The Pennsylvania State University

The Graduate School

College of Medicine

EXPRESSION OF URIDINEDIPHOSPHATE GLUCURONOSYLTRANSFERASE GENES: FOCUS ON ALTERNATIVE SPLICING AND TRANSCRIPTIONAL REGULATORY MECHANISMS THAT CONTRIBUTE TO INTERINDIVIDUAL DIFFERENCES IN DRUG AND CARCINOGEN METABOLISM

A Dissertation in

Pharmacology

by

Nathan R. Jones

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

May 2012

The dissertation of Nathan R. Jones was reviewed and approved* by the following:

Philip Lazarus
Professor of Pharmacology and Public Health Sciences
Dissertation Advisor
Chair of Committee

John P. Richie
Professor of Public Health Sciences and Pharmacology

Thomas E. Spratt
Associate Professor of Biochemistry and Molecular Biology

Richard R. Young
Professor of Supply Chain Management

Jong K. Yun
Associate Professor of Pharmacology

Kent E. Vrana
Elliot S. Vessel Professor and Chair of Pharmacology
Head of the Department of Pharmacology

*Signatures are on file in the Graduate School

ABSTRACT

There is a complex interplay between genetic and environmental factors that determine inter-individual differences in disease disposition and therapeutic response. Drug metabolism pathways have been a major focus of pharmacogenetic studies because inter-individual differences in the expression and activity of these enzymes may cause clinically significant effects on the kinetic properties of various drugs. Because of their role in the metabolism of chemical toxins and carcinogens, genetic differences in drug metabolizing enzymes are also associated with risk of diseases such as cancer. There are many factors governing inter-individual variation in drug-metabolizing enzymes including SNPs, epigenetics, alternative splicing events, transcriptional regulation, and post-translational modifications.

UDP-glucuronosyltransferases (UGTs) play an important role in the metabolism and excretion of endogenous and xenobiotic compounds including drugs and carcinogens. UGT enzymes mediate the phase II conjugation of glucuronic acid to their substrates, thereby increasing substrate polarity and facilitating their excretion. Variations in UGT genes are associated with altered drug metabolism and cancer risk. Some of the genetic factors underlying these associations have been discovered, but often there is wide variability in phenotype within a given genotype. The liver is the organ most commonly associated with metabolism, and most UGTs are expressed in the human liver. Expression in extrahepatic tissues has been less well characterized, even though tissues that form a barrier with the environment, such as aerodigestive and gastrointestinal tract tissues represent an important first line of defense against exposures to xenobiotic compounds.

A better understanding of the inter-individual variability and relative abundance of UGT gene expression in different tissues is important as this helps determine the physiological relevance of each UGT enzyme. While many previous studies have used qualitative reverse

transcription polymerase chain reaction (RT-PCR) for determining which UGT genes are expressed in different tissues, some quantitative analysis of UGT expression has been performed In studies described in this thesis dissertation, real-time PCR was used to quantify the expression of 16 UGT enzymes in multiple specimens of various normal human tissues including lung, liver, larynx, brain, tongue, floor of mouth, tonsil, esophagus, endometrium, and pancreas. Substantial inter-individual variability in expression was observed in both hepatic and extrahepatic tissues. In extrahepatic tissues, unpredictable expression patterns were frequently observed, in which UGT enzymes expressed in some individuals were not expressed in others.

In the liver, there was a high degree of correlation between the expression levels of many UGT enzymes within the same individual, suggesting a common mechanism of transcriptional regulation. The hepatic expression of UGTs is known to be transcriptionally regulated by ligand-activated and liver-enriched transcription factors (LETFs). The hepatic transcriptional regulation of several UGTs has been partially described, with UGT2B10 a notable exception. UGT2B10 exhibits glucuronidation activity against pharmacological substrates (olanzapine), toxins (nicotine), and carcinogens (NNAL), suggesting that inter-individual variability in the expression of this gene may affect both drug metabolism and cancer risk. We hypothesized that transcriptional regulatory mechanisms contribute to the observed interindividual variability in UGT2B10 mRNA expression levels. Through luciferase assays and DNaseI footprint analysis, a region between -180 bp and -250 bp upstream from the transcription start site was identified as the UGT2B10 core promoter element. Subsequent gel shift and supershift analyses determined that the Oct-1 and HNF3α proteins were capable of binding to these elements. Mutation of the HNF3α site in the +27bp/-1948bp UGT2B10 promoter construct led to a 49% decrease in promoter activity, and mutation of the Oct-1 site led to a 61% decrease (p=0.0095 and p=0.0004, respectively). The Oct-1/HNF3α double mutant further reduced UGT2B10 promoter activity to 18% of the WT construct (p=0.0002). The HNF3α protein is an LETF, and future studies will be

conducted to determine whether variants of HNF3α are associated with UGT2B10 expression levels in human liver, as this may be an important contributor to the observed inter-individual differences in expression.

In addition to LETFs, UGTs are also regulated by ligand-activated transcription factors. In particular, UGT genes are known to be regulated by antioxidant response elements (AREs). Phase II enzyme inducers such as L-sulforaphane (SFN) have been shown to induce transcription of UGTs through the nuclear factor-erythroid 2-related factor 2 (nrf2)/ARE pathway. The UGT2B10 enzyme detoxifies carcinogens, making it important to understand how it is being regulated in response to phase II enzyme inducers s as this could influence the carcinogenicity of these compounds. In the current studies, several putative ARE sites were identified in the UGT2B10 promoter and were hypothesized to be regulated by the nrf2/ARE pathway. UGT2B10 expression was found to be significantly decreased in response to SFN treatment (52%, p=0.008). The repression effect was present even after siRNA knockdown of nrf2, indicating that this effect was occurring via an nrf2-independent mechanism. Using luciferase assays, the promoter element responsible for the repression was localized to a 976 bp region between -973 bp and -1948 bp upstream of the transcription start site. Future studies will be needed to elucidate the regulatory factors that are binding and initiating this repression. The repression of UGT2B10 expression via SFN may represent an additional mechanism that is contributing to inter-individual differences in expression.

Another potential regulator of inter-individual differences in glucuronidation is alternative splicing. An alternative exon 5 in the common region of the UGT1A gene cluster leads to the expression of 18 additional mRNA species from this locus. The alternative splice isoforms have a dominant-negative effect on the wild type (WT) isoforms *in vitro.* We hypothesized that inter-individual variation in the relative abundance of WT and splice variant expression affects glucuronidation capacity in human liver, which could carry important

pharmacogenetic implications. In these studies, it was determined that UGT1A splice variants, on average, represent less than 7% of the total UGT1A transcript profile in human liver for all hepatic UGT1A species, with relatively low inter-individual variability in expression between different individuals. A consistent pattern was observed in several extrahepatic tissues as well. UGT1A WT and splice variant expression were both correlated with glucuronidation activity. Sequence alignment of the UGT1A alternative exon 5 with the primate-specific alu transposable element revealed that it is a recent evolutionary event, and it displays a low inclusion rate characteristic of similar exons. Alu-derived exons are usually neutral or only slightly deleterious because the novel, alternatively spliced product represents only a small percentage of the total mRNA species. For this reason, the proteins they encode have often been characterized as nonfunctional, evolutionary intermediates.

This dissertation research has contributed to the understanding of mechanisms underlying UGT gene expression. Studies examining the inter-individual expression patterns of UGT enzymes in human tissues, an analysis of the importance of UGT1A alternative splice variants and how they affect glucuronidation capacity in human liver, and an assessment of the transcriptional regulation of the UGT2B10 promoter by LETFs and SFN were performed in this work. Together these data have improved our knowledge of how inter-individual differences in the expression of metabolizing enzymes are manifested and how these differences may potentially play a role in drug and carcinogen metabolism, personalized medicine, and cancer risk assessment.

Table of Contents

LIST OF FIGURES ix
LIST OF TABLES xi
LIST OF ABBREVIATIONS xii
ACKNOWLEDGEMENTS xv

Chapter 1 REVIEW OF THE LITERATURE 1

1.1 General Background and Significance 2
1.1.1. Drug Metabolism and Genetic Variation 2
1.1.2. Cancer Risk and Genetic Variation 6
1.2. UDP-Glucuronosyltransferases 8
1.2.1. UGT families 9
1.2.2. UGT Function and Localization 12
1.2.3. Importance of UGTs for Metabolizing Endogenous Compounds 14
1.2.4. UGT Pharmacogenetics 15
1.2.5. UGT Variants and Cancer Risk 17
1.2.6. The UGT2B10 D67Y Allele 18
1.3. Expression and transcriptional regulation of UGTs 20
1.3.1 Tissular Distribution and Interindividual Differences in UGT Gene Expression 21
1.3.2. Hepatic Transcriptional Regulation of UGT Gene Expression 23
1.3.3. Extrahepatic Transcriptional Regulation of UGT Genes 27
1.3.4. Xenobiotic Regulation of UGT Gene Expression 28
1.4 Alternative Splicing of UGTs 31
1.4.1. Alternative Splicing and the UGT1A Locus 34
1.4.2 Alternative Splicing and UGT2B Genes 36
1.5. Aims and Hypotheses 37

Chapter 2 ENDOGENOUS CONTROL GENE SELECTION FOR QUANTITATIVE STUDIES OF UGT GENE EXPRESSION IN HUMAN TISSUES 39

2.1 Abstract 40
2.2 Introduction 41
2.3 Methods 43
2.4 Results 46
2.5 Discussion 52

Chapter 3 QUANTIFICATION OF HEPATIC UGT1A SPLICE VARIANT EXPRESSION AND CORRELATION OF UGT1A1 VARIANT EXPRESSION WITH GLUCURONIDATION ACTIVITY 55

3.1 Abstract 56
3.2 Introduction 57
3.3 Methods 58

3.4 Results 65
3.5 Discussion 75

Chapter 4 QUANTIFICATION OF UGT2B GENE EXPRESSION IN HUMAN TISSUES 83

4.1 Abstract 84
4.2 Introduction 84
4.3 Methods 85
4.4 Results 87
4.5 Discussion 98

Chapter 5 ROLE OF HNF3α AND Oct-1 IN THE HEPATIC REGULATION OF UGT2B10 103

5.1 Abstract 104
5.2 Introduction 105
5.3 Methods 106
5.4 Results 112
5.5 Discussion 120

Chapter 6 *IN VITRO* REGULATION OF UGT2B10 EXPRESSION VIA L-SULFORAPHANE TREATMENT 125

6.1 Abstract 126
6.2 Introduction 127
6.3 Methods 128
6.4 Results 130
6.5 Discussion 132

Chapter 7 FUTURE DIRECTIONS AND FINAL CONSIDERATIONS 136

7.1 Conclusions and Future Directions 137
7.2 Final Considerations 143

REFERENCES 148

LIST OF FIGURES

Figure 1.1. Overview of the molecular path traveled by an orally administered drug.....3

Figure 1.2. Formation of a glucuronide conjugate.....8

Figure 1.3. Pylogenetic tree of the UGTs. 9

Figure 1.4. The UGT1 family. 10

Figure 1.5. The UGT2 family. 11

Figure 1.6. Schematic of the structural and functional domains of the UGTs. 13

Figure 1.7. Interactions of LETFs with UGT gene promoters.....23

Figure 1.8. Regulation of extrahepatic UGT1A gene promoters in the gastroinstestinal tract. 27

Figure 1.9. Nrf2-mediated induction of antioxidant response elements (AREs). 30

Figure 1.10. Location of putative nrf2/ARE binding sites in the UGT2B10 promoter. 31

Figure 1.11. The origin of alternative splicing via exonization of intronic sequences. 33

Figure 1.12. Frequency of alternative splicing. 34

Figure 1.13. Schematic representation of the UGT1A gene locus and the exon 5 alternative splicing event.....35

Figure 2.1. Ct values of 31 putative endogenous control genes in 84 human tissues.....49

Figure 2.2 Reverse transcription efficiency of the MT-ATP6 gene.....51

Figure 2.3 qPCR efficiency of the MT-ATP6 gene.....51

Figure 3.1 Schematic of UGT1A expression analysis.....66

Figure 3.2. UGT1A sequence-confirmed amplification products from human liver cDNA ...67

Figure 3.3. Validation of nested PCR methodology.....68

Figure 3.4. Relative mRNA expression levels of UGT1A variants in 10 liver samples.....70

Figure 3.5. Relative expression levels of UGT1A1 v1 and UGT1A1 v2/v3 mRNAs in 58 human liver specimens.....71

Figure 3.6. Correlation of UGT1A1 mRNA expression with HLM glucuronidation activity....73

Figure 3.7. Effect of siRNA treatment of UGT1A_v2 mRNA on raloxifene glucuronidation activity in cell lines....74

Figure 3.8 Sequence alignment of the human UGT1A exon 5b with theprimate-specific alu transposable element consensus sequence. 80

Figure 4.1. UGT2B and MT-ATP6 qPCR standard curves. 88

Figure 4.2. Tissular expression levels of UGT2B genes....91

Figure 4.3. Relative expression levels of UGT2B genes in various tissues....92

Figure 4.4. Relative expression levels of UGT2B genes in matched lung tumor and adjacent normal lung tissue specimens. 96

Figure 5.1. pGL3-Basic luciferase vector....108

Figure 5.2. Schematic representation and nomenclature of the UGT2B10 promoter constructs. 108

Figure 5.3. Schematic representation of the modified 5'RACE procedure used for mapping the UGT2B10 transcription start site. 112

Figure 5.4. UGT2B10 promoter driven luciferase activity in the HepG2 cell line.....113

Figure 5.5. Footprint analysis of the UGT2B10 promoter region....115

Figure 5.6. EMSA analysis of the UGT2B10 promoter 117

Figure 5.7. Gel supershift analysis of the HNF3α and Oct-1 binding sites 118

Figure 5.8. Mutational analysis of the UGT2B10 promoter protein binding sites....119

Figure 5.9. Schematic representation of the UGT2B10 promoter region....120

Figure 6.1. siRNA knockdown of nrf2 gene and effect on UGT gene expression levels.. 131

Figure 6.2. Effect of SFN treatment on UGT2B10 promoter-driven luciferase activity. 132

Figure 6.3. Schematic representation of the UGT2B10 core promoter region and distal repressor element. 133

Figure 7.1. Polymorphisms in the FOXA1 gene.....140

LIST OF TABLES

Table 1.1. UGT glucuronidation activity against selected substrate classes. 12

Table 2.1. Mean Ct value of 31 endogenous control genes in 6 tissue specimens from each of 14 tissue sites. 48

Table 2.2. Average Ct values of 31 endogenous control genes in all tissue samples 50

Table 4.1. Mean* relative UGT2B expression levels in human tissue specimens. 90

Table 4.2. Correlations between UGT mRNA levels in different tissue specimens. 98

Table 5.1. List of primers used for generation of UGT2B10 promoter constructs. 109

Table 5.2. Primers used for generation of UGT2B10 promoter DNaseI footprinting probes. 110

Table 5.3. TESS-predicted protein binding sites within protected footprints. 116

LIST OF ABBREVIATIONS

RACE	rapid amplification of cDNA ends
ACE	angiotensin converting enzyme
ADME	absorption, distribution, metabolism, and excretion
AhR	aryl hydrocarbon receptor
ARE	antioxidant response element
BaP	benzo[a]pyrene
CAR	constitutive androstane receptor
Cdx	caudal-related homeodomain protein
Co-IP	co-immunoprecipitation
CYP	cytochrome P450
DHT	dihydrotestosterone
DMEM	Dubelco's modified eagle medium
EMSA	electrophoretic mobility shift assay
ER	estrogen receptor
FXR	farnesoid X receptor
GSP	gene-specific primer
GST	glutathione-S-transferase
GWAS	genome-wide association study
HCC	hepatocellular carcinoma
HLM	human liver microsome
HNF	hepatocyte nuclear factor
HPLC	high-performance liquid chromatography

Keap1	Kelch-like ECH-associated protein 1
LETF	liver-enriched transcription factor
LXR	liver X receptor
MDR	multi-drug resistance protein
MLM	mouse liver microsome
mRNA	messenger RNA
MT-ATP6	mitochondrial ATP-synthase subunit 6
NAT	N-acetyltransferase
NEAA	non-essential amino acids
NNAL	4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	4-(methylnitrosamino)- 1-(3- pyridyl)-1-butanone
NST	nucleotide sugar transporter
OAT	organic anion transporter
OLZ	olanzapine
ORF	open reading frame
PAH	polycyclic aromatic hydrocarbon
PPAR	peroxisome proliferator-activated receptor
PXR	pregnane X receptor
qPCR	quantitative polymerase chain reaction
RIN	RNA integrity number
RNP	ribonucleoprotein
RQ	relative quantification
rRNA	ribosomal RNA
RT-PCR	reverse-transcription polymerase chain reaction
SAHA	suberoylanilide hydroxamic acid

SCR	scrambled
SFN	sulforaphane
SGA	second generation antipsychotic
siRNA	small, interfering RNA
SNP	single nucleotide polymorphism
snRNA	small nuclear RNA
SULT	sulfotransferase
TE	transposable element
TESS	transcription element search software
TPMT	thiopurine methyltransferase
TSNA	tobacco-specific nitrosamine
UDPGA	5’-diphospho-α-D-glucuronic acid
UGT	uridinediphosphate glucuronosyltransferase
UPLC	ultra-performance liquid chromatography
UTR	untranslated region
VKORC1	vitamin K epoxide reductase complex 1
WT	wild-type
XRE	xenobiotic response element

ACKNOWLEDGEMENTS

I would first like to thank Dr. Philip Lazarus for his years of support and mentorship. He has helped shape me as a scientist and has given me the tools I will need to succeed as I begin my career as an independent investigator. I would also like to thank all past and present members of the Lazarus lab as well as the members of my committee, Dr. Yun, Dr. Young, Dr. Spratt, and Dr. Richie for aiding me in my journey. Finally, I would like to thank my family and friends, especially my wonderful and gracious wife, Anna. Her willingness to care for our daughter, Shelley, and put up with many late nights and long weekends was instrumental in allowing me to achieve this milestone.

Chapter 1

REVIEW OF THE LITERATURE

1.1 General Background and Significance

Genetic and environmental factors work together to determine interindividual differences in disease onset and progression and response to therapeutic agents. While environmental and behavioral factors can be altered, genetic factors are static, and thus represent an important area of research for determining individuals that will be at risk for disease and how they might respond to treatment. It is important to understand a patient's genetic profile and environmental exposures in order to determine the appropriate medical course of action that will maximize efficacy and minimize toxicity. For diseases with complex etiology and vast interindividual variability such as cancer, it is difficult to account for all of the variables that will determine a patient's risk of disease and response to treatment. The science of pharmacogenetics attempts to elucidate the genetic factors behind these differences, and the union of pharmacogenetics and human genomics has greatly increased the rate of discovery of genetic factors that underlie variability in drug response and toxicity as well as interindividual differences in cancer risk [1]. Sequencing of the human genome has enabled some of these goals to come to fruition, and within the next five years, the cost of sequencing an individual's genome will be less than $1000. The impact of this on patient care will take time to develop as we gain more insight into how different genetic profiles affect disease risk and therapeutic outcomes.

1.1.1. Drug Metabolism and Genetic Variation

Pharmacogenetics is the study of how genetic differences lead to differences in the pharmacokinetic and pharmacodynamics properties of a drug (Figure 1.1). Pharmacokinetic properties include absorption, distribution, metabolism, and excretion (ADME), and pharmacodynamic properties include drug-receptor interactions and intracellular signaling

cascades. Genetic differences in any of these pathways can lead to interindividual differences in drug efficacy and side effects [2, 3], and studies of these pathways can lead to the identification of biomarkers for disease and toxicities that can be used in a clinical setting [4, 5]. While diseases are often complex and heterogenous, the goal of pharmacogenetic research is to deliver effective and safe medication to individuals with suitable genotypes.

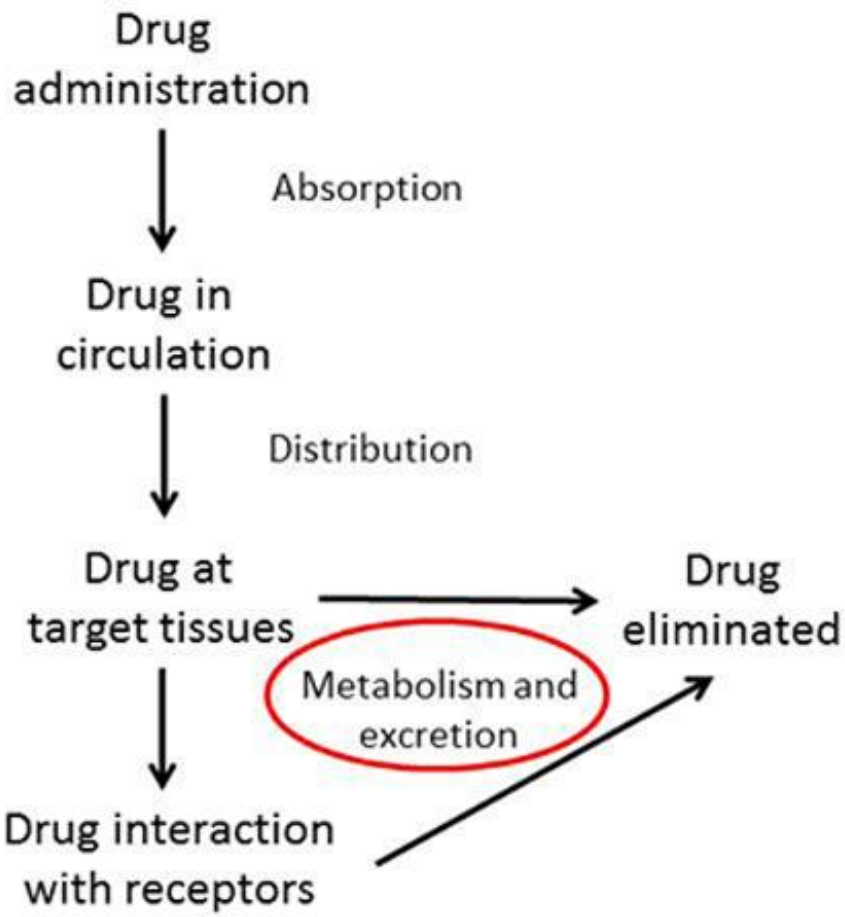

Figure 1.1. Overview of the molecular path traveled by an orally administered drug. Genetic factors that affect any aspects of the pharmacokinetic or pharmacodynamic properties of a drug can influence efficacy and toxicity. Metabolic pathways are the focus of this dissertation (red circle).

As mentioned, variation in any of the pathways involved in pharmacokinetics and pharmacodynamics can affect drug response and toxicity by affecting dose response curves. For

example, warfarin is used to inhibit vitamin K epoxide reductase complex 1 (VKORC1) [6]. The VKOR enzyme converts vitamin K epoxide to vitamin K, which is a cofactor utilized by several enzymes involved in blood clotting. Warfarin blocks the production of vitamin K, leading to reduced clotting rate. Numerous variants of the VKORC1 gene are associated with warfarin resistance and the need for higher dosages [7-10]. Variability in this gene is estimated to be responsible for at least 25% of the interindividual variability in warfarin dose [11]. Other examples of differences in drug targets that affect drug response include angiotensin I converting enzyme (ACE) and ACE inhibitors [12] and the β2-adrenergic receptor and albuterol [13].

Drug metabolism pathways have been a major focus of pharmacogenetic studies, because interindividual differences in the expression and activity of these enzymes cause clinically significant effects on the kinetic properties of various drugs. For example, cytochrome P450 (CYP) enzymes are involved in the metabolism of most clinical drugs. Interindividual differences in the expression and activity of these enzymes can affect plasma concentrations, drug activation, and detoxification [1]. Amongst these enzymes, CYP2D6 metabolizes ~25% of known drugs [14]. There are numerous variant CYP2D6 alleles that have a strong impact on drug response. For example, genetic differences in CYP enzymes have been used to determine the most appropriate breast cancer therapies [15, 16], and there are rapid screening methods available for CYP2D6 genotypes, which physicians can use to adjust dosages of certain therapeutic agents [1]. In the case of tamoxifen, individuals with CYP2D6*3 allele receive no benefit because they do not readily convert it into its more potent metabolite, endoxifen [17]. The importance of CYP2D6 variation was recognized even before genotyping, as individuals were classified into four phenotypic groups with respect to hydroxylation of debrisoquine: ultra-rapid metabolizers, extensive metabolizers, poor metabolizers, and intermediate metabolizers [1]. Individuals within each of these groups have different phenotypes with regard to debrisoquine kinetics and drug response. CYP2C19 is another example of a polymorphic CYP enzyme that is involved in the

metabolism of several commonly administered drugs. Individuals with high activity CYP2C19 alleles do not respond as well to the proton pump inhibitor omeprazole and require higher dosages to receive the same degree of stomach acid neutralization [18]. CYP2C9 variant alleles are associated with greater risk of adverse events from warfarin treatment and require less frequent dosing [19]. More than 20 variants of CYP3A4 have been identified, with varying effects on phenotype and drug response [20, 21].

In addition to CYP enzymes, interindividual differences in several other drug metabolizing enzyme families are important in determining drug response and toxicity. The thiopurine methyltransferase (TPMT) enzyme detoxifies thiopurine chemotherapy drugs, and reduced function variants of this gene are associated with higher toxicity, requiring dose reduction [21]. Reduced function variants of the N-Acetyltransferase 2 (NAT2) gene are associated with reduced clearance of the tuberculosis drug, isoniazid [22], leading to profound neural toxicity.

Variants of the uridinediphosphate glucuronosyltransferase (UGT) gene family are also associated with drug toxicity and efficacy. The most studied example is the UGT1A1*28 allele and toxicity of the colon cancer drug, irinotecan. This allele causes reduced transcription of the UGT1A1 gene, leading to less active protein [23-25]. Individuals homozygous for this allele are at risk of severe side effects from irinotecan treatment, including potentially fatal neutropenia and severe diarrhea [24, 25]. The UGT1A1 enzyme is responsible for detoxification of SN-38, the active metabolite of irinotecan and with less active UGT1A1 protein available, SN-38 accumulates to toxic levels. Dosages must be adjusted accordingly, and genetic screening is recommended for individuals that are candidates for this drug.

1.1.2. Cancer Risk and Genetic Variation

Genome-wide association studies (GWASs) have identified common genetic differences that are associated with risk of different types of cancer [26-29]. Because of their role in the metabolism of chemical toxins and carcinogens, genetic differences in drug metabolizing enzymes are frequently associated with cancer risk [30]. Interindividual variability in metabolic enzymes can alter the carcinogenicity of compounds by affecting how quickly they are metabolized, detoxified, and excreted. They can also affect how quickly the carcinogens are transported into a cell as well as the rate at which they are bioactivated into carcinogenic derivatives [30]. Genetic differences in enzymes involved in any of these pathways can affect carcinogenicity and cancer risk. For example, the CYP family of enzymes is known to be the major group of enzymes involved in the bioactivation of procarcinogens into genotoxic metabolites. Individuals bioactivate procarcinogens at different rates, and those that do so more rapidly are at higher risk for certain cancers because the body cannot detoxify the reactive intermediates quickly enough to prevent DNA damage. This is best exemplified by the decreased risk of tobacco carcinogen-related cancers in individuals with defective CYP2A6 alleles. CYP2A6 bioactivates carcinogens in tobacco smoke and metabolizes the addictive compound nicotine. Genetic differences that lead to a less active enzyme cause significantly reduced risk of lung and head and neck cancers, as individuals with these alleles tend to smoke less due to higher circulating levels of nicotine [31]. Numerous groups have studied this association, predominantly in Asian populations where variant alleles are most prevalent [32-39], and a recent meta-analysis of studies involving the CYP2A6*4 gene deletion found an odds ratio of 0.25 for tobacco-related cancers [40]. Many other studies have looked at interindividual differences in CYP enzymes and found increases in risk for various types of cancer [41], suggesting that interindividual differences in carcinogen bioactivation is an important factor in cancer risk.

In addition to enzymes associated with bioactivation, interindividual differences in carcinogen detoxification enzymes have also been associated with cancer risk. Unlike bioactivation, which is predominantly performed by the CYP450 enzymes, detoxification is catalyzed by several different classes of enzymes, including glutathione-S-transferases (GSTs), UGTs, sulfotransferases (SULTs), and NATs [30]. Factors affecting the expression or activity of these enzyme families have been associated with cancer risk. For example, GSTs are involved in the detoxification of reactive electrophile species caused by tobacco smoke, and the prevalent GSTM1 null allele has been associated with increased risk of lung cancer [42-44]. Interindividual variability in the NAT2 gene causes individuals to be either slow or rapid acetylators. Slow acetylators are at significantly increased risk for bladder cancer, as this enzyme is responsible for the detoxification of a group of potent bladder carcinogens, the aromatic amines [45]. Thus, the slow acetylator phenotype causes a reduced rate of carcinogen clearance and increased carcinogenicity. Interindividual differences in UGT genes have also been associated with various cancers. UGTs perform the majority of detoxification against environmental carcinogens, and interindividual differences in their expression and activity are a major factor in determining cancer risk. The UGT genes are the focus of this dissertation, and a detailed discussion of UGTs and cancer risk is included in subsequent sections.

The majority of pharmacogenetic studies involving drug metabolizing enzymes have focused on genetic variations that affect the structure or activity of the enzyme, but it is becoming increasingly clear that this model is insufficient. There are many factors governing interindividual variation in drug-metabolizing enzyme expression and activity including single nucleotide polymorphisms (SNPs), epigenetics, alternative splicing events, transcriptional regulation, and post-translational modifications [46]. These factors that potentially modify the expression and activity of drug metabolizing enzymes carry important implications for understanding interindividual differences that affect cancer risk and drug metabolism.

1.2. UDP-Glucuronosyltransferases

Endogenous and exogenous compounds are metabolized in two steps in order to make them more readily excreted. The first step, termed Phase I metabolism, creates a functional group on the compound, most commonly a hydroxyl (OH), amine (NH_2), or sulfhydryl (SH). The CYP enzymes are responsible for the vast majority of phase I metabolism, accounting for greater than 80% of the reactions for clinical drugs. The phase two enzymatic reactions involve the conjugation of a polar group to the functional group that was unmasked by the phase I reaction. The addition of the polar moiety makes the compound more water soluble, thereby facilitating excretion in the urine, bile, and feces [47]. The various families of phase II enzymes include GSTs, methyltransferases, NATs, SULTs, and UGTs. Like the CYPs for phase I metabolism, the UGT enzymes catalyze the majority of phase II metabolism, accounting for more than half of the phase II metabolism of clinical drugs [48]. As shown in Figure 1.2, the UGT catalyzed reaction involves the addition of glucuronic acid from a 5'-diphospho-α-D-glucuronic acid (UDPGA) molecule to one of the functional groups unmasked by the phase I reaction [49]. The most common reactions for the UGT family of enzymes involve the conjugation of the glucuronic acid moiety to OH or NH_2 groups. While the majority of human UGTs catalyze O-glucuronidation, UGTs 1A4 and 2B10 are specific for N-glucuronidation to amine groups [49].

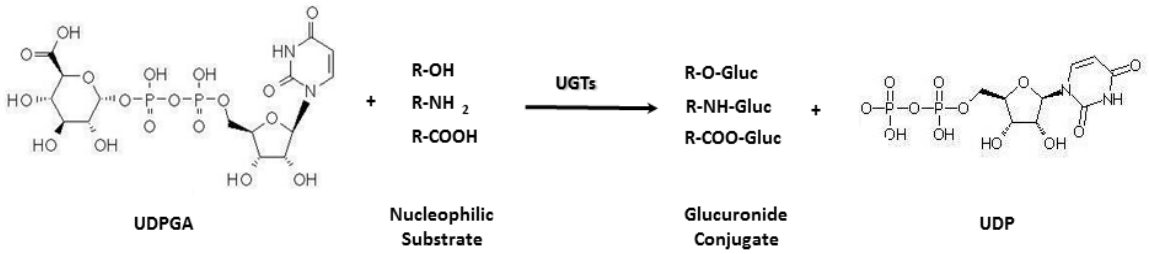

Figure 1.2. Formation of a glucuronide conjugate. The UGT catalyzed reaction involves the addition of a glucuronic acid moiety to a nucleophilic substrate, utilizing the cosubstrate UDPGA. This increases the polarity of the parent compound and facilitates excretion

1.2.1. UGT families

The UGT genes are divided into four families, UGT1, UGT2, UGT3, and UGT8. Families 1, 2, and 3 are known to metabolize drugs, whereas the UGT8 family does not. The DNA homology of the UGT enzymes is at least 40 percent across families and at least 60 percent within a family [47]. The UGT1A and UGT2B families perform the majority of drug metabolism and a dendrogram of these genes is shown in Figure 1.3. UGT genes encode proteins that are 520-540 amino acids in length and share highly conserved C-terminal regions that contain the UDPGA binding site, transmembrane domains, and endoplasmic reticulum (ER) localization signal. The less conserved N-terminal region confers substrate specificity.

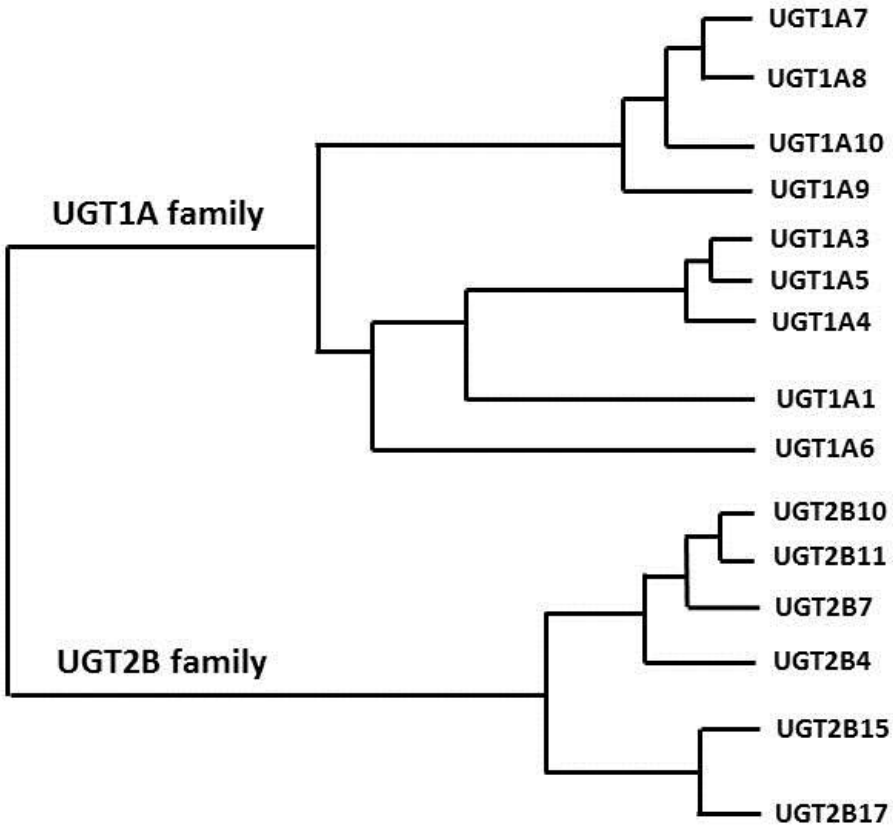

Figure 1.3. Pylogenetic tree of the UGTs. Shown in this dendrogram are those sequences encoding exon 1 regions of the proteins.

The UGT1A gene locus spans more than 200kb of DNA on chromosome 2q37 (Figure 1.4) [50]. Thirteen separate genes are encoded by this locus, of which 9 encode active proteins and 4 encode pseudogenes. The UGT1A genes are combined through exon-sharing, not alternative splicing [51-53]. Each of the UGT1A genes has a unique first exon that confers substrate specificity and shares a common set of exons 2-5 [54]. Even though the UGT1A genes share a substantial portion of the coding sequence, they are each driven by their own promoters, giving them different tissue-specific expression patterns.

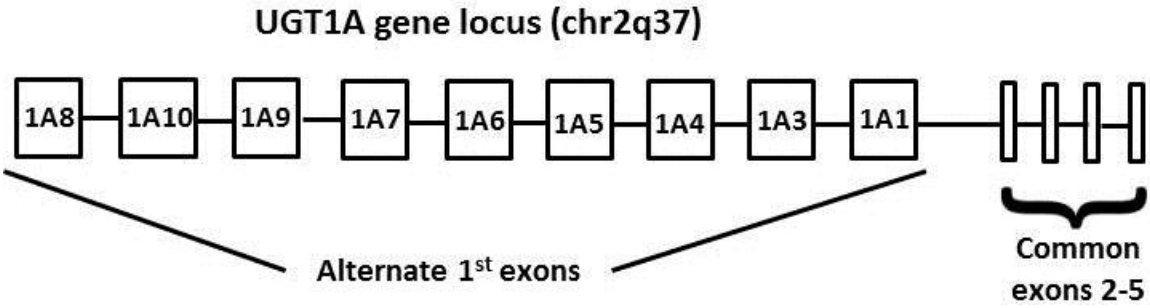

Figure 1.4. The UGT1 family. Alternate UGT1A first exons can be joined to common exons 2-5. Pseudogenes of the UGT1A family are omitted.

The UGT2 family is subdivided into UGT2A and UGT2B genes located in a cluster on chromosome 4q13 (Figure 1.5). UGT2A family members include 2A1, 2A2 and 2A3. For UGT2A1 and 2A2, the first exon is unique and the two genes share exons 2-6, similar to the UGT1A family [55], and UGT2A3 is a unique 6-exon gene. Recent studies in our laboratory have found the UGT2A family to be expressed in a wide range of tissues [56]. Seven functional UGT2B genes and six pseudogenes have been identified, each composed of six exons. The sequence homology is greatest in exons 2-6, which contain the UDPGA binding domain, and the homology is least in exon 1, which confers substrate specificity [57]. The UGT2B locus evolved from whole gene duplication events [58].

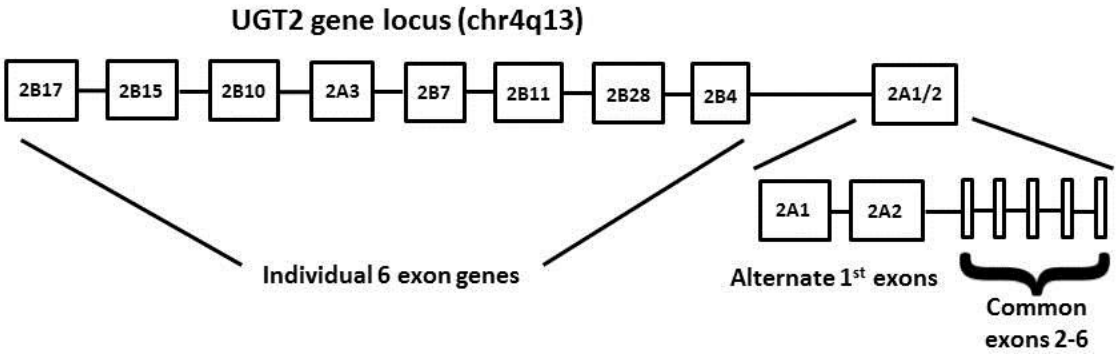

Figure 1.5. The UGT2 family. Each human UGT2B gene consists of 6 exons. The UGT2A1/A2 genes undergo exon sharing similar to the UGT1A locus. Pseudogenes of the UGT2 family are omitted.

The most recently identified UGT family was the UGT3 family, which is located on chromosome 5p13 and is comprised of UGTs 3A1 and 3A2 [59-61]. These enzymes share only around 30% sequence homology with UGT1A and UGT2B families and differ from these families by their use of cosubstrates other than UDPGA. UGT3A1 uses UDP-N-acetylglucosamine while UGT3A2 uses UDP-xylose and UDP-glucose [59, 60]. UGT3A1 has been found to be active against bile acids and estrogens, while UGT3A2 is active against classic UGT substrates such as 4-MU and 1-hydroxypyrene as well as estrogens and biflavones [61].

The UGT8 family is not a drug metabolizing family of UGT enzymes. A single gene, UDP-galactose ceramide galactosyl transferase, is encoded by this locus and is found on chromosome 4q26. This enzyme utilizes UDP-galactose as a cosubstrate [62, 63] and is involved in the biosynthesis of cerebrosides and glycosphingolipids, which are important for myelin biosynthesis. Because of its role in the synthesis of myelin, this enzyme is expressed predominantly in cells of the central nervous system.

1.2.2. UGT Function and Localization

UGT enzymes evolved in order to facilitate the excretion of both endogenous and exogenous substrates. The relative activity of different UGT isoforms against a variety of chemical families is shown in Table 1.1. It has been demonstrated that there is considerable redundancy in UGT substrate specificity [49, 64, 65]. Endogenous substrates include estrogens, androgens, bile acids, eicosanoids, and bilirubin. Glucuronidated substrates are generally inactive, with some notable exceptions, such as morphine-6-glucuronide [66].

Table 1.1. UGT glucuronidation activity against selected substrate classes.

Chemical class	UGTs responsible for glucuronidation
Bilirubin	1A1
Simple phenols	1A1, 1A3, 1A6,, 1A7, 1A8, 1A9, 2A1, 2B15
Complex phenols	1A1, 1A3, 1A6, 1A7, 1A8, 1A9, 2A1, 2B15
Aliphatic alcohols	1A9, 2A1, 2B7
Anthraquinones/flavones	1A1, 1A3, 1A8, 1A9, 2A1, 2B15
Coumarins	1A1, 1A3, 1A6, 1A7, 1A8, 1A9, 2A1, 2B15
Bile acids	1A3, 2B4, 2B7
Carboxylic acids	1A3, 1A9, 2A1
Primary amines	1A4, 1A6, 1A9, 2B10
Secondary amines	1A4, 2B10
Tertiary amines	1A4, 2B10
Heterocyclic amines	1A3, 1A6, 1A8, 1A9, 1A10
Opiods	1A3, 1A8, 2A1, 2B7
C18 steroids	1A1, 1A3, 1A8, 2A1, 2B7
C19 steroids	1A4, 1A8, 2A1, 2B15, 2B17
C21 steroids	1A4, 2A1, 2B15

UGTs are integral membrane proteins that reside in the ER because of a dilysine motif and transmembrane domain near the C-terminus [67, 68]. The transmembrane domain anchors UGT enzymes to the ER membrane [69] while the majority of the protein including the active site resides in the lumen of the ER (Figure 1.6) [70-72]. UGTs have been reported to interact with each other, and co-immunoprecipitation (co-IP) experiments have shown that individual isoforms can interact not only with each other but different UGT isoforms as well [73-75]. The interactions are thought to occur through the transmembrane domains, and *in vitro* studies suggest that protein-protien interactions between UGTs may alter activity and substrate specificity [76]. In addition, the glucuronidation activity of several UGT isoforms has been shown to require phosphorylation [77-80]. The rate-limiting step in UGT catalysis is transport of UDPGA from the cytosol to the lumen of the ER by nucleotide sugar transporters (NSTs) [81-83]. Substrates that have been glucuronidated are transported out of the cell by multi-drug resistance proteins (MDRs) and organic anion transporters (OATs) [84-86].

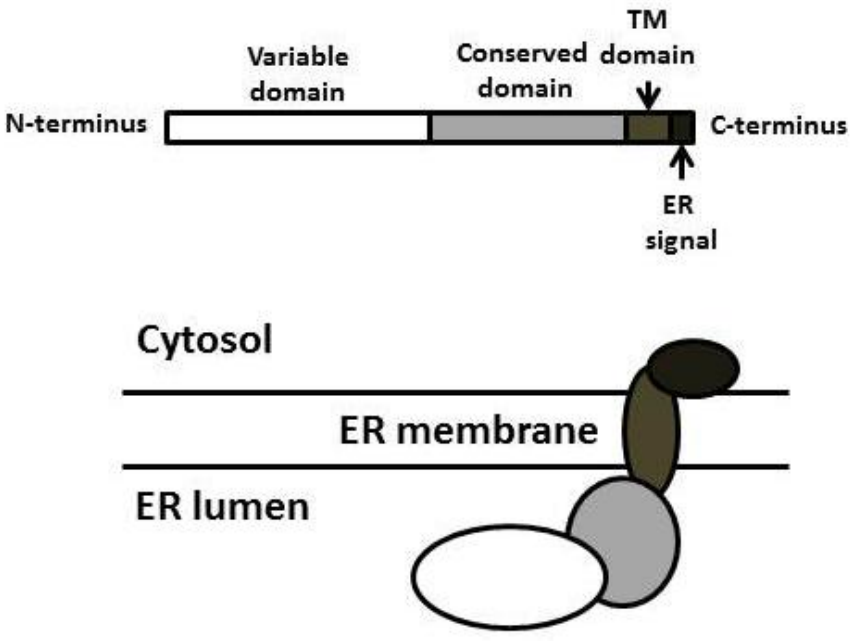

Figure 1.6. Schematic of the structural and functional domains of the UGTs.

Species to species comparison of UGT expression and activity has not been characterized. Even though most rodent species express a number of UGT isoforms, they are generally poor models because few UGTs are truly orthologous between humans and rodents [49]. In addition, putative rodent orthologs have very different substrate specificities and expression patterns [87]. Recently, a chimeric mouse with a humanized liver was developed, and was shown to have the UGT distribution of normal human liver [88]. A transgenic mouse that expresses the UGT1A locus has also been developed [89]. These may prove to be useful animal models for studying the kinetic properties of drugs and carcinogens.

1.2.3. Importance of UGTs for Metabolizing Endogenous Compounds

Defects in glucuronidation of certain endogenous compounds can present a major health risk for the organism. One of the most studied and clinically relevant genetic differences with regard to UGT activity is the UGT1A1 enzyme and glucuronidation of bilirubin, a toxic breakdown product of heme metabolism. Polymorphisms in this gene can lead to reduced activity against bilirubin, which causes accumulation at levels that become toxic to neurons and glial tissue when it crosses the blood brain barrier. Because there is no redundancy for bilirubin metabolism in the UGT metabolic pathway, polymorphisms in the UGT1A1 gene that reduce its activity can be fatal [90]. Hyperbilirubinemia is caused by at least 60 distinct polymorphisms in the UGT1A1 gene, and can lead to variable levels of activity against bilirubin [51]. The most severe cases are called Crigler-Najar syndrome type I, which leads to complete knockout of UGT1A1 activity and rapid accumulation of bilirubin in the neural tissues after birth, causing death in early infancy [51]. Type II Crigler-Najar is less severe and individuals with this genetic disorder can live into adulthood [47]. A third genetic condition that affects bilirubin metabolism is Gilbert's syndrome, present in ~10% of the population [91, 92]. Gilbert's is much less severe

than Crigler Najar, and individuals with Glbert's do not manifest clinical symptoms under normal circumstances, though jaundice can occur in these individuals when they are injured, stressed, or sick [49].

Gilbert's is not caused by a coding region polymorphism in the UGT1A1 gene, but rather a polymorphism in the TATAA box of the promoter, in which there is an additional TA repeat, which leads to reduced transcription and ultimately less functional protein. The wild-type (WT) UGT1A1*1 allele contains 6X TA repeats in the TATAA box, whereas the variant UGT1A1*28 allele contains 7X TA repeats. Reduced transcription of the UGT1A1*28 allele has been reported in several *in vitro* models and *in vivo* tissue samples [23-25]. Other alleles of this gene include one with 8X TA (1A1*37) repeats and one with 5X TA repeats (1A1*36) [93], and it has been found that the more TA repeats that are present, the less transcriptional activity of the promoter [93].

1.2.4. UGT Pharmacogenetics

Because of the redundancy in function of the various UGT isoforms, there are a relatively large number of functional polymorphisms in these proteins that do not cause overt harm to the organism. For example, both copies of the entire UGT2B17 gene is deleted in 10% of Caucasians [94], and the UGT2B10 gene contains a knockout polymorphism at codon 67 that has a minor allele frequency of 10% [95]. While these are examples where the UGT gene is effectively silenced, there are plenty of examples where UGT polymorphisms cause more modest decreases or increases in glucuronidation activity. Depending on the nature of the substitution, SNPs in the UGT genes can also cause increased activity against certain substrates and decreased activity against others [96-100]. Even though these variations do not obviously affect the organism under normal circumstances, many of these differences have been associated with altered drug response

and cancer risk. There have also been reports of copy-number variation in UGT genes that may contribute to interindividual variability in expression and activity [101-103], as well as alternatively spliced variants of the UGT1A, UGT2B4, and UGT2B7 genes [104-106].

Glucuronidation of pharmacological agents is an important function of UGT enzymes. There are numerous genetic differences in UGT genes that cause altered drug metabolism, which in some cases lead to toxicity or subtherapeutic doses [17, 107, 108]. The pharmacogenetic significance of interindividual differences in the expression of UGT genes is exemplified by the UGT1A1*28 allele and its effect on the metabolism of the chemotherapeutic agent irinotecan. Irinotecan is a topoisomerase inhibitor used in the treatment of colon cancer, and its active metabolite SN-38 is predominantly glucuronidated by UGT1A1. Individuals with the UGT1A1*28 allele display reduced clearance of SN-38 allowing it to bioaccumulate to toxic levels, leading to potentially fatal neutropenia and diarrhea [109]. Individuals that are candidates for irinotecan chemotherapy are now routinely screened with UGT1A1 genotyping assays to determine whether they have the variant allele, and dosages are adjusted accordingly [110].

Glucuronidation of the active metabolites of the breast cancer drug tamoxifen is impaired by the UGT2B7 H268Y variant, and this allele may account for some of the observed interindividual variability in patient response and toxicity [111]. UGT enzymes also affect the pharmacokinetic parameters of second generation antipsychotics (SGAs). For example, lower plasma concentrations of olanzapine (OLZ) are seen in individuals with the UGT1A4*3 allele [112]. The lower levels are likely due to faster clearance of the drug, as the *3 allele has higher glucuronidation activity than the wild type allele against certain substrates [112]. Recent studies in our laboratory discovered that the UGT2B10 D67Y allele is associated with reduced glucuronidation activity against OLZ in human liver microsomes (HLMs) [113], and future studies are being proposed to assess its impact in a clinical setting. Metabolism of the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) is affected by the UGT2B17

deletion polymorphism *in vitro*[114]. This is a promising new drug for the treatment of a wide range of different cancers, and the potential clinical significance of this finding is unknown. It may be possible to adjust the dose of this drug based on the reduced activity seen in individuals that are homozygous for the UGT2B17 deletion allele. The pharmacokinetic properties of raloxifene and morphine are also affected by UGT pharmacogenetics [115, 116].

1.2.5. UGT Variants and Cancer Risk

Detoxification of environmental carcinogens is another important function of UGT enzymes. There are many families of mutagenic carcinogens that are detoxified by UGTs. These include heterocyclic amines, aromatic amines, polycyclic aromatic hydrocarbons (PAHs) and tobacco-specific nitrosamines (TSNAs) [117-120]. A major source of exposure to environmental carcinogens is tobacco smoke, and UGT enzymes are involved in the detoxification of many of the potent and abundant carcinogens present in tobacco smoke, including TSNAs and PAHs [100, 121-125]. Variability in the expression and activity of enzymes involved in the detoxification pathway of tobacco smoke carcinogens can affect cancer risk.

For example, low activity alleles of UGT1A7, including the *3 and *4 alleles are associated with increased risk of orolaryngeal cancer risk in Caucasians and African Americans [126]. In addition, a recent meta-analysis of UGT1A7 found an association of the *3 allele with risk of colon cancer [127]. UGT1A7 is an extrahepatic enzyme that is well-expressed in aerodigestive and digestive tract tissues and is highly active against carcinogenic metabolites of benzo[a]pyrene (BaP) [128], so the increase in cancer risk seen in these tissues is consistent with its expression pattern and function. The UGT1A10 E139K allele is associated with increased risk of orolaryngeal cancer [129], and this enzyme is also important in the detoxification of PAHs found in tobacco smoke [122]. The UGT2B17 deletion polymorphism causes an increase in lung

adenocarcinoma risk in women [130]. UGT2B17 is one of the predominant UGT enzymes involved in the detoxification of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), and its mRNA has been detected in human lung tissue. In addition, NNAL has been associated with lung adenocarcinoma in animal models, so this histology-specific association is consistent with its function. Thus, individuals lacking this allele will have less ability to detoxify this carcinogen. It is unclear why the association with lung cancer risk is gender-specific, though there are interesting gender differences in expression of UGT2B17 which may help explain this difference [131]. UGT2A1 is well-expressed in the lung, and the UGT2A1 K308R SNP causes reduced activity against PAHs [56], potentially leading to increased risk of squamous cell carcinoma of the lung (R. Bushey and P. Lazarus, unpublished). This enzyme is involved primarily in the detoxification of PAHs, which have been hypothesized to cause squamous cell carcinoma of the lung [132]. UGT1A8 genetic variants also affect metabolism of tobacco smoke carcinogens [133], and the UGT1A1*28 allele has been linked to breast cancer risk as well as decreased glucruonidation activity against BaP [134, 135].

1.2.6. The UGT2B10 D67Y Allele

A polymorphism in the UGT2B10 allele at codon 67 (D>Y) has a knockout effect on UGT2B10 activity [95, 136-138]. The aspartate residue is predicted to be a part of the catalytic domain of the enzyme, and this polymorphism has been shown to affect UGT2B10 activity against pharmacological substrates (OLZ), toxins (nicotine), and carcinogens (NNAL). It is a good example of genetic variation within the UGT family of genes that affects both drug and carcinogen metabolism.

Glucuronide conjugates of NNAL and other TSNAs account for 60-90% of the total TSNA levels in urine [139, 140]. The glucuronide conjugate of NNAL is non-tumorigenic [141]

and is a major route of detoxification of this potent carcinogen. NNAL can be glucuronidated at the carbinol group or the ring nitrogen, and the *O*-glucuronides (NNAL-*O*-gluc) and *N*-glucuronides (NNAL-*N*-gluc) have an average ratio of 1:1 in smoker's urine [139]. UGT2B10 is the most active UGT with respect to N-glucuronidation of NNAL, and a D67Y polymorphism in the coding region of the gene has a significant impact on glucurondation of this compound in HLMs and is associated with significantly lower levels of the N-gluc in smoker's urine [138, 142]. Factors affecting the polymorphic expression of UGT2B10 may also affect interindividual variability in the carcinogenic potential of NNAL, though no studies have attempted to look at the regulation of UGT2B10 expression.

Nicotine is the only compound present in tobacco that causes addiction. Treatment and prevention of tobacco addiction requires a better understanding of the pharmacological aspects of nicotine addiction, including how the kinetics of nicotine clearance is affected by interindividual differences in metabolism. Nicotine is structurally similar to NNAL, and it has been shown that up to 55% of urinary metabolites of nicotine are phase II glucuronide products, including nicotine-gluc, cotinine-gluc, and 3-OH-cotinie gluc [143]. Studies have shown that N-glucuronidated products of nicotine account for 22% of the urinary metabolites [95, 144], and UGT2B10 is the predominant enzyme involved in nicotine-N-gluc formation in human liver [95]. The UGT2B10 D67Y polymorphism causes a significant decrease in nicotine-N-glucuronide formation in HLMs from individuals with this allele [95]. The UGT2B10 D67Y allele is an important determinant of nicotine metabolism in smokers, and this may potentially modify interindividual differences in the addictiveness of this compound [95]. It is unknown whether other factors affecting the expression and activity of UGT2B10 may affect nicotine glucuronidation pathways.

In addition to the role UGT2B10 plays in the metabolism of NNAL and nicotine, it is also important in the glucuronidation of OLZ [113]. The D67Y allele of UGT2B10 is also

associated with reduced activity of HLMs against this therapeutic agent, and it has been hypothesized that this polymorphism may also be linked to the side effects of SGAs, including severe weight gain [113]. This polymorphism may contribute to the interindividual variability seen in OLZ plasma concentrations, and future studies will likely be conducted to address this issue [145-147]. Other factors that affect the expression and activity of UGT2B10 may also modulate the pharmacokinetic properties of this important antipsychotic agent.

1.3. Expression and transcriptional regulation of UGTs

It is clear that variation in the UGT family of genes is important for modifying interindividual differences in drug response as well as cancer risk. Some of the genetic factors underlying these associations have been discovered, but often there is wide variability in phenotype within a given genotype. This is likely because other factors affecting the expression and activity of the UGTs are playing a role in these differences. It has been hypothesized that interindividual differences in transcriptional regulation contributes to this variability. The UGT1A1*28 allele causes relatively modest interindividual differences in transcriptional activity yet it has very important clinical effects. This may be true for other UGT genes, but little research has been conducted regarding this hypothesis.

The average human gene contains 126 polymorphisms, only 5 of which occur in coding regions [148]. Recent studies have demonstrated that variation in gene regulatory regions is more common than originally anticipated, and several studies have shown that polymorphisms which affect gene expression of drug metabolizing enzymes can have an impact on drug response and toxicity [149, 150]. The following is a review of what is currently known regarding interindividual differences in expression and transcriptional regulation of UGTs.

1.3.1 Tissular Distribution and Interindividual Differences in UGT Gene Expression

At the level of mRNA, each UGT has been shown to have its own tissue-specific expression pattern [151]. Because of the high degree of homology between different UGT isoforms within the same family, antibody development has been slow, and exhaustive studies on protein expression of UGT genes have not been conducted. The primary organ for detoxification in the human body is the liver, and as such the mRNA expression of nearly all UGT1A and UGT2B genes has been described, with 1A7, 1A8 and 1A10 the notable exceptions [152, 153]. Quantitative studies have determined that the expression of UGTs in the liver is high relative to other organs, and it is generally assumed that the majority of physiologically-relevant metabolism occurs here [152, 153].

It is becoming increasingly clear that extrahepatic glucuronidation is also important. For example, in the GI tract, there is high expression and activity of UGTs 1A7, 1A8 and 1A10 [154-156]. This means that even before a xenobiotic compound enters the body, it is already subjected to glucuronidation. Interindividual differences in the expression and activity of these enzymes in the gut can potentially affect how much drug is absorbed in different individuals, as most glucuronide conjugates are too polar to pass through the intestinal epithelium [80]. UGT1A and UGT2B genes are also expressed in a wide array of other aerodigestive tract tissues that are highly exposed to chemical toxins and carcinogens [125]. Their presence in these tissues suggests that, in addition to hepatic metabolism, extrahepatic tissues play an important role as a first line of defense against genotoxic exposures. The importance of *in situ* metabolism is supported by studies that show that UGT-deficient rat skin fibroblasts are significantly more sensitive to micronucleus formation upon treatment with BaP and 4-(methylnitrosamino)- 1-(3-pyridyl)-1-butanone (NNK) than mice with normal UGT activity [157, 158]. The effect is also significant in mice with heterozygous UGT deficiencies, which suggests an extrahepatic

genoprotective role for UGTs against two potent and abundant carcinogens found in tobacco smoke. In addition to tissues that form a barrier between the internal body and the environment, UGTs are also well-expressed in steroid-responsive tissues, including the breast, prostate, and ovaries [153, 159]. In these tissues, UGTs are known to terminate steroid signaling, and high levels of C19 steroid-glucuronides have been found in the human prostate, breast cyst fluid, and ovary follicular fluid [160-162]. They may have a genoprotective role in these tissues as well, considering that several steroid metabolites, such as the catechol estrogens, are genotoxic. A better understanding of the tissular distribution of UGT enzymes is important for understanding where interindividual differences are relevant to cancer risk and pharmacogenetics. The relative levels of UGT mRNA in pooled RNA samples have been assessed by real-time PCR. The expression pattern for each gene varies widely, which is likely due to tissular differences in expression of the transcription factors necessary for UGT expression. Thus, in order to fully understand UGT pharmacogenetics, transcriptional regulatory mechanisms need to be more completely defined. The transcriptional regulation of many UGT genes has been explored in the liver, but only a few of the factors regulating extrahepatic expression have been investigated [163].

There are also large interindividual differences in the expression of UGT isoforms within the same tissue [164], and this may be due to genetic differences in the transcription factors that regulate UGT expression. For example, a quantitative study by Izukawa et al showed that in 25 individual liver specimens, UGT levels varied more than 100-fold for each hepatic UGT species [152]. The interindividual difference in expression was 158-fold for UGT1A1, 666-fold for UGT1A3, 1433-fold for UGT1A4, 700-fold for UGT1A6, 441-fold for UGT1A9, 2050-fold for UGT2B4, 487-fold for UGT2B7, 361-fold for UGT2B10, 359-fold for UGT2B15, and 314-fold for UGT2B17 [152]. This substantial interindividual variability in mRNA levels is likely due to a number of factors, including nutritional, hormonal, and polymorphic regulation of the UGT

family of genes. It may also be due to methodological issues, such as poor choice of internal control genes or poor RNA quality, both of which can affect mRNA quantification via real-time PCR. While some factors that contribute to interindividual variability in UGT gene expression have been elucidated, there is still a poor understanding of how these genes are expressed and regulated throughout the body.

1.3.2. Hepatic Transcriptional Regulation of UGT Gene Expression

Hepatic regulation of UGT gene expression is mediated to a large extent by liver-enriched transcription factors (LETFs). Hepatocyte nuclear factor (HNF) 1α is the most important identified transcription factor for regulating hepatic UGT expression [165-172]. HNF1α is a trans-acting factor required for the activation of many liver-specific genes. It recognizes the consensus binding sequence GTTAATNATTAAC [173], and is a member of the POU subgroup of the homeodomain protein superfamily [163]. It can bind to and activate its consensus sequence as a homodimer or as a heterodimer with HNF1β [174]. HNF1α is predominantly expressed in digestive tract tissues, with highest expression levels occurring in the liver and intestines, and it has been shown to be important for activating transcription of UGT1A and UGT2B family members in liver- and intestine-derived cell lines. The HNF1α site is conserved in sequence and location in all UGT1A gene promoters, though UGTs 1A7, 1A8, and 1A10 are only expressed in extrahepatic tissues (Figure 1.7) [175].

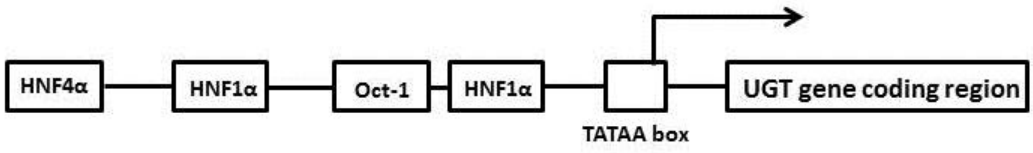

Figure 1.7. Interactions of LETFs with UGT gene promoters. Hepatic transcriptional regulation of UGT gene promoters has been described for several UGT isoforms. HNF1α has been determined to be the most important factor in regulating UGT gene expression.

UGT2B gene expression has been shown to be at least partially under HNF1α transcriptional control. All of the UGT2B gene promoters contain HNF1α sites conserved in sequence and location, and all are expressed in liver [176]. Other transcription factors are known to modify HNF1α-mediated regulation of UGT2B gene expression. For example, a study involving the UGT2B17 promoter identified a pre-B cell homeobox (Pbx) site adjacent to the HNF1α site, and it was discovered to be capable of blocking HNF1α-mediated activation [177]. In another study, octamer transcription factor-1 (Oct-1) was found to bind to the UGT2B7 promoter, but it was found to be incapable of activating transcription. In co-transfection experiments, it was found that Oct-1 can enhance HNF1α-mediated promoter activation [178].

Even though HNF1α regulates the expression of UGTs *in vitro* it may only be part of the story. In HNF1α-/- mice, most genes thought to be dependent on HNF1α expression are only reduced by about 20-30%, including albumin, which is considered to be a paradigm of a gene controlled by HNF1α [179]. The fact that the HNF1α-controlled genes are only modestly down-regulated in its absence suggests that other transcription factors are important for the expression of these genes as well. The notable exception is phenylalanine hydroxylase, which is not expressed at all in HNF1α-/- mice, leading to the observed phenylketonuric phenotype [179].

In addition to HNF1α, some UGT gene promoters have been found to be regulated by HNF4α. UGT1A9 has multiple HNF4α sites in the proximal promoter that are necessary for maximal activation of the promoter [180]. HNF4α is a member of the zinc-finger transcription factor family and binds as a homo- or hetero- dimer to the RGGNCAAAGKTCR consensus DNA sequence [180]. HNF4α levels have been found to correlate with expression levels of UGTs 1A6 and 1A9 in human livers, but not 1A1, 1A3 or 1A4 [181]. In addition, HNF4α-/- mice displayed lower mRNA expression levels of Ugts [182], and small, interfering RNA (siRNA) knockdown of HNF4α in human hepatocytes led to a significant decrease in the expression of several UGTs [183]. HNF4α binding sites are also a major reason for the hepatic expression of UGT1A9, while

the highly homologous isoforms 1A7, 1A8, and 1A10 lack these sites and are only expressed extrahepatically [184].

While HNF4α and HNF1α clearly play important roles in the regulation of hepatic UGT gene expression, other LETFs are likely to play a role as well. For example, it was recently discovered that the UGT2B17 gene is regulated by HNF3α in the human prostate cancer cell line LNCaP. There is also evidence that HNF3α regulates the expression of other metabolizing enzyme families. For example, the HNF3α protein has been found to be important in the basal expression of CYP3A4 in HepG2 cells, along with the Sp1, AP1, and CEBPα proteins [185]. HNF3 family proteins are imperative for normal development of endoderm-derived organs such as the liver, pancreas, and lungs [186]. In the adult, HNF3 family members are predominantly expressed in the liver and pancreas, where they play prominent roles in the regulation of glucose metabolism [187]. HNF3α in particular is encoded by the FOXA1 gene and is a member of the forkhead box family of transcription factors [188]. The consensus binding sequence for HNF3α is A(A/T)TRTT(G/T)RYTY [189]. HNF3α is important in the regulation of gluconeogenic enzymes [190], and in HNF3α-/- mice, there was a 70% decrease in pancreatic proglucagon gene expression leading to perturbations in glucose homeostasis [191]. HNF3α has also been shown to be regulated by nutritional factors and hormones, which may influence interindividual variability in activation of its target genes [192].

Polymorphisms in the genes encoding LETFs may influence interindividual variability in UGT expression and activity. For example, a polymorphic variant of the HNF1α gene leads to the synthesis of a protein that lacks the transactivation domain, making it a non-functional product that dimerizes with the WT protein *in vitro* and acts in a dominant negative fashion to block HNF1α-mediated gene transcription [193]. Individuals with reduced function HNF1α variants are at higher risk for diabetes, because of the role HNF1α plays in the activation of genes involved in glucose metabolism [194]. HNF1α variants also have an effect on UGT gene

expression. In HepG2 cells, activation of the UGT2B17 promoter was reduced by transfection with variant alleles of the HNF1α gene [193]. Little is known about how these variants may affect interindividual variability in glucuronidation capacity.

Polymorphic variations in LETF cognate binding sites within UGT gene promoters may also have a clinically relevant impact on UGT gene expression [195, 196]. A recent study of the UGT2B17 gene promoter revealed that a SNP in the HNF3α binding site of the UGT2B17 promoter (-155A>G) causes 13-fold lower luciferase activity in LNCaP cell line [197]. It was also determined that individuals with this SNP have significantly lower circulating levels of androstane-3α-17β-glucuronide. Although the SNP was not associated with prostate cancer risk, the study suggests that polymorphic regulation of the UGT genes via LETFs can have a significant impact on blood levels of important endogenous compounds, and in this case a potential carcinogen as the initial stages of prostate cancer are androgen-dependent.

The UGT2B7 promoter is highly polymorphic, with 8 frequent polymorphisms in the proximal 1248 bp [193, 198, 199]. These SNPs are inherited as one of three common haplotypes and have a significant impact on promoter activity *in vitro* [198]. With respect to morphine glucuronidation, these haplotypes do not appear to be clinically important, though further studies on UGT2B7 glucuronidation activity are needed before conclusions can be drawn [200]. Polymorphic T-box regions adjacent to the transcriptional initiation site in the promoters of the UGT1A7, 1A8, 1A9, and 1A10 have variable impact on promoter activity. In the UGT1A9 gene, a variant allele with a tenth T in this region leads to a 2.6-fold increase in luciferase activity in HepG2 cells [201]. In contrast, extra T's in the UGT1A10 T-box do not have a significant effect on promoter activity.

1.3.3. Extrahepatic Transcriptional Regulation of UGT Genes

Studies of UGT1A promoters in Caco-2 cells as an *in vitro* model of enterocytes have identified important differences between the UGT1A promoters that account for some of the differences in tissue-specific expression (Figure 1.8). UGTs 1A8 and 1A10 both contain binding sites for the caudal homeodomain transcription factor caudal-related homeodomain protein2 (cdx2), which has been shown to synergistically activate UGT1A promoters in the presence of HNF1α [202]. UGT1A7 is not expressed in the intestines, and it is the only one of the three extrahepatic UGT1A genes that does not contain the intestine-specific cdx2 binding site in the proximal promoter. This is an example of a tissue-specific transcription factor that regulates UGT1A expression in the gastrointestinal tract.

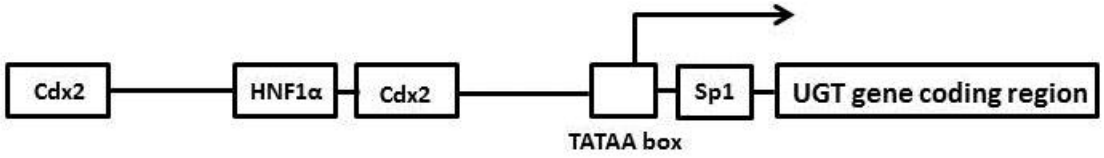

Figure 1.8. Regulation of extrahepatic UGT1A gene promoters in the gastroinstestinal tract. The extrahepatic UGT1A gene promoters have been found to be regulated by tissue-specific transcription factors such as cdx2. HNF1α is also expressed and is important for regulation of UGT gene expression in the gut.

The extrahepatic tissue distribution of UGT2B genes varies, which has been hypothesized to be the result of differences in tissue-specific transcription factor binding sites [176]. Because of their importance in steroid metabolism, several studies have attempted to characterize the transcriptional regulation of UGT2B gene promoters in cell lines derived from steroid-responsive tissues such as the prostate. The UGT2B17 promoter was found to be differentially regulated in the prostate-derived cell line LNCaP compared to the liver-derived cell line HepG2 [167]. Transfection of the pCAT reporter plasmid containing varying lengths of UGT2B17 promoter into HepG2 and LNCaP cell lines resulted in differential transcriptional activation. In LNCaP

cells, deletion of the region between -196 bp and -53 bp resulted in a two-fold decrease in activity while in HepG2 cells it led to a two-fold increase in activity. This led the authors to focus on the proximal ~200 bp for identification of transcription factor binding sites. Mutational analysis using luciferase constructs revealed that the HNF1α site between -40 to -52bp was critical for promoter activation in HepG2 cells but not in LNCaP cells, which do not express HNF1α. It was further shown through co-transfection experiments that HNF1β could activate the UGT2B17 promoter in LNCaP cells, and this activation could occur even when the HNF1α site was mutated. These results provide *in vitro* evidence for tissue-specific regulation of the UGT1A and UGT2B genes, though many of the factors regulating the tissular distribution of the various isoforms remain unknown.

The UGT2B10 promoter is one of the few remaining UGT gene promoters that have not been studied to date. UGT2B10 is important for the detoxification of carcinogens, therapeutic agents, and nicotine [95, 113, 138, 142], making it important to understand how this gene is regulated. One of the main goals of this dissertation research is to characterize how the UGT2B10 gene promoter is regulated and how the regulation may contribute to interindividual differences in expression and activity of this enzyme.

1.3.4. Xenobiotic Regulation of UGT Gene Expression

Interindividual variability in UGT gene expression due to exogenous factors has also been explored. Several ligand-activated transcription factors including the aryl hydrocarbon receptor (AhR), nuclear factor-erythroid 2-related factor 2 (nrf2), constitutive androstane receptor (CAR), pregnane X receptor (PXR), farnesoid X receptor (FXR), liver X receptor (LXR), and peroxisome proliferator-activated receptor (PPAR) are involved in regulating UGT expression [203-213]. In tissues exposed to tobacco carcinogens, the inducible expression of UGT1A genes

via the AhR ligand-activated receptor is of particular importance. Many of these enzymes can be induced by PAHs that bind to AhR and activate xenobiotic response elements (XREs) in the promoters of the genes [209, 214]. UGTs 1A1, 1A6 and 1A9 are induced by various compounds through this element. The induction capabilities are imperative for detoxification of carcinogens, especially considering the fact that many carcinogens induce the enzymes that detoxify them. A 1200 bp deletion polymorphism in the UGT1A10 promoter includes an XRE [215], though the pharmacogenetic implications of this deletion are not fully understood. Because of the role UGT1A10 plays in detoxifying tobacco smoke carcinogens such as PAHs, this deletion polymorphism may influence interindividual differences in expression and activity against these compounds, ultimately impacting an individual's cancer risk.

The XRE in the UGT1A1 promoter is induced by glucocorticoids, carcinogens, antibiotics, and isothiocyanates [216-220]. Interestingly, some compounds like benzo[a]pyrene induce multiple UGTs (UGTs 1A1, 1A6 and 1A9), whereas others such as chrysin will induce specific isoforms (UGT1A1) [216, 221]. Endogenous substrates such as steroids also regulate UGT expression. When treated with dihydrotestosterone (DHT), expression of UGT2B15 is down-regulated and UGT2B11 is up-regulated in human primary prostate epithelial cells [222]. Decreased expression of UGTs 2B15, 2B17, and increased expression of UGT2B11 is also seen in the prostate-derived cell line LNCaP upon treatment with DHT [223, 224]. Estrogen regulation of UGT1A and UGT2B expression has been studied in the estrogen receptor (ER) positive cell line MCF-7. UGT1A10 and UGT2B15 expression were found to be increased upon treatment with 17β-estradiol, and the expression of other UGT2B genes was found to be unchanged [225].

UGT genes are known to be induced by the nrf2 transcription factor. Nrf2 induces expression of phase II genes in response to chemopreventive agents, and it is thought to play an important role in chemoprevention pathways. Supporting this hypothesis is the fact that nrf2

knockout mice are not protected by chemopreventive agents to the same degree as wild type mice [226]. In addition, phase II gene expression is not inducible by chemopreventive agents in nrf2-/- mice, suggesting that this is the predominant transcriptional activator involved in induction by these compounds [226]. Under normal cellular conditions, nrf2 is sequestered in the cytoplasm by the Kelch-like ECH-associated protein 1 (Keap1) [227]. Nrf2 activators are thought to disrupt the interaction between nrf2 and Keap1 by modifying cysteine residues (Figure 1.9) [228].

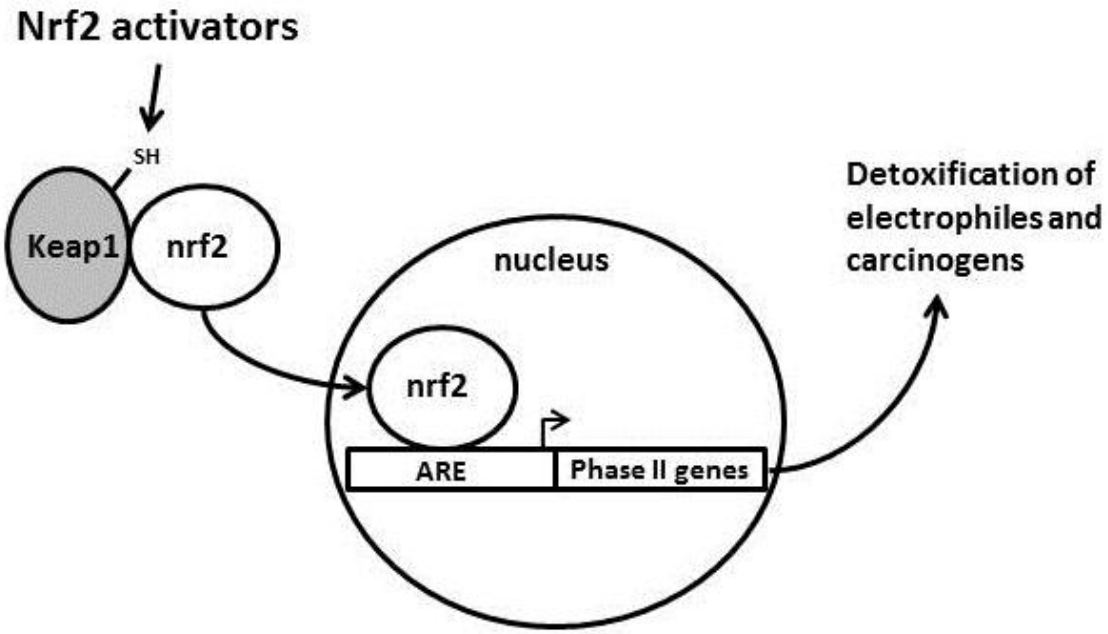

Figure 1.9. Nrf2-mediated induction of antioxidant response elements (AREs). Nrf2 activators modify cysteine residues in the Keap1 protein, disrupting the interaction with nrf2. Free nrf2 protein can then translocate to the nucleus and induce phase II gene transcription.

Sulforaphane (SFN) is the most widely studied phase II enzyme inducer with regard to nrf2 induction pathways. Glucosinolate is a biological precursor of SFN and is abundant in cruciferous vegetables. Much of the chemopreventive effects associated with SFN have been attributed to the induction effects on glutathione S-transferase enzymes [229]. UGTs have also been shown to be induced by SFN [230]. For example, the human UGT2B7 promoter is known to be induced by SFN, and a polymorphism in the nrf2 binding site blocks this induction [231].

Polymorphic variations in the nrf2 binding domain of other UGT gene promoters may have similar effects, though this has not been investigated.

In humans, the induction of UGTs by nrf2 activators may be an important mechanism for reducing carcinogen potency. This is evidenced by the fact that human UGTs that are induced by SFN are known to metabolize tobacco smoke carcinogens such as NNAL and PAHs. Despite the importance of UGT2B10 in the detoxification of NNAL, no studies have looked at potential nrf2-mediated induction of this gene. There are several putative nrf2 binding sites between -1650 bp and -3029 bp of the UGT2B10 promoter that may be active (Figure 1.10). The distal location of putative nrf2 binding sites in the UGT2B10 promoter is consistent with the location of active antioxidant response elements (AREs) in other UGT gene promoters. For example, the UGT1A1 gene promoter contains active ARE elements between -3499 bp to -3210 bp that are induced by SFN treatment *in vitro* and in transgenic mice [230]. The UGT2B7 promoter is also induced by SFN through a distal ARE element [231]. A more complete understanding of the regulation of the UGT2B10 gene promoter will help elucidate potential polymorphic variations in its expression. One of the goals of this dissertation research is to determine whether SFN has an effect on UGT2B10 gene expression.

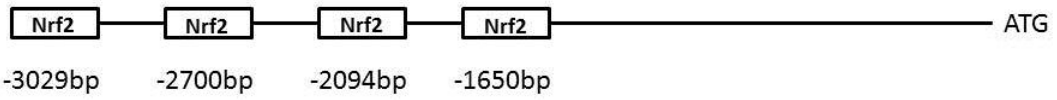

Figure 1.10. Location of putative nrf2/ARE binding sites in the UGT2B10 promoter.

1.4 Alternative Splicing of UGTs

Splicing of mRNA is the process by which intronic sequences are removed and exonic sequences are ligated to from the final, processed transcript [232]. The spliceosome is

responsible for the majority of mRNA splicing in humans and is composed of ribonucleoproteins (RNPs) and small nuclear RNAs (snRNAs) [233]. The recognition of exons and introns in the transcript is controlled by multiple signals, including the 5' splice site, the 3' splice site, the branch site, and the polypyrimidine tract [234]. These four factors provide the majority of exon and intron definition in humans [235]. Other factors regulating splice site selection include cis-acting sequence elements known as splicing silencers and enhancers, both of which can be intronic or exonic [232]. These elements are required for both constitutive and alternative splicing as they aid in recruitment of the spliceosomal proteins to the proper locations [236].

Alternative splicing is the process by which multiple transcripts are generated from the same parent mRNA species [237]. In the human genome, this greatly amplifies genomic diversity, creating novel protein variants with unique functions [238]. Greater than 90% of human genes undergo alternative splicing, and at least five types of alternative splicing have been identified, including exon skipping, alternative 5' splice sites, alternative 3' splice sites, mutually exclusive exons, and intron retention [232]. The choice of alternative splice variants can be regulated by tissue-specific factors, developmental stage, sex determination, physiological processes, and stress factors [239].

One mechanism that has been proposed for the origin of alternative splicing is through the exonization of intronic elements (Figure 1.11) [232]. In humans and other primates, this occurs much more frequently than in other organisms [240]. The most common intronic sequences that are exonized in the human genome are transposable elements (TEs) [240]. Of the many TEs that can be exonized, the primate-specific alu element occurs with the most frequency. There are more than 1 million alu retrotransposition events in the human genome [241]. Alu elements are typically around 300 bp in length and comprise around 11% of the human genome. Most, if not all, alu-derived exons are alternatively spliced and are frequently associated with low inclusion rates in the total transcript distribution [242]. Alu-derived exons are usually neutral or

only slightly deleterious because the novel, alternatively spliced product represents only a small percentage of the total mRNA species [243]. The major form variant upholds the original gene's function, and the minor form variant may evolve its own selective function, often tissue-specific [243]. There are numerous examples of genetic disease arising from constitutive alu splicing [244-246]. Even when incorporated in small amounts, alu-derived exons can be deleterious, such as with the neurofibromatosis type 1, which is caused by a novel alu insertion [246]. In the majority of cases, the exonized alu sequence is neutral and adds diversity to the primate genome, and for this reason, they have often been characterized as nonfunctional, evolutionary intermediates [243]. A bioinformatics approach demonstrated that the inclusion level of alu exons can be predicted based on several factors, such as flanking intron length, 5' splice site strength, 3' splice site strength, and exon length [247], with alu-derived exons generally incorporated at less than 10% of the total transcript abundance (Figure 1.12).

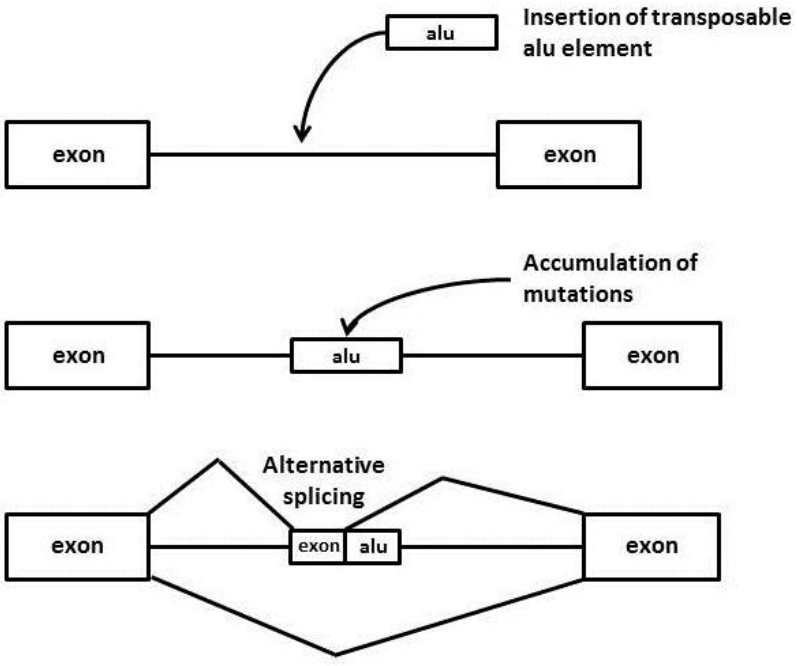

Figure 1.11. The origin of alternative splicing via exonization of intronic sequences.

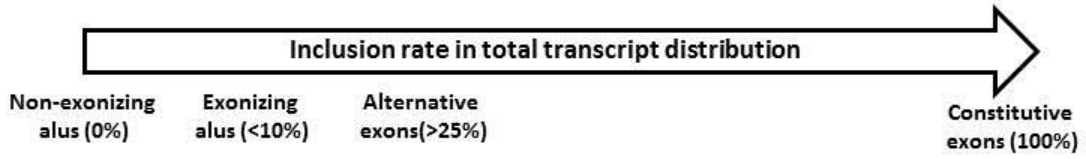

Figure 1.12. Frequency of alternative splicing. Species-specific splice variants, such as those derived from alu elements, tend to have low inclusion rates in mRNA.

1.4.1. Alternative Splicing and the UGT1A Locus

The complexity of the UGT1A gene cluster is amplified by an alternative splicing event at the 3' end of the locus. Each of the 9 UGT1As can also undergo alternative splicing at their 3' ends using an alternate exon 5, resulting in 27 different UGT1A mRNA species with each UGT1A gene encoding three different combinations of 5a and 5b UGT1A exons (Figure 1.13) [105]. UGT1A mRNAs that incorporate only exon 5a are known as variant 1 (v1) and encode the active isoform 1 (i1) proteins [104]. When the alternative exon 5b is incorporated as the last exon (termed variant 2, or v2) or when both exons 5a and 5b are incorporated together (termed variant 3, or v3), the resulting protein is catalytically inactive [104]. The v2 and v3 mRNAs each have the same open reading frame (ORF) and thus encode the same protein [termed isoform 2 (i2), [104]]. *In vitro* co-transfection studies demonstrated a dominant negative effect of i2 proteins on i1 glucuronidation activity [248-251]. In addition, concomitant expression of UGT1A i1 and i2 proteins in HEK293 cells leads to significantly decreased activity against a wide range of substrates [249]. The *in vitro* studies also suggested a direct interaction between the i1 and i2 proteins as the regulatory mechanism as opposed to competition for substrate or co-substrate, evidenced by a decrease in Vmax with little change in Km. This is supported by co-IP studies demonstrating that cell lines overexpressing both isoforms contain homo- and hetero-oligomeric mixtures of i1-i1, i1-i2, and i2-i2 proteins [252]. These studies suggest that the function of the i2 proteins may be to negatively modulate the catalytic activity of the UGT1A i1 proteins. However,

neither quantitative expression of the UGT1A v2 and v3 splice variants nor UGT1A i2-mediated regulation of UGT1A i1 activity has been explored in human tissues.

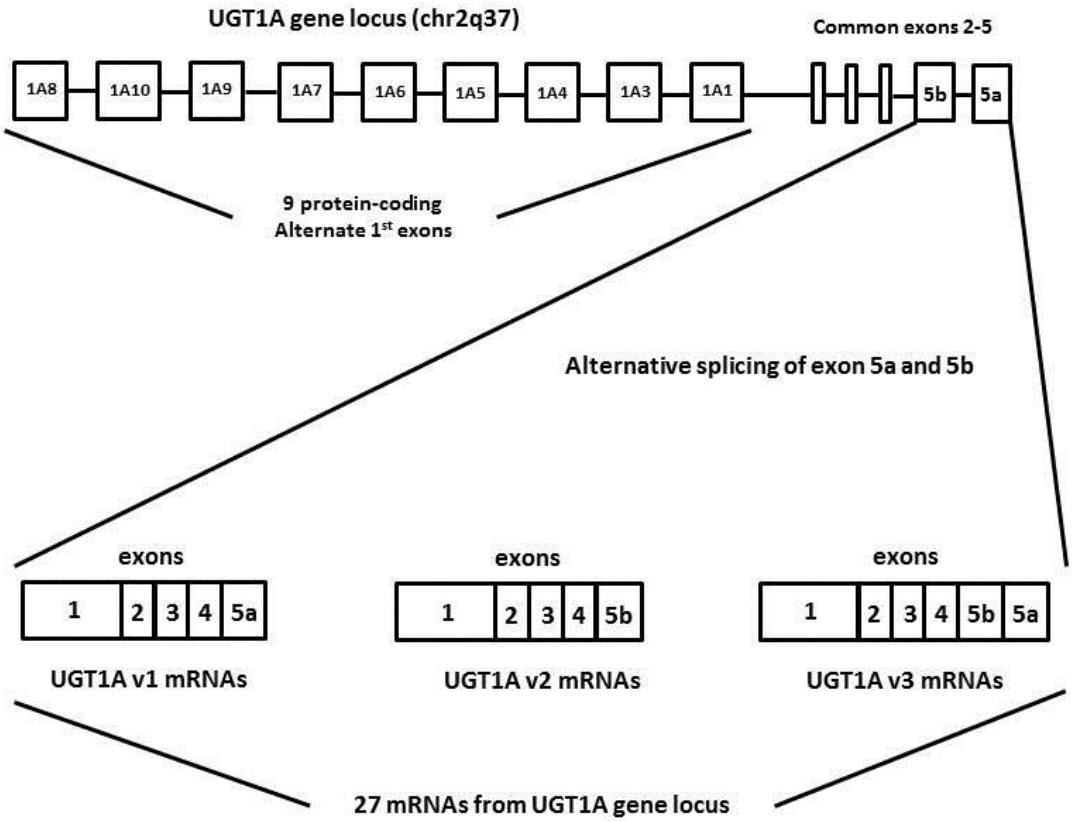

Figure 1.13. Schematic representation of the UGT1A gene locus and the exon 5 alternative splicing event.

It has been hypothesized that this alternative splicing event is a regulatory mechanism that affects cellular glucuronidation capacity against endogenous and exogenous substrates. The i2 proteins are themselves non-functional from a glucuronidation standpoint, but show a dominant-negative effect on the activity of the i1 isoforms when co-expressed. Based on these findings, interindividual variability in the v2/v3 splicing event has been hypothesized to be a major mechanism of interindividual variability in glucuronidation activity. Inter-individual variation in the relative abundance of i1 and i2 protein expression may affect glucuronidation capacity in human tissues, which would carry important pharmacogenetic implications. The i2

proteins have been shown to co-localize with UGT1A i1 proteins to the endoplasmic reticulum [252]. The same group also claimed that there is interindividual variability in the relative abundance of these isoforms between different individuals in both liver and extrahepatic microsomes [252].

The UGT1A splice variants have also been shown to be upregulated in human cancer tissue and cell lines [253]. Both constitutive and alternative splicing events are altered in most types of cancer [254], resulting in changes in the relative abundance of normal and alternatively spliced mRNA species [255, 256]. The mechanisms behind this are not clear, though different conditions are known to affect splice site selection by altering the transcription, translation, and post-translational modification of splicing factors [257]. In general, it is not yet known whether splicing defects themselves contribute to tumor progression or whether these defects are merely a consequence of the malignant phenotype [239].

1.4.2 Alternative Splicing and UGT2B Genes

The UGT2B4 gene is extensively alternatively spliced, with at least eight distinct mRNA species produced from this locus [106]. The alternative splice variants of this gene are also catalytically inactive and regulate the activity of the WT proteins [106]. In addition, it was also recently discovered that the UGT2B7 gene contains encodes additional splice variants that are catalytically inactive and have a dominant-negative effect on UGT2B7-mediated glucuronidation activity [258]. Splice variants of the UGT3A1 and UGT3A2 genes has also been reported [61], but the function of these variants is unknown. The wide abundance of alternative splicing within the UGT family of genes may be another important regulator of interindividual differences in glucuronidation capacity, but this hypothesis has not been explored.

1.5. Aims and Hypotheses

Interindividual variation in the expression of UGT genes in extrahepatic tissues has not been well explored. From this literature review, it is clear that a better understanding of the relative abundance of these genes will help determine where the various isoforms have relevance. In addition, little work has been done regarding the metabolic impact of UGT1A exon 5b splice variants on glucuronidation activity *in vivo*. Because the alternatively spliced UGT1A i2 proteins negatively modulate the activity of the WT i1 proteins, interindividual differences in the expression of the i2 isoforms might have a profound impact on glucuronidation capacity. This is an unexplored area for drug metabolizing enzymes that could prove to be a major determinant in drug and carcinogen metabolism. In addition to interindividual variability in expression and alternative splicing patterns, transcriptional regulation of the UGT genes is another area that has been only partially described. Interindividual differences in transcription of UGT genes affect both drug metabolism and cancer risk. The best example of this is the TATAA box polymorphism of the UGT1A1 gene. The same may be true for other UGT genes, but this hypothesis has not been well-explored. In particular, the UGT2B10 gene has not been characterized with regard to hepatic or ligand-activated transcriptional control mechanisms, both of which account for some of the interindividual differences seen in UGT-mediated metabolism.

There were three major aims of this dissertation research: 1) determine tissular and interindividual differences in expression of UGT1A and UGT2B family members, 2) determine to what extent interindividual differences in UGT1A alternative splicing plays in hepatic metabolism, and 3) determine what liver-enriched and ligand-activated transcription factors are responsible for the hepatic regulation of the UGT2B10 gene. The goal of the first aim was to determine the extent of interindividual variability in UGT expression in different tissues throughout the body. For the second aim, the central hypothesis was that inter-individual

variation in the relative abundance of UGT1A splice variants affects glucuronidation capacity of endogenous and exogenous substrates. For the third aim, the hypothesis was that factors affecting UGT2B10 transcriptional regulation contribute to interindividual differences in expression. As discussed, the UGT2B10 enzyme is important for the metabolism of nicotine, OLZ, and carcinogens like NNK. The importance of UGT2B10 in drug and carcinogen metabolism enzyme was only recently discovered, and factors that affect the expression and regulation of this gene are likely to have important pharmacogenetic significance.

Chapter 2

ENDOGENOUS CONTROL GENE SELECTION FOR QUANTITATIVE STUDIES OF UGT GENE EXPRESSION IN HUMAN TISSUES

2.1 Abstract

The development of real-time PCR has greatly improved accuracy, sensitivity, and reproducibility of gene expression data. Relative quantification (RQ) via the $2^{-\Delta\Delta Ct}$ method is one of the most common strategies used to analyze real-time PCR data. This type of normalization strategy relies on the use of an internal control gene. Proper control gene selection is imperative to obtain accurate RQ values. Thus the internal control gene should be expressed at similar levels in all samples, be unaffected by treatment conditions in the experiment, and occur with similar kinetics during all phases of the RT-qPCR analysis. In order to choose an appropriate control gene for subsequent studies on UGT gene expression in different tissues, it was necessary to experimentally validate specific control genes before conducting the real-time PCR experiments. In this study, 31 different control genes in 6 independent tissue samples from 11 normal tissue sites (liver, lung, colon, larynx, tonsil, tongue, mouth, brain, esophagus, pancreas, and uterus) and 3 tumor tissue sites (liver hepatocellular carcinoma (HCC), lung adenocarcinoma, and colon adenocarcinoma) was measured by real-time PCR. Of the 31 putative control genes assayed mitochondrial ATP-synthase subunit 6 (MT-ATP6) was determined to be among the 4 most stably expressed control gene in 12 out of 14 tissue sites. This study also demonstrated that MT-ATP6 is a stably expressed control gene for evaluation of expression levels between different tissue sites.

2.2 Introduction

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) is an invaluable tool for determining the abundance of specific mRNA species in a biological sample. Several fluorescence based techniques have been developed that make the sensitivity and specificity of this assay far superior to previous methods of gene quantification [259]. One of the most common real-time methods involves the use of two oligonucleotide primers at opposite ends of the amplicon and a probe that anneals in between the primers. The probe is labeled with a fluorescent marker and a quencher molecule. During each cycle of PCR, the 5' exonuclease activity of the polymerase enzyme cleaves apart fluorophore from the probe, leading to less quenching and an increase in overall fluorescence of the sample. The cycle number during which the fluorescence value passes above the background fluorescence value is referred to as the cycle threshold (Ct) of a given sample for the particular cDNA species. The quicker the fluorescence reaches threshold, the smaller the Ct value and the greater the concentration of mRNA in the original sample. Quantitation is accurate over a wide range of mRNA concentrations, and it can even be used to quantitate only a few copies of a specific transcript [260, 261].

Real-time PCR can be subdivided by two main methods of quantitation, absolute and relative. Absolute quantitation utilizes a standard curve of known concentrations of each mRNA species to extrapolate the mRNA copy number in a sample with unknown concentration. This method requires a highly purified cloned amplicon for each mRNA species to be analyzed, so it can be more time consuming than relative quantitation methodologies [262, 263].

There are a variety of methods used for RQ of mRNA expression levels [264-269]. One of the most common methods is via the $2^{-\Delta\Delta Ct}$ comparative Ct method. This method compares the

expression of a target gene to the expression of an internal control gene, which is then normalized to the expression of a calibrator sample. This method requires that the target and control genes amplify with nearly equal efficiency [270], although differences in amplification efficiency can be accounted for using serial dilutions of the given cDNA sample.

The first step in any study that is attempting to use RQ for accurate determination of mRNA levels is to choose an appropriate control gene. An ideal control gene should be expressed at similar levels in all samples, be unaffected by treatment conditions in the experiment, and occur with similar kinetics during all phases of the RT-qPCR analysis [264]. No control gene exists that satisfies all of these criteria all of the time, so it is best to experimentally validate a putative control gene for a given experiment. Common control genes include housekeeping genes (e.g. GAPDH, β-Actin) or ribosomal RNAs (rRNAs), and each has its own associated strengths and weaknesses.

Several different methods have been developed for evaluating the quality of a specific control gene with respect to other putative control genes in a given experiment [263, 268, 271-274]. The GeNorm method is a useful tool for this [275], which can be implemented by first determining the expression level of a number of putative control genes (>10) under each experimental condition. Next, for each control gene, the pairwise variation with all other control genes is calculated and used to generate a gene-stability measure, M, for each control gene [275]. Three to five control genes with the lowest M values can be used under the given experimental conditions to accurately normalize the expression levels via geometric averaging. Vandesompele et al. demonstrated that single control gene normalization can lead to greater than 6-fold error in quantitation and yet greater than 90% of studies utilize a single control gene for normalization purposes [275]. Many studies have shown that some of the most commonly used control genes are remarkably unstable in their expression patterns under different experimental conditions [276-

279]. Numerous studies have attempted to publish data related to these findings for specific experimental conditions and clinical sample sets [271, 280-283].

Tissular expression of control genes can vary widely, due in part to the heterogenous population of cell types, which can make meaningful comparison of expression levels between individual tissue specimens quite cumbersome. In this study, we present expression analysis of 31 different control genes in 6 independent tissue samples from each of 11 normal tissue sites (liver, lung, colon, larynx, tonsil, tongue, mouth, brain, esophagus, pancreas, and uterus) and 3 tumor tissue sites (liver HCC, lung adenocarcinoma, and colon adenocarcinoma). Of the 31 putative control genes assayed, MT-ATP6 was determined to be among the 4 most stably expressed control genes in 12 out of 14 tissue sites. This study also demonstrates that MT-ATP6 is a stably expressed control gene for evaluation of expression levels between different tissue sites.

2.3 Methods

Tissue specimens. The 66 histologically normal tissue specimens utilized in this study were from 11 sites (6 each from the liver, lung, colon, larynx, tonsil, tongue, mouth, brain, esophagus, pancreas, and uterus). The liver, lung, and colon tumor specimens (6 each) were histologically confirmed as HCC, lung adenocarcinoma, and colon adenocarcinoma. All normal and tumor tissue samples were obtained from the Penn State College of Medicine Tissue Bank and were quick frozen within 2 h post-surgery. All protocols involving the collection and analysis of tissue specimens were approved by the Institutional Review Board at Penn State University College of Medicine and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services.

RNA collection and cDNA synthesis. RNA was extracted from all tissue specimens with the Qiagen Rneasy Mini kit (Valencia, CA). Samples were subjected to DNAse I digestion during extraction to prevent genomic DNA contamination. RNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer, and RNA purity was assessed by absorbance ratios A260/A280 (> 1.9) and A260/A230 (>1.8). RNA integrity was determined using an Agilent 2100 Bioanalyzer with Agilent RNA 6000 Nano chips. All samples used for analysis of control gene expression levels had RNA integrity numbers (RINs) greater than 5.0. Reverse transcription was performed using the Invitrogen Superscript First Strand cDNA synthesis kit (Carlsbad, CA) with oligo dT primers and 1μg of starting RNA per sample. A negative control without RNA and a negative control without enzyme were analyzed in parallel.

Quantitative PCR (qPCR). Applied Biosystems endogenous control plates were used for determining the control genes with the greatest expression stability in the human tissue specimens (Carlsbad, CA). These plates included real-time PCR primers and probes for 32 control genes that have been established in the literature. The list of genes is as follows: 18S rRNA, ABL1, ACTB, B2M, CASC3, CDKN1A, CDKN1B, EIF2B1, ELF1, BADD45A, GAPDH, GUSB, HMBS, HPRT1, IPO8, MRPL19, MT-ATP6, PES1, PGK1, POLR2A, POP4, PP1A, PSMC4, PUM1, RPL30, RPL37A, RPLP0, RPS17, TBP, TFRC, UBC, and YWHAZ. 18S rRNA was excluded because it is not a polyadenylated RNA species and oligo (dT) was used for the RT reaction. Quadruplicate real-time PCR reactions were performed for each cDNA sample using a 10 μL final reaction volume containing 5 μL 2X Taqman Universal PCR Master Mix, 4.5 μL of diluted cDNA, and 0.5 μL of gene expression assay. Stock cDNA (50 ng/ul) was diluted 10-fold to obtain a working dilution of 5 ng/ul. A 22.5 ng RNA equivalent of cDNA was used in all reactions. Reactions were performed in 384-well plates using the ABI 7900 HT Sequence

Detection System under the following conditions: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min.

After calculating the mean Ct value and standard deviation for each of the 31 control genes in the different tissues, the expression stability of each gene was determined by ranking the standard deviations from lowest to highest. The expression stability between different tissue sites was then determined by ranking the standard deviation of the control genes in all tissue samples together.

Determination of RT-qPCR kinetics for MT-ATP6 gene and putative target genes. After determining the stability of the MT-ATP6 gene within and across different tissue specimens, it was necessary to determine whether it behaves with similar kinetics through all phases of the RT-qPCR reaction. To test the kinetics during the RT phase, a serial four-fold dilution of pooled human liver RNA (1 µg, 250 ng, 62.5 ng, and 15.6 ng) was reverse transcribed using oligo (dT) primer as described above. The resulting cDNA was then amplified with ABI real-time assays specific for 18S rRNA, MT-ATP6, and UGT2B15. UGT2B15 was used as a target for quantification because it is of much lower abundance than MT-ATP6 and transcripts with different abundance often have very different RT efficiencies. The resulting Ct values were plotted against cDNA dilution, and UGT2B15 mRNA expression levels were normalized to both MT-ATP6 and 18S rRNA to obtain RQ values for each of the RNA dilutions. RT efficiency values for target and control genes were calculated from the slope using the formula: $E_x = [2^{(-1/slope)}-1]$.

The qPCR kinetics of the MT-ATP6 gene was then tested and compared to several putative target genes. 1µg RNA from pooled human liver was reverse transcribed as described above. The resulting cDNA was serially diluted and ABI gene expression assays were used to amplify MT-ATP6, UGT2B7, UGT2B10, UGT2B11, and UGT2B15. The resulting Ct values

were plotted against the cDNA dilution, and the qPCR efficiency values for each of the real-time PCR assays was calculated from the slope using the same formula as used for the RT efficiency calculation.

2.4 Results

Expression stability of control genes in different tissues. The mean Ct values and standard deviations for each control gene in each tissue are shown in Table 2.1. The four most stably expressed control genes in each tissue are shown in bold for each tissue. In the liver, the most stably expressed genes were MT-ATP6, GUSB, POLR2A, and MRPL19. In the liver tumors, the most stable genes were TBP, MT-ATP6, POLR2A, and CDKN1A. In the rest of the tissues, the stability of expression was as follows, lung: MT-ATP6, RPLP0, PES1, HPRT1; lung adenocarcinoma: CASC3, PUM1, ABL1, ELF1; colon: MT-ATP6, POLR2A, PES1, RPS17; colon adenocarcinoma: RPL30, GUSB, MT-ATP6, CASC3; larynx: HPRT1, MT-ATP6, PES1, RPS17; tonsil: MT-ATP6, PES1, RPS17, MRPL19; tongue: MT-ATP6, PP1A, B2M, HPRT1; mouth: MT-ATP6, GUSB, GADD45A, CDKN1B; brain: MT-ATP6, MRPL19, RPL37A, RPS17; esophagus: MT-ATP6, GUSB, B2M, HPRT1; pancreas: MT-ATP6, GUSB, CDKN1B, ELF1; uterus: PES1, PSMC4, POP4, IPO8. The total number of tissues where each control gene was found to be amongst the four most stably expressed control genes is shown in the far right column. Only seven of the 31 control genes were found to be among the most stable control genes in more than two tissue sites. These included MT-ATP6 (12 of 14 sites), PES1 (5 of 14 sites), GUSB (5 of 14 sites), HPRT1 (4 of 14 sites), RPS17 (4 of 14 sites), POLR2A (3 of 14 sites), and MRPL19 (3 of 14 sites).

Expression stability of control genes between different tissues. The expression stability was examined across all tissue sites (Figure 2.2). The data indicate that nearly all of the control genes displayed a range of Ct values greater than 5.0 cycles between different tissue sites. This corresponds to at least a 32-fold (2^5) difference in expression levels between different tissues. The notable exception was the MT-ATP6 gene, which had a range of Ct values less than 4.0 cycles. The MT-ATP6 gene had 25^{th} to 75^{th} percentile of its values separated by less than 2.0 cycles. This is less than a 4-fold difference in expression for the middle 50 percent of the tissue samples. The overall stability of the control genes between different tissues is shown in Table 2.2. The only gene with a standard deviation less than 1.0 cycle is MT-ATP6. Eleven genes displayed moderate expression stability between tissues, with standard deviations less the 2.0 cycles. These included CDKN1B, ELF1, GUSB, HPRT1, MRPL19, PES1, PSMC4, PUM1, RPL30, RPL37A, and RPS17. The remaining 19 genes were determined to have very low expression stability between tissues with standard deviations greater than 2.0.

Table 2.1. Mean Ct value of 31 endogenous control genes in 6 tissue specimens from each of 14 tissue sites.

Control Gene	Liver	Liver HCC	Lung	Lung Adeno	Colon	Colon Adeno	Larynx	Tonsil	Tongue	Mouth	Brain	Eso	Panc	Uterus
E-ABL1	32.6 ± 1.63	33.4 ± 3.28	29.0 ± 0.89	**30.1 ± 0.79**	32.1 ± 2.79	29.7 ± 0.99	29.2 ± 3.21	29.6 ± 2.95	29.2 ± 1.99	30.1 ± 3.25	28.5 ± 2.11	28.4 ± 1.26	31.0 ± 1.28	26.3 ± 0.94
E-ACTB	24.6 ± 2.83	27.0 ± 5.04	20.3 ± 0.34	21.6 ± 1.51	25.6 ± 5.47	20.7 ± 1.68	19.0 ± 1.77	20.3 ± 3.03	21.7 ± 1.67	22.1 ± 3.31	20.6 ± 2.48	19.7 ± 1.91	23.8 ± 1.30	18.9 ± 1.24
E-B2M	24.9 ± 2.14	26.9 ± 4.47	20.0 ± 0.73	21.0 ± 1.57	26.5 ± 5.56	21.2 ± 1.42	22.0 ± 1.68	21.3 ± 3.04	**22.8 ± 0.73**	24.2 ± 3.05	22.8 ± 3.26	**21.4 ± 1.01**	25.5 ± 1.13	21.9 ± 0.56
E-CASC3	32.2 ± 1.10	33.4 ± 3.41	29.0 ± 0.52	**29.2 ± 0.61**	32.1 ± 2.74	**28.5 ± 0.83**	28.4 ± 2.34	29.1 ± 2.52	29.3 ± 2.57	30.5 ± 2.57	29.0 ± 2.86	28.3 ± 1.44	31.1 ± 1.23	26.9 ± 0.33
E-CDKN1A	28.3 ± 2.34	**30.0 ± 2.21**	25.7 ± 1.16	27.3 ± 1.93	30.5 ± 4.04	27.1 ± 0.99	25.8 ± 1.57	28.6 ± 2.44	28.9 ± 0.95	26.6 ± 3.53	27.0 ± 2.39	24.7 ± 2.10	28.6 ± 1.51	26.9 ± 0.87
E-CDKN1B	33.4 ± 1.43	33.7 ± 3.17	30.5 ± 0.79	29.3 ± 1.31	32.7 ± 2.72	27.8 ± 1.22	30.8 ± 3.02	30.7 ± 2.16	32.1 ± 2.53	**33.1 ± 2.47**	31.2 ± 2.60	31.3 ± 1.39	**33.2 ± 0.81**	29.0 ± 0.61
E-EIF2B1	31.1 ± 2.08	33.4 ± 3.95	28.1 ± 0.81	28.9 ± 1.68	32.4 ± 3.47	28.2 ± 1.14	27.4 ± 1.63	28.5 ± 2.33	27.8 ± 3.08	29.9 ± 3.03	27.7 ± 2.77	28.2 ± 1.94	31.2 ± 1.45	27.7 ± 1.01
E-ELF1	29.8 ± 1.76	31.3 ± 3.20	26.9 ± 0.38	**27.4 ± 1.00**	30.8 ± 3.59	26.6 ± 1.29	26.6 ± 2.37	26.6 ± 2.54	28.1 ± 1.27	29.2 ± 3.31	28.6 ± 1.96	26.3 ± 1.51	**29.6 ± 0.85**	26.5 ± 0.73
E-GADD45A	29.9 ± 1.35	32.8 ± 3.84	28.0 ± 0.60	27.9 ± 2.84	32.1 ± 3.19	28.4 ± 1.17	27.7 ± 1.66	29.9 ± 3.17	30.1 ± 1.53	**30.1 ± 2.32**	28.9 ± 4.06	26.9 ± 2.16	31.6 ± 1.52	27.6 ± 0.29
E-GAPDH	26.2 ± 2.09	26.8 ± 3.71	23.6 ± 0.87	21.6 ± 2.67	28.1 ± 5.07	21.9 ± 1.70	21.5 ± 3.68	22.2 ± 3.03	23.0 ± 1.52	23.0 ± 3.12	21.7 ± 2.87	21.6 ± 2.19	26.0 ± 0.92	22.2 ± 0.35
E-GUSB	**27.1 ± 0.68**	30.4 ± 4.10	26.6 ± 0.38	26.7 ± 1.15	31.3 ± 3.61	**26.6 ± 0.48**	27.4 ± 1.36	27.4 ± 2.19	28.0 ± 2.17	**29.1 ± 1.98**	27.2 ± 1.89	**27.4 ± 0.77**	**29.6 ± 0.64**	26.9 ± 0.25
E-HMBS	31.3 ± 1.50	32.7 ± 3.86	29.6 ± 0.43	28.9 ± 1.42	32.8 ± 3.08	27.8 ± 1.39	27.5 ± 1.67	29.0 ± 2.86	29.3 ± 2.12	30.3 ± 3.18	28.6 ± 1.87	28.4 ± 1.64	31.8 ± 0.92	29.3 ± 0.49
E-HPRT1	29.5 ± 1.69	32.0 ± 4.25	**27.8 ± 0.30**	26.9 ± 2.02	32.1 ± 3.75	26.3 ± 1.56	**27.1 ± 0.99**	27.9 ± 2.48	**28.9 ± 0.80**	29.7 ± 2.87	28.3 ± 2.19	**27.4 ± 1.17**	30.5 ± 1.17	27.7 ± 0.26
E-IPO8	31.7 ± 1.76	32.9 ± 3.97	28.0 ± 0.53	28.5 ± 1.36	32.3 ± 3.43	27.9 ± 1.12	27.9 ± 2.11	28.8 ± 2.41	29.1 ± 2.98	30.4 ± 2.82	28.0 ± 2.59	27.8 ± 1.56	31.0 ± 1.11	**27.1 ± 0.22**
E-MRPL19	**29.1 ± 0.94**	30.2 ± 3.71	28.1 ± 0.39	27.3 ± 1.33	31.6 ± 3.49	26.5 ± 0.87	27.1 ± 1.90	**27.8 ± 1.91**	28.3 ± 1.54	28.9 ± 2.51	**27.7 ± 1.06**	27.4 ± 1.56	30.6 ± 1.00	27.5 ± 0.44
E-MT-ATP6	**15.3 ± 0.64**	**17.9 ± 1.33**	**16.2 ± 0.07**	17.9 ± 1.83	**16.8 ± 2.23**	**15.2 ± 0.81**	**15.6 ± 1.14**	**16.0 ± 0.72**	**15.1 ± 0.33**	**16.5 ± 1.84**	**15.7 ± 0.70**	**14.9 ± 0.39**	**16.9 ± 0.43**	15.9 ± 0.43
E-PES1	27.1 ± 1.68	29.4 ± 2.69	**26.2 ± 0.18**	26.1 ± 1.43	**29.8 ± 2.38**	25.1 ± 0.88	**24.7 ± 1.14**	**25.6 ± 1.66**	26.3 ± 0.93	27.0 ± 2.99	26.1 ± 1.77	25.3 ± 1.36	28.3 ± 0.88	**24.7 ± 0.07**
E-PGK1	29.3 ± 1.74	30.5 ± 3.94	25.2 ± 0.35	24.9 ± 1.95	30.6 ± 4.33	24.5 ± 1.90	24.0 ± 1.49	26.0 ± 2.71	26.5 ± 1.38	27.2 ± 3.21	25.7 ± 2.63	25.0 ± 1.98	28.9 ± 1.06	25.4 ± 0.27
E-POLR2A	**31.8 ± 0.88**	**31.7 ± 1.91**	29.4 ± 0.62	29.1 ± 1.39	**32.5 ± 2.31**	30.2 ± 1.10	28.7 ± 3.66	28.6 ± 2.17	29.6 ± 1.49	30.1 ± 3.58	28.8 ± 1.79	28.7 ± 2.18	31.6 ± 1.49	27.6 ± 0.89
E-POP4	31.8 ± 1.24	33.1 ± 4.25	28.3 ± 0.45	28.5 ± 1.48	31.6 ± 3.27	27.2 ± 1.47	27.5 ± 1.41	28.4 ± 2.67	29.0 ± 1.76	30.2 ± 3.62	29.2 ± 2.77	28.0 ± 1.69	31.5 ± 1.04	**28.1 ± 0.18**
E-PPIA	24.5 ± 1.63	26.0 ± 3.92	22.0 ± 0.36	21.9 ± 1.17	27.3 ± 4.07	20.8 ± 0.94	21.1 ± 1.19	22.4 ± 2.65	**22.7 ± 0.39**	23.9 ± 3.03	23.2 ± 2.23	22.0 ± 1.86	25.6 ± 1.01	22.4 ± 0.32
E-PSMC4	29.2 ± 1.51	31.5 ± 4.15	27.5 ± 0.34	26.8 ± 1.73	31.0 ± 3.26	25.9 ± 1.70	26.2 ± 1.36	27.4 ± 2.55	28.3 ± 0.88	28.6 ± 2.80	27.5 ± 2.09	26.8 ± 1.71	30.4 ± 1.09	**26.4 ± 0.11**
E-PUM1	**30.8 ± 1.00**	32.4 ± 3.35	28.2 ± 0.53	**27.8 ± 0.70**	31.5 ± 3.05	27.8 ± 1.02	27.5 ± 2.85	27.8 ± 1.95	28.5 ± 2.29	29.0 ± 2.52	27.7 ± 2.01	27.5 ± 1.53	30.4 ± 1.23	26.5 ± 0.64
E-RPL30	24.1 ± 1.35	25.2 ± 2.96	22.0 ± 0.31	22.5 ± 1.04	27.1 ± 4.01	**21.0 ± 0.34**	21.6 ± 1.39	21.5 ± 2.21	21.8 ± 1.07	23.0 ± 3.52	23.3 ± 1.73	21.7 ± 1.50	24.7 ± 1.05	20.5 ± 1.13
E-RPL37A	23.7 ± 1.55	24.3 ± 2.91	21.6 ± 0.38	21.9 ± 1.09	25.8 ± 3.51	20.5 ± 0.88	20.7 ± 1.33	21.2 ± 2.34	21.3 ± 0.81	22.4 ± 2.99	**22.3 ± 1.35**	21.1 ± 1.54	23.9 ± 0.94	20.0 ± 0.61
E-RPLP0	26.1 ± 2.04	26.5 ± 2.49	**23.0 ± 0.15**	23.1 ± 1.82	27.0 ± 4.00	21.5 ± 1.34	21.1 ± 1.65	22.6 ± 2.69	23.4 ± 2.21	24.0 ± 3.24	23.4 ± 2.29	22.0 ± 2.01	25.5 ± 1.29	21.4 ± 0.82
E-RPS17	24.7 ± 1.57	25.5 ± 2.62	23.0 ± 0.36	23.0 ± 1.42	**26.5 ± 2.51**	22.1 ± 1.06	**21.8 ± 1.16**	**21.9 ± 1.80**	22.9 ± 1.18	23.4 ± 3.07	**24.1 ± 1.47**	22.4 ± 1.59	25.0 ± 1.01	21.3 ± 0.92
E-TBP	30.4 ± 1.62	**29.4 ± 0.53**	27.1 ± 0.44	28.1 ± 1.93	31.9 ± 4.20	26.8 ± 1.26	26.9 ± 1.58	28.2 ± 2.98	28.4 ± 2.57	29.8 ± 2.93	28.4 ± 2.56	27.7 ± 1.65	30.4 ± 1.16	26.9 ± 0.43
E-TFRC	30.5 ± 1.58	31.5 ± 4.79	26.7 ± 0.59	26.3 ± 2.85	31.5 ± 3.79	25.3 ± 3.34	27.4 ± 2.31	28.2 ± 2.44	29.4 ± 1.44	30.5 ± 3.12	28.1 ± 2.21	26.9 ± 2.13	31.3 ± 1.18	26.8 ± 0.26
E-UBC	25.9 ± 1.65	28.6 ± 4.14	23.3 ± 0.47	23.9 ± 1.59	28.3 ± 4.29	23.3 ± 1.25	22.6 ± 1.92	24.4 ± 2.90	24.2 ± 2.11	25.0 ± 3.98	23.8 ± 2.19	22.5 ± 1.93	26.4 ± 1.26	22.5 ± 0.46
E-YWHAZ	33.8 ± 1.41	32.2 ± 4.26	29.3 ± 0.42	28.8 ± 1.13	31.5 ± 3.08	26.7 ± 1.05	27.8 ± 1.94	28.4 ± 2.61	29.4 ± 2.27	29.5 ± 3.59	30.5 ± 1.79	28.3 ± 1.94	31.9 ± 1.42	29.1 ± 0.41

Ct values are given as mean ± std dev; red values indicate when the gene is among the four most stably expressed in the given tissue

*This column indicates the number of tissue sites where the listed control gene was one of the four most stably expressed

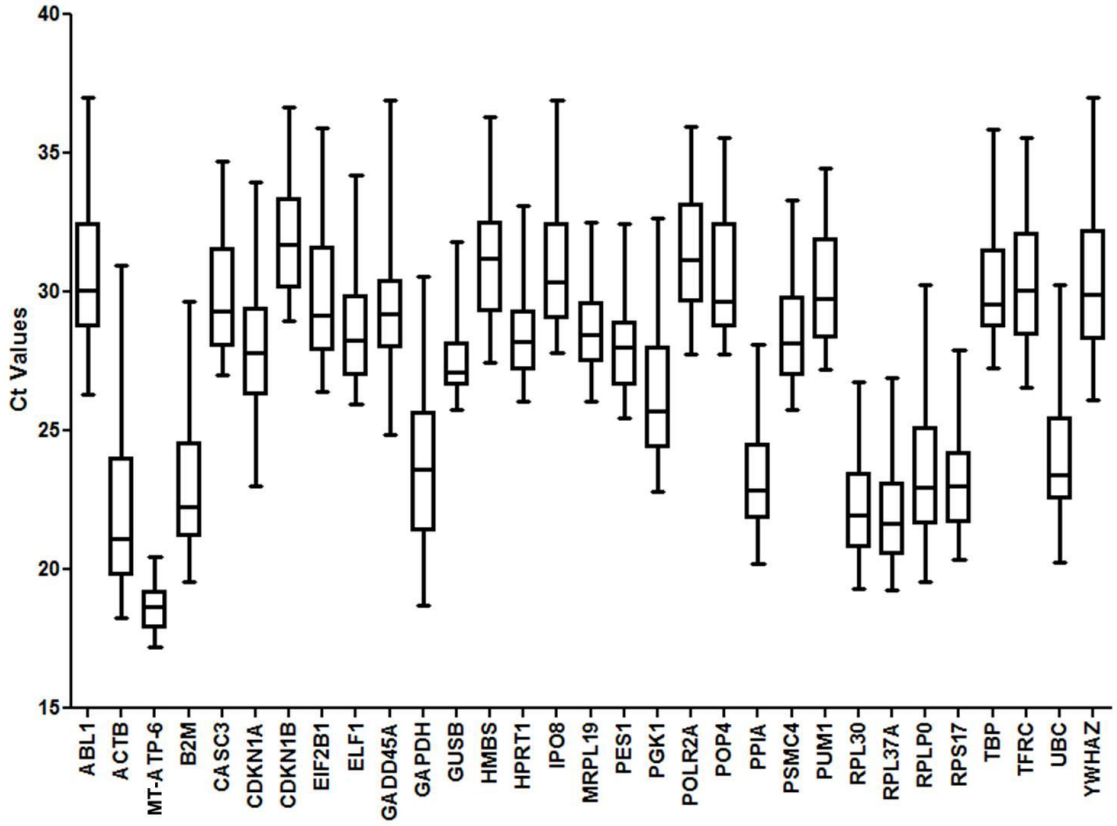

Figure 2.1. Ct values of 31 putative endogenous control genes in 84 human tissue specimens. The range of Ct values of the endogenous control genes in 6 specimens for each of 11 normal tissue and 3 tumor tissue sites is shown. Ct – cycle threshold; boxes, 25th to 75th percentile; whiskers, min to max.

Table 2.2. Average Ct values of 31 putative control genes in human tissues.

Gene	All Tissues
ABL1	30.4 ± 2.29
ACTB	21.8 ± 2.75
MT-ATP-6	**18.6 ± 0.79**
B2M	23.0 ± 2.47
CASC3	29.8 ± 2.17
CDKN1A	27.9 ± 2.35
CDKN1B	31.8 ± 1.98
EIF2B1	29.7 ± 2.21
ELF1	28.6 ± 1.95
GADD45A	29.5 ± 2.47
GAPDH	23.4 ± 2.81
GUSB	27.6 ± 1.43
HMBS	31.2 ± 2.08
HPRT1	28.5 ± 1.76
IPO8	30.9 ± 2.28
MRPL19	28.6 ± 1.59
PES1	28.0 ± 1.67
PGK1	26.2 ± 2.42
POLR2A	31.4 ± 2.10
POP4	30.4 ± 2.20
PPIA	23.3 ± 2.05
PSMC4	28.5 ± 1.96
PUM1	30.0 ± 1.97
RPL30	22.4 ± 1.84
RPL37A	21.8 ± 1.79
RPLP0	23.2 ± 2.37
RPS17	23.0 ± 1.75
TBP	30.2 ± 2.11
TFRC	30.3 ± 2.32
UBC	24.0 ± 2.21
YWHAZ	30.1 ± 2.57

Values are mean Ct ± std dev

Reverse transcription efficiency of MT-ATP6 gene. After determining the high expression stability of the MT-ATP6 gene, the kinetics of the RT reaction was tested using different dilutions of RNA. The reverse transcribed RNA was then subjected to real-time PCR amplification using ABI gene expression assays specific for UGT2B15, MT-ATP6, and 18S rRNA (Figure 2.2). It can be seen in Figure 2.2 that the RT efficiency of the MT-ATP6 gene is nearly identical to the RT efficiency of the much less abundant target gene, UGT2B15, as evidenced by the non-significant difference in slope between the two lines. On the other hand, the RT efficiency for the 18S rRNA gene is much lower than the RT efficiency of both MT-ATP6 and UGT2B15, and the slope of the 18S rRNA line is significantly different from both of them ($p<0.05$).

qPCR efficiency of MT-ATP6 gene. After determining that the MT-ATP6 gene is reverse transcribed with similar efficiency to a much lower abundance transcript, the real-time PCR efficiency of the MT-ATP6 gene was tested and compared to several putative target genes, UGT2B7, UGT2B10, UGT2B11, and UGT2B15 (Figure 2.3). Serial dilutions of pooled cDNA were amplified, and in all cases, the plots of 'average Ct value vs. log2[cDNA] dilution' yielded curves with nearly identical slopes. There were minor differences in real-time PCR efficiency between the different genes, but none

of the slopes were significantly different from each other. These experiments indicated that the MT-ATP6 gene was undergoing all phases of the RT-qPCR reaction with similar kinetics as the lower abundance transcripts.

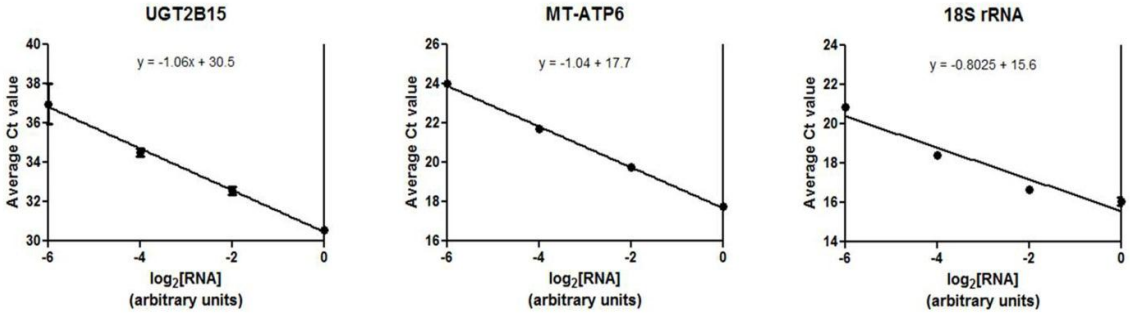

Figure 2.2. Reverse transcription efficiency of the MT-ATP6 gene. Pooled liver RNA was subjected to a serial 4-fold dilution and then amplified in quadruplicate using ABI real-time PCR specific for UGT2B15, MT-ATP6, and 18S rRNA. The resulting average Ct values were plotted against the RNA dilution factor. Linear regression was accomplished with the Graphpad Prism 5.0 software.

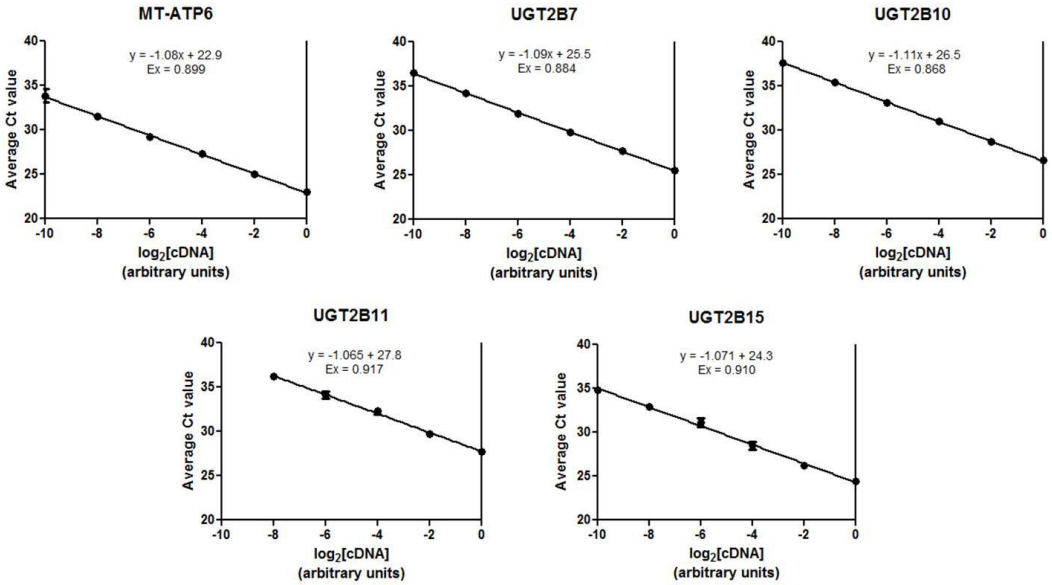

Figure 2.3. qPCR efficiency of the MT-ATP6 gene. 1 µg pooled liver RNA was reverse transcribed and the resulting cDNA was subjected to a serial 4-fold dilution. The resulting cDNA was amplified using ABI real-time PCR assays specific for the UGT2B7, UGT2B10, UGT2B11, UGT2B15, and MT-ATP6 genes. Linear regression was accomplished with the Graphpad Prism 5.0 software.

2.5 Discussion

There are numerous variables that affect the efficiency and reproducibility of RT-qPCR gene expression analysis. Total RNA has previously been used for normalization, with genes such as 18S rRNA frequently used as internal controls for RNA loading. One problem with rRNA molecules is the fact that they are not polyadenylated RNA species, prohibiting the use of oligo dT as an RT primer. This necessitates the use of random hexamers, but because of the relatively high abundance of rRNA species it can lead to different efficiencies of RT priming [284]. This is supported by the finding in this study that 18S Ct values were nearly the same for the two most concentrated RNA samples. This indicates that the 18S rRNA species is so abundant that there is a difficulty in measuring the baseline in the qPCR reaction [284]. Therefore, it is apparent that the use of 18S rRNA as a normalization gene is suboptimal because the kinetics of the reaction occurs with different efficiency for certain phases of the quantitation. On the other hand, MT-ATP6 gene was shown to display similar kinetics compared to various target genes during both the RT and qPCR phase of the RT-qPCR, indicating that it is a viable option as a normalization gene. Mitochondrial genes such as MT-ATP6 are polyadenylated, making it appropriate to use oligo (dT) as an RT primer, allowing for enrichment of the cDNA with polyadenylated RNA species.

Two important caveats exist in the use of a mitochondrial gene for normalization, which are similar to the use of rRNA. First, mitochondrial genes are transcribed by different polymerases than nuclear mRNA species and may be regulated differently [285]. Secondly, the mRNA fraction of a cell may be different in different tissue and cell types than the mitochondrial RNA fraction [285]. The disparity in mitochondrial RNA fraction (5%) may not be nearly as large as with rRNA (90-95%) and mRNA (3-5%), but differences in the relative abundance of mitochondrial and mRNA fractions may affect mRNA quantitation.

Given the potential strengths and weaknesses in the use of one type of normalization gene as compared to another, it is very experiment-specific as to which type of RNA species one should choose as a normalization gene. For example, if comparing within a specific treatment condition or within a single tissue, it is likely appropriate to apply the geometric averaging approach using multiple control genes to estimate the relative levels of different mRNA species in a given set of samples. If instead the goal is to compare relative abundances between different tissue samples from different sites in the body, a gene that is stably expressed across the entire tissue panel and is representative of the total cellular RNA content may be appropriate. In such cases, these studies demonstrate that the MT-ATP6 gene is a strong candidate for a normalization gene and is superior to rRNA species.

There may be tissue sites where MT-ATP6 is expressed lower or higher (e.g. skeletal muscle), which may preclude it as an appropriate choice for gene expression studies. For example, two studies have shown that the transcription of mitochondrial RNAs is reduced in human cancers [286, 287]. Although it has been demonstrated as an optimal control gene in a wide array of tissues throughout the body, it is still important to remember that any study attempting to quantitate expression levels between different tissues should first test a panel of putative control genes on a small subset of samples to ensure that the chosen gene is indeed stably expressed. The data presented in this study can be used to help investigators narrow down a potential list of candidate genes, greatly reducing the time and resources needed to complete the study. To our knowledge, only one other study has discussed the use of MT-ATP6 gene as a control gene for RQ purposes [288]. In that study, it was suggested that MT-ATP6 was superior to other routinely used housekeeping genes such as GAPDH, 18S rRNA, and cyclophilin for quantification of mRNA levels in clinical cancer samples across a diverse tissue panel.

A particular strength of the current study is the high quality of the RNA samples used in the analysis (>5.0 RIN). Because RNA quality affects the efficiency of the enzymatic reactions

differently for transcripts of differing abundance, it can differentially affect the quantification values obtained from samples of differing RNA quality. When isolating RNA from tissues, there can be a great degree of variability in RNA quality among different specimens, even when the samples are processed uniformly and in a highly controlled manner [289]. Several studies have shown that RNA quality can have a profound impact on RT-qPCR data [290]. This is usually true for longer amplicons (>400bp) or rare transcripts and when the RNA quality is seriously compromised (<5.0 RIN) [291]. For the majority of control gene and target gene transcripts, RNA integrity values from 1 (severely degraded) up to 10 (perfectly intact), has little impact on real-time PCR quantification values when an internal control gene is used [291]. This is likely due to the similar level of degradation that occurs for both the target and control genes and further illustrates the importance of using an appropriate internal control gene, as this decreases the likelihood that the data is being compromised due to poor quality RNA.

RQ via RT-qPCR is a powerful tool for assessing the abundance of individual RNA species in a biological sample. The technique is not without its problems, but if proper care is taken in the experimental design phase, it can lead to meaningful and reproducible results. These studies demonstrated that MT-ATP6 is a stably-expressed control gene that can be used for comparison of mRNA expression levels between different tissue sites.

Chapter 3

QUANTIFICATION OF HEPATIC UGT1A SPLICE VARIANT EXPRESSION AND CORRELATION OF UGT1A1 VARIANT EXPRESSION WITH GLUCURONIDATION ACTIVITY

3.1 Abstract

The UGT1A gene cluster encodes nine UGT1A family members via splicing of individual first exons to common exons 2 through 5. Each of these 9 UGT1As can also undergo alternative splicing at their 3' ends using an alternate exon 5, resulting in 27 different UGT1A mRNA species with each UGT1A gene encoding three different combinations of 5A and 5B UGT1A exons. To examine the importance of UGT1A exon 5 splice variants on overall UGT1A activity, a nested quantitative PCR assay was developed to accurately assess the combined expression of exon 5 splice variants (termed 'v2/v3') versus the expression of WT (termed 'v1') for each specific UGT1A. v1 expression was 16-, 17-, 57- and 29-fold higher than that observed for the levels of v2/v3 for UGTs 1A1, 1A4, 1A6, and 1A9, respectively, in normal human liver specimens. In a series of 58 normal human liver specimens, the expression of both UGT1A1 v1 and v2/v3 mRNAs were positively correlated with raloxifene glucuronidation activity in corresponding microsomes prepared from the same specimens ($p<0.0001$, $r^2=0.720$; $p=0.0002$, $r^2=0.241$, respectively), with expression of both variants lower in individuals homozygous for the UGT1A1*28 allele (42% for v1, $p=0.041$; 53% for v2/v3, $p=0.0075$). The expression of UGT1A1 v2/v3 was 1.6-fold higher than v1 ($p=0.03$) in HepG2 cells, and siRNA knockdown of HepG2 v2/v3 increased raloxifene glucuronidation activity by 83%. Together, these data suggest that hepatic UGT1A v2/v3 mRNA species are minor form variants that may not be physiologically relevant in human liver.

3.2 Introduction

UGTs play an important role in the metabolism and excretion of endogenous and xenobiotic compounds including drugs, carcinogens and chemotherapeutic agents. UGTs mediate the conjugation of glucuronic acid to their substrates, thereby increasing the polarity and facilitating excretion of the conjugate in the urine, bile, and feces [292]. Based on sequence homology, the UGTs are divided into two major sub-families. The nine UGT1A family members are encoded by a single locus on chromosome 2q37, where each family member has a unique exon 1 and shares exons 2-5 [293]. The unique first exon confers substrate specificity for the enzyme, while the common region contains the UDPGA cofactor binding domain and other essential motifs [104].

An alternative exon 5 in the common region of the UGT1A gene cluster leads to the expression of 18 additional mRNA species from this locus [105]. UGT1A mRNAs that incorporate only exon 5a are known as variant 1 (v1) and encode the active isoform 1 (i1) proteins [104]. When the alternative exon 5b is incorporated as the last exon (termed variant 2, or v2) or when both exons 5a and 5b are incorporated together (termed variant 3, or v3), the resulting protein is catalytically inactive [104]. The v2 and v3 mRNAs each have the same ORF and thus encode the same protein [termed isoform 2 (i2), [104]]. *In vitro* co-transfection studies demonstrated a dominant negative effect of i2 proteins on i1 glucuronidation activity [248-251]. In addition, concomitant expression of UGT1A1 i1 and i2 proteins in HEK293 cells leads to significantly decreased activity against a wide range of substrates [249]. These studies suggested that the function of the i2 proteins may be to negatively modulate the catalytic activity of the

UGT1A i1 proteins. However, neither expression of the UGT1A v2 and v3 splice variants nor UGT1A i2-mediated regulation of UGT1A i1 activity has been explored in human tissues.

We hypothesized that inter-individual variation in the relative abundance of v1 and v2/v3 mRNA expression affects glucuronidation capacity in human liver. The goal of the present study was to assess the importance of the exon 5 splice variants on overall UGT1A activity by quantifying mRNA expression of the v1 versus v2/v3 UGT1A splice variants in individual human liver specimens, and to determine whether there were correlations with glucuronidation activity in those same specimens.

3.3 Methods

Chemicals and Materials. UDPGA, D,L-2-lysophosphatidyl choline palmital C16:0, and *Escherichia coli* β-glucuronidase were purchased from Sigma (St. Louis, MO). *Pfu* DNA polymerase was purchased from Stratagene and the 1kb-plus DNA ladder, reverse transcription (RT) kits, cell culture reagents, and siRNA reagents were purchased from Invitrogen (Carlsbad, CA). High-performance liquid chromatography (HPLC)-grade solvents were purchased from Fisher Scientific. Suppliers of other reagents included Qiagen (Valencia, CA) for extraction kits, Stratagene (Santa Clara, CA) for Pfu polymerase, Applied Biosystems (Carlsbad, CA) for gene expression assays, Pierce Corp (Rockford, IL) for BCA reagents.

Tissue collection. Liver specimens were collected at the H. Lee Moffitt Cancer Center (Tampa, FL) from 58 individuals undergoing surgery for excision of HCC and were confirmed to be histologically normal as previously described. Tissue microsomes were prepared from the 60 normal liver specimens through differential centrifugation, total protein concentrations were measured with the BCA assay (Pierce Corporation, Rockford, IL), and samples were stored (10–

20 mg protein/ml) at -70°C in 100-µl aliquots [294]. All protocols involving the collection and analysis of tissue specimens were approved by the Institutional Review Board at Penn State University and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services. All samples were isolated and quick-frozen at -70°C within 2 h post-surgery.

RNA collection and cDNA synthesis. RNA was extracted from all liver tissue specimens by the Tissue Procurement Facility at the H. Lee Moffitt Cancer Center, and demographic data were provided for all subjects. Cell line RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Samples were subjected to DNAse I digestion during extraction to prevent genomic DNA contamination. RNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer, and RNA purity was assessed by absorbance ratios A260/A280 (> 1.9) and A260/A230 (>1.8). RNA integrity was determined using an Agilent 2100 Bioanalyzer with Agilent RNA 6000 Nano chips. Two samples with RIN values of 7.9 or less were excluded from further analysis because of degradation. Reverse transcription was performed using the Superscript First Strand cDNA synthesis kit (Invitrogen) with oligo dT primers and 1µg of starting RNA per sample. A negative control without RNA and a negative control without enzyme were analyzed in parallel.

First-step PCR reaction. For the initial amplification of the UGT1A mRNA species, the four exon 1-specific sense primers used were as follows: UGT1A1, 5'-catgctgggaagatactgttga-3' (+75 to +97); UGT1A4, 5'-cttctgctgagatggccaga-3' (-11 to +9); UGT1A6, 5'-tttggggcatggttgtaggt-3' (+59 to +78); and UGT1A9, 5'-ccactggttcaccatgaggt-3' (+108 to +127; locations provided are relative to the ATG translation start site for each UGT). The two exon 5-specific antisense primers were v1, 5'-cgcatgatgttctccttgtaactt-3' (-300 to -276), and v2/v3, 5'-

tgggaagtcagtcatcagtcct-3' (+74 to +96); both relative to the TGA translation stop codon for the v1 and v2/v3 variants, respectively). The v2 and v3 variant mRNA species encode the same non-functional protein, so no attempt was made to distinguish between them. The specificity of each primer set was verified using a 40 cycle-PCR amplification, followed by visualization and extraction of the amplification product, and subsequent sequencing in the forward and reverse directions to confirm its identity. To optimize the first-step PCR reaction, PCR amplification was carried out for 10, 15, or 20 cycles using *Pfu Turbo* DNA polymerase (Stratagene) and a 25 ng RNA equivalent of cDNA. PCR conditions were as follows: 95°C for 2 min, then 10, 15, or 20 cycles of 95°C for 20 seconds, 60°C for 30 seconds, and 72°C for 2 min, and finally 1 cycle of 72°C for 10 min.

Real-time quantitative PCR (qPCR). Following the first-step PCR amplification, cDNA samples were diluted 100-fold and subjected to a Taqman exon 1-specific real-time PCR reaction (Applied Biosystems). Expression levels were normalized to the expression of the MT-ATP6 gene, which exhibited stable expression levels relative to other putative control genes in a subset of normal liver samples (data not shown). Briefly, 32 putative control genes were analyzed in 6 liver specimens using Taqman Express Endogenous Control Plates (Applied Biosystems). MT-ATP6 was determined to be the most stably expressed control gene among the panel, with the lowest standard deviation in amplification cycle number between the six liver tissues examined. Quadruplicate real-time PCR reactions were performed for each cDNA sample using a 10 μL final reaction volume containing 5 μL 2X Taqman Universal PCR Master Mix, 4.5 μL of diluted cDNA, and 0.5 μL of gene expression assay. Reactions were performed in 384-well plates using the ABI 7900 HT Sequence Detection System under the following conditions: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min.

RQ of expression was calculated using the ΔΔCt method. Briefly, ΔCt was calculated as the Ct value of the target gene minus the Ct of the control gene (MT-ATP6). The ΔΔCt was then calculated as the ΔCt of the sample minus the ΔCt of a calibrator sample, in this case the lowest-expressing liver sample. RQ was determined with the formula $(1+E_x)^{-\Delta\Delta Ct}$ where Ex is the amplification efficiency of the given qPCR gene expression assay. Amplification efficiencies were calculated from UGT1A variant standard curves with the formula $E_x=[10^{(-1/slope)}-1]$. Standard curves for the eight UGT1A variants were generated as follows: pooled cDNA from eight livers was amplified for 20-cycles using primer sets specific to each of the eight full-length UGT1A cDNA species as described above. Resulting amplification products were subjected to a serial 10-fold dilution. Serially-diluted UGT1A amplification products were then amplified in quadruplicate for 40-cycles by qPCR, and the average Ct value for each dilution was plotted against the log of the DNA concentration.

Cell line generation. As positive controls for UGT1A variant expression, total RNA from cell lines individually over-expressing UGTs 1A1 v1, 1A4 v1, 1A6 v1 and 1A9 v1, as well as 1A1 v2, 1A4 v2, 1A6 v2, and 1A9 v2, were used to control for any differences in primer pair efficiency that may confound the quantification of the exon 1 and exon 5 splice variants. Generation of the HEK293 cell line over-expressing UGT1A v1 variants has been described previously [295-297]. UGTs 1A1 v2, 1A4 v2, 1A6 v2, and 1A9 v2-over-expressing cell lines were generated by stable transfection using the LipofectAMINE reagent (Invitrogen) according to the manufacturer's protocol. The primers used for cloning the UGT1A v2 variants were as follows: UGT1A1 sense, 5'- ccatggctgtggagtccc-3' (-2 to +16); UGT1A4 sense, 5'-tggcttctgctgagatggccag-3' (-14 to +9); UGT1A6 sense, 5'-ggagccctgtgatttggagagtg-3' (-63 to -40); UGT1A9 sense, 5'-cagttctctgatggcttgca-3' (-10 to +10; all locations listed are relative to the respective ATG translation start site). The antisense primer used for cloning the v2 variants is the

same as listed above for the nested PCR reaction. Briefly, pcDNA3.1/V5-His-TOPO/UGT1A v2 constructs were transfected into HEK293 cells grown to 80% confluence at 37° and 5% CO_2 in Dulbecco's modified Eagle's medium supplemented with 4.5 mM glucose, 10 mM HEPES, 10% fetal bovine serum, 100 U/ml penicillin, and 100 µg/ml streptomycin. At 24 h post-transfection, cells were passaged and grown in geneticin (700 µg/ml medium) for the selection of geneticin-resistant cells, with selection medium changed every 3 to 4 days. For validation of the nested PCR method, RNA (1 µg) isolated from the eight UGT1A v1 and v2/v3 cell lines was reverse transcribed and the resulting cDNA was incubated with primers that amplify a product encompassing the entire UGT1A1 ORF and portions of the 5' and 3' untranslated regions (UTRs). After this initial amplification, the resulting cDNA products were diluted 1:100 and subjected to a second PCR amplification using exon 1-specific real-time PCR assays as described above.

In order to generate a cell line co-expressing UGT1A1_i1 and UGT1A1_i2 protein, UGT1A1 v2 was cloned into the pcDNA 6.2/V5/GW/D-TOPO vector using the cloning primers described above, which amplify the entire coding region of the UGT1A1_i2 protein. This vector was transfected into UGT1A1_i1_over-expressing cells grown to 80% confluence at 37° and 5% CO_2 in Dulbecco's modified Eagle's medium supplemented with 4.5 mM glucose, 10 mM HEPES, 10% fetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin, and 700 µg/ml geneticin. At 24 h post-transfection, cells were passaged and grown in blasticidin for the selection of blasticidin-resistant cells, with selection medium changed every 3 to 4 days.

siRNA knockdown of UGT1A v2/v3 variants in HepG2 cells. Knockdown of the UGT1A v2/v3 species in HepG2 cells was accomplished using Lipofectamine2000 reagent and 10nM of the v2/v3-specific siRNA (5'-CAGCAGUCAGGAAGACAGAUGUGAA -3'). Control cells were transfected with scrambled (SCR) siRNA (5'-UUCGUAACCUGUCUUUCCGUCACUG -

3'). siRNA was designed using the Invitrogen Block-iT siRNA design module on their website. Because adherent HepG2 cells are notoriously difficult to transfect, cells were subjected to neofection, which enables a higher percentage of cells to be exposed to the transfection reagent, which can be a problem with adherent HepG2 cells that tend to layer on top of each other. Briefly, cells were first trypsinized and collected. siRNA and Lipofectamine2000 reagent were then complexed in Invitrogen Optimem reduced serum media according to the manufacturer's protocol. Two mL of the resulting siRNA : lipofectamine complexes were added to 150 mm culture dishes. Finally, $6x10^6$ HepG2 cells were added to each plate so that the confluence of the cells would be ~30% after 24 h. Cells were allowed to grow for 96 h prior to isolation of RNA and homogenate.

Glucuronidation assays. Glucuronidation activities of HLM and HepG2 cell homogenates were performed after an initial incubation of 500 μg HepG2 cell homogenate or 15 μg HLM with alamethicin (50 μg/mg protein) for 15 min in an ice bath. Glucuronidation reactions were then performed in a final reaction volume of 50 μL at 37°C in 50 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 4 mM UDPGA and tested substrate. 1-128 μM raloxifene was used for kinetic analysis in HLM from 6 individual subjects. For glucuronidation rate analysis, 2 μM raloxifene and was utilized for HLM, and 50 μM raloxifene was utilized for HepG2 cell homogenates. For UGT1A1_i1/UGT1A1_i2 co-expression experiments, 200 μg of protein homogenate from cell lines over-expressing either UGT1A1_i1 or UGT1A1 i1+i2 protein was used for kinetic rate determination against the substrate raloxifene (1-200 μM). Reactions were terminated by the addition of 50 μL cold acetonitrile on ice. Mixtures were centrifuged for 10 min at 4°C at 16,100 g and the supernatants were collected.

The analysis of bilirubin glucuronidation in the 58 HLM samples analyzed in the present study had been determined and described previously [134]. Briefly, glucuronidation of bilirubin

was assayed in a 100 μL reaction volume using 0.25 mg of HLM protein incubated for 1 h at 37°C in 50 mM sodium citrate (pH 7.4) containing 0.7 mM bilirubin, 4 mM UDPGA, 10 mM $MgCl_2$,and D,L-2-lysophosphatidyl choline palmital C16:0 (10 μg/100 μg of protein). Reactions were terminated by the addition of an equal volume (100 μL) of 2% ascorbic acid in ethanol.

Raloxifene glucuronidation was analyzed using a Waters ACQUITY ultra-performance liquid chromatography (UPLC) System (Milford, MA) with a 1.7 μ ACQUITY UPLC BEH C18 analytical column (2.1 mm × 50 mm, Waters, Ireland) in series with a 0.2 μm Waters assay frit filter (2.1 mm, Waters, USA). The gradient elution conditions, using a flow rate of 0.3 ml/min, were as follows: starting with 5% acetonitrile and 95% buffer A (5 mM ammonium acetate, pH 5.0) for 1 min, a subsequent linear gradient to 100% acetonitrile over 5 min was performed and then maintained at 100% acetonitrile for 2 min. Raloxifene-glucuronides were confirmed by their stability in 1M NaOH but sensitivity to the treatment of β-glucuronidase. In addition, confirmation of raloxifene glucuronide formation was performed by loading up to 5 μL of incubation product onto an UPLC identical to that described above in tandem with a Waters TQD triple quadrupole MS system. By using a positive mode, the parent compound $[M+H]^+$ peak and their corresponding glucuronide $[M\text{-}Gluc+H]^+$ peaks were characterized.

Bilirubin glucuronides were analyzed by HPLC using a Waters Associates binary pump (model 1525) HPLC system equipped with a dual λ absorbance detector operated at 450 nm, an automatic injector (model WISP 710B) and a β-RAM radioactive flow detector (IN/US, Tampa, FL). Samples were injected onto a Waters μBondapak (3.9 × 300 mm) 5 μC18 column. HPLC separations were performed using solvent A (0.01% trifluoroacetic acid in acetonitrile) and solvent B (0.01% trifluoroacetic acid in water) and the following linear gradient conditions: 0–20 min, 10–90% solvent A*; 2*0–25 min, 90–100% A. The HPLC flow rate was 1 ml/min, and the scintillation fluid flow rate was 4 ml/min. The column was routinely washed with 100% A for 15 min and then equilibrated after every HPLC run with 10% A for at least 20 min. Absorbance (450

nm) and ^{14}C-labeled peaks corresponding to glucuronide conjugates of bilirubin were tentatively identified by relative retention times, and then confirmed by sensitivity to *E. coli* β-glucuronidase treatment (1000 units, overnight incubation at 37°C), using HPLC as described above.

Statistical analyses. Pearson correlation of qPCR data with glucuronidation of raloxifene in HLM was done using GraphPad Prism version 5.00, with significance at $p\text{-value} < 0.05$. The Student's t-test (2-sided) was used to compare raloxifene glucuronidation rates in siRNA-treated versus untreated HepG2 cells. Percent identity and alignment of UGT1A exon 5b with alu consensus sequence was performed using the European Bioinformatics Institute's EMBOSS Stretcher program found at http://www.ebi.ac.uk/.

3.4 Results

Quantification of UGT1A v1 and v2/v3 splice variant mRNA levels. Previous studies examining the relative expression of UGT1A mRNAs prior to the identification of UGT1A exon 5b utilized real-time PCR assays containing exon 1-specific antisense primers [125, 153, 298, 299]. Therefore, quantification values reported in those studies encompassed the entire population of v1, v2, and v3 mRNAs for each UGT1A enzyme. To develop a quantitative assay that could distinguish between the v1 and the v2/v3 mRNA variants for each UGT1A enzyme, a nested real-time PCR technique was designed for quantification purposes because the distance between the unique first exons of each of the UGT1A enzymes and the alternate fifth exons is >1kb, a distance sub-optimal for traditional qPCR techniques. Since for each individual UGT1A the v2 and v3 variant mRNA species encode the same non-functional protein, no attempt was made to distinguish between v2 and v3 transcripts in the present study. This nested real-time PCR approach (see Figure 1) relied on a first-step PCR reaction that was carried out for 20 cycles and

utilized primers specific for eight UGT1A mRNA species (UGT1A1 v1, UGT1A1 v2/v3, UGT1A4 v1, UGT1A4 v2/v3, UGT1A6 v1, UGT1A6 v2/v3, UGT1A9 v1, and UGT1A9 v2/v3). UGTs 1A5, 1A8, and 1A10 were not included because they are expressed at very low or undetectable levels in human liver, and UGT1A3 was excluded because all primer pairs that were developed for this gene resulted in cross-amplification of UGT1A4 (results not shown). A nested real-time assay was developed for the extra-hepatic UGT1A7 [125, 298, 300] as a negative control for these experiments. The second-step PCR reaction was a real-time PCR assay that was performed using UGT1A exon 1-specific primers and probes (Figure 1). To demonstrate the specificity of the chosen primer pairs, pooled cDNA from eight human liver samples were PCR-amplified for 40 cycles. As shown in Figure 2, each primer pair resulted in a single PCR product of the appropriate size, and subsequent DNA sequencing confirmed that each PCR product corresponded to the intended UGT1A isoform (results not shown).

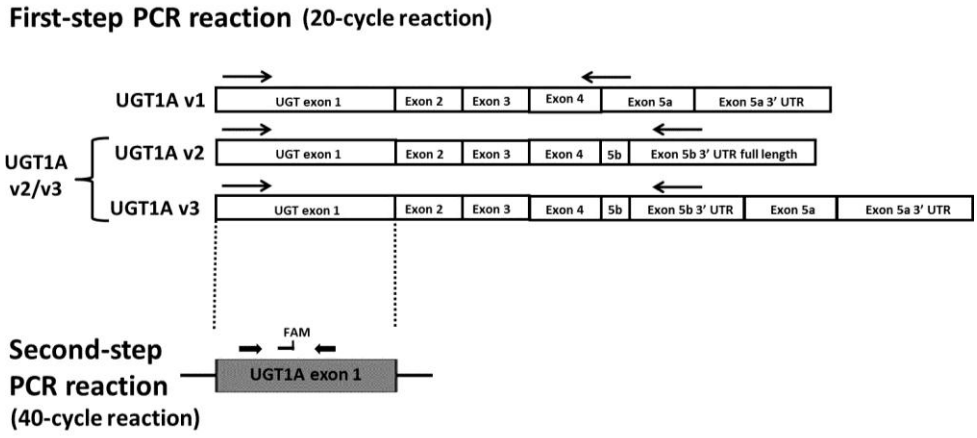

Figure 3.1 **Schematic of UGT1A expression analysis.** The first-step PCR amplification involved a 20-cycle reaction to distinguish between the different exon 5 splice variants. As UGT1A v1 variants encode the active i1 proteins while UGT1A v2 and v3 variants encode the same inactive i2 proteins, UGT1A v2 and v3 were grouped together and designated UGT1A v2/v3 mRNA as previously described [104]. The arrows above the gene represent the approximate location of the primers, with the same antisense primer used for amplification of both UGT1A v2 and v3 mRNA. UGT1A exon 1 is representative of the 9 different UGT1A exon 1 species that can be alternatively spliced to the common UGT1A exons 2-5. The second-step real-time PCR assay was then performed using UGT1A exon 1-specific real-time PCR primers and probes. FAM: approximate location of fluorescent 6-carboxyfluorescein probe; arrows: approximate primer locations.

The same primer pairs were used to amplify the pooled liver cDNA for 20 cycles and none of the UGT1A cDNA species could be visualized by gel electrophoresis (results not shown), indicating that the PCR assay had not yet reached the plateau phase, a requirement for the initial reaction of the nested PCR technique.

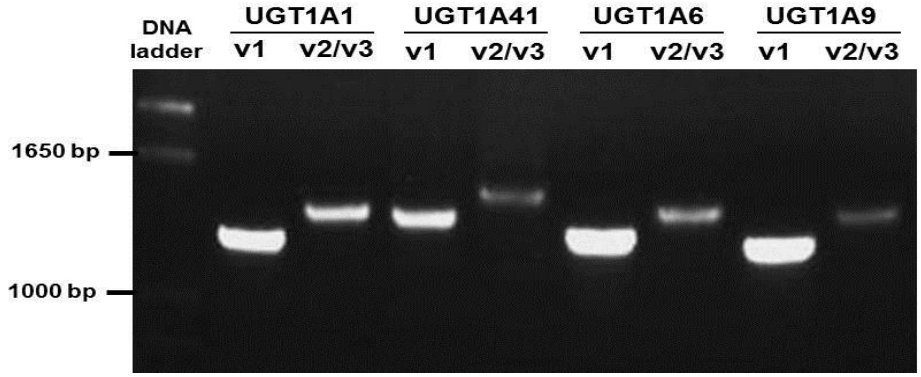

Figure 3.2. UGT1A sequence-confirmed amplification products from human liver cDNA. Pooled cDNA was reverse transcribed from total RNA from eight human liver specimens as described in the Materials and Methods. Sense primers were chosen to exploit sequence differences between the highly homologous UGT1A exon 1 regions. UGT1A isoforms were PCR-amplified for 40 cycles. Expected band sizes: UGT1A1_v1 = 1250 bp; UGT1A1_v2/v3 = 1349 bp; UGT1A4_v1 = 1338 bp; UGT1A4_v2/v3 = 1437 bp; UGT1A6_v1 = 1264 bp; UGT1A6_v2/v3 = 1363 bp; UGT1A9_v1 = 1209 bp; UGT1A9_v2/v3 = 1308 bp.

To assess the efficiency of the different primer sets used during the first-step PCR amplification, total RNA from HEK293 cell lines that over-expressed each of the eight UGT1A variants were used. Following reverse transcription, the nested PCR technique was used to amplify and quantify UGT1A mRNA from each of these cell lines. The first-step PCR amplification was carried out for 10, 15, or 20 cycles using the respective primer pairs for v1 or v2/v3, and each reaction was subsequently amplified in a second-step reaction using the real-time PCR assay described above and in the Materials and Methods. As shown in Figure 3 (panel A), the mean qPCR Ct values were linear for all UGT1A mRNAs for the range of 10- 20 cycles used in the first-step amplification reaction. The maximum efficiency difference between any of the eight primer pairs used in the first reaction was 5.1%.

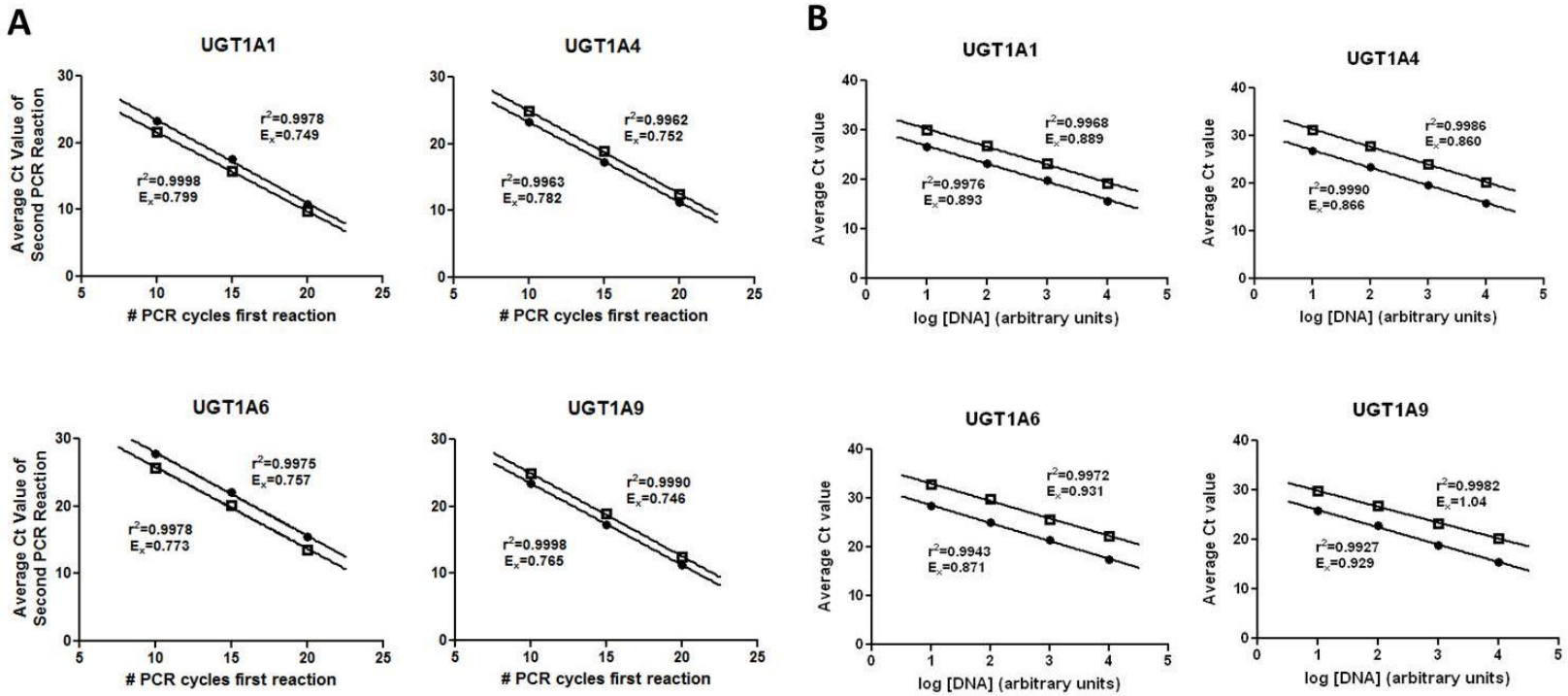

Figure 3.3. Validation of nested PCR methodology. Panel A, Total RNA from UGTs 1A1, 1A4, 1A6, and 1A9 v1- and v2/v3-over-expressing cell lines was reverse transcribed and the resulting cDNA was PCR-amplified for 10-, 15-, or 20-cycles prior to amplification by real-time PCR. Amplification efficiency calculated from the slope of the curves is shown for the amplification of both v1 and v2/v3 for each UGT. Panel B, Pooled cDNA from eight liver samples was PCR-amplified for 20-cycles using primers specific for each of the eight full-length UGT1A variants and the resulting amplification products were serially diluted and analyzed by real-time PCR. Average Ct values were calculated by taking the arithmetic mean of four replicate data points. Data are reported as mean ± standard deviation. ● – v1; □ – v2/v3

In addition to controlling for differences in primer efficiency for the first-step PCR reaction, it was also necessary to control for differences in amplification efficiency of the predesigned ABI qPCR assays used for the second-step PCR amplification. qPCR efficiency was calculated by generating UGT1A v1 and v2/v3 standard curves as described in the Materials and Methods. The calculated qPCR efficiency values were later used to adjust the UGT1A splice variant RQ values in individual liver specimens. The qPCR efficiency values for the different pre-designed ABI assays utilized for the second-step PCR amplification for the analysis of UGT1A enzyme expression ranged from 86%-104% using HEK293 cDNA (Figure 3, panel B).

Relative expression levels of UGT1A v1 and UGT1A v2/v3 mRNA in human liver. The validated nested PCR technique described above was used to determine the relative expression levels of UGT1A mRNA variants in liver samples from 10 individuals. The mean RQ values for the UGT1A v1 species were: UGT1A1, 1233 ± 154; UGT1A4, 645 ± 103; UGT1A6, 228 ± 27; and UGT1A9, 626 ± 79, while the mean RQ values for UGT1A v2/v3 mRNA species were: UGT1A1, 93 ± 19; UGT1A4, 43 ± 5.7; UGT1A6, 6.1 ± 1.4; and UGT1A9, 29 ± 5.3 (Figure 4, panel A). The relative abundance of v1 versus v2/v3, expressed as the log of the ratio of v1:v2/v3 for each of the four different UGT1As analyzed, is shown in Figure 4, panel B. Although some inter-individual differences in the relative abundance of v1 versus v2/v3 mRNAs were observed, v2/v3 mRNA levels did not approach the levels of v1 mRNA for any of the UGT1A species examined. On average, all of the UGT1A v2/v3 mRNA species were expressed at levels that were at least 16-fold lower than that of the corresponding UGT1A v1 mRNA species in the 10 liver specimens examined, with mean v1:v2/v3 mRNA ratios of 16 ± 1.7, 17 ± 2.6, 57 ± 10, and 29 ± 5.2 for UGTs 1A1, 1A4, 1A6, and 1A9, respectively. No expression of either the UGT1A7 v1 or v2/v3 variants were observed in any of the human liver specimens analyzed in this study.

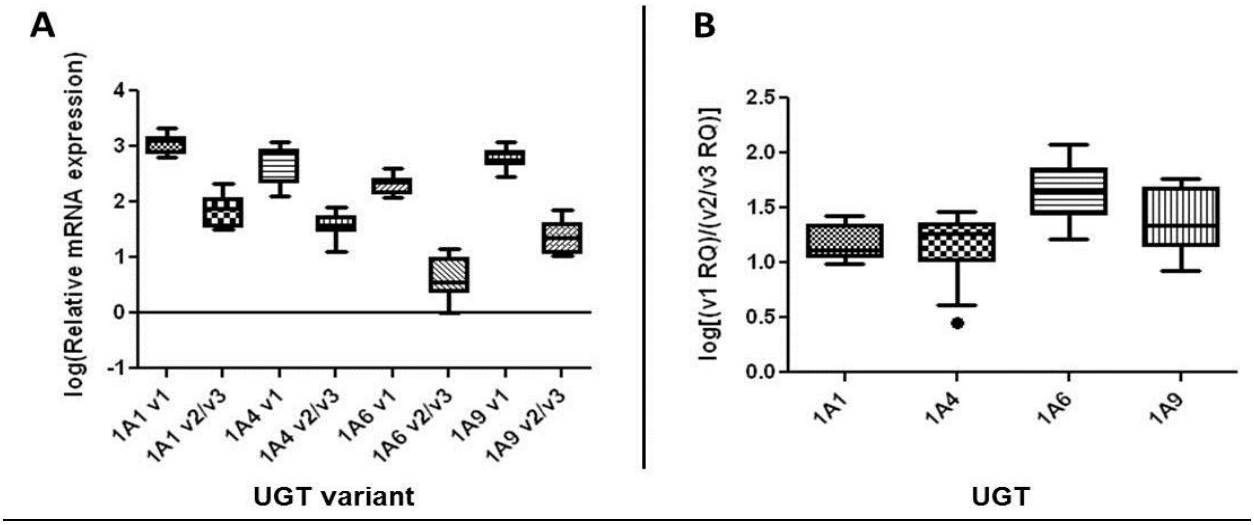

Figure 3.4. Relative mRNA expression levels of UGT1A variants in 10 liver RNA samples. UGT1A v1 and v2/v3 mRNA expression levels were assessed in liver cDNA from ten individuals using the two-step quantitative PCR methodology. **Panel A**, The log of the relative expression levels for the v1 and v2/v3 variants for UGTs 1A1, 1A4, 1A6, and 1A9 are shown. **Panel B**, The relative abundance is expressed as the log of the ratio of v1:v2/v3 for UGTs 1A1, 1A4, 1A6, and 1A9. Boxes, 25^{th}-75^{th} percentile; whiskers, min-max; dot, outlier (>1.5 inter-quartile range outside of lowest quartile).

Expression of UGT1A1 v1 and v2/v3 mRNA and correlation with glucuronidation activity in human liver. The relative expression of UGT1A v1 and v2/v3 was determined in 58 human liver RNA samples. There was a significant correlation between the expression of UGT1A1 v1 mRNA and UGT1A1 v2/v3 mRNA in these samples (Figure 5, $p<0.0001$, $r^2=0.4213$), suggesting that the rate of incorporation of the UGT1A1 alternate exon 5b is tightly regulated in human liver, with relatively low inter-individual variability in the ratio of UGT1A1 v1:v2/v3. The average relative expression ratio of UGT1A1 v1 versus UGT1A1 v2/v3 mRNA was determined to be 17 ± 1.5 in these 58 samples. In all of the liver specimens examined, expression of UGT1A1 v1 was found to be at least 3-fold higher than the expression of the v2/v3 variants, with 3 samples exhibiting a difference of less than 5-fold and 9 samples exhibiting a difference of less than 10-fold (range: 3-fold to 60-fold).

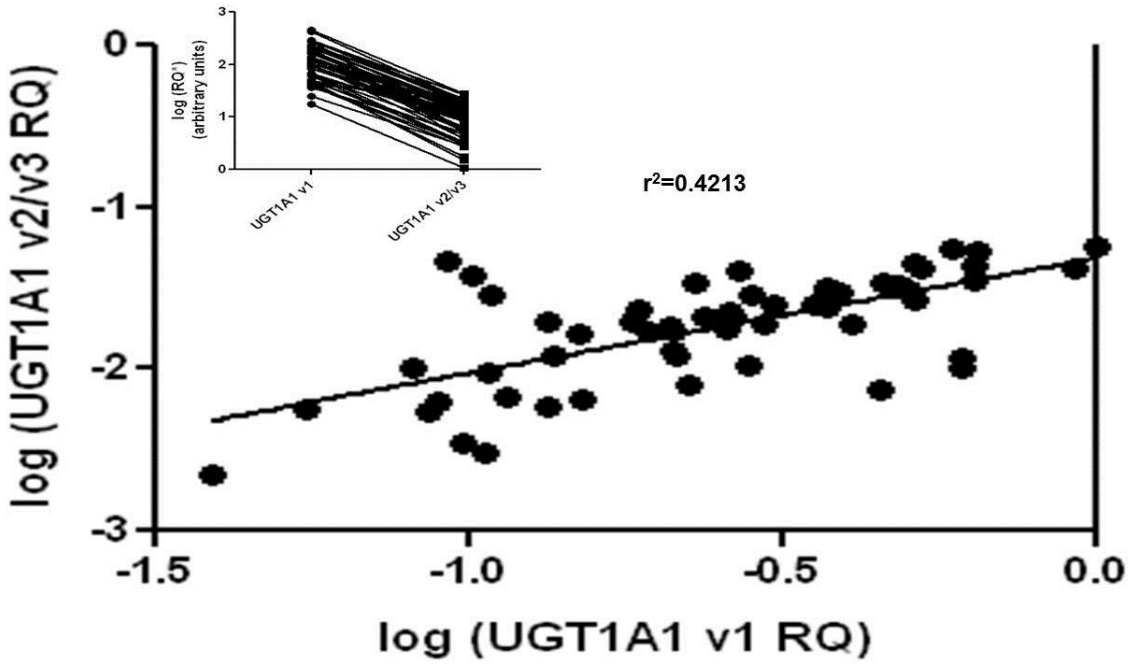

Figure 3.5. Relative expression levels of UGT1A1 v1 and UGT1A1 v2/v3 mRNAs in 58 human liver specimens. The nested qPCR reaction developed in this study was used to determine the relative expression of UGT1A1 v1 and UGT1A1 v2/v3 in total RNA isolated from the liver specimens described in the Materials and Methods. Shown is the correlation between the expression of the v1 and v2/v3 variants in the 58 individual liver specimens. Inset: Lines connect the UGT1A1 v1 RQ value of an individual specimen with the corresponding UGT1A1 v2/v3 RQ value from the same specimen. *RQ: relative quantification.

Raloxifene is glucuronidated extensively by family 1A UGTs, with its major glucuronide metabolite, raloxifene-6-*O*-glucuronide, primarily formed by UGT1A1 [301]. In the present study, UGT1A1 v1 expression was significantly (r^2=0.7195) correlated with raloxifene-6-*O*-glucuronide formation in HLM (Figure 6, panel A). A modest correlation was observed for raloxifene-6-*O*-glucuronide formation and UGT1A1 v2/v3 expression (r^2=0.241), while a minimal correlation was observed with the ratio of v1:v2/v3 expression (r^2=0.117). Similar results were observed when using another UGT1A1-specific substrate, bilirubin, in these experiments (Figure 6, panel B).

The UGT1A1*28 allelic variant contains an additional (TA) dinucleotide repeat in the 'TATAA' box of the UGT1A1 promoter [$(TA)_6$>$(TA)_7$] that has been linked to decreased

expression of the UGT1A1 gene and decreased overall UGT1A1 activity [93, 302, 303]. After stratifying the HLM examined in the present study by UGT1A1 TATAA box genotype, significant correlations were observed between raloxifene-6-*O*-glucuronide formation and UGT1A1 v1 expression for HLM from subjects who exhibited either the (*1/*1), (*1/*28), or (*28/*28) genotypes (Figure 6, panel C). While a modest correlation was observed between raloxifene-6-*O*-glucuronide formation and UGT1A1 v2/v3 expression for HLM from UGT1A1 (*1/*1) subjects (p=0.02, r^2=0.2512), this correlation was not observed for HLM from subjects who exhibited the UGT1A1 (*1/*28) or (*28/*28) genotypes (Figure 6, panel D). The same trends were observed when using bilirubin as substrate (results not shown). As observed in previous studies[23, 91-93, 134, 135, 302, 304], the expression of both UGT1A1 v1 and v2/v3 mRNAs were found to be significantly lower in livers from subjects with the UGT1A1 (*28/*28) genotype (42% lower for v1, p=0.0413; and 53% lower for v2/v3, p=0.0075) as compared to livers from UGT1A1 (*1/*1) subjects.

Effect of UGT1A v2/v3 knockdown on raloxifene glucurondation activities in cell lines. Co-expression of UGT1A1 v1 and v2/v3 had previously been shown to reduce glucuronidation activity *in vitro* [105, 248-251]. In the present study, UGT1A1_i1 and UGT1A1_i2 were co-expressed in HEK293 cells to study the effect of the i2 protein on raloxifene-6-*O*-glucuronide formation. Treatment of cells co-expressing the UGT1A1_i1 and UGT1A1_i2 proteins with SCR siRNA led to a 3.6-fold decrease in the V_{max} of raloxifene-6-*O*-glucuronide formation as compared to the V_{max} of untreated UGT1A1_i1-over-expressing cells (Figure 7, panel A). This pattern was partially reversed (60%) when a v2/v3-specific siRNA was used.

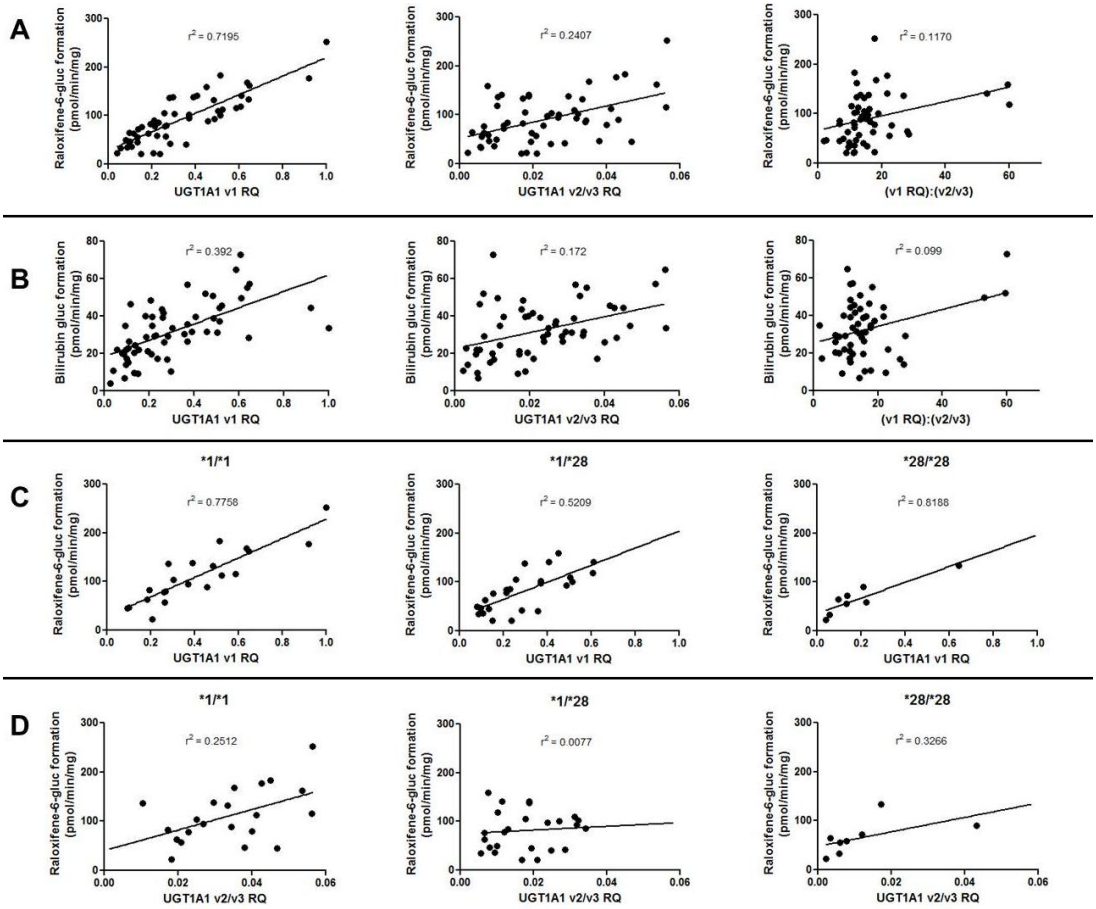

Figure 3.6. Correlation of UGT1A1 mRNA expression with HLM glucuronidation activity. UGT1A1 v1 and UGT1A1 v2/v3 mRNA levels were assayed in 58 human liver specimens and correlated with glucuronidation activity against raloxifene (panel A) and bilirubin (panel B) in HLM derived from the same specimens. Correlations of UGT1A1 v1 (panel C) and UGT1A1 v2/v3 (panel D)

To examine the effects of high levels of endogenous expression of v2/v3 on UGT1A activity, the relative expression of UGT1A v1 versus v2/v3 were examined in the HepG2 cell line. Unlike that observed in human liver tissue, this cancer-derived cell line expressed UGT1A1 v2/v3 mRNA at a level that was 1.6-fold higher than UGT1A1 v1 mRNA (Figure 7, panel B; p=0.030), suggesting that this cell line could serve as a model for high relative endogenous expression of the UGT1A v2/v3 splice variants. siRNA knockdown of UGT1A v2/v3 splice variants resulted in a >95% reduction in UGT1A1 v2/v3 mRNA levels in HepG2 cells (Figure 7,

panel C), resulting in a v1:v2/v3 ratio of 13.6, which was similar to that observed for human liver tissue. This decrease in expression corresponded with a significant (p=0.003) 82.5% increase in raloxifene-6-*O*-glucuronide formation in anti-v2/v3 siRNA-treated HepG2 cells versus SCR-treated controls (Figure 7, panel D).

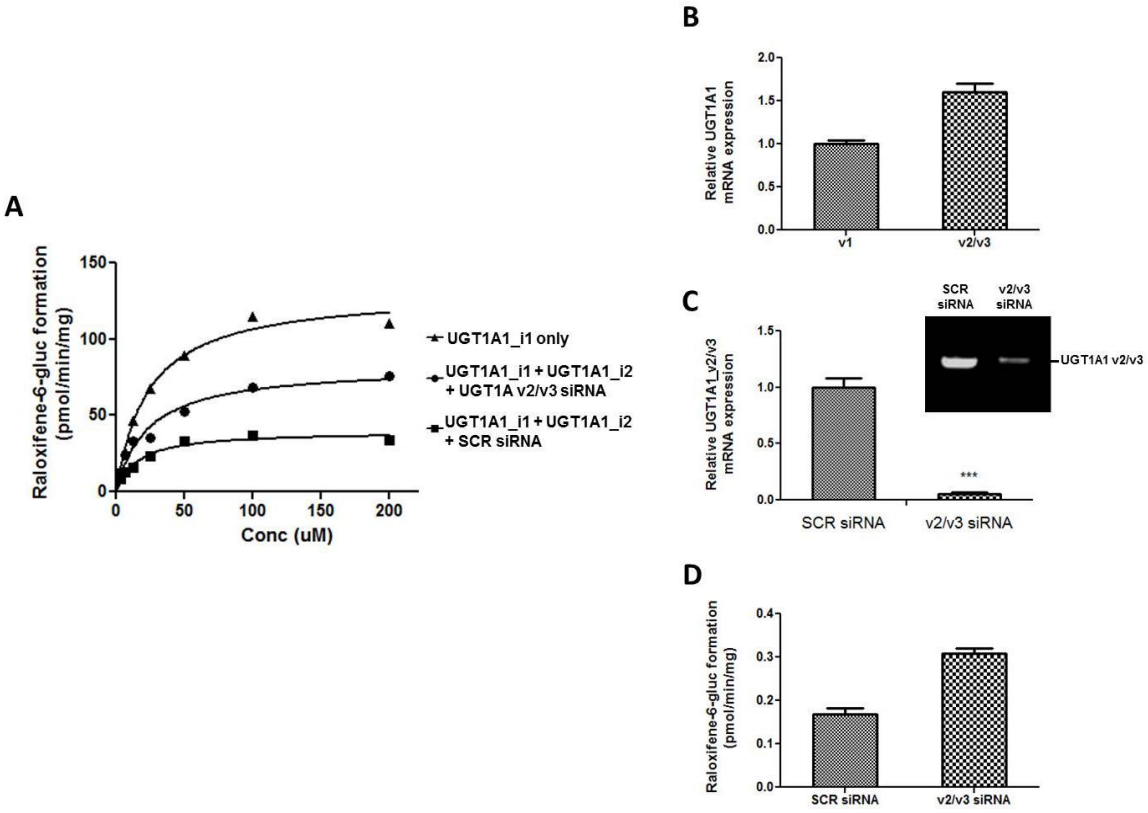

Figure 3.7. Effect of siRNA treatment of UGT1A_v2 mRNA on raloxifene glucuronidation activity in cell lines. Panel A, Representative kinetic curves for raloxifene-6-*O*-glucuronide formation are shown for HEK293 cells over-expressing UGT1A1 v1 alone or in combination with the UGT1A1 v2 variant. Cells co-expressing UGT1A1 v1 and UGT1A1 v2 were treated with either 10 nM SCR or anti-UGT1A1 v2/v3-specific siRNA for 96 h prior to conducting glucuronidation assays. Panels B-D, HepG2 cells. Panel B, The two-step quantitative method described in the Materials and Methods was used to assess the relative endogenous expression of UGT1A1 v1 and UGT1A1 v2/v3 variants in the HepG2 cell line. Panel C, The relative quantification of UGT1A1 v2/v3 mRNA after knockdown with 10nM SCR or anti-v2/v3-specific siRNA. The inset is a representative electrophoresis gel of the RT-PCR of UGT1A1 in HepG2 cells treated with SCR or anti-v2/v3-specific siRNA after 37 cycles of amplification. Panel D, Effect of UGT1A1 v2/v3 knockdown on raloxifene-6-*O*-glucuronide formation in HepG2 cells treated with SCR or anti-v2/v3-specific siRNA.

3.5 Discussion

The goal of the present study was to investigate the importance of UGT1A exon 5 splice variants in overall hepatic UGT1A activity *in vivo* by quantifying expression of the v1 and v2/v3 UGT1A mRNA species in individual human liver specimens and determining whether there were correlations between UGT1A1 variant expression with glucuronidation activity in those same specimens. To accurately assess UGT1A variant expression, it was necessary to design an assay that would permit accurate quantification of the distinct UGT1A mRNAs (v1 and v2/v3) from which the two UGT1A proteins (i1 and i2) can be translated. All previous quantitative expression studies of specific UGT1A enzymes utilized exon 1-specific primers that could not distinguish between the three exon 5 splice variants. Therefore, the quantification of transcripts that encoded active UGT1A protein would have been confounded by the presence of the v2 and v3 mRNA variants. The assay utilized in the present study employed a quantitative nested PCR that relies on a first-step 20 cycle PCR reaction followed by a 40 cycle isoform-specific real-time PCR assay. Because UGT1A v2 and v3 mRNAs encode the same inactive protein [104], a reverse primer was designed to target a shared region of the exon 5b 3'UTR so that both cDNAs could be amplified simultaneously. The reverse primer specific for UGT1A v1 mRNA was designed such that the primer overlapped the exon 4/5a junction, which is the only region completely unique to variant 1. The UGT1A exon 1 sense primers were designed so that the amplification product from the first PCR reaction was specific to the UGT1A enzyme of interest and would contain the binding site for the primers and probes of the subsequent real-time PCR reaction.

Differences in amplification efficiency due to primer characteristics were lessened by employing a primer design strategy that minimized differences in GC content (42%-55%), melting temperature (54.7°C-57.4°C), primer length (20bp-25bp), and hairpin formation ($\Delta G > -2$

kcal/mol). The target sequence and GC content of the different UGT1A amplicons are highly homologous, making it easier to achieve similar amplification efficiency for these cDNA species than for heterogenous DNA sequences. The data from the validation experiments described in this study indicates that this nested real-time PCR methodology preserves the quantitative nature of the real-time PCR assay, making this the first quantitative study of the UGT1A gene locus to simultaneously distinguish between the 5' splice variants, which determine the individual UGT1A species, and the 3' splice variants, which result from incorporation of different fifth exons. Analogous methods may be useful for investigation of other gene loci that can be transcribed with alternate splicing which results in splice variants at both ends of the transcript. The nested PCR methodology results in a small loss of accuracy in favor of increased sensitivity and specificity. However, control experiments demonstrated that the maximal efficiency difference between the primers used for the 20-cycle amplification was 5.3%. After 20 cycles, this would result in a maximum quantitation error of 61%, much lower than the >10-fold difference in relative abundance between the v1 and v2/v3 variants.

The mean hepatic expression of UGT1A1 v1 was found to be 1.9-fold higher than UGT1A4 v1, 5.4-fold higher than UGT1A6 v1, and 2.0-fold higher than UGT1A9 v1, with the order of expression: UGT1A1 > UGT1A4 > UGT1A9 > UGT1A6. The high relative abundance of UGT1A1 relative to other UGT1A species matches closely with previous reports [153, 298, 299]. Our findings show the order of expression of UGT1A9 and UGT1A4 to be reversed from that reported by Ohno et al.; however, this is a minor difference since expression of UGT1A4 was only 3% higher than UGT1A9 in the present study and UGT1A9 was less than two-fold higher than UGT1A4 in the study by Ohno et al. [153]. As proper control selection is imperative for accurate normalization via the $2^{-\Delta\Delta Ct}$ method [261, 264, 267, 284], all UGT1A genes were normalized to MT-ATP6 content, which was determined to be the most stably expressed gene

amongst a panel of 32 putative control genes in a subset of liver samples (Jones et al, submitted for publication).

The mean hepatic expression of the UGT1A v2/v3 splice variants was highest for UGT1A1. These variants were found to be expressed at an average of 2.2-fold higher than UGT1A4 v2/v3, 3.2-fold higher than UGT1A9 v2/v3, and 15-fold higher than UGT1A6 v2/v3. In all cases, UGT1A v1 mRNAs were found to be expressed, on average, at least 16-fold higher than their corresponding v2/v3 mRNAs. Only 2 of the 58 specimens had less than 5-fold difference in the levels of the v1:v2/v3 relative abundance ratio for UGT1A1. This indicates that there is a subset of individuals where the incorporation of exon 5b occurs at a high enough rate to potentially affect glucuronidation capacity. Alternatively, this may also be explained by the fact that the tissue specimens examined were adjacent normal tissue from patients undergoing resection for HCC - while histologically normal, they may have acquired some of the changes associated with the malignant phenotype. This would be consistent with studies by Bellemare et al. demonstrating that both UGT1A i1 and i2 proteins were evident in human liver, but that i2 protein was more extensively expressed in corresponding tumor tissue [305].

The expression of UGT1A1 mRNA has been shown to correlate with UGT1A1_i1 protein expression in human liver [298], and it is known that UGT1A1 mRNA levels are correlated with differences in activity against substrates such as bilirubin and irinotecan [25, 91, 92, 304, 306]. However, previous studies have not differentiated between v1 and v2/v3 mRNA species. To better assess the potential correlation between hepatic glucuronidation activity and the relative abundance of hepatic UGT1A v1 and/or v2/3 mRNA, expression of UGT1A1 variants were examined and correlated with hepatic raloxifene-6-*O*-glucuronide and bilirubin-glucuronide formation activity. Significant positive correlations were observed between UGT1A1 v1 mRNA levels and raloxifene-6-*O*-glucuronide formation or bilirubin glucuronidation in the present study. A modest positive correlation was also observed for UGT1A1 v2/v3. However, if increased v2/v3

expression was affecting hepatic glucuronidation activities, a negative correlation would be expected. These data suggest that since UGT1A1 v1 and v2/v3 expression are likely controlled by the same UGT1A1 promoter, v2/v3 expression levels may, to some extent, be acting as a surrogate marker for v1 expression and UGT1A activity for any given liver specimen. Alternatively, it is possible that UGT1A i2 proteins have a functional role if they exhibit increased stability as compared to their UGT1A i1 counterparts, a pattern consistent with the i2 expression observed by Bellemare et al. in human liver specimens [305]. Future studies will be needed to quantitatively assess protein expression of the i2 proteins and UGT1A regulation at the translational and post-translational levels.

The i2 protein has been previously shown to be expressed in human carcinoma tissue and cell lines, and siRNA knockdown of the v2/v3 variant resulted in increased activity against a range of substrates in colon cancer cell lines [253, 305]. In the current study, it was observed that the human HCC cell line HepG2 exhibited relatively high endogenous expression of exon 5b-containing variants for UGT1A1, with v2/v3 variant expression reaching levels that were slightly greater than that observed for UGT1A1 v1. siRNA-induced silencing of v2/v3 variants in HepG2 cells resulted in a 95% decrease in v2/v3 expression and a corresponding >80% increase in HepG2 activity against raloxifene. In addition, in HEK293 cells co-expressing UGT1A1_i1 and UGT1A1_i2, the addition of anti-v2/v3-specific siRNA also increased raloxifene glucuronidation. Together, these positive *in vitro* data support the apparent lack of an effect by endogenous UGT1A v2/v3 variants on hepatic glucuronidation activity due to low levels of hepatic UGT1A1 v2/v3 expression.

The common TATA box promoter polymorphism associated with the UGT1A1*28 allele is known to reduce transcription of the UGT1A1 gene and decrease the metabolism of bilirubin and other substrates of UGT1A1 such as irinotecan [24, 25, 307]. When the HLMs examined in the present study were stratified by UGT1A1 genotype, a significant correlation between hepatic

UGT1A1 v1 mRNA levels and HLM raloxifene-6-*O*-glucuronide formation was observed for all three UGT1A1 TATAA box genotypes identified. These results were confirmed when using a second UGT1A1-specific substrate, bilirubin. The 42% decrease in hepatic UGT1A1 v1 mRNA expression observed in individuals homozygous for the UGT1A1*28 allele compared to individuals with one or zero copies (*p=0.0413*; data not shown) is comparable to previous reports that have investigated the effect of this genotype on UGT1A1 transcription [93, 135]. Levels of UGT1A1 v2/v3 were found to be 53% lower in individuals homozygous for the UGT1A1*28 allele as compared to individuals with zero copies of this allele (*p=0.0075*; data not shown). These data are consistent with the positive correlation observed between UGT1A1 v1 and v2/v3 expression levels in these liver samples and suggest that the TATAA box polymorphism associated with the UGT1A1*28 allele coordinately regulates the transcription of both the WT v1 and variant v2/v3 mRNAs in human liver.

The low relative expression of v2/v3 variants in human liver is not surprising considering that the v2/v3-specific UGT1A exon 5b is likely a much more recent evolutionary event than originally characterized [104]. Sequence alignment of the human UGT1A1 gene with the rat Ugt1a1 gene reveals that while the exons 1-4 and 5a are highly homologous, there is little homology in the intervening intron containing the alternative exon 5b (http://useast.ensembl.org/Homo_sapiens/Location/Compara_Alignments?align=393&db=core&g=ENSG00000241635&r=2%3A234668894-234681945). The rat does have an alternative exon present in this region; however, it is not truly orthologous with the human exon 5b. The majority of alternative splicing events are not conserved between orthologous genes [242, 308, 309], and the human exon 5b shares 62% homology with a primate-specific repetitive alu transposable element consensus sequence [310] (Figure 8). In fact, the final 180 bp of the exon 5b 3'UTR shares 78% sequence homology with this type of repetitive element. The first half of exon 5b is divergent from the alu consensus sequence, which can occur with older alu elements that are

exonized in the human genome [311]. It is estimated that 4% of human genes contain transposable elements that have been exonized [232, 311, 312], and the most commonly exonized transposable elements are alu elements, which comprise roughly 11% of the genome and account for around 5% of all alternative splicing events [313]. Of the estimated 3,000 'aluternative' exons in the human genome [313], there are few constitutive alu-derived exons, and, in general, alu exons have low retention rates in mRNA [243, 311]. Recently gained exons such as these tend to be included at relatively low levels in spliced mRNA, [311, 314] which is consistent with the pattern observed in the present study for UGT1A exon 5b-incorporated variants.

Figure 3.8 Sequence alignment of the human UGT1A exon 5b with the primate-specific alu transposable element consensus sequence.

The positive correlation observed between UGT1A1 v2/v3 and UGT1A1 v1 expression suggests that there are relatively small inter-individual differences in the rate of exon 5b incorporation into the final UGT1A transcript profile in human liver. Wang et al. reported that alternative splicing events are often tissue-specific, detecting 22,000 such events, whereas inter-individual variation is less common [315]. Given that the liver is the body's primary metabolic organ with a near constant need for high expression and activity of UGT enzymes, there may be little need for UGT1A i2-associated negative regulation of UGT1A i1 protein in liver, but this does not rule out an extra-hepatic, tissue-specific regulatory role for i2 proteins. In a panel of 10 tissues and five cell lines, Wang et al. detected 8,000 alternatively-spliced terminal exons, nearly half of which were found to be expressed in a tissue-specific manner [315]. In addition, human alu-derived exons are often expressed in a tissue- or tumor-specific manner [239, 316, 317]. For example, only the i2 proteins are expressed in the kidney where they may have their own tissue-specific function [305]. Girard et al. has also demonstrated that the kidney and esophagus express only the v2/v3 variants of certain UGT1A species, suggesting that certain tissues may have a 'switch-like' regulation mechanism that favors the expression of some of the non-functional variants [104].

The correlation of UGT1A v1 mRNA levels with glucuronidation activity clearly demonstrates the validity of the methodology employed in this study. Previous studies have disagreed about the extent to which UGT1A gene expression correlates with protein levels and activity [153, 318]. For example, a recent study by Ohtsuki et al. demonstrated that only UGT1A6 mRNA levels correlate with protein levels; however, a previous study by Ohno et al. demonstrated that only UGT1A1 mRNA levels were correlated [153, 318]. The former study relied on a UGT1A1 gene expression assay entirely within exon 1 of the UGT1A1 gene. This is problematic, as it can allow genomic DNA to impact gene expression levels, clearly evidenced by the >100-fold difference in UGT1A1 mRNA levels in the 16 liver specimens utilized. In the

current study, the improved sensitivity and specificity afforded by the nested PCR reaction allowed a more reasonable estimate of the interindividual variability in UGT1A1 expression (<25-fold between 58 specimens). The low interindividual variability in mRNA levels in this study is consistent with the improved correlation between mRNA levels and activity, suggesting that this methodology reduces noise in the quantitative assay.

In summary, this study indicates that UGT1A v1 mRNAs are expressed, on average, at least 16-fold higher than their corresponding UGT1A v2/v3 mRNAs in human liver. There is a strong positive correlation between hepatic UGT1A v1 and v2/v3 expression for all hepatic UGTs tested, suggesting that their expression is controlled by the same transcriptional machinery. The positive correlations observed between the hepatic expression of UGT1A1 v1 as well as UGT1A1 v2/v3 mRNA with hepatic glucuronidation activity suggests that UGT1A v2/v3 expression serves only as a marker for v1 expression and that the expression of UGT1A i2 protein may be too low relative to the UGT1A i1 protein to exert a negative regulatory effect on UGT1A1 glucuronidation activity in human liver. To better understand how this complex gene locus is expressed and regulated throughout the body, future quantitative studies are required to determine organ locations where the relative expression of UGT1A v2/v3 mRNA is high and leads to functional regulation of UGT activity.

Chapter 4

QUANTIFICATION OF UGT2B GENE EXPRESSION IN HUMAN TISSUES

4.1 Abstract

A better understanding of the inter-individual variability and relative abundance of UGT gene expression in hepatic and extrahepatic tissues is important as this helps determine the physiological relevance of each UGT enzyme. In the present study, UGT2B mRNA levels were quantified in normal tissue specimens from lung, liver, larynx, brain, tongue, floor of mouth, tonsil, esophagus, endometrium, and pancreas. UGT2B genes were highly expressed in liver, and UGT2B4 and UGT2B7 were found to be predominantly hepatic enzymes. There was wide variability between individuals in the magnitude of expression in each tissue site. In matched lung tumor and adjacent normal tissue specimens, tumoral expression of UGT2B genes was significantly higher for several isoforms. UGT2B7 was 8.7-fold higher in tumor samples (p=0.0134), , UGT2B11 was 4.3-fold higher (p=0.0019), UGT2B15 was 3.8-fold higher (p=0.017), and UGT2B17 was 4.4–fold higher (p=0.0056). Within the liver, the expression of UGTs 2B4, 2B7, 2B10, and 2B15 were significantly correlated with each other (r^2=0.81–0.97; all p<0.0001), suggesting a common transcriptional regulatory mechanism for these genes. The expression of all UGT2B genes was significantly correlated with each other in tonsil specimens (r^2=0.60–0.94; all p<0.05), and the expression of UGT2B11, UGT2B15, and UGT2B17 were significantly correlated with each other in lung tissue (r2=0.73–0.87; all p<0.01).

4.2 Introduction

UGTs are involved in the detoxification of a wide range of endogenous and exogenous compounds, including many drugs and carcinogenic compounds. Because of this, quantitative determination of the interindividual differences in UGT expression will be helpful in understanding interindividual differences in the metabolism of drugs as well as individual

susceptibility to carcinogenic compounds [48]. The liver is the organ most commonly associated with metabolism, and all UGT2Bs are expressed in the human liver [49, 153, 181, 298]. Expression in extrahepatic tissues including those of the aerodigestive and gastrointestinal tracts has been less well characterized, even though these tissues represent an important first line of defense against exposures to xenobiotic compounds.

Most studies have used qualitative RT-PCR for determining which UGT genes are expressed in different tissues [53, 164, 176, 299, 300, 319-323]. Studies by Ohno and Nakajima have attempted to characterize the relative expression levels in numerous tissues throughout the body [151]. However, in the latter study, pooled RNA was utilized, making it impossible to characterize interindividual differences in expression patterns. The only organ where quantitative levels of expression have been measured with respect to interindividual differences in expression is the liver [151]. In the present study, the relative expression levels of the various UGT2B isoforms were quantified in 132 normal human tissue specimens (23 lung, 26 liver, 8 larynx, 5 brain, 3 tongue, 3 floor of mouth, 23 tonsil, 12 esophagus, 8 endometrium, and 16 pancreas). Expression levels were also determined in 23 matched lung tumors and 5 matched HCCs.

4.3 Methods

Tissue collection. Tissue specimens were obtained from the Penn State College of Medicine Tissue Bank. All samples were isolated and quick-frozen at -70°C within 2 h post-surgery. All protocols involving the collection and analysis of tissue specimens were approved by the Institutional Review Board and Scientific Review Committee at Penn State University and were in accordance with assurances filed with and approved by the United States Department of Health and Human Services.

RNA collection and cDNA synthesis. RNA was extracted from all tissue specimens using the Qiagen RNeasy Mini Kit according to the manufacturer's protocol (Valencia, CA). Samples were subjected to DNAse I digestion during extraction to prevent genomic DNA contamination. RNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer, and RNA purity was assessed by absorbance ratios A260/A280 (> 1.9) and A260/A230 (>1.8). RNA integrity was determined using an Agilent 2100 Bioanalyzer with Agilent RNA 6000 Nano chips. All samples used in this study had RIN scores of 5.0 or higher. Reverse transcription was performed using the Invitrogen Superscript First Strand cDNA synthesis kit with oligo dT primers and 1 µg of starting RNA per sample (Carlsbad, CA). A negative control without RNA and a negative control without enzyme were analyzed in parallel.

Real-time quantitative PCR (qPCR). After the RT reaction, qPCR amplification was carried out using a 25 ng RNA equivalent of cDNA. Expression levels were normalized to the expression of the MT-ATP6 gene, which exhibited stable expression levels relative to other putative control genes in a subset of normal liver samples. Quadruplicate real-time PCR reactions were performed for each cDNA sample using a 10 µL final reaction volume containing 5 µL 2X Taqman Universal PCR Master Mix, 4.5 µL of diluted cDNA, and 0.5 µL of gene expression assay. Reactions were performed in 384-well plates using the ABI 7900 HT Sequence Detection System under the following conditions: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min.

RQ of expression was calculated using the ΔΔCt method. Briefly, ΔCt was calculated as the Ct value of the target gene minus the Ct of the control gene (MT-ATP6). The ΔΔCt was then calculated as the ΔCt of the sample minus the ΔCt of a calibrator sample, in this case the highest-expressing liver sample. RQ was determined with the formula $(1+E_x)^{-\Delta\Delta Ct}$ where Ex is the amplification efficiency of the given qPCR gene expression assay. Amplification efficiencies

were calculated from MT-ATP6 and UGT2B standard curves with the formula $E_x=[2^{(-1/slope)}-1]$. Standard curves for UGT2B and MT-ATP6 cDNA species were constructed using serial dilutions of human liver cDNA. The slope of the resulting 'Ct number vs. log_2cDNA' curve was used to calculate the relative efficiencies of the different real-time PCR assays.

Data analysis. For qPCR data, statistical analyses were done using GraphPad Prism version 5.00. RQ values were log-transformed to obtain a normally distributed data set, and all data are reported based on the geometric mean of RQ values in each tissue. When calculating the mean RQ values, only samples that displayed detectable levels of expression were included in the analysis. Interindividual differences in UGT2B gene expression levels were compared for isoforms expressed in at least ½ of the given tissue specimens from a given tissue site. Similarly, comparison of extrahepatic UGT2B expression to liver expression was determined only for the UGT2B isoforms that were expressed in at least ½ of the tissue specimens within a given tissue site. Expression of UGT2B genes in matched tumor and adjacent normal lung tissue was compared using a paired t-test with significant p-value < 0.05. Spearman's rank method was used for correlative analysis of expressed UGT2B isoforms in human liver, lung, and tonsil tissues.

4.4 Results

Generation of qPCR standard curves. UGT2B and MT-ATP6 standard curves were used to calculate real-time PCR efficiency values (Figure 4.1). The slopes of the various 'Ct vs. log2[cDNA]' curves were different for each gene. The calculated efficiency values ranged from 85.7% to 98.6%. The relative efficiencies were used to correct the RQ values using the formula described in the Materials and Methods.

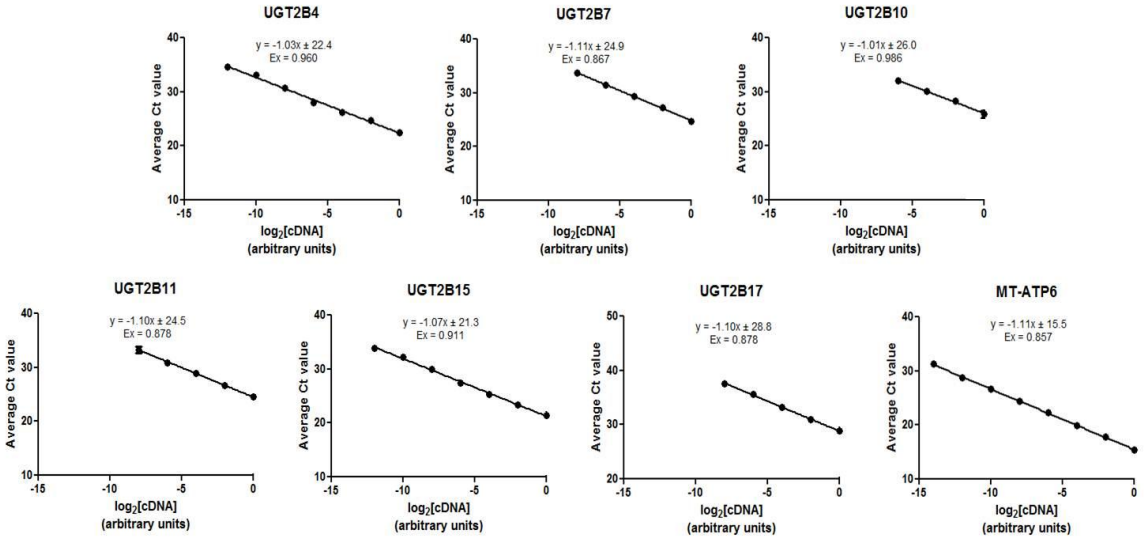

Figure 4.1. UGT2B and MT-ATP6 qPCR standard curves. Pooled cDNA from three liver samples was subjected to a serial 4-fold dilution prior to real-time PCR amplification. The efficiency values of each of the real-time PCR assays was calculated with the formula $E_x=[2^{(-1/slope)}-1]$, as described in the Methods. Average Ct values were calculated by taking the arithmetic mean of four replicate data points. Data are reported as mean ± standard deviation.

Relative expression of UGT2B genes in human tissues. The mean relative expression of UGT2B genes in different tissues is shown in Table 4.1. All UGT2B isoforms were detected in human liver tissue, as previously reported by other groups [153, 231, 298]. The relative expression levels were found to be the highest for UGT2B15, followed by UGT2B4>UGT2B17> UGT2B7>UGT2B10>UGT2B11 (Table 4.1). The mean expression level of UGT2B4 was 1.2–fold higher than UGT2B4, 3.2-fold higher than UGT2B17, 4.0-fold higher than UGT2B7, 4.1–fold higher than UGT2B10, and 26–fold higher than UGT2B11. Only UGTs 2B15 and 2B17 were not expressed in all of the liver samples. Five liver samples showed no detectable UGT2B17 expression and two exhibited no detectable UGT2B15 expression. The result for UGT2B17 was expected since the UGT2B17 whole gene deletion polymorphism is homozygous in 10% of Caucasian populations. While the prevalence of the homozygous UGT2B17 null genotype was higher than expected (5/26, or 19.2%), the sample size was small. The two individual liver samples that did not show quantifiable UGT2B15 expression did amplify above

background, but the Ct value was too large to be in the accurately quantifiable range (above 35 cycles). Interindividual differences ranged in expression by at least 100-fold for all of the UGT2B enzymes analyzed, with UGT2B7 showing the greatest interindividual differences (3673-fold), followed by UGT2B10, UGT2B17, UGT2B4, and finally UGT2B11 (Figure 4.2). The results shown here agree well with previous studies that examined the relative expression levels in individual human liver specimens [298].

Table 4.1. Mean* relative UGT2B expression levels in human tissue specimens.

	UGT2B4		UGT2B7		UGT2B10		UGT2B11		UGT2B15		UGT2B17	
Tissue site	mean	# BLQ	mean	# BLQ	mean	# BLQ	mean	# BLQ	mean	# BLQ	mean	# BLQ
Liver	1.06E-01	0 of 26	3.20E-02	0 of 26	3.19E-02	0 of 26	7.26E-03	0 of 26	1.29E-01	2 of 26	4.08E-02	5 of 26
Liver Tumor	6.99E-03	1 of 5	1.03E-02	1 of 5	2.85E-03	3 of 5	1.19E-02	1 of 5	1.62E-01	3 of 5	1.12E-02	2 of 5
Lung	8.92E-04	22 of 23	2.55E-04	14 of 23	1.99E-03	21 of 23	1.20E-02	2 of 23	1.67E-02	9 of 23	2.75E-02	7 of 23
Lung Tumor	1.86E-03	20 of 23	2.04E-03	10 of 23	2.69E-02	15 of 23	5.29E-02	4 of 23	6.48E-02	7 of 23	1.18E-01	8 of 23
Larnyx	6.82E-04	2 of 8	BLQ	8 of 8	3.70E-03	3 of 8	6.38E-03	2 of 8	1.93E-02	3 of 8	3.07E-02	2 of 8
Esophagus	3.38E-04	8 of 12	2.47E-02	11 of 12	3.71E-03	9 of 12	2.55E-03	5 of 12	5.12E-02	3 of 12	2.01E-02	7 of 12
Tonsil	4.31E-03	8 of 23	BLQ	23 of 23	9.99E-03	8 of 23	2.52E-02	3 of 23	4.91E-02	6 of 23	2.25E-01	1 of 23
Mouth	3.01E-04	0 of 3	BLQ	3 of 3	1.53E-03	0 of 3	4.24E-03	0 of 3	6.45E-03	0 of 3	1.37E-02	0 of 3
Tongue	1.29E-03	2 of 3	BLQ	3 of 3	BLQ	3 of 3	5.13E-03	2 of 3	BLQ	3 of 3	BLQ	3 of 3
Brain	4.01E-04	1 of 5	BLQ	5 of 5	2.70E-03	1 of 5	3.29E-03	0 of 5	9.81E-03	1 of 5	2.65E-02	2 of 5
Pancreas	3.54E-04	13 of 16	6.53E-03	2 of 16	6.90E-04	13 of 16	1.73E-03	10 of 16	6.52E-02	1 of 16	8.11E-03	10 of 16
Endometrium	9.98E-04	4 of 8	BLQ	8 of 8	5.77E-03	4 of 8	5.14E-03	1 of 8	2.69E-02	4 of 8	2.04E-02	4 of 8

BLQ: Below limit of quantification

*Geometric mean, relative to highest expressing liver sample

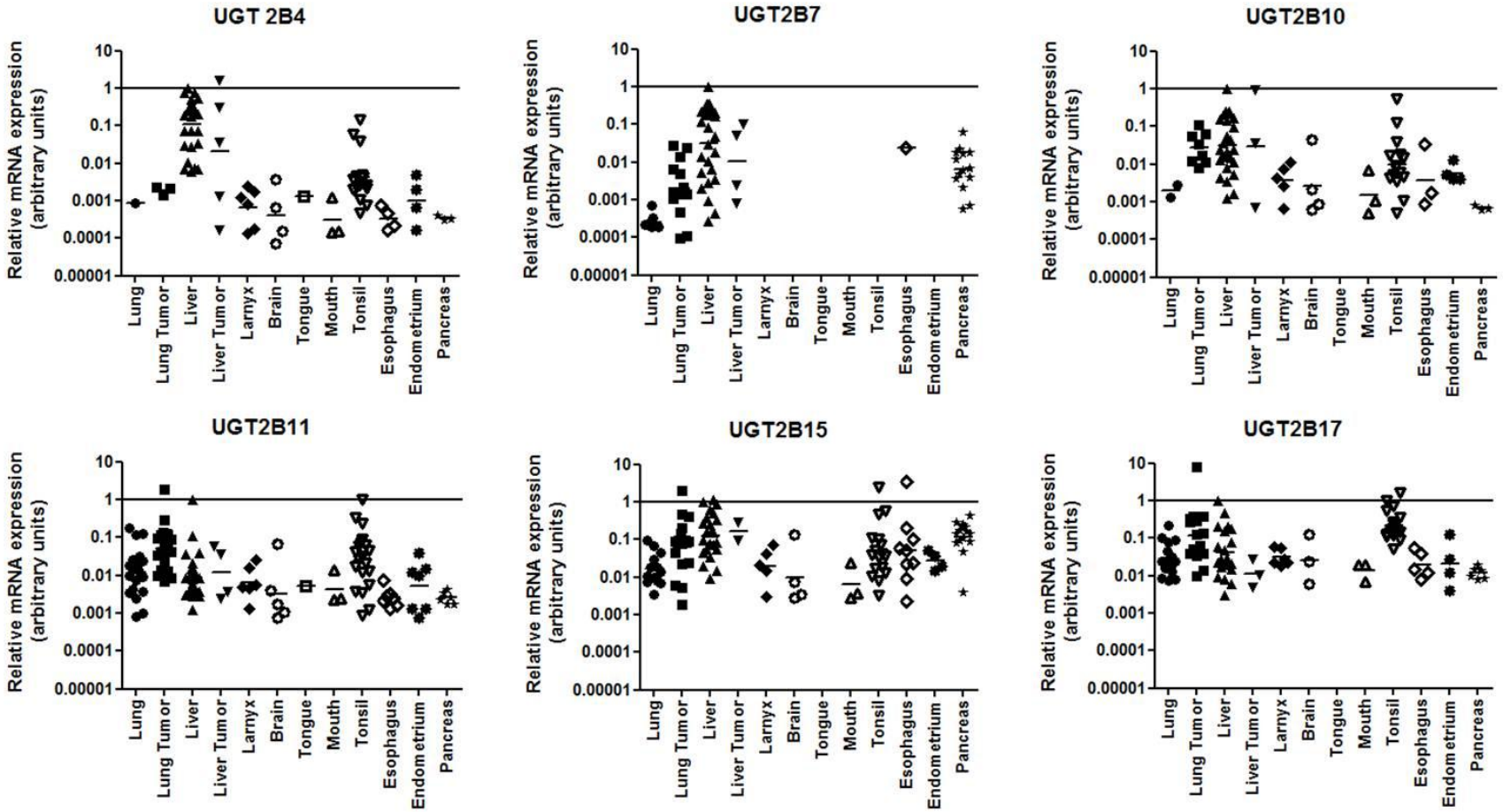

Figure 4.2. Tissular expression levels of UGT2B genes. Individual UGT2B gene expression levels are shown across the tissue panel analyzed in this study. The log of the relative expression levels for the UGT2B genes is shown. The relative abundance was calculated via the $2^{-\Delta\Delta Ct}$ method with the highest expressing liver sample as the calibrator.

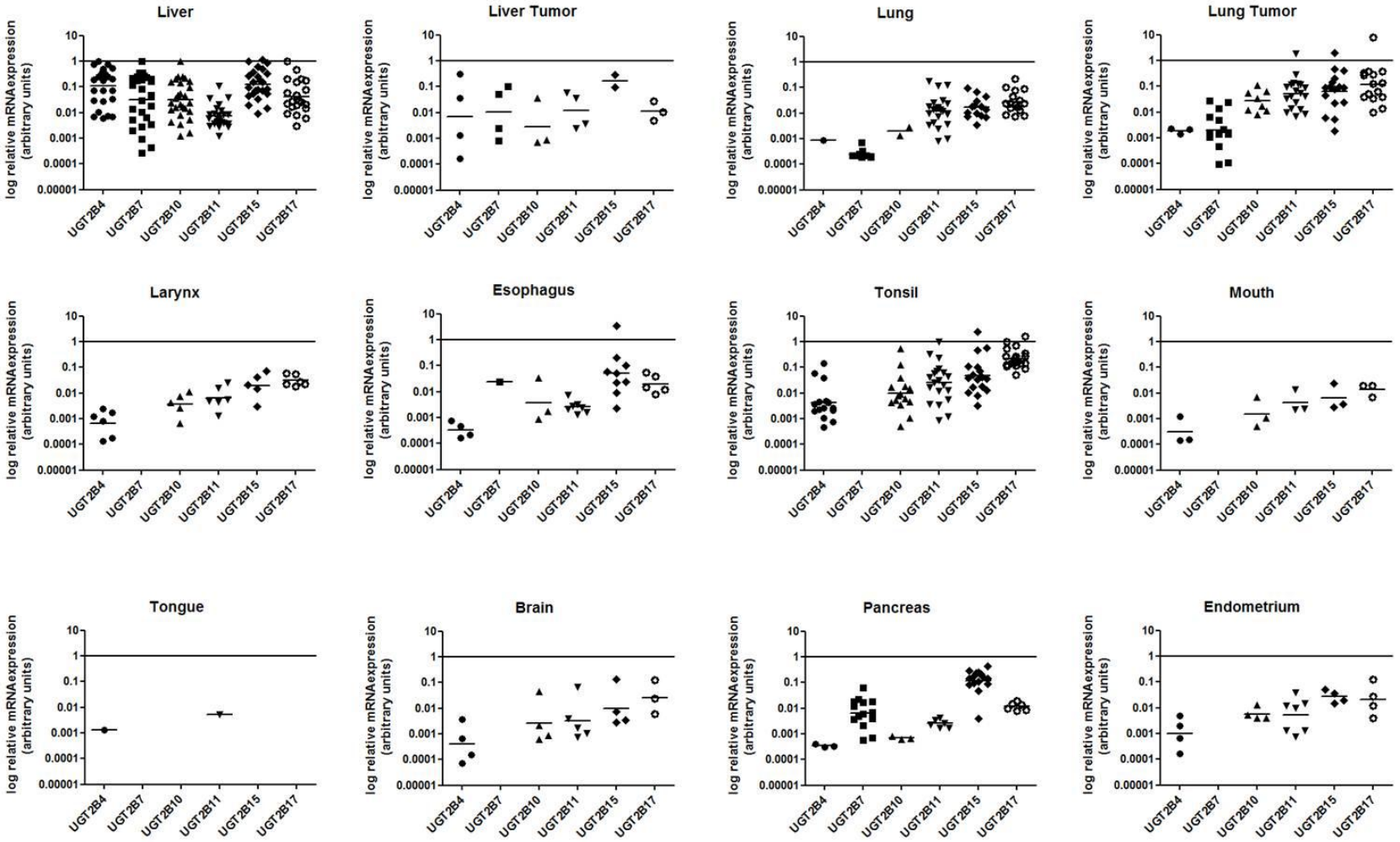

Figure 4.3. Relative expression levels of UGT2B genes in various tissues. UGT2B mRNA expression levels were assessed in 132 normal tissue samples and 28 tumor tissue specimens. The log of the relative expression levels for the UGT2B genes is shown. The relative abundance was calculated via the $2^{-\Delta\Delta Ct}$ method with the highest expressing liver sample as the calibrator.

In human lung tissue, UGT2B17 was found to be expressed the highest, followed by UGT2B15>UGT2B11>UGT2B10>UGT2B7>UGT2B4 (Table 4.1). In the 23 human lung tissues analyzed, only UGT2B11, UGT2B15, and UGT2B17 were detected in at least half of the individual specimens. UGT2B4, UGT2B7, and UGT2B10, predominantly known as liver enzymes, were found to be expressed in one, nine, and two of the specimens, respectively. For the genes expressed in at least half of the specimens, the interindividual difference in expression was 224–fold for UGT2B11, 28-fold for UGT2B15, and 30–fold for UGT2B17 (Figure 4.2). Compared to human liver tissue, UGT2B11 was 1.4–fold higher, UGT2B15 was 7.7 –fold lower, and UGT2B17 was 1.5–fold lower (Figure 4.3).

UGT2B gene expression was detected in the majority of individual larynx specimens. The highest expression was seen for UGT2B17, followed by UGT2B15>UGT2B11>UGT2B10 >UGT2B4 (Table 4.1). UGT2B7 was not detected in any of the individual larynx samples. For the genes expressed in at least half of the specimens, the interindividual difference in expression was 3.1–fold for UGT2B17, 24-fold for UGT2B15, 19-fold for UGT2B11, 17-fold for UGT2B10, and 19–fold for UGT2B4 (Figure 4.2). Compared to human liver tissue, UGT2B17 was 1.3–fold lower, UGT2B15 was 6.7–fold lower, UGT2B11 was 1.4–fold lower, UGT2B10 was 8.6-fold lower, and UGT2B4 was 1554-fold lower (Figure 4.3).

In the 12 human esophagus specimens analyzed, only UGT2B11 and UGT2B15 were found to be expressed in at least half of the specimens, with UGT2B15 the higher expressed of the two (Table 4.1). UGT2B7 was only detected in one specimen, while UGT2B4, UGT2B10, and UGT2B17 were detected in four, three, and five specimens, respectively. For the genes expressed in at least half of the specimens, the interindividual difference in expression was 5.6–fold for UGT2B11 and 1473-fold for UGT2B15 (Figure 4.2). Compared to human liver tissue, UGT2B11 was 3.4–fold lower and UGT2B15 was 2.5–fold lower (Figure 4.3).

With the exception of UGT2B7 (not expressed), UGT2B genes were found to be well expressed in the majority of human tonsil specimens analyzed. The expression was highest for UGT2B17, followed by UGT2B15>UGT2B11>UGT2B10>UGT2B4 (Table 4.1). For the genes expressed in at least half of the specimens, the interindividual difference in expression was 33–fold for UGT2B17, 754-fold for UGT2B15, 1177–fold for UGT2B11, 1096-fold for UGT2B10, and 313–fold for UGT2B4 (Figure 4.2). Compared to human liver tissue, UGT2B17 was 5.5–fold higher, UGT2B15 was 2.6–fold lower, UGT2B11 was 2.9–fold higher, UGT2B10 was 3.2–fold lower, and UGT2B4 was 25–fold lower (Figure 4.3).

The three floor of mouth tissue specimens showed expression of all UGT2B isoforms except UGT2B7. The highest expression was UGT2B17, followed by UGT2B15>UGT2B11> UGT2B10> UGT2B4. For these, the interindividual difference in expression was 2.8–fold for UGT2B17, 8.5-fold for UGT2B15, 6.3–fold for UGT2B11, 13-fold for UGT2B10, and 8.9-fold for UGT2B4 (Figure 4.2). Compared to human liver tissue, UGT2B17 was 2.9–fold lower, UGT2B15 was 20–fold lower, UGT2B11 was 2.1–fold lower, UGT2B10 was 21-fold lower, and UGT2B4 was 352-fold lower (Figure 4.3).

Of the five human brain specimens analyzed, detectable levels of expression were again found for all of the UGT2B genes except UGT2B7. Highest expression was found for UGT2B17, followed by UGT2B15>UGT2B11>UGT2B10> UGT2B4 (Table 4.1). For the genes expressed in at least half of the specimens, the interindividual difference in expression was 21–fold for UGT2B17, 48-fold for UGT2B15, 87-fold for UGT2B11, 74–fold for UGT2B10, and 54–fold for UGT2B4 (Figure 4.2). Compared to human liver tissue, UGT2B17 was 1.5–fold lower, UGT2B15 was 13–fold lower, UGT2B11 was 2.7–fold lower, UGT2B10 was 12–fold lower, and UGT2B4 was 264–fold lower (Figure 4.3).

The UGT2B genes detected in at least half of the 16 pancreas specimens were UGT2B15 and UGT2B7, with UGT2B15 the higher expressed of the two (Table 4.1). UGT2B4, UGT2B10,

UGT2B11, and UGT2B17 were expressed in three, three, six, and six of the pancreas samples examined. For the genes expressed in at least half of the specimens, the interindividual difference in expression was 891–fold for UGT2B15 and 102-fold for UGT2B7 (Figure 4.2). Compared to human liver tissue, UGT2B15 was 1.1–fold lower and UGT2B7 was 4.9–fold lower (Figure 4.3).

With the exception of UGT2B7 (not expressed), all UGT2B genes were detected in at least half of the eight endometrium samples. UGT2B15 was the highest expressed, followed by UGT2B17>UGT2B10>UGT2B11> UGT2B4 (Table 4.1). For the genes expressed in at least half of the specimens, the interindividual difference in expression was 3.5–fold for UGT2B15, 33–fold for UGT2B17, 3.1-fold for UGT2B10, 51–fold for UGT2B11, and 30–fold for UGT2B4 (Figure 4.2). Compared to human liver tissue, UGT2B15 was 4.8–fold lower, UGT2B17 was 2.0–fold lower, UGT2B10 was 5.5–fold lower, UGT2B11 was 1.7 –fold lower, and UGT2B4 was 106–fold lower (Figure 4.3).

Relative UGT2B gene expression in matched tumor and adjacent normal samples. In five matched HCC specimens, the mean expression of the UGT2B genes was not significantly different than the normal human liver specimens. UGT2B4, UGT2B7, and UGT2B11 were not detected in one of the specimens, UGT2B 17 was absent in two of them, while both UGT2B10 and UGT2B15 were absent in three.

In the 23 matched lung tumors (15 adenocarcinoma, 8 squamous cell carcinoma), UGT2B gene expression was significantly higher in the tumor specimens for UGT2B7, UGT2B11, UGT2B15, and UGT2B17 (Figure 4.4). For UGT2B4 and UGT2B10, no comparison could be made because these genes were not found to be expressed in the normal specimens even though they were present in the corresponding tumor samples. Many of the adjacent normal specimens that did not show expression showed detectable expression in the tumor tissue. The mean expression in lung tumor tissue was 2.1–fold higher for UGT2B4, 8.7-fold higher for

UGT2B7 (p=0.0134), 15-fold higher for UGT2B10, 4.3-fold higher for UGT2B11 (p=0.0019), 3.8-fold higher for UGT2B15 (p=0.017), and 4.4–fold higher for UGT2B17 (p=0.0056), as compared to their corresponding normal tissues. UGT2B4 expression was only detected in one of the normal tissue specimens and three of the tumor specimens; UGT2B7 was detected in nine of the normal specimens and 13 of the tumor specimens; UGT2B10 was detected in two of the normal specimens and eight of the tumor specimens; UGT2B11 was detected in 21 of the normal specimens but only 19 of the tumor specimens; UGT2B15 was detected in 14 of the normal specimens and 16 of the tumor specimens; and UGT2B17 was detected in 16 of the normal specimens and 15 of the tumor specimens. For UGT2B17, the same samples that showed no expression in the normal specimens showed no expression for the tumor samples, suggesting that these individuals were homozygous for the UGT2B17 whole gene deletion.

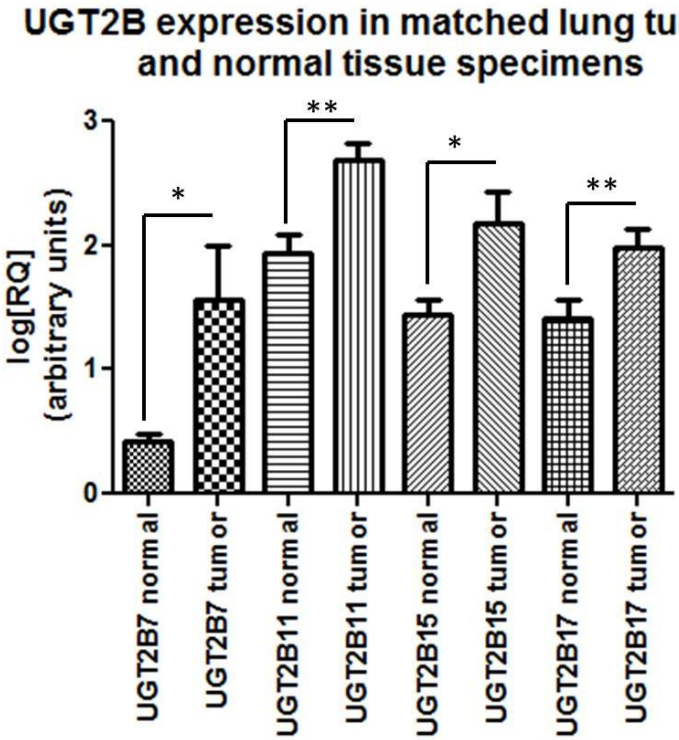

Figure 4.4. Relative expression levels of UGT2B genes in matched lung tumor and adjacent normal lung tissue specimens. Mean UGT2B expression levels were compared between tumor and adjacent normal tissue via a paired t-test. RQ values were calculated via the $2^{-\Delta\Delta Ct}$ method with the lowest expressing lung sample as the calibrator. *p<0.05; **p<0.01

Correlation between human UGT2B gene expression levels in human tissues. Correlative analysis of UGT2B expression in the liver, lung, and tonsils is shown (Table 4.2). These tissues were chosen because there was robust expression of UGT2B genes in a high percentage of specimens. The other tissue sites either did not contain enough samples for meaningful analysis or there was not good expression of the various UGT2B genes. In liver, there was significant correlation between the expression of UGTs 2B4, 2B7, 2B10, and 2B15 ($p<0.0001$). UGT2B11 and UGT2B17 expression levels were not significantly correlated with any of the other UGT2B genes. In the lung, UGT2B15 expression levels were significantly correlated with expression levels of UGTs 2B11 and 2B17 ($p<0.01$). In tonsil, all of the UGT2B genes were significantly correlated with each other ($p<0.05$), with the exception of UGT2B7 which was not expressed at detectable levels in any of the samples.

Table 4.2. Correlations between UGT mRNA levels in different tissue specimens.

Liver

	UGT2B4	UGT2B7	UGT2B10	UGT2B11	UGT2B15	UGT2B17
UGT2B4		0.92****	0.86****	-0.35	0.85****	0.01
UGT2B7			0.86****	-0.36	0.81****	0.08
UGT2B10				-0.23	0.97****	0.33
UGT2B11					-0.24	0.29
UGT2B15						0.20
UGT2B17						

Lung

	UGT2B4	UGT2B7	UGT2B10	UGT2B11	UGT2B15	UGT2B17
UGT2B4		N/A	N/A	N/A	N/A	N/A
UGT2B7			N/A	-0.39	0.54	-0.37
UGT2B10				N/A	N/A	N/A
UGT2B11					0.73**	0.48
UGT2B15						0.87***
UGT2B17						

Tonsil

	UGT2B4	UGT2B7	UGT2B10	UGT2B11	UGT2B15	UGT2B17
UGT2B4		N/A	0.94****	0.91****	0.92****	0.60*
UGT2B7			N/A	N/A	N/A	N/A
UGT2B10				0.76***	0.91****	0.69**
UGT2B11					0.84****	0.54*
UGT2B15						0.75***
UGT2B17						

Data presented are r^2 values from correlation analyses of expressed isoforms using Spearman's rank method.
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001; N/A, not expressed

4.5 Discussion

The UGT2B gene family is located in a cluster on chromosome 4q13 and is responsible for the detoxification of a wide range of endogenous and xenobiotic compounds. Hepatic expression of the UGT2B family members has been described previously [153, 298]. In this study, highest expression was seen for UGT2B15 and UGT2B4 in the liver, followed by

UGT2B7, which agrees well with previous studies [153, 298]. UGT2B4 is predominantly known as a steroid- and bile acid-glucuronidating UGT, so it is needed in high abundance in the liver [321, 324, 325]. UGT2B4 was among the lowest expressed in all of the extrahepatic tissues analyzed in this study. UGT2B7 is another enzyme that is primarily a hepatic enzyme, as it was either not expressed or expressed at very low levels in all of the other tissues examined except the pancreas. UGT2B10 and UGT2B11 were the lowest expressed UGT2B enzymes in human liver.

In lung, UGT2B11 was expressed the highest, though the role of this gene in this tissue is unclear, considering the narrow range of endogenous and exogenous substrates that have been identified for this enzyme [326]. UGT2B15 and UGT2B17 were the only other UGT2B genes expressed in at least half of the lung tissue specimens. High expression of UGT2B17 in lung tissue is consistent with earlier reports that the UGT2B17 gene deletion is associated with lung adenocarcinoma risk [130], since it detoxifies NNAL, the major carcinogenic metabolite of NNK which induces lung adenocarcinoma in animal models. The high expression of UGT2B17 in lung is consistent with a model of carcinogen detoxification in target tissues [327, 328].

The larynx, tonsils, esophagus, and floor of mouth all show similar patterns of expression. Though the relative levels are different between these tissues, they are characterized by similar expression patterns of UGT2B10, UGT2B11, UGT2B15, and UGT2B17, with very low expression of UGT2B4 and little to no expression of UGT2B7. The similarity in expression patterns in these tissues should be expected given their close proximity and similar exposures. Perhaps the most interesting aspect of the expression patterns in these tissues is the expression of UGT2B10 and UGT2B17, the two enzymes that catalyze the majority of NNAL metabolism. Further studies are necessary to determine whether the UGT2B10 D67Y or the UGT2B17 gene deletion polymorphisms are associated with cancer risk in any of these tissues, similar to what was seen with the UGT2B17 gene deletion in the lung [130].

In the five brain specimens analyzed, the expression pattern was also similar to the pattern seen in the upper aerodigestive tract tissues. In the brain, interindividual differences in expression and activity of UGT2B10 may affect the glucuronidation of nicotine. This could impact how an individual responds and becomes addicted to this compound. Further studies are necessary to investigate this possibility.

The pancreas showed a unique pattern of expression, with high expression of UGT2B7 and UGT2B15, and little to no expression of other UGT2B genes. UGT2B7 and UGT2B15 are involved in the metabolism of estrogens and androgens [176, 329], respectively, though their role in pancreatic physiology is unclear.

In many cases, interindividual differences in UGT2B expression were large, making comparisons between people difficult. The wide range of expression differences even within each tissue site agrees well with previous publications [152, 181]. There are several potential reasons for this. Differences in RNA quality can affect transcripts of different abundance differently [290, 291]. Some of the tissue samples had very high quality RNA (RIN>8.0), whereas other specimens were partially degraded (RIN between 5.0 and 8.0). These differences in quality may affect degradation of the UGT2B and MT-ATP6 transcripts differently and are difficult to control. Secondly, most tissues have diverse mixtures of cell types, and depending on which section of tissue is excised, different cell types may be harvested from different specimens. This could have an impact on comparing transcript levels between different tissue specimens in highly heterogenous tissues such as the lung or pancreas.

The higher expression of UGT mRNAs in matched lung tumor tissue may be important for several reasons. First, higher UGT expression in tumor tissue may represent a selective advantage for the tumor by enabling rapid elimination of genotoxic compounds that might otherwise suppress growth of malignant cells. Furthermore, higher UGT gene expression in tumor tissue relative to adjacent normal tissue may increase resistance to chemotherapeutic

agents. The multi-drug resistance phenotype observed for many cancers is known to be caused by overexpression and mutation of drug efflux transporters, such as P-glycoprotein and MDRs [330, 331]. This leads to increased efflux of drugs and lower cellular concentrations of cytotoxic chemotherapy agents. Overexpression of UGT genes may also contribute to this phenotype by increasing the rate of detoxification of chemotherapeutic compounds. Alternatively, UGT gene overexpression in tumor tissues may be caused by differences in the cell types present. For example, within the lung, epithelial cells are known to express the UGT enzymes, and these cells may be overrepresented in the tumor specimens as compared to the adjacent normal tissue. This would lead to a high estimate of UGT gene expression in tumor specimens that may in reality be an artifact of the methodology. In future studies, a technique such as laser capture microdissection may be necessary to compare individual cell types with respect to UGT gene expression levels in matched tumor and adjacent normal specimens.

In the liver, the correlation between expression levels of UGT2B4, UGT2B7, UGT2B10, and UGT2B15 but not UGT2B17 is in agreement with previous studies [298]. This is likely due to a common mechanism of transcriptional regulation for these UGT2B isoforms in human liver. One potential transcription factor that may explain the correlation between these isoforms is HNF1α, which has been shown to be a factor involved in the transcriptional regulation of several UGT2B genes [166, 172, 184, 332, 333]. In lung tissue, UGT2B15 expression levels were significantly correlated with UGT2B17 expression, suggesting a coordinated mechanism of transcriptional regulation for these two genes. The lack of correlation of these two genes in liver suggests that the transcriptional regulation of these genes is occurring by different transcription factors within these different tissues. This makes sense considering that these genes are predominantly controlled by LETFs in the liver, though the mechanisms of transcriptional control in lung remain unknown. In tonsil, all of the expressed isoforms are correlated with each other, suggesting a common regulatory mechanism for the UGT2B gene family in this tissue. The

factors involved in regulation of UGT genes in upper aero-digestive tract tissues have not been explored.

Understanding the relative abundance of UGT gene expression levels in different tissues and different individuals will help direct future studies based on where individual isoforms are likely to have physiological relevance. The results presented in this study may help to better understand how these expression patterns affect the metabolism of a wide range of endogenous, clinical, environmental, or toxic exposures.

Chapter 5

ROLE OF HNF3α AND Oct-1 IN THE HEPATIC REGULATION OF UGT2B10

5.1 Abstract

While several studies have investigated the transcriptional regulation of UGT2B enzymes, there have been no studies focusing on UGT2B10. In the present study, the liver-derived HepG2 cell line was used to examine the regulation of the UGT2B10 gene. In order to determine the regions responsible for UGT2B10 promoter activity, six UGT2B10 promoter constructs of increasing length were cloned into the PGL3 luciferase vector (+27bp/-58bp; +27bp/-121bp; +27bp/-198bp, +27bp/-472bp; +27bp/-973bp; 27bp/-1948bp; relative to the UGT2B10 transcription start site). Promoter constructs consisting of up to 121 bp of proximal UGT2B10 promoter exhibited no significant difference in transcriptional activity compared to the promoterless, negative control luciferase vector. A 7.8-fold and 37-fold increase in luciferase activity was observed for the +27bp/-198bp and +27bp/-472bp constructs (p=0.0004 and p=0.0001, respectively). The two longest constructs (+27bp/-973bp and +27bp/1948bp), were not significantly different from the +27bp/-472bp construct, indicating that the core regulatory element responsible for UGT2B10 promoter activity was located between -121 bp and -472 bp upstream of the transcription start site. DNaseI footprinting was used to identify four distinct protected sequences within this region, with Transcription Element Search Software (TESS) used to identify HNF3α and Oct-1 as putative transcription factors that may bind to these protected regions. Electrophorectic mobility shift assay (EMSA) and gel supershift were used to demonstrate that HNF3α and Oct-1 proteins could bind to these sequences. Mutation of the HNF3α site in the +27bp/-1948bp UGT2B10 promoter construct led to a 49% decrease in promoter activity, and mutation of the Oct-1 site led to a 61% decrease (p=0.0095 and p=0.0004, respectively), as compared to the wild type (WT) +27bp/-1948bp construct. Mutation of both sites led to a 78% decrease in promoter activity (p=0.0002). Mutation of the TATAA box at -

29bp within the +27bp/-1948bp vector led to >95% decrease in UGT2B10 promoter activity, demonstrating that the TATAA box is necessary but not sufficient for UGT2B10 promoter activation. This study demonstrates that the UGT2B10 promoter is regulated by an element between -170bp and -250bp and is dependent on HNF3α and Oct-1 binding sites for maximal activity.

5.2 Introduction

Investigation of the proximal promoter of UGT2B genes has led to the identification of several transcription factor binding sites that partially account for the observed differences in expression seen in different tissues [57, 177, 178, 334]. The difference in expression patterns of UGT2B genes suggests that differential regulation may be occurring at the level of gene transcription, and contributes to differences in glucuronidation rates seen in different tissues [335]. Regulation at the transcriptional level is also a major reason for the differences observed in glucuronidation rates in the same organ between different individuals [164, 336]. These differences influence the ability of an individual to eliminate compounds and reduce their toxicity or carcinogenic potential. There are several examples of polymorphisms in UGT gene promoters that lead to changes in expression and glucuronidation activity [197, 215, 231, 302].

The liver is the predominant organ for glucuronidation, and all of the UGT2B genes are well-expressed in this tissue [176]. While some of the factors regulating expression of UGT2B genes in human liver have been elucidated for UGT2B genes, the transcriptional regulation of the UGT2B10 gene has not been characterized. It is known that UGT2B10 plays a role in the metabolism of therapeutic (OLZ), carcinogenic (NNAL), and toxic compounds (nicotine), so the need for understanding how it is regulated at the transcriptional level is of paramount importance [95, 113, 138, 142]. For many UGT2B genes, the transcriptional regulation has been shown to be

at least partially under HNF1α transcriptional control. All of the UGT2B gene promoters contain HNF1α sites conserved in sequence and location, and all are expressed in liver [176]. Overexpression of HNF1α has been shown to regulate the expression of all UGTs tested, but to highly varying extents [335]. For example, expression of UGT2B10 was previously shown to be only weakly activated by HNF1α overexpression [335].

In the present study, transcriptional regulation of the UGT2B10 gene promoter was assessed in liver-derived HepG2 cell line as a model of human hepatocytes. These studies demonstrate that the UGT2B10 promoter is dependent upon the LETF, HNF3α, as well as the ubiquitous transcription factor Oct-1 by a promoter element located between -170bp and -250bp upstream from the transcription start site.

5.3 Methods

Materials. Restriction enzymes were obtained from New England Biolabs (Beverly, MA). DNaseI was purchased from Invitrogen (Carlsbad, CA). Pfu polymerase was ordered from Stratagene (Santa Clara, CA). [γ-^{32}P]dATP was purchased from Perkin Elmer (Boston, MA). The dual luciferase assay system and corresponding vectors were purchased from Promega (Madison, WI). The chemiluminescent nucleic acid detection module used for the electrophoretic mobility shift assay (EMSA) was purchased from Pierce (Rockford, IL). Oct-1 and HNF3α antibodies were purchased fromAbcam (Cambridge, MA).

Determination of the UGT2B10 Transcription Start Site. The transcription start site of the UGT2B10 gene was mapped using a modified 5'-Rapid Amplification of cDNA Ends (5'RACE) procedure. Briefly, total RNA from normal liver tissue was reverse transcribed using the following UGT2B10 gene-specific primer (GSP): 5'-CCAGCTTCAAATCTCAGATATAAC

TAATCC-3' (-6 to +24 relative to TAG translation stop codon). The resulting cDNA was purified using the Qiagen Minielute PCR Purification Kit (Valencia, CA). Terminal deoxynucleotidyl transferase (TdT) enzyme was used to add a polyA tail to the 3' end of the purified cDNA. The resulting polyA-tailed cDNA was PCR amplified with oligo (dT) and an internal UGT2B10 GSP 5'-AAACTGAGTTGTATCAGCAGAACTGT-3' (+18 to +44 relative to ATG translation start site). Due to the low melting temperature of the oligo (dT) primer during the PCR reaction, it was determined that the optimal conditions for amplification were as follows: 95°C for 2 min; 40 cycles of 95°C for 20 sec, 40°C for 30 sec, 72°C for 2 min; then one cycle of 72°C for 10 min. PCR amplified DNA was resolved on a 1.0% agarose gel, extracted using the Qiagen Gel Extraction kit, and sequenced using the same internal GSP that was used for PCR amplification. The resulting sequence was used to determine the precise location of the UGT2B10 transcription start site and the length of the UGT2B10 5'UTR.

Generation of UGT2B10 Promoter Luciferase Constructs. Variable length segments of the UGT2B10 promoter were cloned into the pGL3 basic luciferase vector using Kpn1 and BglII restriction sites just upstream of the luciferase gene (Figure 5.1). A schematic representation of the UGT2B10 promoter constructs and their nomenclature is shown in Figure 5.2. The primers used for generation of the UGT2B10 promoter constructs are shown in Table 5.1. All sense primers contained a KpnI restriction enzyme site at their 5' ends. The antisense primer was the same for all promoter constructs and contained a BglII restriction enzyme site at its 5' end.. The +27bp/-58 bp construct was generated by ordering the complementary 85bp oligonucleotides with KpnI and BglII restriction sites and annealing them.

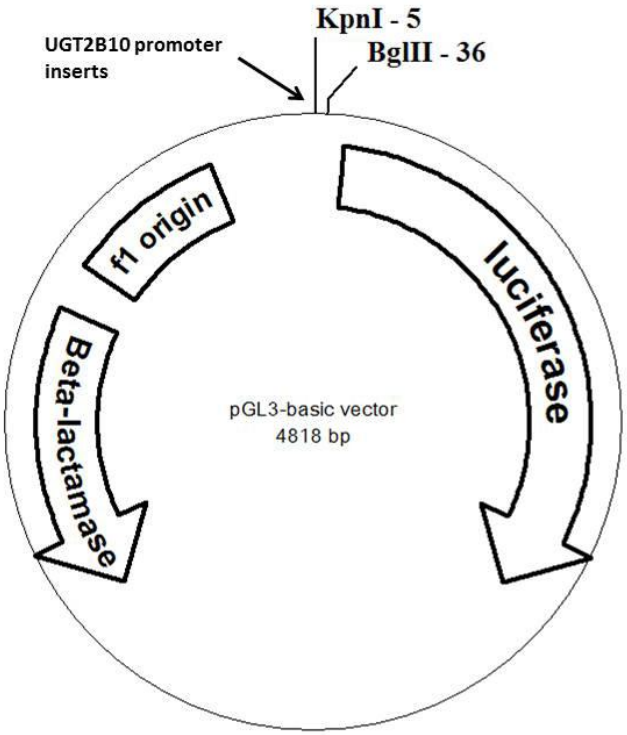

Figure 5.1. pGL3-Basic luciferase vector. The vector above was used for generation of UGT2B10 promoter constructs. The UGT2B10 promoter fragments were inserted in between the KpnI and BglII sites in the cloning region just upstream from the luciferase reporter gene.

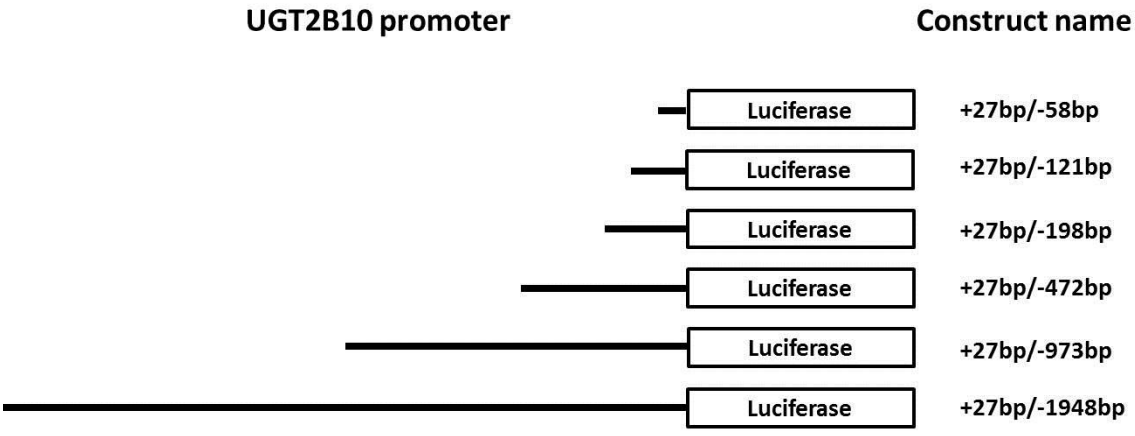

Figure 5.2. Schematic representation and nomenclature of the UGT2B10 promoter constructs. The constructs were named according to their start and end position relative to the UGT2B10 transcription start site.

Table 5.1. List of primers used for generation of UGT2B10 promoter constructs.

Primer name	Primer sequence[a] (5' - 3')	Position[b]
+27bp/-121bp sense	5'-ATACGGTACCGTCTCCTTGCCATCCATGTACTCA-3'	-121 bp to -97 bp
+27bp/-198bp sense	5'-ATACGGTACCGGTCATTAAACTTAGGTGGATGACACC-3'	-198 bp to -171 bp
+27bp/-472bp sense	5'-ATACGGTACCGATTGCTATGTTGGATGCACAAGTCTGTG-3'	-472 bp to -443 bp
+27bp/-973bp sense	5'-ATACGGTACCCATGTCTACAGAAAGCATGTACATCAATG-3'	-973 bp to -942 bp
+27bp/-1948bp sense	5'-ATACGGTACCGCTCTGAGTCCTGTGAGAATGTGC-3'	-1948 bp to -1924 bp
UGT2B10 promoter antisense	5'-AGACAGATCTCCTTGTGCAATGTGATAATTCTTTTCCAGTCAC-3'	+27 bp to -6 bp

[a]KpnI and BglII sites are underlined for sense and antisense primers, respectively.
[b]Relative to transcription start site

Cell culture and Transfection of Luciferase Constructs into HepG2 Cells. HepG2 cells were maintained in Dubelco's modified eagle medium (DMEM) supplemented with 10% FBS, 1% Pen/Strep, 1X non-essential amino acids (NEAA), and L-Glutamine. $1x10^6$ cells were plated 24 h prior to transfection with Lipofectamine2000 reagent, according to manufacturer's protocols. 4 μg of UGT2B10 promoter constructs were transfected into HepG2 cells and assayed for luciferase activity after 24 h. 100 ng of the pRL-TK vector was used as an internal control for transfection efficiency.

DNaseI Footprint Analysis of the UGT2B10 Promoter. HepG2 nuclear extracts were generated by the method of Dignam et al [337]. DNaseI footprinting was performed according to the method of Kroeger and Abraham [338]. Based on the luciferase data, two regions of interest were determined for the optimal activity of the UGT2B10 promoter (+27bp/-198bp and -198bp/-472bp). The primer pairs used for the generation of DNaseI footprinting probes are shown in Table 5.2. Four distinct probes were necessary so there would be forward and reverse orientation of both the +27bp/-198bp and -198bp/-472bp promoter regions. Forward end-labeled and reverse end-labeled probes were generated by PCR amplification under the following conditions: 95°C

for 2 min; 40 cycles of 95°C for 20 sec, 60°C for 30 sec, 72°C for 2 min; then one cycle of 72°C for 10 min. PCR-amplified probes were immobilized on Dynabeads M-280 streptavidin coated beads (Carlsbad, CA). Probes were treated with 100 μg of HepG2 nuclear extract or 100 μg BSA at room temperature for 30 min prior to digestion with 0.0075U DNaseI for 1 min. Digested UGT2B10 promoter probes were run on a 6% sequencing gel at 3000 V for 2.5 h. Sequencing ladders were generated using the USB Sequenase Quick Denature Plasmid Sequencing Kit (Santa Clara, CA) and run in parallel with footprinting probes. Gels were dried for 2 h using a BioRad Gel Dryer with a vacuum pump attachment (Hercules, CA). Dried gels were exposed to a GE phosphorimager screen overnight at room temperature (Piscataway, NJ).

Table 5.2. Primers used for generation of UGT2B10 promoter DNaseI footprinting probes.

Probe	Strand	Primer sequence (5' - 3')	Position[a]
1	Sense	5′^{32}P-GGTCATTAAACTTAGGTGGATGAC-3′	-198 bp to -173 bp
	Antisense	5′biotin-CCTTGTGCAATGTGATAATTCTTTTCC-3′	+27 bp to -1 bp
2	Sense	5′biotin-GGTCATTAAACTTAGGTGGATGAC-3′	-198 bp to -173 bp
	Antisense	5′^{32}P-CCTTGTGCAATGTGATAATTCTTTTCC-3′	+27 bp to -1 bp
3	Sense	5′^{32}P-GATTGCTATGTTGGATGCACAAGTCTGT-3′	-473 bp to -445 bp
	Antisense	5′biotin-TTGCAAATATCGTGTTTTATATAATG-3′	-198 bp to -225 bp
4	Sense	5′biotin-GATTGCTATGTTGGATGCACAAGTCTGT-3′	-473 bp to -445 bp
	Antisense	5′^{32}P-TTGCAAATATCGTGTTTTATATAATG-3′	-198 bp to -225 bp

[a]Relative to transcription start site

EMSA and Gel Supershift Assay. DNaseI footprinting results indicated three regions protected by HepG2 nuclear extract DNA binding proteins. Biotinylated DNA oligomers corresponding to these protected 'footprints' were purchased from IDT DNA technologies:

HNF3α 5'-ATAAAACACGATATTTGCAAGGTCATTAAA-3'; Oct-1 5'-TCTTGACAC TATTATAAAATACATTATATA-3'; TATA 5'-TTTATCTCTGTATATAAGTATGAGA AAGAA-3'; footprint#2 5'-CTTAGGTGGATGACACCAA ATATAAACATA-3'. Mutant competitor oligonucleotides of the putative UGT2B10 promoter protein binding sites were as follows: HNF3α 5'-ATAAAACACGATAAGCGCAAGGTCATTAAA-3'; Oct-1 5'-TCTTGACACTATTACGGAATACATTATATA-3'; TATA 5'-GATTTATCTCTGTACGGAA GTATGAGAAAG-3'; footprint#2 5'- CTTAGGTGGAAACCACCAAATATAAACATA-3' (mutated nucleotides underlined). Sense and antisense oligomers were annealed and used with the Pierce Lightshift EMSA Optimization and Control Kit. Briefly, 25 fmol of biotinlyated promoter oligomers were reacted with 25 μg HepG2 nuclear extract or 25 μg BSA. DNA-protein hybrids were run on a non-denaturing 6% polyacrylamide gel for 1 h and transferred to an Amersham Biosciences positively charged nylon membrane. The DNA probes were crosslinked to the nylon membrane using a Stratagene UV Stratalinker 2400. Crosslinked DNA probes were visualized using the Pierce Chemiluminescent Nucleic Acid Detection Module. Gel supershift assay conditions were identical to EMSA conditions, except that 1 μl of 1:1000 diluted Oct-1 and 1:2000 diluted HNF3α AbCam antibodies were added to the corresponding binding reactions.

UGT2B10 promoter mutational analysis. Site-directed mutagenesis of the +27bp/-1948bp UGT2B10 promoter construct was accomplished using the Stratagene Site-Directed Mutagenesis Kit. Primers used for the generation of UGT2B10 promoter mutants were as follows: HNF3α 5'-ATAAAACACGATAAGCGCAAGGTCATTAAA-3'; Oct-1 5'-TCTTG ACACTATTACGGAATACATTATATA-3'; TATA 5'-GATTTATCTCTGTACGGAAGTATG AGAAAG-3'; footprint#2 5'-CTTAGGTGGAAACCACCAAATATAAACATA-3'. Promoter mutants were assayed for luciferase activity as described above.

Statistical Analyses. Luciferase activity of the UGT2B10 promoter constructs were compared to the empty PGL3-basic vector using the Student's t-test. Luciferase activity of the +27/-1948 bp promoter mutants and double mutants were compared to the WT +27/-1948 bp promoter using the Student's t-test.

5.4 Results

Mapping the UGT2B10 promoter transcription start site. The transcription start site of the UGT2B10 promoter was mapped according to the modified 5'RACE procedure outlined in the Materials and Methods. Using this method, the UGT2B10 promoter was determined to have a 27 bp 5'UTR sequence (Figure 5.3). The length of the 5'UTR was the same as predicted by the UCSC genome browser, confirming the validity of the *in silico* approach.

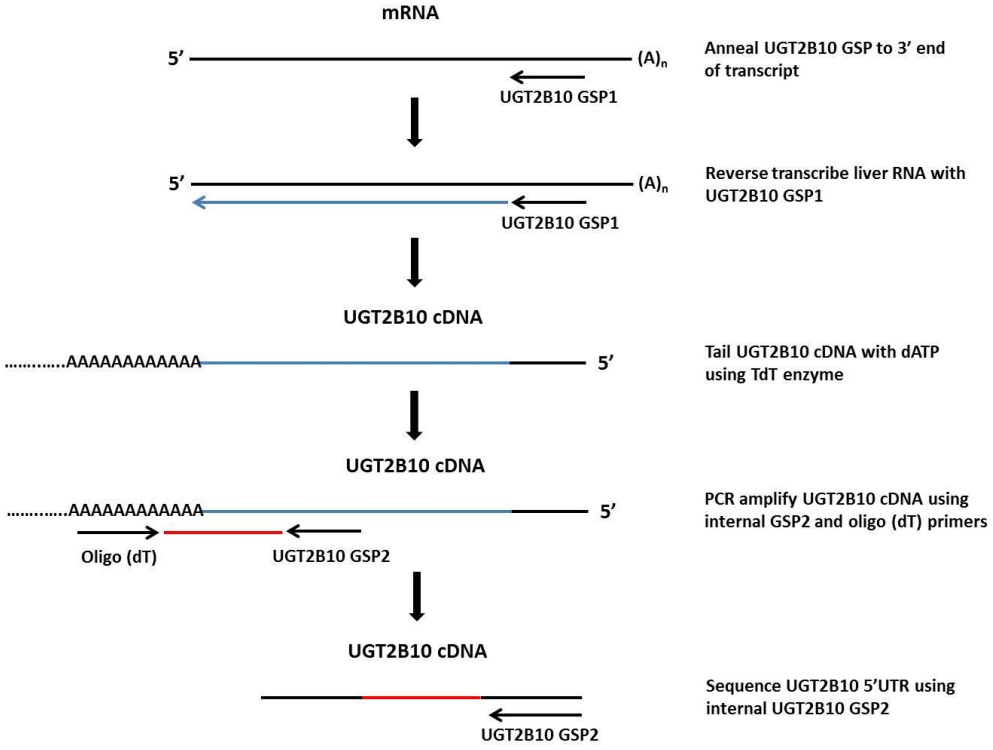

Figure 5.3. Schematic representation of the modified 5'RACE procedure used for mapping the UGT2B10 transcription start site.

Luciferase activity of the UGT2B10 promoter constructs. The variable length UGT2B10 promoter constructs were assayed for luciferase activity and compared to the pGL3-basic empty vector (Figure 5.4). It was determined that the +27bp/-58bp and +27bp/-121bp promoter constructs did not differ in activity from the negative control. The +27bp/-198bp and +27bp/-472bp constructs exhibited a 7.8– and 37-fold increase in activity, respectively, relative to the negative control (*p*-values<0.001). Activity of the +27bp/-472bp construct had a significant increase in activity as compared to the +27bp/-198bp construct (*p*-value=0.001). The activities of the +27bp/-973bp and +27bp/-1948bp constructs did not differ significantly from the +27bp/-472bp construct. Based on this data, the first 499 bp promoter were determined to contain the core promoter activity required for expression of the UGT2B10 gene.

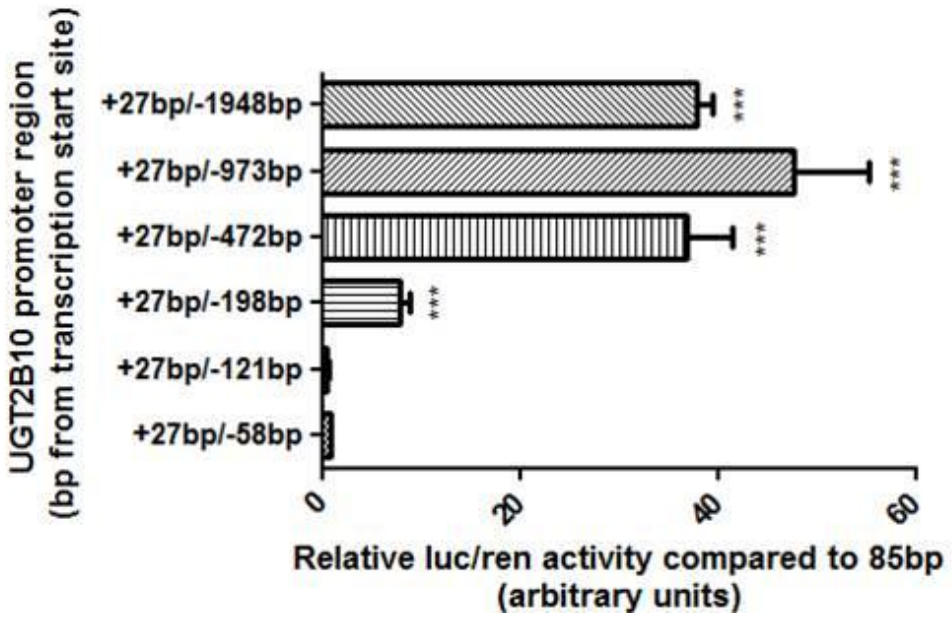

Figure 5.4. UGT2B10 promoter driven luciferase activity in the HepG2 cell line. Luciferase absorbance values were normalized to renilla and the graph depicts the mean ± SE of four separate experiments. ** $p < 0.01$; *** $p<0.001$

DNaseI footprinting to determine regions of DNA-protein interactions. After determining that the +27bp/-198bp and +27bp/-472bp constructs had independent regions necessary for maximal activity of the UGT2B10 promoter, these two promoter regions were

analyzed by DNaseI footprinting to determine the location of DNA-protein interactions (Figure 5.5). It can be seen that the +27bp/-198bp construct contains two distinct protected footprints (Figure 5.5, Panel A). DNaseI footprinting was also used to analyze the promoter region from -198 bp to -472- bp, and there were two footprints located between -198 bp and -240 bp of this region (Figure 5.5, Panel B). The first 225 bp of the UGT2B10 promoter were entered into the TESS to determine potential transcription factor binding sites that may be present in the protected footprint elements (Table 5.3). Footprint #1 corresponds to the TATAA box sequence. The proteins responsible for the second footprint near the end of the 225 bp sequence are less clear but may include binding to AP-1, CAC-binding protein, YY1, LCRF-1, and HOXD8. TESS analysis identified an HNF3α binding site and an Octamer protein binding site within footprints #3 and #4, respectively (Table 5.3).

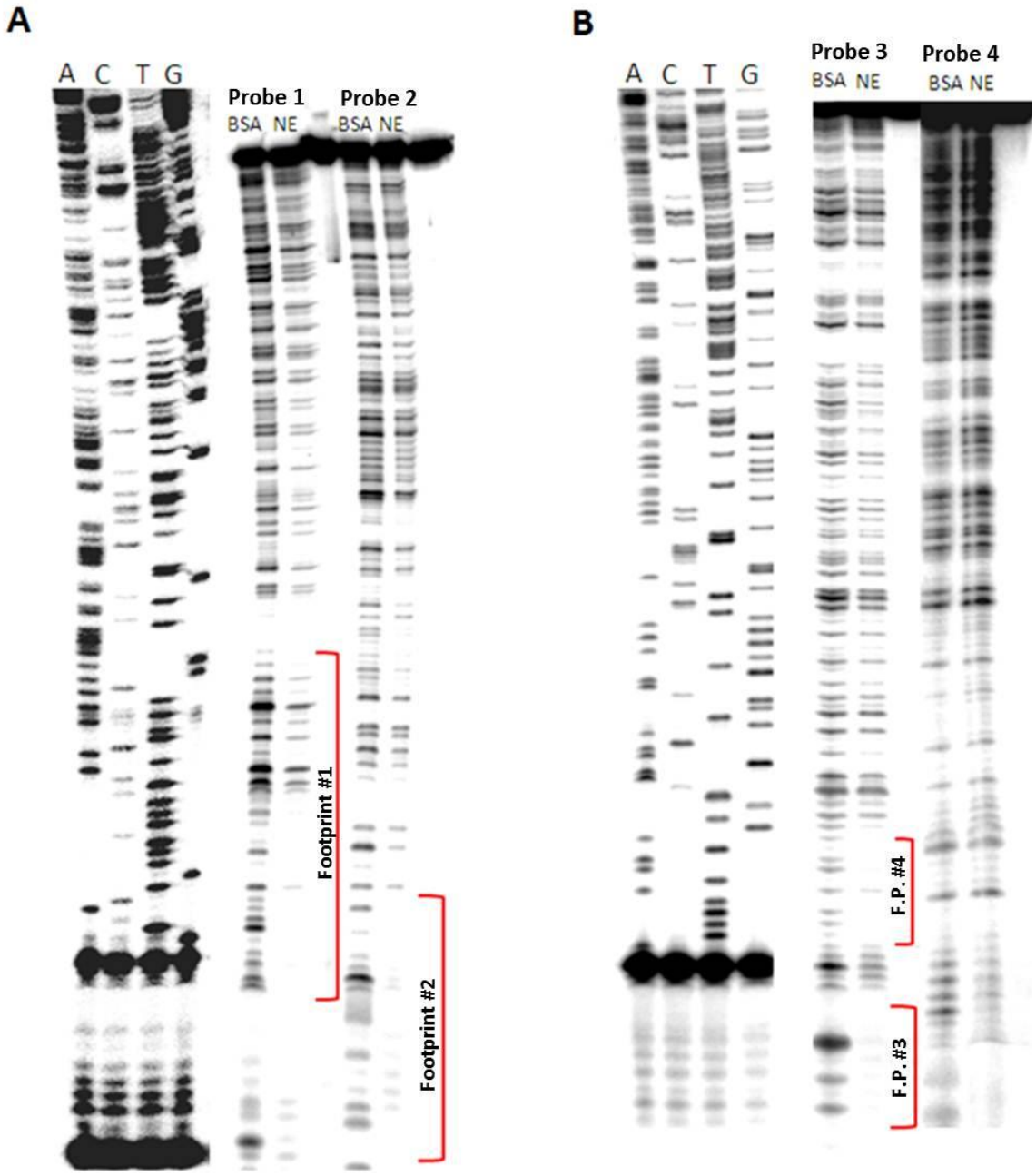

Figure 5.5. Footprint analysis of the UGT2B10 promoter region. Probes were incubated with BSA (100 ug) or HepG2 nuclear extract (NE) (100 ug) for 20 min and then treated with 0.0075U DNaseI for 1 min at RT prior to electrophoresis. 50,000 cpm of ^{32}P-end labeled probe was electrophoresed on a 6% sequencing gel run at 3000 V for 2.5 h. Sequencing ladders were generated from the reverse, end-labeled primer of Probe 1 (Panel A) and Probe 3 (Panel B). From top of gel: Probe 1, -198 bp to +27 bp; Probe 2, +27 bp to -198 bp; Probe 3, -472 bp to -198 bp; Probe 4, -198 bp to -472 bp. F.P.: footprint

Table 5.3. TESS-predicted protein binding sites within protected footprints.

Region	Predicted transcription factors
Footprint #1 (0bp to -50bp)*	TATAA box complex
Footprint #2 (-160bp to -198bp)*	AP-1, CAC-binding protein, YY1
	LCRF-1, HOXD8
Footprint #3 (-198bp to -225bp)*	**HNF3α,** YY1, ELP, ER-alpha
Footprint #4 (-230bp to -240bp)*	**Oct-1**, Oct-2, Oct4, POU2F1
*Relative to transcription start site	

EMSA to confirm specific binding of transcription factors. Next EMSA was used to demonstrate specific binding of transcription factors to the protected sequences identified by footprint analysis (Figure 5.6, Panels A-D). The UGT2B10 biotinylated TATAA box oligomer was used as a positive control for the gel shift assay (Figure 5.6, Panel A). In the presence of HepG2 nuclear extract, the entirety of free biotinylated probe is shifted upwards into distinct DNA:protein hybrids. The shift does not occur in the presence of 200-fold molar excess unlabeled UGT2B10 promoter oligomer, indicating specificity of these interactions. In addition, unlabeled competitor with a mutated TATA sequence does not compete away the shift, further demonstrating the specificity of the DNA:protein hybrids.

The same experiments were carried out for footprint #2 (Figure 5.6, Panel B), the HNF3α binding site (Figure 5.6, Panel C) and the Octamer transcription factor binding site (Figure 5.6, Panel D). In all cases, the presence of a shift towards higher molecular weight was observed for the DNA:protein hybrid, which was competed away in the presence of excess unlabeled competitor oligonucleotide. The specificity of each interaction was demonstrated by the addition of excess unlabeled mutant competitors, which did not compete away the shift to the same degree.

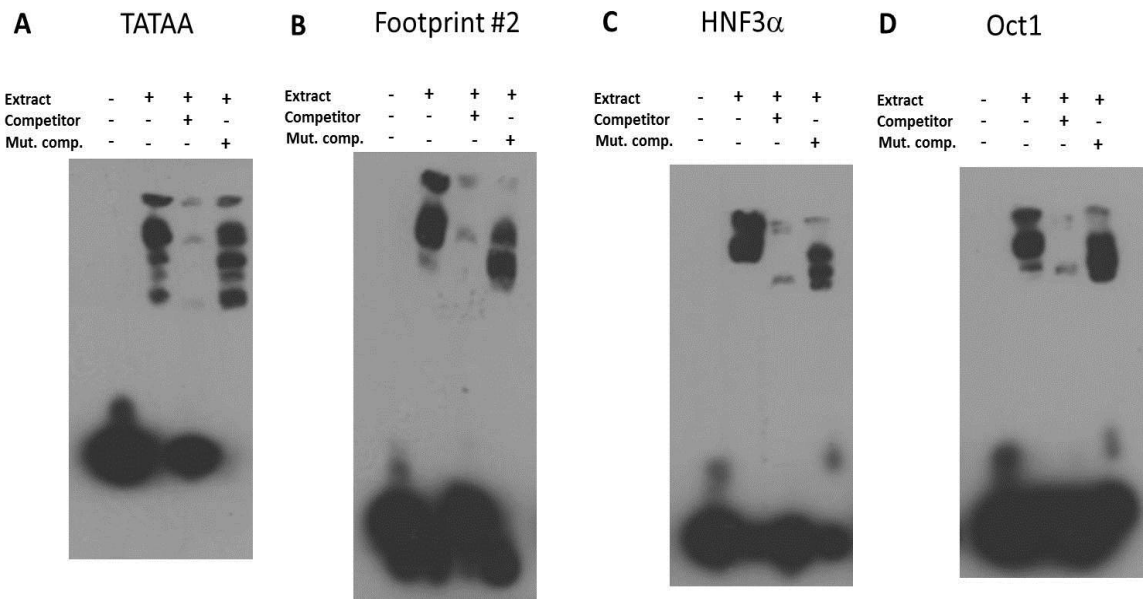

Figure 5.6. EMSA analysis of the UGT2B10 promoter. Biotin-labeled, double-stranded oligonucleotide probes (25 fmol) were treated with 25µg of HepG2 cell line NE in the presence (+) or absence (-) of excess unlabeled competitor probe (5pmol). 200-fold molar excess mutant competitor probes were included to demonstrate the specificity of the DNA:protein interaction. The three EMSA probes utilized in the binding reactions were specific for the TATAA box (Panel A), Footprint #2 (Panel B), HNF3α (Panel C), and Oct-1 (Panel D) binding sites that were identified by the DNAseI footprint analysis.

Gel supershift assay. In order to confirm that the Oct-1 protein and HNF3α protein were present in the expected complexes, antibodies directed against these proteins were added to the corresponding binding reactions (Figure 5.7). A 'supershift' can be observed in the samples that have the antibodies added, indicating that the Oct-1 and HNF3α proteins are present in their respective complexes.

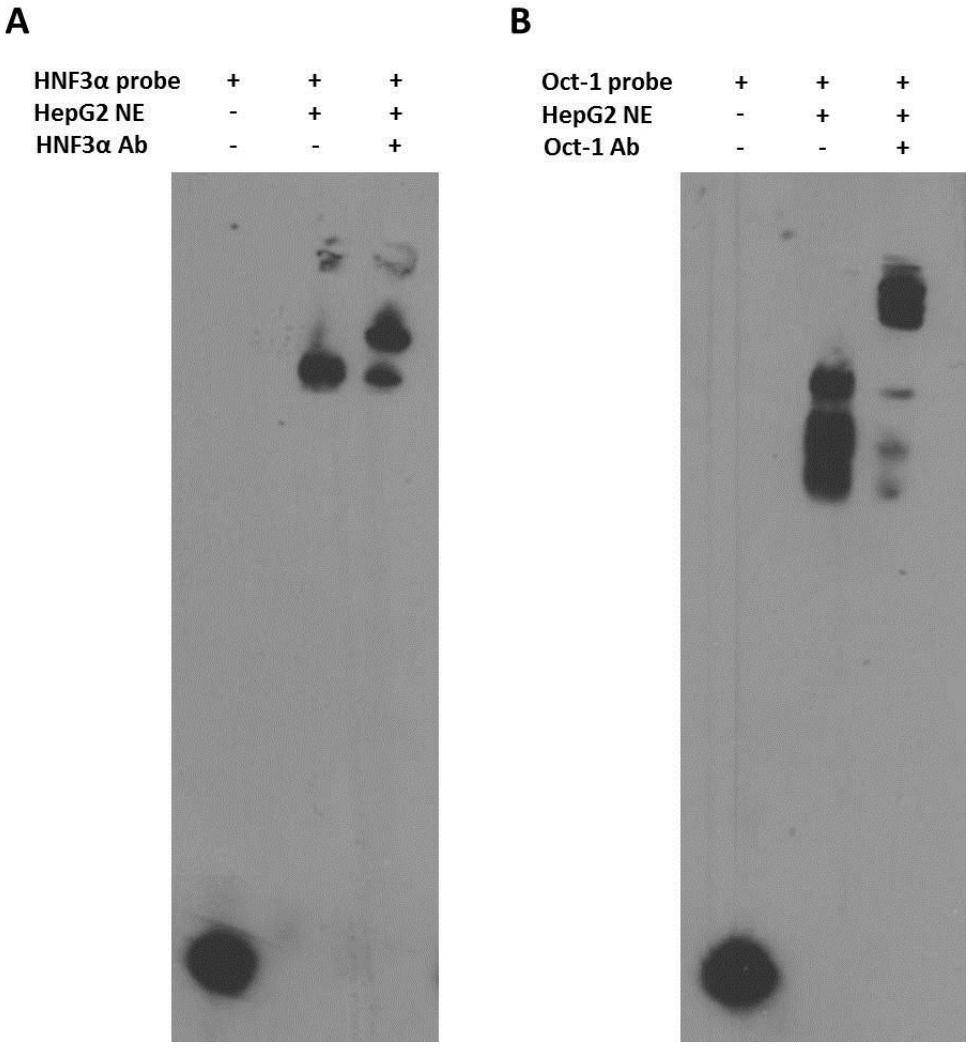

Figure 5.7. Gel supershift analysis of the HNF3α and Oct-1 binding sites. The HNF3α (Panel A) and Oct-1 (Panel B) biotin-labeled oligonucleotides (25fmol) were incubated with 25µg of HepG2 cell NE in the presence (+) or absence (-) of antibodies specific for the HNF3α and Oct-1 proteins. The shift observed in the presence of HNF3α and Oct-1 antibodies indicates that the respective proteins are present in the DNA:protein complex.

Mutational analysis of the UGT2B10 core promoter region. In order to determine the relative contribution of each protected promoter element to UGT2B10 promoter activity, single- and double-mutant constructs of the -27bp/+1948bp UGT2B10 WT promoter were created. Luciferase activity of the TATA box mutant was similar to the negative control (Figure 5.8, Panel A, *p*-value<0.0001). Single mutants corresponding to Footprint #2, the HNF3α, and Oct-1 site

resulted in decreases in promoter activity of 66%, 49%, and 61%, respectively (*p*-values= 0.0003, 0.0095, and 0.0004).

After analysis of the single-site mutants, these sites were mutated in combination to generate UGT2B10 double-mutants (Figure 5.8, Panel B). The footprint #2 (F.P.#2)/Oct-1 mutant resulted in an 82% decrease in promoter activity relative to the WT -27/+1948 bp UGT2B10 promoter construct (*p*-value = 0.0004). Similarly, the F.P.#2/HNF3α double mutant resulted in 81% decrease in activity (*p*-value = 0.003), while the Oct-1/HNF3α combination resulted in a78% decrease in activity (*p*-value = 0.008). All of the combinatorial mutants resulted in lower transcriptional activity than any of the single mutants alone. Taken together, the data from the luciferase, EMSA, and mutational analyses all indicate that UGT2B10 expression is driven by a promoter element between -170 bp and -250 bp upstream of the transcription start site and requires HNF3α and Oct-1 binding sites for maximal activity (Figure 5.9).

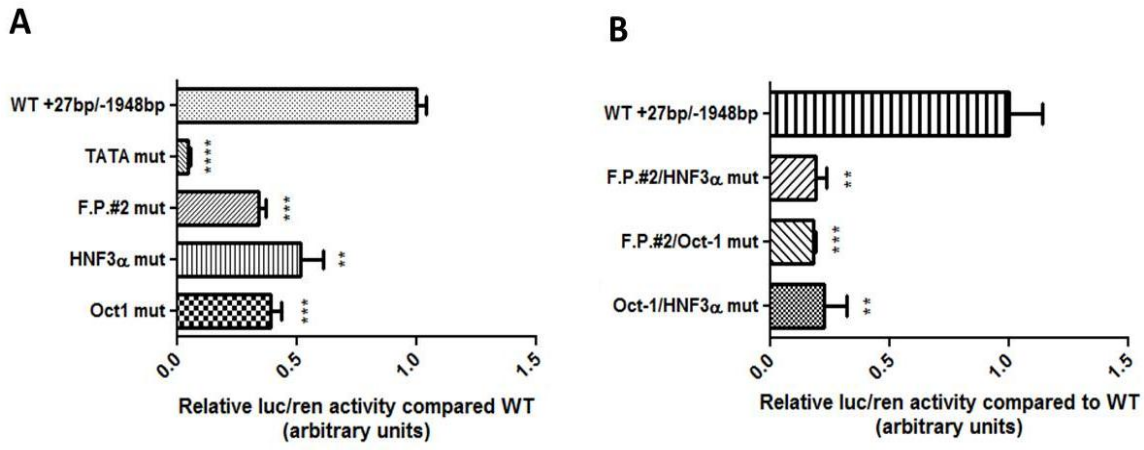

Figure 5.8. Mutational analysis of the UGT2B10 promoter protein binding sites. UGT2B10 promoter mutants generated via SDM were transfected into HepG2 cells and assayed for luciferase activity. Single mutants (Panel A) and double mutants (Panel B) were determined to have lower activity than the WT -27bp/+1948bp promoter construct. ** $p < 0.01$; *** $p<0.001$; **** $p<0.0001$; F.P.: footprint

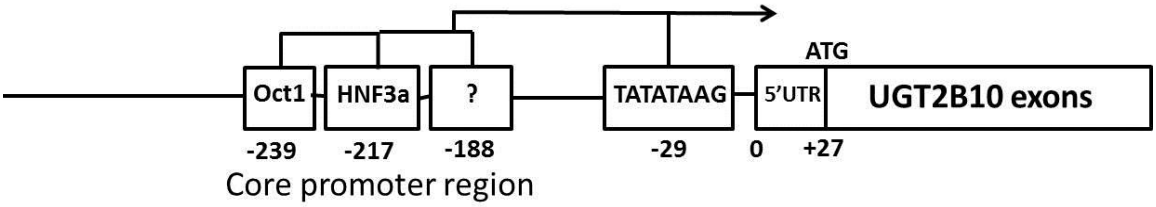

Figure 5.9. Schematic representation of the UGT2B10 promoter region. The UGT2B10 promoter contains a 27bp 5'UTR. Transcription of the UGT2B10 promoter is driven by a promoter element located between -170 bp and -250 bp upstream from the transcription start site.

5.5 Discussion

The UGT2B family of enzymes glucuronidate a diverse range of endogenous and exogenous substrates, including steroids, fatty acids, bile acids, and opiods [339]. All UGT2B genes are expressed in the liver, consistent with its role as a major site of metabolism. The highly variable pattern of expression and activity of UGT2B genes is likely due to differences in regulation at the level of transcription. In the cases studied to date, UGT2B promoter constructs containing the HNF1α binding site display elevated luciferase activity [167, 169, 170, 335, 340]. This site is conserved in sequence and location in all UGT gene promoters. In the present study, the UGT2B10 promoter construct containing only the HNF1α binding site (+27bp/-148bp) did not have significantly different promoter activity from the negative control. This indicated that the transcriptional regulation of the UGT2B10 promoter is somewhat unique relative to other UGT2B gene promoters and is consistent with previous studies by Mackenzie et al. in which overexpression of HNF1α in HepG2 cells increased the expression of most UGT2B genes but not UGT2B10 [335]. The luciferase construct data from the present study indicated that the predominant areas involved in the activation of the UGT2B10 gene promoter were located

between -121bp/-198bp and -198bp/-472bp, as there were two distinct increases in UGT2B10 promoter activity when these constructs were used to drive expression of the luciferase gene.

Footprint analysis revealed four protected regions within these promoter elements. TESS was used to predict proteins that could bind to these protected sequences. The first footprint was likely due to the TATAA box proteins. The second footprint could bind several potential transcription factors, none of which had previously been identified for UGT genes. The third footprint contained a 10 bp HNF3α consensus sequence. HNF3α is a LETF, and these are known to regulate UGT gene promoters [163, 335]. Finally, the fourth footprint was predicted to bind Octamer-related proteins, which have also been shown to regulate UGT gene transcription [172]. EMSA and gel supershift assays demonstrated that HNF3α and Oct-1 could bind to the identified sequences. Mutational analysis of the three distinct DNA-binding regions identified by footprinting analysis revealed that each site contributes to the maxmial transcriptional activity of the UGT2B10 promoter. Double mutant analysis revealed that even when the other two sites are mutated, each of the three sites still contributes a small percentage of promoter activity. This may suggest that the proteins identified do not require direct interaction with each other to perform their function. This is consistent with the fact that HNF3α has only been shown to bind monomerically to DNA to activate transcription [187].

The HNF3α protein is known to regulate hepatic genes involved in metabolism [187, 190, 191], including CYP genes [185]. For example, the HNF3α protein has been found to be important in the basal expression of CYP3A4 in HepG2 cells [185]. A recent study of the UGT2B17 gene promoter revealed that a SNP in the HNF3α binding site of the UGT2B17 promoter (-155A>G) causes a 13-fold reduction in luciferase activity in the LNCaP cell line [197]. That is the only other study that has looked at HNF3α-mediated regulation of UGT genes. In addition, expression of HNF3α is partially controlled by HNF1α, which explains why

UGT2B10 expression is correlated with the expression of other hepatic UGT2B genes (see chapter 4 of this dissertation).

Coordinated regulation between the HNF family of transcription factors and the ubiquitous Oct-1 factor has been reported previously [172, 341]. In the UGT2B7 gene promoter, Oct-1 requires the presence of HNF1α to activate gene transcription [172]. Given that the UGT2B10 gene shares greatest homology with the UGT2B7 gene, it is not surprising that both require the Oct-1 protein for maximal promoter activity. The major difference between the two genes is the fact that the HNF1α protein binding site does not appear to be important for regulating transcription of the UGT2B10 gene, while it is critically important for UGT2B7 gene promoter activity [172].

One potential explanation for the lack of activity of the UGT2B10 promoter containing only the HNF1α binding site may be due to the fact that HNF1α may require the presence of the HNF3α binding site to drive UGT2B10 promoter activity. This is supported by the fact that the UGT1A9 gene promoter requires HNF4α binding sites in order to achieve maximal promoter activity, but the HNF4α sites by themselves are inactive when the HNF1α site is mutated [180]. A good way to test this hypothesis in the future would be to overexpress HNF1α and HNF3α separately and together in a UGT2B10-null cell line, such as HEK293. There is also a possibility that the control of the UGT2B10 promoter by HNF3α instead of HNF1α is specific to HepG2 cells and not indicative of the liver as an organ. While HepG2 cells have been shown to be a good model for studying the regulation of UGT gene expression, there are some noticeable differences in expression patterns relative to normal liver tissue. Using real-time quantitative PCR, it was determined that UGT2B10 expression is highest in HepG2 cells relative to the other UGT2B genes, followed by UGT2B7 and UGT2B4 (Jones and Lazarus, unpublished). This is different from normal liver tissue where UGTs 2B4 and 2B7 are much higher than any of the other hepatic UGT2B genes [152, 153]. Nevertheless, the results from the present study strongly

point to differential regulation of the UGT2B10 promoter in liver as compared to other UGT2B genes.

The UGT2B10 promoter also contains another unknown protein binding region identified by DNaseI footprint analysis (Footprint #2). Mutational analysis of a putative binding site in this region resulted in significantly reduced activity of the luciferase reporter assay relative to the wild type sequence. The percent knockdown from mutation of this binding site was similar to that observed when the HNF3α and Oct-1 binding sites were mutated, indicating that proteins binding in this region are necessary for maximal promoter activity. Gel supershift assays to identify the proteins responsible for the DNA:protein interactions seen by EMSA were not conducted, and future studies will be needed to positively identify the transcription factors that are binding to this sequence. It is also interesting to note that while the UGT2B10 promoter is dependent on the TATAA box for activity, the TATAA box is not sufficient by itself to activate transcription. This is evidenced by the lack of activity of the first two promoter constructs in the luciferase activity assays, both of which contain the TATAA box.

The unique transcriptional regulation of the UGT2B10 promoter relative to other UGT2B genes may be important for several reasons. For example, different therapeutic and environmental agents may differentially affect the induction or repression of HNF3α-controlled promoters as compared to HNF1α-controlled promoters. In addition, genetic differences in the HNF1α gene may not affect expression of UGT2B10 to the same extent as other UGT2B genes, and vice versa for HNF3α. Additional studies are needed to determine how UGT2B10 gene expression is regulated in response to environmental, toxic, and therapeutic agents, and how this compares with other UGT2B genes. From a clinical perspective, this will be important for understanding how administration of different drugs separately or together may affect expression and activity of these detoxifying enzymes, potentially influencing their pharmacokinetic and pharmacodynamics properties [335]. In addition, the HNF3α protein is polymorphic and it has

also been shown to be regulated by nutritional factors and hormones, which may influence interindividual variability in activation of its target genes [192]. All of these factors may contribute to the interindividual differences in expression and activity of this important drug- and carcinogen-metabolizing enzyme. Future studies are needed to determine how *in vitro* transcriptional regulatory mechanisms may contribute to interindividual differences in expression seen *in vivo*.

Chapter 6

IN VITRO REGULATION OF UGT2B10 EXPRESSION VIA L-SULFORAPHANE TREATMENT

6.1 Abstract

The regulation of UGT gene promoters by a wide range of endogenous and exogenous compounds has been well-documented. Despite the importance of UGT2B10 in the detoxification of therapeutic, toxic, and carcinogenic compounds, no studies have examined xenobiotic regulation of this important enzyme. Human UGT enzymes are known to be induced by nrf2 activators, including SFN, through the nrf2/ARE pathway. In the present study, HepG2 cells were treated with SFN in order to determine the effect on UGT2B10 gene expression *in vitro*. It was determined that 10μM SFN treatment leads to a significant 52% decrease in UGT2B10 gene expression (p=0.008). siRNA-induced knockdown of the nrf2 gene failed to eliminate this effect, indicating that the repression of the UGT2B10 promoter is occurring via an nrf2-independent mechanism. Luciferase constructs were used to localize the repressor element to a region between -973 bp and -1948 bp upstream from the UGT2B10 transcription start site. Transfection of luciferase promoter constructs containing +27bp/-472bp or +27bp/-973bp of the UGT2B10 proximal promoter into HepG2 cells showed no significant difference in luciferase activity between SFN-treated and DMSO vehicle control cells. Transfection of the +27bp/-1948bp construct led to a significant 48% decrease in promoter activity in SFN-treated cells (p=0.028). This is the first study to examine xenobiotic regulation of the UGT2B10 promoter as well as the first to report SFN-mediated repression of phase II gene expression.

6.2 Introduction

UGT gene transcription is known to be regulated by xenobiotic compounds, including nrf2 activators, which act through AREs and are dependent upon the nrf2 transcription factor. Nrf2 activators induce transcription of phase two genes via their effect on the nrf2 protein. Nrf2 is normally sequestered in the cytoplasm bound to Keap1, and nrf2 activators interrupt this interaction, allowing nrf2 to translocate to the nucleus and affect gene transcription [226-228]. Glucosinolate is a compound present in cruciferous vegetables which is converted to SFN, a potent chemopreventive agent and inducer of phase II gene expression through the nrf2/ARE pathway [229, 342]. Phase II enzyme inducers such as SFN have been shown to induce transcription of the UGT1A1 gene through AREs adjacent to the XRE in the far upstream UGT1A1 promoter region (-3499bp/-3210bp) [230]. SFN also induces transcription of the UGT2B7 gene through the nrf2/ARE pathway [231]. Little is known regarding the regulation of other UGT genes by SFN.

The UGT2B10 enzyme is important in the detoxification of chemotherapeutic and carcinogenic compounds, such as OLZ, nicotine, and NNAL [95, 113, 138, 142]. It is important to know how the UGT2B10 gene is being regulated in response to xenobiotic compounds as this will influence the kinetic parameters of these compounds, but regulation of the UGT2B10 promoter via exogenous factors has not been studied. Chemopreventive agents are known to regulate the human ortholog of UGT2B10 in mice (Ugt2b34), which was found to be significantly induced by treatment with PBITC and ISC4 [343]. This led to significantly higher activity against NNAL in corresponding mouse liver microsomes (MLMs) [343].

As mentioned in Chapter 1.3, the UGT2B10 promoter contains several putative AREs conserved in sequence and location compared to other UGTs (Figure 1.11). We hypothesized that the UGT2B10 promoter is regulated by SFN treatment via these AREs. In the present study, HepG2 cells were treated with SFN to study the effect on UGT2B10 gene expression. SFN treatment led to a significant decrease in UGT2B10 mRNA expression levels. In order to investigate whether repression of UGT2B10 gene expression was occurring by an nrf2-dependent pathway, nrf2 gene expression was knocked down via siRNA prior to SFN treatment. This is the first study to demonstrate negative regulation of a phase II gene via SFN treatment.

6.3 Methods

RNA isolation and cDNA synthesis. RNA was extracted from HepG2 cells with the Qiagen Rneasy Mini kit (Valencia, CA). Samples were subjected to DNaseI digestion during extraction to prevent genomic DNA contamination. RNA concentrations were determined using a Nanodrop ND-1000 spectrophotometer, and RNA purity was assessed by absorbance ratios A260/A280 (> 1.9) and A260/A230 (>1.8). RNA integrity was determined using an Agilent 2100 Bioanalyzer with Agilent RNA 6000 Nano chips. Reverse transcription was performed using the Invitrogen Superscript First Strand cDNA synthesis kit (Carlsbad, CA) with oligo (dT) primers and 1µg of starting RNA per sample. A negative control without RNA and a negative control without enzyme were analyzed in parallel.

siRNA knockdown of nrf2 in HepG2 cells and treatment with SFN. Knockdown of nrf2 gene expression in HepG2 cells was accomplished using Lipofectamine2000 reagent and 25nM of Invitrogen pre-designed nrf2-specific siRNA species. Control cells were transfected with SCR siRNA (5'-UUCGUAACCUGUCUUUCCGUCACUG-3'). siRNA knockdown was

accomplished using the neofection procedure detailed in Chapter 3.3 for the UGT1A v2/v3 species. $6x10^6$ HepG2 cells were added to each plate such that the confluence of the cells would be ~30% after 24 h. siRNA-transfected cells were allowed to grow for 48 h prior to treatment with DMSO or 10μM SFN. Cells were treated for 48 h with SFN for 48 h prior to RNA isolation.

Quantitative PCR (qPCR). For the UGT1A1 and UGT2B10 genes, quadruplicate real-time PCR reactions were performed for each cDNA sample using a 10 μL final reaction volume containing 5 μL 2X Taqman Universal PCR Master Mix, 4.5 μL of diluted cDNA, and 0.5 μL of gene expression assay. Stock cDNA (50ng/ul) was diluted 10-fold to obtain a working dilution of 5ng/ul. A 22.5 ng RNA equivalent of cDNA was used in all reactions. Reactions were performed in 384-well plates using the ABI 7900 HT Sequence Detection System under the following conditions: 1 cycle at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15 sec and 60°C for 1 min. RQ of expression was calculated using the 2-ΔΔCt method with MT-ATP6 as the endogenous control gene. The calibrator sample was the corresponding DMSO-treated sample for each of the different treatment conditions.

PCR for nrf2 cDNA. To assess the knockdown of the nrf2 mRNA after siRNA treatment, primers specific for nrf2 were designed as follows: FWD: 5'-CGACGGAAAGAGT ATGAGCTGGA-3' and REV: 5'-TCATCTACAAACGGGAATGTCTGCG-3'. Using these primers, nrf2 cDNA was amplified for 40 cycles using *Pfu Turbo* DNA polymerase (Stratagene) and a 25 ng RNA equivalent of cDNA. PCR conditions were as follows: 95°C for 2 min, 40 cycles of 95°C for 20 seconds, 60°C for 30 seconds, and 72°C for 2 min, followed by 1 cycle of 72°C for 10 min. The resulting amplicons were run on a 1.0% agarose gel for 45 min at 100 V.

Cell culture and Transfection of Luciferase Constructs into HepG2 Cells. In order to localize the region required for SFN-mediated UGT2B10 repression, HepG2 cells were treated with 10μM SFN 24 h prior to transfection with UGT2B10 promoter constructs. HepG2 cells were maintained in DMEM supplemented with 10% FBS, 1% Pen/Strep, 1X NEAA, and L-Glutamine. $1x10^6$ cells were plated 24 h prior to transfection with Lipofectamine2000 reagent. 4 μg UGT2B10 promoter constructs were transfected into HepG2 cells and assayed for luciferase activity after 24 h. The pRL-TK vector was used as an internal control for transfection efficiency.

6.4 Results

siRNA knockdown of nrf2 gene expression and effect on UGT gene expression. siRNA directed against the nrf2 gene led to knockdown of nrf2 gene expression in HepG2 cells, as visualized via agarose gel electrophoresis, which persisted for at least 96 h after treatment (Figure 6.1, Panel A). SFN treatment led to a significant increase in UGT1A1 gene expression (79%, p=0.015), but after nrf2 knockdown, SFN was found to have no significant effect on UGT1A1 gene expression (Figure 6.1, Panel B). SFN treatment led to a significant decrease in UGT2B10 mRNA expression (52%, p=0.008). SFN-induced UGT2B10 repression was still observed even after siRNA knockdown of the nrf2 gene (64%, p=0.044) (Figure 6.1, panel C).

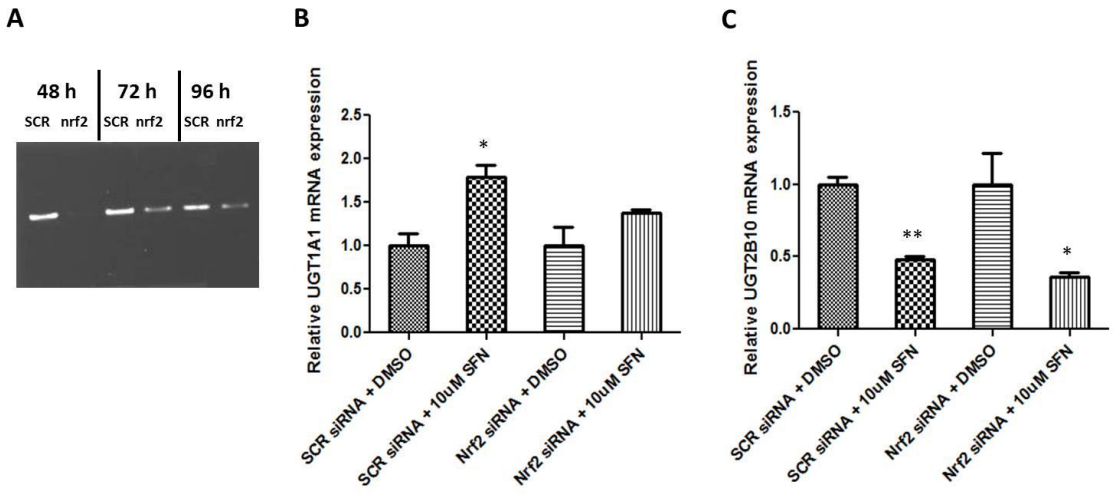

Figure 6.1. siRNA knockdown of nrf2 gene and effect on UGT gene expression levels. HepG2 cells were treated with 25nM nrf2- or SCR-siRNA for 48, 72, and 96 h prior to RNA isolation and PCR analysis (Panel A). HepG2 cells were treated with siRNA for 48 h prior to treatment with 10μM SFN for 48 h, and then assayed for UGT1A1 (Panel B) and UGT2B10 (Panel C) gene expression levels. *p<0.05, **p<0.01; SCR: scrambled.

Characterization of the SFN-responsive element within the UGT2B10 promoter. The luciferase reporter assay was used to identify a region of interest between -973bp and -1948bp in the UGT2B10 promoter that is responsible for the down-regulation of expression seen after SFN treatment (Figure 6.2). UGT2B10 promoter constructs were transfected into HepG2 cells after 24 h treatment with SFN. The +27bp/-472bp and +27bp/-973bp promoter constructs showed no significant difference in promoter activity after treatment with SFN. The +27bp/-1948bp construct showed significantly lower activity after SFN treatment relative to the corresponding DMSO control (48%, p=0.028).

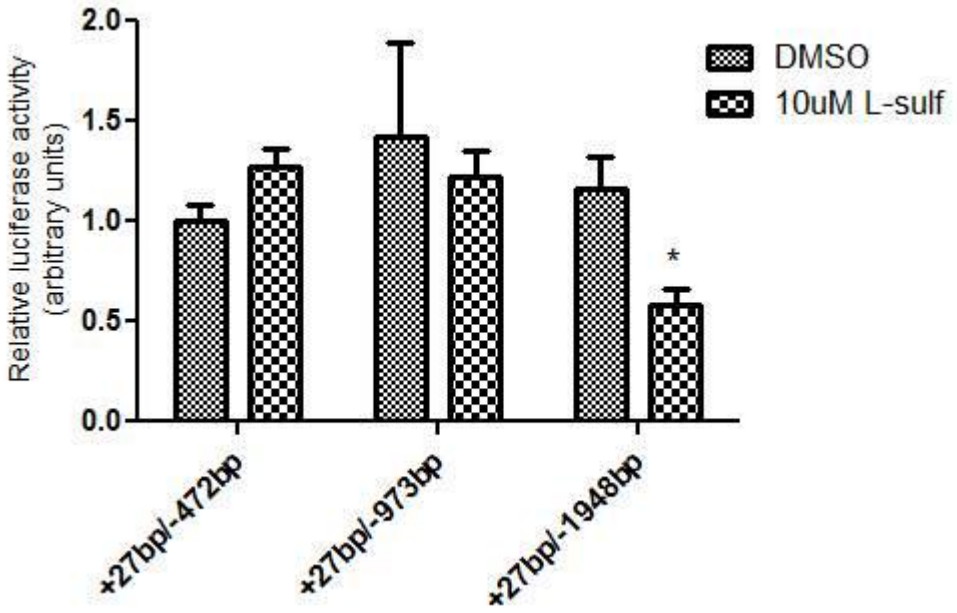

Figure 6.2. Effect of SFN treatment on UGT2B10 promoter-driven luciferase activity. HepG2 cells were treated with 10μM SFN for 24 h prior to transfection with UGT2B10 promoter constructs. Luciferase activity was determined 24 h after transfection. *p=0.0279.

6.5 Discussion

UGT genes are known to be regulated by xenobiotic compounds. An increase in UGT1A1 expression in response to SFN treatment has been reported previously, agreeing well with the current study [230]. After knockdown of the nrf2 gene, this induction effect was markedly reduced, consistent with the finding that UGT1A1 induction via SFN is nrf2-dependent [230]. Interestingly, UGT2B10 expression was found to be significantly reduced in response to SFN treatment. This was surprising considering the fact that only induction of UGT enzymes by SFN has been described previously [230, 231]. The repression effect was present even after silencing nrf2, indicating that this effect is occurring via an nrf2-independent mechanism (Figure 6.3).

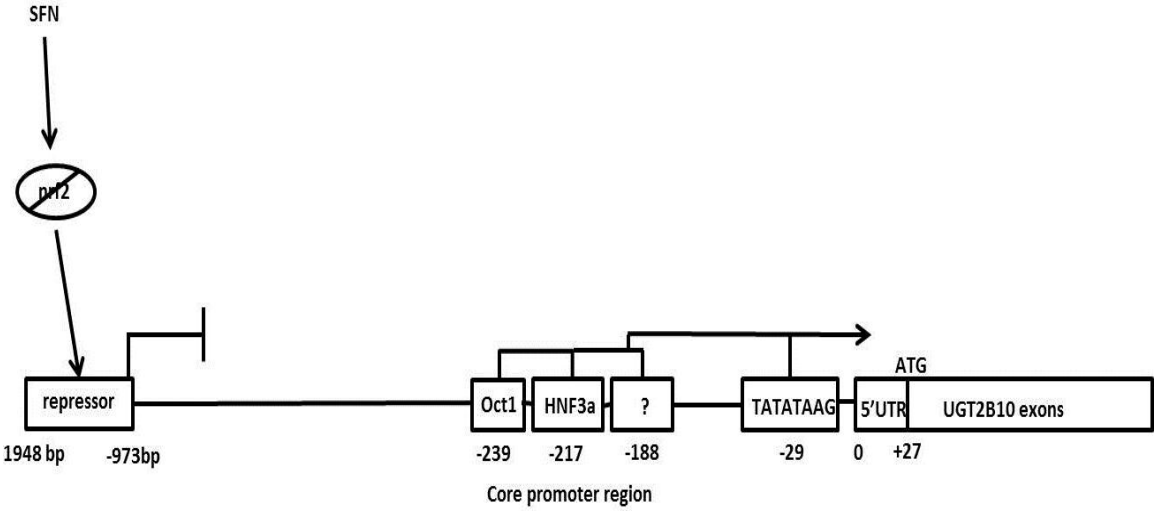

Figure 6.3. Schematic representation of the UGT2B10 core promoter region and distal repressor element. Transcription of the UGT2B10 promoter is driven by a promoter element located between -170 bp and -250 bp upstream from the transcription start site. The UGT2B10 promoter contains an nrf2-independent repressor element between -973 bp and -1948 bp that responds to SFN treatment, though the characteristics of this element have not been elucidated.

The promoter region responsible for the repression was localized to a 976 bp region between -973 bp and -1948 bp upstream of the transcription start site. Future studies will be needed to elucidate the transcription factors that are binding and initiating this repression. We hypothesized that the UGT2B10 promoter would be regulated by SFN via the nrf2/ARE pathway because of four putative AREs located at -1650 bp, -2094 bp, -2700 bp, and -3029 bp upstream of the transcription start site, identified *in silico* using TESS. These distal locations are consistent with regulatory elements in other UGT genes; the UGT1A1 promoter contains active AREs present in a region between -3499 bp and -3210 bp [230]. In addition, the UGT2B10 promoter is most similar to the UGT2B7 promoter (>90% homology), which has been shown to be induced by SFN treatment via the nrf2 pathway [231].

A SNP present in the nrf2 response element of the UGT2B7 gene promoter blocks SFN-mediated induction [231]. Allele-specific differences in response to SFN and oxidative stress

may significantly affect interindividual differences in UGT2B7-mediated glucuronidation activity [231]. There are several SNPs present in the UGT2B10 promoter between -973 bp and -1948 bp, and it is unknown whether any of these may alter the repression effects observed for the UGT2B10 gene in response to SFN treatment. For this reason, the exact nucleotide region that is regulating UGT2B10 gene expression in response to SFN treatment needs to be determined, as there may be very important pharmacogenetic differences between individuals with regard to this element.

The importance of the UGT2B10 enzyme in hepatic metabolism is becoming clearer. Until a couple of years ago, UGT2B10 was mostly thought of as an orphan enzyme with little activity against endogenous and exogenous substrates [326]. Now it is known that UGT2B10 is responsible for a large portion of the glucuronidation of NNAL, nicotine, and OLZ [95, 113, 138, 142]. A genetic knockout polymorphism at codon 67 causes reduced N-glucuronidation capacity against these substrates in HLMs from individuals with this allele [95, 113, 138, 142]. If the repression effect seen in HepG2 cells also occurs *in vivo*, then this may represent an important factor that determines interindividual differences in activity against these compounds. For example, simply eating foods high in phase II enzyme inducers such as SFN may reduce UGT2B10-mediated glucuronidation activity against both carcinogens and therapeutic agents. If true, dietary changes may change the pharmacokinetic profile of a drug like OLZ and have an impact on the drug's side effects. Similarly, reduced glucuronidation activity against nicotine from SFN consumption may alter the rate of nicotine glucuronidation and have an impact on circulating levels of this highly addictive compound. It may also impact the carcinogenicity of a compound like NNAL.

UGT2B10 gene expression was also shown to be repressed by treatment with ISC4 [343] and other chemopreventive agents (data not shown), so regulation of UGT2B10 expression via these compounds is a general phenomenon in HepG2 cells. The mouse ortholog, Ugt2b34, has

been shown to be induced in response to nrf2 activators, including the chemopreventive agent ISC4, and MLMs from ISC4-treated mice have higher glucuronidation activity against the carcinogen NNAL [343]. The induction of UGT activity against NNAL in the mouse may be one of many potential reasons why chemopreventive agents like ISC4 tend to have robust effects in the animal system and yet often have little effect in humans. This hypothesis could be tested by examining urinary metabolites of NNAL in individuals taking a SFN supplement to see if the *N*-glucuronidated product is significantly lower. In addition, the levels of urinary NNAL-*O*-gluc may be elevated by SFN since UGT2B7 is known to be induced by this compound and it is one of three UGTs predominantly responsible for NNAL-*O*-glucuronidation in human liver [120, 123, 142].

Regulation of the UGT2B10 promoter has not been studied adequately to date. This is the first study that has looked at the regulation of the UGT2B10 promoter via xenobiotic compounds. SFN significantly reduces expression of the UGT2B10 gene *in vitro*, and this may at least in part explain interindividual differences in glucuronidation activity against therapeutic agents, carcinogens, and nicotine. Future studies will be needed to determine the exact location of the SFN-responsive element and to address the clinical relevance of UGT2B10 repression via SFN.

Chapter 7

FUTURE DIRECTIONS AND FINAL CONSIDERATIONS

7.1 Conclusions and Future Directions

The UGT enzymes involved in drug metabolism are polymorphic not only within their coding regions, but also within noncoding regulatory regions [46, 344]. In addition to genetic differences, alternative splicing has been hypothesized as a potential contributor to interindividual differences in drug and carcinogen metabolism [104, 105]. This dissertation research assessed interindividual differences in UGT expression and alternative splicing, and identified factors responsible for the transcriptional regulation of the UGT2B10 gene promoter. One goal of this research was to determine the impact of the UGT1A alternative splicing event on hepatic glucuronidation activity. Another goal was to lay the groundwork for future studies that will look at interindividual differences in transcriptional regulation of the UGT2B10 promoter, which may have a significant impact on the metabolism of pharmacologic and carcinogenic compounds.

While this dissertation research has led to the partial description of UGT2B10 promoter regulation via LETFs and SFN, it is important to further characterize how these factors may contribute to interindividual differences in expression. A more complete understanding of the UGT2B10 promoter basal regulation is important in achieving this goal. While it has been shown previously that overexpression of HNF1α in HepG2 cells has no impact on UGT2B10 expression levels [193], it does not necessarily mean that it is not playing a role in regulating expression. Most human UGTs are regulated by HNF1α, and to rule it out as a potential co-regulator of UGT2B10 expression, mutational analysis of the HNF1α site in the UGT2B10 full-length promoter construct would help determine its impact on transcriptional activity. In addition, overexpression of HNF1α and HNF3α individually and in combination in HEK293 cells, which do not endogenously express UGT2B10, would help us gain a better understanding as to the relative contribution of each factor (if any for HNF1α) in regulating UGT2B10 promoter activity. Furthermore, as it is known that most genes that are controlled by HNF1α are only modestly

down-regulated in its absence, siRNA knockdown of the HNF1α gene in HepG2 cells would make it very clear what impact it has on UGT2B10 expression in liver. Analogous experiments involving siRNA knockdown of HNF3α in HepG2 cells would also clarify the hepatic regulation of UGT2B10 expression.

After determining the degree to which UGT2B10 is controlled by HNF3α, a follow-up set of experiments will involve the transfection of HNF3α variants to evaluate the impact on UGT2B10 promoter activity. Because HNF3α is important in the regulation of UGT2B10 expression, variants of this gene may affect interindividual differences in expression. Similar experiments have been conducted by Gregory et al. using HNF1α variants and the UGT2B17 promoter [167]. These studies indicated that HNF1α variants can substantially affect the activity of the UGT2B17 gene promoter, and may be a contributing factor to interindividual variability. The HNF3α gene contains coding region polymorphisms, which may lead to differences in transcriptional activation potential. For example, the HNF3α A83T allele (rs7144658) has a 37% MAF in Caucasians and yet there is no functional data on how this variant affects transcriptional activation of HNF3α target genes (Figure 7.1). By cloning the wild type and variant forms of the HNF3α gene, and studying the impact on UGT2B10 transcription, this may lend some insight into the vast interindividual differences in UGT2B10 expression seen in human liver samples. If this variant does prove to impact UGT2B10 transcription *in vitro*, the high prevalence of the HNF3α A83T allele will make it easy to study its impact on UGT2B10 expression in human liver samples. For example, hepatic UGT2B10 expression levels have been determined in a panel of liver RNA specimens, and genotyping of the rs7144658 SNP would allow correlation between HNF3α genotype and UGT2B10 expression phenotype. If it is found that common HNF3α variants have an impact on UGT2B10 expression in these samples, HNF3α genotype could ultimately be correlated with HLM activity against substrates such as NNAL, OLZ, and nicotine. It has already been reported that the knockout UGT2B10 D67Y allele has a significant impact on

HLM glucuronidation activity against these compounds. Because UGT expression differences are also known to contribute to interindividual differences in metabolism, it is also possible that HNF3α variants are causing clinically relevant differences in UGT2B10-mediated metabolism.

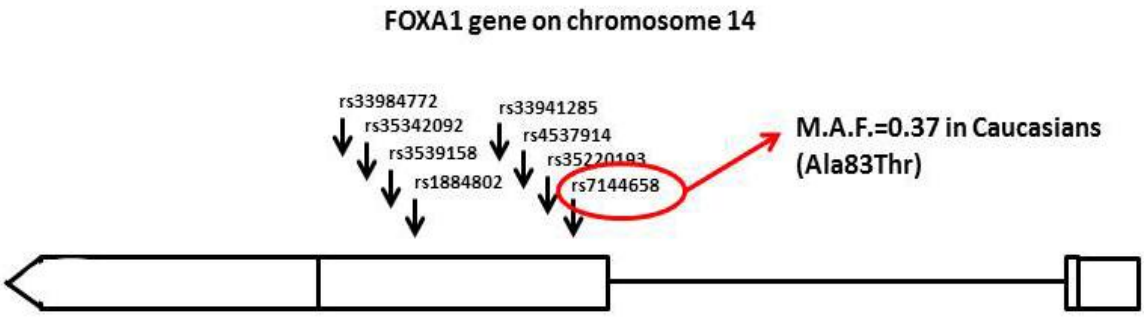

Figure 7.1. Polymorphisms in the FOXA1 gene. The FOXA1 gene encodes the HNF3α protein. Black arrows indicate the position of coding region SNPs. A SNP at codon 83 that results in an alanine to threonine substitution has not been characterized with respect to function (rs7144658, minor allele frequency = 0.37 in Caucasians). Future studies will be conducted to determine the impact of this SNP on UGT2B10 promoter activity *in vitro* and in human liver specimens. Several other SNPs have been identified in the coding region that may impact HNF3α activity as well.

Genetic differences in LETFs and their impact on glucuronidation activity is an area that has barely been touched upon in the literature [193] and represents an important gap in knowledge in the pharmacogenetics of drug metabolizing enzymes. The ultimate goal of such studies would be to prove an effect *in vivo*. If the aforementioned *in vitro* work proves positive, *in vivo* studies could be accomplished by examining HNF3α genotype in relation to urinary metabolites of NNAL and nicotine, both of which are significantly altered by the UGT2B10 D67Y polymorhphism [95, 138]. It is possible that UGT2B10 expression differences also contribute to interindividual differences in urinary metabolites, to which HNF3α variants may be a contributing factor. This could affect the carcinogenicity of NNAL by altering how quickly it is cleared, and may also impact the addictiveness of nicotine for the same reason. In addition to TSNAs and nicotine, HNF3α variants may also have an impact on OLZ metabolism, which could contribute to interindividual differences in drug response and side effects. Furthermore, the HNF3α binding site in the UGT2B10 promoter contains two SNPs identified by the 1000 Genomes' Project. While these SNPs were found to be very rare in the populations studied, they may have a strong impact on promoter activity, as has already been seen for the SNP in the

HNF3α binding site within the UGT2B17 gene promoter, which leads to 13-fold reduced activity in LNCaP cells [197]. A simple mutational analysis of the UGT2B10 promoter to study these variants would help elucidate whether they may be contributing to interindividual expression differences in certain individuals.

In addition to the regulation of UGT2B10 by LETFs, the repression of UGT2B10 expression via SFN treatment *in vitro* may be another mechanism that is contributing to interindividual differences in expression. The next step in this project will be to determine the exact nucleotide sequence that is responsible for the repression. So far, the position has been narrowed down to a region between -973bp and -1948bp upstream from the transcription start site. One way to accomplish this goal would be to use a DNaseI footprinting assay and compare the footprints seen with nuclear extracts from SFN-treated HepG2 cells to the footprints seen with DMSO-treated HepG2 cells. This would enable the identification of the precise location and sequence of the region that is responsible for the repression. No studies have described down-regulation of phase II genes via SFN treatment, and the siRNA experiments demonstrated that the repression is occurring via an nrf2-independent mechanism. Therefore, even though it is a labor intensive approach, a powerful tool such as DNaseI footprinting that enables the identification of the exact position of the binding factors is likely necessary. After identifying the sequence involved in repression, the responsible proteins could be identified first *in silico* using TESS. This could then be confirmed *in vitro* by several potential methods, such as siRNA knockdown of the putative proteins to determine if the repression is lifted. The repressor element could be further confirmed by mutation of this site in the full-length UGT2B10 promoter luciferase construct.

There are several reasons why a repressor element in the UGT2B10 promoter may be important in human metabolism. First, the repressor element itself may be polymorphic. This is true of the UGT2B7 promoter, which has an active nrf2 binding site beginning at -1248 bp

upstream from the transcription start site. A prevalent SNP in the nrf2 binding site blocks SFN-mediated induction of the UGT2B7 promoter in HepG2 cells [231], and given the polymorphic nature of the UGT genes, something similar may be occurring for the UGT2B10 gene. Even if the cognate binding site is not polymorphic, the presence of a repressor element in the UGT2B10 promoter is significant because it means that exogenous factors, such as dietary SFN intake, likely contribute to interindividual differences in expression and UGT2B10-mediated gluruonidation activity. Currently, experiments are being conducted to determine whether SFN-treated HepG2 cells display less N-glucuronidating activity against NNAL and OLZ compared to DMSO-treated HepG2 cells. This will help determine whether UGT2B10 mRNA repression is having a significant impact on protein levels and activity. In order to determine whether this repressor element is having a significant impact on glucuronidation activity *in vivo*, smokers could be administered SFN to see if it leads to a reduction in nicotine and NNAL N-glucuronides in urine. While SFN treatment will likely increase the total glucuronidation capacity of an individual because of its ability to induce several other UGT isoforms, UGT2B10 is the predominant enzyme involved in N-glucuronidation in human liver. Thus, SFN-mediated repression of this gene would be expected to result in less N-glucuronidation capacity. In addition, UGT2B10 is expressed many times higher than UGT1A4 in human liver, the only other hepatic, N-glucuronidating UGT [151].

The repression of a phase II gene promoter in humans via chemopreventive agent treatment may represent an important difference between humans and murine model systems. First of all, UGT2B10 performs primarily N-glucuronidating activity, though this activity is absent in rodents [345]. In addition, the mouse ortholog of UGT2B10, Ugt2b34, is induced by ISC4 treatment, likely contributing to increased activity against NNAL in MLMs [343]. Although the exact isoforms responsible for NNAL glucuronidation in the mouse are not known, the overall activity against an important TSNA is increased after administration of this

chemopreventive agent. The fact that a major human NNAL-glucuronidating UGT is repressed by treatment with a phase II enzyme inducer could be one of many reasons why putative chemopreventive agents are often not as effective in humans as they are in murine models. Future studies to determine the effect on UGT2B10 activity *in vitro* and *in vivo* are warranted in order to clarify this relationship.

7.2 Final Considerations

Genetic variation in the UGT gene family contributes to interindividual differences in human drug and carcinogen metabolism. The exact mechanisms behind these differences have only begun to be elucidated. In this dissertation, a novel quantitation method using an experimentally validated control gene, MT-ATP6, was used to show that there is wide interindividual variability in UGT1A and UGT2B expression in both hepatic and extrahepatic tissues. All of the UGT genes except for UGTs 1A5, 1A7, 1A8, and 1A10 were found to be expressed in the liver, but the extrahepatic tissue distribution of these genes varied widely. Within a given extrahepatic tissue, the isoforms that were expressed varied considerably between individuals. Some of this variation may be due to the heterogeneity of cell types present in the individual tissue samples, but it also shows that previous studies involving pooled RNA from multiple tissue specimens do not give a complete picture as they fail to highlight interindividual differences in UGT expression. The tissue distribution and relative abundance observed in these studies will help direct future studies based on where each isoform is likely to have physiological relevance. While the comparison of UGT isoforms between individuals is difficult, tools such as the RNA Seq Atlas have been recently developed to enable the compilation of gene expression data generated from human tissues and cell lines [346]. This tool will be a useful resource to compare gene expression profiles across a diverse range of tissues and disease states in order to

develop more complete expression profiles. It will be useful in the identification of gene expression signatures in normal and disease states.

The correlation of UGT gene expression in human liver and lung tissue suggests a common genetic regulatory network (GRN) for these genes. GRNs are important for understanding how organisms develop and how pathologies may affect these networks [347]. High-throughput analysis of gene expression data has enabled numerous models to be generated from expression patterns in a variety of physiological conditions. These models are generated using correlations between all pairs of genes to generate co-expression networks [348], improving existing models of cellular physiology. For example, GRNs are now being used to define the pathology of numerous cancers, metabolic disorders, and neurological conditions [349-353]. Recent studies have also combined SNP genotypes with gene expression data to build regulatory networks [354]. In addition, GRNs can also be used to infer clinical phenotypes based on expression patterns observed in the disease state [355]. Using these models, direct comparison of co-expression networks between normal and disease states can be used to identify network components that are differentially connected between the two. This enables the identification of gene expression markers for the disease state. Given the up regulation of UGT gene expression in lung tumors relative to adjacent normal tissues, UGT gene expression patterns may be part of a GRN that is consistently differentially regulated in lung cancer, indicating that these genes may be part of an abnormal gene expression signature in human lung cancer.

GRNs and other systems based approaches for studying disease are becoming more important because regulation of virtually all biological processes is nonlinear and epistatic. Thus, traditional linear models are only appropriate for the most simplistic processes. Identifying relevant interactions in complex GRNs is difficult because many of the genes that show differential connectivity in the disease state may be neutral and merely a consequence of other interactions. Recent studies have shown that connectivity patterns across GRNs can be used to

define different types of cancer, and these connectivity patterns may represent important candidates for therapeutic targeting [355]. For example, gene expression patterns in breast tumors identified 10 hubs of connectivity, at least three of which include proteins that are important in breast cancer invasiveness and malignancy [356-359]. Thus, GRNs can lead to the identification of biologically relevant targets, and they are being used as a novel approach to identify high priority targets underlying complex diseases and biological processes. It is interesting to speculate whether UGT genes, which are upregulated in human lung tumors, are merely part of an abnormal gene expression signature in the disease state or whether they may represent potential targets.

Furthermore, higher UGT gene expression in tumor tissue relative to adjacent normal tissue may lead to increased resistance to chemotherapeutic agents. For example, the multi-drug resistance phenotype observed for many cancers is known to be caused by overexpression and mutation of drug efflux transporters, such as P-glycoprotein and MDRs [330, 331]. This leads to increased efflux of drugs and lower cellular concentrations of cytotoxic chemotherapy agents. Overexpression of UGT genes may also contribute to this phenotype by increasing the rate of detoxification of chemotherapeutic compounds. Further research is required to determine if UGT up-regulation in lung tumor tissue is important in lung tumorigenesis and/or therapeutic resistance.

While the alternative exon 5 in the UGT1A locus was determined to have a low inclusion rate in human tissue samples, it does not rule out the possibility that alternative splicing events in other UGT genes impact interindividual differences in glucuronidation capacity. For example, the UGT2A1 gene contains an alternative splicing event in the middle of the transcript, in which the variant is missing exon 3 (Bushey and Lazarus, unpublished). This exon is highly conserved in rodents and has a high inclusion rate in human tissue samples [56]. While the *in vitro* effects of the UGT2A1 alternative splice variant is similar to the UGT1A locus splice variant, the fact

that it is conserved and expressed at levels comparable to that of the WT species *in vivo* increases the probability for it to have physiological relevance. In addition, the recent identification of alternative splicing events in the UGT2B4 and UGT2B7 genes suggests that alternative splicing is an important mechanism of regulating human glucuronidation pathways is poorly understood [106, 258]. To summarize, although the UGT1A splicing event was determined to be a minor form variant that was not correlated with hepatic glucuronidation activity, other, more highly conserved splice variants in the UGT loci are likely to have pharmacogenetic relevance. Future studies on UGT splicing variants are warranted in order to completely understand the role of alternative splicing in UGT-mediated metabolism.

The way in which transcriptional regulation of UGT genes contributes to interindividual differences in metabolism is poorly understood. The most well-studied example, the UGT1A1 TATAA box polymorphism, has clinical relevance for both endogenous (bilirubin) and exogenous (irinotecan) substrates [110, 302]. The studies presented in this dissertation determined that the UGT2B10 promoter is regulated by the HNF3α protein, which itself is polymorphic and subject to regulation via endogenous and exogenous factors [192]. This finding opens the door to an exciting new area of research, not just for UGT2B10, but UGT genes in general. Very little is known regarding polymorphic LETFs and how they contribute to the vast interindividual expression differences in UGT expression seen in human liver. It is possible that these factors play a greater role in determining glucuronidation capacity than nonsynonymous polymorphisms within UGT coding regions. To our knowledge, only one study has attempted to address this issue. In this study it was found that variants of the HNF1α gene significantly affect transcriptional activity of the UGT2B17 promoter *in vitro* [193]. There is a lot of preliminary work left to be done, but the ultimate goal will be to correlate glucuronidation phenotypes with LETF genotypes for a variety of pharmacological and carcinogenic compounds.

This dissertation research has contributed to the understanding of mechanisms underlying UGT gene expression. Studies examining the inter-individual expression patterns of UGT enzymes in human tissues, an analysis of the importance of UGT1A alternative splice variants and how they affect glucuronidation capacity in human liver, and an assessment of the transcriptional regulation of the UGT2B10 promoter by LETFs and SFN were performed in this work. Together these data have improved our knowledge of how interindividual differences in the expression of metabolizing enzymes are manifested and how these differences may potentially play a role in drug and carcinogen metabolism, personalized medicine, and cancer risk assessment.

REFERENCES

1. Ma, Q. and A.Y. Lu, *Pharmacogenetics, pharmacogenomics, and individualized medicine.* Pharmacological reviews, 2011. **63**(2): p. 437-59.
2. Nebert, D.W., G. Zhang, and E.S. Vesell, *From human genetics and genomics to pharmacogenetics and pharmacogenomics: past lessons, future directions.* Drug metabolism reviews, 2008. **40**(2): p. 187-224.
3. Vesell, E.S., *Therapeutic lessons from pharmacogenetics.* Annals of internal medicine, 1997. **126**(8): p. 653-5.
4. Simon, R., *Development and Validation of Biomarker Classifiers for Treatment Selection.* Journal of statistical planning and inference, 2008. **138**(2): p. 308-320.
5. Wang, S.J., *Utility of adaptive strategy and adaptive design for biomarker-facilitated patient selection in pharmacogenomic or pharmacogenetic clinical development program.* Journal of the Formosan Medical Association = Taiwan yi zhi, 2008. **107**(12 Suppl): p. 19-27.
6. Bell, R.G., *Metabolism of vitamin K and prothrombin synthesis: anticoagulants and the vitamin K--epoxide cycle.* Federation proceedings, 1978. **37**(12): p. 2599-604.
7. Bodin, L., et al., *A vitamin K epoxide reductase complex subunit-1 (VKORC1) mutation in a patient with vitamin K antagonist resistance.* Journal of thrombosis and haemostasis : JTH, 2005. **3**(7): p. 1533-5.
8. Rettie, A.E. and G. Tai, *The pharmocogenomics of warfarin: closing in on personalized medicine.* Molecular interventions, 2006. **6**(4): p. 223-7.
9. Rieder, M.J., et al., *Effect of VKORC1 haplotypes on transcriptional regulation and warfarin dose.* The New England journal of medicine, 2005. **352**(22): p. 2285-93.
10. Rost, S., et al., *Mutations in VKORC1 cause warfarin resistance and multiple coagulation factor deficiency type 2.* Nature, 2004. **427**(6974): p. 537-41.
11. Schwarz, U.I. and C.M. Stein, *Genetic determinants of dose and clinical outcomes in patients receiving oral anticoagulants.* Clinical pharmacology and therapeutics, 2006. **80**(1): p. 7-12.
12. Jacobsen, P., et al., *Angiotensin converting enzyme gene polymorphism and ACE inhibition in diabetic nephropathy.* Kidney international, 1998. **53**(4): p. 1002-6.
13. Lima, J.J., et al., *Impact of genetic polymorphisms of the beta2-adrenergic receptor on albuterol bronchodilator pharmacodynamics.* Clinical pharmacology and therapeutics, 1999. **65**(5): p. 519-25.
14. Williams, J.A., et al., *Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUCi/AUC) ratios.* Drug metabolism and disposition: the biological fate of chemicals, 2004. **32**(11): p. 1201-8.
15. Maitland, M.L. and R.L. Schilsky, *Clinical trials in the era of personalized oncology.* CA: a cancer journal for clinicians, 2011. **61**(6): p. 365-81.
16. Schroth, W., et al., *Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen.* JAMA : the journal of the American Medical Association, 2009. **302**(13): p. 1429-36.

17. Jin, Y., et al., *CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment.* Journal of the National Cancer Institute, 2005. **97**(1): p. 30-9.
18. Furuta, T., et al., *CYP2C19 genotype status and effect of omeprazole on intragastric pH in humans.* Clinical pharmacology and therapeutics, 1999. **65**(5): p. 552-61.
19. Lee, C.R., J.A. Goldstein, and J.A. Pieper, *Cytochrome P450 2C9 polymorphisms: a comprehensive review of the in-vitro and human data.* Pharmacogenetics, 2002. **12**(3): p. 251-63.
20. Miyazaki, M., et al., *Defective activity of recombinant cytochromes P450 3A4.2 and 3A4.16 in oxidation of midazolam, nifedipine, and testosterone.* Drug metabolism and disposition: the biological fate of chemicals, 2008. **36**(11): p. 2287-91.
21. Zhou, S.F., et al., *Clinical pharmacogenetics and potential application in personalized medicine.* Current drug metabolism, 2008. **9**(8): p. 738-84.
22. Evans, D.A., K.A. Manley, and K.V. Mc, *Genetic control of isoniazid metabolism in man.* British medical journal, 1960. **2**(5197): p. 485-91.
23. Guillemette, C., et al., *Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 and association with breast cancer among African Americans.* Cancer Res, 2000. **60**(4): p. 950-6.
24. Iyer, L., et al., *UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity.* Pharmacogenomics J, 2002. **2**(1): p. 43-7.
25. Iyer, L., et al., *Phenotype-genotype correlation of in vitro SN-38 (active metabolite of irinotecan) and bilirubin glucuronidation in human liver tissue with UGT1A1 promoter polymorphism.* Clin Pharmacol Ther, 1999. **65**(5): p. 576-82.
26. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**(7148): p. 1087-93.
27. Gold, B., et al., *Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(11): p. 4340-5.
28. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nature genetics, 2007. **39**(7): p. 870-4.
29. Stacey, S.N., et al., *Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer.* Nature genetics, 2007. **39**(7): p. 865-9.
30. Sheweita, S.A. and A.K. Tilmisany, *Cancer and phase II drug-metabolizing enzymes.* Current drug metabolism, 2003. **4**(1): p. 45-58.
31. Liu, T., et al., *Associations of CYP2A6 genotype with smoking behaviors in southern China.* Addiction, 2011. **106**(5): p. 985-94.
32. Ariyoshi, N., et al., *Genetic polymorphism of CYP2A6 gene and tobacco-induced lung cancer risk in male smokers.* Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2002. **11**(9): p. 890-4.
33. Canova, C., et al., *Genetic associations of 115 polymorphisms with cancers of the upper aerodigestive tract across 10 European countries: the ARCAGE project.* Cancer research, 2009. **69**(7): p. 2956-65.
34. Fujieda, M., et al., *Evaluation of CYP2A6 genetic polymorphisms as determinants of smoking behavior and tobacco-related lung cancer risk in male Japanese smokers.* Carcinogenesis, 2004. **25**(12): p. 2451-8.

35. Gemignani, F., et al., *Development of lung cancer before the age of 50: the role of xenobiotic metabolizing genes.* Carcinogenesis, 2007. **28**(6): p. 1287-93.
36. Miyamoto, M., et al., *CYP2A6 gene deletion reduces susceptibility to lung cancer.* Biochemical and biophysical research communications, 1999. **261**(3): p. 658-60.
37. Tan, W., et al., *Frequency of CYP2A6 gene deletion and its relation to risk of lung and esophageal cancer in the Chinese population.* International journal of cancer. Journal international du cancer, 2001. **95**(2): p. 96-101.
38. Topcu, Z., et al., *CYP2A6 gene deletion reduces oral cancer risk in betel quid chewers in Sri Lanka.* Carcinogenesis, 2002. **23**(4): p. 595-8.
39. Wang, Z., et al., *Etiological study of esophageal squamous cell carcinoma in an endemic region: a population-based case control study in Huaian, China.* BMC cancer, 2006. **6**: p. 287.
40. Rodriguez-Antona, C., et al., *Molecular genetics and epigenetics of the cytochrome P450 gene family and its relevance for cancer risk and treatment.* Human genetics, 2010. **127**(1): p. 1-17.
41. Nebert, D.W., R.A. McKinnon, and A. Puga, *Human drug-metabolizing enzyme polymorphisms: effects on risk of toxicity and cancer.* DNA and cell biology, 1996. **15**(4): p. 273-80.
42. Anttila, S., et al., *Combined effect of CYP1A1 inducibility and GSTM1 polymorphism on histological type of lung cancer.* Carcinogenesis, 1994. **15**(6): p. 1133-5.
43. Daly, A.K., et al., *Genotyping for polymorphisms in xenobiotic metabolism as a predictor of disease susceptibility.* Environmental health perspectives, 1994. **102 Suppl 9**: p. 55-61.
44. Hayashi, S., J. Watanabe, and K. Kawajiri, *High susceptibility to lung cancer analyzed in terms of combined genotypes of P450IA1 and Mu-class glutathione S-transferase genes.* Japanese journal of cancer research : Gann, 1992. **83**(8): p. 866-70.
45. Garcia-Closas, M., et al., *NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses.* Lancet, 2005. **366**(9486): p. 649-59.
46. Bock, K.W., *Functions and transcriptional regulation of adult human hepatic UDP-glucuronosyl-transferases (UGTs): mechanisms responsible for interindividual variation of UGT levels.* Biochemical pharmacology, 2010. **80**(6): p. 771-7.
47. Jancova, P., P. Anzenbacher, and E. Anzenbacherova, *Phase II drug metabolizing enzymes.* Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia, 2010. **154**(2): p. 103-16.
48. Wells, P.G., et al., *Glucuronidation and the UDP-glucuronosyltransferases in health and disease.* Drug metabolism and disposition: the biological fate of chemicals, 2004. **32**(3): p. 281-90.
49. Tukey, R.H. and C.P. Strassburg, *Human UDP-glucuronosyltransferases: metabolism, expression, and disease.* Annual review of pharmacology and toxicology, 2000. **40**: p. 581-616.
50. Ritter, J.K., et al., *A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini.* The Journal of biological chemistry, 1992. **267**(5): p. 3257-61.
51. Nagar, S. and R.P. Remmel, *Uridine diphosphoglucuronosyltransferase pharmacogenetics and cancer.* Oncogene, 2006. **25**(11): p. 1659-72.
52. Radominska-Pandya, A., S. Bratton, and J.M. Little, *A historical overview of the heterologous expression of mammalian UDP-glucuronosyltransferase isoforms over the past twenty years.* Current drug metabolism, 2005. **6**(2): p. 141-60.

53. Tukey, R.H. and C.P. Strassburg, *Genetic multiplicity of the human UDP-glucuronosyltransferases and regulation in the gastrointestinal tract.* Molecular pharmacology, 2001. **59**(3): p. 405-14.
54. Mackenzie, P.I., et al., *Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily.* Pharmacogenetics and genomics, 2005. **15**(10): p. 677-85.
55. Sneitz, N., et al., *Human UDP-glucuronosyltransferase UGT2A2: cDNA construction, expression, and functional characterization in comparison with UGT2A1 and UGT2A3.* Pharmacogenetics and genomics, 2009.
56. Bushey, R.T., et al., *Characterization of UDP-glucuronosyltransferase 2A1 (UGT2A1) variants and their potential role in tobacco carcinogenesis.* Pharmacogenetics and genomics, 2011. **21**(2): p. 55-65.
57. Turgeon, D., et al., *Isolation and characterization of the human UGT2B15 gene, localized within a cluster of UGT2B genes and pseudogenes on chromosome 4.* J Mol Biol, 2000. **295**(3): p. 489-504.
58. Riedy, M., et al., *Genomic organization of the UGT2b gene cluster on human chromosome 4q13.* Pharmacogenetics, 2000. **10**(3): p. 251-60.
59. MacKenzie, P.I., et al., *The novel UDP glycosyltransferase 3A2: cloning, catalytic properties, and tissue distribution.* Molecular pharmacology, 2011. **79**(3): p. 472-8.
60. Mackenzie, P.I., et al., *Identification of UDP glycosyltransferase 3A1 as a UDP N-acetylglucosaminyltransferase.* The Journal of biological chemistry, 2008. **283**(52): p. 36205-10.
61. Meech, R. and P.I. Mackenzie, *UGT3A: novel UDP-glycosyltransferases of the UGT superfamily.* Drug metabolism reviews, 2010. **42**(1): p. 45-54.
62. Bosio, A., et al., *The human gene CGT encoding the UDP-galactose ceramide galactosyl transferase (cerebroside synthase): cloning, characterization, and assignment to human chromosome 4, band q26.* Genomics, 1996. **34**(1): p. 69-75.
63. Coetzee, T., et al., *Molecular cloning, chromosomal mapping, and characterization of the mouse UDP-galactose:ceramide galactosyltransferase gene.* Genomics, 1996. **35**(1): p. 215-22.
64. King, C.D., et al., *UDP-glucuronosyltransferases.* Current drug metabolism, 2000. **1**(2): p. 143-61.
65. Radominska-Pandya, A., et al., *Structural and functional studies of UDP-glucuronosyltransferases.* Drug metabolism reviews, 1999. **31**(4): p. 817-99.
66. Shimomura, K., et al., *Analgesic effect of morphine glucuronides.* The Tohoku journal of experimental medicine, 1971. **105**(1): p. 45-52.
67. Barre, L., et al., *The stop transfer sequence of the human UDP-glucuronosyltransferase 1A determines localization to the endoplasmic reticulum by both static retention and retrieval mechanisms.* The FEBS journal, 2005. **272**(4): p. 1063-71.
68. Jackson, M.R., T. Nilsson, and P.A. Peterson, *Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum.* The EMBO journal, 1990. **9**(10): p. 3153-62.
69. Meech, R. and P.I. Mackenzie, *Determinants of UDP glucuronosyltransferase membrane association and residency in the endoplasmic reticulum.* Archives of biochemistry and biophysics, 1998. **356**(1): p. 77-85.
70. Brierley, C.H. and B. Burchell, *Human UDP-glucuronosyl transferases: chemical defence, jaundice and gene therapy.* BioEssays : news and reviews in molecular, cellular and developmental biology, 1993. **15**(11): p. 749-54.

71. Radominska-Pandya, A., et al., *Structure of UDP-glucuronosyltransferases in membranes.* Methods in enzymology, 2005. **400**: p. 116-47.
72. Shepherd, S.R., et al., *An investigation of the transverse topology of bilirubin UDP-glucuronosyltransferase in rat hepatic endoplasmic reticulum.* The Biochemical journal, 1989. **259**(2): p. 617-20.
73. Fremont, J.J., R.W. Wang, and C.D. King, *Coimmunoprecipitation of UDP-glucuronosyltransferase isoforms and cytochrome P450 3A4.* Molecular pharmacology, 2005. **67**(1): p. 260-2.
74. Meech, R. and P.I. Mackenzie, *UDP-glucuronosyltransferase, the role of the amino terminus in dimerization.* The Journal of biological chemistry, 1997. **272**(43): p. 26913-7.
75. Operana, T.N. and R.H. Tukey, *Oligomerization of the UDP-glucuronosyltransferase 1A proteins: homo- and heterodimerization analysis by fluorescence resonance energy transfer and co-immunoprecipitation.* The Journal of biological chemistry, 2007. **282**(7): p. 4821-9.
76. Iyanagi, T., *Molecular mechanism of phase I and phase II drug-metabolizing enzymes: implications for detoxification.* International review of cytology, 2007. **260**: p. 35-112.
77. Basu, N.K., et al., *The major chemical-detoxifying system of UDP-glucuronosyltransferases requires regulated phosphorylation supported by protein kinase C.* The Journal of biological chemistry, 2008. **283**(34): p. 23048-61.
78. Basu, N.K., L. Kole, and I.S. Owens, *Evidence for phosphorylation requirement for human bilirubin UDP-glucuronosyltransferase (UGT1A1) activity.* Biochemical and biophysical research communications, 2003. **303**(1): p. 98-104.
79. Basu, N.K., et al., *Phosphorylation of a UDP-glucuronosyltransferase regulates substrate specificity.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(18): p. 6285-90.
80. Basu, N.K., et al., *Gastrointestinally distributed UDP-glucuronosyltransferase 1A10, which metabolizes estrogens and nonsteroidal anti-inflammatory drugs, depends upon phosphorylation.* The Journal of biological chemistry, 2004. **279**(27): p. 28320-9.
81. Hauser, S.C., J.C. Ziurys, and J.L. Gollan, *A membrane transporter mediates access of uridine 5'-diphosphoglucuronic acid from the cytosol into the endoplasmic reticulum of rat hepatocytes: implications for glucuronidation reactions.* Biochimica et biophysica acta, 1988. **967**(2): p. 149-57.
82. Muraoka, M., M. Kawakita, and N. Ishida, *Molecular characterization of human UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter, a novel nucleotide sugar transporter with dual substrate specificity.* FEBS letters, 2001. **495**(1-2): p. 87-93.
83. Muraoka, M., et al., *Variety of nucleotide sugar transporters with respect to the interaction with nucleoside mono- and diphosphates.* The Journal of biological chemistry, 2007. **282**(34): p. 24615-22.
84. Cui, Y., et al., *Hepatic uptake of bilirubin and its conjugates by the human organic anion transporter SLC21A6.* The Journal of biological chemistry, 2001. **276**(13): p. 9626-30.
85. Hirohashi, T., H. Suzuki, and Y. Sugiyama, *Characterization of the transport properties of cloned rat multidrug resistance-associated protein 3 (MRP3).* The Journal of biological chemistry, 1999. **274**(21): p. 15181-5.
86. Keppler, D., I. Leier, and G. Jedlitschky, *Transport of glutathione conjugates and glucuronides by the multidrug resistance proteins MRP1 and MRP2.* Biological chemistry, 1997. **378**(8): p. 787-91.
87. Miles, K.K., et al., *An investigation of human and rat liver microsomal mycophenolic acid glucuronidation: evidence for a principal role of UGT1A enzymes and species*

differences in UGT1A specificity. Drug metabolism and disposition: the biological fate of chemicals, 2005. **33**(10): p. 1513-20.
88. Katoh, M., et al., *Expression of human phase II enzymes in chimeric mice with humanized liver.* Drug metabolism and disposition: the biological fate of chemicals, 2005. **33**(9): p. 1333-40.
89. Senekeo-Effenberger, K., et al., *Expression of the human UGT1 locus in transgenic mice by 4-chloro-6-(2,3-xylidino)-2-pyrimidinylthioacetic acid (WY-14643) and implications on drug metabolism through peroxisome proliferator-activated receptor alpha activation.* Drug metabolism and disposition: the biological fate of chemicals, 2007. **35**(3): p. 419-27.
90. Bosma, P.J., et al., *Bilirubin UDP-glucuronosyltransferase 1 is the only relevant bilirubin glucuronidating isoform in man.* J Biol Chem, 1994. **269**(27): p. 17960-4.
91. Bosma, P., J.R. Chowdhury, and P.H. Jansen, *Genetic inheritance of Gilbert's syndrome.* Lancet, 1995. **346**(8970): p. 314-5.
92. Monaghan, G., et al., *Genetic variation in bilirubin UPD-glucuronosyltransferase gene promoter and Gilbert's syndrome.* Lancet, 1996. **347**(9001): p. 578-81.
93. Beutler, E., T. Gelbart, and A. Demina, *Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism?* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 8170-4.
94. Wilson, W., 3rd, et al., *Characterization of a common deletion polymorphism of the UGT2B17 gene linked to UGT2B15.* Genomics, 2004. **84**(4): p. 707-14.
95. Chen, G., et al., *Glucuronidation of nicotine and cotinine by UGT2B10: loss of function by the UGT2B10 Codon 67 (Asp>Tyr) polymorphism.* Cancer research, 2007. **67**(19): p. 9024-9.
96. Bernard, O., et al., *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: the biological fate of chemicals, 2006. **34**(9): p. 1539-45.
97. Coffman, B.L., et al., *Human UGT2B7 catalyzes morphine glucuronidation.* Drug metabolism and disposition: the biological fate of chemicals, 1997. **25**(1): p. 1-4.
98. Innocenti, F., et al., *Epirubicin glucuronidation is catalyzed by human UDP-glucuronosyltransferase 2B7.* Drug metabolism and disposition: the biological fate of chemicals, 2001. **29**(5): p. 686-92.
99. Thibaudeau, J., et al., *Characterization of common UGT1A8, UGT1A9, and UGT2B7 variants with different capacities to inactivate mutagenic 4-hydroxylated metabolites of estradiol and estrone.* Cancer research, 2006. **66**(1): p. 125-33.
100. Wiener, D., et al., *Characterization of N-glucuronidation of the lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) in human liver: importance of UDP-glucuronosyltransferase 1A4.* Drug metabolism and disposition: the biological fate of chemicals, 2004. **32**(1): p. 72-9.
101. Freeman, J.L., et al., *Copy number variation: new insights in genome diversity.* Genome research, 2006. **16**(8): p. 949-61.
102. Menard, V., et al., *Copy-number variations (CNVs) of the human sex steroid metabolizing genes UGT2B17 and UGT2B28 and their associations with a UGT2B15 functional polymorphism.* Human mutation, 2009. **30**(9): p. 1310-9.
103. Xue, Y., et al., *Adaptive evolution of UGT2B17 copy-number variation.* American journal of human genetics, 2008. **83**(3): p. 337-46.
104. Girard, H., et al., *Genetic diversity at the UGT1 locus is amplified by a novel 3' alternative splicing mechanism leading to nine additional UGT1A proteins that act as*

regulators of glucuronidation activity. Pharmacogenet Genomics, 2007. **17**(12): p. 1077-89.
105. Levesque, E., et al., *Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus.* Hepatology, 2007. **45**(1): p. 128-38.
106. Levesque, E., et al., *Extensive splicing of transcripts encoding the bile acid-conjugating enzyme UGT2B4 modulates glucuronidation.* Pharmacogenetics and genomics, 2010. **20**(3): p. 195-210.
107. Liu, C.Y., et al., *UGT1A1*28 polymorphism predicts irinotecan-induced severe toxicities without affecting treatment outcome and survival in patients with metastatic colorectal carcinoma.* Cancer, 2008. **112**(9): p. 1932-40.
108. Paoluzzi, L., et al., *Influence of genetic variants in UGT1A1 and UGT1A9 on the in vivo glucuronidation of SN-38.* Journal of clinical pharmacology, 2004. **44**(8): p. 854-60.
109. Ando, Y., et al., *Polymorphisms of UDP-glucuronosyltransferase gene and irinotecan toxicity: a pharmacogenetic analysis.* Cancer research, 2000. **60**(24): p. 6921-6.
110. Nagar, S. and R.L. Blanchard, *Pharmacogenetics of uridine diphosphoglucuronosyltransferase (UGT) 1A family members and its role in patient response to irinotecan.* Drug metabolism reviews, 2006. **38**(3): p. 393-409.
111. Blevins-Primeau, A.S., et al., *Functional significance of UDP-glucuronosyltransferase variants in the metabolism of active tamoxifen metabolites.* Cancer research, 2009. **69**(5): p. 1892-900.
112. Ghotbi, R., et al., *Carriers of the UGT1A4 142T>G gene variant are predisposed to reduced olanzapine exposure--an impact similar to male gender or smoking in schizophrenic patients.* European journal of clinical pharmacology, 2010. **66**(5): p. 465-74.
113. Erickson-Ridout, K.K., J. Zhu, and P. Lazarus, *Olanzapine metabolism and the significance of UGT1A448V and UGT2B1067Y variants.* Pharmacogenetics and genomics, 2011. **21**(9): p. 539-51.
114. Balliet, R.M., et al., *Characterization of UGTs active against SAHA and association between SAHA glucuronidation activity phenotype with UGT genotype.* Cancer research, 2009. **69**(7): p. 2981-9.
115. Kadiev, E., et al., *Role of pharmacogenetics in variable response to drugs: focus on opioids.* Expert opinion on drug metabolism & toxicology, 2008. **4**(1): p. 77-91.
116. Lazarus, P. and D. Sun, *Potential role of UGT pharmacogenetics in cancer treatment and prevention: focus on tamoxifen and aromatase inhibitors.* Drug metabolism reviews, 2010. **42**(1): p. 182-94.
117. Bock, K.W., *Roles of UDP-glucuronosyltransferases in chemical carcinogenesis.* Critical reviews in biochemistry and molecular biology, 1991. **26**(2): p. 129-50.
118. Malfatti, M.A. and J.S. Felton, *N-glucuronidation of 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) and N_hydroxy-PhIP by specific human UDP-glucuronosyltransferases.* Carcinogenesis, 2001. **22**(7): p. 1087-93.
119. Malfatti, M.A. and J.S. Felton, *Human UDP-glucuronosyltransferase 1A1 is the primary enzyme responsible for the N-glucuronidation of N-hydroxy-PhIP in vitro.* Chemical research in toxicology, 2004. **17**(8): p. 1137-44.
120. Ren, Q., et al., *O-Glucuronidation of the lung carcinogen 4-(methylnitrosamino)-1- (3-pyridyl)-1-butanol (NNAL) by human UDP-glucuronosyltransferases 2B7 and 1A9.* Drug metabolism and disposition: the biological fate of chemicals, 2000. **28**(11): p. 1352-60.
121. Dellinger, R.W., et al., *Glucuronidation of PhIP and N-OH-PhIP by UDP-glucuronosyltransferase 1A10.* Carcinogenesis, 2007. **28**(11): p. 2412-8.

122. Dellinger, R.W., et al., *Importance of UDP-glucuronosyltransferase 1A10 (UGT1A10) in the detoxification of polycyclic aromatic hydrocarbons: decreased glucuronidative activity of the UGT1A10139Lys isoform.* Drug metabolism and disposition: the biological fate of chemicals, 2006. **34**(6): p. 943-9.
123. Lazarus, P., et al., *Genotype-phenotype correlation between the polymorphic UGT2B17 gene deletion and NNAL glucuronidation activities in human liver microsomes.* Pharmacogenetics and genomics, 2005. **15**(11): p. 769-78.
124. Wiener, D., et al., *Correlation between UDP-glucuronosyltransferase genotypes and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone glucuronidation phenotype in human liver microsomes.* Cancer research, 2004. **64**(3): p. 1190-6.
125. Zheng, Z., J.L. Fang, and P. Lazarus, *Glucuronidation: an important mechanism for detoxification of benzo[a]pyrene metabolites in aerodigestive tract tissues.* Drug Metab Dispos, 2002. **30**(4): p. 397-403.
126. Zheng, Z., et al., *Tobacco carcinogen-detoxifying enzyme UGT1A7 and its association with orolaryngeal cancer risk.* Journal of the National Cancer Institute, 2001. **93**(18): p. 1411-8.
127. Lu, P.H., et al., *Genetic polymorphisms of UGT1A7 and cancer risk: evidence from 21 case-control studies.* Cancer investigation, 2011. **29**(10): p. 645-54.
128. Guillemette, C., et al., *Structural heterogeneity at the UDP-glucuronosyltransferase 1 locus: functional consequences of three novel missense mutations in the human UGT1A7 gene*. Pharmacogenetics, 2000. **10**(7): p. 629-44.
129. Elahi, A., et al., *Detection of UGT1A10 polymorphisms and their association with orolaryngeal carcinoma risk.* Cancer, 2003. **98**(4): p. 872-80.
130. Gallagher, C.J., et al., *The UDP-glucuronosyltransferase 2B17 gene deletion polymorphism: sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer*. Cancer Epidemiol Biomarkers Prev, 2007. **16**(4): p. 823-8.
131. Gallagher, C.J., et al., *Sex differences in UDP-glucuronosyltransferase 2B17 expression and activity.* Drug metabolism and disposition: the biological fate of chemicals, 2010. **38**(12): p. 2204-9.
132. Iwagawa, M., et al., *Comparative dose-response study on the pulmonary carcinogenicity of 1,6-dinitropyrene and benzo[a]pyrene in F344 rats.* Carcinogenesis, 1989. **10**(7): p. 1285-90.
133. Huang, Y.H., et al., *Identification and functional characterization of UDP-glucuronosyltransferases UGT1A8*1, UGT1A8*2 and UGT1A8*3.* Pharmacogenetics, 2002. **12**(4): p. 287-97.
134. Fang, J.L. and P. Lazarus, *Correlation between the UDP-glucuronosyltransferase (UGT1A1) TATAA box polymorphism and carcinogen detoxification phenotype: significantly decreased glucuronidating activity against benzo(a)pyrene-7,8-dihydrodiol(-) in liver microsomes from subjects with the UGT1A1*28 variant.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(1): p. 102-9.
135. Guillemette, C., et al., *Association of genetic polymorphisms in UGT1A1 with breast cancer and plasma hormone levels.* Cancer Epidemiol Biomarkers Prev, 2001. **10**(6): p. 711-4.
136. Berg, J.Z., et al., *Nicotine metabolism in African Americans and European Americans: variation in glucuronidation by ethnicity and UGT2B10 haplotype.* The Journal of pharmacology and experimental therapeutics, 2010. **332**(1): p. 202-9.
137. Berg, J.Z., et al., *UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption.* Cancer epidemiology, biomarkers & prevention : a publication of the

American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2010. **19**(6): p. 1423-31.
138. Chen, G., et al., *Identification of a prevalent functional missense polymorphism in the UGT2B10 gene and its association with UGT2B10 inactivation against tobacco-specific nitrosamines.* Pharmacogenetics and genomics, 2008. **18**(3): p. 181-91.
139. Carmella, S.G., et al., *Analysis of N- and O-glucuronides of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) in human urine.* Chemical research in toxicology, 2002. **15**(4): p. 545-50.
140. Stepanov, I. and S.S. Hecht, *Tobacco-specific nitrosamines and their pyridine-N-glucuronides in the urine of smokers and smokeless tobacco users.* Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 2005. **14**(4): p. 885-91.
141. Upadhyaya, P., et al., *Tumorigenicity and metabolism of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol enantiomers and metabolites in the A/J mouse.* Carcinogenesis, 1999. **20**(8): p. 1577-82.
142. Chen, G., et al., *Glucuronidation of tobacco-specific nitrosamines by UGT2B10.* Drug metabolism and disposition: the biological fate of chemicals, 2008. **36**(5): p. 824-30.
143. Chen, G., et al., *Glucuronidation genotypes and nicotine metabolic phenotypes: importance of functional UGT2B10 and UGT2B17 polymorphisms.* Cancer research, 2010. **70**(19): p. 7543-52.
144. Byrd, G.D., et al., *Determination of nicotine N-1-glucuronide, a quaternary N-glucuronide conjugate, in human biological samples.* Drug metabolism and drug interactions, 2000. **16**(4): p. 281-97.
145. Callaghan, J.T., et al., *Olanzapine. Pharmacokinetic and pharmacodynamic profile.* Clinical pharmacokinetics, 1999. **37**(3): p. 177-93.
146. Citrome, L., et al., *Olanzapine plasma concentrations after treatment with 10, 20, and 40 mg/d in patients with schizophrenia: an analysis of correlations with efficacy, weight gain, and prolactin concentration.* Journal of clinical psychopharmacology, 2009. **29**(3): p. 278-83.
147. Wu, T.H., et al., *Pharmacokinetics of olanzapine in Chinese male schizophrenic patients with various smoking behaviors.* Progress in neuro-psychopharmacology & biological psychiatry, 2008. **32**(8): p. 1889-93.
148. Glubb, D.M. and F. Innocenti, *Mechanisms of genetic regulation in gene expression: examples from drug metabolizing enzymes and transporters.* Wiley interdisciplinary reviews. Systems biology and medicine, 2011. **3**(3): p. 299-313.
149. Gradhand, U. and R.B. Kim, *Pharmacogenomics of MRP transporters (ABCC1-5) and BCRP (ABCG2).* Drug metabolism reviews, 2008. **40**(2): p. 317-54.
150. Hines, R.N., et al., *Regulatory polymorphisms and their contribution to interindividual differences in the expression of enzymes influencing drug and toxicant disposition.* Drug metabolism reviews, 2008. **40**(2): p. 263-301.
151. Ohno, S. and S. Nakajin, *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: the biological fate of chemicals, 2009. **37**(1): p. 32-40.
152. Izukawa, T., et al., *Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers.* Drug metabolism and disposition: the biological fate of chemicals, 2009. **37**(8): p. 1759-68.

153. Ohno, S. and S. Nakajin, *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 Metab Dispos, 2009. **37**(1): p. 32-40.
154. Grove, A.D., et al., *Identification of a rat oltipraz-inducible UDP-glucuronosyltransferase (UGT1A7) with activity towards benzo(a)pyrene-7,8-dihydrodiol.* The Journal of biological chemistry, 1997. **272**(3): p. 1621-7.
155. Mojarrabi, B. and P.I. Mackenzie, *Characterization of two UDP glucuronosyltransferases that are predominantly expressed in human colon.* Biochemical and biophysical research communications, 1998. **247**(3): p. 704-9.
156. Watkins, P.B., *The barrier function of CYP3A4 and P-glycoprotein in the small bowel.* Advanced drug delivery reviews, 1997. **27**(2-3): p. 161-170.
157. Kim, P.M. and P.G. Wells, *Genoprotection by UDP-glucuronosyltransferases in peroxidase-dependent, reactive oxygen species-mediated micronucleus initiation by the carcinogens 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and benzo[a]pyrene.* Cancer Res, 1996. **56**(7): p. 1526-32.
158. Vienneau, D.S., U. DeBoni, and P.G. Wells, *Potential genoprotective role for UDP-glucuronosyltransferases in chemical carcinogenesis: initiation of micronuclei by benzo(a)pyrene and benzo(e)pyrene in UDP-glucuronosyltransferase-deficient cultured rat skin fibroblasts.* Cancer Res, 1995. **55**(5): p. 1045-51.
159. Starlard-Davenport, A., B. Lyn-Cook, and A. Radominska-Pandya, *Identification of UDP-glucuronosyltransferase 1A10 in non-malignant and malignant human breast tissues.* Steroids, 2008. **73**(6): p. 611-20.
160. Belanger, A., et al., *Levels of eighteen non-conjugated and conjugated steroids in human breast cyst fluid: relationships with cyst type.* Eur J Cancer, 1990. **26**(3): p. 277-81.
161. Belanger, A., et al., *Determination of nonconjugated and conjugated steroid levels in plasma and prostate after separation on C-18 columns.* Ann N Y Acad Sci, 1990. **595**: p. 251-9.
162. Prevost, J., et al., *Conjugated and unconjugated C-21, C-19 and C-18 steroid concentrations in human follicular fluid from hyperstimulated follicles.* Gynecol Endocrinol, 1987. **1**(4): p. 331-8.
163. Gardner-Stephen, D.A. and P.I. Mackenzie, *Liver-enriched transcription factors and their role in regulating UDP glucuronosyltransferase gene expression.* Current drug metabolism, 2008. **9**(5): p. 439-52.
164. Strassburg, C.P., et al., *Polymorphic gene regulation and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine.* The Journal of biological chemistry, 2000. **275**(46): p. 36164-71.
165. Auyeung, D.J., F.K. Kessler, and J.K. Ritter, *Differential regulation of alternate UDP-glucuronosyltransferase 1A6 gene promoters by hepatic nuclear factor-1.* Toxicology and applied pharmacology, 2003. **191**(2): p. 156-66.
166. Bernard, P., et al., *Activation of the mouse TATA-less and human TATA-containing UDP-glucuronosyltransferase 1A1 promoters by hepatocyte nuclear factor 1.* Molecular pharmacology, 1999. **56**(3): p. 526-36.
167. Gregory, P.A., A.J. Hansen, and P.I. Mackenzie, *Tissue specific differences in the regulation of the UDP glucuronosyltransferase 2B17 gene promoter.* Pharmacogenetics, 2000. **10**(9): p. 809-20.
168. Gregory, P.A., et al., *Regulation of UDP glucuronosyltransferases in the gastrointestinal tract.* Toxicology and applied pharmacology, 2004. **199**(3): p. 354-63.

169. Gregory, P.A., et al., *Coordinate regulation of the human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 genes by hepatocyte nuclear factor 1alpha and the caudal-related homeodomain protein 2.* Molecular pharmacology, 2004. **65**(4): p. 953-63.
170. Gregory, P.A. and P.I. Mackenzie, *The homeodomain Pbx2-Prep1 complex modulates hepatocyte nuclear factor 1alpha-mediated activation of the UDP-glucuronosyltransferase 2B17 gene.* Molecular pharmacology, 2002. **62**(1): p. 154-61.
171. Hansen, A.J., et al., *HNF1 alpha activates the rat UDP glucuronosyltransferase UGT2B1 gene promoter.* DNA and cell biology, 1997. **16**(2): p. 207-14.
172. Ishii, Y., A.J. Hansen, and P.I. Mackenzie, *Octamer transcription factor-1 enhances hepatic nuclear factor-1alpha-mediated activation of the human UDP glucuronosyltransferase 2B7 promoter.* Molecular pharmacology, 2000. **57**(5): p. 940-7.
173. Johnson, P.F., *Transcriptional activators in hepatocytes.* Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research, 1990. **1**(1): p. 47-52.
174. Cha, J.Y., et al., *HNF1 and/or HNF3 may contribute to the tissue specific expression of glucokinase gene.* Exp Mol Med, 2001. **33**(2): p. 59-63.
175. Gregory, P.A., et al., *Regulation of UDP glucuronosyltransferases in the gastrointestinal tract.* Toxicol Appl Pharmacol, 2004. **199**(3): p. 354-63.
176. Turgeon, D., et al., *Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members.* Endocrinology, 2001. **142**(2): p. 778-87.
177. Gregory, P.A. and P.I. Mackenzie, *The homeodomain Pbx2-Prep1 complex modulates hepatocyte nuclear factor 1alpha-mediated activation of the UDP-glucuronosyltransferase 2B17 gene.* Mol Pharmacol, 2002. **62**(1): p. 154-61.
178. Ishii, Y., A.J. Hansen, and P.I. Mackenzie, *Octamer transcription factor-1 enhances hepatic nuclear factor-1alpha-mediated activation of the human UDP glucuronosyltransferase 2B7 promoter.* Mol Pharmacol, 2000. **57**(5): p. 940-7.
179. Pontoglio, M., et al., *Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome.* Cell, 1996. **84**(4): p. 575-85.
180. Gardner-Stephen, D.A. and P.I. Mackenzie, *Hepatocyte nuclear factor1 transcription factors are essential for the UDP-glucuronosyltransferase 1A9 promoter response to hepatocyte nuclear factor 4alpha.* Pharmacogenetics and genomics, 2007. **17**(1): p. 25-36.
181. Aueviriyavit, S., et al., *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 and pharmacokinetics, 2007. **22**(5): p. 391-8.
182. Lu, H., F.J. Gonzalez, and C. Klaassen, *Alterations in hepatic mRNA expression of phase II enzymes and xenobiotic transporters after targeted disruption of hepatocyte nuclear factor 4 alpha.* Toxicological sciences : an official journal of the Society of Toxicology, 2010. **118**(2): p. 380-90.
183. Kamiyama, Y., et al., *Role of human hepatocyte nuclear factor 4alpha in the expression of drug-metabolizing enzymes and transporters in human hepatocytes assessed by use of small interfering RNA.* Drug metabolism and pharmacokinetics, 2007. **22**(4): p. 287-98.
184. Barbier, O., et al., *Hepatic expression of the UGT1A9 gene is governed by hepatocyte nuclear factor 4alpha.* Molecular pharmacology, 2005. **67**(1): p. 241-9.
185. Bombail, V., et al., *Role of Sp1, C/EBP alpha, HNF3, and PXR in the basal- and xenobiotic-mediated regulation of the CYP3A4 gene.* Drug metabolism and disposition: the biological fate of chemicals, 2004. **32**(5): p. 525-35.

186. Friedman, J.R. and K.H. Kaestner, *The Foxa family of transcription factors in development and metabolism.* Cellular and molecular life sciences : CMLS, 2006. **63**(19-20): p. 2317-28.
187. Kaestner, K.H., *The hepatocyte nuclear factor 3 (HNF3 or FOXA) family in metabolism.* Trends in endocrinology and metabolism: TEM, 2000. **11**(7): p. 281-5.
188. Kaestner, K.H., W. Knochel, and D.E. Martinez, *Unified nomenclature for the winged helix/forkhead transcription factors.* Genes & development, 2000. **14**(2): p. 142-6.
189. Peterson, R.S., et al., *Hepatocyte nuclear factor-3 alpha promoter regulation involves recognition by cell-specific factors, thyroid transcription factor-1, and autoactivation.* Cell growth & differentiation : the molecular biology journal of the American Association for Cancer Research, 1997. **8**(1): p. 69-82.
190. Tan, Y., et al., *Adenovirus-mediated increase in HNF-3beta or HNF-3alpha shows differences in levels of liver glycogen and gene expression.* Hepatology, 2002. **35**(1): p. 30-9.
191. Kaestner, K.H., et al., *Inactivation of the winged helix transcription factor HNF3alpha affects glucose homeostasis and islet glucagon gene expression in vivo.* Genes & development, 1999. **13**(4): p. 495-504.
192. Imae, M., et al., *Gene expression of the three members of hepatocyte nuclear factor-3 is differentially regulated by nutritional and hormonal factors.* The Journal of endocrinology, 2000. **167**(1): p. R1-5.
193. Mackenzie, P.I., et al., *Polymorphic variations in the expression of the chemical detoxifying UDP glucuronosyltransferases.* Toxicology and applied pharmacology, 2005. **207**(2 Suppl): p. 77-83.
194. Ellard, S., *Hepatocyte nuclear factor 1 alpha (HNF-1 alpha) mutations in maturity-onset diabetes of the young.* Human mutation, 2000. **16**(5): p. 377-85.
195. Acuna, G., et al., *Pharmacogenetic analysis of adverse drug effect reveals genetic variant for susceptibility to liver toxicity.* The pharmacogenomics journal, 2002. **2**(5): p. 327-34.
196. Girard, H., et al., *Identification of common polymorphisms in the promoter of the UGT1A9 gene: evidence that UGT1A9 protein and activity levels are strongly genetically controlled in the liver.* Pharmacogenetics, 2004. **14**(8): p. 501-15.
197. Hu, D.G., et al., *A novel polymorphism in a forkhead box A1 (FOXA1) binding site of the human UDP glucuronosyltransferase 2B17 gene modulates promoter activity and is associated with altered levels of circulating androstane-3alpha,17beta-diol glucuronide.* Molecular pharmacology, 2010. **78**(4): p. 714-22.
198. Duguay, Y., et al., *A novel functional polymorphism in the uridine diphosphate-glucuronosyltransferase 2B7 promoter with significant impact on promoter activity.* Clinical pharmacology and therapeutics, 2004. **75**(3): p. 223-33.
199. Hirota, T., et al., *Sequence variability and candidate gene analysis in two cancer patients with complex clinical outcomes during morphine therapy.* Drug metabolism and disposition: the biological fate of chemicals, 2003. **31**(5): p. 677-80.
200. Sawyer, M.B., et al., *A pharmacogenetic study of uridine diphosphate-glucuronosyltransferase 2B7 in patients receiving morphine.* Clinical pharmacology and therapeutics, 2003. **73**(6): p. 566-74.
201. Yamanaka, H., et al., *A novel polymorphism in the promoter region of human UGT1A9 gene (UGT1A9*22) and its effects on the transcriptional activity.* Pharmacogenetics, 2004. **14**(5): p. 329-32.
202. Gregory, P.A., et al., *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, 2004. **65**(4): p. 953-63.

203. Barbier, O., et al., *Peroxisome proliferator-activated receptor alpha induces hepatic expression of the human bile acid glucuronidating UDP-glucuronosyltransferase 2B4 enzyme.* The Journal of biological chemistry, 2003. **278**(35): p. 32852-60.
204. Bock, K.W. and C. Kohle, *UDP-glucuronosyltransferase 1A6: structural, functional, and regulatory aspects.* Methods in enzymology, 2005. **400**: p. 57-75.
205. Gardner-Stephen, D., et al., *Human PXR variants and their differential effects on the regulation of human UDP-glucuronosyltransferase gene expression.* Drug metabolism and disposition: the biological fate of chemicals, 2004. **32**(3): p. 340-7.
206. Lankisch, T.O., et al., *Aryl hydrocarbon receptor-mediated regulation of the human estrogen and bile acid UDP-glucuronosyltransferase 1A3 gene.* Archives of toxicology, 2008. **82**(9): p. 573-82.
207. Lee, J.H., J. Zhou, and W. Xie, *PXR and LXR in hepatic steatosis: a new dog and an old dog with new tricks.* Molecular pharmaceutics, 2008. **5**(1): p. 60-6.
208. Lu, Y., et al., *Lithocholic acid decreases expression of UGT2B7 in Caco-2 cells: a potential role for a negative farnesoid X receptor response element.* Drug metabolism and disposition: the biological fate of chemicals, 2005. **33**(7): p. 937-46.
209. Munzel, P.A., et al., *Contribution of the Ah receptor to the phenolic antioxidant-mediated expression of human and rat UDP-glucuronosyltransferase UGT1A6 in Caco-2 and rat hepatoma 5L cells.* Biochemical pharmacology, 2003. **66**(5): p. 841-7.
210. Rae, J.M., et al., *Rifampin is a selective, pleiotropic inducer of drug metabolism genes in human hepatocytes: studies with cDNA and oligonucleotide expression arrays.* The Journal of pharmacology and experimental therapeutics, 2001. **299**(3): p. 849-57.
211. Sugatani, J., et al., *Transcriptional regulation of human UGT1A1 gene expression: activated glucocorticoid receptor enhances constitutive androstane receptor/pregnane X receptor-mediated UDP-glucuronosyltransferase 1A1 regulation with glucocorticoid receptor-interacting protein 1.* Molecular pharmacology, 2005. **67**(3): p. 845-55.
212. Tobin, J.F. and L.P. Freedman, *Nuclear receptors as drug targets in metabolic diseases: new approaches to therapy.* Trends in endocrinology and metabolism: TEM, 2006. **17**(7): p. 284-90.
213. Verreault, M., et al., *The liver X-receptor alpha controls hepatic expression of the human bile acid-glucuronidating UGT1A3 enzyme in human cells and transgenic mice.* Hepatology, 2006. **44**(2): p. 368-78.
214. Yueh, M.F., et al., *Involvement of the xenobiotic response element (XRE) in Ah receptor-mediated induction of human UDP-glucuronosyltransferase 1A1*. The Journal of biological chemistry, 2003. **278**(17): p. 15001-6.
215. Balliet, R.M., et al., *UDP-glucuronosyltransferase 1A10: activity against the tobacco-specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol, and a potential role for a novel UGT1A10 promoter deletion polymorphism in cancer susceptibility.* Drug metabolism and disposition: the biological fate of chemicals, 2010. **38**(3): p. 484-90.
216. Galijatovic, A., et al., *Induction of UDP-glucuronosyltransferase UGT1A1 by the flavonoid chrysin in Caco-2 cells--potential role in carcinogen bioinactivation.* Pharmaceutical research, 2001. **18**(3): p. 374-9.
217. Iwuchukwu, O.F., R.J. Tallarida, and S. Nagar, *Resveratrol in combination with other dietary polyphenols concomitantly enhances antiproliferation and UGT1A1 induction in Caco-2 cells.* Life sciences, 2011. **88**(23-24): p. 1047-54.
218. Kanou, M., et al., *Stimulation of transcriptional expression of human UDP-glucuronosyltransferase 1A1 by dexamethasone.* Molecular biology reports, 2004. **31**(3): p. 151-8.

219. Sugatani, J., et al., *The induction of human UDP-glucuronosyltransferase 1A1 mediated through a distal enhancer module by flavonoids and xenobiotics.* Biochemical pharmacology, 2004. **67**(5): p. 989-1000.
220. Svehlikova, V., et al., *Interactions between sulforaphane and apigenin in the induction of UGT1A1 and GSTA1 in CaCo-2 cells.* Carcinogenesis, 2004. **25**(9): p. 1629-37.
221. Bonzo, J.A., A. Belanger, and R.H. Tukey, *The role of chrysin and the ah receptor in induction of the human UGT1A1 gene in vitro and in transgenic UGT1 mice.* Hepatology, 2007. **45**(2): p. 349-60.
222. Chouinard, S., et al., *Isoform-specific regulation of uridine diphosphate-glucuronosyltransferase 2B enzymes in the human prostate: differential consequences for androgen and bioactive lipid inactivation.* Endocrinology, 2006. **147**(11): p. 5431-42.
223. Guillemette, C., D.W. Hum, and A. Belanger, *Regulation of steroid glucuronosyltransferase activities and transcripts by androgen in the human prostatic cancer LNCaP cell line.* Endocrinology, 1996. **137**(7): p. 2872-9.
224. Guillemette, C., et al., *Differential regulation of two uridine diphospho-glucuronosyltransferases, UGT2B15 and UGT2B17, in human prostate LNCaP cells.* Endocrinology, 1997. **138**(7): p. 2998-3005.
225. Harrington, W.R., S. Sengupta, and B.S. Katzenellenbogen, *Estrogen regulation of the glucuronidation enzyme UGT2B15 in estrogen receptor-positive breast cancer cells.* Endocrinology, 2006. **147**(8): p. 3843-50.
226. Zhang, D.D., *Mechanistic studies of the Nrf2-Keap1 signaling pathway.* Drug metabolism reviews, 2006. **38**(4): p. 769-89.
227. Zhang, D.D., et al., *Ubiquitination of Keap1, a BTB-Kelch substrate adaptor protein for Cul3, targets Keap1 for degradation by a proteasome-independent pathway.* The Journal of biological chemistry, 2005. **280**(34): p. 30091-9.
228. Zhang, D.D. and M. Hannink, *Distinct cysteine residues in Keap1 are required for Keap1-dependent ubiquitination of Nrf2 and for stabilization of Nrf2 by chemopreventive agents and oxidative stress.* Molecular and cellular biology, 2003. **23**(22): p. 8137-51.
229. McWalter, G.K., et al., *Transcription factor Nrf2 is essential for induction of NAD(P)H:quinone oxidoreductase 1, glutathione S-transferases, and glutamate cysteine ligase by broccoli seeds and isothiocyanates.* The Journal of nutrition, 2004. **134**(12 Suppl): p. 3499S-3506S.
230. Yueh, M.F. and R.H. Tukey, *Nrf2-Keap1 signaling pathway regulates human UGT1A1 expression in vitro and in transgenic UGT1 mice.* The Journal of biological chemistry, 2007. **282**(12): p. 8749-58.
231. Nakamura, A., et al., *Genetic polymorphisms in the 5'-flanking region of human UDP-glucuronosyltransferase 2B7 affect the Nrf2-dependent transcriptional regulation.* Pharmacogenetics and genomics, 2008. **18**(8): p. 709-20.
232. Kim, E., A. Goren, and G. Ast, *Alternative splicing: current perspectives.* BioEssays : news and reviews in molecular, cellular and developmental biology, 2008. **30**(1): p. 38-47.
233. Zhou, Z., et al., *Comprehensive proteomic analysis of the human spliceosome.* Nature, 2002. **419**(6903): p. 182-5.
234. Graveley, B.R., *Alternative splicing: increasing diversity in the proteomic world.* Trends in genetics : TIG, 2001. **17**(2): p. 100-7.
235. Lim, L.P. and C.B. Burge, *A computational analysis of sequence features involved in recognition of short introns.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(20): p. 11193-8.

236. Sanford, J.R., J. Ellis, and J.F. Caceres, *Multiple roles of arginine/serine-rich splicing factors in RNA processing.* Biochemical Society transactions, 2005. **33**(Pt 3): p. 443-6.
237. Chabot, B., *Directing alternative splicing: cast and scenarios.* Trends in genetics : TIG, 1996. **12**(11): p. 472-8.
238. Maniatis, T. and B. Tasic, *Alternative pre-mRNA splicing and proteome expansion in metazoans.* Nature, 2002. **418**(6894): p. 236-43.
239. Mersch, B., et al., *SERpredict: detection of tissue- or tumor-specific isoforms generated through exonization of transposable elements.* BMC genetics, 2007. **8**: p. 78.
240. Sela, N., et al., *Comparative analysis of transposed element insertion within human and mouse genomes reveals Alu's unique role in shaping the human transcriptome.* Genome biology, 2007. **8**(6): p. R127.
241. Walters, R.D., J.F. Kugel, and J.A. Goodrich, *InvAluable junk: the cellular impact and function of Alu and B2 RNAs.* IUBMB life, 2009. **61**(8): p. 831-7.
242. Modrek, B. and C.J. Lee, *Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss.* Nature genetics, 2003. **34**(2): p. 177-80.
243. Krull, M., J. Brosius, and J. Schmitz, *Alu-SINE exonization: en route to protein-coding function.* Molecular biology and evolution, 2005. **22**(8): p. 1702-11.
244. Ferlini, A., et al., *A novel Alu-like element rearranged in the dystrophin gene causes a splicing mutation in a family with X-linked dilated cardiomyopathy.* American journal of human genetics, 1998. **63**(2): p. 436-46.
245. Knebelmann, B., et al., *Splice-mediated insertion of an Alu sequence in the COL4A3 mRNA causing autosomal recessive Alport syndrome.* Human molecular genetics, 1995. **4**(4): p. 675-9.
246. Wallace, M.R., et al., *A de novo Alu insertion results in neurofibromatosis type 1.* Nature, 1991. **353**(6347): p. 864-6.
247. Schwartz, S., et al., *Alu exonization events reveal features required for precise recognition of exons by the splicing machinery.* PLoS computational biology, 2009. **5**(3): p. e1000300.
248. Bellemare, J., et al., *Alternatively Spliced Products of the UGT1A Gene Interact with the Enzymatically Active Proteins to Inhibit Glucuronosyltransferase Activity In Vitro.* Drug Metab Dispos. **38**(10): p. 1785-9.
249. Bellemare, J., et al., *Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions.* J Biol Chem. **285**(6): p. 3600-7.
250. Bellemare, J., et al., *Alternative-splicing forms of the major phase II conjugating UGT1A gene negatively regulate glucuronidation in human carcinoma cell lines.* Pharmacogenomics J, 2009.
251. Benoit-Biancamano, M.O., et al., *Deferiprone glucuronidation by human tissues and recombinant UDP glucuronosyltransferase 1A6: an in vitro investigation of genetic and splice variants.* Drug Metab Dispos, 2009. **37**(2): p. 322-9.
252. Bellemare, J., et al., *Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions.* The Journal of biological chemistry, 2010. **285**(6): p. 3600-7.
253. Bellemare, J., et al., *Alternative-splicing forms of the major phase II conjugating UGT1A gene negatively regulate glucuronidation in human carcinoma cell lines.* The pharmacogenomics journal, 2010. **10**(5): p. 431-41.
254. Kalnina, Z., et al., *Alterations of pre-mRNA splicing in cancer.* Genes, chromosomes & cancer, 2005. **42**(4): p. 342-57.

255. Wang, Z., et al., *Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer.* Cancer research, 2003. **63**(3): p. 655-7.
256. Xu, Q. and C. Lee, *Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences.* Nucleic acids research, 2003. **31**(19): p. 5635-43.
257. Stamm, S., *Signals and their transduction pathways regulating alternative splicing: a new dimension of the human genome.* Human molecular genetics, 2002. **11**(20): p. 2409-16.
258. Menard, V., et al., *Transcriptional diversity at the UGT2B7 locus is dictated by extensive pre-mRNA splicing mechanisms that give rise to multiple mRNA splice variants.* Pharmacogenetics and genomics, 2011. **21**(10): p. 631-41.
259. Higuchi, R., et al., *Kinetic PCR analysis: real-time monitoring of DNA amplification reactions.* Bio/technology, 1993. **11**(9): p. 1026-30.
260. Morrison, T.B., J.J. Weis, and C.T. Wittwer, *Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification.* Biotechniques, 1998. **24**(6): p. 954-8, 960, 962.
261. VanGuilder, H.D., K.E. Vrana, and W.M. Freeman, *Twenty-five years of quantitative PCR for gene expression analysis.* Biotechniques, 2008. **44**(5): p. 619-26.
262. Gerard, C.J., et al., *Improved quantitation of minimal residual disease in multiple myeloma using real-time polymerase chain reaction and plasmid-DNA complementarity determining region III standards.* Cancer research, 1998. **58**(17): p. 3957-64.
263. Pfaffl, M.W., et al., *Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations.* Biotechnology letters, 2004. **26**(6): p. 509-15.
264. Freeman, W.M., S.J. Walker, and K.E. Vrana, *Quantitative RT-PCR: pitfalls and potential.* Biotechniques, 1999. **26**(1): p. 112-22, 124-5.
265. Gentle, A., F. Anastasopoulos, and N.A. McBrien, *High-resolution semi-quantitative real-time PCR without the use of a standard curve.* Biotechniques, 2001. **31**(3): p. 502, 504-6, 508.
266. Liu, W. and D.A. Saint, *A new quantitative method of real time reverse transcription polymerase chain reaction assay based on simulation of polymerase chain reaction kinetics.* Analytical biochemistry, 2002. **302**(1): p. 52-9.
267. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
268. Pfaffl, M.W., *A new mathematical model for relative quantification in real-time RT-PCR.* Nucleic acids research, 2001. **29**(9): p. e45.
269. Wong, M.L. and J.F. Medrano, *Real-time PCR for mRNA quantitation.* Biotechniques, 2005. **39**(1): p. 75-85.
270. Labos, E., *Codes, operations, measurements and neural networks.* Bio Systems, 2000. **58**(1-3): p. 9-18.
271. Andersen, C.L., J.L. Jensen, and T.F. Orntoft, *Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.* Cancer research, 2004. **64**(15): p. 5245-50.
272. Huggett, J., et al., *Real-time RT-PCR normalisation; strategies and considerations.* Genes and immunity, 2005. **6**(4): p. 279-84.
273. Larionov, A., A. Krause, and W. Miller, *A standard curve based method for relative real time PCR data processing.* BMC bioinformatics, 2005. **6**: p. 62.

274. Pfaffl, M.W., G.W. Horgan, and L. Dempfle, *Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR.* Nucleic acids research, 2002. **30**(9): p. e36.
275. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome biology, 2002. **3**(7): p. RESEARCH0034.
276. Spanakis, E., *Problems related to the interpretation of autoradiographic data on gene expression using common constitutive transcripts as controls.* Nucleic acids research, 1993. **21**(16): p. 3809-19.
277. Suzuki, T., P.J. Higgins, and D.R. Crawford, *Control selection for RNA quantitation.* Biotechniques, 2000. **29**(2): p. 332-7.
278. Thellin, O., et al., *Housekeeping genes as internal standards: use and limits.* Journal of biotechnology, 1999. **75**(2-3): p. 291-5.
279. Warrington, J.A., et al., *Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes.* Physiological genomics, 2000. **2**(3): p. 143-7.
280. Biederman, J., J. Yee, and P. Cortes, *Validation of internal control genes for gene expression analysis in diabetic glomerulosclerosis.* Kidney international, 2004. **66**(6): p. 2308-14.
281. Cicinnati, V.R., et al., *Validation of putative reference genes for gene expression studies in human hepatocellular carcinoma using real-time quantitative RT-PCR.* BMC cancer, 2008. **8**: p. 350.
282. Liu, D.W., S.T. Chen, and H.P. Liu, *Choice of endogenous control for gene expression in nonsmall cell lung cancer.* The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology, 2005. **26**(6): p. 1002-8.
283. Saviozzi, S., et al., *Selection of suitable reference genes for accurate normalization of gene expression profile studies in non-small cell lung cancer.* BMC cancer, 2006. **6**: p. 200.
284. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. **3**(7): p. RESEARCH0034.
285. Solanas, M., R. Moral, and E. Escrich, *Unsuitability of using ribosomal RNA as loading control for Northern blot analyses related to the imbalance between messenger and ribosomal RNA content in rat mammary tumors.* Analytical biochemistry, 2001. **288**(1): p. 99-102.
286. Abril, J., et al., *Altered expression of 12S/MT-RNR1, MT-CO2/COX2, and MT-ATP6 mitochondrial genes in prostate cancer.* The Prostate, 2008. **68**(10): p. 1086-96.
287. Dmitrenko, V., et al., *Reduction of the transcription level of the mitochondrial genome in human glioblastoma.* Cancer letters, 2005. **218**(1): p. 99-107.
288. Gerard, C.J., L.M. Andrejka, and R.A. Macina, *Mitochondrial ATP synthase 6 as an endogenous control in the quantitative RT-PCR analysis of clinical cancer samples.* Molecular diagnosis : a journal devoted to the understanding of human disease through the clinical application of molecular biology, 2000. **5**(1): p. 39-46.
289. Bustin, S.A., et al., *Quantitative real-time RT-PCR--a perspective.* Journal of molecular endocrinology, 2005. **34**(3): p. 597-601.
290. Imbeaud, S., et al., *Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces.* Nucleic acids research, 2005. **33**(6): p. e56.

291. Fleige, S. and M.W. Pfaffl, *RNA integrity and the effect on the real-time qRT-PCR performance.* Molecular aspects of medicine, 2006. **27**(2-3): p. 126-39.
292. Dutton, G.F., *Glucuronidation of drugs and other compounds*1980, Boca Raton, Fla.: CRC Press. 268 p.
293. Gong, Q.H., et al., *Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus.* Pharmacogenetics, 2001. **11**(4): p. 357-68.
294. Coughtrie, M.W., B. Burchell, and J.R. Bend, *Purification and properties of rat kidney UDP-glucuronosyltransferase.* Biochem Pharmacol, 1987. **36**(2): p. 245-51.
295. Olson, K.C., et al., *Functional characterization of low-prevalence missense polymorphisms in the UDP-glucuronosyltransferase 1A9 gene.* Drug Metab Dispos, 2009. **37**(10): p. 1999-2007.
296. Ethell, B.T., et al., *Use of cloned and expressed human UDP-glucuronosyltransferases for the assessment of human drug conjugation and identification of potential drug interactions.* Drug Metab Dispos, 2001. **29**(1): p. 48-53.
297. Coffman, B.L., et al., *Cloning and stable expression of a cDNA encoding a rat liver UDP-glucuronosyltransferase (UDP-glucuronosyltransferase 1.1) that catalyzes the glucuronidation of opioids and bilirubin.* Mol Pharmacol, 1995. **47**(6): p. 1101-5.
298. Izukawa, T., et al., *Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers.* Drug Metab Dispos, 2009. **37**(8): p. 1759-68.
299. Nakamura, A., et al., *Expression of UGT1A and UGT2B mRNA in human normal tissues and various cell lines.* Drug Metab Dispos, 2008. **36**(8): p. 1461-4.
300. Beaulieu, M., et al., *Isolation and characterization of a human orphan UDP-glucuronosyltransferase, UGT2B11.* Biochem Biophys Res Commun, 1998. **248**(1): p. 44-50.
301. Trdan Lusin, T., J. Trontelj, and A. Mrhar, *Raloxifene glucuronidation in human intestine, kidney, and liver microsomes and in human liver microsomes genotyped for the UGT1A1*28 polymorphism.* Drug metabolism and disposition: the biological fate of chemicals, 2011. **39**(12): p. 2347-54.
302. Bosma, P.J., et al., *The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome.* N Engl J Med, 1995. **333**(18): p. 1171-5.
303. Burchell, B. and R. Hume, *Molecular genetic basis of Gilbert's syndrome.* Journal of gastroenterology and hepatology, 1999. **14**(10): p. 960-6.
304. Bosma, P.J., *Inherited disorders of bilirubin metabolism.* J Hepatol, 2003. **38**(1): p. 107-17.
305. Bellemare, J., et al., *Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans.* The Journal of pathology, 2011. **223**(3): p. 425-35.
306. Hoskins, J.M., et al., *UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters.* J Natl Cancer Inst, 2007. **99**(17): p. 1290-5.
307. Ando, Y., et al., *UGT1A1 genotypes and glucuronidation of SN-38, the active metabolite of irinotecan.* Ann Oncol, 1998. **9**(8): p. 845-7.
308. Pan, Q., et al., *Alternative splicing of conserved exons is frequently species-specific in human and mouse.* Trends in genetics : TIG, 2005. **21**(2): p. 73-7.
309. Sorek, R., R. Shamir, and G. Ast, *How prevalent is functional alternative splicing in the human genome?* Trends in genetics : TIG, 2004. **20**(2): p. 68-71.
310. Price, A.L., E. Eskin, and P.A. Pevzner, *Whole-genome analysis of Alu repeat elements reveals complex evolutionary history.* Genome research, 2004. **14**(11): p. 2245-52.

311. Sorek, R., G. Ast, and D. Graur, *Alu-containing exons are alternatively spliced.* Genome research, 2002. **12**(7): p. 1060-7.
312. Nekrutenko, A. and W.H. Li, *Transposable elements are found in a large number of human protein-coding genes.* Trends in genetics : TIG, 2001. **17**(11): p. 619-21.
313. Ast, G., *How did alternative splicing evolve?* Nature reviews. Genetics, 2004. **5**(10): p. 773-82.
314. Lev-Maor, G., et al., *The birth of an alternatively spliced exon: 3' splice-site selection in Alu exons.* Science, 2003. **300**(5623): p. 1288-91.
315. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.
316. Amit, M., et al., *Biased exonization of transposed elements in duplicated genes: A lesson from the TIF-IA gene.* BMC molecular biology, 2007. **8**: p. 109.
317. Lin, L., et al., *Diverse splicing patterns of exonized Alu elements in human tissues.* PLoS genetics, 2008. **4**(10): p. e1000225.
318. Ohtsuki, S., et al., *Simultaneous absolute protein quantification of transporters, cytochromes P450, and UDP-glucuronosyltransferases as a novel approach for the characterization of individual human liver: comparison with mRNA levels and activities.* Drug metabolism and disposition: the biological fate of chemicals, 2012. **40**(1): p. 83-92.
319. Beaulieu, M., et al., *Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids.* The Journal of biological chemistry, 1996. **271**(37): p. 22855-62.
320. Levesque, E., et al., *Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene.* Pharmacogenetics, 1997. **7**(4): p. 317-25.
321. Levesque, E., et al., *Characterization and substrate specificity of UGT2B4 (E458): a UDP-glucuronosyltransferase encoded by a polymorphic gene.* Pharmacogenetics, 1999. **9**(2): p. 207-16.
322. Strassburg, C.P., M.P. Manns, and R.H. Tukey, *Expression of the UDP-glucuronosyltransferase 1A locus in human colon. Identification and characterization of the novel extrahepatic UGT1A8.* The Journal of biological chemistry, 1998. **273**(15): p. 8719-26.
323. Strassburg, C.P., et al., *Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue.* Molecular pharmacology, 1997. **52**(2): p. 212-20.
324. Fournel-Gigleux, S., et al., *Expression of a human liver cDNA encoding a UDP-glucuronosyltransferase catalysing the glucuronidation of hyodeoxycholic acid in cell culture.* FEBS letters, 1989. **243**(2): p. 119-22.
325. Pillot, T., et al., *Glucuronidation of hyodeoxycholic acid in human liver. Evidence for a selective role of UDP-glucuronosyltransferase 2B4.* The Journal of biological chemistry, 1993. **268**(34): p. 25636-42.
326. Turgeon, D., et al., *Glucuronidation of arachidonic and linoleic acid metabolites by human UDP-glucuronosyltransferases.* Journal of lipid research, 2003. **44**(6): p. 1182-91.
327. Kim, P.M. and P.G. Wells, *Genoprotection by UDP-glucuronosyltransferases in peroxidase-dependent, reactive oxygen species-mediated micronucleus initiation by the carcinogens 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and benzo[a]pyrene.* Cancer research, 1996. **56**(7): p. 1526-32.
328. Vienneau, D.S., U. DeBoni, and P.G. Wells, *Potential genoprotective role for UDP-glucuronosyltransferases in chemical carcinogenesis: initiation of micronuclei by*

benzo(a)pyrene and benzo(e)pyrene in UDP-glucuronosyltransferase-deficient cultured rat skin fibroblasts. Cancer research, 1995. **55**(5): p. 1045-51.
329. Beaulieu, M., et al., *Isolation and characterization of a simian UDP-glucuronosyltransferase UGT2B18 active on 3-hydroxyandrogens.* Journal of molecular biology, 1998. **275**(5): p. 785-94.
330. Masanek, U., G. Stammler, and M. Volm, *Modulation of multidrug resistance in human ovarian cancer cell lines by inhibition of P-glycoprotein 170 and PKC isoenzymes with antisense oligonucleotides.* Journal of experimental therapeutics & oncology, 2002. **2**(1): p. 37-41.
331. Volm, M., *Multidrug resistance and its reversal.* Anticancer research, 1998. **18**(4C): p. 2905-17.
332. Gardner-Stephen, D.A. and P.I. Mackenzie, *Isolation of the UDP-glucuronosyltransferase 1A3 and 1A4 proximal promoters and characterization of their dependence on the transcription factor hepatocyte nuclear factor 1alpha.* Drug metabolism and disposition: the biological fate of chemicals, 2007. **35**(1): p. 116-20.
333. Gregory, R.T., *Presidential address: The state of the union of vascular surgery, 2004.* Vascular, 2004. **12**(6): p. 354-8.
334. Gregory, P.A., et al., *The caudal-related homeodomain protein Cdx2 and hepatocyte nuclear factor 1alpha cooperatively regulate the UDP-glucuronosyltransferase 2B7 gene promoter.* Pharmacogenet Genomics, 2006. **16**(7): p. 527-36.
335. Mackenzie, P.I., D.G. Hu, and D.A. Gardner-Stephen, *The regulation of UDP-glucuronosyltransferase genes by tissue-specific and ligand-activated transcription factors.* Drug metabolism reviews, 2010. **42**(1): p. 99-109.
336. Mojarrabi, B., R. Butler, and P.I. Mackenzie, *cDNA cloning and characterization of the human UDP glucuronosyltransferase, UGT1A3.* Biochemical and biophysical research communications, 1996. **225**(3): p. 785-90.
337. Carey, M.F., C.L. Peterson, and S.T. Smale, *Dignam and Roeder nuclear extract preparation.* Cold Spring Harbor protocols, 2009. **2009**(12): p. pdb prot5330.
338. Kroeger, K.M. and L.J. Abraham, *Magnetic bead purification of specific transcription factors using mutant competitor oligonucleotides.* Analytical biochemistry, 1997. **250**(1): p. 127-9.
339. Radominska-Pandya, A., et al., *Structural and functional studies of UDP-glucuronosyltransferases.* Drug Metab Rev, 1999. **31**(4): p. 817-99.
340. Mackenzie, P.I., et al., *Steroid UDP glucuronosyltransferases: characterization and regulation.* The Journal of endocrinology, 1996. **150 Suppl**: p. S79-86.
341. Zhou, D.X. and T.S. Yen, *The ubiquitous transcription factor Oct-1 and the liver-specific factor HNF-1 are both required to activate transcription of a hepatitis B virus promoter.* Molecular and cellular biology, 1991. **11**(3): p. 1353-9.
342. Zhang, Y., et al., *A major inducer of anticarcinogenic protective enzymes from broccoli: isolation and elucidation of structure.* Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(6): p. 2399-403.
343. Crampsie, M.A., et al., *Phenylbutyl isoselenocyanate modulates phase I and II enzymes and inhibits 4-(methylnitrosamino)-1-(3-pyridyl)- 1-butanone-induced DNA adducts in mice.* Cancer prevention research, 2011. **4**(11): p. 1884-94.
344. Thomas, S.S., et al., *Genetic variability, haplotypes, and htSNPs for exons 1 at the human UGT1A locus.* Human mutation, 2006. **27**(7): p. 717.
345. Chiu, S.H. and S.W. Huskey, *Species differences in N-glucuronidation.* Drug metabolism and disposition: the biological fate of chemicals, 1998. **26**(9): p. 838-47.

346. Krupp, M., et al., *RNA-Seq Atlas - A reference database for gene expression profiling in normal tissue by next generation sequencing.* Bioinformatics, 2012.
347. Huynh-Thu, V.A., et al., *Inferring regulatory networks from expression data using tree-based methods.* PloS one, 2010. **5**(9).
348. Eisen, M.B., et al., *Cluster analysis and display of genome-wide expression patterns.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(25): p. 14863-8.
349. Duncan, S.A., et al., *Regulation of a transcription factor network required for differentiation and metabolism.* Science, 1998. **281**(5377): p. 692-5.
350. Hernandez, P., et al., *Evidence for systems-level molecular mechanisms of tumorigenesis.* BMC Genomics, 2007. **8**: p. 185.
351. Lim, J., et al., *A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration.* Cell, 2006. **125**(4): p. 801-14.
352. Mani, K.M., et al., *A systems biology approach to prediction of oncogenes and molecular perturbation targets in B-cell lymphomas.* Molecular systems biology, 2008. **4**: p. 169.
353. Rhodes, D.R. and A.M. Chinnaiyan, *Integrative analysis of the cancer transcriptome.* Nature genetics, 2005. **37 Suppl**: p. S31-7.
354. Chu, J.H., et al., *A graphical model approach for inferring large-scale networks integrating gene expression and genetic polymorphism.* BMC systems biology, 2009. **3**: p. 55.
355. Chu, J.H., et al., *Quantifying differential gene connectivity between disease states for objective identification of disease-relevant genes.* BMC systems biology, 2011. **5**: p. 89.
356. Kennedy, R.D., et al., *BRCA1 and c-Myc associate to transcriptionally repress psoriasin, a DNA damage-inducible gene.* Cancer research, 2005. **65**(22): p. 10265-72.
357. Margheri, F., et al., *Systemic sclerosis-endothelial cell antiangiogenic pentraxin 3 and matrix metalloprotease 12 control human breast cancer tumor vascularization and development in mice.* Neoplasia, 2009. **11**(10): p. 1106-15.
358. Moon, A., et al., *Global gene expression profiling unveils S100A8/A9 as candidate markers in H-ras-mediated human breast epithelial cell invasion.* Molecular cancer research : MCR, 2008. **6**(10): p. 1544-53.
359. Rhee, D.K., S.H. Park, and Y.K. Jang, *Molecular signatures associated with transformation and progression to breast cancer in the isogenic MCF10 model.* Genomics, 2008. **92**(6): p. 419-28.

CURRICULUM VITAE

NAME	POSITION TITLE
Jones, Nathan, R	Graduate Assistant, Department of Pharmacology Penn State College of Medicine
EMAIL ADDRESS njones1@hmc.psu.edu	

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
Penn State University, University Park	B.Sc.	2002-2006	Biochem and Molec Bio
Penn State University, Harrisburg	MBA	2007-2009	Business Admin
Penn State University, College of Medicine	Ph.D.	2006-2012	Pharmacology

A. Positions and Honors

2007-2012: Graduate Assistant, Laboratory of Dr. Philip Lazarus, Department of Pharmacology, Penn State College of Medicine, Hershey, Pennsylvania
2009: Class of 1983 Endowed Scholarship, Penn State College of Medicine Alumni Society
2008: Karl H. Beyer Scholarship, Penn State College of Medicine, Awarded to the graduate student with the highest GPA after completion of all coursework
2007: Class of 1975 Endowed Scholarship, Penn State College of Medicine Alumni Society
2006-2009: University Graduate Fellowship, Penn State College of Medicine
2004: Undergraduate Internship, Walter Reed Army Institute of Research, Division of Biochemistry Department of Molecular Pharmacology, Silver Spring, MD

B. Publications

1. **Jones NR**, Spratt TE, Berg AS, Muscat JE, Lazarus P, Gallagher CJ. Association studies of *Excision Repair Cross-Complementation Group 1* (*ERCC1*) haplotypes with lung and head and neck cancer risk in a Caucasian population. Cancer Epidemiology. 2011 35:175-81.
2. Mikse OR, Blake DC Jr, **Jones NR**, Sun YW, Amin S, Gallagher CJ, Lazarus P, Weisz J, Herzog CR. FOXO3 Encodes a Carcinogen-Activated Transcription Factor Frequently Deleted in Early-Stage Lung Adenocarcinoma. Cancer Res. 2010 Aug 1;70(15):6205-15. Epub 2010 Jul 14.
3. Gimbor M, **Jones NR**, Desai D, Das A, Aliaga C, Lazarus P, Amin S, Sharma AK. Phenyl butyl isoselenocyanate (ISC4) given orally to A/J mice reaches target tissues of lung and liver and results in time-dependent inhibition of Phase I CYP enzyme activity, induction of Phase II gene expression, and inhibition of NNK-induced DNA adduct formation. *Cancer Prevention Research* In press.
4. **Jones NR,** Sun D, Freeman W, Lazarus P. Quantification of hepatic UGT1A splice variant expression and correlation of UGT1A1 variant expression with hepatic glucuronidation activity Submitted to *J Pharmacol Exper Ther.*
5. **Jones NR**, Lazarus P. Role of HNF3α and Oct-1 in the hepatic transcriptional regulation of UGT2B10. (*In preparation*).
6. **Jones NR,** Lazarus P. Quantification of UGT2B gene expression in human tissues. (*In preparation*).
7. **Jones NR,** Lazarus P. Mitochondrial ATP synthase subunit 6 as an endogenous control gene for the quantitative RT-PCR analysis of human tissue samples. (*In preparation*).
8. Sun D, **Jones NR**, Lazarus P. Characterization of raloxifene glucuronidation and the importance of the UGT1A1*28 and UGT1A8 codon 173 variants in *in vivo* raloxifene metabolism. (In preparation)

C. Conference Presentations

1. **Jones NR,** Freeman W, Lazarus P. Quantitative and correlative analysis of variant UGT1A mRNA expression in human liver. *AACR Annual Meeting, 2010.*

2. **Jones NR**, Spratt TE, Muscat JE, Lazaurs P, and Gallagher CJ. Association studies of Excision Repair Cross-Complementation Group 1 (ERCC1) haplotypes with lung and oral cancer risk in a Caucasian population. *AACR Frontiers in Cancer Prevention Meeting*. 2007.