CDX2 and HNF4 $\alpha$  in the presence or absence of DN-HNF4 $\alpha$ . As shown in Figure 3.11A, HNF4 $\alpha$  (alone and in combination with CDX2) induced -1kb *UGT1A9* promoter activity, and co-transfection with DN-HNF4 $\alpha$  modestly reduced this induction, suggesting that DN-HNF4 $\alpha$  may compete with wildtype HNF4 $\alpha$  for binding to the HNF4 $\alpha$ REs within -1kb-*UGT1A9* promoter. The -0.19k-*UGT1A9* promoter was not induced by wildtype HNF4 $\alpha$  alone but was synergistically induced by the combination CDX2 and wildtype HNF4 $\alpha$ . Interestingly, DN-HNF4 $\alpha$  did not inhibit this synergistic induction of the -0.19k-*UGT1A9* promoter (Figure 3.11B). Similarly, DN-HNF4 $\alpha$  did not inhibit synergistic induction of the -0.19k-*UGT1A8* promoter by the combination of CDX2 and wildtype HNF4 $\alpha$ , although it did very slightly reduce synergistic induction of the -0.19bk *UGT1A10* promoter (Figure 3.11B). Overall, the observation that DN HNF4 $\alpha$  can inhibit the activity of wildtype HNF4 $\alpha$  at the distal *UGT1A9* HNF4 $\alpha$  elements (i.e. in the -1kb *UGT1A9* promoter) but is less effective or ineffective in inhibiting synergistic induction via the proximal CDX2/HNF4 $\alpha$  composite element (i.e. in the -0.19kb *UGT1A8-1A10* promoters) suggests that HNF4 $\alpha$  functions differently at these different elements. Future investigation is required to obtain an understanding of the structure-function relationships that may underlie these promoter-specific differences in HNF4 $\alpha$  activity terms of DNA-binding or co-factor binding.

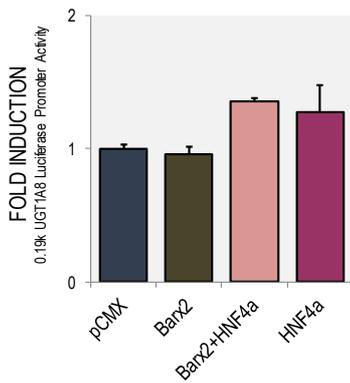


**Figure 3.11. The CDX2/HNF4 $\alpha$  composite element does not mediate dominant negative (DN) HNF4 $\alpha$  action.**

A. DN-HNF4 $\alpha$  (HNF4 $\alpha$ RR76/77) reduces the ability of wild-type (WT) HNF4 $\alpha$  alone or in combination with CDX2 to activate the -1kb-UGT1A9 promoter and to a lesser degree the -1kb UGT1A10 promoter, but not the -1kb-UGT1A8 promoter. B. DN-HNF4 $\alpha$  does not significantly inhibit the synergistic induction of the short -0.19kb-UGT1A8, -1A9 and -1A10 promoter by the combination of HNF4 $\alpha$  and CDX2. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### 3.3.5 The intestinal homeobox factor Barx2 does not synergise with HNF4 $\alpha$

Many homeobox proteins bind to a common recognition motif in DNA, these motifs are usually AT-rich and typically contain a core motif resembling the sequence TAAT. The specificity of different homeobox proteins for their target genes is likely determined by combinatorial interactions with other factors, as well as their own temporospatially restricted expression patterns. In addition to CDX2, a number of other homeobox genes are expressed in the intestinal epithelial lineage including the BAR-class homeobox protein Barx2. CDX2 expression is highest at the distal end of the GIT and Barx2 is expressed more proximally (Duprey *et al.*, 1988, Herring *et al.*, 2001, Walters *et al.*, 1997, Sander and Powell, 2004). In addition, while CDX2 expression increases towards the villus of the epithelial cells, Barx2 is abundant in the crypt and decreases as cells progress along the crypt-villus axis, suggesting that it may play a role in progenitor cell proliferation (Sander and Powell, 2004). Both CDX2 and Barx2 homeobox genes are reported to share target genes associated with cell adhesion such as L1-cadherin (Hinoi *et al.*, 2002, Sander and Powell, 2004), both are expressed in Caco-2 cells (Sander and Powell, 2004), and both are considered as markers for increased risk of colorectal cancer (Saandi *et al.*, 2013, Mi *et al.*, 2016). Other ongoing projects in our laboratory investigate the role of Barx2 in stem cell niches, hence a number of reagents for studying Barx2 function were readily available. Because we were interested in what factors might control the modest level of UGT expression observed in proliferating crypt cells, we decided to test whether Barx2 could induce intestinal UGT promoter activity, either alone or in combination with HNF4 $\alpha$ . In Caco-2 cells, we co-transfected the 0.19k-UGT1A8 promoter with Barx2, HNF4 $\alpha$ , or a combination of Barx2 and HNF4 $\alpha$ . As shown in Figure 3.12, the -0.19kb-UGT1A8 promoter did not respond to the expression of Barx2 or HNF4 $\alpha$  alone, nor to the combination of Barx2 and HNF4 $\alpha$ . This data suggests that the CDX2/HNF4 $\alpha$  composite element specifically mediates the activity of CDX2 and not other intestinal homeobox factors.



**Figure 3.12. The combination of Barx2 and HNF4 $\alpha$  does not induce the UGT1A8 promoter.**

The -0.19k-UGT1A8 promoter construct was co-transfected with HNF4 $\alpha$  and Barx2. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test, all values are means  $\pm$  SD.

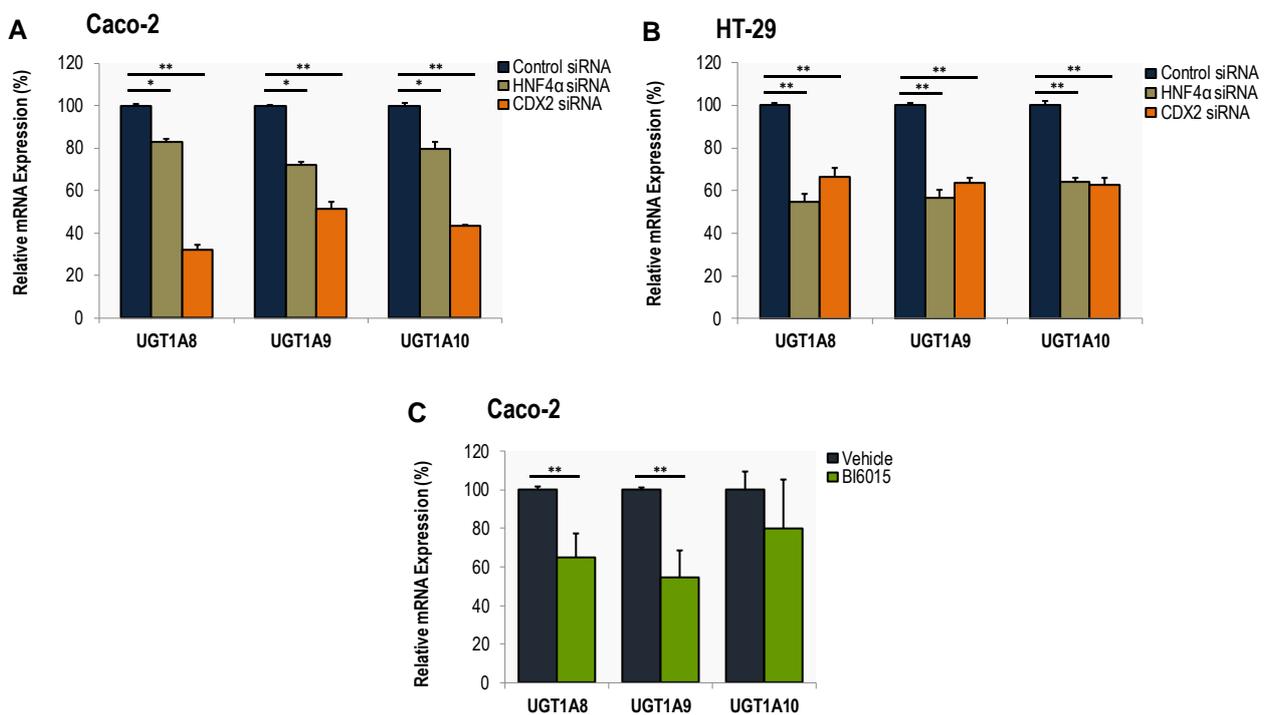
### 3.3.6 Regulation of endogenous UGT1A8-1A10 by CDX2 and HNF4 $\alpha$ in Caco-2 cells

CDX2 has been previously reported as a regulator of UGT1A8 expression, however, the previous study only examined *UGT1A8* promoter-reporter constructs and did not show induction of the endogenous UGT1A8 mRNA (Gregory *et al.*, 2004a). In this project, it was crucial to define the role of CDX2 and HNF4 $\alpha$  in regulation of endogenous UGT1A8, UGT1A9 and UGT1A10 mRNAs.

We determined that Caco2 cells express moderate-high levels of both HNF4 $\alpha$  and CDX2, hence we elected to use siRNA-mediated knockdown of these factors to assess their roles in regulation of the *UGT* genes. The efficacy of the siRNA in knockdown of CDX2 and HNF4 $\alpha$  is presented in section 3.3.7. As shown in Figure 3.13A, transfection of siRNA targeting HNF4 $\alpha$  in Caco-2 cells decreased the endogenous level of UGT1A8, -1A9 and 1A10 mRNAs by around 20-30%, whilst siRNA targeting CDX2 generated a 50-70% decrease of all three UGT mRNAs. To complement these data, we used another colon cancer cell line HT-29 that express a high level of UGT1A8-1A10 mRNAs (over 500x higher compared to that in Caco-2; see Figure 4.22 in Chapter 4). Knockdown of CDX2 and HNF4 $\alpha$  using siRNA reduced endogenous UGT1A8-1A10 mRNA levels by around 40-50% (Figure 3.13B), supporting our results from Caco-2 cells.

To further support the conclusion that endogenous *UGT1A8-1A10* genes are regulated by HNF4 $\alpha$ , we used the HNF4 $\alpha$  inhibitor BI6015 (Kiselyuk *et al.*, 2012). Treatment of Caco-2 cells with 2.5  $\mu$ M

BI6015 for 48 hours reduced all three intestinal UGT mRNAs, although the reduction was only significant for UGT1A8 and -1A9 (Figure 3.13C). Further studies to assess the mechanism of action of BI6015 with respect to HNF4 $\alpha$  and CDX2 level are presented in the section below (3.3.7). Overall these data indicate that both CDX2 and HNF4 $\alpha$  are required to maintain the expression level of endogenous *UGT1A8*, *-1A9* and *-1A10* genes in intestinal-derived Caco-2 and HT-29 cells, which is consistent with our model that UGT1A8-1A10 expression is determined by the CDX2/HNF4 $\alpha$  regulatory axis during intestinal development and homeostasis.



**Figure 3.13. Inhibition of endogenous UGT1A8-1A10 expression by inhibition of HNF4 $\alpha$  or CDX2 expression or function.**

A, B. In Caco-2 cells (A) and HT-29 cells (B), expression of endogenous UGT1A8-1A10 is decreased by siRNAs targeting CDX2 or HNF4 $\alpha$ . **C.** Treatment of Caco-2 cells with HNF4 $\alpha$  inhibitor BI6015 for 48 hours decreases UGT1A8 and UGT1A9 but not UGT1A10 mRNA levels. The mRNA expression was analysed by q-RT-PCR and normalised to 18S rRNA. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### 3.3.7 A regulatory feedback loop involving CDX2 and HNF4 $\alpha$ is confirmed at the mRNA level

As described in the Introduction to this Chapter, CDX2 and HNF4 $\alpha$  not only interact functionally at target genes, but they also regulate each other's expression (Olsen *et al.*, 2012); CDX2 binds and regulates the HNF4 $\alpha$  promoter (Verzi *et al.*, 2013), and reciprocally, HNF4 $\alpha$  binds and regulates the CDX2 promoter (Saandi *et al.*, 2013, Boyd *et al.*, 2009). Given that CDX2 and HNF4 $\alpha$  are key factors in intestinal development as well as homeostasis, understanding maintenance of these factors is important. Moreover, many studies have shown that in colorectal cancers (CRCs), the CDX2 level is reduced (Hinoi *et al.*, 2001, Subtil *et al.*, 2007, Brabletz *et al.*, 2004) and this is typically followed by a decrease of HNF4 $\alpha$ . The loss of both factors is likely to facilitate colorectal cancer progression by promoting cellular de-differentiation (Saandi *et al.*, 2013).

We sought to confirm if CDX2 and HNF4 $\alpha$  regulate each other in the intestinal derived cell lines used in our study. As shown in Figure 3.14, knockdown of CDX2 with siRNA in both Caco-2 and HT-29 cell lines decreased not only CDX2 mRNA, but also HNF4 $\alpha$  mRNA by 60-80%.

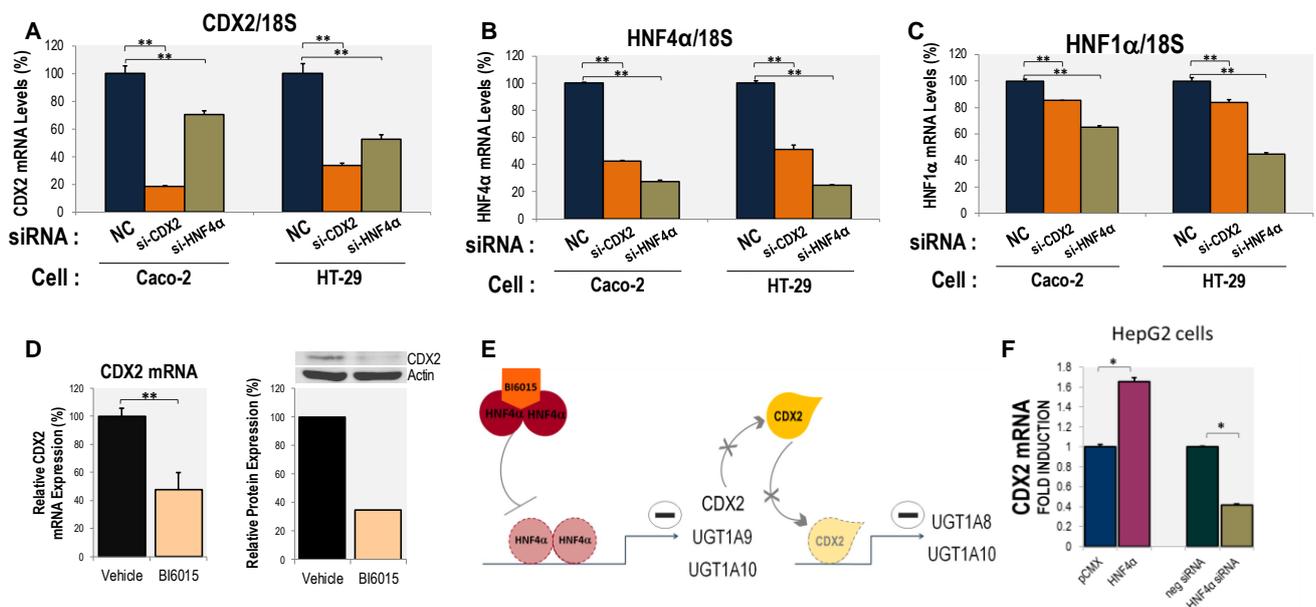
Reciprocally, siRNA targeting HNF4 $\alpha$  reduced levels of both HNF4 $\alpha$  and CDX2 by around 35-80%. We also measured the mRNA level of a well-known HNF4 $\alpha$  direct target gene, HNF1 $\alpha$  (Hansen *et al.*, 2002). As expected, HNF4 $\alpha$  knockdown decreased HNF1 $\alpha$  mRNA level by ~40% in Caco-2 cells and ~60% in HT-29 cells. CDX2 siRNA also produced a modest decrease in HNF1 $\alpha$  endogenous expression by ~20% in both cell lines, likely due to the reduction in HNF4 $\alpha$  levels (Figure 3.14C). Overall these data indicate that reciprocal positive regulation of HNF4 $\alpha$  and CDX2 occurs in both Caco-2 and HT-29 cells

The capability of the HNF4 $\alpha$  inhibitor BI6015 to decrease HNF4 $\alpha$  target genes has been reported (Kiselyuk *et al.*, 2012); while the mechanism of its action is not fully defined, studies suggest that it may reduce the stability of HNF4 $\alpha$  protein and/or prevent its binding to DNA. Given that HNF4 $\alpha$  regulates CDX2, we tested whether CDX2 level was affected by BI6015. Treatment of Caco-2 cells for 48-72 hours at 2.5  $\mu$ M reduced the CDX2 mRNA level by ~60%, and consistent with that, BI6015 decreased CDX2 protein level by around 65% (Figure 3.14D). This is the first

evidence showing BI6015 is capable of affecting HNF4 $\alpha$  target genes, including CDX2, in an intestinal cell context. In Figure 3.14E, we present a model for the inhibitory effect of BI6015 on UGT expression both directly via inhibition of HNF4 $\alpha$  and indirectly via inhibition of CDX2. This chemical might be useful to investigate HNF4 $\alpha$ -mediated gene regulation *in vivo* using mouse models of intestinal development or regeneration.

Finally, we applied both HNF4 $\alpha$  overexpression and siRNA-mediated HNF4 $\alpha$  knockdown in liver-derived (hepatocellular carcinoma) HepG2 cells to determine whether HNF4 $\alpha$  may be a general regulator of CDX2 expression (i.e. context independent). As shown in Figure 3.14F, HNF4 $\alpha$  overexpression could modestly induce CDX2 mRNA expression in HepG2 cells and surprisingly, HNF4 $\alpha$  knockdown could reduce the very low level of CDX2 mRNA expressed in these cells.

Overall these data are consistent with previous reports (Saandi *et al.*, 2013, Boyd *et al.*, 2009) that HNF4 $\alpha$  is a robust regulator of CDX2 in various cellular contexts.



**Figure 3.14. CDX2 and HNF4 $\alpha$  regulate each other's expression in Caco-2 and HT-29 colon cancer cell lines.**

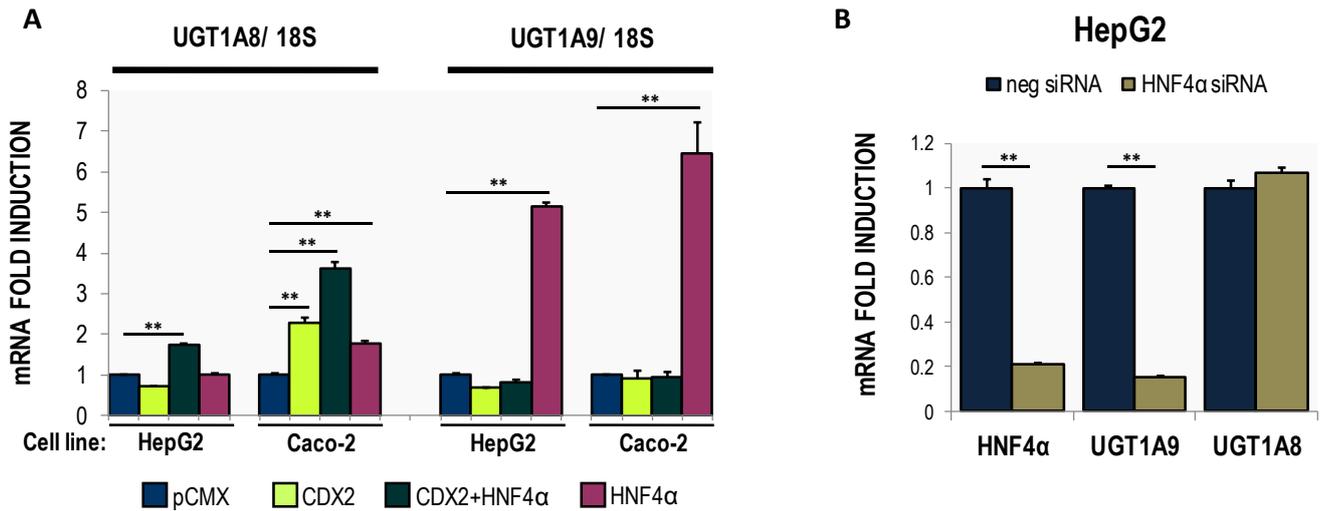
A. Transfection of siRNA targeting HNF4 $\alpha$  decreases CDX2 mRNA level. B. Transfection of siRNA targeting CDX2 decreases HNF4 $\alpha$  mRNA level. C. Transfection of siRNA targeting HNF4 $\alpha$  produces a decrease in mRNA level of HNF1 $\alpha$ , a direct target of HNF4 $\alpha$ . D. HNF4 $\alpha$  inhibitor BI6015 decreases mRNA level of CDX2 and protein. E. A proposed model of BI6015 action shows that BI6015 inhibits HNF4 $\alpha$  binding to the DNA directly inhibiting HNF4 $\alpha$  target genes including CDX2 and UGTs; reduction in CDX2 level further inhibits UGT gene expression. F. Induction of CDX2 level by HNF4 $\alpha$  in HepG2 cells. The mRNA expression was analysed by qRT-PCR and normalised to 18S rRNA. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

### 3.3.8 Differential CDX2 and HNF4 $\alpha$ functions in the regulation of UGT1A8 and -1A9 in hepatic vs intestinal contexts

Our findings have revealed a novel CDX2/HNF4 $\alpha$  composite element that is shared between *UGT1A8*, -1A9 and -1A10 and mediates a similar synergistic transcriptional mechanism in all three promoters. Given that the -70CDX2 element in the *UGT1A9* promoter was previously defined as non-functional, the binding of CDX2 and HNF4 $\alpha$  to the newly identified composite element provides a mechanism for intestinal induction of *UGT1A9* by CDX2. In addition, the -1kb *UGT1A9* promoter contains upstream HNF4 $\alpha$  binding sites that were previously defined as functional in *UGT1A9*, but non-functional in the *UGT1A8* and -1A10 promoters due to sequence divergences (Barbier *et al.*, 2005). The latter was suggested to be responsible for the lack of hepatic *UGT1A8* and -1A10 expression (Barbier *et al.*, 2005). To better understand the mechanism of intestinal specificity, we compared regulation of *UGT1A9* and *UGT1A8* by CDX2 and HNF4 $\alpha$  at the mRNA level.

Using both hepatic (HepG2) and intestinal (Caco-2) cell lines, we first examined whether overexpression of CDX2 and HNF4 $\alpha$  had different effects on the levels of endogenous *UGT1A8* and -1A9 mRNAs. As presented in Figure 3.15A, in Caco-2 cells, CDX2 induced *UGT1A8* mRNA and there was a synergistic induction CDX2 and HNF4 $\alpha$  similar to that seen in our promoter analyses (Figure 3.15A). In HepG2 cells, in which the expression levels of CDX2 and *UGT1A8* are very low (barely detectable), overexpression of CDX2 or HNF4 $\alpha$  alone had no effect on *UGT1A8* mRNA level while the combination of CDX2 and HNF4 $\alpha$  produced a very slight induction of *UGT1A8* mRNA.

In both HepG2 and Caco-2 cells, overexpression of HNF4 $\alpha$  alone induced *UGT1A9* mRNA level by around 6-7 fold. In contrast, CDX2 alone had no effect on *UGT1A9* expression and, similar to the promoter assay, the combination of CDX2 and HNF4 $\alpha$  reduced activation by HNF4 $\alpha$  (Figure 3.15A). In further support of these findings, in HepG2 cells, HNF4 $\alpha$  siRNA had no effect on *UGT1A8* mRNA, but dramatically reduced *UGT1A9* mRNA levels (Figure 3.15B).



**Figure 3.15. UGT1A8 and -1A9 are differently regulated by CDX2 and HNF4 $\alpha$  in hepatic and intestinal cell models.**

A. In HepG2 cells, UGT1A8 mRNA can be induced by overexpression of both CDX2 and HNF4 $\alpha$ , but not by CDX2 or HNF4 $\alpha$  alone. In Caco-2 cells, CDX2 alone induces UGT1A8 mRNA and synergy with HNF4 $\alpha$  further increases UGT1A8 mRNA level. In both HepG2 and Caco-2, endogenous UGT1A9 mRNA can only be induced by overexpression of HNF4 $\alpha$ ; moreover, expression of CDX2 inhibits induction by HNF4 $\alpha$ . B. In HepG2 cells, siRNA mediated inhibition of HNF4 $\alpha$  reduces UGT1A9 mRNA but shows no effect on UGT1A8 level. The mRNA expression was analysed by qRT-PCR and normalised to 18S rRNA. For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means  $\pm$  SD.

Overall, these data are broadly consistent with our promoter-reporter data and indicate that the endogenous *UGT1A8* gene requires CDX2 for induction by HNF4 $\alpha$ , which is consistent with its lack of hepatic expression. In contrast, HNF4 $\alpha$  can induce *UGT1A9* mRNA expression in a CDX2-independent manner, which is consistent with its robust hepatic expression. The ability of CDX2 to inhibit induction of UGT1A9 by HNF4 $\alpha$  might be explained by the formation of CDX2-HNF4 $\alpha$  complexes (that preferentially bind to the composite element) at the expense of HNF4 $\alpha$  complexes that can bind and activate at the consensus HNF4 $\alpha$  elements upstream within the *UGT1A9* promoter. It is also likely that CDX2 directs chromatin modifications that control promoter accessibility to HNF4 $\alpha$  (Verzi *et al.*, 2010, San Roman *et al.*, 2015).

To better define the HNF4 $\alpha$  elements that are involved in *UGT1A9* induction by HNF4 $\alpha$  in intestinal cells, we compared *UGT1A9* promoter constructs carrying mutations in three different predicted HNF4 $\alpha$  binding elements (see Figure 3.16). The HNF4 $\alpha$  element at -801 to -789 bp from the

*UGT1A9*-TSS (marked as element A in Figure 3.16A) overlaps with the -808/-795 PPAR-binding element previously reported to be functional in hepatic *UGT1A9* regulation (Barbier *et al.*, 2003). The HNF4 $\alpha$  element at -334 to -322bp from *UGT1A9*-TSS (marked as element B in Figure 3.16A) was previously shown to play a role in hepatic *UGT1A9* regulation by HNF4 $\alpha$  (Barbier *et al.*, 2005), the absence of this functional element was proposed to explain the lack of *UGT1A8* and *-1A10* expression in the liver. The HNF4 $\alpha$  element at -57 bp from the *UGT1A9* TSS (marked as element C in Figure 3.16A) is our newly identified CDX2/HNF4 $\alpha$  composite element (also designated -57HNF4 $\alpha$ RE). Promoter-reporter constructs containing the -1kb-*UGT1A9* promoter region with mutations in various combinations of these HNF4 $\alpha$  elements (Figure 3.16A) were tested for activation by transfection of a HNF4 $\alpha$  expression plasmid in Caco2 cells.

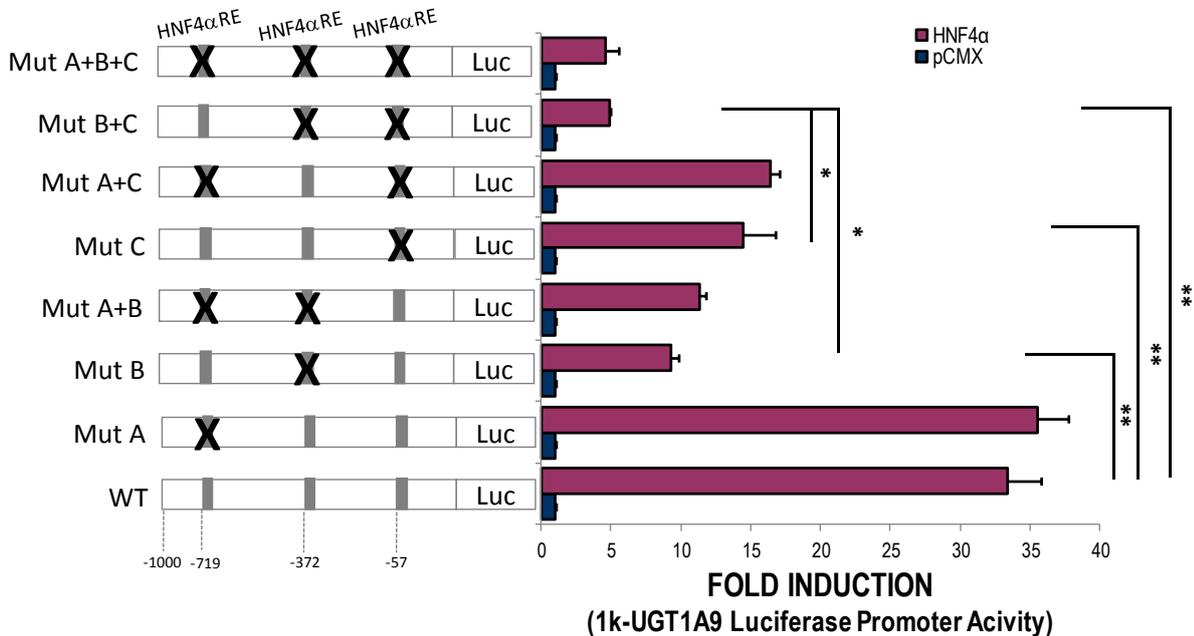
As shown in Figure 3.16B, HNF4 $\alpha$  induced wild-type -1kb *UGT1A9* promoter activity by 33 fold. When the -801HNF4 $\alpha$ RE was mutated, the level of induction by HNF4 $\alpha$  did not change, suggesting that this predicted HNF4 $\alpha$ RE was not functional. Mutating the -334HNF4 $\alpha$ RE decreased the ability of HNF4 $\alpha$  to induce promoter activity by ~70%, and mutating -57HNF4 $\alpha$ RE decreased the ability of HNF4 $\alpha$  to induce promoter activity by ~60%; when both motifs were mutated, the promoter induction was reduced by ~80% (Figure 3.16B). Previous work by Barbier *et al.* showed in HepG2 cells that loss of the -334HNF4 $\alpha$ RE completely abolished promoter induction (Barbier *et al.*, 2005). In contrast while we found this element to be important, even after mutation of both the -334bp and -57bp elements, the promoter could be induced by HNF4 by around 5 fold . These data imply that there are other functional HNF4 $\alpha$  elements within the 1kb *UGT1A9* promoter that contribute to its activation by HNF4 $\alpha$  in Caco2 cells. Further work would be required to identify these elements and to determine how they might interact with CDX2 mediated regulation.

UGT1A9 HNF4α RE	Wild type (WT) Sequence	Mutant Sequence
- 801 to - 789 <sup>A)</sup>	T G A C C T C A G G G A G	T G A T T T C A C G G A G
- 334 to - 322 <sup>B)</sup>	G G G A C A A A T T C C A	G G G A T G A A T T C C A
- 57 to - 45 <sup>C)</sup>	T C T A T T G G G T C A	T C T A T T G G A A T C A

<sup>A)</sup> Identified by Barbier et al; 2003; overlaps DR1 motif.

<sup>B)</sup> Identified by Barbier et al; 2004.

<sup>C)</sup> Identified in this study.



**Figure 3.16. HNF4α activation via predicted and functional motifs within 1kb of the *UGT1A9* promoter region in the intestinal Caco-2 cell line.**

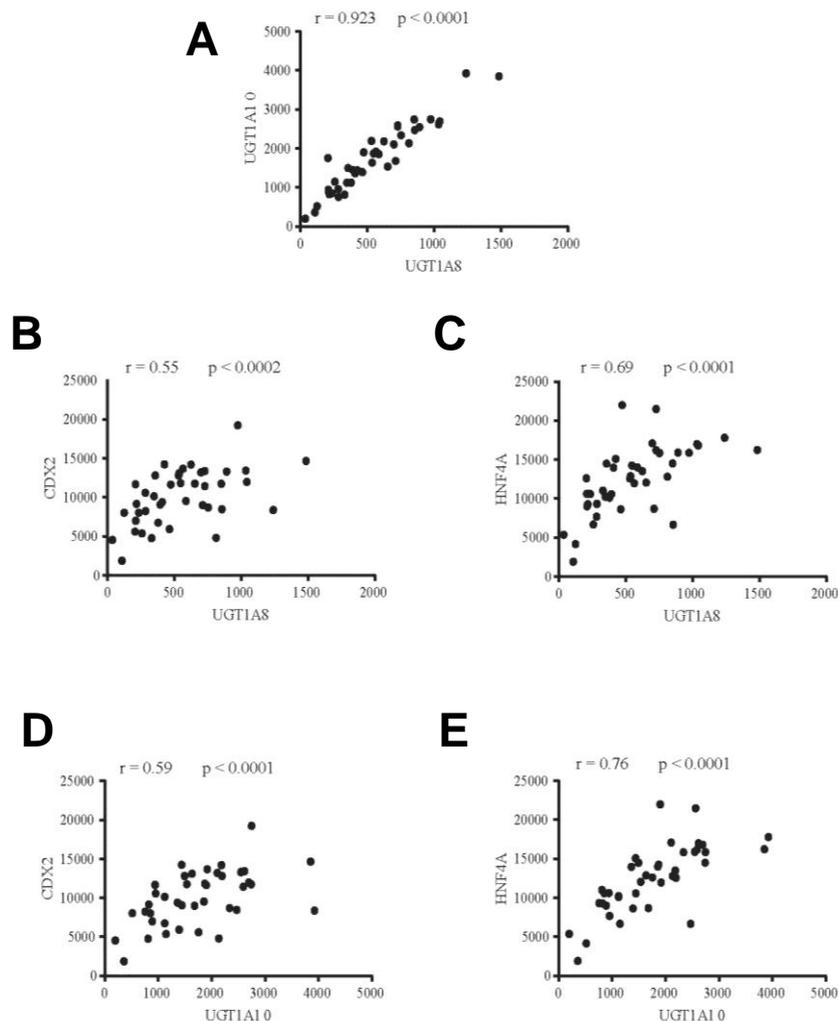
A. Table showing three functional HNF4α binding sites in the *UGT1A9* promoter region (-801bp, -334bp, and -57bp). Mutant -1kb-*UGT1A9* promoter constructs were generated with mutated bases indicated by red font.

B. 1k-*UGT1A9* wild-type and mutant constructs were co-transfected with HNF4α plasmids in Caco-2 cells. The predicted -801bp HNF4αRE that overlaps PPAR-RE identified by (Barbier, et al, 2003) is not required for induction by HNF4α. The -334HNF4αRE that was previously reported (Barbier *et al.*, 2005) is functional, but mutation does not completely abolish induction by HNF4α. The new -57bp HNF4αRE identified in this study is functional but mutation does not completely abolish induction by HNF4α. Relative luciferase activity was normalised to the promoter-less pGL3 basic vector and induction by transfection of HNF4α expression vector was expressed as fold over empty vector pCMX (set as a value of 1). For each dataset, n= 3 or greater; (\* indicates statistical significant differences using one-way ANOVA, Tukey's post-hoc test at \*P < 0.05; \*\* P < 0.001. All values are means ± SD.

### 3.3.9 Data analysis from TCGA database supports positive correlation of *UGT1A8* and -1A10 expression with *CDX2* and *HNF4α* levels in human colon samples

To gain insight into the importance of *CDX2* and *HNF4a* for *UGT1A8-10* regulation in an *in vivo* context, we took advantage of the TCGA database that records gene expression data from thousands of human normal and cancer samples. Using the colon adenocarcinoma dataset (COAD), we investigated the correlation between the expression level of *UGT1A8/-1A10* and the

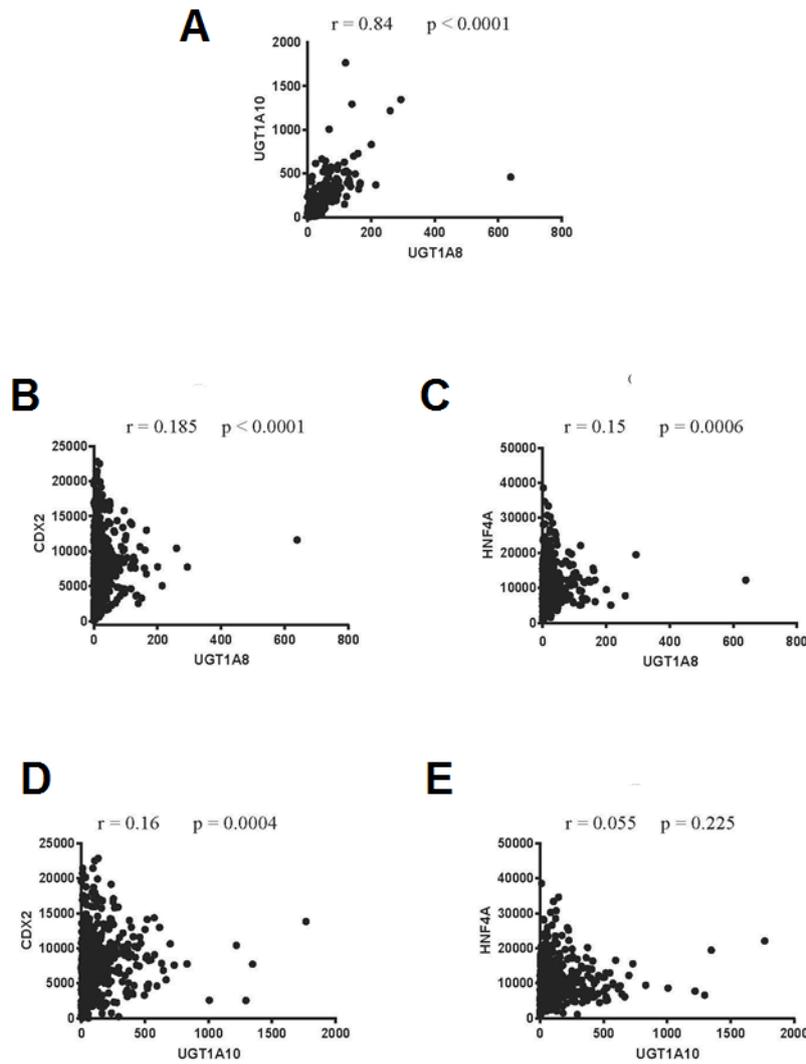
level of CDX2 and HNF4 $\alpha$ . This analysis was performed by colleagues Dr. Dong Gui Hu and Dr. Robyn Meech. Figure 3.17 shows the correlation analysis in normal colon samples (n= 41). The mRNA level of UGT1A8 and UGT1A10 was shown to be extremely tightly correlated (A). A very robust correlation was also shown for both UGT1A8 and -1A10 with the level of CDX2 and HNF4 $\alpha$  (B, C, D, E).



**Figure 3.17. Correlation analysis of UGT1A8-1A10, CDX2 and HNF4 $\alpha$  levels in normal colon samples (n=41) from the Colon Adenocarcinoma (COAD) dataset of The Cancer Genome Atlas (TCGA).**

A. In normal colon samples, UGT1A8 and UGT1A10 levels are extremely tightly correlated. B, C. UGT1A8 shows a robust correlation with levels of both CDX2 (B) and HNF4 $\alpha$  (C). D, E. UGT1A10 shows a robust correlation with levels of both CDX2 (D) and HNF4 $\alpha$  (E).

When we examined the colon cancer sample set (n= 480), correlation between UGT1A8 and -1A10 level was still strong although less robust; and correlation of both UGTs with the level of CDX2 and HNF4 $\alpha$  was weaker but still statistically significant, except for correlation between UGT1A10 and HNF4 $\alpha$  (Figure 3.18). Consistent with CDX2 deregulation during colon cancer progression, which co-occurs with HNF4 $\alpha$  decrease (Saandi *et al.*, 2013), alteration of CDX2 and HNF4 $\alpha$  levels in colon cancer may affect the maintenance of UGT1A8 and -1A10 expression. The CDX2 and HNF4 $\alpha$  partnership is important for regulation of proliferation and differentiation in intestinal cells (San Roman *et al.*, 2015); however these functions could be dysregulated in cancer and although different tumours may express similar levels of CDX2 and HNF4 $\alpha$ , these may confer differing downstream programs. The heterogeneous nature of this dysregulation in different tumours may explain why there is less robust correlation of CDX2 and HNF4 $\alpha$  with UGT1A8 and -1A10 in colon cancer samples. We have also gained some insights into the role of these factors in regulation of UGTs in cancer by analysis of cancer cell lines. The Caco-2 cell model presents a relatively differentiated phenotype, whilst HCT116 colon cancer cells show a completely undifferentiated phenotype (Yeung *et al.*, 2010). We found that Caco-2 cells have moderate expression of UGT1A8-1A10 and high levels of CDX2 and HNF4 $\alpha$ , whereas HCT116 cells have no UGT1A8-1A10, no HNF4 $\alpha$  and low CDX2 levels (data not shown). Thus at least in these cancer cell line models, there remains a correlation of UGT1A8-10 level with CDX2 and HNF4 $\alpha$  expression.



**Figure 3.18. Analysis of UGT1A8-1A10, CDX2 and HNF4 $\alpha$  levels in colon cancer samples (n=480) from the TCGA dataset.**

A. UGT1A8 and UGT1A10 levels are tightly correlated. B, C. UGT1A8 shows a significant correlation with levels of both CDX2 (B) and HNF4 $\alpha$  (C). D, E. UGT1A10 shows a significant correlation with the level of CDX2 (D) but not HNF4 $\alpha$  (E).

### 3.4 Conclusions

Recent studies in mouse models indicate that intestinal development during embryogenesis, and maintenance of intestinal homeostasis in adults, involve CDX2 and HNF4 $\alpha$  acting in partnership to regulate expression of numerous intestine-specific genes. In this study, we show for the first time that UGT1A8, -1A9 and 1A10 expression in intestinal cells is also regulated by the combinatorial action of CDX2 and HNF4 $\alpha$ , potentially determining tissue-specific patterning of these UGTs in the intestine.

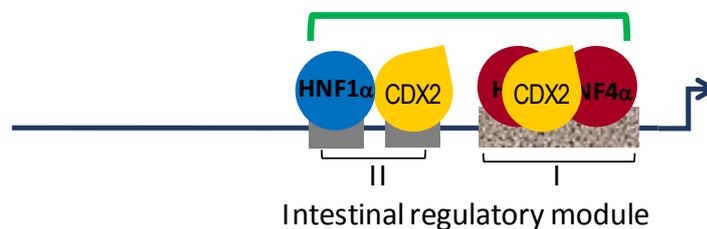
The major finding of this study is the identification of a novel composite regulatory element located in the proximal promoters of *UGT1A8*, *-1A9* and *-1A10* that binds to both CDX2 and HNF4 $\alpha$ . The element is fully conserved in these three promoters, and mediates their similar regulation in an intestinal-cell context. Originally predicted in the *UGT1A8* promoter using bioinformatic analysis as a likely HNF4 $\alpha$  binding motif (and hence designated the *UGT1A8*-44HNF4 $\alpha$ RE), we initially hypothesized that CDX2 was recruited to the element indirectly via its interaction with HNF4 $\alpha$ . However, subsequent data from EMSA and mutational analysis dismissed this hypothesis and revealed that the element in fact contains a cryptic CDX2 binding motif as well as the HNF4 $\alpha$  binding motif. This element mediates synergistic induction of the *UGT1A8*, *-1A9*, and *-1A10* proximal promoter regions by CDX2 and HNF4 $\alpha$ . Furthermore the synergy requires both the CDX2 and HNF4 $\alpha$  motifs to be intact; the loss of one motif cannot be compensated for by the other, strongly suggesting concomitant occupation by CDX2 and HNF4 $\alpha$ . Although CDX2 and HNF4 $\alpha$  have been shown by ChIP-Seq analysis to bind in close proximity within chromatin (Verzi *et al.*, 2010), such studies could not precisely determine the relative position of CDX2 and HNF4 $\alpha$  binding sites, therefore, to our knowledge, this is the first study reporting integration of CDX2 and HNF4 $\alpha$  within a very short sequence of 12 nucleotides.

It is known that CDX2, HNF1 $\alpha$ , and HNF4 $\alpha$  interact in a transcriptional regulatory network during intestinal cell differentiation (Olsen *et al.*, 2012). Among these three factors, HNF4 $\alpha$  and HNF1 $\alpha$  are expressed both in the liver and intestine, whereas CDX2 is expressed only in the intestine and is also regarded as the intestinal master regulator.

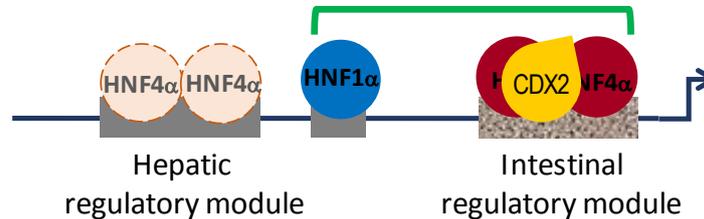
Previous work by our group identified a CDX2 binding element at around -70bp in the *UGT1A8* promoter that mediates combinatorial regulation by CDX2 and HNF1 $\alpha$  (Gregory *et al.*, 2004a). This CDX2 element does not mediate synergistic induction by CDX2 and HNF4 $\alpha$  (as shown in this study). Moreover, this element was reported to be mutated in *UGT1A9* and unable to effectively bind CDX2 (Gregory *et al.*, 2004a). This finding left an unresolved the problem of how *UGT1A9* is induced in intestinal cells. Our findings in this current study have largely resolved this problem by

showing that the novel conserved CDX2/HNF4 $\alpha$  composite element can mediate induction of UGT1A8, -1A9 and -1A10 in intestinal cells by CDX2 independently of the -70CDX2RE. A cartoon depicting UGT1A8, -1A9 and -1A10 regulation by CDX2 and HNF4 $\alpha$  is presented in Figure 3.19. The model defines two elements in the *UGT1A8* and -1A10 promoters that can mediate regulation by CDX2 in intestinal cells: module I is the novel composite CDX2/HNF4 $\alpha$ RE that mediates CDX2/HNF4 $\alpha$  synergy; module II is the -70CDX2RE that is involved in CDX2/HNF1 $\alpha$  synergy (Figure 3.19A).

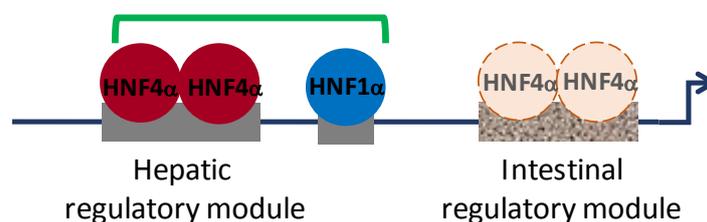
**A UGT1A8/10 – intestinal expression (High level of CDX2)**



**B UGT1A9 – intestinal expression (High level of CDX2)**



**C UGT1A9 – hepatic expression (Absence/low level of CDX2)**



**Figure 3.19. A model showing differential regulation of specific intestinal UGT1A8 and hepatic/intestinal UGT1A9 by CDX2 and HNF4 $\alpha$ .**

A. In the *UGT1A8/10* proximal promoters, a two part intestinal module that includes (I) the new HNF4 $\alpha$ /CDX2 composite element (-44bp in *UGT1A8*), and (II) the previously defined CDX2 (-70bp) and HNF1 $\alpha$  sites (-100bp). When CDX2 is high (e.g. intestine), the proximal HNF4 $\alpha$ /CDX2 composite element (I) recruits HNF4 $\alpha$  and CDX2 and mediates cooperative regulation by these two factors, the upstream elements (II)

likely 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 the -372bp HNF4 $\alpha$ RE. When CDX2 is high (e.g. intestine), HNF4 $\alpha$  may be recruited mainly to the intestinal module where it can also cooperate with HNF1 $\alpha$  (green bracket). C. When CDX2 is low/absent (e.g. liver), HNF4 $\alpha$  may be recruited mainly to the hepatic module where it can also cooperate with HNF1 $\alpha$  (green bracket). Chromatin structures set during cell fate restriction may help determine the relative accessibility of these modules.

*UGT1A9* is the only UGT within the *UGT1A8-1A10* cluster that is expressed in the liver. Within the -1kb *UGT1A9* promoter, there are two regulatory modules which we propose are differentially active in the hepatic and intestinal contexts (Figure 3.19B and C). When CDX2 is absent or low, such as in liver cells, we suggest that HNF4 $\alpha$  forms complexes that bind to the distal/upstream HNF4 $\alpha$  elements (defined as a 'hepatic regulatory module') to regulate *UGT1A9* (Figure 3.19C). When CDX2 is abundant and HNF4 $\alpha$  levels are limiting such as in intestinal cells, HNF4 $\alpha$  forms complexes with CDX2 that may bind preferentially to the conserved proximal CDX2/HNF4 $\alpha$  composite element (defined here as an 'intestinal module') (Figure 3.19B). The ability of CDX2 to direct HNF4 $\alpha$  binding to the 'intestinal regulatory module' may also involve modifying chromatin structure and thus binding site accessibility. Lower HNF4 $\alpha$  levels in the intestine may also partly explain why *UGT1A9* expression in the intestine is lower as a percentage of total UGT expression than in the liver (Guillemette *et al.*, 2014).

In this study, inhibition of CDX2 or HNF4 $\alpha$  was associated with reduction of UGT mRNA levels in intestinal cells. Low expression of intestinal UGTs may play a role in colon cancer progression associated with CDX2/HNF4 $\alpha$  deregulation. Our preliminary findings in human colon samples reveal that the correlation of intestinal UGT levels with CDX2 and HNF4 $\alpha$  levels is less robust in colon cancer samples (Figure 3.17) than in normal colon (Figure 3.18). However, further investigation is required to confirm or refute this idea.

High expression of CDX2 and HNF4 $\alpha$  is associated with intestinal cell differentiation and this is the context in which UGTs are also most highly expressed. However, there is also moderate UGT expression in proliferating crypt cells and one of our broader interests is determining factors that might control its expression in this context. We did not find any evidence that Barx2, a homeobox

factor expressed in crypt cells, could regulate UGT1A8 in Caco-2 cells. However more sophisticated experimental models that represent the intestine structurally, functionally and physiologically are required to fully understand the developmental patterning of intestinal UGTs. This work could use transgenic mice carrying the promoter regions of human intestinal UGTs as described in Chapter 5, or the previously described 'humanised' UGT mice that carry the entire human *UGT1* locus. In addition, human or transgenic mouse intestinal organoids could be developed in which transcription factor levels can be manipulated using tool such as viral transduction. Studies described in Chapter 5 begin to address the viability of these more complex models for studying UGT regulation and function.

Chromatin structures are not only programmed during cell fate determination, but also regulated dynamically by dietary compounds (e.g. butyrate, flavonoids and certain drugs); moreover there is recent evidence that such compounds may also regulate HNF4 $\alpha$  directly (Hwang-Verslues and Sladek, 2010, Li *et al.*, 2015). The overarching goal of this project was to provide a better understanding of the regulators of developmental/constitutive UGT expression in the GIT (i.e. studies described in this Chapter), as well as the mechanisms involved in inducible expression by exogenous factors (studies described in Chapter 4).

Ultimately, the findings presented here could have important consequences in the area of pharmacogenomic and personalized health care. Understanding the complex interplay of developmental programs and exogenous signals that underlie the wide inter-individual variation in *UGT1A8-1A10* mRNA levels seen in adult intestine is valuable for researchers physicians and the pharmaceutical industry for the identification of biomarker genes that can predict drug response and toxicity.

Oral administration is the most frequently used, convenient and economic route of drug administration, however, drug bioavailability is influenced by gastro-intestinal metabolism, which shows considerable interindividual variation. A better understanding of this variation in intestinal drug metabolism and detoxification capacity, in which intestinal UGTs play a significant role, could ultimately lead to improvements in individualizing drug treatment plans for patients.

# CHAPTER 4. *UGT1A8* GENE REGULATION IN RESPONSE TO FLAVONOIDS, BUTYRATE INVOLVEMENT AND MECHANISM ANALYSIS

## 4.1 Introduction

The gastrointestinal tract (GIT) is the organ where foods, drugs and other ingested xenobiotics first encounter the detoxification mechanisms mediated largely by UGT enzymes. Among the vast number of active UGTs residing in the digestive tract, the exclusively extrahepatically expressed UGT enzyme, UGT1A8, plays a major role in intestinal metabolism. As described in Chapter 1, UGT1A8 plays an important role in the metabolism of carcinogenic materials (Jefcoate *et al.*, 2000, Zhu and Conney, 1998, Mojarrabi and Mackenzie, 1998, Cheng *et al.*, 1998, Nowell *et al.*, 1999, Gagné *et al.*, 2002, Thibaudeau *et al.*, 2006) and low UGT1A8 activity has been associated with a elevated risk of GIT cancers, such as oesophageal and colorectal cancer (Dura *et al.*, 2012, Wang *et al.*, 2013).

On the other hand, high pre-systemic glucuronidation of flavanoids by UGT1A8 has been reported to contribute to the low bioavailability of these dietary compounds (Mojarrabi and Mackenzie, 1998, Cheng *et al.*, 1999). In general, flavonoids have protective effects inhibiting carcinogenesis as well as other diseases (reviewed in (Thilakarathna and Rupasinghe, 2013)). Given the complex multifactorial roles of glucuronidation in the GIT, it is important to understand the how local UGT expression is dynamically controlled by chemicals in GIT. In the study described in this Chapter, we proposed to investigate a potential feedback loop mechanism where flavonoids may induce UGT1A8 expression by acting as ligands for transcription factors that activate the UGT1A8 gene promoter, leading to increase of UGT1A8 activity towards its substrates. Involvement of flavonoids in induction of UGT1A8 expression would link the purported cancer-protective properties of flavonoids to the prominent role that UGTs play in detoxification of carcinogens and protection against carcinogenesis (detail description in Chapter 1, section 1.6.1).

An overarching paradigm guiding our studies into UGT enzymes is that these enzymes form part of a temporally responsive chemical detoxification feedback system. This notion has been

comprehensively reviewed by Bock (Bock, 2012), who described how UGTs such as UGT1A1, UGT2B4 and UGT2B7 are regulated by their own substrates, creating regulatory feedback circuits. Ligand-activated transcription factors/nuclear receptors (NRs) are a class of transcription factors that are activated by binding to small molecules including xenobiotics. Once activated, these NRs can induce expression of genes involved in xenobiotic metabolism, thus providing increased detoxification capacity (Bock, 2010). Several such NRs are known to directly regulate UGT genes. For example, baicalein and 3-hydroxyflavone induce UGT1A1 via the aryl hydrocarbon receptor (Majid *et al.*) and pregnane X receptor (PXR) respectively (Hiura *et al.*, 2014).

As described in Chapter 1, section 1.4.3, intestinal expression of the UGT1A family is localised to the intestinal epithelial surface and is concentrated in villi which have the most exposure to the luminal content, including dietary xenobiotics. Differentiated Caco-2 cells have been described as developing morphological and biochemical features similar to human enterocytes (Tremblay *et al.*, 2006). Hence this culture system has proven useful to model enterocyte functions and gene expression mechanisms (Soutoglou and Talianidis, 2002, Halbleib *et al.*, 2007). In studies described in this Chapter, the differentiated Caco-2 culture model was used to identify xenobiotics that may regulate UGT expression; with particular focus on the regulation of UGT1A8 by flavonoids.

UGT1A8 was first isolated in 1998, by Strassburg *et al.*, who noted its high sequence similarity to other exclusively extrahepatic UGTs: UGT1A8 is 93.8% similar to UGT1A7 and 90.2% to UGT1A10. As mentioned in Chapter 1, this homology extends to the proximal promoter regions. However, UGT1A8 shows a highly specific pattern of expression mainly at the distal end of the gastrointestinal tract (distal small intestine and the colon), and unlike UGT1A7 and UGT1A10, UGT1A8 is absent from the most proximal region of GI tract (stomach) (Strassburg *et al.*, 1998). The high level of expression and activity of UGT1A8 in colon may play an important role in protecting luminal cells from prolonged exposure to xenobiotics in this context. The colon is also known to host the densest population of microbiota, with bacteria present as early as 3 days post-natally (Yatsunenکو *et al.*, 2012). Microbes produce small molecules that may be substrates for UGTs, or that may modulate UGT expression.

One of the significant functions of colon microbiota is their role in saccharolytic fermentation of food substrates containing fibre, which include fruit, grains, cereals, legumes and vegetables.

Fermentation produces favourable end products in the form of short chain fatty acids (SCFA), including acetate, propionate and butyrate (Gibson, 2004). These products provide essential energy for intestinal cells, as well as potentially modulating their functions by acting as signalling molecules. Butyrate is of particular interest to this study, due to its ability to inhibit histone deacetylases leading to increased histone acetylation and chromatin accessibility and thus increased gene expression (Candido *et al.*, 1978). Butyrate has long been recognized as possessing anti-cancer activity. *In vitro* studies showed that butyrate inhibited the growth of colon cancer cell lines (Kim *et al.*, 1980, Whitehead *et al.*, 1986), and *in vivo* studies indicate that butyrate suppresses colon cancer formation (McIntyre *et al.*, 1993). In addition, studies in colon adenocarcinoma cells, SW480 and SW620, revealed that butyrate promotes expression of phenotypic markers of differentiation (Kim *et al.*, 1980). This property inducing differentiation has also been utilised experimentally to induce colon cancer cells, especially in Caco-2 to exhibit are more differentiated enterocyte-like phenotype (Mariadason *et al.*, 2001). Yamashita *et al.* developed a rapid differentiation protocol for the Caco-2 cell line using butyrate, which produced cultures exhibiting comparable differentiation characteristics to those generated using the previously developed long-term (21 day) post-confluent culture method (Yamashita *et al.*, 2002).

The effect of butyrate on the induction of the metabolism enzymes glutathione S-transferases (GST) has been reported (Pool-Zobel *et al.*, 2005); however only a few studies have linked the activity of gut microbiota with xenobiotic metabolism (Ruan *et al.*, 2015, Takahiro *et al.*, 2009). A very recent study (Van Rymenant *et al.*, 2017) reported that chronic exposure of Caco-2 cells to short-chain fatty acids and microbiota-derived phenolic compounds, hesperetin (HT) and ferulic acid (FA), increased the expression of transporters (MCT-1, MCT-4 and ABCG2). In this study changes in UGT1A protein levels were not observed; however use of a pan-UGT1A antibody that detects all UGT1A isoforms could masked changes in specific UGT1A proteins.

The studies described in this Chapter sought to better understand the role of flavonoids and the microbiota-produced signalling molecule butyrate in regulation of UGT1A8 expression. We hypothesized that certain flavonoids could modulate transcriptional regulation of UGT1A8 via ligand-activated transcription factors, and that butyrate could enhance the ability of these factors to induce UGT1A8 by modulating chromatin accessibility. Such regulatory mechanisms would provide a link between the anti-cancer effects of flavonoids and butyrate and the detoxification/elimination of cancer-promoting chemicals by glucuronidation.

Using the Caco-2 cell line as a model, the aims of this study are:

1. To screen the *UGT1A8* promoter for induction by a range of chemicals (flavonoids and steroids) in a differentiated Caco-2 cell model and identify the most effective inducer(s).
2. To determine the effect of butyrate on UGT1A8 promoter activity, and its contribution to UGT1A8 promoter upregulation by inducer(s) identified in Aim 1.
3. To identify specific ligand-activated transcription factor(s) that mediate the upregulation of UGT1A8 by inducer(s) identified in Aim 1.

## **4.2 Methods**

### **4.2.1 Caco-2 cells and culture**

Two stocks of Caco-2 cells were used in studies described in this Chapter. The first was the parental Caco-2 cell line obtained from ATCC (HTB-37) and maintained at moderately low passage numbers (11-35). The second was a transgenic Caco-2 cell line that had been stably transfected with a *UGT1A8* promoter-reporter construct; this transgenic cell line was unavoidably maintained at higher passage numbers (35-70). The reporter construct contained a 7 kb segment of the UGT1A8 promoter upstream of the TSS linked to a firefly luciferase reporter; the stable cell line (designated -7k-UGT1A8 promoter Caco-2 cells) was developed in our laboratory previously by Dr. Dong Gui Hu. Caco-2 cells were cultured as described in general methods (Section 2.2.1) in DMEM supplemented with 10% FBS and 1x NEAA (termed complete DMEM media).

#### **4.2.1.1 Post-confluent Caco-2 differentiation protocol (21-day protocol)**

Caco-2 cells were initially seeded at  $5 \times 10^3$  cells/well in a 48-well plate in 250  $\mu$ L complete DMEM media and allowed to become confluent. They were then cultured post-confluency for 21 days. The culture medium was replaced every 2 days for the first week of culture and then daily until day 21 of culture (Hughes *et al.*, 1987). After 21 days, the post-confluent differentiated Caco-2 cells were ready for harvest or for other treatments.

#### **4.2.1.2 Sodium butyrate-induced Caco-2 differentiation protocol**

Differentiation protocols using sodium butyrate were adapted from (Yamashita *et al.*, 2002). On day 1, Caco-2 cells were seeded in high-density at  $1 \times 10^5$  cells/well of a 48-well plate in complete DMEM media. The medium was replaced on day 3 with complete media containing sodium butyrate (NaB) at a final concentration of 3 mM and grown for another 48 hours or until day 5 in the NaB supplemented media. Cells were then ready for harvest or for other treatments.

#### **4.2.2 Chemical treatment of Caco-2 cells**

Unless otherwise stated, all flavonoid and steroid chemical compounds used for treatment of Caco-2 cells were prepared at a final concentration of 10  $\mu$ M in DMSO or ethanol solvent. Cells were incubated with complete DMEM media containing the chemical compound for 48 hours prior to assay. The antagonist agents in this study, which were ICI 182780 and GW9662, were incubated with the cells for 1-hour prior to flavonoid treatment. Vehicle controls were exposed to the same final concentration of vehicle as in the treatment condition.

#### **4.2.3 Luciferase assays**

Forty-eight hours post treatment, transgenic -7kb *UGT1A8* promoter Caco-2 cells were assayed for firefly luciferase activity to quantify the activity of the stably integrated *UGT1A8* promoter. Prior to assay, cells were lysed by adding 75  $\mu$ L of 1x passive lysis buffer (PLB) to each well of the 48-well plate; the lysate was prepared and firefly luciferase activity was analysed as described in Chapter 2. Relative luciferase activity was calculated as the ratio of luciferase activity to total protein concentration of the lysate (determined using the Bradford assay). Normalization to total protein was necessary because a reference reporter gene (such as *Renilla* luciferase) was not present in

the cell line. All treatments and assays were performed in duplicate.

#### **4.2.4 Bradford assay for the determination of protein concentration**

The total protein in the cell lysates used for luciferase assay was measured using the Bio-Rad Protein Assay which is a variant of the Bradford dye-binding method (Bradford, 1976). Ten  $\mu\text{L}$  of diluted protein lysate (in PLB) was transferred to a 96-well microplate. Two hundred  $\mu\text{L}$  of 1:5 diluted dye reagent was added to each well and mixed by pipetting, followed by 5 minutes incubation at room temperature. Absorbance was measured at 595 nm wavelength in a plate reader (DTX 880 Multimode Detector; Beckman Coulter). Protein concentration was determined by plotting the absorbance of the sample onto a standard curve generated using of bovine serum albumin (BSA) dilutions of known concentration (also measured at 595 nm). The concentrations of BSA protein standards were 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5 mg/mL.

#### **4.2.5 Plasmids**

##### **4.2.5.1 Luciferase constructs**

All of the wild-type *UGT1A8*, *-1A9* and *-1A10* promoter-reporter constructs in pGL3-Basic were previously made in our laboratory by Dr. Philip Gregory as described in (Gregory *et al.*, 2003). PPRE X3-TK-luc was obtained from Addgene (plasmid #1015, a gift from Bruce Spiegelman (Kim *et al.*, 1998)).

##### **4.2.5.2 Expression vectors**

The PPAR $\gamma$ -pCMX vector was obtained from Dr. Michael Downes at the Scripps Research Institute, La Jolla, California. PGC-1 $\alpha$ -pcDNA4-myc vector was obtained from Addgene (plasmid 10974) (Ichida *et al.*, 2002). GFP-pMM043 was a gift from Dr. Michael Michael (Flinders Centre for Innovation in Cancer, Flinders University, South Australia). VP 16-PPAR $\alpha$  and RXR $\alpha$ -pCMX vectors were made previously in our laboratory by Dr. Dione Gardner-Stephen. NF-YA-pSG5 was a gift from Dr. Roberto Mantovani (Department of Biomolecular Sciences and Biotechnology, University of Milan, Italy) (Mantovani *et al.*, 1994). The PPAR $\gamma$ -2A-RXR $\alpha$ -pcDNA3 and PPAR $\alpha$ -2A-RXR $\alpha$ -pcDNA3 vectors were generated as detailed in the subsequent section.

#### **4.2.5.3 Generation of bicistronic expression constructs for PPAR $\gamma$ -2A-RXR $\alpha$ and PPAR $\alpha$ -2A-RXR $\alpha$ in the pcDNA3 expression vector backbone**

A PPAR $\gamma$ /RXR $\alpha$  and PPAR $\alpha$ /RXR $\alpha$  expression cassette were cloned into pcDNA3 vector containing a 2A peptide sequence. This allows two genes to be expressed from the same promoter in the same vector (the CMV promoter in pcDNA3). The resulting expression cassette structure was PPAR $\gamma$ -2A-RXR $\alpha$  and PPAR $\alpha$ -2A-RXR $\alpha$ .; the 2A peptide mediates bicistronic expression of two genes by inducing ribosomal peptide-bond skipping (self-cleavage) during translation. The result is high expression of two different protein products from a single transcription event with approximately 1:1 ratio (reviewed in (Hutson *et al.*, 2014)).

The 2A-pcDNA3 vector used in this study originally carried mCherry and Barx2 cDNA sequences (mCherry/2A/Barx2-pcDNA3) and was made by Dr. Julie-Ann Hulin. Using restriction sites flanking these genes, the mCherry and mBarx2 inserts were removed from the vector leaving the 2A sequence intact. Human RXR $\alpha$  cDNA was amplified by Phusion DNA Polymerase (Thermo Fisher Scientific) using RXR $\alpha$  F EcoRV and RXR $\alpha$  R XbaI primers (see Table 4.1) and a human RXR $\alpha$ -pCMX plasmid as template. The purified PCR product was digested with EcoRV and XbaI and ligated directionally into the EcoRV and XbaI sites of the empty 2A-pcDNA3 vector, thus creating 2A-RXR $\alpha$ -pcDNA3 plasmid.

Subsequently, PPAR $\gamma$  was amplified by Phusion DNA Polymerase from mouse kidney cDNA, using PPAR $\gamma$  F KpnI and PPAR $\gamma$  R MfeI primers (see Table 4.1) (MfeI is compatible for ligation to an *EcoRI* site). The purified PCR product was digested with KpnI and MfeI and ligated into the *KpnI* and *EcoRI* sites of 2A-RXR $\alpha$ -pcDNA3, generating the PPAR $\gamma$ -2A-RXR $\alpha$ -pcDNA3 vector.

Mouse PPAR $\alpha$  was amplified from mPPAR $\alpha$ -pCMX plasmid, using PPAR $\alpha$  F BamHI and PPAR $\alpha$  R MfeI primers (see Table 4.1). The purified PCR product was digested with BamHI and MfeI and then ligated into the *BamHI* and *EcoRI* sites of 2A-RXR $\alpha$ -pcDNA3, creating the PPAR $\alpha$ -2A-RXR $\alpha$ -pcDNA3 vector. All primers used for this cloning are listed in Table 4.1.

#### **4.2.5.4 Generation of the -1kb +143bp UGT1A8 promoter/exon 1 construct and mutant variants**

A 143 bp fragment of *UGT1A8* Exon-1 containing a direct repeat (DR) corresponding to a putative PPAR binding site was cloned at the 3' side of the -1kb-*UGT1A8*-promoter fragment in the existing-1kb *UGT1A8* pGL3 construct. Both wild-type and mutant 143 bp exon1 fragments were amplified by PCR (Phusion DNA polymerase; Thermo Fisher Scientific) using human genomic DNA (Sigma) as the template and specifically designed cloning primers. The primers for the wild-type fragment were 1A8 Ex1-143 wt forward and reverse primers, whereas primers for mutant fragment were 1A8 Ex1-143 mut forward and reverse primers (Table 4.1). PCR products corresponding to *UGT1A8*-Exon1 wild-type and mutant sequences were digested with PstI and NcoI restriction enzymes and ligated into the PstI and NcoI sites of the 1k-*UGT1A8*-pGL3 vector.

#### **4.2.6 Mutagenesis of PPAR response element (PPAR-RE) in the UGT1A8 promoter**

Mutation of the predicted PPAR-RE in the -1kb-*UGT1A8* pGL3 vector or in the -1kb *UGT1A9* pGL3 vector was performed using the site-directed mutagenesis (SDM) method described in Chapter 2. The PPAR-RE at location -818 to -805 was mutated from "TCACCACTGACCT" to "TCATTA CTGATTT". The corresponding site in 1k-*UGT1A9* at location -719 to -709 (Barbier *et al.*, 2003) was mutated from "TCACCTCTGACCT" to "TCATTTCTGATTT". In addition, based on a report by Barbier *et al.*, a PPAR 'half-site' sequence predicted at location -50 to -45 in the proximal promoter of *UGT1A9* was mutated from "GGGTCA" to "GAATCA" in the 1kb *UGT1A9* pGL3 vector. All primers used for SDM protocol are listed in Table 4.1.

#### **4.2.7 Fluorescence activated cell sorting (FACS) of Vp 16-PPAR $\alpha$ / GFP co-transfection**

Transfection of VP 16-PPAR $\alpha$  vector in Caco-2 cells was performed using GFP as a co-transfectant to enable FACSorting of transfected cells based on GFP expression level. VP 16-PPAR $\alpha$ / GFP expressing cells were FACSorted on a BD FACSAria (BD Biosciences, San Jose, CA) and directly collected in a 1.5 mL eppendorf tube containing 1 mL TRIzol® Reagent for RNA extraction.

#### 4.2.8 Cell Electroporation

Caco-2 cells were harvested by trypsinisation and counted. An aliquot of  $1.5 \times 10^6$  cells was transferred to a 1.5 mL Eppendorf tube and spun at 1000 rpm for 5 minutes to collect the cell pellet that was then resuspended in 400  $\mu$ L isoosmolar electroporation buffer (IEB), containing  $\text{KH}_2\text{PO}_4$  0.3 mM, KCl 25 mM,  $\text{K}_2\text{HPO}_4$  0.85 mM and myo-inositol 280 mOsmol/kg, (recipe adapted from (Deora *et al.*, 2007)). The cell suspension was transferred to Bio-Rad Electroporation Cuvette with a 0.4 cm gap, and 3  $\mu$ g plasmid DNA was added to the cells and mixed by pipetting prior to electroporation on a Bio-Rad MicroPulser™ Electroporator. The settings for electroporation were capacitance at 675  $\mu$ F and voltage at 250 mV and were optimized by electroporation of the GFP expression plasmid and imaging using the EVOS® FL fluorescence microscope. Electroporated cells were placed into a well of a 6-well plate and 1.5ml of complete DMEM media was added. Twenty-four hours post electroporation, the media was changed to remove dead cells. The cells were grown for another 48 hours before harvested in TRIzol® reagent for RNA extraction.

#### 4.2.9 Lentiviral packaging and transduction

Lentiviral transduction of Caco-2 cells was tested as a method to obtain high numbers of cells carrying desired expression vectors. The GFP/pTiger FIV-based lentiviral plasmid was used to test packaging and transduction efficiency.

To package the lentivirus,  $2 \times 10^5$  HEK293T cells in T25 flask (70-90% confluent) were transfected with 3 $\mu$ g of the GFP/pTiger lentiviral vector, 2  $\mu$ g of FIV Gag-Pol vector and 1.5  $\mu$ g of vesicular-stomatitis virus (VSV)-G vector. The plasmid DNA was mixed in 200  $\mu$ L serum free DMEM and incubated at room temperature for 5 minutes before adding 16  $\mu$ L of Lipofectamine® 2000 also diluted in 200  $\mu$ L serum free DMEM; after a further 20 minutes incubation at room temperature the transfection complexes were added to the HEK293T cells. The media was replaced 6 hours after transfection and the supernatant containing lentiviral particles was collected after 48 hours and either used immediately or snap frozen in dry ice and stored at  $-80^\circ\text{C}$ . After two rounds of virus collection, the HEK293T virus-producing cells were discarded appropriately.

To transduce Caco-2 cells with lentivirus, Caco-2 cells were seeded the day before transduction in a 6-well culture plate at a density of  $2 \times 10^5$  cells/well. Lentiviral supernatant was prepared by centrifuging at 2000 rpm for 5 minutes to pellet any remaining packaging cells/debris. The supernatant was transferred to a new tube and polybrene was added to a final concentration of 4  $\mu\text{g}/\text{mL}$ . The media from the Caco-2 plate was replaced with viral supernatant, followed by centrifugation of the plate at 3500 rpm for 2 hours at 32°C (spinfection). Following the spinfection, viral media was replaced with DMEM complete media (10% FBS). To check transfection efficiency, Caco-2 cells were observed under an EVOS®FL microscope with a GFP filter at 48 hours post-transduction.

#### 4.2.10 Oligonucleotides

**Table 4.1. Oligonucleotides used in studies described in Chapter 4.**

Name	Sequence (5' → 3')
Cloning	
1A8 Ex1-143 wt F	GCCCTGCAGTTCTCTCATGGCTCGC
1A8 Ex1-143 wt R	GCCCCATGGGCACTACCAGCAGCT
1A8 Ex1-143 mut F	GCCCTGCAGTTCTCTCATGGCTCGC
1A8 Ex1-143 mut R	GCCCCATGGGCACTACCAGCAGCTTCCCTGCCTCAGCAA AGTTTCAGGTTAGCAGC
1A8 mut PPRE full F	CATCATTACTGATTTTCAGGGAGTGCCCAG
1A8 mut PPRE full R	CCCTGAAATCAGTAATGATGTGCATCTTTGTGT
1A9 mut PPRE full F	GACATCATTTCTGATTTACGGAGTGCTCAGCAGACTG
1A9 mut PPRE full R	CACTCCGTGAAATCAGAAATGATGTCAACTTTGTG
1A9 mut PPRE half F	ATTGGAATCAGGTTTTGTGCCTGTAGTTC
1A9 mut PPRE half R	CCTGATTCCAATAGAGGGCGTGTATTTATCCTG
PPAR $\alpha$ F BamHI	CGGGATCCATGGTGGACACAGAG
PPAR $\alpha$ R EcoRI	CGCAATTGGTACATGTCTCTGTAGATCTC
PPAR $\gamma$ F KpnI	GGGGTACCATGGGTGAACTCTGGGAG
PPAR $\gamma$ R MfeI	CCGCAATTGATACAAGTCCTTGTAGATCTCC
RXR $\alpha$ F EcoRV	CGGATATCATGGACACCAAACATTTCTGCC
RXR $\alpha$ R XbaI	GCTCTAGACTAAGTCATTTGGTGCG
qRT-PCR	
18S F	CGATGCTCTTAGCTGAGTGT
18S R	GGTCCAAGAATTTACCTCT

UGT1A8 F	CTGCTGACCTGTGGCTTTGCT
UGT1A8 R	CCATTGAGCATCGGCGAAAT
UGT1A9 F	GAGGAACATTTATTATGCCACCG
UGT1A9 R	GCAACAACCAAATTGATGTGT
UGT1A10 F	CCTCTTTCCTATGTCCCAATGA
UGT1A10 R	GCAACAACCAAATTGATGTGTG
hH-Ferritin F (hFTH1)	CTGGCTTGGCGGAATATCT
hH-Ferritin R (hFTH1)	CCCGAGGCTTAGCTTTCATT
hABCB1 F	GCCATCAGTCCTGTTCTTGG
hABCB1 R	GCTTTTGCATACGCTAAGAGTTC
hABCG2 F	TGCAACAGGAAACAATCCTTGT
hABCG2 R	GATCGATGCCCTGCTTTACC
hCDX2 F	ATCACCATCCGGAGGAAAAG
hCDX2 R	TGCGGTTCTGAAACCAGATT
hHNF4 $\alpha$ F	CAGCACTCGAAGGTCAAGCTA
hHNF4 $\alpha$ R	ACGGGGGAGGTGATCTGT
hPPAR $\gamma$ F	GCTTCATGACAAGGGAGTTTC
hPPAR $\gamma$ R	ACTCAAACCTGGGCTCCATAAAG
hER $\beta$ F	TGGTCGTGTGAAGGATGTAAG
hER $\beta$ R	ACTTCTCTGTCTCCGCACAAG
NF-YA F	CAATTCAGGAGGGATGGTCA
NF-YA R	GAGAGGCTCTTCTTCAAGCATC
L-FABP F	CACCTTCCAACCTGAACCACTG
L-FABP R	TGATCCAAAACGAATTCACG
ChIP-qPCR	
ChIP negative control locus For	ACATACTCAGATGGAAATGAGAA
ChIP negative control locus Rev	AGCTCAACATTCTGCTGAAC
UGT1A8-100 For	TTTTGGTACCTCAAAAAATGATACTC
UGT1A8-100 Rev	AGCCACGCGTGAACCTGCAGCCCGAGC
EMSA	
LUEGO-Tag	GTGCCCTGGTCTGG
UGT1A8-50 DR1 half-site wt Top	GTGCCCTGGTCTGGCGCCCTCTATTGGGGTCAGGTTTT
UGT1A8-50 DR1 half-site mut Top	GTGCCCTGGTCTGGCGCCCTCTATTGAAGTCAGGTTTT

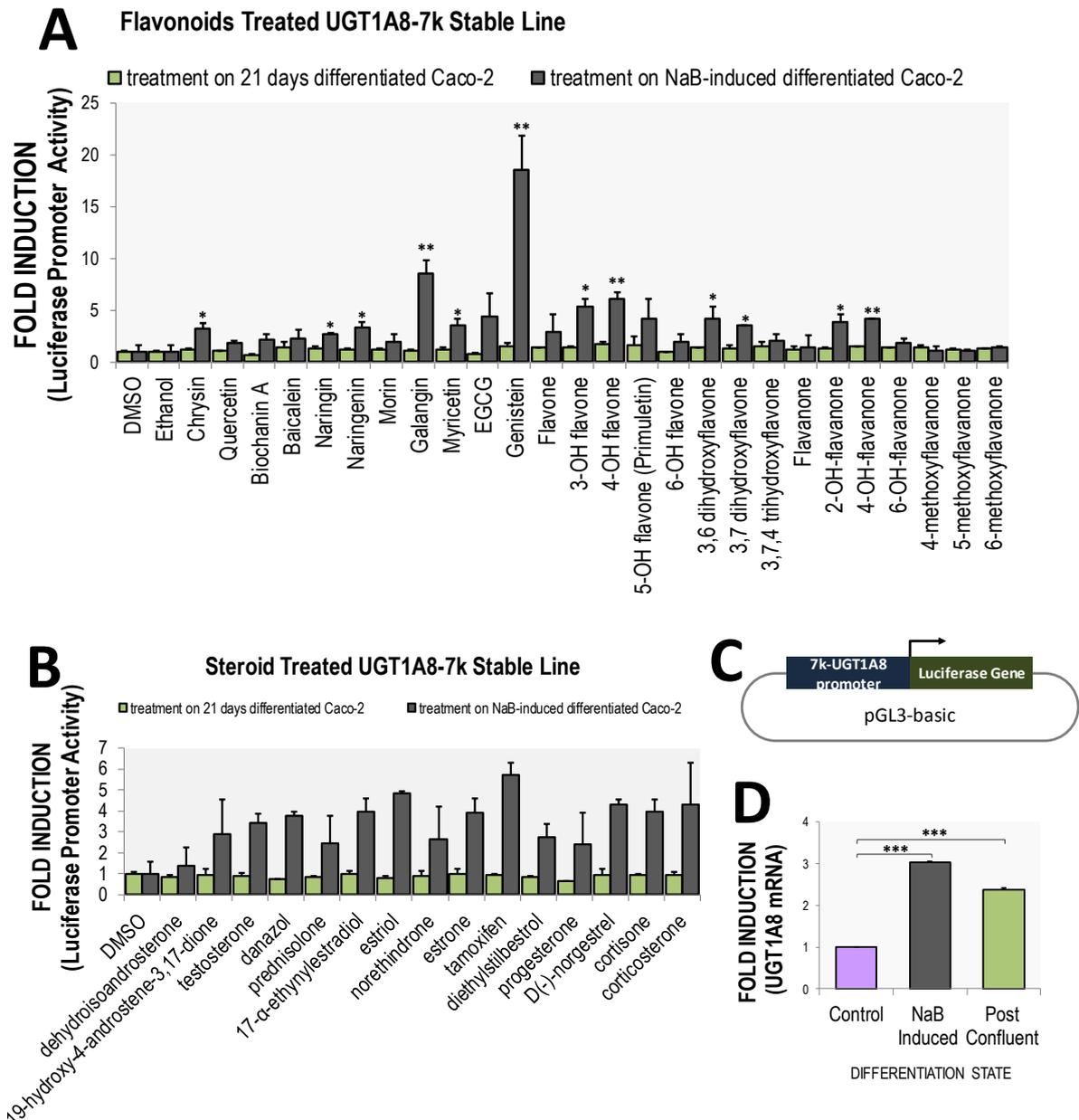
## 4.3 Results and Discussion

### 4.3.1 Screening of *UGT1A8* promoter response to flavonoids identifies genistein as a potent inducer of promoter activity in NaB-induced differentiated Caco-2 cells

To identify chemicals (flavonoids and steroids) that can induce *UGT1A8* promoter activation we used a stably transfected Caco-2 cell line carrying the -7kb *UGT1A8* promoter linked to the luciferase reporter (Figure 4.1C). The -7kb-*UGT1A8* stable line was differentiated to represent mature intestinal enterocyte by two differentiation methods: the 21 day-post confluent differentiation (21d-diff) and sodium butyrate-induced rapid differentiation (NaB-diff) methods. Consistent with the pattern of UGT expression in intestinal epithelium, which increases in differentiated cells at the villi surface by 4 fold compared to crypt cells (Dubey and Singh, 1988), and a report showing that genes involved in xenobiotic and drug metabolism increase during Caco-2 differentiation (Mariadason *et al.*, 2002), we found that both differentiation methods increased the basal level of *UGT1A8* mRNA (Figure 4.1D), with the NaB diff method producing slightly greater increase (3-fold) than the 21d-diff method (2.4-fold).

A total of 26 flavonoid compounds were tested at a concentration of 10  $\mu$ M for their ability to induce the *UGT1A8* promoter in differentiated Caco-2 cells. The flavonoid concentration (10  $\mu$ M) was selected as it was considered at low range and non-toxic for Caco-2 cells, this was supported by recent findings of Fang *et al* which showed that the viability of Caco-2 cells was unaffected at flavonoid concentrations of up to 40  $\mu$ M (Fang *et al.*, 2017). Thus, we could eliminate cell viability factors in interpreting the effects of flavonoids on *UGT1A8* promoter activity. Activity assays were performed 48 hours post-treatment. Figure 4.1A shows that flavonoid exposure could stimulate -7kb-*UGT1A8* promoter activity only in cells differentiated by NaB induction (NaB-diff), whereas the 21d-diff Caco-2 cells were insensitive to the treatment. This insensitivity of the 21d-diff cells might have been caused in part by impermeant tight junctions that develop during the differentiation process. The permeability of tight junctions formed after NaB-induced differentiation of Caco-2 cells appears to be more variable when compared to the 21-day differentiation method (Yamashita *et al.*, 2002). Gomez *et al.* found that the tight junctions of differentiated Caco-2 cells are impermeable to mannitol (Gómez *et al.*, 1999), although it is unclear if they would be impermeable

to flavonoids. It is also possible that NaB more effectively induces regulatory factors that mediate promoter induction by flavonoids; and/or makes the *UGT1A8* promoter construct more accessible to these regulatory factors via chromatin modification.



**Figure 4.1. Screening for activation of *UGT1A8* promoter by flavonoid and steroid in differentiated Caco-2 cells.**

A. Flavonoids at 10  $\mu$ M were screened for their ability to induce the -7kb-*UGT1A8* promoter in Caco-2 cells differentiated according to the 21-day post confluent protocol or the NaB-induced protocol. Genistein (4',5,7-trihydroxyisoflavone) generated the highest activation in NaB-induced differentiation conditions. B. Steroid were screened for their ability to induce the -7kb-*UGT1A8* promoter in Caco-2 cells under both differentiation

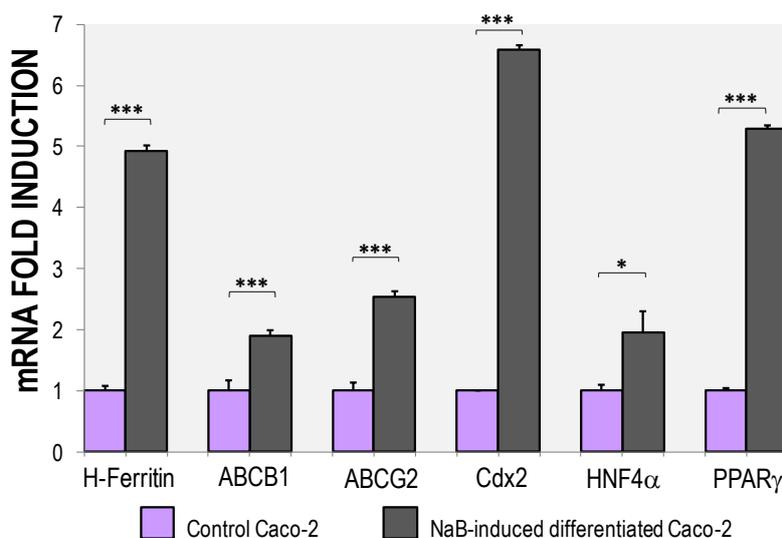
conditions. C. Schematic construct of the -7kb segment of the *UGT1A8* promoter cloned into the Luciferase reporter pGL3basic. D. Confirmation that *UGT1A8* mRNA level is increased in Caco-2 cells by both differentiation protocols. Relative luciferase activities are shown as fold induction over vehicle and normalised to the amount of total protein. The mRNA expression was analysed by real-time reverse-transcriptase PCR and normalised to 18S rRNA. Data are average of duplicates from 3 experiments. T-test \* $P < 0.05$ ; \*\* $P < 0.005$ , \*\*\* $P < 0.001$ .

In NaB induced differentiated Caco-2 cells, our results indicated that exposure to 18 out of 26 flavonoids increased *UGT1A8* activation significantly in the following order: Flavanone < Morin < 6-OH flavone < 3,7,4 trihydroxyflavone < Naringin < Flavone < Naringenin < 3,7 dihydroxyflavone < Myricetin < 2-OH flavanone < 4-OH flavanone < 3,6 dihydroxyflavone < 5-OH flavone < epigallocatechin gallate (EGCG) < 3-OH flavone < 4-OH flavone < galangin < genistein (Figure 4.1A). Genistein was the most potent inducer of the *UGT1A8* promoter inducing it ~18 fold over the vehicle condition. Interestingly, Tang *et al.* reported that at a concentration of 10  $\mu\text{M}$  or greater, the largest contributor to genistein glucuronidation is *UGT1A8*, while at a low concentration, genistein was glucuronidated more by *UGT1A9* (Tang *et al.*, 2009). Thus, the induction of *UGT1A8* promoter by genistein suggests a possible regulatory feedback circuit.

*UGT1A8* is known to also metabolise steroid compounds (Cheng *et al.*, 1999), thus we decided to test whether steroids also mediate promoter activation. Again, only in cells differentiated using the NaB-diff protocol could stimulation of the *UGT1A8* promoter be observed (Figure 4.1B). Tamoxifen (TAM) produced the highest induction of the *UGT1A8* promoter, with an activation of ~6 fold over vehicle condition, followed by estriol (5 fold) (Figure 4.1B). Sun *et al.* reported that, together with *UGT2B7* and *UGT1A10*, *UGT1A8* is active in TAM glucuronidation (Sun *et al.*, 2007). These data suggest that the *UGT1A8* promoter is regulated by estrogenic compounds that function through the estrogen receptor (ER) signalling pathway.

As discussed above and presented in Figure 4.1A and Figure 4.1B, modulation of *UGT1A8* promoter activity was observed only when chemical treatment was applied to Caco-2 cells differentiated by NaB induction. We propose that flavonoid and steroid molecules regulate the promoter via ligand-activated transcription factors. During the differentiation process, NaB alters expression of differentiation-associated genes by inducing histone hyperacetylation (Mariadason *et*

*al.*, 2001), in a manner that also depends on cell density (Davie, 2003). We demonstrated that several differentiation markers were activated in butyrate-induced differentiation, confirming the state of differentiated Caco-2 (Figure 4.2). H (heavy)-ferritin, a gene involved in iron cellular traffic and intracellular metabolism, has been associated with phenotypic differentiation in Caco-2 cells (Bevilacqua *et al.*, 1995). We found that H-ferritin mRNA was increased ~5-fold by NaB-induced differentiation, a similar increase to that shown by Bevilacqua *et al.* Drug efflux transporters ABCB1 and ABCG2 also showed small increases of around 2 fold.



**Figure 4.2. Differentiation markers increase following sodium butyrate-induced differentiation.**

mRNA levels of differentiation markers were analysed by qRT-PCR in NaB-induced differentiated Caco-2 and non-differentiated Caco-2 cells (set as 1). The results are normalised to 18S rRNA. Data is in duplicate from 2 experiments. T-test \* $P < 0.05$ ; \*\* $P < 0.005$ ; \*\*\* $P < 0.001$ .

The expression of transcription factors known to be involved in UGT regulation, CDX2, HNF4α and peroxisome proliferator-activated receptor gamma (PPARγ) were also analysed in NaB-induced differentiated Caco-2 cells and undifferentiated cells (Figure 4.2). We found that CDX2 mRNA was significantly upregulated by around 6 fold in differentiated Caco-2 cells. CDX2 overexpression is known to promote upregulation of intestinal differentiation markers (Suh and Traber, 1996). The expression of HNF4α was also shown to increase by around 2 fold following NaB induced differentiation; this might be related to its regulation by CDX2 (see Chapter 3). PPARγ is a member

of the nuclear receptor family, and has been linked to differentiation of colon cancer cells including Caco-2 cells (Kitamura *et al.*, 1999), inhibiting cell growth and inducing differentiation markers (Wächtershäuser *et al.*, 2000). PPAR $\gamma$ , but not PPAR $\alpha$ , was previously found to be induced during Caco-2 differentiation (Huin *et al.*, 2002). We confirmed here that PPAR $\gamma$  mRNA was induced during NaB-induced differentiation of Caco-2 cells (Figure 4.2).

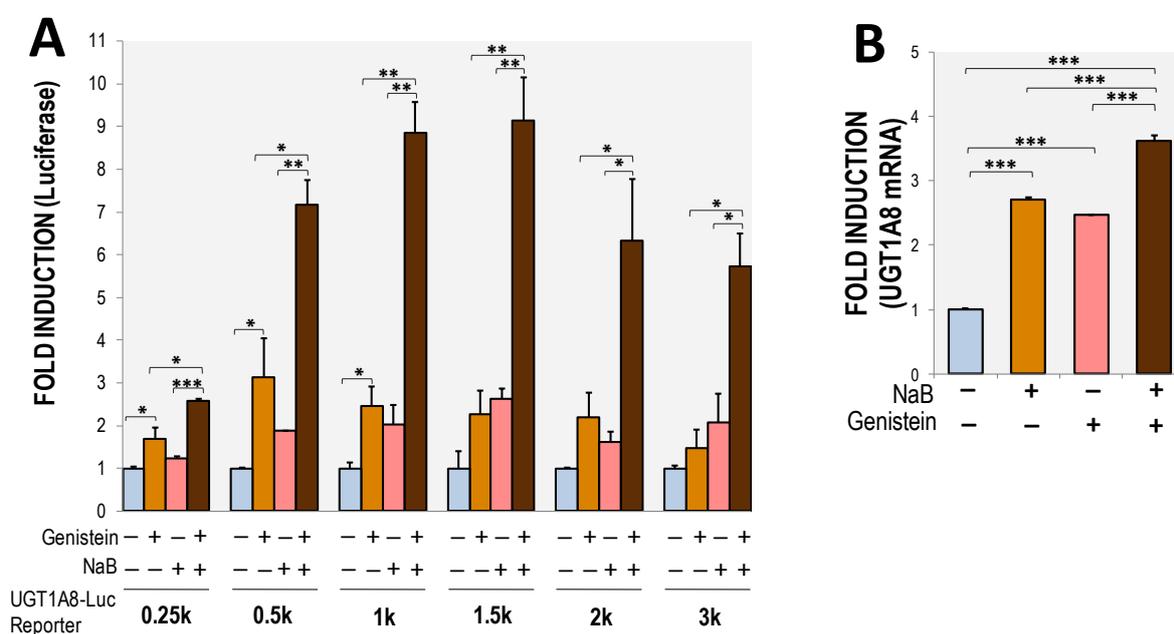
#### **4.3.2 Genistein and butyrate synergistically induce *UGT1A8* promoter activity and *UGT1A8* mRNA expression**

In the NaB-induced differentiated Caco-2 model, genistein was identified as the most effective compound for inducing *UGT1A8* promoter activity amongst the flavonoids and steroids screened. It was possible that NaB treatment enhanced induction by genistein because it induced a differentiated state that includes induction of effectors of genistein action. It is also possible that NaB had more immediate effects on the ability of the *UGT1A8* promoter to be transcriptionally activated. To further examine these possibilities, we transiently transfected *UGT1A8* promoter constructs in cells seeded at low density ( $5 \times 10^4$  cells/well of a 48 well-plate) such that differentiation would not occur.

We also used a number of *UGT1A8* promoter deletion constructs in order to determine the elements within the promoter that contribute to the induction by genistein. The constructs range from -0.25kb to -3kb of the promoter, and were previously made by Dr. Phil Gregory. Figure 4.3A shows how six *UGT1A8* promoter constructs (-0.25kb, -0.5kb, -1kb, -1.5kb, -2kb and -3kb) responded to simultaneous genistein and butyrate exposure. The -0.25kb, -0.5kb and -1kb *UGT1A8* promoters showed small but significant induction by genistein alone (approximately 2 or 3 fold relative to control). Furthermore, synergic induction by genistein and NaB was detected for all promoter constructs. The -1kb and -1.5kb *UGT1A8* promoters showed the greatest induction of around 9 fold. This was less than that previously demonstrated for the integrated -7kb-*UGT1A8* promoter under NaB-induced differentiated conditions, suggesting that density-induced differentiation and distal regulatory elements may contribute to induction. Taken together, Figure 4.3 shows a similar pattern of *UGT1A8* promoter activation by genistein alone or by the synergistic

action of genistein and butyrate, which suggests that important elements responsible for genistein action may be located within -1kb of the *UGT1A8* TSS.

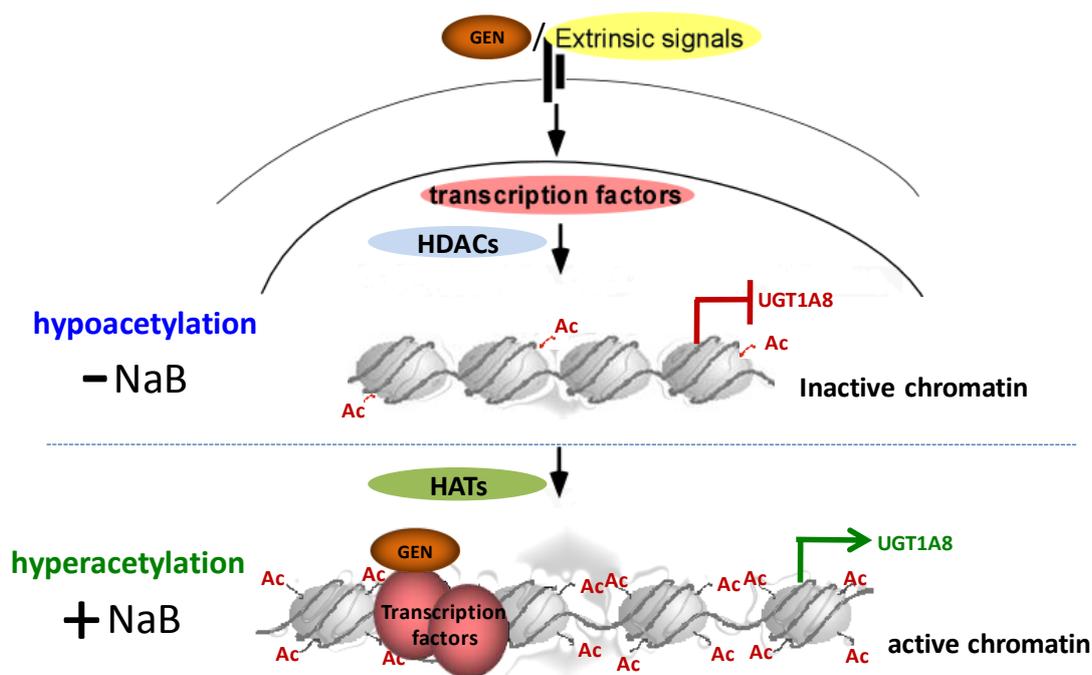
To determine whether *UGT1A8* mRNA is synergistically induced by genistein and NaB, we performed mRNA analysis. Genistein and NaB were each able to increase *UGT1A8* mRNA levels, and a combination of both chemicals generated further induction (Figure 4.3B). Thus, this mRNA data is consistent with promoter analysis in indicating that genistein and NaB regulate *UGT1A8* gene expression.



**Figure 4.3. Identification of genistein as an inducer of *UGT1A8* promoter activity and gene expression.**

A. Genistein activates *UGT1A8* promoter constructs and interacts synergistically with NaB. Fold induction of luciferase activity is presented relative to vehicle treated cells (set as 1). B. *UGT1A8* mRNA expression is increased by genistein, NaB, and by genistein and NaB. The mRNA expression was normalised to 18S rRNA. Data are in duplicate from a minimum of 3 experiments. T-test \* $P < 0.05$ ; \*\* $P < 0.005$ ; \*\*\* $P < 0.001$ .

Based on the evidence that both the *UGT1A8* promoter and mRNA are induced by genistein and NaB, and the capacity of NaB to act as a HDAC inhibitor, we propose that histone acetylation enhances the access of a genistein-activated transcriptional complex to the *UGT1A8* promoter (Figure 4.4). Which transcription factor(s) mediate the activity of genistein remain to be determined and is addressed in the subsequent section.



**Figure 4.4. Schematic model for the synergistic regulation of the *UGT1A8* promoter by NaB and genistein.**

NaB alters chromatin structure by inhibition of histone deacetylases (HDACs). In the absence of NaB, HDACs are active and mediate removal of acetyl groups from histone lysine residues, which results in compacted chromatin. Inhibition of HDACs by NaB allows histone acetyltransferases (HATs) to hyperacetylate histones and decompact chromatin (Delage and Dashwood, 2008). This may facilitate access of a-genistein-responsive transcription factor complex to the *UGT1A8* promoter to induce gene transcription.

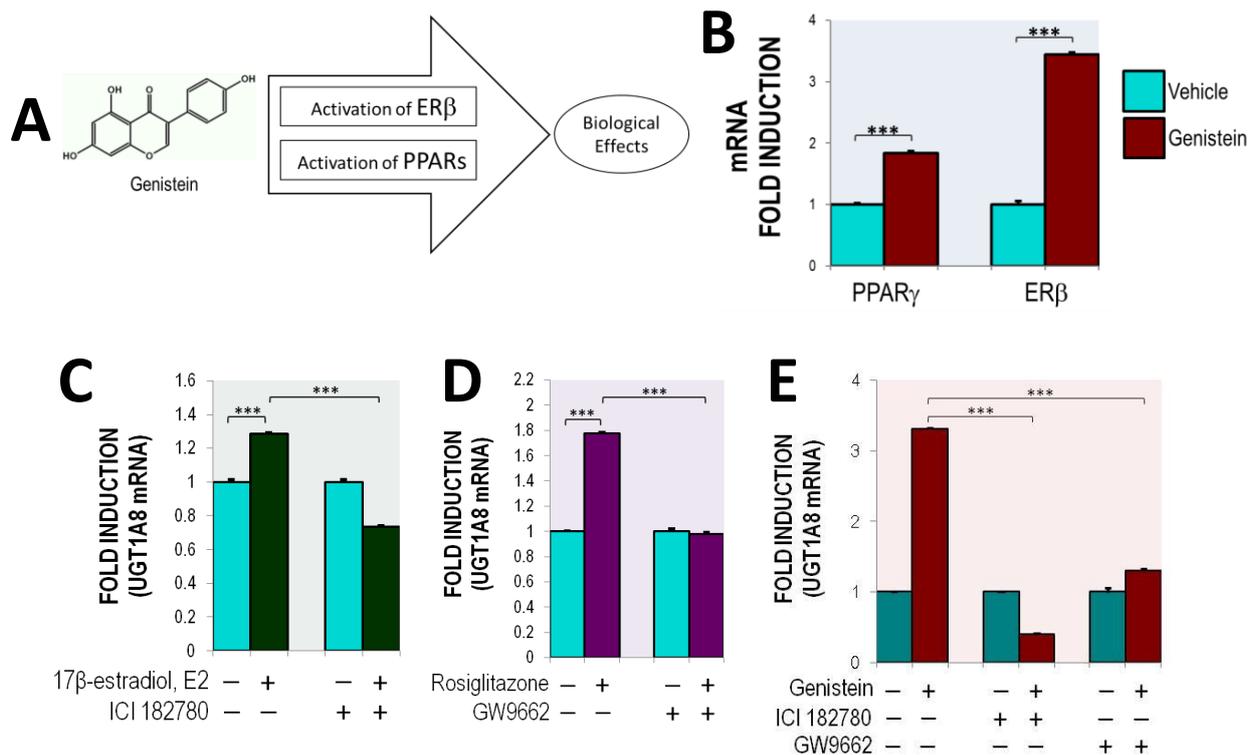
### 4.3.3 Mechanistic analysis of *UGT1A8* regulation by genistein

As the main isoflavone source in the human diet, soy foods, such as tofu and soy milk, are rich in the isoflavone genistein (Murphy *et al.*, 1999). Genistein has significant physiological effects including as a phyto-estrogen. Having a close structural similarity to 17 $\beta$ -estradiol, genistein is able to bind to both estrogen receptor (ER)- $\alpha$  and ER $\beta$  to modulate estrogen-responsive gene expression (Casanova *et al.*, 1999, Nikov *et al.*, 2000). The binding affinity of genistein for ER $\alpha$  is poor compared to estradiol (E2) at approximately 4% that of E2, although genistein has been reported to show ER $\alpha$  selective activity in some contexts (Barkhem *et al.*, 1998). In contrast, the binding affinity of genistein for ER $\beta$  is relatively high (~87% that of E2) (Kuiper *et al.*, 1998), suggesting that this receptor is the main mediator of genistein's estrogenic effects.

Although genistein is considered a phyto-estrogen, it can also bind other NRs such as PPARs (Ricketts *et al.*, 2005) and the role of genistein as a PPAR agonist has been summarized in a review by Patel and Barnes ((Patel and Barnes, 2010) and references therein). A survey of literature reveals many studies supporting the idea that isoflavone action is predominantly mediated by PPAR $\gamma$  (Chacko *et al.*, 2007, Dang, 2009, Miyake *et al.*, 2009, Chacko *et al.*, 2005, Dang and Lowik, 2005, Pallauf *et al.*, 2017). Moreover, genistein was shown to bind the PPAR $\gamma$  ligand binding domain (LBD) leading to a transcriptionally favourable conformation (Salam *et al.*, 2008). However, there is also evidence that soy-isoflavone action can be mediated by PPAR $\alpha$  (Ricketts *et al.*, 2005), although there is a lack of information regarding binding mechanism to PPAR $\alpha$ .

Studies in this section examined whether ER $\beta$ , PPAR $\gamma$  or PPAR $\alpha$  may be involved in regulation of UGT1A8 by genistein in Caco-2 cells (see Figure 4.5A). As presented in Figure 4.5B, ER $\beta$  and PPAR $\gamma$  mRNA levels were upregulated by genistein. To investigate whether these factors might mediate the response of the *UGT1A8* promoter to genistein, we used the chemical inhibitors ICI 182780, an ER antagonist, and GW9662, a PPAR $\gamma$  potent non-competitive antagonist. In addition, agonists were used to further confirm selectivity. 17 $\beta$ -estradiol was used as an ER agonist; although 17 $\beta$ -estradiol binds both ER $\alpha$  and ER $\beta$ , this agent has been shown to be competitive to genistein (Barkhem *et al.*, 1998). Rosiglitazone was used as a PPAR $\gamma$ -selective agonist.

Treatment of cells with either ICI 182780 or GW9662 at 10  $\mu$ M abolished the induction of UGT1A8 mRNA by genistein (Figure 4.5E). In addition, both 17 $\beta$ -estradiol and rosiglitazone were able to induce UGT1A8 expression and this effect was negated by their antagonists (Figure 4.5C and D). These data suggest that may act via both ER and PPAR $\gamma$  to regulate UGT1A8 in Caco-2.



**Figure 4.5. Potential mechanism of UGT1A8 mRNA induction by genistein via PPAR $\gamma$  and ER $\beta$ .**

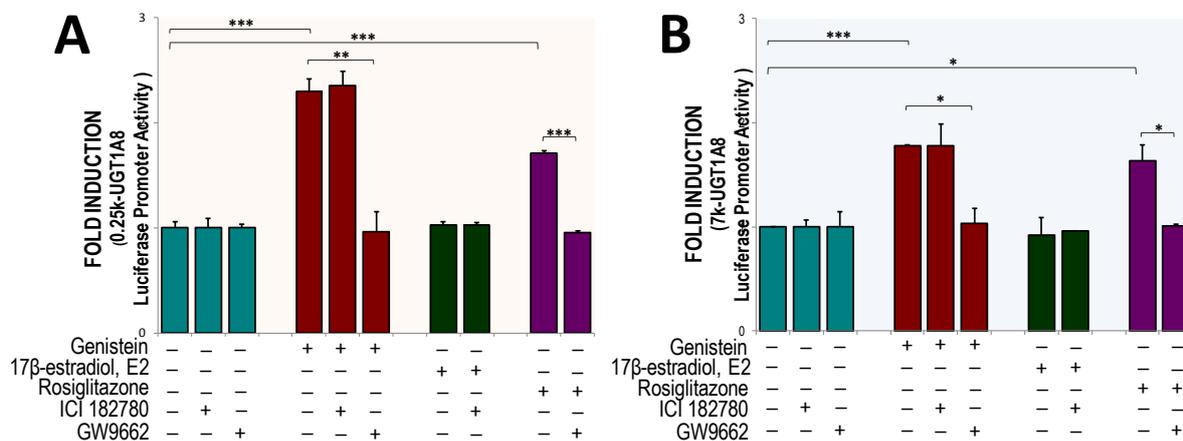
A. Genistein is known to regulate gene expression via both PPAR and ER $\beta$  signalling pathways to target multiple biological processes (Patel and Barnes, 2010). B. Genistein treatment of low density undifferentiated Caco-2 cells induced endogenous PPAR $\gamma$  and ER- $\beta$  mRNA. C and D. UGT1A8 mRNA is increased by an ER agonist (17 $\beta$ -estradiol/E2) and a PPAR $\gamma$  agonist (Rosiglitazone); the ER $\beta$  PPAR $\gamma$  and antagonists/inhibitors, ICI 182780 and GW9662 respectively, negated the effect. E. UGT1A8 mRNA induction by genistein is abolished by both ICI 182780 and GW9662. Genistein, ICI 182780 and GW9662 were used at 10  $\mu$ M, 17 $\beta$ -estradiol at 1 nM, and rosiglitazone at 1  $\mu$ M. Level of UGT1A8 mRNA in vehicle treated Caco-2 cells was set as control at a value of 1. The mRNA expression was normalised to 18S rRNA. Data are in duplicate from two experiments. T-test\*\*\*P<0.001.

In order to identify response elements of ER and PPAR within the *UGT1A8* promoter that may be responsible for genistein action in Caco-2 cells, two UGT1A8 promoter constructs (-0.25kb and a long -7kb) were first screened for responses to genistein, 17 $\beta$ -estradiol, and rosiglitazone; ICI 182780 (ER antagonist) and GW9662 (PPAR $\gamma$  antagonist) were then used to determine whether they could block induction.

Caco-2 cells were transiently transfected with the *UGT1A8* promoter constructs in low density without NaB to observe *UGT1A8* promoter response to genistein alone. Figure 4.6 shows that both the short and long *UGT1A8* promoters were activated by genistein although the fold induction was lower than when NaB was included, again suggesting that butyrate-significantly contributes to

genistein action. GW9662 was able to block activation of both the short and long promoters by genistein, suggesting that PPAR $\gamma$  is involved in mediating genistein action. Moreover, the promoters were activated by PPAR $\gamma$  agonist rosiglitazone and this was also blocked by GW9662.

Treatment of cells with 17 $\beta$ -estradiol did not activate either the short or long *UGT1A8* promoter construct. In addition, the ER antagonist had no effect on genistein-mediated induction of the promoters. Given that the ER antagonist did inhibit genistein-mediated induction of *UGT1A8* mRNA, it seems likely that ER may be involved in induction by genistein via element(s) located in a more distal region of the promoter. Contribution of ER to a post-transcriptional mechanism that increases *UGT1A8* mRNA levels is another possible explanation for our observations.



**Figure 4.6. Mechanistic analysis of *UGT1A8* regulation by genistein at the promoter level.**

A and B. Short -0.25kb- and long -7kb-*UGT1A8* promoter luciferase reporter constructs were transiently transfected in Caco-2 cells. pRL-null *Renilla* reporter vector was used to normalise for transfection efficiency. Cells were treated with 10  $\mu$ M genistein, 1 nM 17 $\beta$ -estradiol, or 10  $\mu$ M rosiglitazone. 10  $\mu$ M ICI 182780 or GW9662 were used to inhibit ER or PPAR $\gamma$  activity respectively. Both promoter constructs showed induction by genistein and PPAR $\gamma$  agonist rosiglitazone but not by 17 $\beta$ -estradiol; only the PPAR $\gamma$  inhibitor GW9662 blocked genistein activation. Data are in duplicate from 3 experiments. T-test \*P<0.05; \*\*P<0.005; \*\*\*P<0.001.

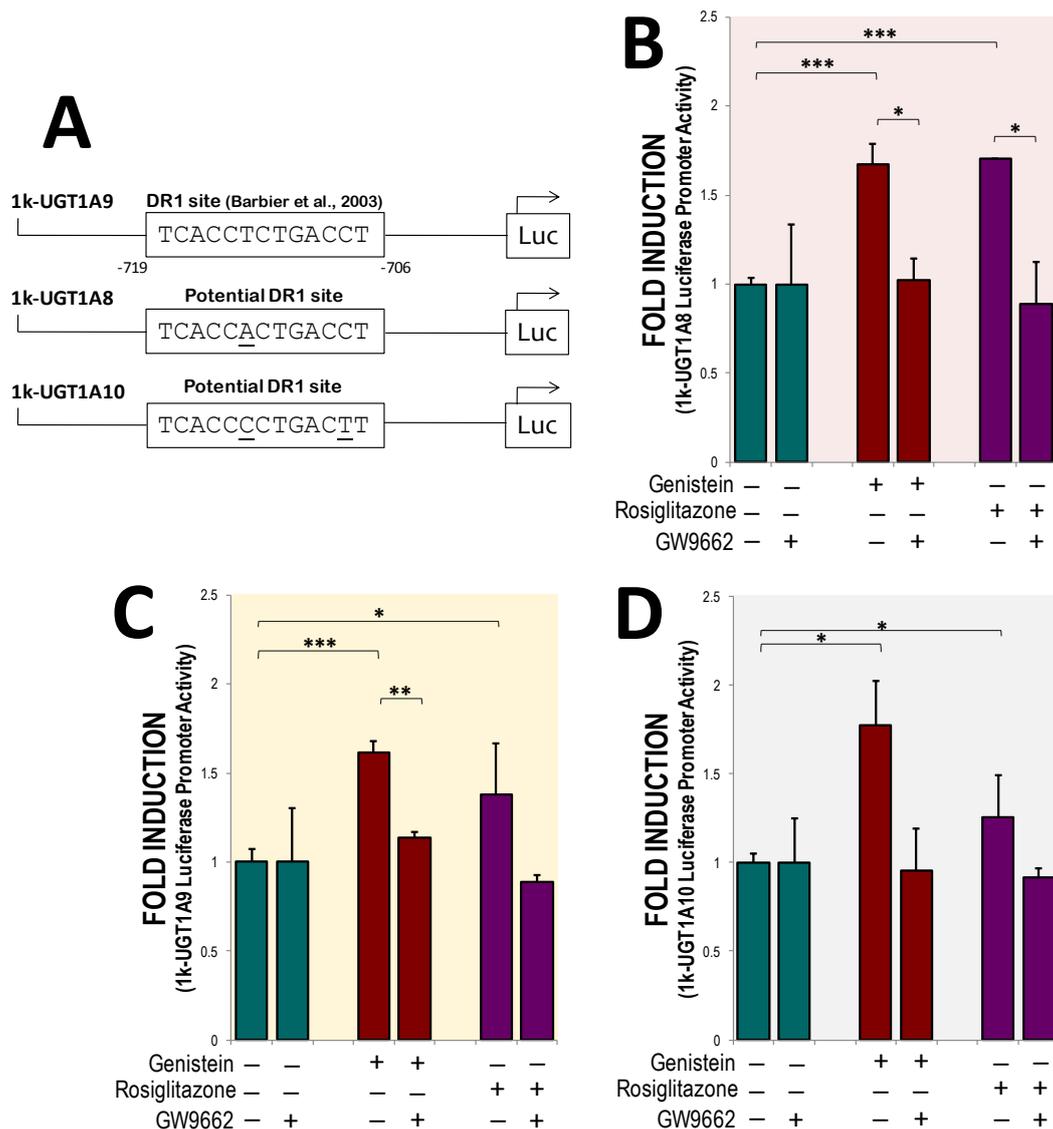
Our *UGT1A8* promoter assay results suggest the presence of a PPAR $\gamma$  response element in the proximal -0.25 promoter region. To date, only one study (Barbier *et al.*, 2003) has identified a functional PPAR binding site in a *UGT1A* gene, and this is located at -719 to -706 of the *UGT1A9* promoter. Since the *UGT1A8-1A10* gene cluster shows a high degree of sequence similarity

(>75%) within the -1k promoter region (Gregory, 2004), we investigated whether genistein can activate UGT1A9 and UGT1A10 in a PPAR $\gamma$ -dependent manner. For this study, -1kb *UGT1A8*, -1A9 and -1A10 promoter constructs were used. The putative PPAR binding site (called direct repeat DR1) in UGT1A9 and its sequence alignment in the three UGT promoters is shown in Figure 4.7.



**Figure 4.7. A putative PPAR binding site within *UGT1A8*, -1A9 and -1A10 1kb promoters.**

Results from the -1kb *UGT1A8*, -1A9 and -1A10 promoter analysis in Caco-2 cells (Figure 4.8) indicate that genistein significantly induced each of the -1kb promoter constructs, and that the PPAR $\gamma$  antagonist GW9662 attenuated induction by genistein. Rosiglitazone treatment also activated the -1kb *UGT1A8* promoter, and its effect was blocked by GW9662. Although not significant, rosiglitazone also showed a trend towards induction of the -1kb-*UGT1A9* and -1A10 promoters. These data suggest that genistein may induce these promoters by acting as a PPAR $\gamma$  ligand.

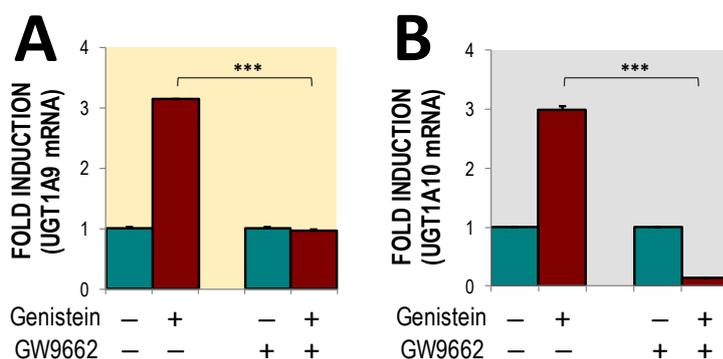


**Figure 4.8. Genistein activates the -1kb *UGT1A8*, -1A9 and -1A10 promoters, likely via PPAR $\gamma$ .**

A. Barbier *et al.* identified an element (DR1) at -719 to -706 in the *UGT1A9* promoter that responds to PPAR. This element is conserved in *UGT1A8* and -1A10 with only 1-2 nt mismatches (underlined). B-D. Genistein treatment, rosiglitazone treatment and PPAR $\gamma$  inhibition (GW9662) were performed in Caco-2 cells transiently transfected with the -1kb-*UGT1A8* (B), -1kb-*UGT1A9* (C) or -1kb-*UGT1A10* (D) promoter reporter constructs. Data are in duplicate from a minimum of 3 experiments. T-test\* $P < 0.05$ ; \*\* $P < 0.005$ ; \*\*\* $P < 0.001$ .

To determine whether the endogenous *UGT1A9* and -1A10 genes are regulated by genistein, we analysed the *UGT1A9* and -1A10 mRNA levels following genistein treatment with or without PPAR $\gamma$  inhibition using GW9662 (Figure 4.9). We found that genistein induced *UGT1A9* and -1A10 mRNA around 3-fold compared to vehicle treatment. Treatment with GW9662 abolished induction of both *UGT1A9* and -1A10 by genistein. These data suggest that the three intestinal UGT genes

*UGT1A8*, *-1A9* and *-1A10* are regulated by genistein, and this is at least partially mediated by PPAR $\gamma$ . *UGT1A9* was previously identified as both a PPAR $\alpha$  and PPAR $\gamma$  target gene (Barbier *et al.*, 2003). *UGT1A9* is also the dominant hepatic UGT involved in glucuronidating genistein (Pritchett *et al.*, 2008). Meanwhile both *UGT1A8* and *UGT1A9* may play a role in glucuronidation of genistein in intestine (Tang *et al.*, 2009). We suggest that genistein induces *UGT1A8* and *UGT1A9* expression through the PPAR $\gamma$  signalling pathway as part of a feedback mechanism that increases its own clearance.



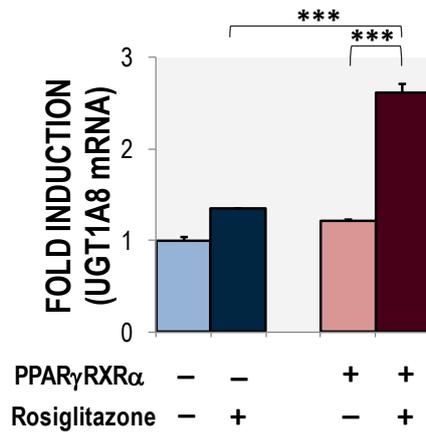
**Figure 4.9. Analysis of *UGT1A9* and *-1A10* mRNA induction by genistein.**

Genistein induces mRNA expression of (A) *UGT1A9* and (B) *UGT1A10*. PPAR $\gamma$  antagonist GW9662 abolishes genistein action. The mRNA expression was normalised to 18S rRNA. Data are in duplicate from 2 experiments. T-test\*\*\*P<0.001.

#### 4.3.4 Analysis of *UGT1A8-1A10* gene regulation by PPAR $\gamma$

To examine whether PPAR $\gamma$  positively regulates endogenous *UGT1A8*, a plasmid which allows bicistronic expression of PPAR $\gamma$  and retinoid X receptor-alpha (RXR $\alpha$ ) was transfected into Caco-2 cells in the presence or absence of the PPAR $\gamma$  ligand rosiglitazone. The bicistronic expression plasmid was used because PPAR $\gamma$  binds to its response elements within gene promoters as obligate heterodimers with RXR (Ahmadian *et al.*, 2013),

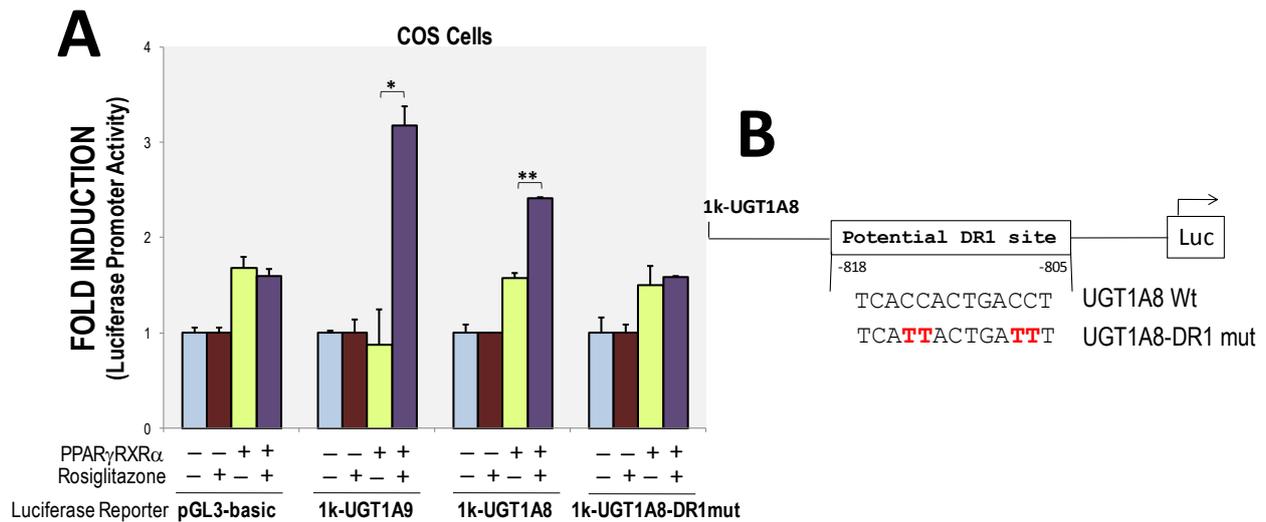
Caco-2 cells transfected with the PPAR $\gamma$  and RXR $\alpha$  expression plasmid and treated with PPAR $\gamma$  agonist rosiglitazone show greater increase in *UGT1A8* transcript levels than cells transfected with empty vector (Figure 4.10). This data confirms that PPAR $\gamma$  positively activates the endogenous *UGT1A8* gene.



**Figure 4.10. PPAR $\gamma$ -RXR $\alpha$  overexpression enhances the ability of a PPAR $\gamma$  agonist to induce UGT1A8 mRNA expression in Caco-2 cells.**

Caco-2 cells were transfected with mPPAR $\gamma$ 2ARXR $\alpha$ -pcDNA3 vector in the presence and absence of PPAR $\gamma$  activator rosiglitazone (10  $\mu$ M). pcDNA3 was used as the empty vector control. The mRNA target expression was normalised to 18S rRNA. Data are in duplicate from 2 experiments. T-test \*\*\*P<0.001.

While studies to this point clearly show that genistein action on the *UGT1A8-1A10* -1kb promoters (and likely -0.25kb promoters) is mediated by PPAR $\gamma$ , the identity of the PPAR $\gamma$  binding element within these promoters remains undefined. To begin to address this, we co-transfected -1kb-*UGT1A8* and *UGT1A9* promoter constructs with mPPAR $\gamma$ 2ARXR $\alpha$  in COS-7 cells to replicate the previously published study that defined a functional PPAR binding element (DR1) at -719 to -706nt in the *UGT1A9* promoter (Barbier *et al.*, 2003). As shown in Figure 4.11A, PPAR $\gamma$  and RXR $\alpha$  overexpression together with rosiglitazone treatment increased *UGT1A9* promoter activity by 3-fold; this result was similar to that previously published (Barbier *et al.*, 2003). PPAR $\gamma$  and RXR $\alpha$  overexpression combined with rosiglitazone treatment also increased *UGT1A8* promoter activity by 2.4-fold. Mutation of the putative PPAR (DR1) element at location -818 to -805nt in the *UGT1A8* promoter (Figure 4.11B) reduced induction of the *UGT1A8* promoter, suggesting that rosiglitazone-liganded PPAR $\gamma$  mediates induction of the *UGT1A8* promoter via the DR1 element.



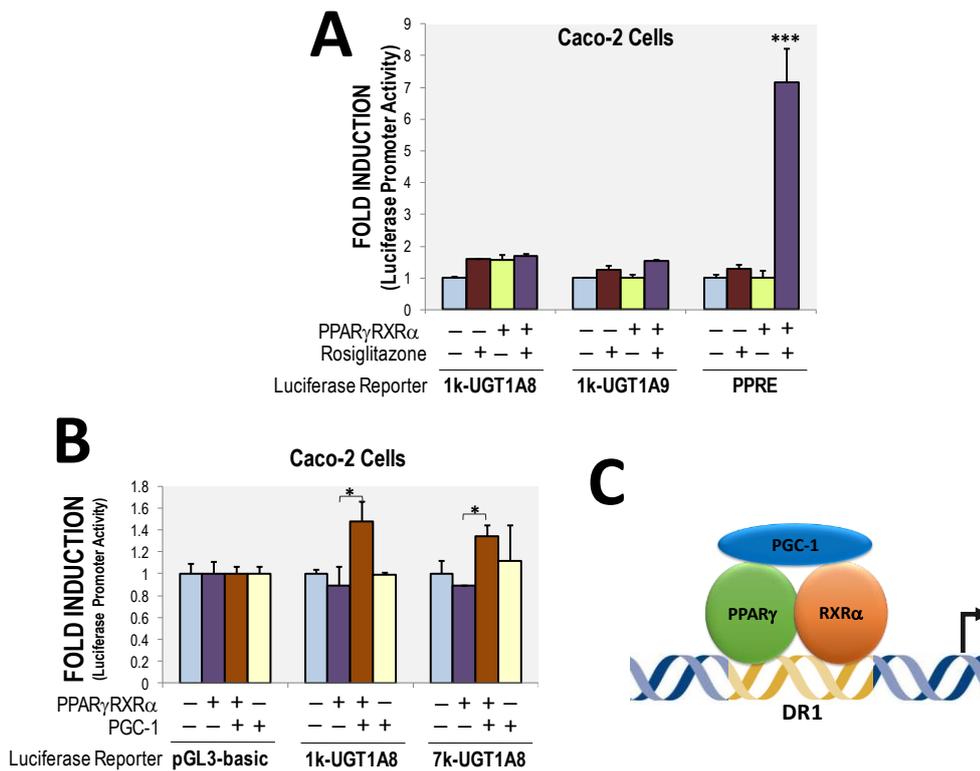
**Figure 4.11. PPAR $\gamma$  induces *UGT1A8* promoter activity via a PPRE at location -818 to -805nt in COS-7 cells.**

A. The wild-type -1kb *UGT1A8* and *UGT1A9* promoters as well as a mutated version of the -1kb *UGT1A8* promoter lacking the predicted PPAR binding element (DR1) were co transfected into COS-7 cells with the mPPAR $\gamma$ 2ARXR $\alpha$ -pcDNA3 expression vector. At 24-hours post transfection, rosiglitazone (10  $\mu$ M) or vehicle was added. Cells were harvested after 48 hours of treatment for luciferase assay. pRL-null *Renilla* reporter vector was used to normalise for transfection efficiency. The data is presented relative to the pcDNA3 control transfection condition (set at a value of 1). Data are in duplicate from 2 experiments. T-test \*P<0.05; \*\*P<0.005. B. Schematic of the mutated-1kb-*UGT1A8* promoter reporter construct. The predicted PPAR binding element (DR1) within the promoter is indicated with the mutated bases shown in red font.

The *UGT1A8* and *UGT1A9* promoter study shown in Figure 4.11A was repeated in Caco-2 cells to determine whether the regulatory mechanisms would be conserved between different cell models.

In the Caco-2 cell line promoter study, we employed three promoter constructs: the -1kb-*UGT1A8* and *UGT1A9* promoters, and a synthetic PPRE reporter construct that acted as a positive control.

As presented in Figure 4.12, activation by rosiglitazone-liganded PPAR $\gamma$  was only observed for the synthetic PPRE reporter construct in Caco-2 cells (Figure 4.12A). There was no activation of the -1kb *UGT1A8* and *UGT1A9* promoters by rosiglitazone-liganded PPAR $\gamma$  in Caco-2 cells. This was in contrast to results obtained in COS-7 cells (see Figure 4.11A). When an expression vector for the PPAR $\gamma$  co-activator-1 (PGC-1) was co-transfected, a modest (<1.5-fold) promoter activation was observed (Figure 4.12B). This data suggests that liganded PPAR $\gamma$  may require additional factors to regulate *UGT1A8* promoter activity in Caco-2 cells and these factors may include PGC-1 (Figure 4.12C).



**Figure 4.12. PPAR<sub>γ</sub> may require the co-activator PGC-1 to induce UGT1A8 promoter activity in Caco-2 cells.**

A. The -1kb-*UGT1A8* and *UGT1A9* promoter constructs are not induced by expression of PPAR<sub>γ</sub>-RXR<sub>α</sub> and treatment with PPAR<sub>γ</sub> agonist rosiglitazone in Caco-2 cells. The synthetic PPRE construct is induced by expression PPAR<sub>γ</sub>-RXR<sub>α</sub> and treatment with rosiglitazone. B. The -1kb and -7kb-*UGT1A8* promoter constructs are weakly induced by co-expression of PPAR<sub>γ</sub>-RXR<sub>α</sub> and coactivator PGC-1 in Caco-2 cells. pRL-null *Renilla* reporter vector was used to normalise for transfection efficiency of luciferase reporter. The data is presented relative to the pcDNA3 control transfection condition (set at a value of 1). Data are in duplicate from minimum of 3 experiments. T-test \*P<0.05; \*\*\*P<0.001. C. Schematic representation of PPAR<sub>γ</sub> action in Caco-2 cells; PPAR<sub>γ</sub>-RXR<sub>α</sub> dimers recruit PGC-1 to help activate the target gene.

There may be multiple explanations for why the *UGT1A8* promoter responds differently to rosiglitazone-liganded PPAR<sub>γ</sub> in COS-7 cells relative to Caco-2 cells (Figure 4.12A). One possibility is that other factors required for PPAR<sub>γ</sub> function at the *UGT1A8* promoter (including but not limited to PGC-1) are limiting in Caco-2 cells.

In addition, it is important to note that the binding sites for PPAR<sub>γ</sub> and HNF4<sub>α</sub> are very similar both corresponding to the direct-repeat1 (DR-1) configuration of AGGTCAXAGGTCA, leading to the possibility of competition for binding of these factors to the same DNA elements. The DR1 element located at -719nt/-706nt in the *UGT1A9* promoter (which corresponds to the DR1 located at -818nt/-805nt in the *UGT1A8* promoter) has been reported to bind to HNF4<sub>α</sub>, supporting the potential for

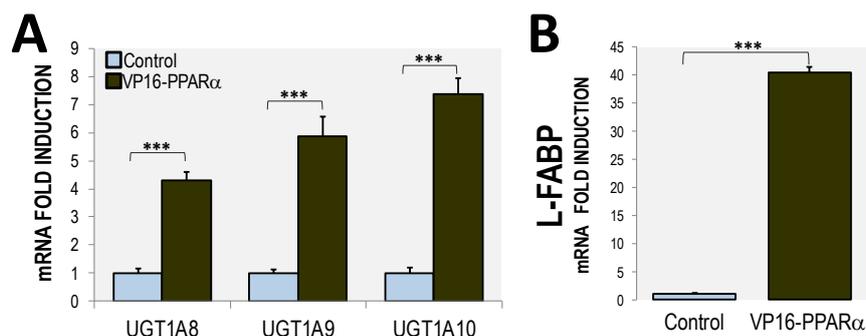
PPAR $\gamma$  and HNF4 $\alpha$  competition. The absence of HNF4 $\alpha$  in COS7 cells (Jiang *et al.*, 1995), compared to its abundance in Caco-2 cells, might provide another explanation for why PPAR $\gamma$  can induce the *UGT1A9* and *UGT1A8* promoters via the DR1 element in COS7 cells but not in Caco-2 cells.

Finally, other cell-type specific variables such as nuclear receptor phosphorylation (which can alter ligand binding, DNA binding, and recruitment of co-factors) could be responsible for cell-type specific promoter activation. Several kinases are known to phosphorylate PPARs in different contexts (Diradourian *et al.*, 2005). In particular, (Chen *et al.*, 2003) have shown that in the colon cancer cell line HT-29, the PPAR $\gamma$  ligand ciglitazone induces ERK1/2 activity, resulting in PPAR $\gamma$  phosphorylation which affects transcriptional activity. It is conceivable that genistein has as yet unknown effects on PPAR $\gamma$  phosphorylation and this would be a useful future direction for this study.

#### **4.3.5 Potential role of PPAR $\alpha$ in regulating *UGT1A8-1A10* gene expression**

Like PPAR $\gamma$ , PPAR $\alpha$  is also known to mediate soy-isoflavone action (Ricketts *et al.*, 2005). In this section we investigated whether PPAR $\alpha$  could induce *UGT1A8* mRNA expression. Caco-2 cells were transfected with an expression construct encoding a VP16-PPAR $\alpha$  fusion protein. This protein contains the PPAR $\alpha$  DNA binding region linked to a constitutive activation domain (VP16). The VP16 activation domain promotes assembly of an initiation complex and recruitment of coactivators to activate gene transcription; hence PPAR $\alpha$  ligand treatment is not required. The VP16-PPAR $\alpha$  expression plasmid was co-transfected with a GFP-expression plasmid so that transfected Caco-2 cells could be identified and collected by FACS sorting. Thus, only GFP/VP16-PPAR $\alpha$ -expressing cells were collected for RNA analysis.

As shown in Figure 4.13, VP16-PPAR $\alpha$  induced endogenous *UGT1A8*, *-1A9* and *-1A10* mRNA levels by between 4 and 7 fold. A well-known PPAR $\alpha$  target liver fatty acid-binding protein (L-FABP) (Landrier *et al.*, 2004), was increased around 40 fold. This data suggests that PPAR $\alpha$  binds to and regulates the *UGT1A8-1A10* genes.



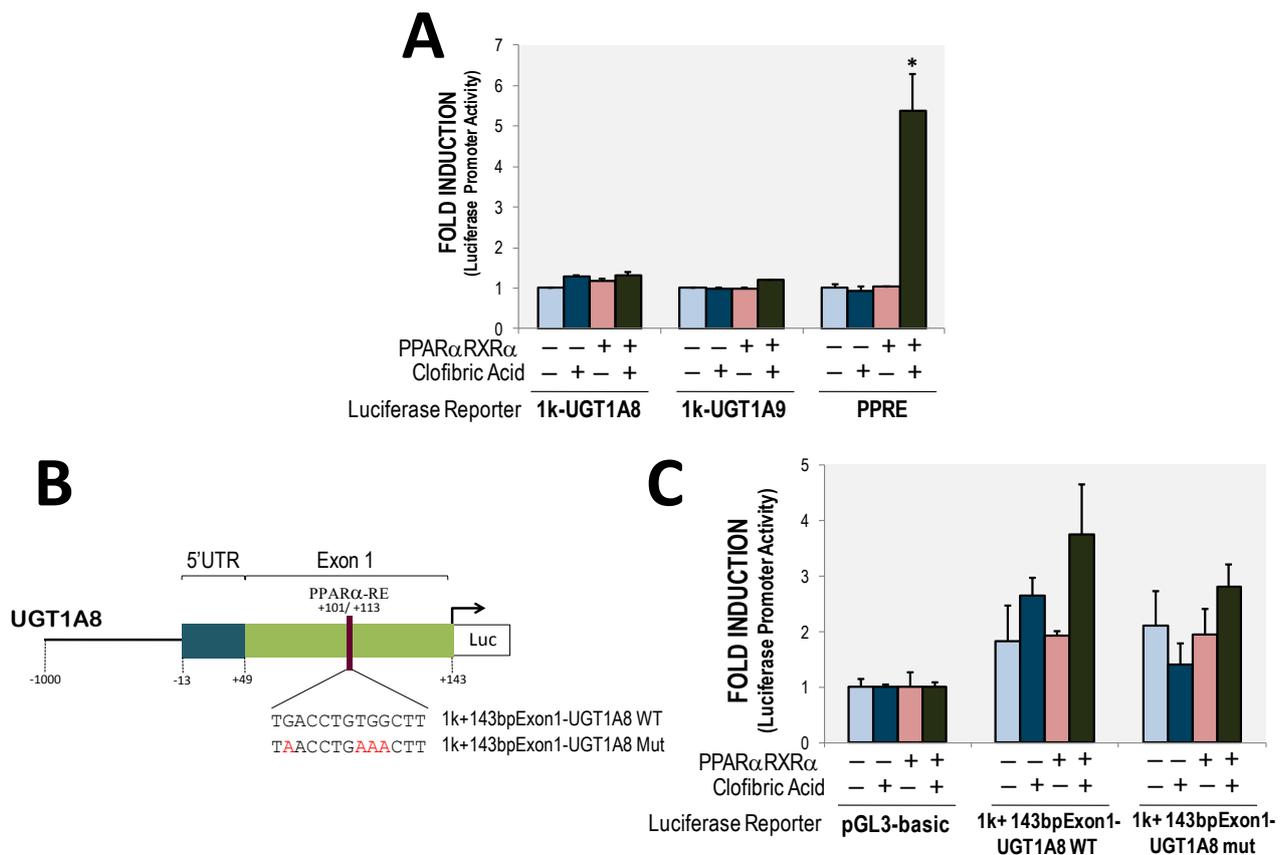
**Figure 4.13. VP16-PPAR $\alpha$  induces UGT1A8, UGT1A9, UGT1A10 and L-FABP mRNA expression.**

Caco-2 cells were transfected with VP16-PPAR $\alpha$  and GFP expression vectors. FACS sorting was used to enrich expressing cells. A. UGT1A8, -1A9 and -1A10 mRNA levels assessed by qRT-PCR. B. L-FABP mRNA level assessed by qRT-PCR. Gene expression was normalised to 18S rRNA and then to the control (empty vector) condition. Data are in duplicate from two experiments. T-test \*\*\*P<0.001.

Next, we transiently co-transfected Caco2 cells with the -1kb-*UGT1A8* and -1A9 promoter constructs and the mPPAR $\alpha$ 2ARXR $\alpha$  expression plasmid in the presence or absence of ligand (500  $\mu$ M of clofibric acid). PPAR $\alpha$ 2ARXR $\alpha$  and clofibric acid had no effect on the *UGT1A8* or -1A9 promoters (Figure 4.14A). However, the synthetic PPRE reporter construct was activated over 5 fold by the combination of PPAR $\alpha$ 2ARXR $\alpha$  and clofibric acid (Figure 4.14A). These results echo those seen with transfection of the PPAR $\gamma$ 2ARXR $\alpha$  expression plasmid in Caco2 cells (see section 4.3.4 of this Chapter); similar explanations can be postulated for the failure of PPAR $\alpha$  and PPAR $\gamma$  to induce the UGT promoters in Caco-2 cells (see section 4.3.4).

The studies to this point indicated that PPAR $\alpha$  and PPAR $\gamma$  activated by high affinity ligands (rosiglitazone and clofibric acid) could induce the endogenous UGT1A8-1A10 mRNA in Caco2 cells. But these factors could not induce the various *UGT* promoter constructs that were tested in Caco-2 cells. We have proposed that induction in Caco-2 cells requires additional protein factors/modifications that are lacking in these cells. However, it is also possible that *UGT* promoter regulation in these cells requires cis elements that are not contained within the tested promoter regions. Bioinformatic prediction of transcription factor binding sites in the *UGT1A8* promoter using the Genome Browser at the University of California Santa Cruz (UCSC) and SABiosciences' Text Mining Application (<http://www.sabiosciences.com>) revealed the presence of a potential PPAR $\alpha$

binding site location between +101 to +113 bp from the *UGT1A8* TSS (i.e. overlapping part of exon-1). This segment downstream of the TSS is not included in the existing promoter constructs that were tested, To investigate any role of this predicted PPAR $\alpha$  site in *UGT1A8* promoter activation in Caco-2 cells, we cloned an extra 143 bp sequence including the predicted PPAR $\alpha$  binding site at the 3' side of the existing -1kb promoter fragment in the -1kb-*UGT1A8* pGL3 construct (Figure 4.14B).



**Figure 4.14. *UGT1A8* and *-1A9* -1kb promoters do not respond to ligand-activated PPAR $\alpha$  in Caco-2 cells.**

A. PPAR $\alpha$ -RXR $\alpha$  was coexpressed in Caco-2 cells with the -1kb-*UGT1A8* and *UGT1A9* promoters or with a synthetic PPRE reporter. Cells were treated with PPAR $\alpha$  agonist clofibrilic acid for 48 hours. Only the PPRE construct responded to PPAR $\alpha$  activation B. Schematic showing the proximal 143bp spanning the 5'UTR and part of exon-1 of the *UGT1A8* gene that was inserted into the -1kb-*UGT1A8*-pGL3 vector (generating -1kb+143bpExon1-*UGT1A8*). The additional 143bp contains a predicted PPAR $\alpha$  response element located at +101 to +113. Wildtype and mutant promoter constructs were made, with red font indicating the mutated nucleotides in the predicted PPAR binding site. C. The 1k+143bpExon1-*UGT1A8* wildtype and mutant promoters were co-transfected with PPAR $\alpha$ /RXR $\alpha$  in the presence and absence of clofibrilic acid. Empty pCDNA3 vector was used as a control. pRL-null *Renilla* reporter vector was used to normalise for transfection efficiency of luciferase reporter. Data are in duplicate from minimum of 3 experiments. T-test\*P<0.05.

The new -1kb+143bp *UGT1A8* promoter construct (containing the additional 143bp segment that overlaps the new predicted PPAR $\alpha$  binding site) did appear to be slightly responsive to expression of mPPAR $\alpha$ 2ARXR $\alpha$  in the presence of clofibrilic acid, but unfortunately the results were variable between replicates and not statistically significant (Figure 4.14C). Hence, we cannot determine whether this element plays any significant role in the observed induction of endogenous *UGT1A8* mRNA expression by PPAR $\alpha$ .

Ultimately, the conclusion drawn from the studies in this section is that expression of the endogenous *UGT1A8-1A10* genes is regulated by genistein by a mechanism that appears to involve PPAR and ER factors. These *UGT* genes are also regulated by PPAR factors when they are activated by classical ligands (rosiglitazone, clofibrilic acid). However, we are unable to define discrete functional PPAR binding elements that mediate the induction of the *UGT1A8-1A10* promoters by PPAR factors in Caco2 cells. Moreover, the elements involved in activation by PPAR in Caco-2 cells may be different to the elements involved in other cell types (such as COS7).

#### **4.3.6 PPAR $\alpha$ and PPAR $\gamma$ attenuate HNF4 $\alpha$ -mediated *UGT1A9* activation**

It has been already mentioned that the major intestinal nuclear receptors PPARs and HNF4 $\alpha$  can interact and may also compete for binding to common DNA regulatory motifs. To investigate these phenomenon further, we studied the -1kb *UGT1A9* promoter that contains the previously identified DR1 PPAR response element located at -719 to -706 nt (Barbier *et al.*, 2003) as well as a the novel CDX2/HNF4 $\alpha$  composite element around -50 to -45nt that we characterized extensively in Chapter 3 (designated in Chapter 3 as the -47HNF4 $\alpha$ RE). Both of these elements are predicted to be able to bind to PPAR factors and also HNF4 $\alpha$ . We compared the activity of the wild type -1kb *UGT1A9* promoter to that of variants in which either the DR1 PPAR response element located at -719 to -706 nt had been mutated, or the CDX2/HNF4 element (for the purpose of this studies in this section now designated as as a DR1 half site) at -50 to -45nt had been mutated. Figure 4.15A shows the sequence elements of these two elements with the mutated nucleotides indicated.



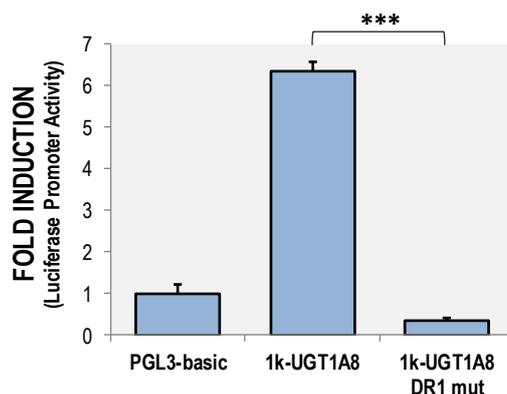
since mutating this element only reduces this activation by about half. This is in contrast to studies in HepG2 cells, which identified one functional HNF4 $\alpha$ -RE within the 1kb-*UGT1A9* promoter (Barbier *et al.*, 2005).

Figure 4.15C and B confirm our previous data (Figure 4.14A and Figure 4.12A), showing that PPAR $\alpha$  or  $\gamma$  alone does not activate the -1kb *UGT1A9* promoter. However, when PPAR $\alpha$  or  $\gamma$  are coexpressed with HNF4 $\alpha$ , *UGT1A9* promoter activation by HNF4 $\alpha$  is reduced. In the 1kb-*UGT1A8* promoter with mutated DR1 half-motifs, activation by HNF4 $\alpha$  was abolished by coexpression of PPAR $\alpha$  and  $\gamma$ . The synthetic PPRE construct was used as a control and, as expected, the response of the promoter to PPAR $\alpha$  and  $\gamma$  in the presence of ligand was not affected by HNF4 $\alpha$  (Figure 4.15D).

Overall, these data suggest that: 1) the element at -50 to -45nt binds and mediates induction by HNF4 $\alpha$ , and this binding may be influenced by PPAR factors; 2) the DR1 PPAR element at -719 to 706 nt that was previously shown to bind PPAR factors in other contexts (Barbier *et al.*, 2003) does not mediate induction by PPAR factors in Caco-2 cells, this element also does not mediate any significant induction by HNF4 $\alpha$ . One possible pathway for PPAR-to influence HNF4 $\alpha$  binding or activity is post-transcriptionally as discussed further in the Conclusions section. Future studies will be required to determine how binding of HNF4 $\alpha$  is influenced by PPAR factors and the extent to which these factors form a regulatory network.

#### **4.3.7 The putative DR1 PPAR response element in the *UGT1A8* promoter is important for transcriptional activation**

Using Caco-2 cells, we analysed the importance of the putative PPAR response element in the *UGT1A8* promoter at position -818 to -805 nt, which corresponds to the functional DR1 PPAR response element in *UGT1A9*. As shown in Figure 4.16, mutating this element decreases *UGT1A8* promoter activity below basal level. This result suggests that, although we did not observe PPAR-mediated *UGT1A8* promoter activation in Caco-2 cells via this DR1 site, the element is important for basal *UGT1A8* promoter activity. It is possible that the binding of endogenous PPAR $\alpha$  and  $\gamma$ , which are abundant in Caco-2 cells, to the DR1 element contributes to basal promoter activity.



**Figure 4.16. A DR1 PPAR response element is important for *UGT1A8* basal promoter activity.**

Caco-2 cells were transfected with wild-type -1kb *UGT1A8* promoter construct and a variant in which the DR1 element was mutated, and subsequently assayed for luciferase expression. Promoter-less pGL3 basic was used as negative control. Data are in duplicate from minimum of 3 experiments. T-test\*\*\*P<0.001.

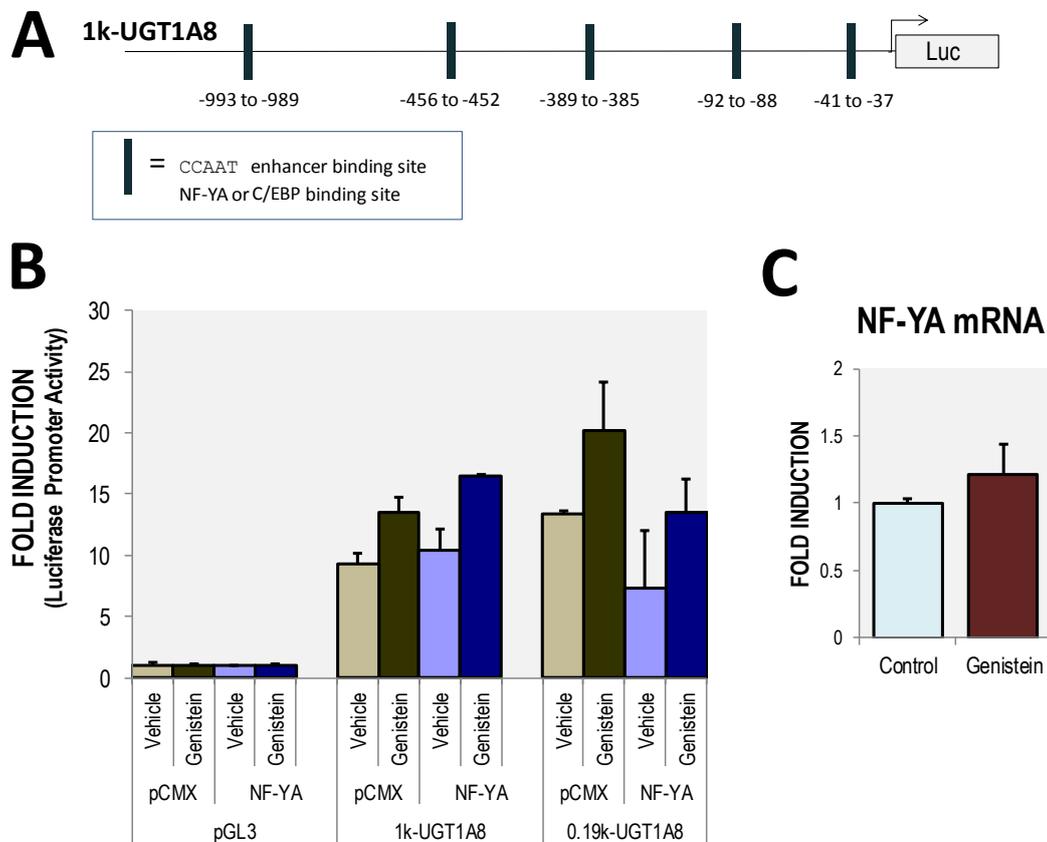
#### 4.3.8 Neither NF-YA, nor C/EBP mediates genistein action in *UGT1A8* regulation

As previously mentioned, genistein is the component of dietary soybeans that has been primarily associated with a reduced risk of cancer, possibly due to its ability to induce growth arrest (Choi *et al.*, 1998, Matsukawa *et al.*, 1993, Kuzumaki *et al.*, 1998) by targeting genes involved in the cell cycle. During this study, we noticed that Caco-2 cell growth was reduced after genistein treatment.

A growth arrest and DNA damage-inducible gene 45 (*gadd45*) gene has been reported to be regulated by genistein in a prostate cancer cell line via nuclear factor-Y (NF-Y) (Oki *et al.*, 2004). Genistein induces NF-Y to bind a response element (CCAAT) in the *gadd45* promoter and induce transcriptional activity (Oki *et al.*, 2004). NF-Y mediated genistein action via was also reported by Shimizu *et al.* in regulation of bone sialoprotein (BSP), a protein involved in bone metastasis (Shimizu and Ogata, 2002).

Like the BSP and *gadd45* promoters, sequence analysis of the -1kb *UGT1A8* promoter revealed a considerable number of CCAAT sites, as schematically represented in Figure 4.17A. To examine whether NF-Y mediates *UGT1A8* promoter activation in Caco-2 cells, we co-transfected NF-Y subunit A (NF-YA) with two *UGT1A8* promoter constructs into the Caco-2 cells in the presence and absence of genistein. The first promoter construct used in the assay was the -1kb-*UGT1A8*

construct, which has 5 CCAAT sites, we also tested the -0.19kb-UGT1A8 construct that contains 2 CCAAT sites. As shown in Figure 4.17B, overexpressing NF-YA did not activate the *UGT1A8* promoters, and genistein does not appear to activate NF-YA to induce the promoter. We also analysed NF-YA mRNA expression in Caco-2 cells following genistein treatment; in contrast to the findings of Oki *et al.*, who reported an increase in NF-YA transcripts (Oki *et al.*, 2004), genistein had no effect on NF-YA levels in Caco-2 (Figure 4.17C). Thus, induction of UGT1A8 expression in Caco-2 cells by genistein is unlikely to be mediated by NF-YA.

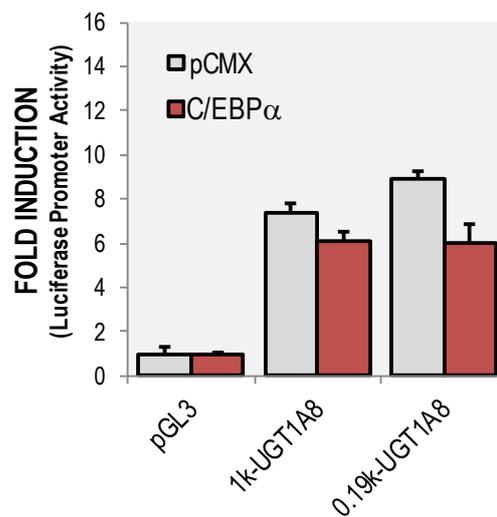


**Figure 4.17. NF-YA does not mediate induction of the *UGT1A8* promoter by genistein.**

A. Schematic diagram of -1kb-UGT1A8 luciferase construct with the positions of CCAAT binding sites indicated. B. Caco-2 cells were co-transfected with NF-YA and -1kb and -0.19kb UGT1A8 promoter constructs, in the presence or absence of genistein (48-hours, 10  $\mu$ M). The relative luciferase activity of UGT1A8 promoter construct is indicated as fold-induction over pGL3 basic (promoter-less) (set as value of 1). Internal control pRL-null was used to normalise the luciferase activity. C. Caco-2 cells were treated with 10  $\mu$ M genistein for 48 hours and assayed for NF-YA mRNA expression. The mRNA target expression was normalised to 18S rRNA. Data are in duplicate from 3 experiments.

Another factor that can bind to CCAAT-box motifs is the CCAAT/enhancer-binding protein alpha (C/EBP $\alpha$ ) (Johnson and McKnight, 1989). C/EBP $\alpha$  is expressed at high level in the intestine, mainly in the proximal end; it is detected only in differentiated cells of intestinal villi, not in the crypt region (Chandrasekaran and Gordon, 1993). C/EBP $\alpha$  has been identified to form a regulatory feedback-loop with PPAR $\gamma$ , with the two transcription factors up-regulating each other to maintain a differentiated state (Ramji and Pelagia, 2002).

To examine the role of C/EBP $\alpha$  in *UGT1A8* promoter activity, we co-transfected C/EBP $\alpha$  with the -1kb and -0.19kb *UGT1A8* promoter constructs in Caco-2 cells. As shown in Figure 4.18, C/EBP $\alpha$  did not increase promoter activity. Thus, overall it is unlikely that induction of the *UGT1A8-1A10* genes by genistein is mediated by CCAAT motifs as it is in other genistein-regulated genes.



**Figure 4.18. C/EBP $\alpha$  over-expression in Caco-2 cells does not induce *UGT1A8* promoter activation.**

Caco-2 cells were co-transfected with C/EBP $\alpha$  and -1kb and -0.19kb *UGT1A8* promoter constructs. The relative luciferase activity of the *UGT1A8* promoter construct is presented as fold-induction over promoter-less pGL3 basic construct (set as value of 1). Internal control pRL-null (*Renilla*) was used to normalise firefly luciferase activity. Data are in duplicate from 3 experiments.

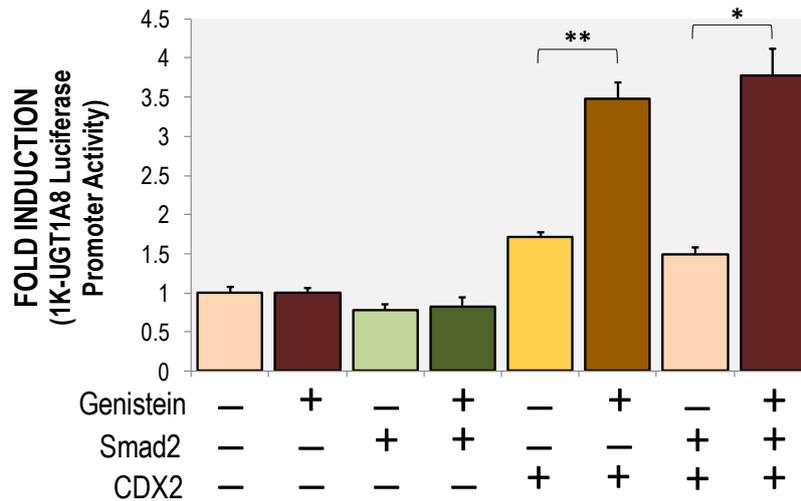
#### 4.3.9 Possible roles for CDX2 and SMAD2 in *UGT1A8* regulation

Interaction of genistein with numerous transcription factor and subsequent modulation of their transcriptional activity has been reported (reviewed in (Banerjee *et al.*, 2008)). A study by Yu *et al.* showed that genistein can induce phosphorylation of Smad proteins (Smad2 and -3), and facilitate

the translocation of phosphorylated Smad (pSmad) from cell surface to nucleus to bind a Smad binding element (SBE) and activate transforming growth factor beta 1 (TGF- $\beta$ 1), an intestinal expressed gene that induces growth arrest and apoptosis of colon cancer cells (Yu *et al.*, 2005).

Another study (Mizutani *et al.*, 2011) found that the SBE is bound by pSMAD with a low affinity, and interaction with other transcription factors is required to induce transcriptional activation in a tissue-specific manner. In human oesophageal cells, pSMAD interacts with CDX2 forming a transcriptional complex to bind and activate the promoter of Mucin-2 (*MUC-2*), a goblet cell secreted protein (Mari *et al.*, 2014).

To explore the possibility that induction of *UGT1A8* by genistein involves SMAD factors in interaction with CDX2, we cotransfected SMAD2 and CDX2 expression plasmids with the 1k-*UGT1A8* promoter construct, in the presence or absence of genistein. As shown in Figure 4.19, CDX2 increased *UGT1A8* promoter activity by approximately 1.7 fold and treatment with 10  $\mu$ M genistein for 48 hours demonstrated further induction to 3.5 fold. SMAD2 had no effect on the *UGT1A8* promoter, alone or in combination with CDX2, with or without genistein. This study used a genistein concentration of 10  $\mu$ M, which may not be adequate to phosphorylate SMAD2. In their study, Yu *et al.* used 60  $\mu$ M of genistein to induce SMAD phosphorylation (Yu *et al.*, 2005) . Unfortunately, time constraints prevented us from repeating the studies at a higher dose of genistein. Nevertheless, an interesting interaction between genistein and CDX2 was detected, with the two synergising to induce *UGT1A8* promoter activation. The genistein and CDX2 synergy found here is supported by a recent study (Du *et al.*, 2016) showing that a synthetic genistein derivative acts as an inducer for CDX2 binding to the APC and Axin2 promoters in colon cancer cells (Du *et al.*, 2016).



**Figure 4.19. UGT1A8 promoter response to genistein treatment in combination with SMAD2 and CDX2 overexpression.**

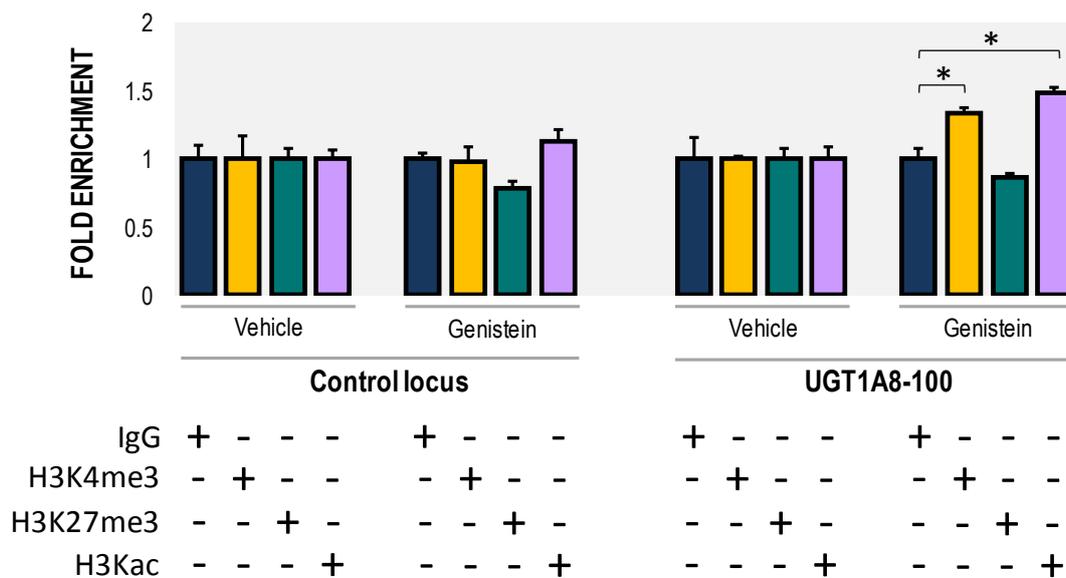
Caco-2 cells were co-transfected with SMAD2, CDX2 or a combination of SMAD2 and CDX2 expression plasmid and the -1kb-UGT1A8 promoter constructs. Cells were treated with 10  $\mu$ M genistein for 48 hours. The relative luciferase activity of UGT1A8 promoter construct is indicated as fold-induction over the promoter-less pGL3 basic construct (set as value of 1). Internal control pRL-null (*Renilla*) was used to normalise the firefly luciferase activity. Data are in duplicate from 3 experiments. T-test\* $P < 0.05$ ; \*\* $P < 0.005$ .

#### 4.3.10 Genistein action involves altered histone modifications in the *UGT1A8* proximal promoter region

In parallel with the studies described above that sought to uncover transcription factors that mediate induction of the *UGT* genes by genistein, we also examined whether genistein mediates epigenetic modifications at target loci. Genistein has been reported to be able to inhibit DNA methyltransferase activity, resulting in an increase in hypermethylation events and is associated with induction of tumour suppressor genes (Qin *et al.*, 2009, Majid *et al.*, 2009, Meeran *et al.*, 2010). Reduced DNA methylation is accompanied by changes in histone methylation and acetylation that render chromatin more accessible.

To assess the potential of genistein to induce chromatin modifications at target loci, we used ChIP to examine the distribution of activating and repressive histone modifications within the proximal *UGT1A8* promoter region. Caco-2 cells were treated with genistein treatment (48 hours, 10  $\mu$ M) and ChIP was performed using the antibodies that detect histone H3 lysine 4 tri-methylation (H3K4me3), histone H3 lysine acetylation (H3Kac); and histone H3 lysine 27 tri-methylation (H3K27me3).

The ChIP sample was analysed using qPCR with primers that bind to the UGT1A8 proximal (100 bp upstream from the UGT1A8 start site) as well as a control non-target locus (hch12-gene desert ChIP). As demonstrated in Figure 4.20, genistein modestly increased the level of the activating modifications H3K4me3 and H3Kac by 1.3 fold and 1.5 fold respectively within the 100 bp promoter region of *UGT1A8*. No change was observed for the repressive histone marker, H3K27me3. These results suggest that genistein might increase the accessibility of the UGT1A8 proximal promoter to transcription factors. These data are in broad accordance with a report that genistein induced H3 and H4 acetylation resulting in upregulation of *p21* and *p16* in human prostate cancer cells (Majid *et al.*, 2008). Overall, our data suggest that both genistein and butyrate may induce UGT1A8 gene expression at the epigenetic level, although clearly butyrate is a much more potent epigenetic regulator than genistein.



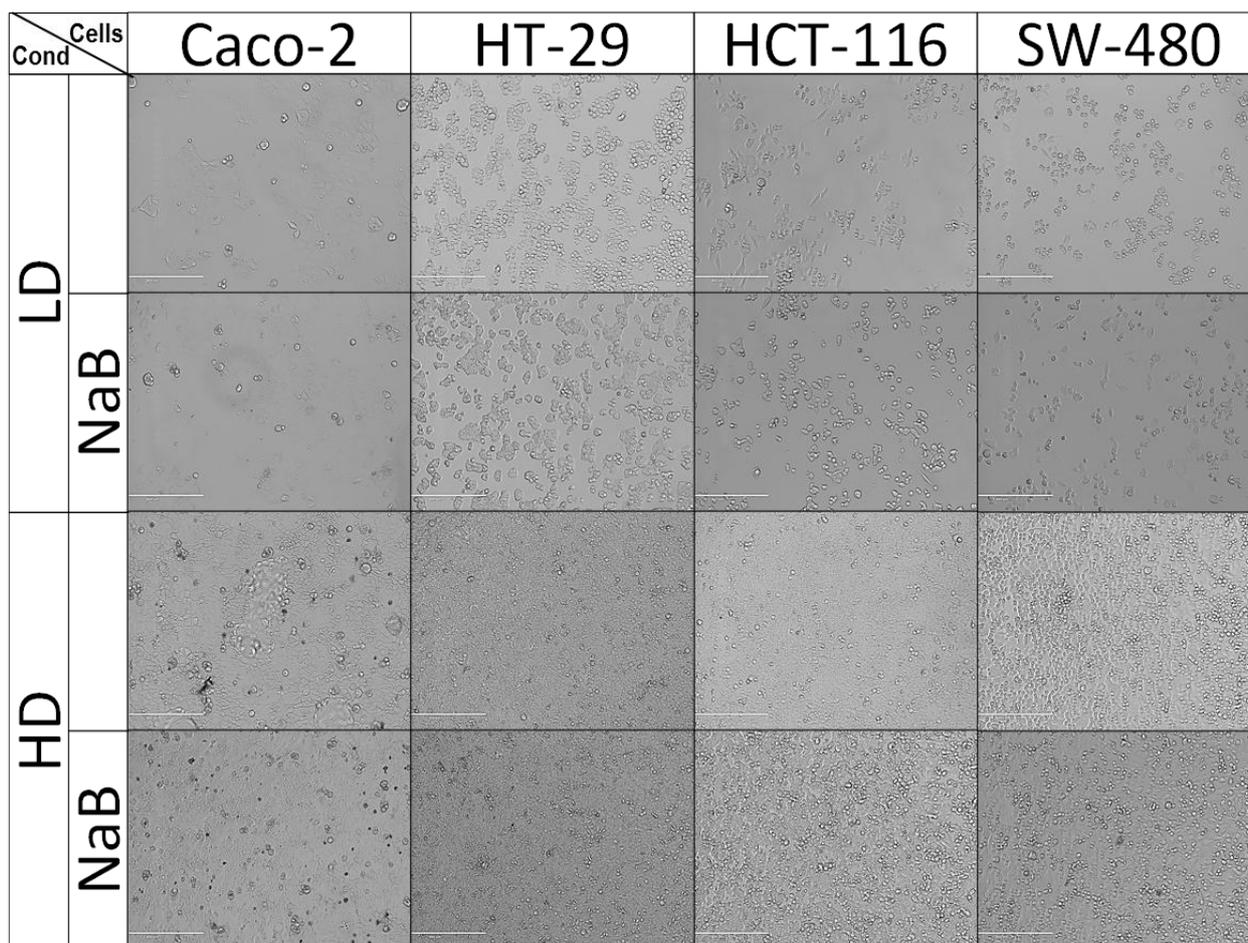
**Figure 4.20. Genistein induces H3K4me3 and H3Kac enrichment at the *UGT1A8* promoter.**

Caco-2 cells were treated with 10  $\mu$ M genistein or DMSO (vehicle control) for 48 hours and subjected to ChIP using antibodies for histone modifications (H3K4me3, H3K27me3 and H3Kac). PCR amplification of precipitated genomic DNA samples detected enrichment of the proximal *UGT1A8* promoter region in H3K4me3 and H3Kac ChIP samples from cells treated with genistein. Enrichment of the UGT1A8-100 region was calculated by normalisation to the control locus (hch12 desert ChIP), followed by normalisation over thenon-immune IgG control ChIP condition and then over vehicle treatment, which was set to a value of 1. Data are in duplicate from 2 experiments. T-test\* $P < 0.05$ .

#### 4.3.11 Comparison of gene expression in human colonic cancer cell lines

As described in the literature review (Chapter 1, Section 1.5.1), Caco-2 cells are the most commonly used colon-derived cell line to study intestinal UGTs. We mainly utilised Caco-2 cells in this project to study the regulation of *UGT1A8-1A10* genes; however, we consider it likely that some outcomes may have been different in other model cell lines with differing basal levels of relevant regulatory factors. To begin to understand the variation in the expression of these regulatory factors in colon-derived cell lines, we cultured 4 colonic cell lines (Caco-2, HT-29, HCT-116 and SW-480) in 4 culture conditions (low density and high-density culture in the presence and absence of the NaB) (Figure 4.21). The Caco-2, HT-29, and HCT-116 cells grew more rapidly than SW-480 cells. All cell lines showed retarded growth in response to 3 mM NaB but little cell death (Figure 4.21). At higher concentration (5mM NaB), cell death was observed with minimal cell survival (data not shown). This is consistent with previous reports using colorectal cancer cell lines, and may involve both autophagy and apoptosis (Zhang *et al.*, 2016). NaB treatment altered the morphology of Caco-2, HT-29 and SW-480 cells, both at low and high-density seeding cells, producing a more flattened shape and larger area. This was consistent with a report that colon cancer cells are more flattened when reaching differentiation (Kim *et al.*, 1980, Fok *et al.*, 2012). NaB has been shown to induce morphological HT-29 and HCT-116 differentiation (Saldanha *et al.*, 2014). We found that HCT-116 cells displayed a more spindle-like shape following NaB treatment, a different morphological change compared to the other three cell lines; nevertheless, the HCT-116 changes were consistent with previous reports of their differentiation process (Saldanha *et al.*, 2014).

Caco-2 cultures forms enterocyte-like cells when fully differentiated; this was the only cell line to show spontaneous formation of distinct domes once confluent, which is consistent with villus formation by enterocytes (Figure 4.21 shows Caco-2 cells in high-density seeding without NaB treatment).



**Figure 4.21. Growth and morphological variation of colon cancer cell lines.**

Four colon cancer cell lines, Caco-2, HT-29, HCT-116 and SW-480, were cultured in different conditions in low or high-density. Cells were treated with 3 mM differentiation agent, NaB, 48 hours after trypsinisation and seeding (control = water). LD= Low Density ( $5 \times 10^3$  cells/well of 48 well-plate); HD= High Density ( $1 \times 10^5$  cells/well of 48 well-plate); NaB= sodium butyrate. All images were taken using an EVOS® FL microscope in bright field lighting 24 hours after NaB (or water control) treatment using 10x magnification. Scale bar, 400  $\mu$ m.

Total RNA was extracted from all cell lines cultured as described above: i.e. Caco-2, HT-29, HCT-116 and SW-480 at low and high-density, with or without NaB treatment. The basal expression of UGT1A8, -1A9 and -1A10, were compared. As shown in Figure 4.22, differential expression of UGT1A8, -1A9 and -1A10 was observed. Interestingly, we found that the level of these three UGTs in the SW-480 cell line was extremely low; it could not be detected by the qPCR cycle threshold, and high-density condition slightly induced the basal level. When NaB was added, regardless of the seeding density, UGT1A8, -1A9 and -1A10 mRNA levels were increased around 3-5 fold (Figure 4.22D). A similar pattern was observed in HCT-116 cell line when NaB was used to treat

the cells; it induced the three UGTs at around 30 fold higher compared to that in SW-480 (Figure 4.22B). However, unlike UGT1A8, -1A9 and -1A10 mRNA in SW-480 which was induced by high-density seeding, the expression level of these genes were not affected by high-density seeding of HCT-116 cells (Figure 4.22B), suggesting that activation of UGT1A8, -1A9 and -1A10 is differentially regulated in SW-480 and HCT-116 cells, which was not examined further in this study. The extremely low basal levels of UGT1A8, -1A9 and -1A10 caused these cell lines to be less suitable models to examine UGT1A8, -1A9 and -1A10 expression in enterocytes.

Conversely, we observed very high basal expression of UGT1A8, -1A9 and -1A10 in HT-29 cells (Figure 4.22C), at over 600 times higher than that in Caco-2 cells (Figure 4.22A). Interestingly, while UGT mRNAs in Caco-2 were induced by high-density seeding and NaB treatment, the level of these UGTs (-1A8, -1A9 and -1A10) were not affected by both high-density and NaB, which suggest that the level of UGT expression in HT-29 may have reached the maximum level in the most accessible chromatin conformation.

We next examined the levels of CDX2 and HNF4 $\alpha$  mRNA in the Caco-2 and HT-29 cell lines under different conditions. As shown in Figure 4.22F, CDX2 mRNA in Caco-2 cells was increased by high-density seeding and butyrate treatment alone; the highest level was reached when high-density seeding and NaB treatment were combined. Interestingly, despite high mRNA levels of UGT1A8, -1A9 and -1A10 in HT-29 cells, the CDX2 level in HT-29 cells is very low, over 150 fold lower than that in Caco-2. However, NaB treatment induced CDX2 mRNA in HT-29 significantly to over 60 fold, suggesting that accessibility for transcription complexes to bind and activate CDX2 expression in HT-29 was modulated by NaB. This NaB-mediated induction of HNF4 $\alpha$  expression was not observed in HT-29 cells (Figure 4.22C).

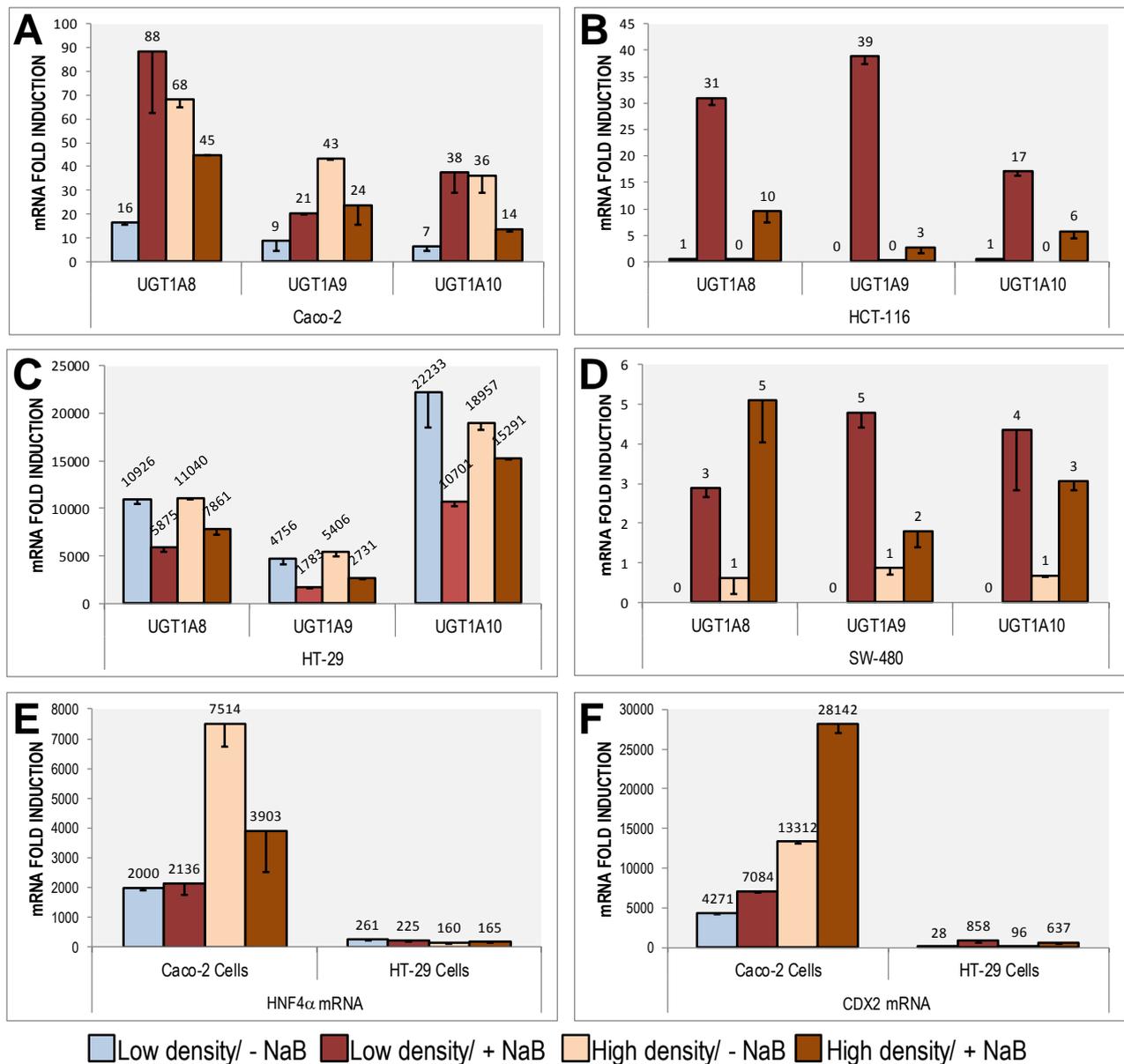
Figure 4.22E shows that in Caco-2 cells, high-density seeding induced HNF4 $\alpha$  mRNA by 4 fold relative to low-density seeding, whereas the addition of NaB to high-density cells surprisingly decreased the induction due to high-density. The induction of HNF4 $\alpha$  by high density culture plus NaB here was similar to HNF4 $\alpha$  induction by NaB-induced differentiation of Caco-2 (Figure 4.2).

Our results indicate that NaB did not affect HNF4 $\alpha$  levels in low-density Caco-2 cells; however NaB is reported to reduce HNF4 $\alpha$  binding to its target DNA sequence (Rada-Iglesias *et al.*, 2007). Moreover, as HNF4 $\alpha$  may also be self-regulated (Lu, 2016); reduced HNF4 $\alpha$  binding may decrease its own expression, which in this case, may explain the mechanism of NaB-mediated HNF4 $\alpha$  mRNA reduction.

Despite high UGT1A8, -1A9 and -1A10 mRNA levels in HT-29 cells, surprisingly, the level of HNF4 $\alpha$  expression in this cell was 10 times lower compared to that in Caco-2 (Figure 4.22E). Consistent with HNF4 $\alpha$  in Caco-2, NaB did not appear to affect HNF4 $\alpha$  levels in HT-29 cells. However, different to Caco-2, in which cell confluency induced HNF4 $\alpha$  expression, high-density HT-29 culture did not affect the endogenous HNF4 $\alpha$  level. As described in Chapter 3, although the level of CDX2 and HNF4 $\alpha$  in HT-29 cells is low, knocking down their expression decreased the UGT1A8 mRNA level significantly, confirming UGT1A8 transcriptional regulation by Cdx2 and HNF4 $\alpha$ . We hypothesise that these transcription factors are able to readily access their binding sites in the *UGT1A8* promoter, recruit the transcriptional complex and activate gene transcription. Further investigation in the future using ChIP analysis would be required to support this hypothesis.

Some of our studies examined whether co-transfection of the -1kb *UGT1A8* promoter with PPAR $\gamma$  and PPAR $\alpha$  in HT-29 cells could reveal activation of the promoter that was not evident in Caco-2 (data not shown). However, we found that experimental replications did not show consistent results; future study is required to address UGT1A8 regulation in HT-29 cells.

Although Caco2 cells effectively form a differentiated enterocyte that are useful in the study of intestinal UGTs, this cell type shows poor transfection efficiency. Poor efficiency can reduce the magnitude of response of endogenous genes to transfected effector plasmids, because only a small percentage of cells carrying the endogenous target gene receive the expression vector and also because the number of copies of expression vector per cell is often low. The following section discusses several methods that were tested to improve the transfection efficiency in Caco-2 cells.



**Figure 4.22. Basal mRNA expression of UGT1A8, -1A9, -1A10, CDX2 and HNF4α in human colonic cell lines.**

A. Caco-2. B. HCT-116. C. HT-29. and D. SW-480 cell lines were cultured in low or high density, and treated with 3 mM NaB or water control for 48 hours. RNAs were extracted from the cell culture 24 hours post-NaB treatment. qRT-PCR analysis was performed on mRNA expression of UGT1A8, -1A9 and -1A10. E. HNF4α and F. CDX2 mRNA expression in Caco-2 and HT-29 was analysed. LD= Low Density ( $5 \times 10^3$  cells/well of 48 well-plate); HD= High Density ( $1 \times 10^5$  cells/well of 48 well-plate); NaB= sodium butyrate. All values were normalised to 18S rRNA expression. Relative target gene expression level was calculated using the  $(2^{-(\text{Ct value})}) \times 10^8$ . Values were measured in duplicate and are derived from 3 experiments. Data is expressed as the mean  $\pm$  SE (bars).

### 4.3.12 Optimizing Caco-2 transfection methods

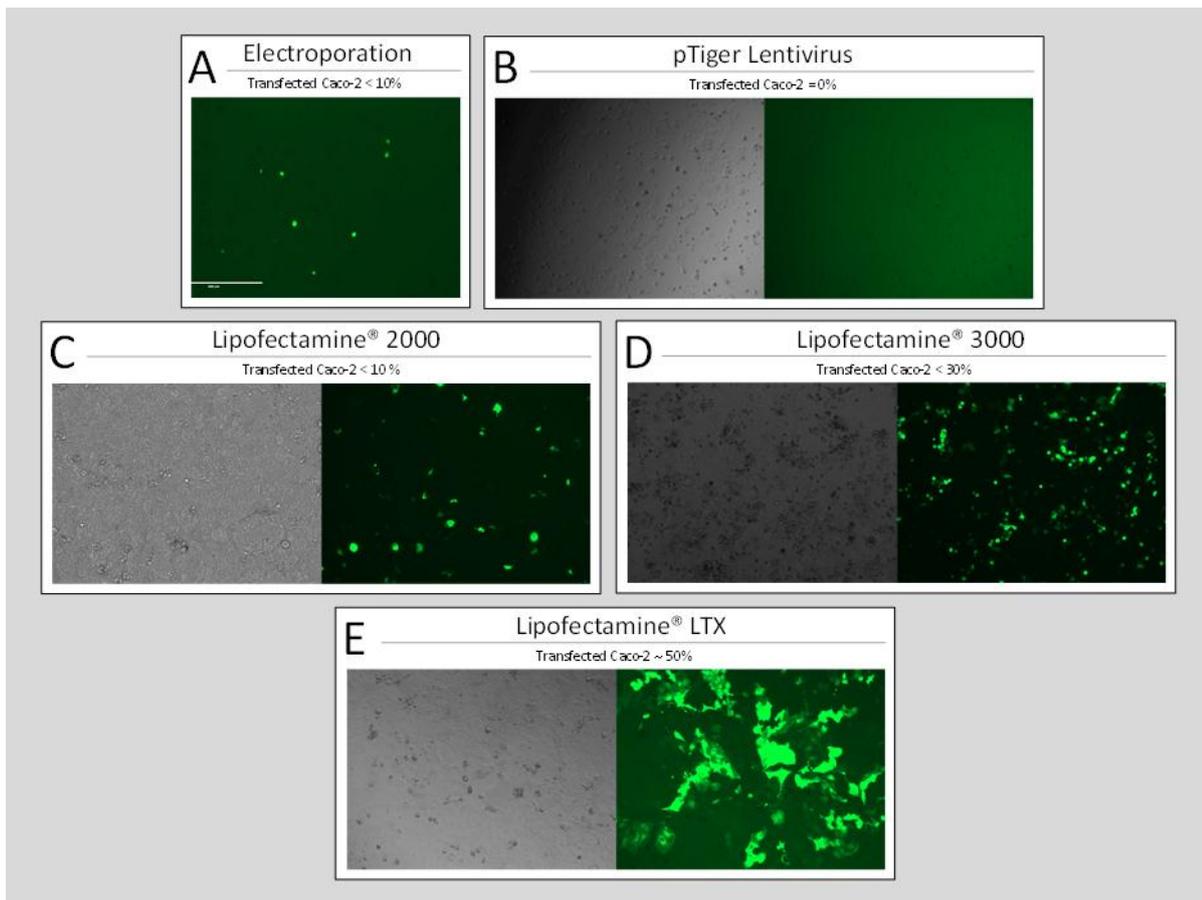
The comparison of colonic cell lines discussed above (4.3.11) confirms that among the cell line models tested here, Caco-2 remains the best choice to represent a differentiated enterocyte-like model in which to study regulation of the intestinal UGT genes. As discussed, transfection efficiency is often a major factor in determining a successful experiment and unfortunately the Caco-2 cell line is relatively resistant to standard lipid-based transfection. Therefore, during our study we tested several methods to improve transfection of Caco-2 cells. Most of luciferase reporter transfection experiments used Lipofectamine®2000 as the transfection agent, although we were aware that the transfection efficiency was quite low, usually 10% or less (Figure 4.23A). However poor efficiency is offset by the fact that the promoter-reporter constructs and the effector constructs that are tested for modulation of promoter activity, are all taken up by the same subset of cells.

Poor transfection efficiency is a problem when assessing the response of endogenous genes (i.e. at mRNA level) to a transfected effector plasmid. If the target gene is present in 100% of the cells but the effector is only present in 10% of cells, the response of the target to the effector will be minimal when measured across the whole cell population. To maximise efficiency, we used the reverse transfection method where cells were transfected in suspension for all of our studies (luciferase promoter response and mRNA analysis). In addition, we tested the ability of a variety of liposomal transfection agents to transfect a GFP-expressing plasmid by imaging and counting GFP-expressing cells. Of the three Lipofectamine variants, Lipofectamine®2000, Lipofectamine®3000 and Lipofectamine®LTX, we found that Lipofectamine® LTX provided the most efficient transfection at up to 50%, followed by Lipofectamine®3000. Lipofectamine®2000 was the least efficient transfection agent and also caused the most cell death necessitating transfection of cells at higher density which further reduced efficiency (see Figure 4.23).

Interestingly, siRNAs transfected much more efficiently than plasmids with all reagents and in fact Lipofectamine®2000 was very effective for transfection of siRNAs as judged using a fluorophore labelled siRNA and also by the percentage knockdown of the targeted gene (e.g. CDX2 or HNF4).

In general, transfection by electroporation is considered a very efficient procedure to introduce DNA into bacterial, plant and mammalian cells. We tested electroporation procedure for Caco-2 cells with an optimised electroporation buffer called iso-osmolar electroporation buffer (IEB). Although we observed very efficient transfection in two initial experiments (data shown in Chapter 3), ultimately, we could not obtain consistently high efficiencies whilst maintaining good cell survival. Future work could further optimise electroporation parameters such as cell health prior to electroporation, cuvette type, buffer composition, pulse settings and temperature.

We also tested the ability of feline immunodeficiency virus (FIV) based lentiviruses to transduce Caco-2 cells. Lentivirus was packaged using a GFP-pTiger viral expression plasmid and transduced into Caco-2 cells. We packaged virus under standard conditions that had previously proven to be efficient and compared spinfection of the Caco-2 cells with normal transduction (without spinning). Surprisingly, Caco2 cells were transduced very poorly under all conditions (Figure 4.23B). As a control we transduced HEK293T and this indicated a high viral titre (not shown). Thus we concluded that Caco-2 may be inherently resistant the FIV based lentivirus. Other virus types were not tested during this study.



**Figure 4.23. Optimization of transfection methods for Caco-2 cells.**

Multiple methods and reagents were tested to transfect Caco-2 cells using GFP-plasmids to assess transfection efficiency. All images were taken using an EVOS® FL microscope 48 hours post-transfection at 10x magnification. A. Representative example of Caco-2 cells electroporated in IEB (isoosmolar electroporation buffer) showing a low transfection efficiency of <10%. B. Caco-2 cells were unexpectedly resistant to pTiger lentivirus transduction showing no transduced cells. C. Representative image of Caco-2 transfection using Lipofectamine® 2000; transfection efficiency was <10%. D. Representative image of Caco-2 transfection using Lipofectamine® 3000; transfection efficiency was 10-30% E. Representative image of Caco-2 transfection using Lipofectamine® LTX ; this was the most effective reagent with good cell survival and transfection efficiency up to 50%.

## 4.4 Conclusions

Through screening of flavonoid-mediated UGT1A8 promoter induction in an enterocyte model of differentiated Caco-2, we discovered that genistein is highly effective at activating the *UGT1A8* promoter in NaB-induced differentiated Caco-2 cells. While NaB was shown in this study to be capable of increasing UGT1A8 expression on its own, the effect of genistein and NaB at the promoter and mRNA level was strongly synergistic. The use of NaB in generating a differentiated enterocyte-like Caco-2 model is somewhat physiologically relevant because intestinal microbiota naturally produces butyrate and it is an important regulator of intestinal homeostasis *in vivo*.

Therefore, this study begins to address the significant contribution of microbiota to the activities of dietary flavonoid compounds in the regulation of intestinal detoxification pathways mediated by UGTs. At high concentrations (10  $\mu$ M), genistein was reported to be mainly conjugated by UGT1A8 in intestinal microsomes (Tang *et al.*, 2009) thus induction of UGT1A8 by genistein could create a feedback loop that increases genistein metabolism and clearance.

The major part of the studies described in this Chapter sought to determine the mechanism(s) of induction of the *UGT1A8* gene by genistein. It appeared that this mechanism was very complex and no single mediator could be defined. Genistein induced UGT1A8 mRNA levels and this was blocked by both a PPAR antagonist and an ER antagonist, suggesting that both pathways might be involved. Varying length UGT1A8 promoter constructs (up to -7kb) also showed induction by genistein; however, this was only blocked by the PPAR antagonist and not the ER antagonist, suggesting that any contribution by the ER pathway was mediated by elements outside of this promoter region (or were post-transcriptional). In agreement with this, we further found that the *UGT1A8* promoter was induced by a classical PPAR agonist but not by estradiol.

We found that both short (-0.25kb) and long (7kb) UGT1A8 promoter constructs could mediate a response to the classical PPAR $\gamma$  ligand rosiglitazone, and also to genistein, and both responses were blocked by a PPAR antagonist. This result might be expected focus the search for a PPAR response element within the proximal -0.25kb region. However, the only previously functionally defined PPAR response element identified in the *UGT1A8-1A10* genes was located approximately -700 to -800nt upstream of the transcription start site (Uhlén *et al.*) in the *UGT1A9* gene, moreover we suspected that multiple PPAR REs might be present within these promoters. Hence, we focussed on identifying elements that might mediate the response of the -1kb *UGT1A8-1A10* promoters to PPAR factors in Caco-2 cells.

Subsequent analysis indeed showed that the all three of the -1kb *UGT1A8-1A10* promoters were induced by PPAR $\gamma$  ligand rosiglitazone, and also by genistein, and both responses were blocked by a PPAR antagonist. Moreover, the PPAR response element at approximately -700 to -800nt upstream of the transcription start site in the *UGT1A8* and -1A9 promoters appeared to be at least

partly responsible for the activity of PPAR when liganded by rosiglitazone. We did not ultimately confirm whether this element was involved in induction of the promoters by genistein.

Much of our analytical strategy was based on previous reports of the regulation of UGT1A9 by PPAR factors that had been determined in non-intestinal contexts including COS7 cells and hepatocytes. However, we found that overexpression of PPAR factors (even with bicistronic overexpression of RXR) did not mediate induction of the *UGT1A8* and *-1A9* promoters in Caco-2 cells which was in contrast to results reported previously in other cell models. The fact that PPAR $\gamma$  is abundant in differentiated intestinal cells including Caco2 cells might explain why expression of additional PPAR factors had no effect in Caco-2 cells. The excess PPAR protein may not be functional due to limiting amounts of other essential cofactors. In support of this idea, an increase in *UGT1A8* promoter activity was achieved when the PGC-1 coactivator was coexpressed with PPAR $\gamma$ . PPAR activity is complex and influenced by interactions of numerous other factors (Triff *et al.*, 2013, Paziienza *et al.*, 2012, Su *et al.*, 2007). Ultimately it is likely that multiple DNA elements are involved in induction of the *UGT1A8-1A10* promoters by genistein via PPAR factors, and these remain to be defined.

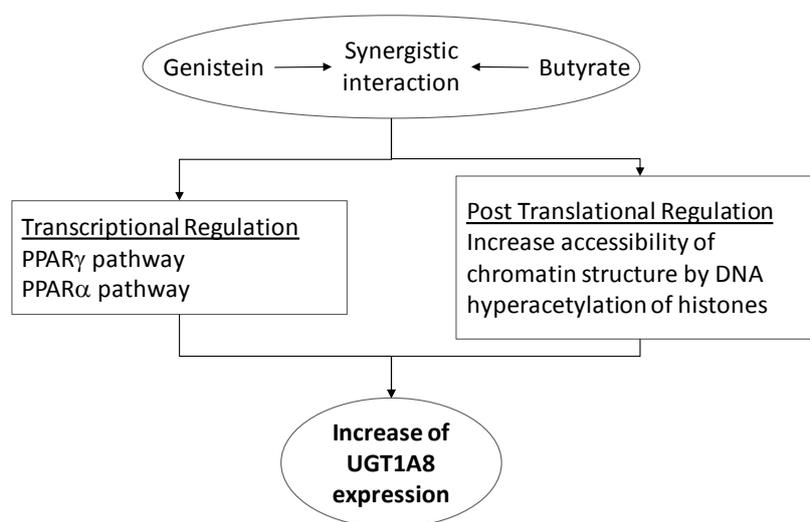
Given that the PPAR factors and HNF4 $\alpha$  bind to very similar DNA sequences and that PPAR $\alpha$  and PPAR $\gamma$  can interact with HNF4 $\alpha$  (Martinez-Jimenez *et al.*, 2010), the possibility of competition between PPAR factors and HNF4 $\alpha$  for binding to the previously defined PPAR response element (at approximately -700 to -800nt) and the HNF4 $\alpha$  binding element (at approximately -45bp) in the *UGT1A9* promoter was explored. This work led to the conclusion that PPAR factors might be able to inhibit the binding or activity of HNF4 $\alpha$  at this promoter. These studies support the idea that the constitutive regulatory mechanisms that define developmental patterning of UGTs (see Chapter 3) overlap with the inducible regulatory mechanisms that allow response to chemical insults.

The activities of PPAR factors and HNF4 $\alpha$  are controlled by post-translational modifications including phosphorylation. In a recent study, ERK1/2 activation-dependent HNF4 $\alpha$  phosphorylation was shown to inhibit HNF4 $\alpha$ -dependent hepatic gene regulation (Vetř *et al.*, 2017). Since PPAR

phosphorylation also involves ERK1/2 activation, this pathway may provide a functional intersection between PPAR and HNF4 $\alpha$  activities. The PPARs are a doubled-edge sword, having the potential to either stimulate or suppress carcinogenesis (Panigrahy *et al.*, 2009). However, expression of PPAR $\gamma$  has been associated with good prognosis in colorectal cancer (Ogino *et al.*, 2009), and this might be viewed as consistent with its apparent ability to induce *UGT1A8-1A10* genes in response to certain chemical ligands.

Our studies also revealed that genistein can increase the abundance of activating chromatin modifications at the *UGT1A8* proximal promoter. Similarly, butyrate is a HDAC inhibitor that promotes histone hyperacetylation at differentiation-associated genes in intestinal cells (Delage and Dashwood, 2008). Thus, both genistein and butyrate may act via epigenetic mechanisms, although the effect of genistein on histone modification is modest and it is unlikely that this is the main mechanism by which it induces *UGT1A8-1A10* gene expression. Figure 4.24 presents a schematic summary of the likely modes of action of genistein and butyrate in regulating *UGT1A8* gene expression, based on results obtained in this study.

Understanding how genistein intake, from supplements or foods (e.g. soy), might impact on intestinal UGT activity could lead to strategies to optimize its beneficial activities in promoting intestinal-health and cancer protection. Moreover, it may give more insight into interactions that could occur between genistein and drugs that are metabolized by UGT1A8.



**Figure 4.24. Schematic representation of the mode of action of genistein and butyrate in regulating UGT1A8.**

Genistein, butyrate, and their synergistic interaction were investigated in this study and it was revealed that their mode of action in regulating UGT1A8 expression. This was mediated transcriptionally at least in part by PPAR $\gamma$  and  $-\alpha$ , as well as via histone modification that creates more accessible chromatin facilitating binding of transcriptional activators.

This study has also provided more insight into how by-products of microbiota activity (e.g butyrate) may regulate intestinal UGTs. Research focusing on the effects of intestinal microbiota on gene regulation has been emerging for the last two decades; however, data specifically on *UGT* gene regulation has been minimal. Further investigation of the interaction of the microbiome with intestinal xenobiotic metabolism will require a more physiologically native model such as transgenic mice or intestinal organoids. In Chapter 5 we begin to explore intestinal organoid culture from UGT1A8 transgenic mice as a system to better understand intestinal UGT function and regulation.

Overall, based on results in this Chapter, we conclude that flavonoids and butyrate produced by microbiota could protect the gut from toxic substances and carcinogens by induction of intestinal UGTs, particularly UGT1A8. This is likely to be an important mechanism by which dietary consumption of flavonoid-containing and high fibre foods mediate their anti-cancer effects.

With the increasing interest in modulating the microbiome to control not just intestinal, but systemic health, the findings of this Chapter have potential for translational benefit. Modulation of the microbiome may be achieved by giving probiotics, which are specific bacterial

strains/combinations, delivered orally or otherwise, or using prebiotics that promote growth of 'healthy' gut bacteria and include resistant starch which is fermented in colon to produce butyrate. Moreover, there is growing interest in postbiotics, which are secondary metabolites produced by microbiota, such as short chain fatty acids, that might be provided exogenously as a drug/supplement/functional food. UGT enzymes may metabolize a number of dietary compounds that regulate microbiota, as well as many of their secondary metabolites. As interest increases in modulating both the intestinal biome and metabolome via diet- or drug-based approaches, it is clear that a better understanding of intestinal UGT regulation and function will be required. It is very likely that UGTs will become valuable as biomarkers or targets for personalized monitoring and/or modulation of intestinal health.

# CHAPTER 5. EXPLORATION OF ENTEROID AND ORGANOID CULTURE FROM UGT1A8/ LUCIFERASE-TRANSGENIC MICE

## 5.1 Introduction

The human intestinal tract is one of the main sites of xenobiotic metabolism and detoxification, and this is largely mediated by the enzymes encoded by the *UGT1A8-1A10* gene cluster with the UGT1A8 isoform playing a prominent role. To date, the Caco-2 human colon cancer cell line has been our model of choice to study UGT1A8 gene regulation and function. Analysis of literature suggests that this model is also widely used by other groups studying intestinal drug and xenobiotic metabolism. To provide a greater understanding of the physiological and pharmacological roles of UGT1A8 *in vivo*, studies would ideally be performed in a model that more closely represents the human intestine, both structurally and functionally. Since the complexity of the intestinal system has been difficult to model *in vitro* using cell cultures, studies have used animal models such as mice. The Human *UGT* genes are quite poorly conserved in non-primate species in terms of sequence, function, and regulation, thus it is not possible to simply study mouse *Ugt* homologues of human *UGT* genes. Instead, transgenic mice carrying human UGT gene loci have been developed (often called 'humanised' UGT mice). In our laboratory we generated a 'humanized' UGT transgenic mouse (UGT1A8/Luc)-that carries a 13-kb segment of the human *UGT1A8* promoter region linked to a luciferase reporter. The expression of the *UGT1A8* promoter in *UGT1A8/Luc*-mice has been shown by bioluminescence in the digestive tract, as described in Chapter 1, section 1.5.3.

At the outset of this project we tentatively planned to use these mice to study the regulation of *UGT1A8* gene in the intact intestine. Two studies were considered: the first was analysis of *UGT1A8* induction by flavonoids and genistein *in vivo* to complement the *in vitro* studies described in Chapter 4; the second was analysis of the constitutive regulation of UGT1A8 by HNF4  $\alpha$  using the HNF4  $\alpha$  inhibitor BI6015 *in vivo* to complement the *in vitro* studies described in Chapter 3. However, due to time constraints, studies that involved treatments of large numbers of live mice became infeasible. We thus decided to use our UGT1A8/Luc mice as a source of intestinal tissue

to develop protocols for *ex vivo* culture of intestinal structures known as organoids. Such organoid cultures could represent a valuable and experimental manipulable intermediate model between simple 2D cell cultures and intact animals.

Self-organised three-dimensional (3D) organoid culture derived from intestinal stem cells as developed by Sato *et al.*, has excited multiple research disciplines for its capability to bridge the gap between *in vitro* and *in vivo* models (Sato *et al.*, 2009). It offers the prospect of an 'easy to use' physiologically native model for numerous areas of study, such as intestinal biology, development and disease, drug discovery, metabolism, toxicology, gene therapy, and regenerative medicine (Meneses *et al.*, 2016). Intestinal organoids/enteroids are described as intestinal rudiments that reproduce at least partly the complex architecture and functions of the intestine.

Organoids have been used extensively to study intestinal stem cell self-renewal, proliferation and differentiation and are hence a powerful tool to investigate developmental patterning of intestinal gene expression. Development of a UGT1A8/Luc-organoid system from UGT1A8/Luc-transgenic mice would allow us to study the roles of various developmental regulators including CDX2 and HNF4 $\alpha$  in regulation of *UGT1A8* during migration of cells from the crypt to the villus, by manipulating the levels of these factors (refer to Chapter 3). Because the organoid undergoes renewal of the luminal epithelial layer including producing an appropriate mixture of luminal cell types, it is also likely to be useful for modelling the normal absorptive and detoxifying function of the mature intestinal lumen. These mature organoids could be used to examine the effect of dietary flavonoids and the intestinal environmental factor butyrate on *UGT1A8* regulation, as well as on organoid growth (see Chapter 4).

Use of organoid technology in toxicology is still a very new area, and the first application of organoids to the study of *Ugt1* null mouse intestine (*Ugt1* <sup>$\Delta$ EC</sup>) was reported in 2017. These studies assessed the toxicity of camptothecin-11 in organoids lacking *Ugt1* expression in intestinal epithelial cells (Lu *et al.*, 2017). This study by Lu *et al.* gives us confidence that establishing an organoid system from our 'humanised' mouse to study UGT1A8 is feasible and likely to be beneficial.

### 5.1.1 Intestinal stem cells and identification of markers

As discussed previously, intestinal epithelium undergoes continual homeostatic self-renewal during which stem cells residing in the crypt give rise proliferating progenitor cells that begin to differentiate as they migrate up to villi where they mature (Marshman *et al.*, 2002). Structurally, the crypt compartment forms epithelial invaginations that were named the crypts of Lieberkühn after their discoverer, Jonathan Nathanael Lieberkühn (1711–1756) (Clevers, 2013). Two populations of intestinal stem cells are identified at the crypt base: the proliferative crypt base columnar (CBC) stem cells, and the +4 population (Leushacke and Barker, 2014). The CBC cells are the active engines producing proliferating progenitors (transit-amplifying (TA) cells) as well as maintaining the supply of new CBC. These cells are frequently turned over, undergoing stochastic symmetric cell division every 24 hours. The +4 cells are DNA-label-retaining cells (LRCs) located directly above the terminally differentiated Paneth cells (Meneses *et al.*, 2016, Leushacke and Barker, 2014). They are damage-resistant, and their cellular division is very infrequent but they can be induced to produce CBC cells in response to injury or catastrophic loss of the CBC cell pool.

Molecular markers and transgenic lineage-tracing techniques are used to specify CBC and +4 stem cells. The CBC are marked by expression of the Wnt target genes *Lgr5* together with Musashi-1, Prominin1/CD133 and *Smoc2*; whereas markers of +4 cells include *Bmi-1*, *Hopx*, *mTert* and *Lrig1* (Leushacke and Barker, 2014, Meneses *et al.*, 2016, Clevers, 2013). However, the specificity of these markers for CBC and +4 cells is controversial. For example, a subset of quiescent +4 label-retaining cells were found to express the CBC marker *Lgr5*. In pathological conditions such as inflammation, these cells revert to CBC capable of extensive proliferation and can contribute to intestinal tissue repair (Buczacki *et al.*, 2013). *LGR5+* cancer stem cells (*LGR5+* CSC) have also been identified in human colorectal cancer cells, which are responsible for CSC self-renewing capacity and support tumour growth (Shimokawa *et al.*, 2017). Overall, the *Lgr5* gene marks stem cell plasticity in crypt populations, and defines self-renewal and multipotency properties.

### 5.1.2 Organoid culture of intestinal epithelium

Identification of *Lgr5* as an intestinal stem cell marker led to the development of *ex vivo* organoid cultures that can be grown from a single stem cell, or a crypt fragment containing stem cells. When derived from the small intestine, these organoids are termed enteroids, and when derived from the colon they are termed colonoids (Stelzner *et al.*, 2012).

Organoids provide a long-term culture system for primary intestine because they contain self-renewing stem cells. Protocols for organoid culture have been established fairly recently. In 2009, Ootani and colleagues described expansion of small and large intestinal fragments containing epithelial and mesenchymal cells from neonatal mice for several months in culture. This occurred in growth factor-independent conditions, by incorporating an air-liquid interphase with an underlying stromal element that secretes relevant growth factors. The organoids were sphere-like and contained all major intestinal cell types, with showed proliferation and multilineage differentiation (Ootani *et al.*, 2009). In the same year, Sato *et al.* designed a Matrigel based culture system which allowed the generation of epithelial enteroids from an isolated *Lgr5*<sup>+</sup> stem cell population. Their methodology allowed for long term propagation (1.5 years). Their enteroids were also spherical with multiple crypt-like regions containing *Lgr5*<sup>+</sup> cells intermingled with Paneth cells as well as proliferative TA population that gave rise to differentiated cells of all intestinal lineages that ultimately shed into the lumen after approximately 5 days in culture. The architecture of the organoid involves an apical surface facing the lumen and a basolateral domain in contact with the Matrigel and culture media (Sato *et al.*, 2009). No significant changes in cell karyotype occurred over long-term passaging (Sato *et al.*, 2009).

The crypt-villus units generated by Sato and colleagues are self-organising small intestinal epithelial organoid structures, which surprisingly were built in the absence of mesenchymal cells. This indicates an independency from growth support provided by the submucosal niche (Sato *et al.*, 2009). Their protocols for growing the enteroids however, depend absolutely on the presence of essential components of secreted growth factors, including R-spondin1, noggin and epidermal growth factor (EGF), with Matrigel to support the base of the 3-dimensional culture (Sato *et al.*,

2009). Matrigel is a mixture of proteins mimicking extracellular matrix that is derived from Englebreth-Holm-Swarm mouse tumor cells. It is rich in collagen, laminin ( $\alpha 1$  and  $\alpha 2$ ), and enactin and provides basement membrane-like support for stem cells (Hughes *et al.*, 2010, Sasaki *et al.*, 2002).

The critical factor in driving crypt proliferation in enteroid culture is the activation of Wnt signalling. At least four Wnt genes (*Wnt4*, *Wnt6*, *Wnt11* and *Wnt14b*) are specifically expressed in the intestinal crypt, Wnt signalling is also reported to control the differentiation of intestinal secretory cells (Pinto *et al.*, 2003). Addition of the Wnt activator R-spondin1 is essential for viable organoid culture and Barker *et al.* found that R-spondin1 acts as a ligand for Lgr5, a stem cell marker and Wnt target gene (Barker *et al.*, 2007). Crypt formation can be repressed by bone morphogenetic protein (BMP) signalling; to inhibit BMP signalling, noggin is required in the organoid growth factor cocktail (Haramis *et al.*, 2004). Noggin is also important to maintain Lgr5 expression (Sato *et al.*, 2011). EGF addition is required to induce intestinal crypt proliferation and regulate intestinal function (Dignass and Sturm, 2001). The protocol by Sato *et al.* also suggests addition of Rho-associated protein kinase (ROCK) inhibitor to prevent anoikis when single stem cells are passaged (Watanabe *et al.*, 2007). ROCK inhibitor also acts as a synthetic Notch ligand, which is necessary for stem cell maintenance (Sato *et al.*, 2009). A summary of growth factors and supplements used in enteroid culture is presented in Table 5.1.

**Table 5.1. Growth factors and reagents for enteroid culture.**

Reagent	Function
<b><u>Extracellular matrix</u></b>	
<b>Matrigel</b>	Structural support for 3-dimensional growth of intestinal epithelial stem cells (IESCs) through the laminin ( $\alpha 1$ and $\alpha 2$ chains) and collagen enriched extracellular matrix (Sasaki <i>et al.</i> , 2002, Hughes <i>et al.</i> , 2010).
<b><u>Growth factor</u></b>	
<b>R-spondin 1 (Rspo1)</b>	Inducer of intestinal crypt proliferation through $\beta$ -catenin activation and stabilisation (Kim <i>et al.</i> , 2005).

<b>Noggin 1</b>	Inducer for expansion of crypt numbers (Haramis <i>et al.</i> , 2004) and maintenance of Lgr5 expression through BMP signal inhibition (Sato <i>et al.</i> , 2011).
<b>Epidermal growth factor (EGF)</b>	Potent stimulator of intestinal cell proliferation and promotes intestinal regulatory function (e.g. enzyme production, transport system, and self-repair capability) (Dignass and Sturm, 2001).
<b><u>Media supplement</u></b>	
<b>B27</b>	Supplement for serum free media, to help long term-survival of culture growth (Brewer and Cotman, 1989).
<b>N2</b>	Supplement for serum free media, to help initial commitment, differentiation and survival of post-mitotic cells in culture (Johe <i>et al.</i> , 1996).
<b>HEPES</b>	Maintains physiological pH of organoid growth media.
<b>Y-27632</b>	A ROCK inhibitor, preventing dissociation-induced apoptosis (anoikis) of IESCs culture and the loss of Notch signalling (Watanabe <i>et al.</i> , 2007, Sato <i>et al.</i> , 2009).

**Table adapted from “Identification, Isolation and Culture of Intestinal Epithelial Stem Cells from Murine Intestine”, by A.D Gracz, B.J. Puthoff, and S.T. Magness, *Somatic Stem Cells: Methods and Protocols*, p.89. With permission from © Springer.**

Although organoids can be formed from single stem cells and maintained long term with appropriate growth factor support, (Pastuła *et al.*, 2016) recently reported that organoid co-culture with other cell types from intestine (such as neurons, myofibroblasts and collagen) improves organoid growth and survival. Their findings confirm the importance of the mesenchymal element of the lamina propria in 3D organoid culture (Lahar *et al.*, 2011, Lei *et al.*, 2014). Understanding the mesenchymal-epithelial interaction in the intestinal stem cell niche may also enable identification of the biomolecular mechanism of early tumorigenesis (Pastuła *et al.*, 2016).

Overall, protocols for organoid preparation have become both more standardized (including the development of commercial culture kits) but also more flexible recognizing that different types of co-culture have different requirements for additional of trophic factors.

### 5.1.3 Aims

Given the time constraints of this project and the labour-intensive nature of organoid culture, the goals of this part of the project were quite modest. We sought to establish a mouse organoid culture method in our laboratory and generate and maintain sufficient numbers of organoids to develop protocols for transfection or transduction, and for analysis of gene expression (RNA or protein analysis). Achieving these goals would allow us to manipulate known regulators of intestinal *UGT* expression and assess the effects of these manipulations on the *UGT1A8/Luc* reporter gene activity. Ultimately, we also hoped to transfer the knowledge gained here to human organoid cultures so that a better understanding of the regulation of the main intestinal UGTs in a physiologically-relevant yet readily manipulable human system could be achieved. This would provide a critical foundation for understanding inter-individual variation in drug metabolism and intestinal detoxification capacity.

Studies in this chapter aim to:

1. Generate an enteroid/organoid model from the *UGT1A8/Luc*-mice.
2. Observe the enteroid/organoid phenotype upon successful culture.
3. Explore methods for DNA delivery into organoids and for detection of gene expression by immunostaining.
4. Assess the expression of the *UGT1A8/Luc* transgene and other intestinal specific genes in organoids.

## 5.2 Methods

### 5.2.1 Development and maintenance of the *UGT1A8/Luc* transgenic mice

The *UGT1A8/Luc*-transgenic mice were developed in 2007 by Prof. Peter Mackenzie, Dr. Dong Gui Hu, and Dr Dione Gardner-Stephen at Flinders University. The transgenic mice carry the 13-kb promoter region of the human *UGT1A8* gene linked to a luciferase reporter gene. Bioluminescence imaging to assess firefly luciferase expression was performed by Ms. Joanna Gillis, Flinders University. The *UGT1A8/Luc*-transgenic mice (Ethics approval AWC#825-13), were maintain at

Flinders University until 2011 and then transferred to University of South Australia (UniSA) Animal Facility (Ethics approval BC61-12), under the supervision of Dr. Michael Ward (School of Pharmacy and Medical Sciences, University of South Australia). All mice handling and procedures including euthanasia were approved by the Animal Welfare Committee (AWC) and the Institutional Biosafety Committee (IBC) of UniSA.

## **5.2.2 Preparation of Matrigel, culture plates, and media**

### **5.2.2.1 Matrigel preparation**

Matrigel must be prepared prior to crypt isolation. Matrigel is in liquid form at 2-8°C and solidifies when it warms above this temperature. Aliquots of Matrigel were frozen at -20°C; when required for organoid culture the tubes were thawed on ice and kept on ice until use. Tubes for resuspending the crypt pellet with Matrigel were also pre-chilled on ice to avoid premature gelling.

### **5.2.2.2 Preparation of enteroid/organoid culture media**

Basal media for organoid culture was prepared using Advanced-DMEM/F12 containing 2 mM of GlutaMax, 10 mM of HEPES buffer, 0.5U/mL of penicillin/streptomycin, 1 x of N2 supplement and 1 x of B-27 supplement. To prepare organoid growth media, 50 ng/mL of Mouse Recombinant EGF, 100 µg/mL of Mouse Recombinant Noggin and 1 µg/mL of Human Recombinant Rspodin1 were added to the basal media. ROCK-inhibitor Y-27632 was added to the growth media at 10 µM for crypt plating and enteroid/organoid passaging. Supplemented media was stable for 2 weeks at 4°C.

A commercial organoid media, IntestiCult™ Organoid Growth Medium (Mouse), was also used in this project. The IntestiCult™ media was prepared as instructed in the product manual.

### **5.2.2.3 Preparation of culture plates**

Prior to plating crypts in Matrigel culture, culture plates were pre-incubated at 37°C to allow Matrigel droplets to gel rapidly forming a dome in the well.

### **5.2.2.4 Preparation of enteroid/organoid freezing media**

Enteroid/organoid freezing media contained 10% DMSO and 10% of Foetal Bovine Serum (FBS) and 80% advanced-DMEM/F12.

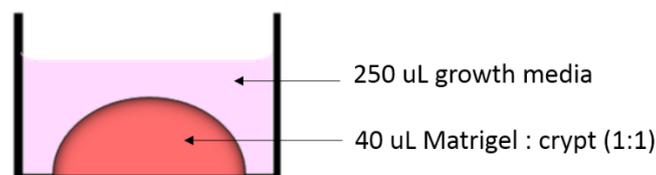
### 5.2.3 Isolation of mouse intestinal crypts

Adult *UGT1A8/Luc*-mice (16-24 weeks) were euthanised by carbon dioxide asphyxiation and the chest cavity opened to access the intestinal tract. Ten cm of small intestine was harvested and placed in cold PBS. Membranes, blood vessels and fat were removed completely from the intestinal exterior. The intestine was flushed with cold PBS until it rinsed clean and then cut longitudinally. The open intestinal sheets (intestinal lumen facing up) were gently washed 3 times in a petri dish with cold PBS. Clean intestine was then cut into 2 mm pieces and put into a tube containing 10 mL cold PBS. Using 10 mL pipettes pre-wetted with FBS (to prevent sticking), intestinal sections were washed by gently pipetting up and down 3x to dislodge single cells and all other debris. The pieces were allowed to settle by gravity and the supernatant containing debris of the villus fraction was then removed as much as possible. Another 10 mL of cold PBS was added and the washing process was repeated 15 to 20 times or until the PBS was clear. The intestinal pieces were resuspended in 12 mL of ice-chilled enzyme-free dissociation buffer (Gibco®) and incubated for 20 minutes at room temperature on a rocking platform in order to 'loosen' the crypt from intestinal tissue. Supernatant of the first incubation was removed using a pipette by previously settling intestinal pieces by gravity for 30 seconds. To dislodge the crypts from intestinal pieces, 10 mL of 0.1% BSA in PBS was added, and the intestinal pieces pipetted up and down three times. The supernatant containing crypts was collected and labelled as fraction one. The steps to dislodge the crypts by pipetting in 0.1% BSA/PBS were repeated five times until five fractions of crypt solution were generated. Each fraction was observed under the microscope for the presence of crypts. The third and fourth fractions typically contained the most crypts.

To collect the crypts, each fraction was spun at 290 x g for 5 minutes at 4°C., the supernatant was decanted and 1 mL of 0.1% BSA/PBS was added to resuspend the crypt pellet, and then transferred into 1.5 mL tubes and labelled. The crypts were centrifuged again at 200 x g for 3 minutes at 4°C, the supernatant was removed and the pellet containing crypts was resuspended in 100 µL of room temperature complete growth media.

### 5.2.4 Enteroid/organoid culture from isolated crypts

Isolated crypt suspensions in growth media were mixed with Matrigel in a 1:1 ratio by gently pipetting up and down ten times to fully resuspend the mixture. Care was taken to avoid introducing bubbles. Forty  $\mu\text{L}$  of Matrigel: crypt mixture was applied to each well of a prewarmed 48-well plate. The droplet of Matrigel: crypt mixture was plated into dome formation in the centre of the well, as illustrated in Figure 5.1, followed by incubation at  $37^\circ\text{C}$  for 15-30 minutes to allow the Matrigel to completely polymerise. An aliquot of 250  $\mu\text{L}$  of complete media containing 10  $\mu\text{M}$  Y27623 was overlaid on the Matrigel dome in each well. The crypts were cultured in a  $\text{CO}_2$  incubator ( $37^\circ\text{C}$ , 5%  $\text{CO}_2$ ) and media was replaced every 4 days with growth media without Y27623. Passaging of the organoid was performed after 7-10 days of culture to avoid organoid over-growth and excessive accumulation of shed cell debris in the lumen that may lead to organoid death.



**Figure 5.1. Illustration of a Matrigel dome of enteroid/organoid culture, in a single well of 48 well-plate.**

Forty  $\mu\text{L}$  of Matrigel: crypt was plated and polymerised at  $37^\circ\text{C}$ , and 250  $\mu\text{L}$  growth media added after the Matrigel is set. The dome formation of Matrigel in the centre of the well allows pipette tips to reach the bottom of the well to remove and replace the media without disturbing the dome.

### 5.2.5 Passaging of the organoid culture

The organoid culture was passaged after 7-10 days of culture when they turned dark in the center indicating accumulation of shed cells in the lumen. Passaging of organoids was performed by removing the growth media and adding 1 mL of enzyme-free dissociation buffer, then incubating for 1 minute at room temperature followed by pipetting (using FBS-pre-wetted tips) up and down to break up the Matrigel dome and organoids. The same pipette tip was used to transfer the organoid suspension into a 10 mL canonical tube, to which another 1 mL of dissociation buffer was added. The tube was incubated at room temperature for 15 minutes on a rocking platform, followed by

centrifugation at 290 x g, 4°C, for 5 minutes. The supernatant was discarded by decantation, and the cell pellet was washed with basal medium and centrifuged at 290 x g for 5 minutes at 4°C to collect the pellet. At this point, the pellet could be frozen in freezing media (refer to section 5.2.2.4) if desired. Following dissociation of the organoids into single cells, the washed pellet (from a single well organoid culture) was resuspended in 80 µL of room temperature growth medium, and 80 µL of pre-thawed Matrigel was added to the suspension. The total 160 µL mixture was distributed across 4 wells by applying a 40 µL droplet of the mixture into a dome in the centre of each well. After the Matrigel dome was set by incubation at 37°C, 250 µL growth media containing Y27623 was added to each well and the passaged organoid was cultured in the CO<sub>2</sub> incubator (37°C, 5% CO<sub>2</sub>).

### **5.2.6 Freezing organoids**

Organoids could be frozen 3 days post-passaging. Organoids were dissociated into single cells as described above (section 5.2.5). The pellet of single cells was resuspended in freezing media and transferred to a labelled cryovial tube. One mL of freezing media was used for every 3 wells of organoids collected. The organoid stock must be frozen slowly to -80°C, so the cryotube was wrapped in several layers of paper towel before being stored at -80°C for 24 hours, and then transferred to liquid nitrogen storage.

### **5.2.7 Enteroid/organoid transfection**

#### **5.2.7.1 Transfection of plasmid DNA**

Transfection of enteroids/organoids was tested using Lipofectamine® 2000, with the experiments performed in accordance with the manufacturer's standard protocol. For each well of enteroid/organoid culture in Matrigel, a complex of 1 µg pMM045-GFP expression plasmid combined with 5 µL of Lipofectamine was applied. To prepare the transfection complex, 1 µg of pMM045-GFP plasmid and 5 µL of Lipofectamine® 2000 reagent were each diluted with 50 µL of advanced-DMEM/F12 serum free media in separate tubes and incubated at room temperature for 5 minutes. The DNA and Lipofectamine mixtures were combined and incubated for a further 20 minutes. Transfection of organoids was performed three days post passaging. The

enteroid/organoid media was replaced with fresh growth media and Lipofectamine/DNA complex was added to the well. Three days after transfection, enteroids/organoids were observed using an EVOS® FL microscope.

### **5.2.7.2 siRNA transfection**

siRNA transfection in organoids were tested using Lipofectamine® 2000. Transfection was either performed in dissociated organoid single cells, or in intact organoids in Matrigel on the third day after organoid passaging. For both methods, a complex of Lipofectamine® 2000 with siRNA was prepared according to the manufacturers' instructions. In detail, 5 pmol of FAM labelled siRNA and 5 µL Lipofectamine® 2000 were each diluted in 50 µL of advanced-DMEM/F12 serum free media and incubated for 5 minutes at room temperature. Both mixes were combined and incubated for a further 20 minutes before being added to the cells/organoids.

Protocols for single cell organoid transfection was adapted from (Schwank *et al.*, 2013). After organoid dissociation, single cells from 1 well of organoid culture were resuspended in 250 µL growth medium containing Y27632, and plated in 1 well of a 48 well-plate, and the Lipofectamine® 2000 + FAM labelled siRNA complex was added to the cells. The plate was centrifuged at 600 x g for 1 hour at 32°C and then incubated for 4 hours at 37°C. Cells were transferred to 1.5 mL tubes and spun at 1000 x g, 5 min. Cell pellets were resuspended in 40 µL of the Matrigel : growth media (1:1) mixture and plated as described in the organoid passaging protocol.

siRNA transfection of growing organoids in Matrigel was performed at day 3 post-passaging as follows. The prepared complex of Lipofectamine® 2000 + FAM labelled siRNA in fresh growth media was added to the well containing organoid culture in Matrigel. Forty-eight hours post-transfection, organoids were observed using an EVOS® FL microscope.

## **5.2.8 Organoid transduction**

### **5.2.8.1 Lentiviral packaging and production**

The GFP/pTiger lentivirus plasmid was packaged in HEK293T cells as described in Chapter 4 to produce the lentiviral particles for transduction of organoids. Virus was frozen at -80 °C prior to use.

### **5.2.8.2 Transduction**

Organoid transduction was performed by spinfection three days after organoid passaging. Frozen virus suspension was thawed to room temperature and spun at 1000 rpm for 5 minutes to remove any packaging cells in the suspension. The supernatant was transferred to a new tube and Polybrene was added to a final concentration of 4 µg/mL. To infect the organoid with the virus, organoid growth media was removed from the organoid culture, and virus suspension containing polybrene was overlaid on top of the Matrigel (organoid culture), and the plate was spun at 3500 rpm for 2 hours at 32°C (spinfection). Following the spinfection, media containing the virus was removed and replaced with organoid growth medium. Organoids were observed using an EVOS® FL microscope three days after transduction.

### **5.2.9 Mouse tissue and enteroid/organoid RNA extraction, reverse transcription and qRT-PCR**

#### **5.2.9.1 RNA extraction from mouse tissue**

Mouse small intestine and colon tissue that was not used for organoid production was stored in RNALater solution at 4 °C. Subsequently, total RNA was extracted from the tissue using TRIzol® Reagent (Thermo Fisher Scientific). Two hundred µL of Trizol was added and a homogenisation pestle was used to grind the tissues; another 800 µL of Trizol was then added and incubated for 5 minutes room temperature before addition of 200 µL chloroform. RNA was extracted and precipitated as described in Chapter 2. The RNA pellet was air-dried and resuspended in 100 µL RNase-free H<sub>2</sub>O and heated at 65°C in a dry heating block for 5 minutes to aid RNA solubilisation. RNA concentration was measured using NanoDrop™ microvolume spectrophotometers as described in Chapter 2 and RNA samples were kept at -80°C until required.

#### **5.2.9.2 Enteroid/organoid RNA extraction**

Extraction of total RNA from enteroids/organoids was performed using TRIzol® Reagent (Thermo Fisher Scientific) with glycogen addition in accordance to the manufacturer's manual. Growth media was removed from organoid cultures and 1x cold PBS was applied on top of the Matrigel, the enteroids/organoids were released from the Matrigel by pipetting using 1 mL pipette tips. This was followed by centrifugation to collect the pellet. One mL of TRIzol® reagent was then added to

each tube of enteroid/organoid and incubated at room temperature for 10 minutes on a rocking platform. The TRIzol® suspension containing cell lysate was transferred to a 1.5 mL tube and glycogen was added to a final concentration of 250 µg/mL (12.5 µL of 20 mg/mL glycogen stock). The subsequent RNA extraction and precipitation steps were as described in Chapter 2. Each RNA pellet was resuspended in 20 µL RNase-free H<sub>2</sub>O and heated at 65°C for 5 minutes to aid RNA solubilisation. RNA concentration was measured using NanoDrop™ microvolume spectrophotometers as described in Chapter 2 and RNA samples were kept at -80°C until required.

### **5.2.9.3 Reverse-transcription (cDNA synthesis)**

cDNA synthesis was performed using RNA from both tissues and enteroids/organoids as described in Chapter 2.

### **5.2.9.4 Quantitative real-time PCR**

Quantitative real-time PCR analysis of gene expression levels in cDNA samples was performed as described in Chapter 2. The  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001) was applied for quantification. The Student's t-test used to assess the statistical significance of the results.

### **5.2.10 Whole mount immunofluorescence (IF) of organoids**

Whole organoid immunostaining methods were tested using growing organoid cultures in Matrigel. On the day of staining, the media was removed and the Matrigel dome washed with 1x PBS, followed by fixation with 4% paraformaldehyde in PBS for 15 minutes at room temperature. Organoids were washed after fixation with 1x PBS containing 0.1 M glycine, and gentle rocking of the plate for 10 minutes at room temperature, three times. Further washing with 1x IF wash buffer solution (PBS containing 0.2% Triton-X and 0.05% Tween) was performed three times for 10 minutes. After washing, the organoids were incubated in blocking buffer containing 3% horse serum in IF wash buffer for 1.5 hours on a rocking platform at room temperature. The blocking buffer was removed and replaced with primary antibody diluted in blocking buffer, at a final concentration of 4 µg/mL, for 4 hours gently rocking or overnight on the bench. Primary antibody solution was aspirated and organoids were washed again with 1x IF buffer three times, with 10 minutes of gentle rocking at room temperature for each wash. After the third wash, the IF wash

solution was removed from the organoid culture and replaced with secondary antibody solution diluted in blocking buffer at a final concentration of 10 µg/mL, incubated for 1 hour at room temperature, with gentle rocking. After removal of the secondary antibody solution, the organoid culture was washed with 1x IF buffer twice for 20 minutes each, again with gentle rocking. A final wash with 1x PBS for 10 minutes was performed before imaging. Images were taken using an EVOS®FL microscope.

### 5.2.11 Oligonucleotides

**Table 5.2. Oligonucleotides used for qRT-PCR analysis of organoids and mouse tissue.**

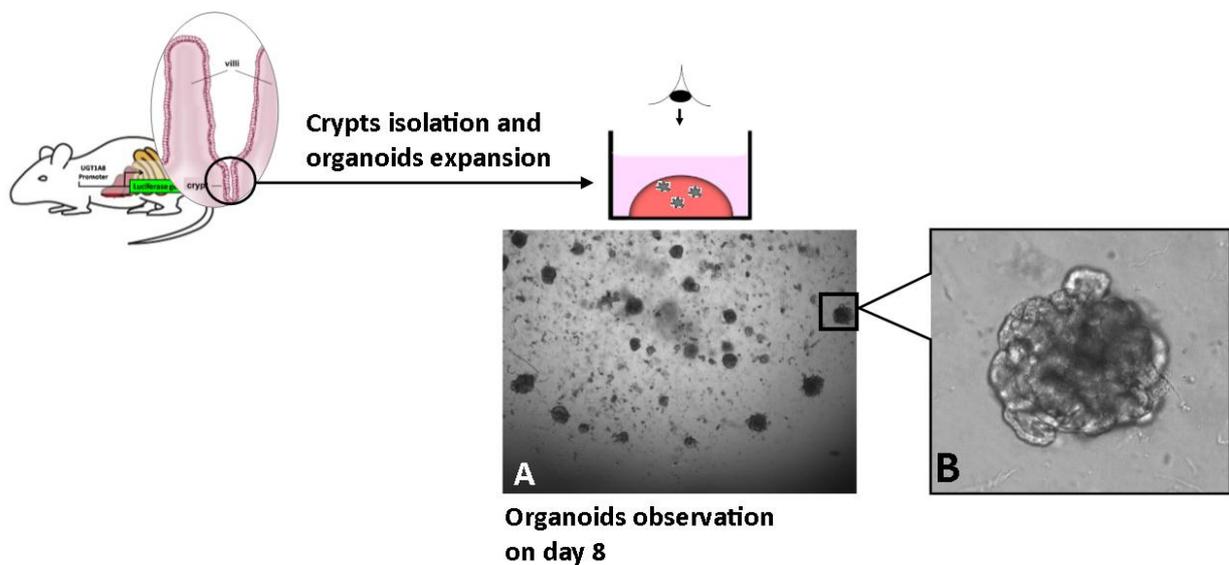
Name	Sequence (5' → 3')
qRT-PCR	
mCdx2 F	CACCATCAGGAGGAAAAGTGA
mCdx2 R	CTGCGGTTCTGAAACCAAAT
mHNF4 $\alpha$ F	CAGCAATGGACAGATGTGTGA
mHNF4 $\alpha$ R	TGGTGATGGCTGTGGAGTC
Ugt1A7c F	TGTGATGCCCAATGTGATCT
Ugt1A7c R	CAGAGGCGTTGACATAGGC
mPPAR $\gamma$ F	AAGATGGAGTCCTCATCTCAGA
mPPAR $\gamma$ R	ACTCAAACCTGGGCTCCATAAAG
Firefly Luciferase F	TACGATTTTGTGCCAGAGTCC
Firefly Luciferase R	ATCCGGAATGATTTGATTGCC

## 5.3 Results and discussion

### 5.3.1 Generation of enteroids and colonoids from isolated crypts

To generate *UGT1A8*/Luc-mouse enteroids using a protocol adapted from (Sato *et al.*, 2009). The small intestinal crypts obtained from *UGT1A8*/Luc-mice were cultured in Matrigel as per the protocol in section 5.2.4. Multiple crypts were plated in one drop of Matrigel for each well, and overlaid with complete media containing R-spondin, noggin, EFG and ROCK inhibitor Y-27632. Formation of crypt derived enteroids was determined to be successful following visual examination. One day after the crypts were plated, they closed at the broken end and displayed a cup-like

morphology; in the next few days they formed closed spherical enteroids and began to form buds, indicating that Wnt signalling was activated. We replaced the media with complete media without Y-27632 every 2-3 days. On days 5-10, the enteroids were well developed and ready for either harvesting or passaging. Figure 5.2 shows enteroid formation in Matrigel culture from isolated crypts of *UGT1A8/Luc*-mice. Enteroid morphology on day 8 was recorded, with a darker enteroid lumen observed, indicating the lumen had filled with shed enterocytes, and these enteroids were starting to distort. Passaging of the enteroids was performed approximately every 10 days in a 1 : 5 ratio. Enteroids were dissociated mechanically by pipetting and re-plated. We found that addition of Y-27632 is essential during passaging, as single dissociated enteroid cells failed to be maintained in the absence of ROCK inhibitor Y-27632. This confirmed the required media conditions described in (Sato *et al.*, 2009).



**Figure 5.2. Successful enteroid formation from crypts of *UGT1A8/Luc*-mice.**

Small intestinal crypts were isolated and suspended in Matrigel, and plated with complete enteroid medium, supplemented with R-spondin 1, epidermal growth factor (EGF), and Noggin. A. A representative field of view from above a well of enteroid culture at day 8 at magnification of 2x. The number of enteroids were not counted. B. Higher magnification of an enteroid at 10x. The budding formations demonstrated continuous generation of crypt-villus structures reminiscent of mammalian intestinal epithelium. Images were taken using an EVOS® FL microscope in bright field lighting.

Encouraged by the success of enteroid culture, we attempted generation of *UGT1A8/Luc*-mouse colon organoids by applying the exact same culture conditions used for the mouse small intestinal

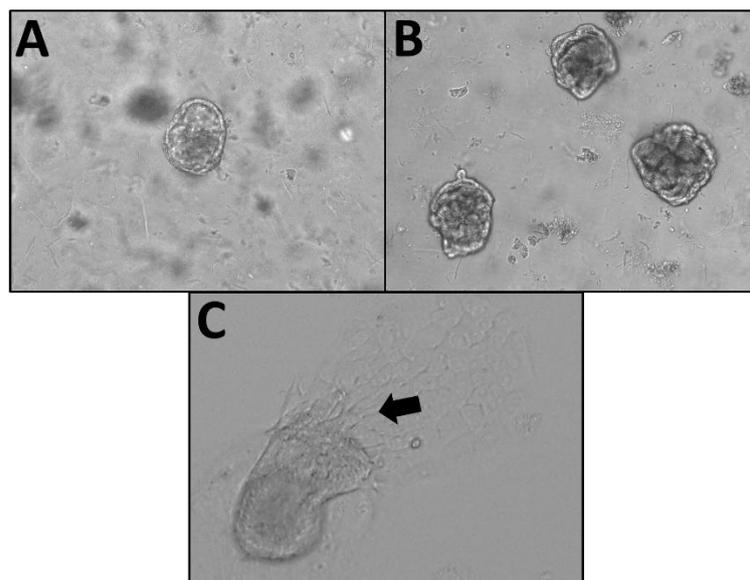
organoids. We isolated colon crypts that were sufficient for plating in 1 well of Matrigel culture. Interestingly, 2 days after culture initiation, while the small intestinal crypt showed growth and changes into sphere-like structures, colon crypts maintained their original crypt morphology. Shortly after day 3, despite media being refreshed, these colon crypts disintegrated and failed to survive. This failed attempt to culture colon crypts may indicate that Wnt signalling was not sufficiently activated to initiate the growth of the colon crypt, confirming that stem cell survival in organoid systems is determined by adequate expression of Wnt signals (Sato *et al.*, 2011). Intestinal crypts on the other hand, constitutively express Wnt in their buds (organoid crypt structure), with Wnt being actively produced by Paneth cells within (Sato *et al.*, 2011). Attempting *UGT1A8/Luc*-colon organoid formation using Wnt-supplemented culture media was not achievable during the time frame of this project.

Despite the unsuccessful generation of *UGT1A8/Luc*-mouse colonoid cultures, we have succeeded in maintaining the enteroid cultures for over 4 months before freezing them for long term storage stocks.

### **5.3.2 Observation of structures during organoid culture**

The National Institute of Health (NIH) Intestinal Stem Cell Consortium has proposed a standardised nomenclature for intestinal *in vitro* culture, which defines the structures observed at various stages of culture (Stelzner *et al.*, 2012). In our preparation of *UGT1A8/Luc*-mouse intestinal organoid culture, we observed three structures of these defined structures (Figure 5.3). The first structure was the “enterosphere” (Figure 5.3A), which was observed 1 day after crypt plating or after passaging the mature organoids. The enterosphere is a sphere-like structure with distinct borders derived from either small intestinal crypts or single stem cells. Based on our observation, especially after organoids were dissociated for passaging, there were variations in enterosphere sizes and the length of time that enterosphere persisted in the culture. These variations may relate to the extent of Wnt-related signalling in the structure (Lahar *et al.*, 2011) and the amount of growth factors in the media. The second structure observed was the “enteroid” (Figure 5.3B), a multilobulated structure produced by continuing enterosphere growth. This 3-

dimensional “mini-intestine” structure is defined as an intestinal epithelial crypt-villus unit independent from the cellular elements of the mesenchymal niche (Zachos *et al.*, 2016, Stelzner *et al.*, 2012). The third structure observed in the culture was “reconstituted intestinal organoid” cluster that includes small intestinal epithelial structures adherent to intestinal fibroblasts in a co-culture system (Pastuła *et al.*, 2016) (Figure 5.3C). The “reconstituted intestinal organoid” structure was found unintentionally during the second attempt to culture the *UGT1A8*/Luc-organoids. Unfortunately, we initially defined this structure as an organoid culture “contaminated” with mesenchymal cells and did not take care to characterize it properly by identifying the mesenchymal cell identity (i.e. fibroblast or intestinal subepithelial myofibroblast (ISEMF)). Based on the appearance, we assumed that both fibroblast and ISEMF might both be present in the organoid co-culture (Figure 5.3C). The presence of fibroblasts/ISEMFs affected the morphological structure of the epithelial component, which is discussed in the following section (5.3.3).



**Figure 5.3. Representative images of structures observed during *UGT1A8*/Luc-mice organoid expansion.**

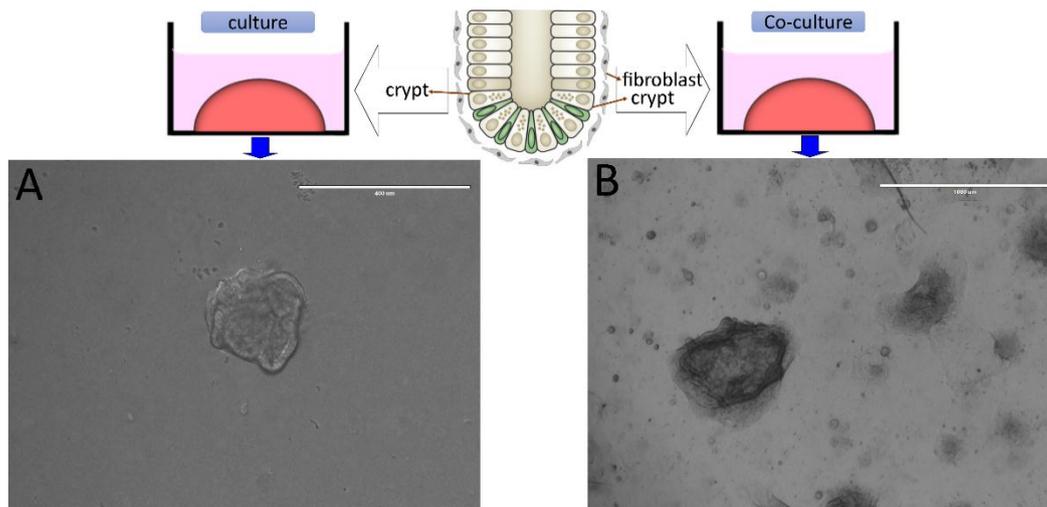
Terms used for structure identification are based on the nomenclature for intestinal *in vitro* culture (Stelzner *et al.*, 2012). During intestinal epithelial expansion, three structures were identified in this study. A. Enterosphere, a spherical structure, formation started within 3-24 hours at the beginning of crypt culture. B. Enteroid, a mature organoid formation that was observed after 3-7 days of culture, and C. Reconstituted intestinal organoid, a multicellular structure formed from intestinal epithelial and mesenchymal cell elements. The culture consists of intestinal epithelial and intestinal fibroblast/ISEMFs. The flat ISEMFs are indicated by the arrow. All images were taken using an EVOS® FL microscope in bright field lighting at 10x magnification.

### 5.3.3 Organoid formation in the presence or absence of intestinal fibroblasts/ISEMFs

It is well described that components of the niche at the base of intestinal crypts support the regulation and maintenance of intestinal epithelial stem cells (IESCs), primarily mediated by Wnt signalling (Yeung *et al.*, 2011). On one of our attempts to culture the *UGT1A8*/Luc-mouse intestinal epithelial cells, we unintentionally co-isolated fibroblasts/ISEMFs during crypt isolation, and therefore co-cultured crypts with fibroblasts/ISEMFs population. We were subsequently able to compare two conditions, which were epithelial cultures in the presence and absence of fibroblasts/ISEMFs. The same Matrigel based culture technique and media were used to maintain both conditions. We demonstrated that the presence of fibroblasts/ISEMFs, encouraged formation of larger of epithelial structures (Figure 5.4). Unlike the epithelial-only enteroids that generated a closed spherical crypt-villus structure with distinct crypt buds, the 'reconstituted organoids' that included fibroblasts/ISEMFs generated large 'open' epithelial structure that had long invaginated walls. This suggests the capacity of fibroblasts/ISEMFs to provide a structural frame for organoids which is different to what the Matrigel provides. The extracellular matrix (ECM) and collagen that were synthesised by the fibroblasts may contribute to shaping the organoid phenotype.

Although the budding was hardly seen in 'reconstituted organoids', they survived several rounds of passaging, the epithelial organoid component grew much larger when compared to enteroids. This suggests that the stemness was maintained and that they contained many proliferating cells; moreover, it is likely that the fibroblasts/ISEMFs produce trophic factors that support epithelial growth.

As mentioned above, these 'reconstituted organoids' were grown unintentionally, and we did not apply any specific treatment to these cultures, therefore they were also grown in the Matrigel base. However, we noticed that the Matrigel was degraded in this culture condition over time. This may be caused by matrix metalloproteinase (MMPs) production by the fibroblasts/ISEMFs (Pender *et al.*, 1997).



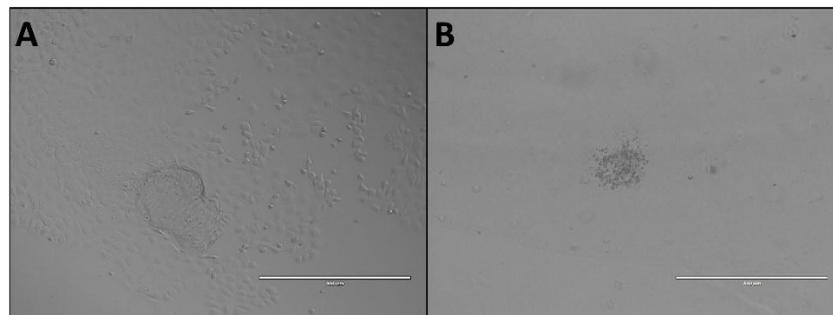
**Figure 5.4. Comparison of organoid culture in the presence and absence of fibroblasts/ISEMFs.**

Images are representative morphology of organoid culture from isolated *UGT1A8/Luc*-mice small intestinal crypts with and without fibroblasts/ISEMFs, in a complete organoid medium, within the same range of culture time. A. represents organoid culture in the absence of fibroblasts which shows distinct sphere/budding morphology. Scale bars are 400  $\mu\text{m}$ , at magnification of 10x. B. shows the growth of 'reconstituted organoid' in the presence of fibroblasts/ISEMFs which support growth of larger, open organoids in various irregular shapes. Scale bars are 1000  $\mu\text{m}$ , at magnification of 4x. All images were taken using an EVOS® FL microscope in bright field lighting.

Although intestinal organoid culture of stromal cells generated a larger, open an elongated epithelial structure, our findings showed that these fibroblasts/ISEMFs mesenchymal elements were not capable of supporting the organoid growth in the absence of exogenous growth factors (R-spondin, Noggin and EGF) (Figure 5.5). Thus, these unintentionally isolated fibroblasts/ISEMFs are deemed non-supportive: while they likely produce factors that enhance growth, they cannot fully replace the exogenous growth factors. Interestingly, we observed that the fibroblast/ISEMF layer could not survive when the epithelial cells were removed or died, indicating that mutual cell-cell communication between fibroblasts of the mesenchymal niche and epithelial cells are essential for the survival of both cell types. Further work is required to characterise the signalling factors involved in regulating the survival of both epithelial and stromal cells.

The characteristics of the non-supportive stromal cells isolated in this project are consistent with findings of Lei *et al*, they categorised two types of fibroblasts and ISEMFs used in intestinal mouse organoid co-culture: 'supportive' and 'non-supportive'. The supportive stromal element comprised ISEMFs isolated from infant mice aged less than 1 week; these ISEMFs support organoid growth

in the absence of growth factors in the media. Supportive ISEMFs uniquely express high levels of Rspo-2 but not Rspo-1, and Rspo-2 is more potent Wnt activator than Rspo-1 through binding to Lgr5 (Lei *et al.*, 2014). Adult and embryonic fibroblasts were found to be non-supportive of epithelial organoid growth in the absence of added R-spondin (Lei *et al.*, 2014). The *UGT1A8/Luc* mice used in our project were adults. Based on Lei *et al.*'s study, adult ISEMFs are non-supportive since they do not exhibit high enough R-spondin levels to support epithelial organoid growth, therefore exogenous growth factors, particularly Rspo, are required (Lei *et al.*, 2014).



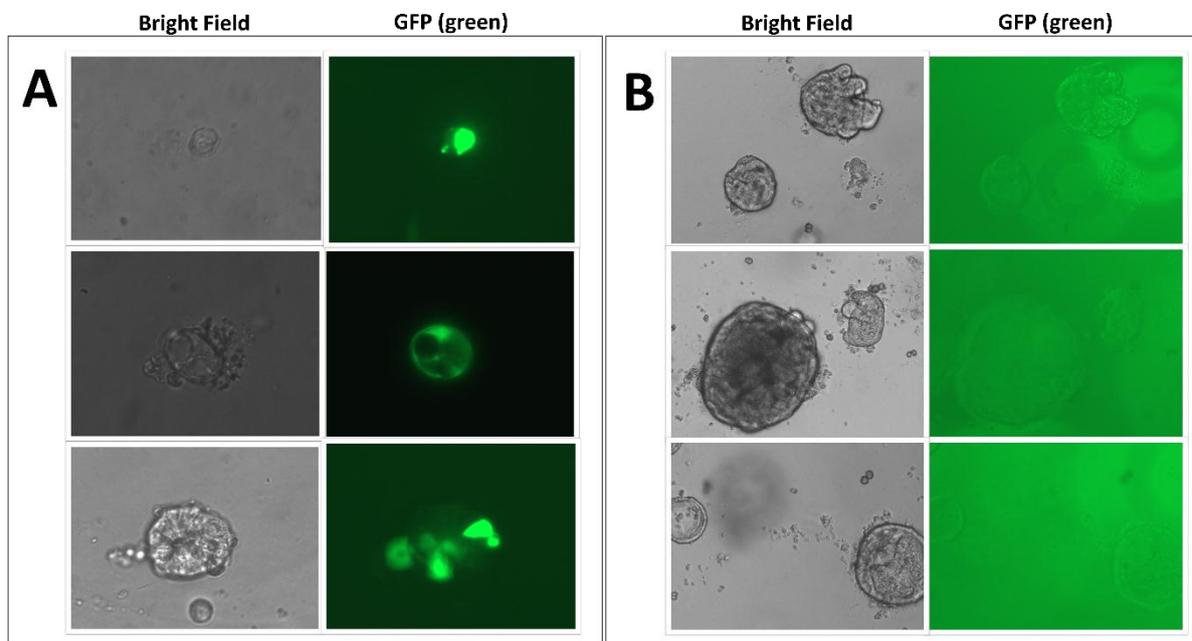
**Figure 5.5. Organoid growth in the presence of fibroblasts/ISEMFs cannot be maintained without exogenous growth factors.**

A. Reconstituted intestinal organoid survives normally in growth factor supplemented media, whereas B. Reconstituted organoid stops growing in the absence of growth factors and leads to fibroblast death. All scale bars are 400 µm. Images were taken using an EVOS® FL microscope in bright field lighting at 10x magnification.

### 5.3.4 Exploration of gene delivery methods in *UGT1A8/Luc*-mouse intestinal enteroids

In general, to understand gene function, expression and regulation, it is critical to be able to genetically manipulate the experimental model system. The genetic manipulation includes gene overexpression and gene silencing/knockdown. Hence, our first task after successfully growing the *UGT1A8/Luc*-enteroid culture, was to identify which gene transfer methods may be effective in enteroid culture. We tested two methods for transfer of an expression construct (using a GFP-expression construct to assess efficiency): DNA-plasmid: liposomal transfection and lentiviral transduction.

We first transfection/transduction of intact enteroids/enterospheres. For lentiviral transduction, we performed a spin-transduction protocol to increase efficiency. This method however was not as successful as the lipofection method, which we performed based on standard Lipofectamine® 2000 protocols. Figure 5.6 shows the comparison of both DNA delivery protocols (lipofection and lentiviral). Although we could not determine the precise transfection efficiency using Lipofectamine® 2000, we were able to confirm that the intact enteroid is amenable to transfection using Lipofectamine® 2000. Although recently a successful lentiviral transduction protocol has been demonstrated in organoids (Van Lidth de Jeude *et al.*, 2015), lipofection has the advantage of ease of use.



**Figure 5.6. GFP transfection into enteroids by lipofection and lentiviral transduction.**

Enteroids were split, resuspended in Matrigel and cultured for 72 hours before transfections. Observation of transfection efficacy was carried out after 72 hours using bright field and fluorescence microscopy. A. Enteroids were observed to be successfully transfected using Lipofectamine®2000 in the Matrigel culture based on GFP expression. B. Transduction using pTiger-GFP lentivirus in Matrigel-embedded organoid culture was observed to be unsuccessful as it did not result in any GFP expression.

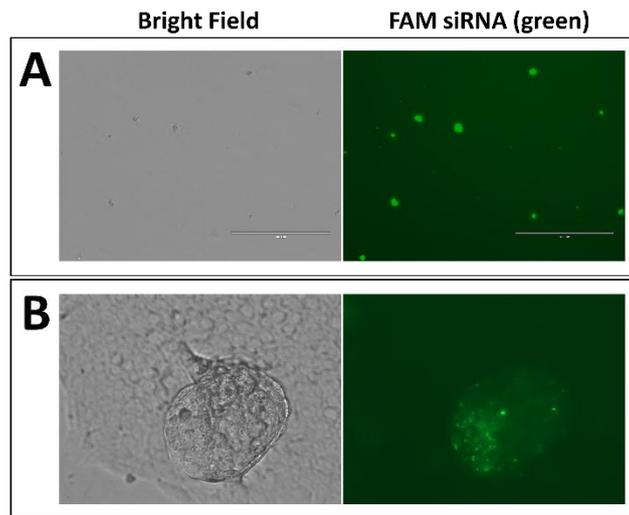
Lipofectamine® 2000 has been the liposomal reagent of choice in various studies in the organoid system (Schwank *et al.*, 2013, Drost *et al.*, 2015) even though efficiencies do not appear to be very high, i.e. Schwank *et al.* found a 2.5% transfection efficiency using this reagent. Another method that has been used fairly successfully is electroporation as published by Fujii *et al.* (Fujii *et al.*,

2015). They demonstrated an efficient gene transfer to organoids by modification of the electroporation method, including the addition of 1.25% (v/v) DMSO to the culture medium and treatment with GSK-3 inhibitor (CHIR99021) a few days before dissociating the organoid and conducting electroporation. They found the electroporation protocol to be superior to lipofection methods. Unfortunately, due to time constraints we were not able to test the electroporation methods in this project.

Next, we tested whether siRNA transfection could be achieved the same liposomal reagent Lipofectamine® 2000. Reconstituted organoids were utilised for siRNA transfections. A fluorescent labelled, fluorescein amidite (FAM) siRNA was used for these experiments. We tried two approaches for transfection: in the first we transfected dissociated single organoid cells and then plating them in a Matrigel base; in the second approach, the dissociated organoid cells were grown in Matrigel for 3 days into cyst or organoid phenotypes before transfection.

Observation of transfection following cell dissociation showed what appeared as successful FAM siRNA transfection with essentially all cells showing uptake of the siRNA 24 hours post-transfection (Figure 5.7A). These cells however did not grow into organoids in Matrigel even after 7 days, suggesting that transfection may have killed the cells. This problem was also encountered when transfecting plasmid into dissociated cells (not shown).

FAM siRNA transfection of intact organoids on the other hand, was successful, with transfected siRNA clearly observed in the organoids (Figure 5.7B). However, siRNA transfection efficiency was not determined in this study. To the best of our knowledge, exploration of siRNA transfection in the enteroid/organoid using Lipofectamine® 2000 has not been reported previously. Collectively, DNA or RNA delivery into *UGT1A8*/Luc-mouse organoids appear to be possible using lipofection methods, albeit it with varying efficiencies.



**Figure 5.7. Transfection of fluorescent labelled, fluorescein amidite (FAM) siRNA into organoids.**

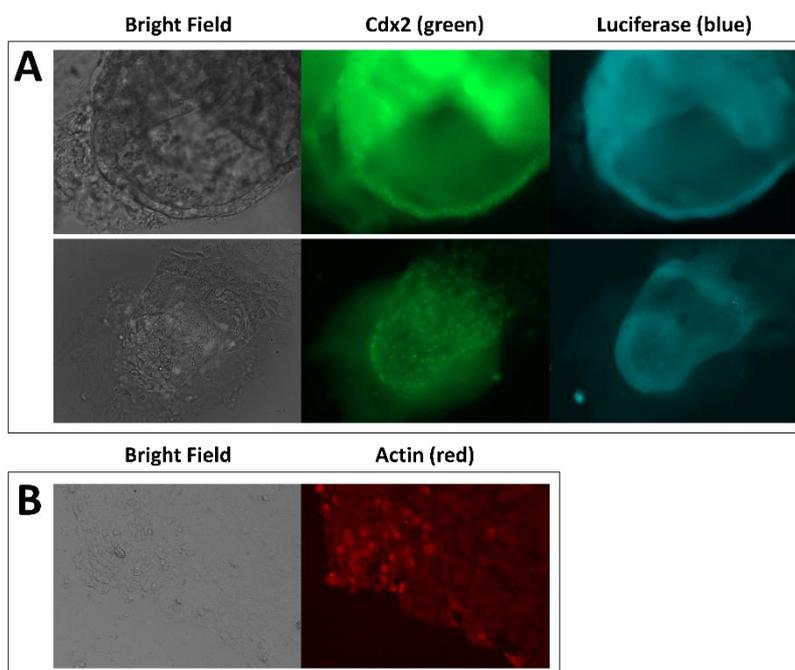
A. A representative image of FAM siRNA transfected single cells of dissociated organoids. Organoids were split into single cells, transfected with siRNA, and plated in Matrigel to grow. B. FAM siRNA transfection into growing organoid. Organoids were grown in the Matrigel in the presence of fibroblast for 72 hours and transfected with FAM siRNA. All images were taken 24 hours post-transfection, at 40x magnification using an EVOS® FL microscope.

### 5.3.5 Immunofluorescent staining of organoid

Immunofluorescent staining was performed on a ‘reconstituted organoid’ culture derived from *UGT1A8/Luc*-mice. We attempted to visualise the expression of the firefly-luciferase protein, which is proportional to the *UGT1A8* gene promoter activity using anti-luciferase antibody (Figure 5.8A). Immunostaining with CDX2 antibody was also performed (Figure 5.8A).

Two reconstituted organoids on day 7 of growth (top and bottom Figure 5.8) were fixed and stained using primary anti-CDX2 antibody (green), firefly luciferase antibody (blue) and flurophore-labelled secondary antibodies. Organoid staining was imaged using a green filter for CDX2 and a blue filter for luciferase (Figure 5.8A). CDX2 appeared to be localised within the nuclei, and positive CDX2 expression was observed along the line of the organoid cyst wall consistent with the presence of organoid intestinal epithelium. *UGT1A8/Luc* expression appeared to be broadly expressed across the structure and cytoplasmic.

Intestinal fibroblasts (in monolayer) from a *UGT1A8/Luc* ‘reconstituted organoid’ culture were stained with anti-actin antibody, and cytoplasmic expression of actin was observed, confirming fibroblast identity (Figure 5.8B).



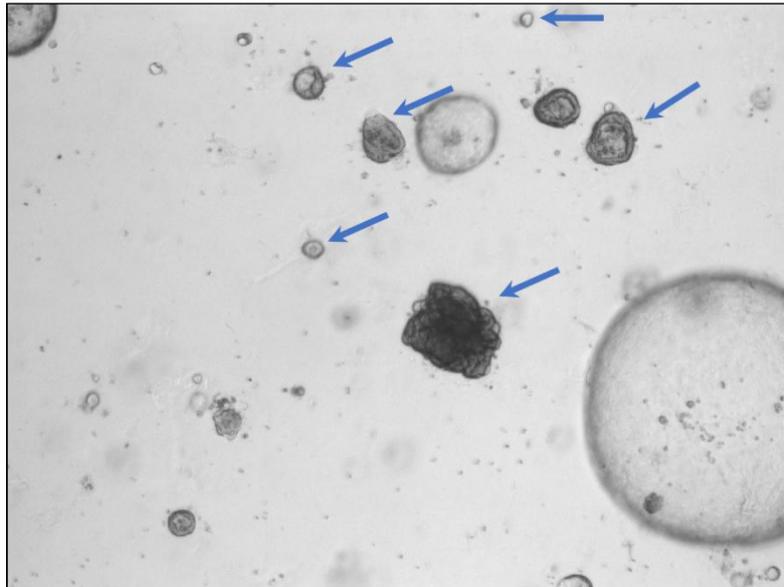
**Figure 5.8. Immunofluorescence of organoids and intestinal fibroblast.**

Immunofluorescent staining was performed in a Matrigel embedded organoid. A. Representative images of *UGT1A8/Luc*-mice organoids expressing Cdx2 protein localised in the nuclei, and firefly luciferase in the cell membranes. The top and bottom images are organoids cultured in different wells. B. Immunostaining of intestinal fibroblasts expressing actin. All images were taken using an EVOS® FL microscope.

Overall, these data suggest that antibodies can penetrate Matrigel to mediate immunofluorescent staining of the embedded organoids; however future work may be required to optimize the use of different antibodies.

### 5.3.6 Challenges and opportunities in organoid culture

We have shown that organoid culture has great potential as an ‘ex vivo’ intestinal model system to study UGTs. Technical difficulties encountered during *UGT1A8/Luc*-enteroid/organoid culture are worthy of discussion here. A major concern is the unequal enteroid growth both in primary culture and during secondary passaging where dissociation of organoids does not reliably produce single cell suspensions (Figure 5.9). The lack of uniformity in organoid growth may affect the reproducibility of experimental manipulations designed to alter cell growth and response.

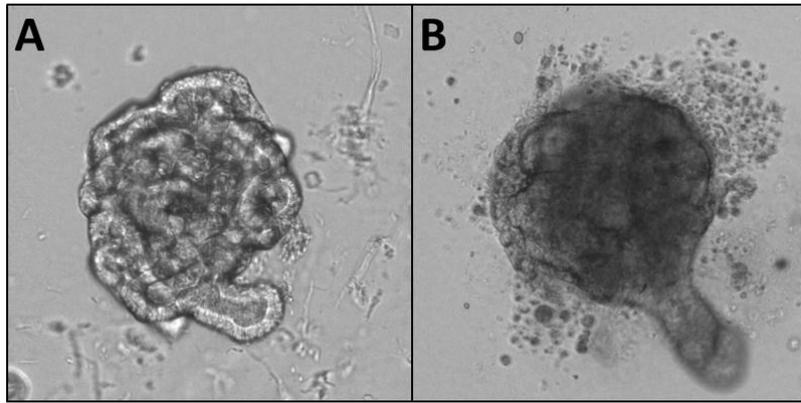


**Figure 5.9. Uneven organoid dissociation into single cells results in unequal organoid growth.**

Blue arrows show various organoid structures which were observed after organoid passage. Image was taken using an EVOS® FL microscope in bright field lighting at 4x magnification.

Another problem is natural progression of organoids towards a non-variable state in culture. Figure 5.10 shows the viability of two different organoids; the appearance of the healthy organoid was marked by a clear lining of epithelial coherence, whereas the appearance of the “disrupted” organoid was darker, and filled with dying or dead cells and disturbed epithelial lining.

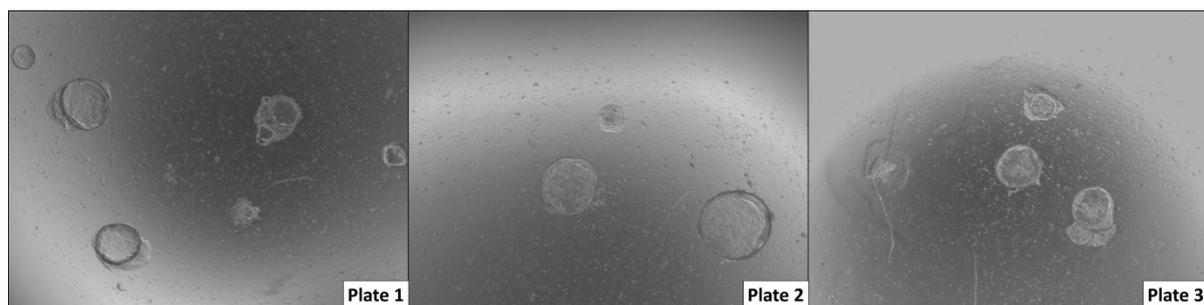
Unfortunately, the organoids do not appear to all ‘age’ at the same rate, and the same culture may show both viable and non-viable organoids. The presence of unhealthy organoids can comprise the health of neighbouring organoids, hence once dying organoids are observed, the culture must be passaged to avoid further deterioration.



**Figure 5.10. Comparison of two stages of organoid viability.**

A. A representative image of a healthy and viable organoid. B. A representative image of an organoid with many dead/dying cells. Both types of organoids may appear in the same culture well and cause difficulty in quantification in an experimental treatment. Images were taken at 10x magnification using an EVOS® FL microscope in bright field lighting.

Regular supplementation of media with fresh growth factors is important as lowered growth factor concentrations prevent the enterosphere from growing into the enteroid or organoid (Figure 5.11). Deficient R-spondin, noggin and EFG cause less activated Wnt signalling, and consistent with previous report (Lahar *et al.*, 2011), we found that if enteroid/organoid culture is not sufficiently supported by growth factors, they may maintain their cyst-structure indefinitely without appropriate growth and differentiation.



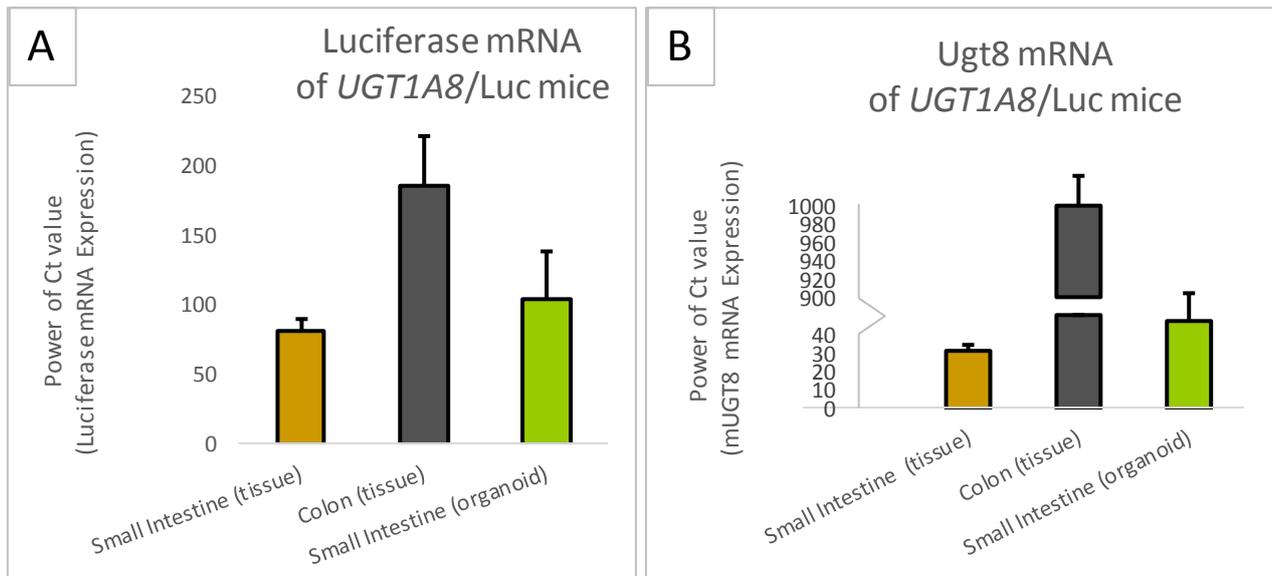
**Figure 5.11. Persistent enterosphere structure after 4 weeks culture in growth factor deficient media.**

Enterosphere structures observed in three wells were persistent as cysts without significant growth; this is caused by insufficient growth factors in the media. Images were taken at 2x magnification using an EVOS® FL microscope in bright field lighting.

### 5.3.7 Detection of UGT1A8/luciferase expression in *UGT1A8/Luc* tissues and enteroids.

We performed qRT-PCR analysis of luciferase expression in transgenic *UGT1A8/Luc*-mice enteroids to investigate the activity of the *UGT1A8* promoter via the linked firefly-luciferase reporter transgene. The mouse small intestine and colon tissues were also analysed for luciferase expression. Unfortunately, only RNA from *UGT1A8/Luc* small intestinal enteroids was able to be collected, as our *UGT1A8/Luc* colonoid culture was not successful. The successfully grown enteroids were harvested on day 7 for RNA analysis. We demonstrated that luciferase mRNA was expressed in all *UGT1A8/Luc* mouse tissue samples, as well as the derived enteroids (Figure 5.12A). Consistent with *UGT1A8* expression along the human gastrointestinal tract, which increases towards the colon (Cheng *et al.*, 1998), our result shows that expression of the *UGT1A8/Luc* reporter in the mouse colon is higher than in the small intestine. Furthermore, we detected a level of *UGT1A8/Luc* expression in the enteroid that is comparable to that in the small intestinal tissue. This result may indicate a similar level of cell maturity in the enteroid as in the adult mouse intestine.

In this RNA-analysis study, we also compared the expression pattern of the *UGT1A8/Luc* reporter to the expression of mouse *Ugt8*. This gene was chosen because of its similar activity between human (*UGT8*) and mouse (*ugt8*) and its significant expression in the intestinal tract (small intestine and colon) (Meech *et al.*, 2015). *Ugt8* mRNA analysis in mouse tissues indicates that its colonic expression is dramatically higher than its small intestinal expression (Figure 5.12B), whereas in human, colonic expression of *UGT8* was reported to be equivalent to the small intestinal expression (Meech *et al.*, 2015). Interestingly however, consistent with *UGT1A8/Luc* expression, expression of *Ugt8* in mouse enteroids is comparable to that in mouse small intestinal tissue.



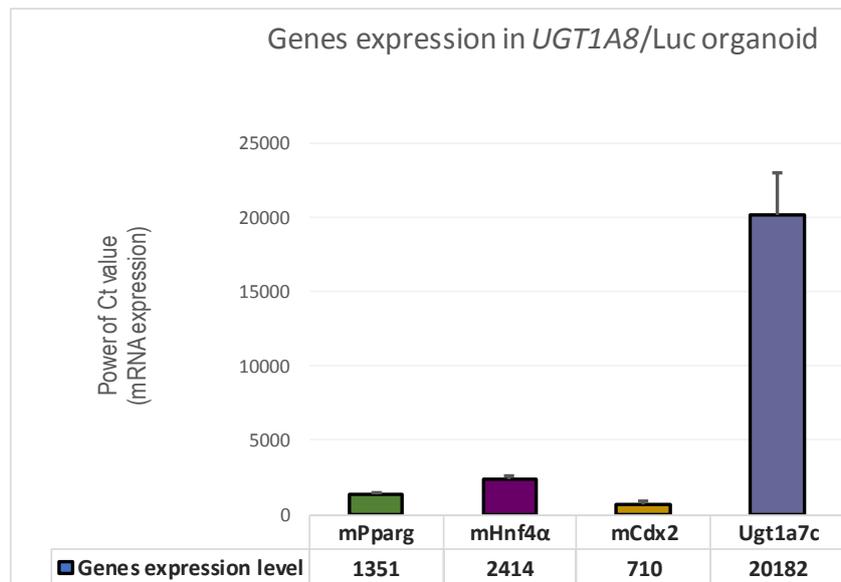
**Figure 5.12. Expression of *UGT1A8/Luc* reporter and *Ugt8* mRNA in tissues and enteroids from *UGT1A8/Luc*-mice.**

A. Expression of the human *UGT1A8/Luc* reporter (luciferase RNA) in mouse small intestinal enteroid is comparable to that the mouse intestinal tissue. *UGT1A8/Luc* expression in colon tissue is higher than in small intestinal tissue. B. A similar pattern of intestinally expressed *Ugt8* to *UGT1A8/Luc* was displayed. The expression level was calculated using the  $2^{-Ct}$  method ( $\times 1000000$ ). Average of expression in a minimum of 4 mice in replicated experiments was taken. All values were normalised by internal 18s (mouse) rRNA expression.

### 5.3.8 Expression of mouse-*Cdx2*, *Hnf4 $\alpha$* , *Ppary*, and *Ugt1a7c* in the *UGT1A8/Luc* enteroids

The expression of intestinal genes in mature enteroids was assessed by qRT-PCR. Mouse *Cdx2*, *Hnf4 $\alpha$* , *Ppary* and *Ugt1a7c* were all expressed (Figure 5.13). High *Cdx2* and *Hnf4 $\alpha$*  expression in the enteroids reflects active intestinal epithelial differentiation and turnover, and the critical roles of these two factors in these processes (Boyd *et al.*, 2009).

*Ppary* transcripts were also high in the enteroid (Figure 5.13), which is consistent with its important role in intestinal cell migration, as well as epithelial cell differentiation (Chen *et al.*, 2006).



**Figure 5.13. Expression of mouse mPPAR $\gamma$ , mHNF4 $\alpha$ , mCdx2 and mouse Ugt1a7c in UGT1A8/Luc organoids.**

The expression level was calculated using the  $2^{-Ct}$  method (\*1000000). Average of expression in a minimum of three organoid cultures was calculated. All values were normalised by internal 18s (mouse) rRNA expression.

Finally, we examined the expression of the mouse *Ugt* gene which most corresponds to human *UGT1A8*. In a study by (Buckley and Klaassen, 2007), *Ugt1a7c* expression was identified as being absent in the liver and high in the gastrointestinal tract, with the *Ugt1a7c* transcripts highest in the colon. This pattern suggests that *Ugt1a7c* may have similar functions in the mouse to human *UGT1A8*. We found that *Ugt1a7c* expression was very high in *UGT1A8*/Luc-enteroids (Figure 5.13), which may suggest a significant function in mouse small intestine.

## 5.4 Conclusion

These studies have shown that establishment of enteroid/organoid cultures from *UGT1A8*/Luc mice is highly feasible, and that they provide a valuable model for studying intestinal development and homeostasis. We successfully cultured intestinal crypts with the ability to self-organise into enteroids which are independent from any mesenchymal element (Figure 5.3). We also demonstrated that organoids can be cultured together with intestinal fibroblasts which promotes more robust growth, even though the adult mouse fibroblasts are technically 'non-supportive' as

they cannot fully replace exogenous growth factors, confirming previous reports from (Lei *et al.*, 2014).

We found that transfection using a liposomal agent is applicable to the enteroid/organoid model within Matrigel culture, both for DNA and RNA delivery. We found that the organoid survival rate was higher when the DNA-delivery was conducted in stable growing organoids rather than dissociated single cells. However, the transfection efficiency was low and not every cell of the organoid can be transfected. One possible solution to this problem is to transfect intact organoids and later dissociate into single cells and grow new organoids: any organoids derived from transfected stem cells should carry the transfected plasmid in all cells. Such organoids could be isolated using GFP as a marker. In addition, other reagents could be tested to determine if they can transfect dissociated cells with less toxicity, this optimisation should include testing the electroporation method. Analysis of FAM labelled siRNA uptake suggests that siRNA may effectively target isolated cells and organoids; however assessing the effectiveness with which gene knockdown can be achieved will require by performing qRT-PCR analysis of a siRNA-target gene.

Expression of *UGT1A8*/Luc in the enteroid system was confirmed by qRT-PCR analysis as well as immunofluorescent staining. The gene expression profile of enteroids is reported as being similar to freshly isolated crypt. We discovered a comparable expression of *UGT1A8*/Luc mRNA level between enteroid and adult mouse small intestinal tissue, which suggests that the enteroid is a good model to study *UGT1A8* gene regulation. High expression of the intestinal specific regulators *Cdx2*, *Hnf4 $\alpha$* , *Ppar $\gamma$*  in the enteroid further support the idea that enteroids can be used to study the factors that control the developmental patterning of the *UGT1A8* transgene, as well the patterning of relevant mouse *Ugt* genes such as *Ugt8* and *Ugt1a7c*.

We also showed that protein expression can be examined in enteroids using immunofluorescence staining of the enteroid/organoid embedded in Matrigel; although replication and further controls are required to confirm the specificity of immunostaining performed in these conditions.

Overall, we have established the fundamental methodology for intestinal organoid culture as well as genetic manipulation of organoids and analysis of gene expression in organoids. These can be applied to future studies of the *UGT1A8*/Luc transgenic mouse. These studies would complement findings reported in Chapter 3 and 4 as discussed in the Introduction to this Chapter. In addition, human intestinal organoids should be generated to assess whether the transgene in the *UGT1A8*/Luc mouse faithfully recapitulates the expression and regulation of the endogenous *UGT1A8* gene. As well as studying *UGT* regulation, knockdown or deletion (e.g. using CRISPR) of intestinal *UGTs* in organoid systems would provide a powerful model to assess the functions of these *UGTs* in intestinal cell detoxification and drug metabolism, and in cancer.

It is frequently observed that failures in translating basic molecular research into clinical outcomes are due to lack of clinically relevant models. Organoid models have been widely embraced as a valuable intermediate between cellular, animal, and clinical studies that may improve the translatability of preclinical research. Moreover, many laboratories are focussed on producing organoid cultures from large numbers of patients that can be used in drug-screening in order to find the most effective therapy for a given patient. This work is beginning to realize the goal of personalized drug-targeting and development. Our establishment of organoid cultures to study *UGT* expression, regulation, and ultimately function, provides a platform to better address the bases of variation in drug metabolic capacity, and is likely to contribute to more effective pharmacological validation and optimization, including personalized therapy.

## CHAPTER 6. GENERAL DISCUSSION AND CONCLUSION

UGT enzymes in the gastrointestinal tract play a critical role as a first line barrier against ingested xenobiotics and in the pre-systemic metabolism of drugs. This project focussed on mechanisms that control the constitutive and inducible expression of *UGT* genes in the intestine. Understanding these two aspects of regulation is of considerable importance as the former ensures that UGTs are expressed in the correct tissue-specific pattern, and the latter allows them to be dynamically responsive to changed demand for detoxification capacity. The mechanisms underlying the constitutive intestinal expression of the *UGT1A8-1A10* genes were investigated in studies described in Chapter 3. The inducible regulation of *UGT1A8* by dietary chemicals, particularly flavonoids, was investigated in studies described in Chapter 4. To complement this work, we established a sophisticated organoid system to further investigate intestinal UGT regulation and function in studies described in Chapter 5.

The expression of UDP-glucuronosyltransferases (UGTs) is well known to be regulated in a tissue-specific fashion. While regulation of hepatic UGT expression by liver-specific transcription factors has been extensively studied, less is known about control of intestinal expression by developmental regulators. All UGT1A family members are present to some degree in the gastrointestinal tract although only *UGT1A7*, *UGT1A8* and *UGT1A10* are specific to this tissue, of these *UGT1A7* is confined to stomach while *UGT1A8* and *-1A10* are expressed in intestine. *UGT1A9*, while a member of the same gene cluster and abundant in intestine, is also expressed in liver.

The expression of intestinal UGTs is highest in the differentiated epithelial cells that are exposed to the luminal contents and lower in undifferentiated cells of the intestinal crypts. At the beginning of this study, the role of CDX2 in specifying UGT expression in intestinal epithelial cells had been relatively well defined. Known as the intestinal master regulator, CDX2 regulates intestinal development during embryogenesis, and maintains cellular proliferation and differentiation during intestinal epithelial renewal (Ralston and Rossant, 2008, Gao *et al.*, 2009, Silberg *et al.*, 2000).

Studies in mouse show that Cdx2 partners with HNF4 $\alpha$  to regulate intestinal specific genes (Olsen *et al.*, 2012); both factors interact and co-occupy target promoters in a combinatorial transcriptional mechanism (Verzi *et al.*, 2013, San Roman *et al.*, 2015).

While CDX2 was shown to regulate *UGT1A8* and *-1A10* via a CDX2 response element (CDX2-RE) that is conserved in these two genes but not conserved in *UGT1A9* (Gregory *et al.*, 2004a). (Gregory *et al.*, 2004b). However, previous studies had not defined any role for HNF4 $\alpha$  in intestinal regulation of these genes. In this thesis project, we provided the first evidence of cooperative regulation of *UGT1A8*, *-1A9* and *-1A10* by CDX2 and HNF4 $\alpha$ . We found that CDX2 and HNF4 $\alpha$  were required to maintain endogenous expression of *UGT1A8*, *-1A9* and *-1A10*, in Caco-2 cells. At the promoter level, our major finding was identification of a novel 12 nt Cdx2/HNF4 $\alpha$  synergistic composite element located in the proximal promoters of these three genes. EMSA analysis showed that the element was bound by both CDX2 and HNF4 $\alpha$  via two adjacent but discrete motifs. This is the first evidence showing CDX2 and HNF4 $\alpha$  integration in such a short sequence (12 nt), and is different from the previously described synergistic interaction of CDX2 and HNF1 $\alpha$  that is dependent on two more distantly spaced elements (Cdx2-RE at -70 bp and HNF1 $\alpha$ -RE at -105 bp from TSS) (Gregory *et al.*, 2004a).

An additional significant finding was that the novel CDX2/HNF4 $\alpha$  composite element is functional in the *UGT1A9* promoter and mediates regulation of this gene by CDX2. Given that the previously identified CDX2 RE in the *UGT1A8* and *-1A10* promoters was found to be non-functional in *UGT1A9*, it has been unclear how intestinal regulation of *UGT1A9* is determined. Our new findings resolve this quandry.

Our studies of *UGT1A9*, which is both hepatic and intestinal, also gave greater insight into how HNF4 $\alpha$  might control hepatic and intestinal regulation via different regulatory modules. It appears that the CDX2/HNF4 $\alpha$  composite element does not function in hepatic cells where CDX2 is absent. Instead, in this context, upstream HNF4 $\alpha$  elements bind and mediate regulation of the *UGT1A9* promoter. These upstream HNF4 $\alpha$  elements are non-functional in *UGT1A8*, *-1A10*, thus these

genes are not expressed in liver. Hence we define an 'intestinal module' where CDX2 and HNF4 $\alpha$  interact to act synergistically and a 'liver module' where HNF4 $\alpha$  functions either alone or with other partners such as HNF1 $\alpha$  (Barbier *et al.*, 2005, Gardner-Stephen and Mackenzie, 2007).

One model that is consistent with our results is that the presence of CDX2 changes the function of HNF4 $\alpha$  such that it preferentially binds the 'intestinal module' rather than the 'liver module' in the UGT1A9 promoter. This expands our understanding of intestinal vs hepatic UGT patterning involving co-regulation by tissue specific developmental factors.

The inducible regulation of UGTs by ligand-activated transcription factors (TFs) allows dynamic response to chemical insults (Hu *et al.*, 2014a). Many such ligands are also UGT substrates thus providing a direct feedback response (Bock, 2012). Substances contained in foods consumed daily may be metabolised as substrates by intestinal UGTs and potentially also regulate the level of UGTs. Flavonoids are thought to be responsible for many of the beneficial effects of diets rich in fruits and vegetables. The activities of UGT enzymes contribute to the low bioavailability of flavonoids (Wu *et al.*, 2011, Gao and Hu, 2010). To understand how flavonoid compounds may affect intestinal UGT regulation, we studied UGT1A8 which has intense activity towards flavonoids (Mojarrabi and Mackenzie, 1998, Cheng *et al.*, 1999). A differentiated Caco-2 cell model was used to represent mature intestinal epithelial cells. Caco-2 cells were differentiated in two ways: traditionally by 21 day culture, and by using sodium butyrate (NaB) as inducer for more rapid differentiation (Yamashita *et al.*, 2002). NaB was also considered relevant to the intestinal environment where microbiota endogenously produce butyrate and it may play a role in colon cancer prevention by promoting cellular differentiation (Candido *et al.*, 1978, Kim *et al.*, 1980). To date, despite the great contribution butyrate makes to metabolic activities in the intestine, information that links butyrate with xenobiotic metabolism, particularly involving UGT enzymes, has been very limited.

We found that flavonoids only induced UGT1A8 promoter activity in butyrate-induced differentiated Caco-2 cells and not in 21-day long-term differentiated cultures. This might have been due to

different characteristics of the differentiated cultures, or due to a direct effect of the butyrate on increasing the accessibility of the *UGT1A8* promoter to ligand-dependent transcription factors. In assaying both *UGT1A8* promoter activation and RNA analysis, we found that genistein had the greatest capability in inducing *UGT1A8* expression and that genistein and NaB acted synergistically. From a functional standpoint, this finding suggests that ingestion of genistein together with dietary fibre that promotes butyrate production in the intestine might increase *UGT1A8* expression. This could promote the metabolism of carcinogenic materials and reduce the risk of carcinogenesis. Confirmation of this model would require studies *in vivo*, although it could also be informed by studies using cancer-derived organoid models, including the emerging technology of organoid/microbiota co-culture (Williamson *et al.*, 2018).

Potential regulatory circuits involving genistein and *UGT1A8* may not be surprising, as Tang *et al.* reported that *UGT1A8* is the most potent UGT in genistein glucuronidation at a concentration of 10  $\mu$ M and over (Tang *et al.*, 2009). Hence, our data suggests that although *UGT1A8* is responsible for the low bioavailability of genistein, reciprocally, genistein exposure significantly increases *UGT1A8* expression. Given that flavonoids such as genistein have anti-cancer properties, whether the increased metabolism of flavonoids due to induction of *UGT1A8* would offset any protective effects mediated by enhanced clearance of carcinogens is unclear.

A major component of our studies was identifying the mechanism by which genistein regulates the *UGT1A8* gene. Many studies have reported that the transcription factors PPARs ( $-\alpha$  and  $-\gamma$ ) and ER $\beta$  mediate gene transcription activation by genistein (Kuiper *et al.*, 1998, Ricketts *et al.*, 2005, Chacko *et al.*, 2005, Salam *et al.*, 2008, Pallauf *et al.*, 2017). Our studies using PPAR and ER antagonists suggested that induction of *UGT1A8* mRNA by genistein may involve both classes of factors. However studies at the promoter level did not reveal any role for ER hence subsequent studies focussed in PPAR factors. Interestingly, previous report suggests that genistein acts via ER $\beta$  at a low concentrations of  $<1 \mu$ M and via PPAR $\gamma$  at higher concentrations (Dang, 2009, Patel and Barnes, 2010), hence our results may relate to the fact that genistein was used at high concentration in our studies.

The use of PPAR agonists and overexpression of PPAR $\alpha$  and  $\gamma$  genes confirmed that the UGT1A8, -1A9 and -1A10 genes are all responsive to PPAR, and likely have functional PPAR response elements (PPREs) in their proximal promoters. However, extensive studies failed to define this element in the context of Caco-2 cells. This was inconsistent with studies in COS7 cells where we were able to show the effect of PPAR $\gamma$  on *UGT1A8* promoter induction was mediated by a PPRE at -818 bp. This PPRE corresponds to the PPRE identified previously by Barbier *et al* in *UGT1A9* that also functioned in COS7 in their studies (Barbier *et al.*, 2003). It is intriguing that this PPRE did not mediate the response to PPAR factors in Caco-2 cells, and suggests that nuclear receptors function differently in different cellular contexts. We explored whether other intestinal factors might be required to cooperate with PPAR in Caco-2 cells, either at a transcriptional or post-transcriptional level. We saw evidence for antagonistic interaction between HNF4 $\alpha$  and PPAR $\gamma$  in regulation of *UGT1A9* promoter activity. One possible explanation for this effect is modulation of post-transcriptional mechanisms in association with ERK1/2 phosphorylation. PPAR $\gamma$  phosphorylation is known to involve ERK1/2 activation (Chen *et al.*, 2003), and a recent report showed that HNF4 $\alpha$ -dependent gene regulation is inhibited by phosphorylation related-ERK1/2 activation (Vetř *et al.*, 2017). However, this interaction might mediate a tissue specific effect of PPAR remains unclear.

We also found evidence that genistein can modulate the epigenetic status at the proximal promoter of *UGT1A8*, shown by enrichment of activating histone modifications. This effect might cooperate with the known effect of butyrate on inhibiting histone deacetylases (HDACs) (Davie, 2003). Thus, it is possible that the synergistic action of genistein and butyrate involves chromatin modification. Overall however, we consider that more work is required to fully understand how genistein regulates the *UGT1A8-1A10* genes in the intestinal context.

Our studies used the Caco-2 cell line model. However, even when differentiated this model does not perfectly represent intestinal enterocytes; moreover, it cannot model the lineage progression of intestinal stem cells and transit amplifying cells. Study of UGTs in this context would provide a better understanding of how these progenitor cells might be protected from chemical insults,

particularly since these are considered the cells of origin for cancer. While mice represent a physiologically relevant system to study the intestinal genes, the human *UGT* genes do not have direct homologs in mice. Our laboratory previously generated transgenic 'humanised' mice carrying 13kb of the *UGT1A8* promoter linked to a luciferase reporter that could be used in our studies. We elected to use these mice to establish the intestinal enteroid/organoid model in our laboratory.

Using the available protocols (Sato *et al.*, 2009), we isolated intestinal crypts of *UGT1A8/Luc* mice and generated two types of intestinal organoids. One was a self-organised 'enteroid' culture which is independent of any mesenchymal element; the other was a 'reconstituted intestinal organoid' (shorted to 'organoid'), which is a co-culture of epithelial and mesenchymal (fibroblast) components (Stelzner *et al.*, 2012). Our enteroids were closed spheres that showed distinctive budding events; whereas organoids cultured with fibroblasts formed larger elongated epithelial structures that did not close, suggesting fibroblast contribution to shaping the organoid. While our organoid co-culture was not fully self-supportive (i.e. it required external growth factors), it grew more robustly than the enteroids. Our findings were in agreement with previous reports that adult fibroblasts (unlike neonatal) are a 'non-supportive' element. The phenotype of our mesenchymal component resembled subepithelial myofibroblasts (ISEMFs); however, we did not confirm this with marker analysis. ISEMFs from neonatal mice are known to express high level of *Rspo-2* and is thus a supported matrix in organoid culture (Lei *et al.*, 2014). Both of our cultures (enteroid and organoid) however, showed self-renewal capacity while maintaining their stem-ness during organoid passages.

As genetic manipulation is essential in the study of transcriptional regulatory mechanisms, we tested the effectiveness of DNA and RNA delivery systems using the simple protocol of lipofection and also lentiviral transduction. Surprisingly, Lipofectamine® 2000 was modestly effective at introducing plasmid DNA and very effective at introducing siRNA into organoids. While the transfection efficiency was higher in dissociated organoid cells, the survival rate of these cells was low. Furthermore, although lentiviral transfection generally shows efficient transfection, our study showed that mature organoids were resistant to the lentiviral spinfection, and again, dissociating

organoid cells prior to spinfection caused low survival rates. Optimisation of these transfection methods and techniques is required in the future and other methods such as electroporation should also be tested. Currently however, our data suggest that siRNA can be effectively introduced and further studies will assess the efficacy of target gene knockdown.

It was important to determine whether the level of gene expression in organoids grown from *UGT1A8*/Luc-mice paralleled expression levels in tissue. We found very good correspondence of *UGT1A8*/Luc reporter gene expression in intestinal organoids and intestinal tissue by analysis of luciferase mRNA level. We also showed that two endogenous mouse intestinal *Ugts*, *Ugt8* and *Ugt1a7c*, and the intestinal regulators *Ppar $\gamma$* , *Hnf4 $\alpha$*  and *Cdx2* were all expressed equivalently in organoids and intestine tissue. This data sets the stage for studies of *UGT1A8* regulation in organoids by knockdown of these factors, or chemical inhibition, as well as potentially overexpression studies. As mentioned previously, there is also the potential to study the role of dietary compounds and the microbiome in regulation of UGT expression and function using organoids. Indeed, this model might be ideal to study the complex yet critical interplay of intrinsic (transcriptional), endobiotic (including microbiome derived), and xenobiotic (dietary or pharmacological) regulators in determining the metabolic capacity of the gut.

Studies to examine the consequences of loss of UGT function in the organoid model will require a different mouse model – such as the humanized *UGT1* mouse that carries the entire human *UGT1* locus (protein coding and regulatory regions of all *UGT1* genes). Knockdown of specific *UGT* genes using siRNA in organoids from these mice could allow their contribution to the intestinal metabolism of particular chemicals to be assessed, including clinically important drugs, as well as potentially the consequences of their activity (or lack thereof) for carcinogenesis.

In summary, this project uncovered important new aspects of both constitutive and inducible regulation of intestinal UGTs. In practical terms, our study supports the idea that activation of intestinal UGTs contributes to the protective effect of foods containing-flavonoids, combined with the activity of butyrate producing-microbiota, by enhancing detoxification capacity. This may prompt the development of UGTs as biomarkers or even targets in the developing fields of

microbiome modulation and functional foods. We also show how organoids may provide an easily manipulated *in vitro* model system of normal intestine that could be used to achieve a deeper understanding of intestinal UGT regulation and function, which may play an important role in drug development as well as the emerging field of personalized drug selection.

# APPENDICES

## Appendix 1

Chemicals	Supplier
<b>Cell treatment chemicals</b>	
3,6-dihydroxyflavone	Sigma Chemical Co, St Louis, MO, USA
3,7-dihydroxyflavone	Sigma Chemical Co
17- $\alpha$ -ethynylestradiol	Sigma Chemical Co
19-hydroxy-4-androstene-3,17-dione	Sigma Chemical Co
2'-hydroxyflavanone	Sigma Chemical Co
4'-hydroxyflavanone	Sigma Chemical Co
6'-hydroxyflavanone	Sigma Chemical Co
3'-hydroxyflavone	INDOFINE Chemical Company, Inc., Hillsborough, NJ, USA
4'-hydroxyflavone	INDOFINE Chemical Company, Inc.
5'-hydroxyflavone (Primuletin)	Sigma Chemical Co
6'-hydroxyflavone	Sigma Chemical Co
4'-methoxyflavanone	Sigma Chemical Co
5'-methoxyflavanone	Sigma Chemical Co
6'-methoxyflavanone	Sigma Chemical Co
3,7,4-trihydroxyflavone	Sigma Chemical Co
Baicalein	Aldrich Chemical Company Inc., Milwaukee, WI, USA
Biochanin A (5,7-dihydroxy-4'-methoxyisoflavone)	Sigma Chemical Co
Chrysin	Aldrich Chemical Company Inc.
Corticosterone	Sigma Chemical Co
Cortisone	Sigma Chemical Co
Danazol	Sigma Chemical Co
Dehydroisoandrosterone	Sigma Chemical Co
Diethylstilbestrol	Sigma Chemical Co
(-)-Epigallocatechin gallate	Sigma Chemical Co
Estradiol	Sigma Chemical Co
Estriol	Sigma Chemical Co
Estrone	Sigma Chemical Co
Flavanone	Sigma Chemical Co
Flavone (2-phenyl-4H-1-benzopyran-4-one)	Sigma Chemical Co
Galangin	Sigma Chemical Co
Genistein	Sigma Chemical Co
Morin	Sigma Chemical Co

Myricetin	Sigma Chemical Co
(±)-naringenin (4',5,7-trihydroxyflavanone)	Sigma Chemical Co
Naringin (4',5,7-trihydroxyflavanone 7-rhamnoglucoside)	Sigma Chemical Co
Norethindrone	Sigma Chemical Co
D(-)-norgestrel	Sigma Chemical Co
Prednisolone	Sigma Chemical Co
Progesterone	Sigma Chemical Co
Quercetin (3,3',4',5,7-pentahydroxyflavone) dehydrate	Sigma Chemical Co
Tamoxifen	Sigma Chemical Co
Testosterone	Sigma Chemical Co
BI6015	Cayman Chemical, Ann Arbor, MI, USA
Clofibric acid	Sigma-Aldrich, St Louis, MO, USA
GW9662	Tocris Bioscience, Bristol, UK
Rosiglitazone	Sigma-Aldrich
<b>Buffer chemicals</b>	
Acetic acid	Ajax Finechem, Seven Hills, NSW, Australia
Boric acid	Fluka Analytical (Honeywell), Morris Plains, NJ, USA
Bovine serum albumin (BSA) solution (100 mg/ml)	New England Biolabs (Mackenzie <i>et al.</i> ), Beverley, MA, USA
Bromophenol blue	Sigma Chemical Co
Dimethyl sulfoxide (DMSO)	Merck, Darmstadt, Germany
dithiothreitol (dTT)	Sigma-Aldrich
Ethylenediaminetetra-acetic acid, di-sodium salt (EDTA)	Biochemicals, Gymea, NSW, Australia
Glycerol	Amresco, Solon, OH, USA
Glycine	Amresco
HCl	VWR International, Radnor, PA, USA
Isopropanol	Chem-Supply, Gillman, SA, Australia
KCl	Amresco
KH <sub>2</sub> PO <sub>4</sub>	Amresco
Methanol	RCI Labscan, Bangkok, Thailand
MgCl <sub>2</sub>	Amresco
myo-inositol	Merck
Na <sub>2</sub> HPO <sub>4</sub>	Ajax Finechem
NaCl	Biochemicals
Nonidet P-40	Fluka Analytical (Honeywell)

Proteinase K	NEB and Thermo Fisher Scientific
Sodium dodecyl sulphate (SDS)	A.G. Scientific, San Diego, CA, USA
Tris[hydroxymethyl]aminomethane	Astral Scientific, Taren Point, NSW, Australia
Triton-X	Sigma-Aldrich
Xylene cyanol FF	Sigma-Aldrich
<b>Mammalian tissue culture</b>	
Dulbecco's modified Eagle's medium (DMEM)	Invitrogen/Life Technologies, Carlsbad, CA, USA
Foetal calf serum	HyClone (GE Healthcare), Chicago, IL, USA, and Thermo Fisher Scientific
MEM non-essential amino acids	Invitrogen (Thermo Fisher Scientific)
MEM sodium pyruvate	Invitrogen (Thermo Fisher Scientific)
Tissue culture flasks and plates	Nunc, Roskilde, Denmark
Trypan blue	Sigma-Aldrich
Trypsin-EDTA	Invitrogen (Thermo Fisher Scientific)
<b>Organoid culture</b>	
Advanced-DMEM/F12	Gibco (Thermo Fisher Scientific)
B-27 supplement	Gibco
Enzyme Free Cell Dissociation Buffer Glutamax	Gibco
HEPES	Gibco
Human recombinant Rspodin1	PeproTech, Rocky Hill, NJ, USA
IntestiCult™ Organoid Growth Medium (Mouse)	Stemcell Technologies, Vancouver, BC, Canada
Matrigel®	Sigma-Aldrich
Mouse recombinant EGF	PeproTech
Mouse recombinant Noggin	PeproTech
N2 supplement	Gibco
Penicillin/Streptomycin	Gibco
Y-27632	Sigma-Aldrich
<b>Transfection and Reporter Gene Assays</b>	
Dual-luciferase Reporter Assay System	Promega, Madison, WI, USA
Lipofectamine® 2000	Invitrogen (Life Technologies)
<b>Electroporation</b>	
poly (deoxyinosinic-deoxycytidylic) (poly (dI-dC))	Sigma-Aldrich
<b>Bacterial Culture</b>	
Agar	Amresco

Ampicillin	Aspen Pharmacare, KwaZulu-Natal, South Africa
Kanamycin	Sigma-Aldrich
Luria Broth (LB) EZMix	Amresco
<b>DNA Detection, Purification and Modification</b>	
30% Acrylamide/Bis solution 19:1	Bio-Rad, Hercules, CA, USA
Agarose	Astral Scientific
Ethidium bromide	Amresco
QIAGEN Plasmid Midiprep kit	Qiagen, Hilden, Germany
QIAprep Spin Miniprep kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
QIAquick PCR Purification kit	Qiagen
Quick Ligation kit	NEB
Restriction enzymes	NEB
<b>RNA Purification and cDNA Synthesis</b>	
Amplification grade DNaseI	NEB
Chloroform	VWR
NxGen M-MuLV Reverse Transcriptase	Lucigen, Middleton, WI, USA
NxGen RNase Inhibitor	Lucigen
TRIzol	Thermo Fisher Scientific, Scoresby, VIC, Australia
<b>Polymerase Chain Reaction (PCR)</b>	
Deoxynucleotide-triphosphate mix (dNTP)	NEB
GoTaq qPCR master mix	Promega
Oligonucleotides	Geneworks or Integrated DNA Technologies
Phire HotStart DNA Polymerase	Thermo Scientific
Phusion High-Fidelity DNA Polymerase	Thermo Scientific
<b>Western Blot</b>	
30% Acrylamide/Bis solution (29:1)	Bio-Rad
Ammonium persulphate	Amresco
BioRad Protein Assay Reagent	Bio-Rad
Complete Proteinase Inhibitor tablets	Roche Diagnostics, Mannheim, Germany
N,N,N',N'-Tetramethyl-1-,2-diaminomethane (Temed)	Sigma-Aldrich
Skim milk powder	Fonterra Brands, NZ
SuperSignal West Pico chemiluminescent (ECL) HRP substrate	Thermo Fisher Scientific

Trans-blot nitrocellulose membrane	Bio-Rad
Tween-20	Astral Scientific
<b>Chromatin Immunoprecipitation (ChIP)</b>	
ChIP grade Protein G magnetic beads	Cell Signaling Technology, Danvers, MA, USA
Formaldehyde	Sigma-Aldrich
Protein A sepharose beads	GE Healthcare
<b>Antibodies</b>	
CDX2	Biogenex, San Ramon, CA, USA
DAPI (4',6-diamidino-2-phenylindole)	Sigma-Aldrich
Histone H3K acetylation	Millipore, Burlington, MA, USA (06-599)
Histone H3K4 tri-methylation	Millipore (07-473)
Histone H3K27 tri-methylation	Millipore (07-449)
HNF4 $\alpha$	Santa Cruz Biotechnology, Santa Cruz, CA, USA (sc-6556)
Luciferase	Promega
Normal rabbit IgG	Cell Signaling Technology (2779S)

## Appendix 2

<u>General buffers (1x working solution)</u>	<u>Chromatin Immunoprecipitation</u>
<b>Phosphate buffered saline (PBS)</b> 137 mM NaCl 2.7 mM KCl 10 mM Na <sub>2</sub> HPO <sub>4</sub> 1.8 mM KH <sub>2</sub> PO <sub>4</sub> pH 7.4	<b>Lysis Buffer 1 (Cell Lysis)</b> 1% Nonidet P-40 (NP-40) 15 mM Tris 0.5 mM EGTA 15 mM NaCl 60 mM KCl 300 mM Sucrose 0.5 mM β-mercaptoethanol pH 8.0
<b>Tris-acetate EDTA electrophoresis buffer (TAE)</b> 40 mM Tris (pH 7.6) 20 mM acetic acid 1 mM EDTA	<b>Lysis Buffer 2 (Nuclear Lysis)</b> 1% SDS 10 mM EDTA 50 mM Tris pH 8.0
<b>Tris-borate EDTA electrophoresis buffer (TBE)</b> 89 mM Tris 89 mM boric acid 2 mM EDTA pH 8	<b>Dilution Buffer</b> 0.01% SDS 1% Triton X-100 1.2 mM EDTA 16.7 mM Tris 150 mM NaCl pH 8.0
<b>SDS-PAGE running buffer</b> 25 mM Tris 192 mM glycine 0.1% SDS pH 8.3	<b>High Salt Wash Buffer</b> 0.1% SDS 1% Triton X-100 1 mM EDTA 20 mM TrisCl 500 mM NaCl pH 8.0
<b>SDS-PAGE transfer buffer</b> 25 mM Tris 192 mM glycine pH 8.3 20% methanol	<b>LiCl Buffer</b> 1% NP-40 1% deoxycholic acid sodium salt 1 mM EDTA
<b>Tris-buffered saline (TBS)</b> 10 mM Tris 150 mM NaCl pH 8	

<p><b>SDS loading sample buffer (Laemmli buffer)</b></p> <p>50 mM TrisHCL pH 6.8  10% SDS  30% glycerol  5% β-mercaptoethanol  0.02% bromophenol blue</p>	<p>10 mM Tris  250 mM LiCl  pH 8.0</p> <p><b>TE Buffer</b></p> <p>1 mM EDTA  10 mM Tris  pH 8.0</p>
<p><b><u>Electrophoretic Mobility Shift Assay (EMSA)</u></b></p> <p><b>Hypotonic Lysis Buffer</b></p> <p>20 mM Tris  10 mM MgCl<sub>2</sub>  10 mM KCl  1 mM EDTA  10% glycerol  1% Triton X-100  1 mM DTT  10 mM glycerol-2-phosphate  pH 7.4</p> <p><b>Nuclear Extract Buffer</b></p> <p>20 mM HEPES  420 mM NaCl  5 mM EDTA  10% glycerol  10 mM glycerol-2-phosphate</p> <p><b>EMSA Buffer</b></p> <p>10 mM Tris  100 mM NaCl  1 mM MgCl<sub>2</sub>  pH 8  20% glycerol</p>	<p><b>Elution Buffer</b></p> <p>20 mM Tris  5 mM EDTA  50 mM NaCl  1% SDS  pH 8.0</p> <p><b><u>Electroporation</u></b></p> <p><b>Isoosmolar electroporation buffer (IEB)</b></p> <p>0.3 mM KH<sub>2</sub>PO<sub>4</sub>  25 mM KCl  0.85 mM K<sub>2</sub>HPO<sub>4</sub>  280 mOsmol/kg myo-inositol</p>
	<p><b><u>Western Blotting</u></b></p> <p><b>Radioimmunoprecipitation assay buffer (Chen <i>et al.</i>)</b></p> <p>50 mM Tris-HCl, pH 8.0  150 mM NaCl  1% Nonidet P-40  0.5% sodium deoxycholate  0.1% sodium deodecyl sulphate (SDS)</p>

## Appendix 3

### MOLPHARM/2017/110619 Published Online

molpharm@aspet.org

Fri 3/9/2018 3:55 AM

Siti Mubarakah; [julieann.hulin@flinders.edu.au](mailto:julieann.hulin@flinders.edu.au); [peter.mackenzie@flinders.edu.au](mailto:peter.mackenzie@flinders.edu.au);  
[ross.mckinnon@flinders.edu.au](mailto:ross.mckinnon@flinders.edu.au); [alex.haines@flinders.edu.au](mailto:alex.haines@flinders.edu.au); [donggui.hu@flinders.edu.au](mailto:donggui.hu@flinders.edu.au);  
[robyn.meech@flinders.edu.au](mailto:robyn.meech@flinders.edu.au)

Inbox

MS ID#: MOLPHARM/2017/110619

MS TITLE: Cooperative regulation of intestinal UDP-glucuronosyltransferases by CDX2 and HNF4 $\alpha$  is mediated by a novel composite regulatory element

We are pleased to announce that the above manuscript has been published online in the Fastforward section of the Journal. Please visit <http://molpharm.aspetjournals.org> to view the published manuscript.

Thank you,  
Molecular Pharmacology Editorial Office

**MOL #110619**

**Cooperative regulation of intestinal UDP-glucuronosyltransferases 1A8, -1A9 and 1A10 by CDX2 and HNF4 $\alpha$  is mediated by a novel composite regulatory element**

Nurul Mubarakah, 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)

Running Title: Regulation of UGTs by CDX2 and HNF4 $\alpha$

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](mailto: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

## 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 HNF4 $\alpha$ , and mediates synergistic activation by these factors. We also show that HNF4 $\alpha$  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 HNF4 $\alpha$  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.

## 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, acetaminophen, ketoprofen, mycophenolic acid, raloxifen, resveratrol, and quercetin (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 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 Hnf4 $\alpha$  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 Hnf4 $\alpha$  to adjacent sites in chromatin (Verzi et al., 2013). Simultaneous deletion of both *Hnf4 $\alpha$*  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 HNF4 $\alpha$  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 (HNF1 $\alpha$ ) 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 *UGT1A9* regulation by CDX2 unresolved.

The current study shows that *UGT1A8*, *-1A9* and *-1A10* expression is programmed by the CDX2/HNF4 $\alpha$  regulatory axis, and identifies a novel composite promoter element that mediates synergistic activation by these factors. Further we propose a model for regulation of intestinal/hepatic *UGT1A9* by both CDX2 and HNF4 $\alpha$  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<sub>2</sub>. Cells were plated into 48-well plates at a density of 4 × 10<sup>4</sup> cells/well and transfected the following day with 0.2  $\mu$ g of each pGL3basic promoter- reporter construct and 0.02  $\mu$ g of the *Renilla reniformis* vector pRL-null (Promega, Madison, WI) using 2  $\mu$ l/well Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen). For cotransfections, 0.2  $\mu$ g of HNF4 $\alpha$ , CDX2, or both HNF4 $\alpha$  and CDX2 expression vectors (effectors) were added to the above reaction mix and normalized to a total of 0.4  $\mu$ g DNA with empty expression vector pCMV5, before incubation with 1.2  $\mu$ l/well Lipofectamine 2000. After 48h, the cells were harvested in 50  $\mu$ l of 1 $\times$  passive lysis buffer and 20  $\mu$ l assayed for 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 HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  protein levels after siRNA transfection, cell lysates were subjected to immunoblotting analysis using anti-CDX2, anti-HNF4 $\alpha$  and  $\beta$ -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](http://www.lifetechnologies.com)); after DNase treatment, cDNA was synthesized using NxGen M-MuLV reverse transcriptase (Lucigen, Wisconsin, [www.lucigen.com](http://www.lucigen.com)) and random primers (New England Biolabs, Ipswich, Massachusetts, [www.neb.com](http://www.neb.com)). Quantitative RT-PCR was performed using a Corbett Rotorgene (Qiagen, Venlo, Limburg, Netherlands, [www.qiagen.com](http://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 HNF4 $\alpha$  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 10 $\mu$ g of antibody. Rabbit antibodies against HNF4 $\alpha$  (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  $\mu$ l of Tris-EDTA buffer and 2  $\mu$ l 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)

Caco2 cells were transfected with the HNF4 $\alpha$  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 HNF4 $\alpha$  (sc-6556; Santa Cruz Biotechnology) and CDX2 (Biogenex, San Ramon, CA) at 1  $\mu$ g 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, HNF4 $\alpha$ ) 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

### *Synergistic regulation of the UGT1A8 promoter by CDX2 and HNF4 $\alpha$ .*

In previous work we showed that HNF1 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  binding motif in the proximal region of *UGT1A8* (at -44bp). To test whether HNF4 $\alpha$  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, HNF4 $\alpha$ , or the combination of CDX2 and HNF4 $\alpha$ . As shown in Figure 1B, the promoters were not transactivated by HNF4 $\alpha$  alone, however they were each transactivated by CDX2. Moreover the combination of HNF4 $\alpha$  and CDX2 synergistically activated all three promoter constructs (Figure 1B).

The ability of HNF4 $\alpha$  and CDX2 to synergize on all three *UGT1A8* promoter constructs suggested that the new predicted HNF4 $\alpha$ RE at -44bp and the CDX2RE contained within the proximal region (-190bp from the TSS) are primarily involved in synergy. Consistent with this idea, mutation of the proximal (-44bp) HNF4 $\alpha$ RE within the 1kb promoter construct ablated the synergistic induction by CDX2 and HNF4 $\alpha$  (Figure 1C); ablation of distal (-811bp) HNF4 $\alpha$ RE had no effect (not shown).

The proximity of the -44bp HNF4 $\alpha$ RE to the previously identified CDX2 binding site (at -70bp) (Gregory et al., 2004), suggested that this CDX2 site mediates the synergy with HNF4 $\alpha$ . To test this idea, we mutated the -70bp CDX2 site within the -190bp *UGT1A8* promoter construct, and tested for induction by CDX2, HNF4 $\alpha$ , or the combination of CDX2 and HNF4 $\alpha$ . Unexpectedly, whilst this mutation prevented induction by CDX2 alone, there was still synergistic activation by CDX2 and HNF4 $\alpha$  (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 HNF4 $\alpha$  (Figure 1E). These data indicate that both the -70bp CDX2RE and the Sp1/Inr element are redundant for HNF4 $\alpha$ /CDX2 synergy.

#### **Identification of a novel composite element that binds both CDX2 and HNF4 $\alpha$**

It was previously reported that HNF4 $\alpha$  interacts with CDX2 (Verzi et al., 2010), thus we considered the possibility that the *UGT1A8* -70bp CDX2 element is redundant for HNF4 $\alpha$ /CDX2 synergy (Figure 1D) because CDX2 might be recruited directly to the *UGT1A8* -44bp HNF4 $\alpha$ RE via interaction with HNF4 $\alpha$ . To examine this possibility, we performed EMSA with a probe corresponding to the -44bp HNF4 $\alpha$ RE. Nuclear extracts from cells transfected with HNF4 $\alpha$  alone, or the combination of HNF4 $\alpha$  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 HNF4 $\alpha$ RE probe was also used as a positive control.

As shown in Figure 2A, HNF4 $\alpha$  formed a strong complex on the consensus HNF4 $\alpha$ RE that was supershifted by HNF4 $\alpha$  antibody (lanes 1, 2). The HNF4 $\alpha$  extract formed a comparatively weaker complex on the -44bp HNF4 $\alpha$ RE probe (lane 3) but mutation of the HNF4 $\alpha$  core recognition motif prevented this complex from forming (lane 4) indicating specificity; blockade of this complex with HNF4 $\alpha$  antibody is also shown in Supplemental Figure 2. Of note, previous studies showed that binding of HNF4 $\alpha$  to the functional upstream HNF4 $\alpha$ REs in *UGT1A9* was also much weaker than to a consensus HNF4 $\alpha$  probe (Gardner- Stephen and Mackenzie, 2007). Extracts containing both HNF4 $\alpha$  and CDX2 formed an additional faster migrating complex on the -44bp HNF4 $\alpha$ RE probe (lanes 7, 8) that they did not form on the consensus HNF4 $\alpha$ RE probe (lanes 5, 6). This complex was not ablated by mutation of the core HNF4 $\alpha$  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 HNF4 $\alpha$ RE 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 HNF4 $\alpha$ RE 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 HNF4 $\alpha$  RE probe independently of HNF4 $\alpha$ . We also examined whether HNF4 $\alpha$  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 HNF4 $\alpha$  (lanes 3, 4). This result indicates that while CDX2 binds to the new element that we have designated the -44bp HNF4 $\alpha$ RE, HNF4 $\alpha$  does not bind to the previously defined -70bp CDX2RE; this finding is consistent with the redundancy of the -70bp CDX2RE for CDX2- HNF4 $\alpha$  synergy (Fig 1D).

Overall, these data suggest that the -44bp HNF4 $\alpha$ RE, which we have identified as mediating a novel synergistic response to HNF4 $\alpha$  and CDX2, binds to both HNF4 $\alpha$  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 HNF4 $\alpha$ RE probe and performed EMSA with extracts containing both HNF4 $\alpha$  and CDX2. As

shown in Figure 3B, mutation of the HNF4 $\alpha$  motif blocked formation of the HNF4 $\alpha$  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 HNF4 $\alpha$  respectively (Figure 3C). The -44bp HNF4 $\alpha$  probe formed both the CDX2 and HNF4 complexes (lane 1); a consensus HNF4 $\alpha$ RE competitor blocked formation of the HNF4 $\alpha$  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 HNF4 $\alpha$  complex (lane 3), whereas the -44bp HNF4 $\alpha$ RE (self) competitor blocked both complexes (lane 4). A -44bp HNF4 $\alpha$ RE competitor with a mutated HNF4 $\alpha$  motif did not block the HNF4 $\alpha$  complex but reduced the CDX2 complex (lane 5), in contrast a -44bp HNF4 $\alpha$ RE competitor with a mutated CDX2 motif had little effect on the CDX2 complex but blocked the HNF4 $\alpha$  complex (lanes 6, 7). These data further confirm that the -44bp HNF4 $\alpha$ RE is a composite of two motifs that likely mediate adjacent binding of HNF4 $\alpha$  and CDX2.

To assess the functional significance of the cryptic CDX2 motif in the -44bp HNF4 $\alpha$ RE, we mutated this motif in the context of the -190bp *UGT1A8* promoter construct (Figure 4A) and assessed activation by CDX2 and HNF4 $\alpha$ . Mutation of the cryptic CDX2 motif inhibited the synergistic activation of the promoter by CDX2 and HNF4 $\alpha$  as effectively as mutating the HNF4 $\alpha$  motif, showing that both motifs are required for synergy (Figure 4B).

#### **A conserved CDX2 /HNF4 $\alpha$ composite binding element regulates *UGT1A8*, -1A9 and -1A10 promoters in Caco2 cells.**

The sequence of the -44bp HNF4 $\alpha$ RE in *UGT1A8* is fully conserved in the *UGT1A9* and -1A10 proximal promoters (Figure 5A). Consistent with this conservation, chromatin immunoprecipitation (ChIP) assays using HNF4 $\alpha$  antibody indicates that this region of all three promoters recruits HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$ ; whereas *UGT1A9* may be regulated differently.

To confirm that *UGT1A10* is regulated in an equivalent manner to *UGT1A8*, we

mutated the equivalent HNF4 $\alpha$ /CDX2 composite element in *UGT1A10* (-47bp HNF4 $\alpha$ RE 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 $\alpha$  than by either factor alone, and the synergy was ablated by mutation of the -47bp HNF4 $\alpha$ RE (Figure 5C).

We next mutated the equivalent HNF4 $\alpha$ /CDX2 composite element in *UGT1A9* (-57bp HNF4 $\alpha$ RE 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 HNF4 $\alpha$  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 HNF4 $\alpha$  was lost when the -57bp HNF4 $\alpha$ RE was mutated, indicating that the HNF4 $\alpha$ /CDX2 composite element in the proximal *UGT1A9* promoter can function similarly to that in *UGT1A8* and *UGT1A10*. The longer -1kb *UGT1A9* promoter was transactivated by HNF4 $\alpha$  alone (unlike the -1kb *UGT1A8* promoter), presumably due to the previously described functional upstream HNF4 $\alpha$  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 HNF4 $\alpha$  reduced activation of the -1kb promoter relative to HNF4 $\alpha$  alone. This latter result may indicate competition between binding of HNF4 $\alpha$  to the upstream HNF4 $\alpha$ REs and the proximal -57bp HNF4 $\alpha$ RE, as discussed later.

Overall, the data presented here indicate that the new HNF4 $\alpha$ /CDX2 composite element can mediate HNF4 $\alpha$ /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 HNF4 $\alpha$ regulate endogenous *UGT1A8*, -1A9 and -1A10 in Caco2 cells**

Given the clear role for CDX2 and HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  and CDX2 siRNAs in reducing their target mRNA and protein levels is shown in

Supplemental Figure 3. As shown in Figure 6A, HNF4 $\alpha$  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 $\alpha$  and CDX2 than Caco2 cells. Both HNF4 $\alpha$  and CDX2 siRNA produced a 40-50% decrease of all three genes. Treatment of cells with the HNF4 $\alpha$  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 and HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  to the *UGT1A8* promoter would be affected by knockdown of endogenous CDX2. We transfected the HNF4 $\alpha$  expression plasmid with either CDX2-siRNA or scrambled control-siRNA, and then performed ChIP using antibodies to CDX2 and HNF4 $\alpha$ . Binding of exogenous HNF4 $\alpha$  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).

HNF4 $\alpha$  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 HNF4 $\alpha$  mRNA levels (Supplemental Figure 3). Interestingly however, HNF4 $\alpha$  siRNA reduced not only HNF4 $\alpha$  levels but also CDX2 levels (Supplemental Figure 3 and 4), and a similar result was seen after treatment of cells with the HNF4 $\alpha$  inhibitor BI6015 (Supplemental Figure 3). The regulation of CDX2 expression by HNF4 $\alpha$  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 HNF4 $\alpha$  interact, thus it is plausible that knocking down HNF4 $\alpha$  affects CDX2 autoregulation.

#### **Regulation of *UGT1A9* by CDX2 and HNF4 $\alpha$ is mechanistically different than regulation of *UGT1A8* and *-1A10*.**

Our data using different length promoter constructs suggests that the regulation of *UGT1A9* by HNF4 $\alpha$  and CDX2 has two distinct components. The composite HNF4 $\alpha$ /CDX2 element shared between *UGT1A8*, *-1A9* and *-1A10* appears to mediate mechanistically similar synergistic regulation of all three proximal promoters. However, the HNF4 $\alpha$  sites

located further upstream in the *UGT1A9* promoter (that are not conserved in *UGT1A8* and *-1A10*) appear to mediate independent regulation of this gene by HNF4 $\alpha$ . 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 HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  together (Figure 7A). CDX2 could not induce *UGT1A9* mRNA in either Caco2 or HepG2 cells, either alone or together with HNF4 $\alpha$ . In contrast, HNF4 $\alpha$  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 HNF4 $\alpha$ . In contrast, HNF4 $\alpha$  can increase *UGT1A9* mRNA expression in a CDX2-independent manner. In further support of these findings, in HepG2 cells, HNF4 $\alpha$  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 HNF4 $\alpha$  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 HNF4 $\alpha$  (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 HNF4 $\alpha$  was weaker, albeit still statistically significant, for all comparisons except for *UGT1A10* and HNF4 $\alpha$  (Supplemental Figure 5)

## **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 HNF1 $\alpha$  were shown to play important roles in intestinal cell expression of *UGT1A8* (Gregory et al., 2004). HNF1 $\alpha$  also regulates oesophageal cell expression of *UGT1A7* in cooperation with HNF4 $\alpha$  (Ehmer et al., 2010). The sole member of this cluster that is

expressed in liver, *UGT1A9*, is regulated in liver cells by *HNF4 $\alpha$*  and this also involves cooperation with *HNF1 $\alpha$*  (Barbier et al., 2005; Gardner-Stephen and Mackenzie, 2007). Recent genome-wide binding studies have revealed that *CDX2* and *HNF4 $\alpha$*  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/HNF4 $\alpha$*  regulatory nexus.

A major finding of our study was the identification of a new composite 12nt element that binds to both *CDX2* and *HNF4 $\alpha$*  in the *UGT1A8-1A10* proximal promoters. *CDX2* and *HNF4 $\alpha$*  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 *HNF4 $\alpha$* , and confirm that the cryptic TATT motif within the element recruits *CDX2* directly. The relative positions of *CDX2* and *HNF4 $\alpha$*  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 *HNF4 $\alpha$*  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 *HNF4 $\alpha$* RE) with extracts containing both *CDX2* and *HNF4 $\alpha$*  produced two distinct complexes that migrated equivalently to the complexes formed with separate *CDX2* and *HNF4 $\alpha$*  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 *HNF4 $\alpha$*  recruitment to the *UGT1A8* proximal promoter requires the presence of *CDX2* is also consistent with the previous report that *CDX2* promotes binding of *HNF4 $\alpha$*  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 *HNF1 $\alpha$*  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 *HNF4 $\alpha$ /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 *HNF4 $\alpha$* . We also showed that the -70bp *CDX2* motif in the *UGT1A8* promoter that mediated *HNF1 $\alpha$ /CDX2* synergy (Gregory et al., 2004), is not involved in *HNF4 $\alpha$ /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 *HNF4 $\alpha$* REs but include the proximal (approximately -57bp) *HNF4 $\alpha$ /CDX2* composite element, we observed the same *HNF4 $\alpha$ /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 *HNF4 $\alpha$* REs. In this model, *HNF4 $\alpha$ /CDX2* heterodimers activate *UGT1A9* through the proximal composite element whilst *HNF4 $\alpha$*  homodimers activate through the upstream *HNF4 $\alpha$* REs (Figure 9B, C). The observation that the -1kb *UGT1A9* promoter construct was activated more by *HNF4 $\alpha$*  alone than by co-expression of *HNF4 $\alpha$*  and *CDX2* (Figure 5D), suggests that the upstream *HNF4 $\alpha$* REs can mediate greater activation than the proximal element.

One observation that is not consistent with the model described above is that co-expression of *HNF4 $\alpha$*  and *CDX2* did not increase levels of endogenous *UGT1A9* mRNA in Caco2 cells (Figure 7). The result implies that overexpressed *HNF4 $\alpha$ /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 *HNF4 $\alpha$*  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 *HNF4 $\alpha$*  protein (a proposed mechanism of action) and we postulate that it may inhibit the ability of *HNF4 $\alpha$*  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 that we are currently pursuing in order to better understand the roles of the distal and proximal *HNF4 $\alpha$*  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 HNF4 $\alpha$  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.

#### 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, Pearnau 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, Radomska-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) HNF4 $\alpha$  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 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.

## 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.

## LEGENDS FOR FIGURES

**Figure 1.** Synergistic regulation of the *UGT1A8* promoter by CDX2 and HNF4 $\alpha$ . **A.** Schematic of the 1kb *UGT1A8* promoter region showing the positions of three predicted CDX2 binding sites, two potential HNF4 $\alpha$  binding sites, and the Sp1/Inr element, +1 indicates the transcription start site (see text). **B.** CDX2 and HNF4 $\alpha$  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) HNF4 $\alpha$  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 HNF4 $\alpha$  binding to the *UGT1A8* -44bp HNF4 $\alpha$ RE **A.** Lanes 1, 2, 5, 6: an HNF4 $\alpha$  consensus probe incubated with extracts containing either HNF4 $\alpha$  or HNF4 $\alpha$ +CDX2; addition of anti-HNF4 $\alpha$  antibody (lane 5) inhibits complex formation. Lanes 3, 4: wildtype or mutated *UGT1A8* -44bp HNF4 $\alpha$  probes incubated with extracts containing HNF4 $\alpha$ . Lanes 7, 8: wildtype or mutated *UGT1A8* -44bp HNF4 $\alpha$  probes incubated with extracts containing HNF4 $\alpha$ +CDX2. **B.** Lanes 1, 2: the *UGT1A8* -44bp HNF4 $\alpha$ RE 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+HNF4 $\alpha$ , without (lanes 1, 3) or with (lanes 2, 4) addition of CDX2 antibody. Lane 5: consensus HNF4 $\alpha$  probe incubated with extracts containing CDX2+HNF4 $\alpha$ .

**Figure 3.** EMSA mutational analysis of HNF4 $\alpha$  and CDX2 binding to the *UGT1A8* -44bp HNF4 $\alpha$ RE. **A.** Sequence of the *UGT1A8* -44bp HNF4 $\alpha$ RE showing the predicted HNF4 $\alpha$  and CDX2 binding motifs. **B.** Lane 1: consensus HNF4 $\alpha$  probe incubated with extracts containing HNF4 $\alpha$ +CDX2. Lane 2: the 44bp HNF4 $\alpha$ RE wildtype probe incubated with extract containing HNF4 $\alpha$ +CDX2. Lanes 3-5: -44bp HNF4 $\alpha$ RE probes with mutation of either the HNF4 $\alpha$  motif (lane 3), or the CDX2 motif (lanes 4, 5) incubated with extracts containing HNF4 $\alpha$ +CDX2. **C.** The -44bp HNF4 $\alpha$ RE wildtype probe incubated with extract containing HNF4 $\alpha$ +CDX2 without (lane 1), or with (lanes 2-7) various unlabelled competitor oligonucleotides. Lane 2: consensus HNF4 $\alpha$ RE competitor. Lane 3: -70bp CDX2RE competitor. Lane 4: wildtype -44bp HNF4 $\alpha$ RE competitor. Lane 5: -44bp HNF4 $\alpha$ RE competitor with mutation of the

HNF4 $\alpha$  motif. Lane 6,7: -44bp HNF4 $\alpha$ RE competitor with mutation of the CDX2 motif.

**Figure 4.** Mutation of either the CDX2 or HNF4 $\alpha$  motif within the *UGT1A8* -44bp HNF4 $\alpha$ RE prevents synergistic promoter activation. **A.** Schematic showing mutations generated in the CDX2 and HNF4 $\alpha$  motifs within the -44bp HNF4 $\alpha$ RE in the *UGT1A8* -190 promoter construct. **B.** Regulation of the *UGT1A8* promoter constructs by CDX2, HNF4 $\alpha$ , or CDX2+HNF4 $\alpha$ . 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 HNF4 $\alpha$ /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 HNF4 $\alpha$ /CDX2 composite binding element. **B.** CHIP- qPCR analysis testing binding of HNF4 $\alpha$  to regions spanning the HNF4 $\alpha$ /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, HNF4 $\alpha$ , or CDX2+HNF4 $\alpha$ . **D.** Regulation of the -0.19kb and -1kb *UGT1A9* promoter constructs by CDX2, HNF4 $\alpha$ , or CDX2+HNF4 $\alpha$ . 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 HNF4 $\alpha$  and/or CDX2. **A.** Transfection of Caco2 cells with either HNF4 $\alpha$  or CDX2 siRNA decreases the level of *UGT1A8*, -*1A9*, and -*1A10* mRNAs. **B.** Transfection of HT29 cells with either HNF4 $\alpha$  siRNA or CDX2 siRNA and measurement of *UGT1A8*, -*1A9*, and -*1A10* mRNA levels. **C.** Treatment of Caco2 cells with HNF4 $\alpha$  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 HNF4 $\alpha$  in hepatic and intestinal cell lines. **A.** Transfection of HepG2 and Caco2 cells with HNF4 $\alpha$  and CDX2 expression plasmids and measurement of *UGT1A8* and -*1A9* mRNA levels. **B.** Transfection of HepG2 cells with HNF4 $\alpha$  siRNA and measurement of HNF4 $\alpha$ , *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 HNF4 $\alpha$  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 HNF4 $\alpha$  (C) in normal colon samples. **D, E.** Correlation of *UGT1A10* with levels of CDX2 (D) and HNF4 $\alpha$  (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.

**Figure 9.** A model for the regulation of *UGT1A8/1A10* and *UGT1A9* by HNF4 $\alpha$  and CDX2. **A.** In the *UGT1A8/1A10* proximal promoters, a two part intestinal module that includes (I) the new HNF4 $\alpha$ /CDX2 composite element (-44bp in *UGT1A8*), and (II) the previously defined CDX2 (-70bp) and HNF1 $\alpha$  sites (-100bp). When CDX2 is high (e.g. intestine), the proximal HNF4 $\alpha$ /CDX2 composite element (I) recruits HNF4 $\alpha$ /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 HNF4 $\alpha$ REs. When CDX2 is high (e.g. intestine), HNF4 $\alpha$  forms heterodimers with CDX2 that bind the intestinal module; these may also cooperate with HNF1 $\alpha$  (green bracket). **C.** When CDX2 is low/absent (e.g. liver), HNF4 $\alpha$  forms homodimers that bind the hepatic module, these may also cooperate with HNF1 $\alpha$  (green bracket). Chromatin architecture may help determine the relative accessibility of these modules.

Fig 1

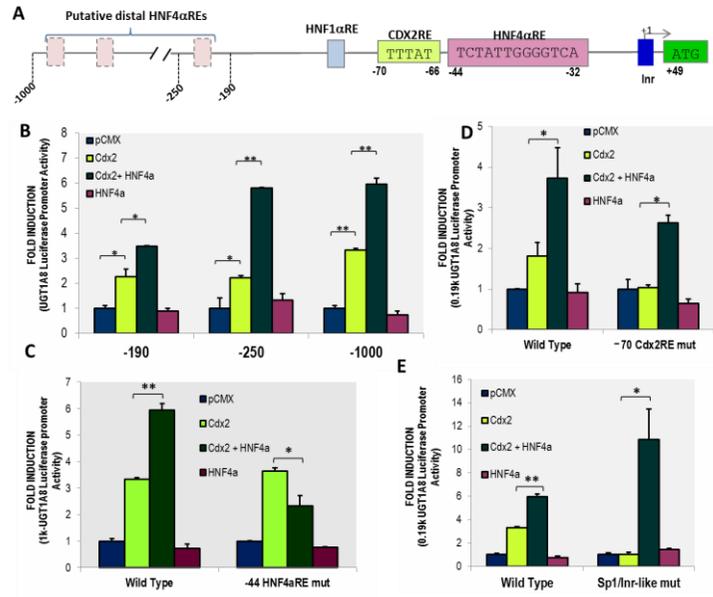


Fig 2

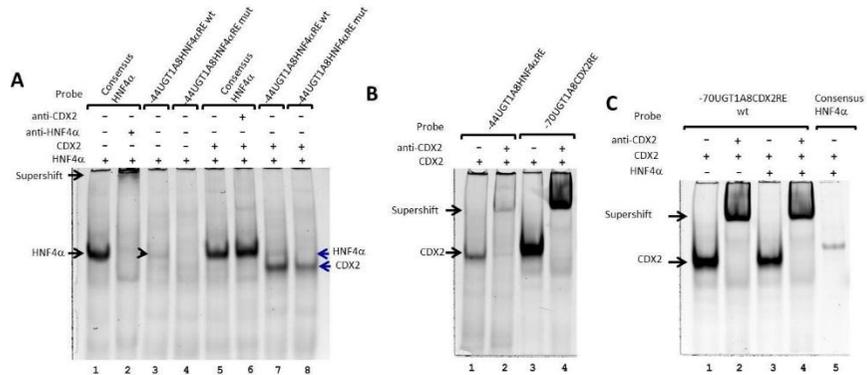
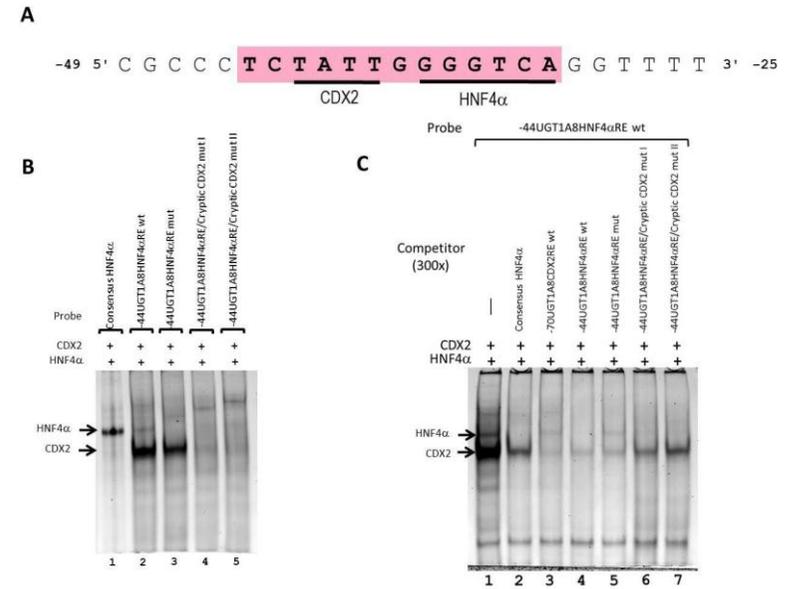
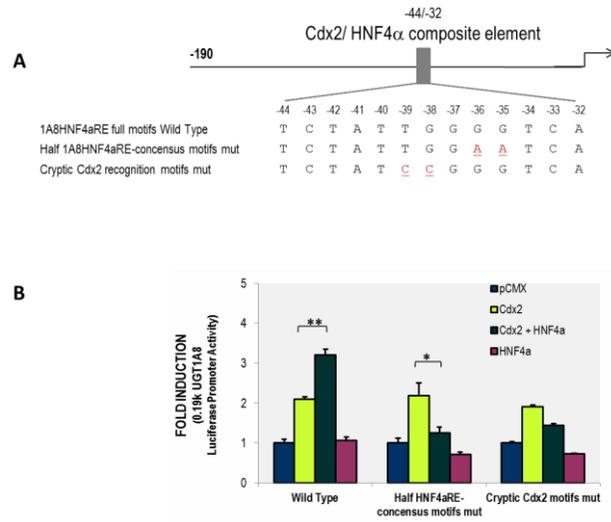


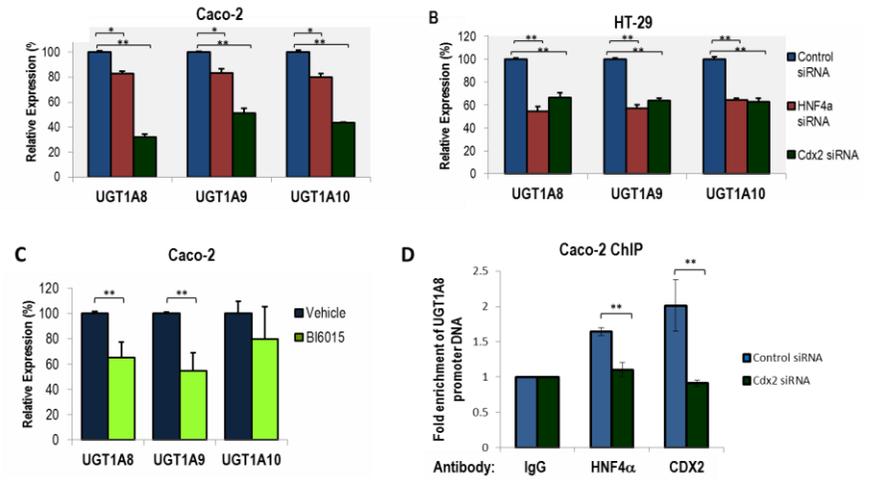
Fig 3



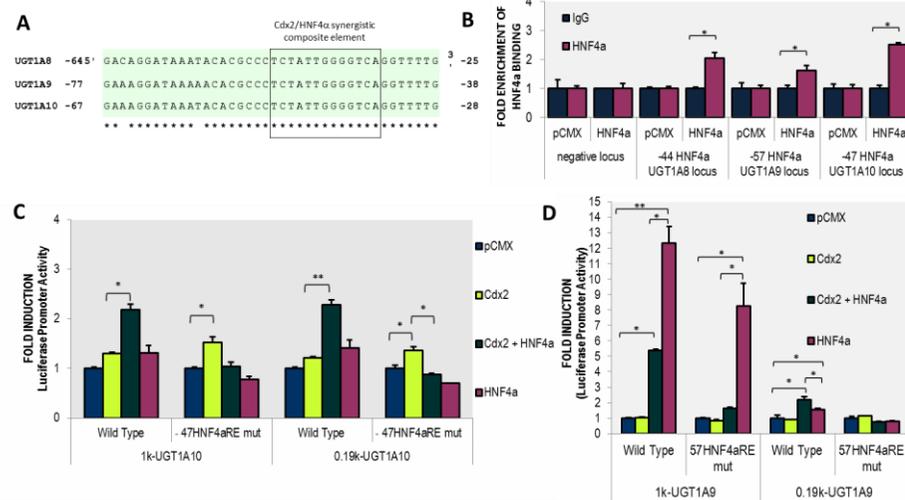
**Fig 4**



**Fig 6**



**Fig 5**



**Fig 7**

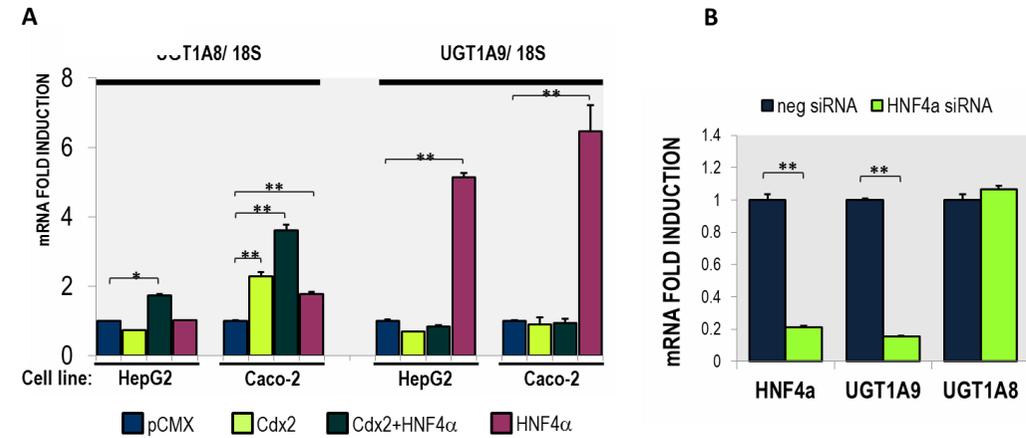


Fig 8

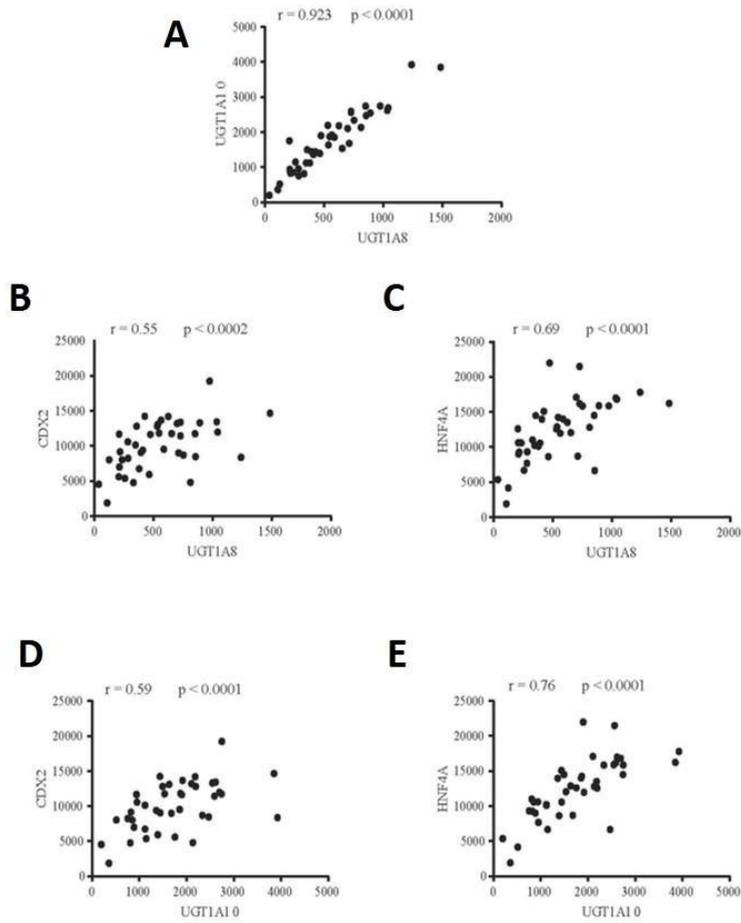
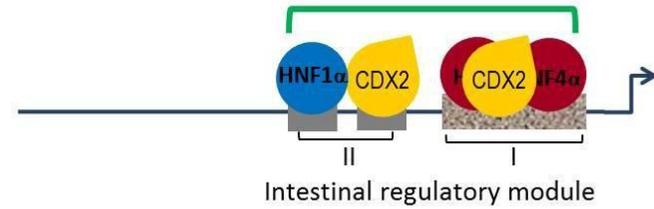
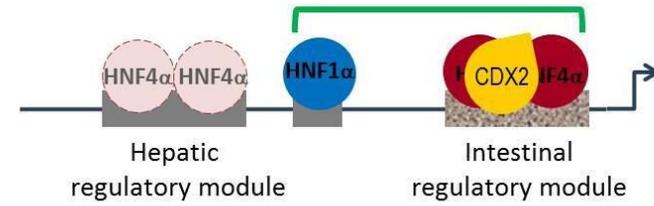


Fig 9

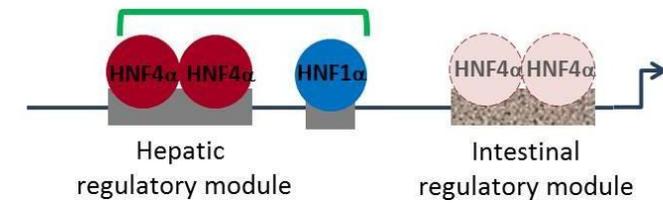
**A** UGT1A8/10 – intestinal expression (High level of CDX2)



**B** UGT1A9 – intestinal expression (High level of CDX2)



**C** UGT1A9 – hepatic expression (Absence/low level of CDX2)



## REFERENCES

- AHMADIAN, M., SUH, J. M., HAH, N., LIDDLE, C., ATKINS, A. R., DOWNES, M. & EVANS, R. M. 2013. PPAR [gamma] signaling and metabolism: the good, the bad and the future. *Nature medicine*, 99, 557-566.
- ANTONIO, L., XU, J., LITTLE, J. M., BURCHELL, B., MAGDALOU, J. & RADOMINSKA-PANDYA, A. 2003. Glucuronidation of catechols by human hepatic, gastric, and intestinal microsomal UDP-glucuronosyltransferases (UGT) and recombinant UGT1A6, UGT1A9, and UGT2B7. *Archives of biochemistry and biophysics*, 411, 251-261.
- ANZENBACHER, P. & ZANGER, U. M. 2012. *Metabolism of Drugs and Other Xenobiotics*, Hoboken, GERMANY, John Wiley & Sons, Incorporated.
- APRILE, S., DEL GROSSO, E. & GROSA, G. 2010. Identification of the human UDP-glucuronosyltransferases involved in the glucuronidation of combretastatin A-4. *Drug Metabolism and Disposition*, 38, 1141-1146.
- ARGIKAR, U. A. 2012. Unusual glucuronides. *Drug Metabolism and Disposition*, 40, 1239-1251.
- BANERJEE, S., LI, Y., WANG, Z. & SARKAR, F. H. 2008. Multi-targeted therapy of cancer by genistein. *Cancer letters*, 269, 226-242.
- BARBIER, O., GIRARD, H., INOUE, Y., DUEZ, H., VILLENEUVE, L., KAMIYA, A., FRUCHART, J.-C., GUILLEMETTE, C., GONZALEZ, F. J. & STAELS, B. 2005. Hepatic expression of the UGT1A9 gene is governed by hepatocyte nuclear factor 4 $\alpha$ . *Molecular pharmacology*, 67, 241-249.
- BARBIER, O., VILLENEUVE, L., BOCHER, V., FONTAINE, C., TORRA, I. P., DUHEM, C., KOSYKH, V., FRUCHART, J.-C., GUILLEMETTE, C. & STAELS, B. 2003. The UDP-glucuronosyltransferase 1A9 enzyme is a peroxisome proliferator-activated receptor  $\alpha$  and  $\gamma$  target gene. *Journal of Biological Chemistry*, 278, 13975-13983.
- BARKER, N. 2014. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature reviews Molecular cell biology*, 15, 19-33.
- BARKER, N., VAN ES, J. H., KUIPERS, J., KUJALA, P., VAN DEN BORN, M., COZIJNSEN, M., HAEGEBARTH, A., KORVING, J., BEGTHEL, H. & PETERS, P. J. 2007. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, 449, 1003-1007.
- BARKHEM, T., CARLSSON, B., NILSSON, Y., ENMARK, E., GUSTAFSSON, J.-Å. & NILSSON, S. 1998. Differential response of estrogen receptor  $\alpha$  and estrogen receptor  $\beta$  to partial estrogen agonists/antagonists. *Molecular pharmacology*, 54, 105-112.
- BARROS, R., DA COSTA, L. T., PINTO-DE-SOUSA, J., DULUC, I., FREUND, J.-N., DAVID, L. & ALMEIDA, R. 2011. CDX2 autoregulation in human intestinal metaplasia of the stomach: impact on the stability of the phenotype. *Gut*, 60, 290-298.
- BASU, S., KAUFMAN, B. & ROSEMAN, S. 1968. Enzymatic synthesis of ceramide-glucose and ceramide-lactose by glycosyltransferases from embryonic chicken brain. *Journal of Biological Chemistry*, 243, 5802-5804.
- BECK, F., CHAWENGSAKSOPHAK, K., WARING, P., PLAYFORD, R. J. & FURNESS, J. B. 1999. Reprogramming of intestinal differentiation and intercalary regeneration in Cdx2 mutant mice. *Proceedings of the National Academy of Sciences*, 96, 7318-7323.

- BECK, F., ERLER, T., RUSSELL, A. & JAMES, R. 1995. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. *Developmental Dynamics*, 204, 219-227.
- BENOIT, Y. D., PARÉ, F., FRANCOEUR, C., JEAN, D., TREMBLAY, E., BOUDREAU, F., ESCAFFIT, F. & BEAULIEU, J.-F. 2010. Cooperation between HNF-1 $\alpha$ , Cdx2, and GATA-4 in initiating an enterocytic differentiation program in a normal human intestinal epithelial progenitor cell line. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 298, G504-G517.
- BERNARD, O. & GUILLEMETTE, C. 2004. The main role of UGT1A9 in the hepatic metabolism of mycophenolic acid and the effects of naturally occurring variants. *Drug Metabolism and Disposition*, 32, 775-778.
- BERNARD, O., TOJCIC, J., JOURNAULT, K., PERUSSE, L. & GUILLEMETTE, C. 2006. Influence of nonsynonymous polymorphisms of UGT1A8 and UGT2B7 metabolizing enzymes on the formation of phenolic and acyl glucuronides of mycophenolic acid. *Drug metabolism and disposition*, 34, 1539-1545.
- BEVILACQUA, M., FANIELLO, M., QUARESIMA, B., TIANO, M., PIGNATA, S., RUSSO, T., CIMINO, F. & COSTANZO, F. 1995. Transcriptional activation of the H-ferritin gene in differentiated Caco-2 cells parallels a change in the activity of the nuclear factor Bbf. *Biochemical Journal*, 311, 769-773.
- BHAGWAT, S., HAYTOWITZ, D. B. & HOLDEN, J. M. 2014. USDA database for the flavonoid content of selected foods, Release 3.1. *US Department of Agriculture: Beltsville, MD, USA*.
- BIGGIN, M. D. 2011. Animal transcription networks as highly connected, quantitative continua. *Developmental cell*, 21, 611-626.
- BLAUT, SCHOEFER & BRAUNE 2003. Transformation of flavonoids by intestinal microorganisms. *International journal for vitamin and nutrition research*, 73, 79-87.
- BOCK, K. W. 2010. Functions and transcriptional regulation of adult human hepatic UDP-glucuronosyl-transferases (UGTs): mechanisms responsible for interindividual variation of UGT levels. *Biochem Pharmacol*, 80, 771-7.
- BOCK, K. W. 2012. Human UDP-glucuronosyltransferases: Feedback loops between substrates and ligands of their transcription factors. *Biochemical pharmacology*, 84, 1000-1006.
- BOCK, K. W. & KÖHLE, C. 2009. Topological aspects of oligomeric UDP-glucuronosyltransferases in endoplasmic reticulum membranes: advances and open questions. *Biochemical pharmacology*, 77, 1458-1465.
- BOUDREAU, F., RINGS, E. H., VAN WERING, H. M., KIM, R. K., SWAIN, G. P., KRASINSKI, S. D., MOFFETT, J., GRAND, R. J., SUH, E. R. & TRABER, P. G. 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. *Journal of Biological Chemistry*, 277, 31909-31917.
- BOURASSA, M. W., ALIM, I., BULTMAN, S. J. & RATAN, R. R. 2016. Butyrate, neuroepigenetics and the gut microbiome: can a high fiber diet improve brain health? *Neuroscience letters*, 625, 56-63.

- BOYD, M., BRESSENDORFF, S., MØLLER, J., OLSEN, J. & TROELSEN, J. T. 2009. Mapping of HNF4 $\alpha$  target genes in intestinal epithelial cells. *BMC gastroenterology*, 9, 68.
- BRABLETZ, T., SPADERNA, S., KOLB, J., HLUBEK, F., FALLER, G., BRUNS, C. J., JUNG, A., NENTWICH, J., DULUC, I. & DOMON-DELL, C. 2004. Down-Regulation of the Homeodomain Factor Cdx2 in Colorectal Cancer by Collagen Type I. *Cancer research*, 64, 6973-6977.
- BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, 72, 248-254.
- BRAHE, L. K., ASTRUP, A. & LARSEN, L. H. 2013. Is butyrate the link between diet, intestinal microbiota and obesity-related metabolic diseases? *Obesity reviews*, 14, 950-959.
- BREWER, G. J. & COTMAN, C. W. 1989. Survival and growth of hippocampal neurons in defined medium at low density: advantages of a sandwich culture technique or low oxygen. *Brain research*, 494, 65-74.
- BRILL, S. S., FURIMSKY, A. M., HO, M. N., FURNISS, M. J., LI, Y., GREEN, A. G., GREEN, C. E., IYER, L. V., BRADFORD, W. W. & KAPETANOVIC, I. M. 2006. Glucuronidation of trans-resveratrol by human liver and intestinal microsomes and UGT isoforms. *Journal of pharmacy and pharmacology*, 58, 469-479.
- BUCKLEY, D. B. & KLAASSEN, C. D. 2007. Tissue-and gender-specific mRNA expression of UDP-glucuronosyltransferases (UGTs) in mice. *Drug metabolism and disposition*, 35, 121-127.
- BUCZACKI, S. J., ZECCHINI, H. I., NICHOLSON, A. M., RUSSELL, R., VERMEULEN, L., KEMP, R. & WINTON, D. J. 2013. Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature*, 495, 65-69.
- CAI, H., NGUYEN, N., PETERKIN, V., YANG, Y. S., HOTZ, K., LA PLACA, D. B., CHEN, S., TUKEY, R. H. & STEVENS, J. C. 2010. A humanized UGT1 mouse model expressing the UGT1A1\* 28 allele for assessing drug clearance by UGT1A1 dependent glucuronidation. *Drug Metabolism and Disposition*, dmd. 109.030130.
- CANDIDO, E. P. M., REEVES, R. & DAVIE, J. R. 1978. Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell*, 14, 105-113.
- CASANOVA, M., YOU, L., GAIDO, K. W., ARCHIBEQUE-ENGLE, S., JANSZEN, D. B. & HECK, H. D. A. 1999. Developmental effects of dietary phytoestrogens in Sprague-Dawley rats and interactions of genistein and daidzein with rat estrogen receptors alpha and beta in vitro. *Toxicological sciences: an official journal of the Society of Toxicology*, 51, 236-244.
- CATTIN, A.-L., LE BEYEC, J., BARREAU, F., SAINT-JUST, S., HOULLIER, A., GONZALEZ, F. J., ROBINE, S., PINÇON-RAYMOND, M., CARDOT, P. & LACASA, M. 2009. Hepatocyte nuclear factor 4 $\alpha$ , a key factor for homeostasis, cell architecture, and barrier function of the adult intestinal epithelium. *Molecular and cellular biology*, 29, 6294-6308.
- CHACKO, B. K., CHANDLER, R. T., D'ALESSANDRO, T. L., MUNDHEKAR, A., KHOO, N. K., BOTTING, N., BARNES, S. & PATEL, R. P. 2007. Anti-inflammatory effects of isoflavones are dependent on flow and human endothelial cell PPAR $\gamma$ . *The Journal of nutrition*, 137, 351-356.

- CHACKO, B. K., CHANDLER, R. T., MUNDHEKAR, A., KHOO, N., PRUITT, H. M., KUCIK, D. F., PARKS, D. A., KEVIL, C. G., BARNES, S. & PATEL, R. P. 2005. Revealing anti-inflammatory mechanisms of soy isoflavones by flow: modulation of leukocyte-endothelial cell interactions. *American Journal of Physiology-Heart and Circulatory Physiology*, 289, H908-H915.
- CHANDRASEKARAN, C. & GORDON, J. I. 1993. Cell lineage-specific and differentiation-dependent patterns of CCAAT/enhancer binding protein alpha expression in the gut epithelium of normal and transgenic mice. *Proceedings of the National Academy of Sciences*, 90, 8871-8875.
- CHANG, P. V., HAO, L., OFFERMANN, S. & MEDZHITOV, R. 2014. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proceedings of the National Academy of Sciences*, 111, 2247-2252.
- CHAWENGSAKSOPHAK, K., DE GRAAFF, W., ROSSANT, J., DESCHAMPS, J. & BECK, F. 2004. Cdx2 is essential for axial elongation in mouse development. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 7641-7645.
- CHEN, F., WANG, M., O'CONNOR, J. P., HE, M., TRIPATHI, T. & HARRISON, L. E. 2003. Phosphorylation of PPAR $\gamma$  via active ERK1/2 leads to its physical association with p65 and inhibition of NF- $\kappa$ B. *Journal of cellular biochemistry*, 90, 732-744.
- CHEN, L., NECELA, B. M., SU, W., YANAGISAWA, M., ANASTASIADIS, P. Z., FIELDS, A. P. & THOMPSON, E. A. 2006. Peroxisome proliferator-activated receptor  $\gamma$  promotes epithelial to mesenchymal transformation by Rho GTPase-dependent activation of ERK1/2. *Journal of Biological Chemistry*, 281, 24575-24587.
- CHEN, S., BEATON, D., NGUYEN, N., SENEKEO-EFFENBERGER, K., BRACE-SINNOKRAK, E., ARGIKAR, U., REMMEL, R. P., TROTTIER, J., BARBIER, O. & RITTER, J. K. 2005. Tissue-specific, inducible, and hormonal control of the human UDP-glucuronosyltransferase-1 (UGT1) locus. *Journal of Biological Chemistry*, 280, 37547-37557.
- CHEN, W. S., MANOVA, K., WEINSTEIN, D. C., DUNCAN, S. A., PLUMP, A. S., PREZIOSO, V. R., BACHVAROVA, R. F. & DARNELL, J. 1994. Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes & development*, 8, 2466-2477.
- CHENG, Z., RADOMINSKA-PANDYA, A. & TEPHLY, T. R. 1998. Cloning and expression of human UDP-glucuronosyltransferase (UGT) 1A8. *Archives of Biochemistry and Biophysics*, 356, 301-305.
- CHENG, Z., RADOMINSKA-PANDYA, A. & TEPHLY, T. R. 1999. Studies on the substrate specificity of human intestinal UDP-glucuronosyltransferases 1A8 and 1A10. *Drug Metabolism and Disposition*, 27, 1165-1170.
- CHOI, Y. H., ZHANG, L., LEE, W. & PARK, K. 1998. Genistein-induced G2/M arrest is associated with the inhibition of cyclin B1 and the induction of p21 in human breast carcinoma cells. *International journal of oncology*, 13, 391-397.
- CLAUS, S. P., ELLERO, S. L., BERGER, B., KRAUSE, L., BRUTTIN, A., MOLINA, J., PARIS, A., WANT, E. J., DE WAZIERS, I. & CLOAREC, O. 2011. Colonization-induced host-gut microbial metabolic interaction. *MBio*, 2, e00271-10.
- CLEVERS, H. 2013. The intestinal crypt, a prototype stem cell compartment. *Cell*, 154, 274-284.

- CLIFFORD, M. 2004. Diet-derived phenols in plasma and tissues and their implications for health. *Planta medica*, 70, 1103-1114.
- COLLIER, A. C., YAMAUCHI, Y., SATO, B. L., ROUGÉE, L. R. & WARD, M. A. 2014. UDP-glucuronosyltransferase 1a enzymes are present and active in the mouse blastocyst. *Drug Metabolism and Disposition*, 42, 1921-1925.
- COURT, M. H., ZHANG, X., DING, X., YEE, K. K., HESSE, L. M. & FINEL, M. 2012. Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica*, 42, 266-277.
- CREAMER, B., SHORTER, R. & BAMFORTH, J. 1961. The turnover and shedding of epithelial cells Part I The turnover in the gastro-intestinal tract. *Gut*, 2, 110-116.
- CSALA, M., STAINES, A. G., BÁNHEGYI, G., MANDL, J., COUGHTRIE, M. W. & BURCHELL, B. 2004. Evidence for multiple glucuronide transporters in rat liver microsomes. *Biochemical pharmacology*, 68, 1353-1362.
- DAHL, J. A. & COLLAS, P. 2008. A rapid micro chromatin immunoprecipitation assay (ChIP). *Nature protocols*, 3, 1032-1045.
- DANG, Z. C. 2009. Dose-dependent effects of soy phyto-oestrogen genistein on adipocytes: mechanisms of action. *Obesity Reviews*, 10, 342-349.
- DANG, Z. C. & LOWIK, C. 2005. Dose-dependent effects of phytoestrogens on bone. *Trends in Endocrinology & Metabolism*, 16, 207-213.
- DAVIDSON, E. H. & LEVINE, M. S. 2008. Properties of developmental gene regulatory networks. *Proceedings of the National Academy of Sciences*, 105, 20063-20066.
- DAVIE, J. R. 2003. Inhibition of histone deacetylase activity by butyrate. *The Journal of nutrition*, 133, 2485S-2493S.
- DELAGE, B. & DASHWOOD, R. H. 2008. Dietary manipulation of histone structure and function. *Annu. Rev. Nutr.*, 28, 347-366.
- DEORA, A. A., DIAZ, F., SCHREINER, R. & RODRIGUEZ-BOULAN, E. 2007. Efficient electroporation of DNA and protein into confluent and differentiated epithelial cells in culture. *Traffic*, 8, 1304-1312.
- DIGNASS, A. U. & STURM, A. 2001. Peptide growth factors in the intestine. *European journal of gastroenterology & hepatology*, 13, 763-770.
- DIRADOURIAN, C., GIRARD, J. & PÉGORIER, J.-P. 2005. Phosphorylation of PPARs: from molecular characterization to physiological relevance. *Biochimie*, 87, 33-38.
- DROST, J., VAN JAARVELD, R. H., PONSIOEN, B., ZIMBERLIN, C., VAN BOXTEL, R., BUIJS, A., SACHS, N., OVERMEER, R. M., OFFERHAUS, G. J. & BEGTHEL, H. 2015. Sequential cancer mutations in cultured human intestinal stem cells. *Nature*, 521, 43.
- DU, Q., WANG, Y., LIU, C., WANG, H., FAN, H., LI, Y., WANG, J., ZHANG, X., LU, J. & JI, H. 2016. Chemopreventive activity of GEN-27, a genistein derivative, in colitis-associated cancer is mediated by p65-CDX2- $\beta$ -catenin axis. *Oncotarget*, 7, 17870.

- DUBEY, R. K. & SINGH, J. 1988. Localization and characterization of drug-metabolizing enzymes along the villus-crypt surface of the rat small intestine—II: Conjugases. *Biochemical pharmacology*, 37, 177-184.
- DUNCAN, S. A., MANOVA, K., CHEN, W. S., HOODLESS, P., WEINSTEIN, D. C., BACHVAROVA, R. F. & DARNELL, J. 1994. Expression of transcription factor HNF-4 in the extraembryonic endoderm, gut, and nephrogenic tissue of the developing mouse embryo: HNF-4 is a marker for primary endoderm in the implanting blastocyst. *Proceedings of the National Academy of Sciences*, 91, 7598-7602.
- DUNCAN, S. A., NAGY, A. & CHAN, W. 1997. Murine gastrulation requires HNF-4 regulated gene expression in the visceral endoderm: tetraploid rescue of Hnf-4 (-/-) embryos. *Development*, 124, 279-287.
- DUPREY, P., CHOWDHURY, K., DRESSLER, G., BALLING, R., SIMON, D., GUENET, J. & GRUSS, P. 1988. A mouse gene homologous to the Drosophila gene caudal is expressed in epithelial cells from the embryonic intestine. *Genes & development*, 2, 1647-1654.
- DURA, P., SALOMON, J., TE MORSCH, R. H., ROELOFS, H. M., KRISTINSSON, J. O., WOBBER, T., WITTEMAN, B. J., TAN, A., DRENTH, J. P. & PETERS, W. H. 2012. High enzyme activity UGT1A1 or low activity UGT1A8 and UGT2B4 genotypes increase esophageal cancer risk. *International journal of oncology*, 40, 1789.
- DUTTON, G. J. 1980. *Glucuronidation of drugs and other compounds* [Online]. Boca Raton, Fla.: CRC Press.
- ENGLE, M., GOETZ, G. & ALPERS, D. 1998. Caco-2 cells express a combination of colonocyte and enterocyte phenotypes. *Journal of cellular physiology*, 174, 362-369.
- EVANS, W. E. & RELLING, M. V. 1999. Pharmacogenomics: translating functional genomics into rational therapeutics. *science*, 286, 487-491.
- FANG, Y., CAO, W., XIA, M., PAN, S. & XU, X. 2017. Study of Structure and Permeability Relationship of Flavonoids in Caco-2 Cells. *Nutrients*, 9.
- FINEL, M. & KURKELA, M. 2008. The UDP-glucuronosyltransferases as oligomeric enzymes. *Current drug metabolism*, 9, 70-76.
- FOGH, J., WRIGHT, W. C. & LOVELESS, J. D. 1977. Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *Journal of the National Cancer Institute*, 58, 209-214.
- FOK, K. L., CHUNG, C. M., YI, S. Q., JIANG, X., SUN, X., CHEN, H., CHEN, Y. C., KUNG, H.-F., TAO, Q. & DIAO, R. 2012. STK31 maintains the undifferentiated state of colon cancer cells. *Carcinogenesis*, 33, 2044-2053.
- FUJII, M., MATANO, M., NANKI, K. & SATO, T. 2015. Efficient genetic engineering of human intestinal organoids using electroporation. *Nature protocols*, 10, 1474.
- GAGNÉ, J.-F., MONTMINY, V., BELANGER, P., JOURNAULT, K., GAUCHER, G. & GUILLEMETTE, C. 2002. Common human UGT1A polymorphisms and the altered metabolism of irinotecan active metabolite 7-ethyl-10-hydroxycamptothecin (SN-38). *Molecular pharmacology*, 62, 608-617.
- GALIJATOVIC, A., OTAKE, Y., WALLE, U. K. & WALLE, T. 2001. Induction of UDP-glucuronosyltransferase UGT1A1 by the flavonoid chrysin in Caco-2 cells—potential role in carcinogen bioinactivation. *Pharmaceutical research*, 18, 374-379.

- GAO, N., WHITE, P. & KAESTNER, K. H. 2009. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Developmental cell*, 16, 588-599.
- GAO, S. & HU, M. 2010. Bioavailability challenges associated with development of anti-cancer phenolics. *Mini reviews in medicinal chemistry*, 10, 550-567.
- GARDNER-STEPHEN, D. A. & MACKENZIE, P. I. 2007. Hepatocyte nuclear factor1 transcription factors are essential for the UDP-glucuronosyltransferase 1A9 promoter response to hepatocyte nuclear factor 4 $\alpha$ . *Pharmacogenetics and genomics*, 17, 25-36.
- GARDNER-STEPHEN, D. A. & MACKENZIE, P. I. 2008. Liver-enriched transcription factors and their role in regulating UDP glucuronosyltransferase gene expression. *Current drug metabolism*, 9, 439-452.
- GARRISON, W. D., BATTLE, M. A., YANG, C., KAESTNER, K. H., SLADEK, F. M. & DUNCAN, S. A. 2006. Hepatocyte nuclear factor 4 $\alpha$  is essential for embryonic development of the mouse colon. *Gastroenterology*, 130, 19. e1-19. e.
- GHARAVI, N. & EL-KADI, A. O. 2005. tert-Butylhydroquinone is a novel aryl hydrocarbon receptor ligand. *Drug metabolism and disposition*, 33, 365-372.
- GIBSON, G. R. 2004. Fibre and effects on probiotics (the prebiotic concept). *Clinical Nutrition Supplements*, 1, 25-31.
- GÓMEZ, S., DEL MONT LLOSAS, M., VERDÚ, J., ROURA, S., LLORETA, J., FABRE, M. & DE HERREROS, A. G. A. 1999. Independent regulation of adherens and tight junctions by tyrosine phosphorylation in Caco-2 cells. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1452, 121-132.
- GONG, Q.-H., CHO, J. W., HUANG, T., POTTER, C., GHOLAMI, N., BASU, N. K., KUBOTA, S., CARVALHO, S., PENNINGTON, M. W. & OWENS, I. S. 2001. Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics and Genomics*, 11, 357-368.
- GONZALEZ, F. J. 2008. Regulation of hepatocyte nuclear factor 4 $\alpha$ -mediated transcription. *Drug metabolism and pharmacokinetics*, 23, 2-7.
- GRABINGER, T., LUKS, L., KOSTADINOVA, F., ZIMBERLIN, C., MEDEMA, J. P., LEIST, M. & BRUNNER, T. 2014. Ex vivo culture of intestinal crypt organoids as a model system for assessing cell death induction in intestinal epithelial cells and enteropathy. *Cell death & disease*, 5, e1228.
- GREGORY, P. A. 2004. *Transcriptional Regulation of UDP-Glucuronosyltransferases in Extrahepatic Tissues*. PhD, Flinders University of South Australia.
- GREGORY, P. A., GARDNER-STEPHEN, D. A., LEWINSKY, R. H., DUNCLIFFE, K. N. & MACKENZIE, P. I. 2003. Cloning and Characterization of the Human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 Gene Promoters DIFFERENTIAL REGULATION THROUGH AN INITIATOR-LIKE REGION. *Journal of Biological Chemistry*, 278, 36107-36114.
- GREGORY, P. A., LEWINSKY, R. H., GARDNER-STEPHEN, D. A. & MACKENZIE, P. I. 2004a. Coordinate regulation of the human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 genes by hepatocyte nuclear factor 1 $\alpha$  and the caudal-related homeodomain protein 2. *Molecular pharmacology*, 65, 953-963.

- GREGORY, P. A., LEWINSKY, R. H., GARDNER-STEPHEN, D. A. & MACKENZIE, P. I. 2004b. Regulation of UDP glucuronosyltransferases in the gastrointestinal tract. *Toxicology and applied pharmacology*, 199, 354-363.
- GUILLEMETTE, C. 2003. Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *The pharmacogenomics journal*, 3, 136-158.
- GUILLEMETTE, C., LÉVESQUE, E., HARVEY, M., BELLEMARE, J. & MENARD, V. 2010. UGT genomic diversity: beyond gene duplication. *Drug metabolism reviews*, 42, 24-44.
- GUILLEMETTE, C., LÉVESQUE, É. & ROULEAU, M. 2014. Pharmacogenomics of human uridine diphospho-glucuronosyltransferases (UGTs) and clinical implications. *Clinical Pharmacology & Therapeutics*, 96, 324-339.
- GÜNTHER, C., MARTINI, E., WITTKOPF, N., AMANN, K., WEIGMANN, B., NEUMANN, H., WALDNER, M. J., HEDRICK, S. M., TENZER, S. & NEURATH, M. F. 2011. Caspase-8 regulates TNF-[agr]-induced epithelial necroptosis and terminal ileitis. *Nature*, 477, 335-339.
- HALBLEIB, J. M., SÄÄF, A. M., BROWN, P. O. & NELSON, W. J. 2007. Transcriptional modulation of genes encoding structural characteristics of differentiating enterocytes during development of a polarized epithelium in vitro. *Molecular biology of the cell*, 18, 4261-4278.
- HANAHAHAN, D., JESSEE, J. & BLOOM, F. R. 1991. [4] Plasmid transformation of Escherichia coli and other bacteria. *Methods in enzymology*, 204, 63-113.
- HANSEN, A. J., LEE, Y.-H., GONZALEZ, F. J. & MACKENZIE, P. I. 1997. HNF1  $\alpha$  Activates the Rat UDP Glucuronosyltransferase UGT2B1 Gene Promoter. *DNA and cell biology*, 16, 207-214.
- HANSEN, A. J., LEE, Y.-H., STERNECK, E., GONZALEZ, F. J. & MACKENZIE, P. I. 1998. C/EBP $\alpha$  Is a Regulator of the UDP Glucuronosyltransferase UGT2B1 Gene. *Molecular Pharmacology*, 53, 1027-1033.
- HANSEN, S. K., PÁRRIZAS, M., JENSEN, M. L., PRUHOVA, S., EK, J., BOJ, S. F., JOHANSEN, A., MAESTRO, M. A., RIVERA, F. & EIBERG, H. 2002. Genetic evidence that HNF-1 $\alpha$ -dependent transcriptional control of HNF-4 $\alpha$  is essential for human pancreatic  $\beta$  cell function. *The Journal of clinical investigation*, 110, 827.
- HARAMIS, A.-P. G., BEGTHEL, H., VAN DEN BORN, M., VAN ES, J., JONKHEER, S., OFFERHAUS, G. J. A. & CLEVERS, H. 2004. De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine. *Science*, 303, 1684-1686.
- HERRING, B. P., KRIEGEL, A. M. & HOGGATT, A. M. 2001. Identification of Barx2b, a serum response factor-associated homeodomain protein. *Journal of Biological Chemistry*, 276, 14482-14489.
- HINOI, T., LUCAS, P. C., KUICK, R., HANASH, S., CHO, K. R. & FEARON, E. R. 2002. CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. *Gastroenterology*, 123, 1565-1577.
- HINOI, T., TANI, M., LUCAS, P. C., CACA, K., DUNN, R. L., MACRI, E., LODA, M., APPELMAN, H. D., CHO, K. R. & FEARON, E. R. 2001. Loss of CDX2 expression and microsatellite instability are prominent features of large cell minimally differentiated carcinomas of the colon. *The American journal of pathology*, 159, 2239-2248.

- HIURA, Y., SATSU, H., HAMADA, M. & SHIMIZU, M. 2014. Analysis of flavonoids regulating the expression of UGT1A1 via xenobiotic receptors in intestinal epithelial cells. *Biofactors*, 40, 336-345.
- HU, D. G., MEECH, R., MCKINNON, R. A. & MACKENZIE, P. I. 2014a. Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug metabolism reviews*, 46, 421-458.
- HU, D. G., ROGERS, A. & MACKENZIE, P. I. 2014b. Epirubicin upregulates UDP glucuronosyltransferase 2B7 expression in liver cancer cells via the p53 pathway. *Molecular pharmacology*, 85, 887-897.
- HUANG, Y.-H., GALIJATOVIC, A., NGUYEN, N., GESKE, D., BEATON, D., GREEN, J., GREEN, M., PETERS, W. H. & TUKEY, R. H. 2002. Identification and functional characterization of UDP-glucuronosyltransferases UGT1A8\* 1, UGT1A8\* 2 and UGT1A8\* 3. *Pharmacogenetics and Genomics*, 12, 287-297.
- HUGHES, C. S., POSTOVIT, L. M. & LAJOIE, G. A. 2010. Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics*, 10, 1886-1890.
- HUGHES, T., SASAK, W. V., ORDOVAS, J., FORTE, T., LAMON-FAVA, S. & SCHAEFER, E. 1987. A novel cell line (Caco-2) for the study of intestinal lipoprotein synthesis. *Journal of Biological Chemistry*, 262, 3762-3767.
- HUIN, C., SCHOHN, H., HATIER, R., BENTEJAC, M., ANTUNES, L., PLÉNAT, F., BUGAUT, M. & DAUÇA, M. 2002. Expression of peroxisome proliferator-activated receptors alpha and gamma in differentiating human colon carcinoma Caco-2 cells. *Biology of the Cell*, 94, 15-27.
- HUTSON, T. H., KATHE, C., MENEZES, S. C., ROONEY, M.-C., BUELER, H. & MOON, L. D. F. 2014. The use of an adeno-associated viral vector for efficient bicistronic expression of two genes in the central nervous system. *Axon Growth and Regeneration: Methods and Protocols*, 189-207.
- HWANG-VERSLUES, W. W. & SLADEK, F. M. 2010. HNF4 $\alpha$ —role in drug metabolism and potential drug target? *Current opinion in pharmacology*, 10, 698-705.
- ICHIDA, M., NEMOTO, S. & FINKEL, T. 2002. Identification of a specific molecular repressor of the peroxisome proliferator-activated receptor  $\gamma$  coactivator-1  $\alpha$  (PGC-1 $\alpha$ ). *Journal of Biological Chemistry*, 277, 50991-50995.
- ICHIKAWA, S., SAKIYAMA, H., SUZUKI, G., HIDARI, K. & HIRABAYASHI, Y. 1996. Expression cloning of a cDNA for human ceramide glucosyltransferase that catalyzes the first glycosylation step of glycosphingolipid synthesis. *Proceedings of the National Academy of Sciences*, 93, 4638-4643.
- ISMAIL, S., AZIAH HANAPI, N., AB HALIM, M. R., UCHAIPICHAT, V. & MACKENZIE, P. I. 2010. Effects of *Andrographis paniculata* and *Orthosiphon stamineus* Extracts on the Glucuronidation of 4-Methylumbelliferone in Human UGT Isoforms. *Molecules*, 15, 3578-3592.
- IWUCHUKWU, O. F., TALLARIDA, R. J. & NAGAR, S. 2011. Resveratrol in combination with other dietary polyphenols concomitantly enhances antiproliferation and UGT1A1 induction in Caco-2 cells. *Life sciences*, 88, 1047-1054.
- IZUKAWA, T., NAKAJIMA, M., FUJIWARA, R., YAMANAKA, H., FUKAMI, T., TAKAMIYA, M., AOKI, Y., IKUSHIRO, S.-I., SAKAKI, T. & YOKOI, T. 2009. Quantitative Analysis of UDP-

Glucuronosyltransferase (UGT) 1A and UGT2B Expression Levels in Human Livers. *Drug Metabolism and Disposition*, 37, 1759-1768.

- JAMES, R., ERLER, T. & KAZENWADEL, J. 1994. Structure of the murine homeobox gene *cdx-2*. Expression in embryonic and adult intestinal epithelium. *Journal of Biological Chemistry*, 269, 15229-15237.
- JEFCOATE, C. R., LIEHR, J. G., SANTEN, R. J., SUTTER, T. R., YAGER, J. D., YUE, W., SANTNER, S. J., TEKMAL, R., DEMERS, L. & PAULEY, R. 2000. Tissue-specific synthesis and oxidative metabolism of estrogens. *JNCI Monographs*, 2000, 95-112.
- JIANG, G., NEPOMUCENO, L., HOPKINS, K. & SLADEK, F. M. 1995. Exclusive homodimerization of the orphan receptor hepatocyte nuclear factor 4 defines a new subclass of nuclear receptors. *Molecular and Cellular Biology*, 15, 5131-5143.
- JIAO, L., KRAMER, J. R., RUGGE, M., PARENTE, P., VERSTOVSEK, G., ALSARRAJ, A. & EL-SERAG, H. B. 2013. Dietary intake of vegetables, folate, and antioxidants and the risk of Barrett's esophagus. *Cancer Causes & Control*, 24, 1005-1014.
- JOHE, K. K., HAZEL, T. G., MULLER, T., DUGICH-DJORDJEVIC, M. M. & MCKAY, R. 1996. Single factors direct the differentiation of stem cells from the fetal and adult central nervous system. *Genes & development*, 10, 3129-3140.
- JOHNSON, P. F. & MCKNIGHT, S. L. 1989. Eukaryotic transcriptional regulatory proteins. *Annual review of biochemistry*, 58, 799-839.
- JULLIEN, N. & HERMAN, J.-P. 2011. LUEGO: a cost and time saving gel shift procedure. *Biotechniques*, 51, 267-269.
- KALTHOFF, S., EHMER, U., FREIBERG, N., MANNS, M. P. & STRASSBURG, C. P. 2010a. Coffee induces expression of glucuronosyltransferases by the aryl hydrocarbon receptor and Nrf2 in liver and stomach. *Gastroenterology*, 139, 1699-1710. e2.
- KALTHOFF, S., EHMER, U., FREIBERG, N., MANNS, M. P. & STRASSBURG, C. P. 2010b. Interaction between oxidative stress sensor Nrf2 and xenobiotic-activated aryl hydrocarbon receptor in the regulation of the human phase II detoxifying UDP-glucuronosyltransferase 1A10. *Journal of Biological Chemistry*, 285, 5993-6002.
- KARDASSIS, D., PARDALI, K. & ZANNIS, V. I. 2000. SMAD proteins transactivate the human ApoCIII promoter by interacting physically and functionally with hepatocyte nuclear factor 4. *Journal of Biological Chemistry*, 275, 41405-41414.
- KELLY, C. J., ZHENG, L., CAMPBELL, E. L., SAEEDI, B., SCHOLZ, C. C., BAYLESS, A. J., WILSON, K. E., GLOVER, L. E., KOMINSKY, D. J. & MAGNUSON, A. 2015. Crosstalk between microbiota-derived short-chain fatty acids and intestinal epithelial HIF augments tissue barrier function. *Cell host & microbe*, 17, 662-671.
- KEMP, D. C., FAN, P. W. & STEVENS, J. C. 2002. Characterization of raloxifene glucuronidation in vitro: contribution of intestinal metabolism to presystemic clearance. *Drug Metabolism and Disposition*, 30, 694-700.
- KIANG, T. K., ENSOM, M. H. & CHANG, T. K. 2005. UDP-glucuronosyltransferases and clinical drug-drug interactions. *Pharmacology & therapeutics*, 106, 97-132.

- KIM, J. B., WRIGHT, H. M., WRIGHT, M. & SPIEGELMAN, B. M. 1998. ADD1/SREBP1 activates PPAR $\gamma$  through the production of endogenous ligand. *Proceedings of the national academy of sciences*, 95, 4333-4337.
- KIM, K.-A., KAKITANI, M., ZHAO, J., OSHIMA, T., TANG, T., BINNERTS, M., LIU, Y., BOYLE, B., PARK, E. & EMTAGE, P. 2005. Mitogenic influence of human R-spondin1 on the intestinal epithelium. *Science*, 309, 1256-1259.
- KIM, Y. S., TSAO, D., SIDDIQUI, B., WHITEHEAD, J. S., ARNSTEIN, P., BENNETT, J. & HICKS, J. 1980. Effects of sodium butyrate and dimethylsulfoxide on biochemical properties of human colon cancer cells. *Cancer*, 45, 1185-1192.
- KISELYUK, A., LEE, S.-H., FARBER-KATZ, S., ZHANG, M., ATHAVANKAR, S., COHEN, T., PINKERTON, A. B., YE, M., BUSHWAY, P. & RICHARDSON, A. D. 2012. HNF4 $\alpha$  antagonists discovered by a high-throughput screen for modulators of the human insulin promoter. *Chemistry & biology*, 19, 806-818.
- KITAMURA, S., MIYAZAKI, Y., SHINOMURA, Y., KONDO, S., KANAYAMA, S. & MATSUZAWA, Y. 1999. Peroxisome proliferator-activated receptor gamma induces growth arrest and differentiation markers of human colon cancer cells. *Jpn J Cancer Res*, 90, 75-80.
- KOBAYASHI, T., SLEEMAN, J. E., COUGHTRIE, M. W. & BURCHELL, B. 2006. Molecular and functional characterization of microsomal UDP-glucuronic acid uptake by members of the nucleotide sugar transporter (NST) family. *Biochemical Journal*, 400, 281-289.
- KOKAWA, Y., KISHI, N., JINNO, H., TANAKA-KAGAWA, T., NARIMATSU, S. & HANIOKA, N. 2013. Effect of UDP-glucuronosyltransferase 1A8 polymorphism on raloxifene glucuronidation. *European Journal of Pharmaceutical Sciences*, 49, 199-205.
- KUIPER, G. G., LEMMEN, J. G., CARLSSON, B., CORTON, J. C., SAFE, S. H., VAN DER SAAG, P. T., VAN DER BURG, B. & GUSTAFSSON, J.-A. K. 1998. Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor  $\beta$ . *Endocrinology*, 139, 4252-4263.
- KUO, C. J., CONLEY, P. B., HSIEH, C.-L., FRANCKE, U. & CRABTREE, G. R. 1990. Molecular cloning, functional expression, and chromosomal localization of mouse hepatocyte nuclear factor 1. *Proceedings of the National Academy of Sciences*, 87, 9838-9842.
- KUZUMAKI, T., KOBAYASHI, T. & ISHIKAWA, K. 1998. Genistein induces p21<sup>Cip1</sup>/WAF1 expression and blocks the G1 to S phase transition in mouse fibroblast and melanoma cells. *Biochemical and biophysical research communications*, 251, 291-295.
- L'AURELLE, A. J., OETTING, W. S., BASU, S., PRAUSA, S., MATAS, A. & JACOBSON, P. A. 2008. Pharmacogenetic effect of the UGT polymorphisms on mycophenolate is modified by calcineurin inhibitors. *European journal of clinical pharmacology*, 64, 1047.
- LAHAR, N., LEI, N. Y., WANG, J., JABAJI, Z., TUNG, S. C., JOSHI, V., LEWIS, M., STELZNER, M., MARTÍN, M. G. & DUNN, J. C. 2011. Intestinal subepithelial myofibroblasts support in vitro and in vivo growth of human small intestinal epithelium. *PLoS one*, 6, e26898.
- LANDRIER, J.-F., THOMAS, C., GROBER, J., DUEZ, H., PERCEVAULT, F., SOUIDI, M., LINARD, C., STAELS, B. & BESNARD, P. 2004. Statin induction of liver fatty acid-binding protein (L-FABP) gene expression is peroxisome proliferator-activated receptor- $\alpha$  dependent. *Journal of Biological Chemistry*, 279, 45512-45518.

- LEI, N. Y., JABAJI, Z., WANG, J., JOSHI, V. S., BRINKLEY, G. J., KHALIL, H., WANG, F., JAROSZEWICZ, A., PELLEGRINI, M. & LI, L. 2014. Intestinal subepithelial myofibroblasts support the growth of intestinal epithelial stem cells. *PLoS One*, 9, e84651.
- LÉPINE, J., BERNARD, O., PLANTE, M., TÊTU, B., PELLETIER, G., LABRIE, F., BÉLANGER, A. & GUILLEMETTE, C. 2004. Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. *The Journal of Clinical Endocrinology & Metabolism*, 89, 5222-5232.
- LEUSHACKE, M. & BARKER, N. 2014. Ex vivo culture of the intestinal epithelium: strategies and applications. *Gut*, gutjnl-2014-307204.
- LEVESQUE, E., DELAGE, R., BENOIT-BIANCAMANO, M. O., CARON, P., BERNARD, O., COUTURE, F. & GUILLEMETTE, C. 2007. The impact of UGT1A8, UGT1A9, and UGT2B7 genetic polymorphisms on the pharmacokinetic profile of mycophenolic acid after a single oral dose in healthy volunteers. *Clinical Pharmacology & Therapeutics*, 81, 392-400.
- LI, J., INOUE, J., CHOI, J.-M., NAKAMURA, S., YAN, Z., FUSHINOBU, S., KAMADA, H., KATO, H., HASHIDUME, T. & SHIMIZU, M. 2015. Identification of the flavonoid luteolin as a repressor of the transcription factor hepatocyte nuclear factor 4 $\alpha$ . *Journal of Biological Chemistry*, 290, 24021-24035.
- LIANG, S.-C., GE, G.-B., LIU, H.-X., ZHANG, Y.-Y., WANG, L.-M., ZHANG, J.-W., YIN, L., LI, W., FANG, Z.-Z. & WU, J.-J. 2010. Identification and characterization of human UDP-glucuronosyltransferases responsible for the in vitro glucuronidation of daphnetin. *Drug Metabolism and Disposition*, 38, 973-980.
- LIU, M., CHEN, S., YUEH, M.-F., WANG, G., HAO, H. & TUKEY, R. H. 2016. Reduction of p53 by knockdown of the UGT1 locus in colon epithelial cells causes an increase in tumorigenesis. *CMGH Cellular and Molecular Gastroenterology and Hepatology*, 2, 63-76. e5.
- LIU, X., TAM, V. H. & HU, M. 2007. Disposition of flavonoids via enteric recycling: determination of the UDP-glucuronosyltransferase (UGT) isoforms responsible for the metabolism of flavonoids in intact Caco-2 TC7 cells using siRNA. *Molecular pharmaceuticals*, 4, 873.
- LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>  $\Delta\Delta$ CT method. *methods*, 25, 402-408.
- LU, H. 2016. Crosstalk of HNF4 $\alpha$  with extracellular and intracellular signaling pathways in the regulation of hepatic metabolism of drugs and lipids. *Acta Pharmaceutica Sinica B*, 6, 393-408.
- LU, W., RETTENMEIER, E., PASZEK, M., YUEH, M.-F., TUKEY, R. H., TROTTIER, J., BARBIER, O. & CHEN, S. 2017. Crypt organoids culture as an in vitro model in drug metabolism and cytotoxicity studies. *Drug Metabolism and Disposition*, dmd. 117.075945.
- LUUKKANEN, L., TASKINEN, J., KURKELA, M., KOSTIAINEN, R., HIRVONEN, J. & FINEL, M. 2005. Kinetic characterization of the 1A subfamily of recombinant human UDP-glucuronosyltransferases. *Drug Metabolism and Disposition*, 33, 1017-1026.
- MACKENZIE, P. I. 2000. Identification of uridine diphosphate glucuronosyltransferases involved in the metabolism and clearance of mycophenolic acid. *Therapeutic drug monitoring*, 22, 10-13.

- MACKENZIE, P. I., BOCK, K. W., BURCHELL, B., GUILLEMETTE, C., IKUSHIRO, S.-I., IYANAGI, T., MINERS, J. O., OWENS, I. S. & NEBERT, D. W. 2005a. Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenetics and genomics*, 15, 677-685.
- MACKENZIE, P. I., GARDNER-STEPHEN, D. A. & MINERS, J. O. 2010a. 4.20 - UDP-Glucuronosyltransferases\* A2 - McQueen, Charlene A. *Comprehensive Toxicology (Second Edition)*. Oxford: Elsevier.
- MACKENZIE, P. I., GREGORY, P. A., LEWINSKY, R. H., YASMIN, S., HEIGHT, T., MCKINNON, R. & GARDNER-STEPHEN, D. A. 2005b. Polymorphic variations in the expression of the chemical detoxifying UDP glucuronosyltransferases. *Toxicology and applied pharmacology*, 207, 77-83.
- MACKENZIE, P. I., HJELMELAND, L. M. & OWENS, I. S. 1984. Purification and immunochemical characterization of a low-pI form of UDP glucuronosyltransferase from mouse liver. *Archives of biochemistry and biophysics*, 231, 487-497.
- MACKENZIE, P. I., HU, D. G. & GARDNER-STEPHEN, D. A. 2010b. The regulation of UDP-glucuronosyltransferase genes by tissue-specific and ligand-activated transcription factors. *Drug metabolism reviews*, 42, 99-109.
- MACKENZIE, P. I., OWENS, I. S., BURCHELL, B., BOCK, K. W., BAIROCH, A., BELANGER, A., GIGLEUX, S. F., GREEN, M., HUM, D. W. & IYANAGI, T. 1997. The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics and Genomics*, 7, 255-269.
- MACKENZIE, P. I., ROGERS, A., ELLIOT, D. J., CHAU, N., HULIN, J.-A., MINERS, J. O. & MEECH, R. 2011. The novel UDP glycosyltransferase 3A2: cloning, catalytic properties, and tissue distribution. *Molecular pharmacology*, 79, 472-478.
- MACKENZIE, P. I., ROGERS, A., TRELOAR, J., JORGENSEN, B. R., MINERS, J. O. & MEECH, R. 2008. Identification of UDP glycosyltransferase 3A1 as a UDP N-acetylglucosaminyltransferase. *Journal of Biological Chemistry*, 283, 36205-36210.
- MAJID, S., DAR, A. A., AHMAD, A. E., HIRATA, H., KAWAKAMI, K., SHAHRYARI, V., SAINI, S., TANAKA, Y., DAHIYA, A. V. & KHATRI, G. 2009. BTG3 tumor suppressor gene promoter demethylation, histone modification and cell cycle arrest by genistein in renal cancer. *Carcinogenesis*, 30, 662-670.
- MAJID, S., KIKUNO, N., NELLES, J., NOONAN, E., TANAKA, Y., KAWAMOTO, K., HIRATA, H., LI, L. C., ZHAO, H. & OKINO, S. T. 2008. Genistein induces the p21WAF1/CIP1 and p16INK4a tumor suppressor genes in prostate cancer cells by epigenetic mechanisms involving active chromatin modification. *Cancer research*, 68, 2736-2744.
- MANTOVANI, R., LI, X.-Y., PESSARA, U., VAN HUISJDIJNEN, R. H., BENOIST, C. & MATHIS, D. 1994. Dominant negative analogs of NF- $\kappa$ B. *Journal of Biological Chemistry*, 269, 20340-20346.
- MARI, L., MILANO, F., PARIKH, K., STRAUB, D., EVERTS, V., HOEBEN, K. K., FOCKENS, P., BUTTAR, N. S. & KRISHNADATH, K. K. 2014. A pSMAD/CDX2 complex is essential for the intestinalization of epithelial metaplasia. *Cell reports*, 7, 1197-1210.
- MARIADASON, J. M., ARANGO, D., CORNER, G. A., ARAÑES, M. J., HOTCHKISS, K. A., YANG, W. & AUGENLICHT, L. H. 2002. A gene expression profile that defines colon cell maturation in vitro. *Cancer Research*, 62, 4791-4804.

- MARIADASON, J. M., VELCICH, A., WILSON, A. J., AUGENLICHT, L. H. & GIBSON, P. R. 2001. Resistance to butyrate-induced cell differentiation and apoptosis during spontaneous Caco-2 cell differentiation. *Gastroenterology*, 120, 889-899.
- MARKEY, S. P. 2002. *Pathways of drug metabolism*, National Institutes of Health Clinical Center.
- MARSHMAN, E., BOOTH, C. & POTTEN, C. S. 2002. The intestinal epithelial stem cell. *Bioessays*, 24, 91-98.
- MARTINEZ-JIMENEZ, C. P., KYRMIZI, I., CARDOT, P., GONZALEZ, F. J. & TALIANIDIS, I. 2010. Hepatocyte nuclear factor 4 $\alpha$  coordinates a transcription factor network regulating hepatic fatty acid metabolism. *Molecular and cellular biology*, 30, 565-577.
- MATSUKAWA, Y., MARUI, N., SAKAI, T., SATOMI, Y., YOSHIDA, M., MATSUMOTO, K., NISHINO, H. & AOIKE, A. 1993. Genistein arrests cell cycle progression at G2-M. *Cancer Research*, 53, 1328-1331.
- MCGINNIS, W. & KRUMLAUF, R. 1992. Homeobox genes and axial patterning. *Cell*, 68, 283-302.
- MCINTYRE, A., GIBSON, P. & YOUNG, G. 1993. Butyrate production from dietary fibre and protection against large bowel cancer in a rat model. *Gut*, 34, 386-391.
- MEECH, R. & MACKENZIE, P. I. 1997. Structure and function of uridine diphosphate glucuronosyltransferases. *Clinical and experimental pharmacology and physiology*, 24, 907-915.
- MEECH, R. & MACKENZIE, P. I. 2010. UGT3A: novel UDP-glycosyltransferases of the UGT superfamily. *Drug metabolism reviews*, 42, 45-54.
- MEECH, R., MINERS, J. O., LEWIS, B. C. & MACKENZIE, P. I. 2012. The glycosidation of xenobiotics and endogenous compounds: versatility and redundancy in the UDP glycosyltransferase superfamily. *Pharmacology & therapeutics*, 134, 200-218.
- MEECH, R., MUBAROKAH, N., SHIVASAMI, A., ROGERS, A., NAIR, P. C., HU, D. G., MCKINNON, R. A. & MACKENZIE, P. I. 2015. A novel function for UDP glycosyltransferase 8: galactosidation of bile acids. *Molecular pharmacology*, 87, 442-450.
- MEERAN, S. M., AHMED, A. & TOLLEFSBOL, T. O. 2010. Epigenetic targets of bioactive dietary components for cancer prevention and therapy. *Clinical epigenetics*, 1, 101.
- MENESES, A., SCHNEEBERGER, K., KRUITWAGEN, H. S., PENNING, L. C., VAN STEENBEEK, F. G., BURGNER, I. A. & SPEE, B. 2016. Intestinal Organoids—Current and Future Applications. *Veterinary Sciences*, 3, 31.
- MI, Y., ZHAO, S., ZHANG, W., ZHANG, D., WENG, J., HUANG, K., SUN, H., TANG, H., ZHANG, X. & SUN, X. 2016. Down-regulation of Barx2 predicts poor survival in colorectal cancer. *Biochemical and biophysical research communications*, 478, 67-73.
- MIKSITS, M., MAIER-SALAMON, A., VO, N., PHUONG, T., SULYOK, M., SCHUHMACHER, R., SZEKERES, T. & JÄGER, W. 2010. Glucuronidation of piceatannol by human liver microsomes: Major role of UGT1A1, UGT1A8 and UGT1A10. *Journal of Pharmacy and Pharmacology*, 62, 47-54.
- MINERS, J. O., SMITH, P. A., SORICH, M. J., MCKINNON, R. A. & MACKENZIE, P. I. 2004. Predicting human drug glucuronidation parameters: application of in vitro and in silico modeling approaches. *Annu. Rev. Pharmacol. Toxicol.*, 44, 1-25.

- MITCHELMORE, C., TROELSEN, J. T., SPODSBERG, N., SJÖSTRÖM, H. & NOREN, O. 2000. Interaction between the homeodomain proteins Cdx2 and HNF1 $\alpha$  mediates expression of the lactase-phlorizin hydrolase gene. *Biochemical Journal*, 346, 529-535.
- MIYAKE, A., TAKEDA, T., ISOBE, A., WAKABAYASHI, A., NISHIMOTO, F., MORISHIGE, K.-I., SAKATA, M. & KIMURA, T. 2009. Repressive effect of the phytoestrogen genistein on estradiol-induced uterine leiomyoma cell proliferation. *Gynecological Endocrinology*, 25, 403-409.
- MIZUMA, T. 2009. Intestinal glucuronidation metabolism may have a greater impact on oral bioavailability than hepatic glucuronidation metabolism in humans: a study with raloxifene, substrate for UGT1A1, 1A8, 1A9, and 1A10. *International journal of pharmaceuticals*, 378, 140-141.
- MIZUTANI, A., KOINUMA, D., TSUTSUMI, S., KAMIMURA, N., MORIKAWA, M., SUZUKI, H. I., IMAMURA, T., MIYAZONO, K. & ABURATANI, H. 2011. Cell type-specific target selection by combinatorial binding of Smad2/3 proteins and hepatocyte nuclear factor 4 $\alpha$  in HepG2 cells. *Journal of Biological Chemistry*, 286, 29848-29860.
- MOJARRABI, B. & MACKENZIE, P. I. 1998. Characterization of two UDP glucuronosyltransferases that are predominantly expressed in human colon. *Biochemical and biophysical research communications*, 247, 704-709.
- MURAOKA, M., KAWAKITA, M. & ISHIDA, N. 2001. Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity. *FEBS letters*, 495, 87-93.
- MURPHY, P. A., SONG, T., BUSEMAN, G., BARUA, K., BEECHER, G. R., TRAINER, D. & HOLDEN, J. 1999. Isoflavones in retail and institutional soy foods. *Journal of agricultural and food chemistry*, 47, 2697-2704.
- NAKAMURA, A., NAKAJIMA, M., YAMANAKA, H., FUJIWARA, R. & YOKOI, T. 2008. Expression of UGT1A and UGT2B mRNA in Human Normal Tissues and Various Cell Lines. *Drug Metabolism and Disposition*, 36, 1461-1464.
- NIKOV, G. N., HOPKINS, N. E., BOUE, S. & ALWORTH, W. L. 2000. Interactions of dietary estrogens with human estrogen receptors and the effect on estrogen receptor-estrogen response element complex formation. *Environmental Health Perspectives*, 108, 867.
- NOAH, T. K., DONAHUE, B. & SHROYER, N. F. 2011. Intestinal development and differentiation. *Experimental cell research*, 317, 2702-2710.
- NOWELL, S. A., MASSENGILL, J. S., WILLIAMS, S., RADOMINSKA-PANDYA, A., TEPHLY, T. R., CHENG, Z., STRASSBURG, C. P., TUKEY, R. H., MACLEOD, S. L. & LANG, N. P. 1999. Glucuronidation of 2-hydroxyamino-1-methyl-6-phenylimidazo [4, 5-b] pyridine by human microsomal UDP-glucuronosyltransferases: identification of specific UGT1A family isoforms involved. *Carcinogenesis*, 20, 1107-1114.
- O'DWYER, P. J. & CATALANO, R. B. 2006. Uridine diphosphate glucuronosyltransferase (UGT) 1A1 and irinotecan: practical pharmacogenomics arrives in cancer therapy. *Journal of Clinical Oncology*, 24, 4534-4538.
- OGINO, S., SHIMA, K., BABA, Y., NOSHO, K., IRAHARA, N., KURE, S., CHEN, L., TOYODA, S., KIRKNER, G. J. & WANG, Y. L. 2009. Colorectal cancer expression of peroxisome proliferator-activated receptor  $\gamma$  (PPARG, PPAR $\gamma$ ) is associated with good prognosis. *Gastroenterology*, 136, 1242-1250.

- OHNO, S. & NAKAJIN, S. 2009. Determination of mRNA expression of human UDP-glucuronosyltransferases and application for localization in various human tissues by real-time reverse transcriptase-polymerase chain reaction. *Drug Metabolism and Disposition*, 37, 32-40.
- OKI, T., SOWA, Y., HIROSE, T., TAKAGAKI, N., HORINAKA, M., NAKANISHI, R., YASUDA, C., YOSHIDA, T., KANAZAWA, M. & SATOMI, Y. 2004. Genistein induces Gadd45 gene and G2/M cell cycle arrest in the DU145 human prostate cancer cell line. *FEBS letters*, 577, 55-59.
- OLIVEIRA, E. & WATSON, D. 2000. In vitro glucuronidation of kaempferol and quercetin by human UGT-1A9 microsomes. *FEBS letters*, 471, 1-6.
- OLSEN, A. K., BOYD, M., DANIELSEN, E. T. & TROELSEN, J. T. 2012. Current and emerging approaches to define intestinal epithelium-specific transcriptional networks. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 302, G277-G286.
- OOTANI, A., LI, X., SANGIORGI, E., HO, Q. T., UENO, H., TODA, S., SUGIHARA, H., FUJIMOTO, K., WEISSMAN, I. L. & CAPECCHI, M. R. 2009. Sustained in vitro intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nature medicine*, 15, 701-706.
- PALLAUF, K., DUCKSTEIN, N., HASLER, M., KLOTZ, L.-O. & RIMBACH, G. 2017. Flavonoids as Putative Inducers of the Transcription Factors Nrf2, FoxO, and PPAR $\gamma$ . *Oxidative Medicine and Cellular Longevity*, 2017.
- PANIGRAHY, D., KAIPAINEN, A., KIERAN, M. W. & HUANG, S. 2009. PPARs: a double-edged sword in cancer therapy? *PPAR research*, 2008.
- PARVIZ, F., MATULLO, C., GARRISON, W. D., SAVATSKI, L., ADAMSON, J. W., NING, G., KAESTNER, K. H., ROSSI, J. M., ZARET, K. S. & DUNCAN, S. A. 2003. Hepatocyte nuclear factor 4 $\alpha$  controls the development of a hepatic epithelium and liver morphogenesis. *Nature genetics*, 34, 292-296.
- PASTUŁA, A., MIDDELHOFF, M., BRANDTNER, A., TOBIASCH, M., HÖHL, B., NUBER, A. H., DEMIR, I. E., NEUPERT, S., KOLLMANN, P. & MAZZUOLI-WEBER, G. 2016. Three-dimensional gastrointestinal organoid culture in combination with nerves or fibroblasts: a method to characterize the gastrointestinal stem cell niche. *Stem cells international*, 2016.
- PATEL, H., HEWER, A., PHILLIPS, D., HAYES, J., WOLF, C. & CAMPBELL, F. 1997. Metabolic competence and susceptibility of intestinal epithelium to genotoxic injury during regeneration. *Carcinogenesis*, 18, 2171-2177.
- PATEL, R. P. & BARNES, S. 2010. Isoflavones and PPAR Signaling: A Critical Target in Cardiovascular, Metastatic, and Metabolic Disease. *PPAR Research*, 2010, 10.
- PAZIENZA, V., VINCIGUERRA, M. & MAZZOCCOLI, G. 2012. PPARs signaling and cancer in the gastrointestinal system. *PPAR research*, 2012.
- PELCLOVÁ, D., URBAN, P., PREISS, J., LUKAS, E., FENCLOVÁ, Z., NAVRÁTIL, T., DUBSKÁ, Z. & SENHOLDOVÁ, Z. 2006. Adverse health effects in humans exposed to 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin (TCDD). *Reviews on environmental health*, 21, 119.
- PENDER, S., TICKLE, S. P., DOCHERTY, A., HOWIE, D., WATHEN, N. C. & MACDONALD, T. T. 1997. A major role for matrix metalloproteinases in T cell injury in the gut. *The Journal of Immunology*, 158, 1582-1590.

- PERICLEOUS, M., MANDAIR, D. & CAPLIN, M. E. 2013. Diet and supplements and their impact on colorectal cancer. *Journal of gastrointestinal oncology*, 4, 409-423.
- PETRI, N., TANNERGREN, C., HOLST, B., MELLON, F. A., BAO, Y., PLUMB, G. W., BACON, J., O'LEARY, K. A., KROON, P. A. & KNUTSON, L. 2003. Absorption/metabolism of sulforaphane and quercetin, and regulation of phase II enzymes, in human jejunum in vivo. *Drug Metabolism and Disposition*, 31, 805-813.
- PINTO, D., GREGORIEFF, A., BEGTHEL, H. & CLEVERS, H. 2003. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & development*, 17, 1709-1713.
- PINTO, M., ROBINE-LEON, S., APPAY, M., KEDINGER, M., TRIADOU, N., DUSSAULX, E., LACROIX, B., SIMON-ASSMAN, P., HAFFEN, K., FOGH, J. & ZWEIBAUM, A. 1983. Enterocyte-like differentiation and polarization of the human colon carcinoma cell line Caco-2 in culture. *Biol Cell* 47, 323-330.
- POOL-ZOBEL, B. L., SELVARAJU, V., SAUER, J., KAUTENBURGER, T., KIEFER, J., RICHTER, K. K., SOOM, M. & WÖLFL, S. 2005. Butyrate may enhance toxicological defence in primary, adenoma and tumor human colon cells by favourably modulating expression of glutathione S-transferases genes, an approach in nutrigenomics. *Carcinogenesis*, 26, 1064-1076.
- PRITCHETT, L. E., ATHERTON, K. M., MUTCH, E. & FORD, D. 2008. Glucuronidation of the soyabean isoflavones genistein and daidzein by human liver is related to levels of UGT1A1 and UGT1A9 activity and alters isoflavone response in the MCF-7 human breast cancer cell line. *The Journal of nutritional biochemistry*, 19, 739-745.
- PUDDU, A., SANGUINETI, R., MONTECUCCO, F. & VIVIANI, G. L. 2014. Evidence for the gut microbiota short-chain fatty acids as key pathophysiological molecules improving diabetes. *Mediators of inflammation*, 2014.
- QIN, W., ZHU, W., SHI, H., HEWETT, J. E., RUHLEN, R. L., MACDONALD, R. S., ROTTINGHAUS, G. E., CHEN, Y.-C. & SAUTER, E. R. 2009. Soy isoflavones have an antiestrogenic effect and alter mammary promoter hypermethylation in healthy premenopausal women. *Nutrition and cancer*, 61, 238-244.
- RADA-IGLESIAS, A., ENROTH, S., AMEUR, A., KOCH, C. M., CLELLAND, G. K., RESPUELA-ALONSO, P., WILCOX, S., DOVEY, O. M., ELLIS, P. D. & LANGFORD, C. F. 2007. Butyrate mediates decrease of histone acetylation centered on transcription start sites and down-regulation of associated genes. *Genome research*, 17, 708-719.
- RADOMINSKA-PANDYA, A., CZERNIK, P. J., LITTLE, J. M., BATTAGLIA, E. & MACKENZIE, P. I. 1999. Structural and functional studies of UDP-glucuronosyltransferases. *Drug metabolism reviews*, 31, 817-899.
- RADOMINSKA-PANDYA, A., LITTLE, J. M., PANDYA, J. T., TEPHLY, T. R., KING, C. D., BARONE, G. W. & RAUFMAN, J.-P. 1998. UDP-glucuronosyltransferases in human intestinal mucosa. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism*, 1394, 199-208.
- RADOMINSKA-PANDYA, A., OUZZINE, M., FOURNEL-GIGLEUX, S. & MAGDALOU, J. 2005. Structure of UDP-glucuronosyltransferases in membranes. *Methods in enzymology*, 400, 116-147.

- RALSTON, A. & ROSSANT, J. 2008. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Developmental biology*, 313, 614-629.
- RAMÍREZ, J., MIRKOV, S., HOUSE, L. K. & RATAIN, M. J. 2015. Glucuronidation of OTS167 in humans is catalyzed by UDP-glucuronosyltransferases UGT1A1, UGT1A3, UGT1A8, and UGT1A10. *Drug Metabolism and Disposition*, 43, 928-935.
- RAMJI, D. P. & PELAGIA, F. 2002. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochemical Journal*, 365, 561-575.
- RANGA, A., GJOREVSKI, N. & LUTOLF, M. P. 2014. Drug discovery through stem cell-based organoid models. *Advanced drug delivery reviews*, 69, 19-28.
- REN, B., ROBERT, F., WYRICK, J. J., APARICIO, O., JENNINGS, E. G., SIMON, I., ZEITLINGER, J., SCHREIBER, J., HANNETT, N. & KANIN, E. 2000. Genome-wide location and function of DNA binding proteins. *Science*, 290, 2306-2309.
- RICKETTS, M.-L., MOORE, D. D., BANZ, W. J., MEZEI, O. & SHAY, N. F. 2005. Molecular mechanisms of action of the soy isoflavones includes activation of promiscuous nuclear receptors. A review. *The Journal of nutritional biochemistry*, 16, 321-330.
- RITTER, J. K. 2007. Intestinal UGTs as potential modifiers of pharmacokinetics and biological responses to drugs and xenobiotics. *Expert opinion on drug metabolism & toxicology*, 3, 93-107.
- ROUSSET, M. 1986. The human colon carcinoma cell lines HT-29 and Caco-2: two in vitro models for the study of intestinal differentiation. *Biochimie*, 68, 1035-1040.
- ROWLAND, A., MINERS, J. O. & MACKENZIE, P. I. 2013. The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *The international journal of biochemistry & cell biology*, 45, 1121-1132.
- RUAN, J.-Q., LI, S., LI, Y.-P., WU, W.-J., LEE, S. M.-Y. & YAN, R. 2015. The presystemic interplay between gut microbiota and orally administered calycosin-7-O- $\beta$ -D-glucoside. *Drug Metabolism and Disposition*, 43, 1601-1611.
- RUEFER, C. E., GERHÄUSER, C., FRANK, N., BECKER, H. & KULLING, S. E. 2005. In vitro phase II metabolism of xanthohumol by human UDP-glucuronosyltransferases and sulfotransferases. *Molecular nutrition & food research*, 49, 851-856.
- SAANDI, T., BARAILLE, F., DERBAL-WOLFROM, L., CATTIN, A. L., BENAHMED, F., MARTIN, E., CARDOT, P., DUCLOS, B., RIBEIRO, A., FREUND, J. N. & DULUC, I. 2013. Regulation of the tumor suppressor homeogene Cdx2 by HNF4[alpha] in intestinal cancer. *Oncogene*, 32, 3782-3788.
- SALAM, N. K., HUANG, T. H. W., KOTA, B. P., KIM, M. S., LI, Y. & HIBBS, D. E. 2008. Novel PPAR-gamma agonists identified from a natural product library: A virtual screening, induced-fit docking and biological assay study. *Chemical biology & drug design*, 71, 57-70.
- SALDANHA, S. N., KALA, R. & TOLLEFSBOL, T. O. 2014. Molecular mechanisms for inhibition of colon cancer cells by combined epigenetic-modulating epigallocatechin gallate and sodium butyrate. *Experimental cell research*, 324, 40-53.
- SAN ROMAN, A. K., ARONSON, B. E., KRASINSKI, S. D., SHIVDASANI, R. A. & VERZI, M. P. 2015. Transcription factors GATA4 and HNF4A control distinct aspects of intestinal

- homeostasis in conjunction with transcription factor CDX2. *Journal of Biological Chemistry*, 290, 1850-1860.
- SANCHEZ, R. & KAUFFMAN, F. 2010. Regulation of Xenobiotic Metabolism in the Liver. *Comprehensive Toxicology*, 9, 109-128.
- SANDER, G. R. & POWELL, B. C. 2004. Expression of the homeobox gene *barx2* in the gut. *Journal of Histochemistry & Cytochemistry*, 52, 541-544.
- SASAKI, T., GILTAY, R., TALTS, U., TIMPL, R. & TALTS, J. F. 2002. Expression and distribution of laminin  $\alpha 1$  and  $\alpha 2$  chains in embryonic and adult mouse tissues: an immunochemical approach. *Experimental cell research*, 275, 185-199.
- SATO, T. & CLEVERS, H. 2013. Growing self-organizing mini-guts from a single intestinal stem cell: mechanism and applications. *Science*, 340, 1190-1194.
- SATO, T., STANGE, D. E., FERRANTE, M., VRIES, R. G., VAN ES, J. H., VAN DEN BRINK, S., VAN HOUT, W. J., PRONK, A., VAN GORP, J. & SIERSEMA, P. D. 2011. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology*, 141, 1762-1772.
- SATO, T., VRIES, R. G., SNIPPERT, H. J., VAN DE WETERING, M., BARKER, N., STANGE, D. E., VAN ES, J. H., ABO, A., KUJALA, P. & PETERS, P. J. 2009. Single Lgr5 stem cells build crypt villus structures in vitro without a mesenchymal niche. *Nature*, 459, 262-265.
- SCHWANK, G., ANDERSSON-ROLF, A., KOO, B.-K., SASAKI, N. & CLEVERS, H. 2013. Generation of BAC transgenic epithelial organoids. *PloS one*, 8, e76871.
- SHAW, L. M. & NOWAK, I. 1995. Mycophenolic Acid: Measurement and Relationship to Pharmacologic Effects. *Therapeutic drug monitoring*, 17, 685-689.
- SHIMIZU, E. & OGATA, Y. 2002. Activation of bone sialoprotein gene transcription by flavonoids is mediated through an inverted CCAAT box in ROS 17/2.8 cells. *Journal of cellular biochemistry*, 86, 35-44.
- SHIMOKAWA, M., OHTA, Y., NISHIKORI, S., MATANO, M., TAKANO, A., FUJII, M., DATE, S., SUGIMOTO, S., KANAI, T. & SATO, T. 2017. Visualization and targeting of LGR5+ human colon cancer stem cells. *Nature*, 545, 187-192.
- SISS, M.-H., LE BON, A.-M., CANIVENC-LAVIER, M.-C., AMIOT, M.-J., SABATIER, S., AUBERT, S. Y. & SUSCHETET, M. 1996. Flavonoids of Honey and Propolis: Characterization and Effects on Hepatic Drug-Metabolizing Enzymes and Benzo [a] pyrene- DNA Binding in Rats. *Journal of Agricultural and Food Chemistry*, 44, 2297-2301.
- SILBERG, D. G., SWAIN, G. P., SUH, E. R. & TRABER, P. G. 2000. Cdx1 and cdx2 expression during intestinal development. *Gastroenterology*, 119, 961-971.
- SMITH, R. E. 2013. *Medicinal Chemistry-Fusion of Traditional and Western Medicine*, Bentham Science Publishers.
- SOUTOGLU, E. & TALIANIDIS, I. 2002. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science*, 295, 1901-1904.
- STAHL, N., JUREVICS, H., MORELL, P., SUZUKI, K. & POPKO, B. 1994. Isolation, characterization, and expression of cDNA clones that encode rat UDP-galactose: Ceramide galactosyltransferase. *Journal of neuroscience research*, 38, 234-242.

- STEGMANN, A., HANSEN, M., WANG, Y., LARSEN, J. B., LUND, L. R., RITIÉ, L., NICHOLSON, J. K., QUISTORFF, B., SIMON-ASSMANN, P. & TROELSEN, J. T. 2006. Metabolome, transcriptome, and bioinformatic cis-element analyses point to HNF-4 as a central regulator of gene expression during enterocyte differentiation. *Physiological genomics*, 27, 141-155.
- STELZNER, M., HELMRATH, M., DUNN, J. C., HENNING, S. J., HOUCHEM, C. W., KUO, C., LYNCH, J., LI, L., MAGNESS, S. T. & MARTIN, M. G. 2012. A nomenclature for intestinal in vitro cultures. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 302, G1359-G1363.
- STRASSBURG, C., STRASSBURG, A., KNEIP, S., BARUT, A., TUKEY, R., RODECK, B. & MANNS, M. 2002. Developmental aspects of human hepatic drug glucuronidation in young children and adults. *Gut*, 50, 259-265.
- STRASSBURG, C. P., KNEIP, S., TOPP, J., OBERMAYER-STRAUB, P., BARUT, A., TUKEY, R. H. & MANNS, M. P. 2000. Polymorphic gene regulation and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine. *Journal of Biological Chemistry*, 275, 36164-36171.
- STRASSBURG, C. P., MANNS, M. P. & TUKEY, R. H. 1998. Expression of the UDP-glucuronosyltransferase 1A locus in human colon identification and characterization of the novel extrahepatic UGT1A8. *Journal of Biological Chemistry*, 273, 8719-8726.
- STRASSBURG, C. P., OLDHAFFER, K., MANNS, M. P. & TUKEY, R. H. 1997. Differential Expression of the UGT1A Locus in Human Liver, Biliary, and Gastric Tissue: Identification of UGT1A7 and UGT1A10 Transcripts in Extrahepatic Tissue. *Molecular Pharmacology*, 52, 212-220.
- SU, W., BUSH, C. R., NECELA, B. M., CALCAGNO, S. R., MURRAY, N. R., FIELDS, A. P. & THOMPSON, E. A. 2007. Differential expression, distribution, and function of PPAR- $\gamma$  in the proximal and distal colon. *Physiological Genomics*, 30, 342-353.
- SUBTIL, C., GUÉRIN, E., SCHNEIDER, A., CHENARD, M.-P., MARTIN, E., DOMON-DELL, C., DULUC, I., BRABLETZ, T., KEDINGER, M. & DUCLOS, B. 2007. Frequent rearrangements and amplification of the CDX2 homeobox gene in human sporadic colorectal cancers with chromosomal instability. *Cancer letters*, 247, 197-203.
- SUH, E., CHEN, L., TAYLOR, J. & TRABER, P. G. 1994. A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Molecular and Cellular Biology*, 14, 7340-7351.
- SUH, E. & TRABER, P. G. 1996. An intestine-specific homeobox gene regulates proliferation and differentiation. *Mol Cell Biol*, 16, 619-25.
- SUN, D., JONES, N. R., MANNI, A. & LAZARUS, P. 2013. Characterization of raloxifene glucuronidation. Potential role of UGT1A8 genotype on raloxifene metabolism in vivo. *Cancer Prevention Research*, canprevres. 0448.2012.
- SUN, D., SHARMA, A. K., DELLINGER, R. W., BLEVINS-PRIMEAU, A. S., BALLIET, R. M., CHEN, G., BOYIRI, T., AMIN, S. & LAZARUS, P. 2007. Glucuronidation of active tamoxifen metabolites by the human UDP glucuronosyltransferases. *Drug Metabolism and Disposition*, 35, 2006-2014.
- SWANSON, H. I. 2015. Drug metabolism by the host and gut microbiota: a partnership or rivalry? *Drug Metabolism and Disposition*, 43, 1499-1504.

- TAKAHIRO, T., KANNA, O., YOSHIDA, T., IKARASHI, N., KIYOMI, I. & SUGIYAMA, K. 2009. Ciprofloxacin suppresses Cyp3a in mouse liver by reducing lithocholic acid-producing intestinal flora. *Drug metabolism and pharmacokinetics*, 24, 201-208.
- TANG, L., SINGH, R., LIU, Z. & HU, M. 2009. Structure and concentration changes affect characterization of UGT isoform-specific metabolism of isoflavones. *Molecular pharmaceutics*, 6, 1466-1482.
- TANG, L., YE, L., SINGH, R., WU, B., LV, C., ZHAO, J., LIU, Z. & HU, M. 2010. Use of glucuronidation fingerprinting to describe and predict mono- and dihydroxyflavone metabolism by recombinant UGT isoforms and human intestinal and liver microsomes. *Molecular pharmaceutics*, 7, 664-679.
- TARAVIRAS, S., MONAGHAN, A. P., SCHÜTZ, G. & KELSEY, G. 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mechanisms of development*, 48, 67-79.
- TEPHLY, T. R. & BURCHELL, B. 1990. UDP-glucuronosyltransferases: a family of detoxifying enzymes. *Trends in Pharmacological Sciences*, 11, 276-279.
- TERRY, P., GIOVANNUCCI, E., MICHELS, K. B., BERGKVIST, L., HANSEN, H., HOLMBERG, L. & WOLK, A. 2001. Fruit, vegetables, dietary fiber, and risk of colorectal cancer. *Journal of the National Cancer Institute*, 93, 525-533.
- THIBAUDEAU, J., LÉPINE, J., TOJCIC, J., DUGUAY, Y., PELLETIER, G., PLANTE, M., BRISSON, J., TÊTU, B., JACOB, S. & PERUSSE, L. 2006. Characterization of common UGT1A8, UGT1A9, and UGT2B7 variants with different capacities to inactivate mutagenic 4-hydroxylated metabolites of estradiol and estrone. *Cancer research*, 66, 125-133.
- THILAKARATHNA, S. H. & RUPASINGHE, H. 2013. Flavonoid bioavailability and attempts for bioavailability enhancement. *Nutrients*, 5, 3367-3387.
- TREMBLAY, E., AUCLAIR, J., DELVIN, E., LEVY, E., MÉNARD, D., PSHEZHETSKY, A. V., RIVARD, N., SEIDMAN, E. G., SINNETT, D. & VACHON, P. H. 2006. Gene expression profiles of normal proliferating and differentiating human intestinal epithelial cells: A comparison with the Caco-2 cell model. *Journal of cellular biochemistry*, 99, 1175-1186.
- TRIFF, K., KONGANTI, K., GADDIS, S., ZHOU, B., IVANOV, I. & CHAPKIN, R. S. 2013. Genome-wide analysis of the rat colon reveals proximal-distal differences in histone modifications and proto-oncogene expression. *Physiological genomics*, 45, 1229-1243.
- TUKEY, R. H. & STRASSBURG, C. P. 2000. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annual review of pharmacology and toxicology*, 40, 581-616.
- TUKEY, R. H. & STRASSBURG, C. P. 2001. Genetic multiplicity of the human UDP-glucuronosyltransferases and regulation in the gastrointestinal tract. *Molecular Pharmacology*, 59, 405-414.
- UCHAIPICHAT, V., MACKENZIE, P. I., GUO, X.-H., GARDNER-STEPHEN, D., GALETIN, A., HOUSTON, J. B. & MINERS, J. O. 2004. Human UDP-glucuronosyltransferases: isoform selectivity and kinetics of 4-methylumbelliferone and 1-naphthol glucuronidation, effects of organic solvents, and inhibition by diclofenac and probenecid. *Drug Metabolism and Disposition*, 32, 413-423.

- UHLÉN, M., FAGERBERG, L., HALLSTRÖM, B. M., LINDSKOG, C., OKSVOLD, P., MARDINOGLU, A., SIVERTSSON, Å., KAMPF, C., SJÖSTEDT, E. & ASPLUND, A. 2015. Tissue-based map of the human proteome. *Science*, 347, 1260419.
- VAN LIDTH DE JEUDE, J., VERMEULEN, J., MONTENEGRO-MIRANDA, P. S., VAN DEN BRINK, G. R. & HEIJMANS, J. 2015. A protocol for lentiviral transduction and downstream analysis of intestinal organoids. *Journal of visualized experiments: JoVE*.
- VAN RYMENANT, E., ABRANKÓ, L., TUMOVA, S., GROOTAERT, C., VAN CAMP, J., WILLIAMSON, G. & KERIMI, A. 2017. Chronic exposure to short-chain fatty acids modulates transport and metabolism of microbiome-derived phenolics in human intestinal cells. *The Journal of Nutritional Biochemistry*, 39, 156-168.
- VANHOOK, A. M. 2015. Butyrate benefits the intestinal barrier. *Sci. Signal.*, 8, ec135-ec135.
- VERZI, M. P., SHIN, H., HE, H. H., SULAHIAN, R., MEYER, C. A., MONTGOMERY, R. K., FLEET, J. C., BROWN, M., LIU, X. S. & SHIVDASANI, R. A. 2010. Differentiation-specific histone modifications reveal dynamic chromatin interactions and partners for the intestinal transcription factor CDX2. *Developmental cell*, 19, 713-726.
- VERZI, M. P., SHIN, H., SAN ROMAN, A. K., LIU, X. S. & SHIVDASANI, R. A. 2013. Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. *Molecular and cellular biology*, 33, 281-292.
- VETŐ, B., BOJCSUK, D., BACQUET, C., KISS, J., SIPEKI, S., MARTIN, L., BUDAY, L., BÁLINT, B. L. & ARÁNYI, T. 2017. The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2. *PloS one*, 12, e0172020.
- VOGIATZOGLU, A., MULLIGAN, A. A., LENTJES, M. A., LUBEN, R. N., SPENCER, J. P., SCHROETER, H., KHAW, K.-T. & KUHNLE, G. G. 2015. Flavonoid intake in European adults (18 to 64 years). *PloS one*, 10, e0128132.
- WÄCHTERSCHÄUSER, A., LOITSCH, S. M. & STEIN, J. 2000. PPAR- $\gamma$  is selectively upregulated in Caco-2 cells by butyrate. *Biochemical and biophysical research communications*, 272, 380-385.
- WALLIG, M. A. 2004. Glucuronidation and susceptibility to chemical carcinogenesis. *Toxicological sciences*, 78, 1-2.
- WALTERS, J., HOWARD, A., RUMBLE, H., PRATHALINGAM, S. R., SHAW-SMITH, C. J. & LEGON, S. 1997. Differences in expression of homeobox transcription factors in proximal and distal human small intestine. *Gastroenterology*, 113, 472-477.
- WANG, M., QI, Y.-Y., CHEN, S., SUN, D.-F., WANG, S., CHEN, J., LI, Y.-Q., HAN, W. & YANG, X.-Y. 2012. Expression of UDP-glucuronosyltransferase 1A, nuclear factor erythroid-E2-related factor 2 and Kelch-like ECH-associated protein 1 in colonic mucosa, adenoma and adenocarcinoma tissue. *Oncology letters*, 4, 925-930.
- WANG, M., SUN, D.-F., WANG, S., QING, Y., CHEN, S., WU, D., LIN, Y.-M., LUO, J.-Z. & LI, Y.-Q. 2013. Polymorphic expression of UDP-glucuronosyltransferase UGT1A gene in human colorectal cancer. *PLoS One*, 8, e57045.
- WANG, S. W., KULKARNI, K. H., TANG, L., WANG, J. R., YIN, T., DAIDOJI, T., YOKOTA, H. & HU, M. 2009. Disposition of flavonoids via enteric recycling: UDP-glucuronosyltransferase (UGT) 1As deficiency in Gunn rats is compensated by increases in UGT2Bs activities. *Journal of Pharmacology and Experimental Therapeutics*, 329, 1023-1031.

- WATANABE, K., UENO, M., KAMIYA, D., NISHIYAMA, A., MATSUMURA, M., WATAYA, T., TAKAHASHI, J. B., NISHIKAWA, S., NISHIKAWA, S.-I. & MUGURUMA, K. 2007. A ROCK inhibitor permits survival of dissociated human embryonic stem cells. *Nature biotechnology*, 25, 681-686.
- WATANABE, Y., NAKAJIMA, M. & YOKOI, T. 2002. Troglitazone glucuronidation in human liver and intestine microsomes: high catalytic activity of UGT1A8 and UGT1A10. *Drug Metabolism and Disposition*, 30, 1462-1469.
- WHITEHEAD, R., YOUNG, G. & BHATHAL, P. 1986. Effects of short chain fatty acids on a new human colon carcinoma cell line (LIM1215). *Gut*, 27, 1457-1463.
- WICKING, C., SIMMS, L. A., EVANS, T., WALSH, M., CHAWENGSAKSOPHAK, K., BECK, F., CHENEVIX-TRENCH, G., YOUNG, J., JASS, J. & LEGGETT, B. 1998. CDX2, a human homologue of *Drosophila* caudal, is mutated in both alleles in a replication error positive colorectal cancer. *Oncogene*, 17, 657-659.
- WILLIAMS, J. A., HYLAND, R., JONES, B. C., SMITH, D. A., HURST, S., GOOSEN, T. C., PETERKIN, V., KOUP, J. R. & BALL, S. E. 2004. Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUC<sub>i</sub>/AUC) ratios. *Drug Metabolism and Disposition*, 32, 1201-1208.
- WILLIAMS, R. T. 1959. *Detoxification mechanisms : the metabolism and detoxification of drugs, toxic substances and other organic compounds*, London, Chapman and Hall.
- WILLIAMSON, I. A., ARNOLD, J. W., SAMSA, L. A., GAYNOR, L., DISALVO, M., COCCHIARO, J. L., CARROLL, I., AZCARATE-PERIL, M. A., RAWLS, J. F., ALLBRITTON, N. L. & MAGNESS, S. T. 2018. A High-Throughput Organoid Microinjection Platform to Study Gastrointestinal Microbiota and Luminal Physiology. *Cell Mol Gastroenterol Hepatol*, 6, 301-319.
- WU, B., KULKARNI, K., BASU, S., ZHANG, S. & HU, M. 2011. First-pass metabolism via UDP-glucuronosyltransferase: a barrier to oral bioavailability of phenolics. *Journal of pharmaceutical sciences*, 100, 3655-3681.
- XIE, X., LI, J., WANG, H., LI, H., LIU, J., FU, Q., HUANG, J., ZHU, C., ZHONG, G. & WANG, X. 2015. Associations of UDP-glucuronosyltransferases polymorphisms with mycophenolate mofetil pharmacokinetics in Chinese renal transplant patients. *Acta Pharmacologica Sinica*, 36, 644-650.
- YAMASHITA, S., KONISHI, K., YAMAZAKI, Y., TAKI, Y., SAKANE, T., SEZAKI, H. & FURUYAMA, Y. 2002. New and better protocols for a short-term Caco-2 cell culture system. *Journal of pharmaceutical sciences*, 91, 669-679.
- YATSUNENKO, T., REY, F. E., MANARY, M. J., TREHAN, I., DOMINGUEZ-BELLO, M. G., CONTRERAS, M., MAGRIS, M., HIDALGO, G., BALDASSANO, R. N. & ANOKHIN, A. P. 2012. Human gut microbiome viewed across age and geography. *nature*, 486, 222.
- YEUNG, T. M., CHIA, L. A., KOSINSKI, C. M. & KUO, C. J. 2011. Regulation of self-renewal and differentiation by the intestinal stem cell niche. *Cellular and Molecular Life Sciences*, 68, 2513-2523.
- YEUNG, T. M., GANDHI, S. C., WILDING, J. L., MUSCHEL, R. & BODMER, W. F. 2010. Cancer stem cells from colorectal cancer-derived cell lines. *Proceedings of the National Academy of Sciences*, 107, 3722-3727.

- YU, Z., TANG, Y., HU, D. & LI, J. 2005. Inhibitory effect of genistein on mouse colon cancer MC-26 cells involved TGF- $\beta$ 1/Smad pathway. *Biochemical and biophysical research communications*, 333, 827-832.
- ZACHOS, N. C., KOVBASNJUK, O., FOULKE-ABEL, J., IN, J., BLUTT, S. E., DE JONGE, H. R., ESTES, M. K. & DONOWITZ, M. 2016. Human enteroids/colonoids and intestinal organoids functionally recapitulate normal intestinal physiology and pathophysiology. *Journal of Biological Chemistry*, 291, 3759-3766.
- ZHANG, J., YI, M., ZHA, L., CHEN, S., LI, Z., LI, C., GONG, M., DENG, H., CHU, X., CHEN, J., ZHANG, Z., MAO, L. & SUN, S. 2016. Sodium Butyrate Induces Endoplasmic Reticulum Stress and Autophagy in Colorectal Cells: Implications for Apoptosis. *PLoS One*, 11, e0147218.
- ZHANG, L., LIN, G. & ZUO, Z. 2007. Involvement of UDP-glucuronosyltransferases in the extensive liver and intestinal first-pass metabolism of flavonoid baicalein. *Pharmaceutical research*, 24, 81.
- ZHOU, Q., ZHENG, Z., XIA, B., TANG, L., LV, C., LIU, W., LIU, Z. & HU, M. 2010. Use of isoform-specific UGT metabolism to determine and describe rates and profiles of glucuronidation of wogonin and oroxylin A by human liver and intestinal microsomes. *Pharmaceutical research*, 27, 1568-1583.
- ZHU, B. T. & CONNEY, A. H. 1998. Functional role of estrogen metabolism in target cells: review and perspectives. *Carcinogenesis*, 19, 1-27.