Pharmacology & Therapeutics 204 (2019) 107414

Contents lists available at ScienceDirect

Pharmacology & Therapeutics

journal homepage: www.elsevier.com/locate/pharmthera

The UGTome: The expanding diversity of UDP glycosyltransferases and its impact on small molecule metabolism

Dong Gui Hu [1], J.ulie-Ann Hulin [1], Pramod C. Nair, Alex Z. Haines, Ross A. McKinnon, Peter I. Mackenzie, Robyn Meech *

Department of Clinical Pharmacology and Flinders Cancer Centre, Flinders University College of Medicine and Public Health, Flinders Medical Centre, Bedford Park, South Australia, Australia

ARTICLE INFO

Available online 21 October 2019

ABSTRACT

The UDP glycosyltransferase (UGT) superfamily of enzymes is responsible for the metabolism and clearance of thousands of lipophilic chemicals including drugs, toxins and endogenous signaling molecules. They provide a protective interface between the organism and its chemical-rich environment, as well as controlling critical signaling pathways to maintain healthy tissue function. UGTs are associated with drug responses and interactions, as well as a wide range of diseases including cancer. The human genome contains 22 UGT genes; however as befitting their exceptionally diverse substrate ranges and biological activities, the output of these UGT genes is functionally diversified by multiple processes including alternative splicing, post-translational modification, homo- and hetero-oligomerization, and interactions with other proteins. All UGT genes are subject to extensive alternative splicing generating variant/truncated UGT proteins with altered functions including the capacity to dominantly modulate/inhibit cognate full-length forms. Heterotypic oligomerization of different UGTs can alter kinetic properties relative to monotypic complexes, and potentially produce novel substrate specificities. Moreover, the recently profiled interactions of UGTs with non-UGT proteins may facilitate coordination between different metabolic processes, as well as providing opportunities for UGTs to engage in novel 'moonlighting' functions. Herein we provide a detailed and comprehensive review of all known modes of UGT functional diversification and propose a UGTome model to describe the resulting expansion of metabolic capacity and its potential to modulate drug/xenobiotic responses and cell behaviours in normal and disease contexts.

Contents

Abbreviations: bp, base pair; CN, Crigler-Najjar; CYP, cytochrome P450; ER, endoplasmic reticulum; FRET, fluorescence resonance energy transfer; ORF, open reading frame; UDPGA, UDP-glucuronic acid; UGT, UDP-glycosyltransferase/glucuronosyltransferase; UTR, untranslated region.

* Corresponding author at: Department of Clinical Pharmacology, Flinders University College of Medicine and Public Health, Flinders University, Bedford Park, South Australia, Australia.

E-mail address: robyn.meech@flinders.edu.au (R. Meech).

[1] Denotes joint first authors.

1. Introduction: the relevance of glycosidation in pharmacology, toxicology, metabolic homeostasis, and disease

1.1. Metabolic transformation and clearance of small molecules

The human body is exposed to a myriad of small chemicals from the environment, diet, and in the form of prescription and other drugs. Many of these chemicals are lipophilic in nature and may accumulate

https://doi.org/10.1016/j.pharmthera.2019.107414
0163-7258/

in tissues to toxic levels unless effectively detoxified and cleared. In addition, many small lipophilic molecules are produced within the body either as byproducts of metabolism, or to perform important biological functions including cell signaling. These endogenous molecules typically also require clearance and/or must be maintained within precise concentration ranges to ensure normal tissue function. The detoxification and clearance of such molecules is mediated by a metabolic cascade that includes oxidation, reduction and conjugation reactions. Conjugation reactions typically involve the covalent attachment of polar groups that render the target more water soluble and often charged, facilitating transport and elimination. Of the enzyme families known to catalyze conjugation, the largest and arguably most important are the UDP glycosyltransferases (UGTs).

UGTs conjugate an exceptionally diverse range of small lipophilic molecule with sugars including glucuronic acid, glucose, xylose, N-acetylglucosamine or galactose; although glucuronic acid is most commonly used in mammals. The general glycosidation mechanism in which the sugar moiety from an activated nucleotide sugar donor is transferred to a suitable acceptor group on the aglycone is shown in Fig. 1A. Conjugation by UGTs is known to be critical to the clearance of many thousands of chemicals from the body including environmental toxins and carcinogens, clinically used drugs, bioactive dietary components, and endogenously produced metabolites and ligands for signaling molecules.

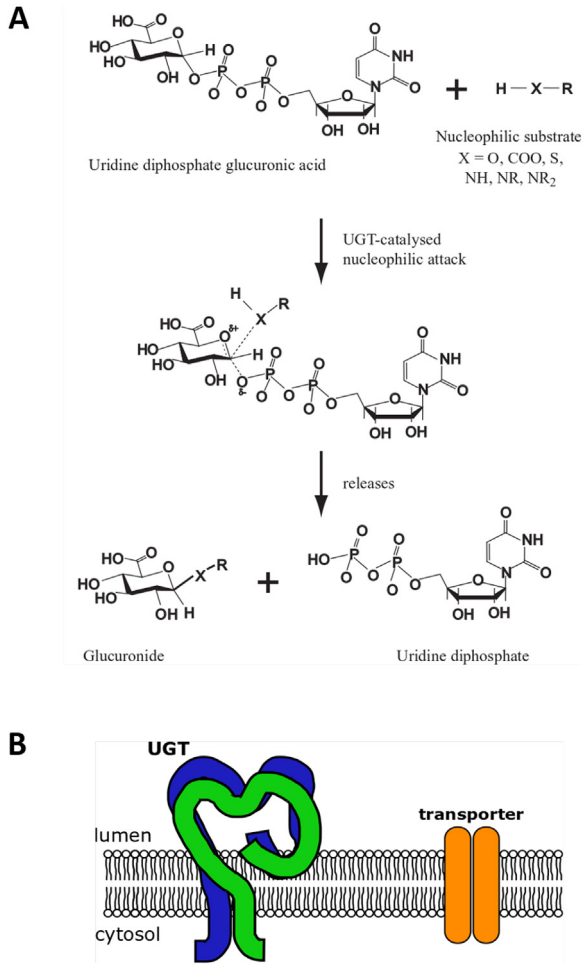


Fig. 1. A. The glycosidation reaction mechanism shown using UDPGA and a generic acceptor as exemplars. **B.** The generally accepted model for UGT enzyme topology in the ER including potential for dimerization.

1.2. The UGT superfamily

The human genome contains 22 genes encoding distinct UGT enzymes. These enzymes form part of a much larger superfamily of UDP-glycosyltransferases (GT-1) that spans essentially all plant, animal and bacterial taxa (Campbell, Davies, Bulone, & Henrissat, 1997). The broad evolutionary conservation of this family reflects the fundamental nature of the biological requirement to control levels of small lipophilic molecules within cellular organisms. It should be noted that the members of the GT-1 superfamily all transfer glycosyl groups to relatively small lipophilic chemicals and are distinct in sequence and structure from the families of enzymes that transfer sugars to macromolecules such as proteins. In particular, the GT-1 superfamily is distinguished by structural features such as protein size (around 500 amino acids) and structural fold (GT—B), as well as a characteristic 'signature sequence' in their C-terminal half that is involved in UDP sugar binding (Meech & Mackenzie, 1997a, 1997b; Mackenzie et al., 2005a, 2005b).

The UGT gene superfamily is believed to have arisen via a series of gene duplication events and subsequent divergence, with substrate specificities evolving to help organisms exploit different environmental niches (Mackenzie et al., 2005a, 2005b). Human UDP-glycosyltransferases are divided into four families (UGT1, UGT2, UGT3, UGT8); members within each family share more than 45 % amino acid identity but are ≤45 % identical to UGTs of other families. The human UGT2 family is further divided into subfamilies, UGT2A and UGT2B with members of each sharing ≥70 % amino acid identity. The human UGT3 family contains two members with 78 % identity, while the UGT8 family has only one member, formerly called ceramide galactosyltransferase (CGT).

The UGT families show a distinctive pattern of UDP-sugar co-factor preference. The members of the UGT1 and UGT2 subfamilies show preferential usage of UDP-glucuronic acid (UDPGA), leading to their historical designation in literature as UDP-glucuronosyltransferases. The human UGT3A1 enzyme uses UDP-*N*-acetylglucosamine, while UGT3A2 uses UDP-glucose and UDP-xylose (Mackenzie et al., 2008; Meech & Mackenzie, 2010; Mackenzie et al., 2011). UGT8 uses only UDP-galactose (Meech et al., 2014). The human *UGT*s are well conserved across primate species. However, direct orthologues are only identifiable for a subset of human *UGT*s in lower mammals such as rodents, reflecting relatively rapid divergence over evolutionary time.

1.3. UGTs are ER resident type I membrane proteins

All vertebrate UGT proteins are about 530 amino acids in length and are comprised of two domains: the N-terminal domain (approximately the N-terminal half of the polypeptide) is proposed to be involved in substrate selection and the more conserved C-terminal domain is known to bind the UDP sugar. Not only is the N-terminal half of the protein more variable between isoforms than the C-terminal half, it is proposed to be more flexible based on molecular modelling studies and its resistance to crystallization (Miley et al., 2007). Early studies showed that mammalian UGTs are integral membrane proteins localized in the endoplasmic reticulum (ER) and to a lesser degree in the contiguous nuclear envelope (Shepherd, Baird, Hallinan, & Burchell, 1989; Meech & Mackenzie, 1998; Radominska-Pandya et al., 2002). UGTs contain a signal peptide at the N-terminus which is involved in insertion into the ER and is subsequently cleaved from the mature protein, and a stretch of approximately 17 hydrophobic amino acids near the C-terminus which is predicted to span the lipid bilayer (Mackenzie & Owens, 1984; Iyanagi et al., 1986). The accepted membrane topology, with most of the protein located in the ER lumen and only a short C-terminal tail region extending into the cytoplasm, is shown in Fig. 1B. The cytoplasmic tail contains an ER retention/retrieval signal, although it has been reported that N-terminal regions of the protein may also play a role in membrane and ER retention (Barré, Magdalou, Netter, Fournel-Gigleux, & Ouzzine, 2005; Ouzzine, Magdalou, Burchell, &

Fournel-Gigleux, 1999). The general topology of UGTs is consistent with activity assays as well as structure-function studies that demonstrated that UGTs must be associated with the membrane for optimal activity. These studies also gave rise to the concept of latency, whereby activity is increased by disruption of the microsomal membrane, which is in part due to vesicularization during microsome preparation (Wishart & Fry, 1980; Shepherd et al., 1989; Coughtrie, Blair, Hume, & Burchell, 1991).

1.4. UGTs influence drug response and disease risk via metabolism of diverse exogenous and endogenous substrates

Human UGT enzymes are expressed in a wide range of organs and tissues, but most isoforms show greatest abundance in liver, kidney and intestinal tract, reflecting their roles in detoxification. Each UGT enzyme can conjugate a large number (potentially thousands) of different compounds and there is also considerable overlap in substrate selectivities between different enzymes. The resulting redundancy is thought to play a role in reducing the impact of genetic and regulatory variation in UGTs. Despite the robustness provided by redundancy, there are numerous examples where variation in UGT function has adverse consequences, including disease, drug toxicity, and drug-drug or drug-xenobiotic/endobiotic interactions. These functions have been reviewed recently in detail (Meech et al., 2019); hence they will only be briefly summarized here.

The only disease known to result entirely from UGT dysfunction is unconjugated hyperbilirubinemia. Bilirubin is a breakdown product of heme, which is released from ageing red blood cells and is toxic at high concentrations inducing jaundice and potentially neonatal brain damage called kernicterus. UGT1A1 is the only enzyme that glucuronidates bilirubin and thus facilitates its excretion in bile (Bosma et al., 1994). Numerous genetic variants are known to reduce *UGT1A1* gene expression resulting in the mild unconjugated hyperbilirubinemia called Gilbert's syndrome, which affects ~3–9 % of individuals of European ancestry (Bosma et al., 1995). In addition, inactivating mutations in *UGT1A1* result in the rarer severe unconjugated hyperbilirubinemia called Crigler Najjar (CN) syndrome (Servedio et al., 2005).

Another important class of endogenous substrates for UGTs are steroid hormones that act as ligands for nuclear receptors. Of particular importance is the glucuronidation of androgens and estrogens in steroid dependent tissues such as prostate and breast. Because steroid glucuronides do not bind to nuclear receptors and glucuronidation is largely irreversible in these tissues, it is considered an important mechanism to terminate steroid (mainly androgen) signalling. Consistent with this function, several steroid conjugating UGTs have been linked to breast and/or prostate cancer risk and progression (Chouinard et al., 2008; Gauthier-Landry, Belanger, & Barbier, 2015; Hu, Mackenzie, McKinnon, & Meech, 2016; Sun, Liu, McCloskey, & Lazarus, 2011; Yao et al., 2010; Hu, Selth et al., 2016; Zhu et al., 2018). The conjugation of exogenous chemicals by UGTs is also linked to cancer risk; many UGTs are expressed at interfaces with the environment such as airways and gut, where they detoxify carcinogenic xenobiotics. Genetic variation in some UGTs that metabolize carcinogenic or tumour promoting chemicals (including benzo(a)pyrene, heterocyclic amines and nitrosamines from smoke and cooked/cured meat) have been linked to cancer (Butler et al., 2005; Kua et al., 2012).

Several UGTs conjugate clinically important drugs and thus variation in their expression or function can contribute to drug efficacy and play a role in drug-drug or drug-xenobiotic/endobiotic interactions. One important example is the ability of UGT1A1 to conjugate the active metabolite of the anti-cancer drug irinotecan. A low activity allele of UGT1A1 is associated with impaired glucuronidation and hence increased irinotecan toxicity (Ichikawa et al., 2008); which previously led the US Food and Drug Administration (FDA) to recommend *UGT1A1* genotyping of patients prior to irinotecan treatment (O'Dwyer & Catalano, 2006), although this is not in common practice.

It is clear from the above examples (and many more reviewed in (Meech et al., 2019)), that mechanisms that contribute to inter-individual variability in UGT expression and function can be relevant for predicting patient drug response, and for potentially predicting the risk of diseases that are linked to UGT function, such as cancer. Our specific focus in this review is on mechanisms that change the functions of UGT proteins, rather than the abundance of protein produced from each UGT gene. Such variation arises primarily from alternative splicing, UGT oligomerization, and other protein-protein interactions. We will not discuss here genetic variation, nor changes in expression levels that arise due to transcriptional regulation and miRNA-mediated post-transcriptional regulation as these have all been extensively reviewed relatively recently (2016a, Guillemette, Levesque, Harvey, Bellemare, & Menard, 2010; Hu, Meech, McKinnon, & Mackenzie, 2014; Hu, Selth et al., 2016; Hong, Guoxiu, Guangji, & Haiping, 2018; Nakano & Nakajima, 2018; Meech et al., 2019).

2. Expanding the diversity of UGTs through alternative splicing of UGT transcripts

2.1. Human UGT gene organization

The human genome contains 22 functional *UGT* genes: nine *UGT1A* family members, three *UGT2A* genes, seven *UGT2B* genes, two *UGT3A* genes and *UGT8A* (Mackenzie et al., 1997, 2005a; Mackenzie et al., 2005b). There are also multiple pseudogenes and gene remnants, some of which may play a role in the diversification of the UGTome. The structures of the gene loci corresponding to the four *UGT* families are shown in Fig. 2A. The human *UGT1A* locus spans approximately 200 kb on chromosome 2q37 and contains 13 unique exons (exon 1), which encode the N-terminal half of each UGT1A form, and a set of 4 common exons (exons 2–5) that are shared by all UGT1A forms (Gong et al., 2001; Ritter et al., 1992). Thus, UGT1A enzymes have unique N-terminal ends (comprised of ~290 amino acids) but identical C-terminal regions (comprised of 245 amino acids) (Fig. 2A). Four *UGT1A* first exons contain mutations that disrupt their open reading frame and hence are designated pseudogenes (*UGT1A2P*, *UGT1A11P*, *UGT1A12P*, and *UGT1A13P*). The nine functional UGT1 proteins (*1A1, 1A3, 1A4 1A5, 1A6, 1A7, 1A8, 1A9, 1A10*) can be parsed into homology clusters, e.g. *UGT1A3*, *UGT1A4*, and *UGT1A5* showing greater than 93 % overall identity, and *UGT1A7, 1A8, 1A9, 1A10* showing from 89 to 95% identity. All vertebrate *UGT1* loci show the unusual shared exon organization, suggesting that this genomic structure appeared early in the evolution of UGTs (Mackenzie et al., 2005a, 2005b).

The nine *UGT2A* and *UGT2B* genes (*2A1/2, 2A3*, *2B4, 2B7, 2B10, 2B11, 2B15, 2B17, 2B28*) are clustered on chromosome 4 at position 4q13. The *UGT2A1* and *UGT2A2* genes are derived from a single locus containing two isoform-specific exons 1 followed by a set of shared exons similar to the *UGT1A* locus (Jedlitschky, Cassidy, Sales, Pratt, & Burchell, 1999; Mackenzie et al., 2005a, 2005b; Sneitz et al., 2009) (Fig. 2A). In contrast, the *UGT2A3* gene and all seven functional *UGT2B* genes are derived from discreet gene loci with no shared exons (Court, Hazarika, Krishnaswamy, Finel, & Williams, 2008; Mackenzie et al., 2005a, 2005b). There are also five *UGT2B* pseudogenes and various gene remnants scattered throughout the locus (Fig. 2A). The first two exons within each *UGT2B* gene encode approximately the first 290 amino acids of the protein; this is approximately the same region encoded by the unique exon 1 of each *UGT1* gene and includes the aglycone-recognition domain (see Fig. 2B).

The two *UGT3A* genes are tandemly arranged on Chromosome 5 at position 5p13.2 (Meech & Mackenzie, 2010); there is no *UGT3B* subfamily. Interestingly, the region encoding the N-terminal, aglycone-recognition domain of *UGT3A1* and *UGT3A2* is made up of four exons (Fig. 2B). The sole human *UGT8* gene is located on chromosome 4q26.

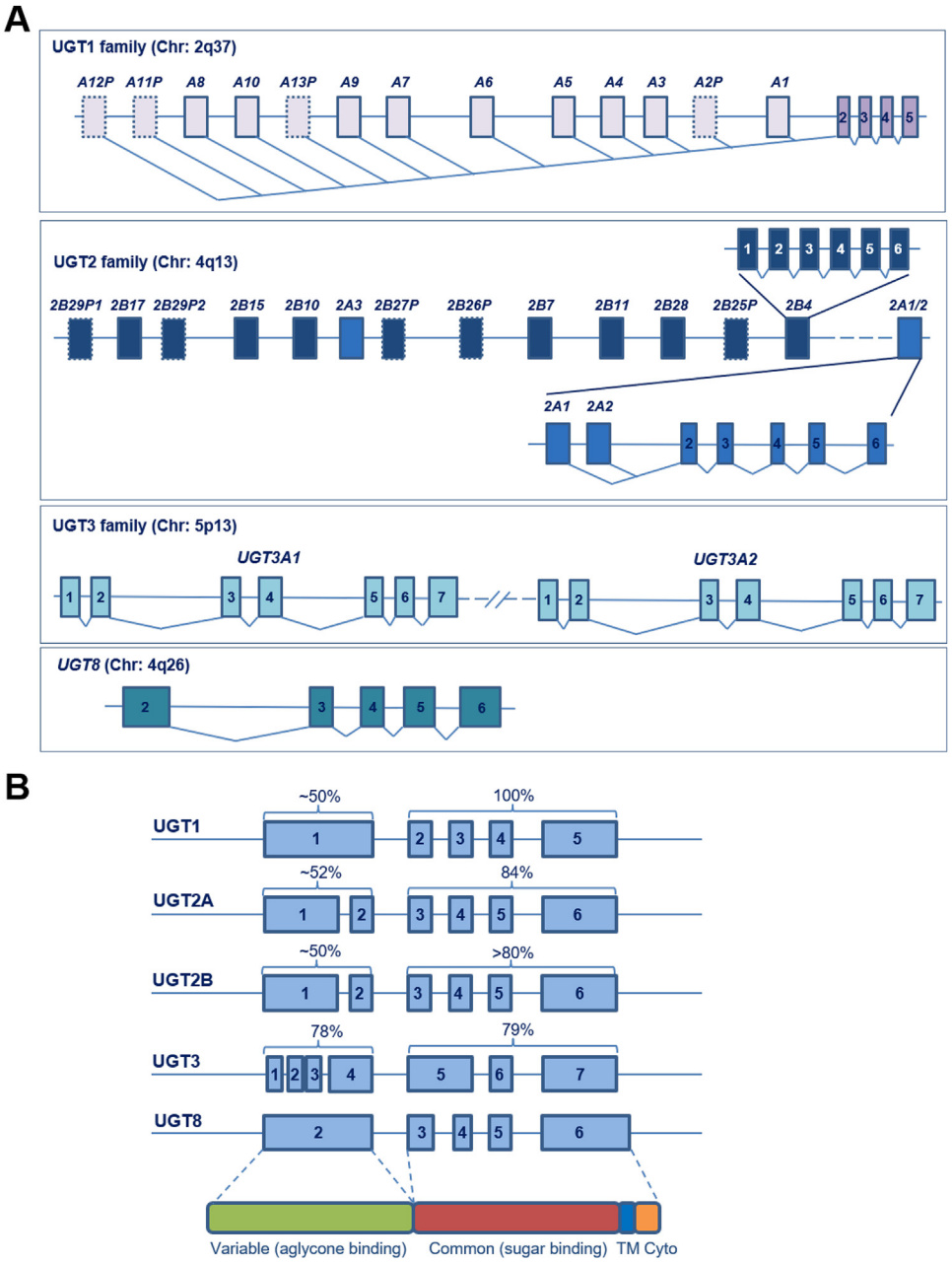


Fig. 2. A. Exonic structures of the gene loci corresponding to the four *UGT* families, see text for details. Note that only canonical/constitutive exons are shown whilst alternatively-used exons are omitted for simplicity. Pseudogenes are designated *P* and are represented by boxes with broken borders. **B.** The regions of the UGT protein encoded by each exon vary between UGT families. The N-terminal half (~240aa) of the protein that is believed to mediate substrate recognition is encoded by exon 1 in the *UGT1* genes, by exons 1 and 2 in the *UGT2* genes, by exons 1–4 in the *UGT3* genes, and by exon 2 in the *UGT8* gene. Note that only protein coding exons are shown. For all gene families except *UGT8*, the first coding exon is designated exon 1 as it is considered the site of constitutive transcription initiation in the majority of tissues. For *UGT8*, the first coding exon is designated exon 2 as non-coding upstream exons (designated exons 1) are considered the site of constitutive transcription initiation in the tissues where UGT8 is most abundantly expressed. See Fig. 4 for more details of the arrangement of the non-coding *UGT8* exons.

As mentioned, *UGT1* and *UGT2* gene orthologues and paralogues show much greater sequence conservation within the region encoding the C-terminal half of the protein than the region encoding the N-terminal half. This reflects the role of the N-terminal domain in recognition of diverse aglycone substrates, whilst the C-terminal domain binds the common UDP-sugar UDPGA (Fig. 2B). In contrast, sequence conservation between the paralogous human *UGT3A* subfamily members is similar in the N-terminal and C-terminal halves (Fig. 2). This may relate to the fact that the UGT3A1 and UGT3A2 enzymes have different sugar specificities (Mackenzie et al., 2008, 2011; Meech, Miners, Lewis, & Mackenzie, 2012). *UGT8* orthologues are found in all vertebrate species and are more highly conserved than other *UGT* orthologues; moreover, gene orthologues show a similar degree of sequence identity in the C-terminal and N-terminal halves. This may be explained by their very narrow aglycone specificity and the critical nature of these substrates: to date the UGT8 enzymes in humans and mice are only known to galactosidate ceramides and bile acids. Conjugation of ceramide produces galactosyl-ceramide which is essential for formation of myelin and loss of UGT8 function in mice leads to dysmyelination and death within a few weeks of birth (Coetzee et al., 1996).

2.2. Pre-mRNA splicing

Most eukaryotic protein-coding genes are multi-exon; following transcription, introns are removed from precursor messenger RNAs (pre-mRNAs) and exons are ligated to form mature mRNAs in a process termed pre-mRNA splicing (Park, Pan, Zhang, Lin, & Xing, 2018). In humans, pre-mRNA splicing is primarily carried out by the major spliceosome, consisting of five small nuclear ribonucleoprotein complexes (snRNPs), U1, U2, U4, U5 and U6 (Wahl, Will, & Luhrmann, 2009). Splicing is tightly controlled by *cis*-elements within the pre-mRNAs and *trans*-acting factors that bind to them. Canonical *cis*-elements within pre-mRNAs include the 5' splice sites (5'SSs), 3' splice sites (3'SSs), branch site (A), and polypyrimidine tract [Y(n)]. Canonical exons are generally flanked by the dinucleotide "AG" at the 5'-end of the exon (3'SS) and the dinucleotide "GT" at the 3'-end of the exon (5'SS), often referred to as the "GT-AG" rule (Breathnach, Benoist, O'Hare, Gannon, & Chambon, 1978; Mount, 1982). Splicing requires the assembly of the major spliceosome on every intron of the pre-mRNA and its subsequent disassembly upon intron excision (Lee & Rio, 2015). Spliceosome assembly involves the initial binding of the U1 snRNP, U2 snRNP, and U2 auxiliary factors (U2AF) to the respective 5'SS, the branch site and 3'SS, and the subsequent recruitment of the U4/U5/U6 tri-snRNP complex to form a catalytically active spliceosome (Lee & Rio, 2015). Pre-mRNAs may also harbour *cis*-elements such as exonic splicing enhancers (ESEs) or silencers (ESSs), and intronic splicing enhancers (ISEs) or silencers (ISSs). Serine/Arginine-rich proteins (SR proteins), a family of splicing factors, frequently promote splicing via binding to ESEs and ISEs. Many members of heterogeneous nuclear ribonucleoproteins (hnRNPs) are also splicing factors that can repress splicing via their binding to ESSs and ISSs (Dvinge, Kim, Abdel-Wahab, & Bradley, 2016; Park et al., 2018).

2.3. Alternative splicing and its deregulation in cancer

The human genome has at least 20,687 protein-coding genes (GENCODE annotation, V7), most of which are multi-exon with at least two constitutive exons and a varying number of alternative exons (Consortium, 2012). Pre-mRNAs of nearly 95 % of human multi-exon genes are alternatively spliced generating multiple transcripts (Pan, Shai, Lee, Frey, & Blencowe, 2008; Wang et al., 2008). As depicted in Fig. 3, there are at least eight types of alternative splicing events that involve 1) cassette alternative exons, 2) alternative 5' splice sites, 3) alternative 3' splice sites, 4) intron retention, 5) mutually exclusive alternative exons, 6) alternative promoter and first exons, 7) alternative poly (A) site and terminal exon, or 8) exitrons (also termed internal exonic intronization) (Blencowe, 2006; Dvinge et al., 2016; Marquez, Brown, Simpson, Barta, & Kalyna, 2012; Marquez, Hopfler, Ayatollahi, Barta, & Kalyna, 2015; Roy, Haupt, & Griffiths, 2013; Sibley et al., 2015; El Marabti & Younis, 2018; Park et al., 2018). In addition to these basic types of alternative splicing events, two or more types of alternative splicing events (termed complex alternative splicing events) frequently occur simultaneously within single pre-mRNAs of many human genes, generating chaotic transcripts (Park et al., 2018; Vaquero-Garcia et al., 2016). The Encyclopedia of DNA Elements (ENCODE) project (GENCODE annotation, V7) reveals that human protein-coding genes have on average 6.3 alternatively spliced transcripts (3.9 different protein-coding transcripts) per locus (Consortium, 2012). To support these findings, genome-wide high-resolution mass spectrometry analyses have shown that approximately 37 % of human multi-exon protein-coding genes have multiple protein isoforms (Kim et al., 2014). Alternative splicing is regulated in a cell-type-, tissue-, organ- and developmental-stage-specific manner (Park et al., 2018; Wang et al., 2008). Cumulative evidence supports a vital role of combinatorial regulation of alternative splicing networks for the development of human organs, including the brain, heart, skeletal muscle and liver (Baralle & Giudice, 2017). Therefore, alternative splicing represents a general mechanism that greatly expands the complexity of the human transcriptome and proteome (Liu et al., 2017).

Alternative splicing is frequently deregulated in cancers. The molecular mechanisms are not fully understood but could relate to 1) altered expression levels and somatic mutations of genes coding for splicing factors and components of the spliceosome (Climente-Gonzalez, Porta-Pardo, Godzik, & Eyras, 2017; Dvinge et al., 2016; El Marabti & Younis, 2018) and 2) somatic mutations disrupting splicing regulatory *cis*-elements in pre-mRNA (Jung et al., 2015; Supek, Minana, Valcarcel, Gabaldon, & Lehner, 2014; Scotti & Swanson, 2016). Modulators of alternative splicing that target *cis*-elements within pre-mRNAs (e.g. oligonucleotides, RNA binding inhibitors/activators) or *trans*-activating factors (e.g. splicing factor kinase inhibitors) are promising therapeutic agents for cancer therapy (Bates, Morris, Oltean, & Donaldson, 2017).

2.4. Constitutive splicing of UGT genes and tissue-specific UGT expression profiles

Constitutive splicing predominates for most human genes and produces mature mRNAs coding for wild-type functional proteins. For the *UGT1* genes, each isoform-specific first exon is transcribed from a separate isoform-specific promoter (Hu et al., 2014) and then constitutively spliced to the four shared exons as outlined in Section 2.2 (Mackenzie et al., 2005a, 2005b). Similarly, the two first exons of the *UGT2A1/2* locus are each transcribed from a unique promoter and constitutively spliced to the five shared exons (Mackenzie et al., 2005a, 2005b). While *UGT2A3*, the *UGT2B* genes, the *UGT3A* genes and *UGT8* are also transcribed from unique promoters, no exon sharing occurs during their constitutive splicing.

Cumulative studies using quantitative real-time RT-PCRs with isoform-specific primers have shown tissue-specific UGT expression profiles and inter-individual variability in constitutive UGT mRNA levels in human tissues (Nishimura & Naito, 2006; Nakamura, Nakajima, Yamanaka, Fujiwara, & Yokoi, 2008; Izukawa et al., 2009; Ohno & Nakajin, 2009; Court, 2010; Court et al., 2012; Schaefer et al., 2012). Isoform-specific transcriptional and post-transcriptional regulation is believed to contribute to this tissue-specific UGT expression pattern (Hu et al., 2014). Further comprehensive expression analysis can be performed using RNA-sequencing (RNA-seq) data accessible via the Human Protein Atlas (HPA; https://www.proteinatlas.org). The expression levels (TPM or RPKM) of 22 *UGT* genes in 43 human tissues representing HPA and/or GTEx (The Genotype-Tissue Expansion) RNA-seq datasets (Consortium, 2013) are shown in Table 1 and discussed further below.

UGT8 is expressed at widely varying levels in 42 of the 43 tissues shown in Table 1 (except for parathyroid gland); the highest expression is found in the brain (cerebral cortex, hippocampus, hypothalamus, caudate, cerebellum) while many other tissues such as kidney, duodenum, rectum, small intestine, colon, thyroid gland, gallbladder and stomach show moderate expression. The high expression of UGT8 in the brain reflects its aforementioned role in myelin formation, and its expression in the gastrointestinal tract could be relevant to its role of bile acid metabolism (Meech et al., 2015).

UGT1A1 is expressed in 36/43 tissues with high expression in liver, small intestine, and duodenum, and moderate expression in urinary bladder, esophagus, kidney, and colon. Five *UGT* genes (*1A6*, *1A10*, *2B7*, *2B15*, *2B17*) are expressed in more than 20 tissues. *UGT1A6* is highly expressed in urinary bladder, kidney, liver, and moderately in esophagus, skin, and stomach. *UGT1A10* is highly expressed throughout the gastroenteral tract, including esophagus, stomach, duodenum, small intestine, colon, and rectum. *UGT2B7* is highly expressed in kidney, liver, and small intestine. *UGT2B15* is highly expressed in gallbladder, liver, and moderately in small intestine, colon, duodenum, stomach, and rectum. *UGT2B17* is highly expressed in small intestine, colon, duodenum, and moderately in rectum, liver, appendix, and smooth muscle.

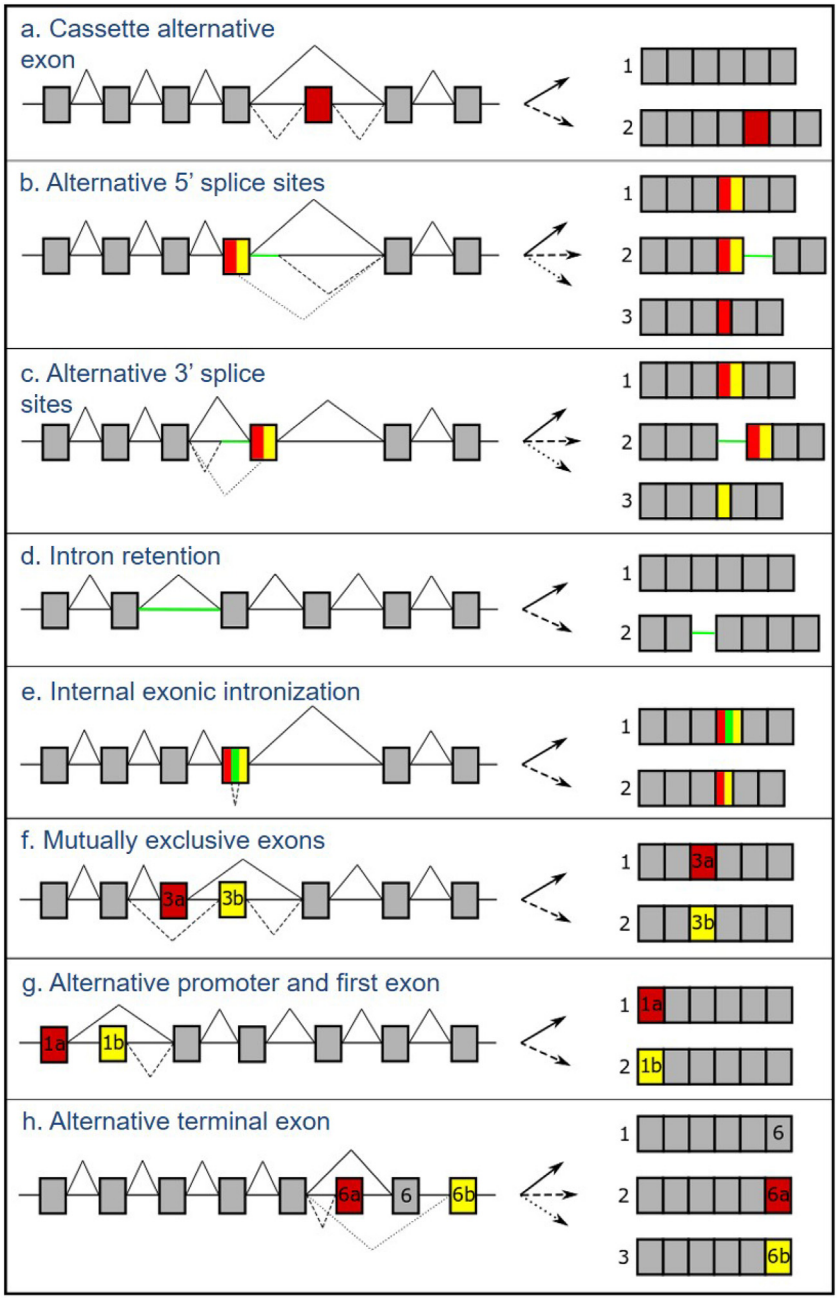


Fig. 3. Schematic depicting the different forms of alternative splicing discussed in Section 2. Key: a, cassette alternative exon; b, alternative 5′ splice sites; c, alternative 3′ splice sites; d, intron retention; e, internal exonic intronization (exitrons); f, mutually exclusive exons; g, alternative promoter and first exon; h, alternative terminal exon.

Five *UGT* genes (*1A7*, *2A3*, *2B11*, *3A1*, *3A2*) are expressed in at least 15 human tissues. *UGT1A7* is highly expressed in esophagus. *UGT2A3* is highly expressed in small intestine, kidney, and gallbladder, and moderately in rectum, colon, and liver. *UGT2B11* is highly expressed in the breast. *UGT3A1* and *UGT3A2* are differentially expressed in human tissues: *UGT3A1* is abundant in kidney, liver, duodenum, testis, and small intestine whereas *UGT3A2* is highly expressed in skin, testis, bone marrow, spleen, and breast. The remaining *UGT*s have their highest expression in liver (*1A3*, *1A4*, *2B4*, *2B10*, *2A1*), kidney (*1A9*), breast (*2B28*), urinary bladder (*1A8*), duodenum (*1A5*), and pituitary gland (*2A2*).

Most of the 22 *UGT*s are expressed in drug-metabolizing tissues such as liver, kidney, duodenum and small intestine. Only one gene (*2A2*) is not expressed in liver, three (*1A4*, *1A7*, *2B28)* are absent from kidney, five (*2B4*, *2B10*, *2B28*, *2A1*, *2A2*) are absent from duodenum and six (*1A9, 2B4*, *2B10*, *2B28*, *2A1, 3A2*) are absent from small intestine.

2.5. Alternative splicing of UGT genes and functions of alternatively spliced UGT transcripts

In addition to constitutive exons that encode wild-type UGT enzymes, all *UGT*s have alternative exons and generate alternatively spliced transcripts. A recent *UGT*-targeted RNA-seq analysis has revealed the transcriptomic landscape of all human *UGT1A, UGT2A* and *UGT2B* genes in normal and cancerous metabolic (e.g. liver, kidney, intestine and colon) and hormone-sensitive (e.g. prostate, breast, and uterus) tissues (Tourancheau et al., 2016). This study identified 234 novel exon-exon junctions and predicted over 130 structurally and functionally diverse novel UGT variant transcripts. However, only 3 % of the alternative splicing events were common to all tissues analysed, suggesting that variant UGT transcripts are expressed in a tissue-specific manner. The expression of variant UGT transcripts was also

Table 1
mRNA expression profiles of UGT genes in human tissues.

Tissues	mRNA levels of UGT genes in human tissues																						
		1A1	1A3	1A4	1A5	1A6	1A7	1A8	1A9	1A10	2B4	2B7	2B10	2B11	2B15	2B17	2B28	2A1	2A2	2A3	3A1	3A2	8	No %
Adipose tissue	TPM	0.1							0.2	0.1	0.8	0.3		0.1		0.3							0.7	8
	RPKM																							
Adrenal gland	TPM	0.2					0.8															1	0.9	4
	RPKM						0.1																	
Appendix	TPM	0.3				0.2		0.8		4.4		0.4		0.2	0.2	49				2			5.4	10
	RPKM														3.7									
Bone marrow*	TPM	0.4																				3.2	0.8	3
	RPKM																							
Breast	TPM	0.3				1	0.5				0.1	2.9		320	0.3	0.9	267			3.1	0.1	1.4	4.9	13
	RPKM											0.2		0.6	0.3	0.1	0.2						0.5	
Caudate#	TPM																						12	1
	RPKM																							
Cerebellum#	TPM																						3.9	1
	RPKM																							
Cerebral cortex	TPM	0.2				0.1	0.2		0.4			1.2				0.3					0.1		101	8
	RPKM																						6.3	
Cervix, Uterine	TPM	0.2				0.3						48		0.2	0.2	2.8			0.2				9.2	8
	RPKM											0.7				0.4							0.9	
Colon	TPM	3	0.1			1.8	0.1	16	0.4	49		3.7	0.1	0.1	18.8	372				23		0.1	22	15
	RPKM	1.1								8		0.5			18	103				11			3.1	
Duodenum*	TPM	199	2.9	6.7	7.5	6	0.2	2.8	0.1	80		69		0.7	17.9	371				82	9.5	0.1	27	17
	RPKM																							
Endometrium	TPM	0.4				0.1				0.2		15		0.1	0.1	2.4					0.1		5.7	9
	RPKM																						0.2	
Epididymis*	TPM	0.5	0.1			3.1						29		1.1		1.8							0.4	7
	RPKM																							
Esophagus	TPM	1.2				15	61.4	0.2		16	0.2	0.2			0.1	0.1							1.4	10
	RPKM	1.5				0.6	5.4			3.6													0.5	
Fallopian tube	TPM	0.2				0.1						79		0.8	0.1	0.4						0.2	6.1	8
	RPKM											1.7				0.1							0.2	
Gall bladder*	TPM	1.6	0.4	2.1	3.9	5.3	0.1	10.3		34		16	3.4	18	110	1	21			46		0.2	15	17
	RPKM																							
Heart Muscle	TPM	0.1									8.3												0.2	3
	RPKM										3.3													
Hippocampus#	TPM																						25	1
	RPKM																							
Hypothalamus#	TPM																					0.1	14	2
	RPKM																							
Kidney	TPM	4.3	0.2			75.4	0.4	0.4	465	0.3	1.3	1326	0.1	3.8	0.1	0.3		8.1	1	47	53	2.6	30	19
	RPKM	1.4				2.3	0.1					64		0.2						14	5.2	0.5	6.4	
Liver	TPM	155	38.2	212	0.9	72	0.2	0.2	93	0.1	655	825	304	5.1	107	76	0.1	5.5		11	29	0.3	0.1	21
	RPKM	18.8		30.6		2.9					133	78	73	1.3	51	8.8				6	3.9			
Lung	TPM	0.3				1					0.8	0.6				0.3			1.5				1.6	7
	RPKM										0.5												0.2	
Lymph node*	TPM	0.3								0.1					0.3	1.8				0.1			5.4	6
	RPKM																							
Ovary	TPM	0.3																			0.2		0.6	3
	RPKM																							
Pancreas	TPM											2.1			0.4	0.2							1.3	4
	RPKM											3			5.7								1.9	
Parathyroid gland*	TPM	0.3														0.1				3.5				3

(continued on next page)

Table 1 (*continued*)

Tissues	mRNA levels of UGT genes in human tissues																					
		1A1	1A3	1A4	1A5	1A6	1A7	1A8	1A9	1A10	2B4	2B7	2B10	2B11	2B15	2B17	2B28	2A1	2A2	2A3	3A1	3A2	8	No %
	RPKM																							
Pituitary gland#	TPM																	1.6	5.6		0.2	0.7	0.1	5
	RPKM																							
Placenta*	TPM	0.2				0.1						0.2								8			0.4	5
	RPKM																							
Prostate	TPM	0.3	0.4			0.8		0.1	0.1		0.4	0.5			0.3	0.2					0.5	0.7	0.7	12
	RPKM	0.1																					0.1	
Rectum*	TPM	2.7				1.5	0.1	9.7	0.2	39		7.8	0.3		6.2	151				24			26	12
	RPKM									0.2														
Salivary gland	TPM	0				0.2	0.4			0.2		0.7		0.8		0.1	1.4						2.7	9
	RPKM	0.2				0.1						1.4				0.1							2.5	
Seminal vesicle*	TPM	0.5				5				0.1		40		0.2	0.1	1.1		0.2		0.1	0.3		0.9	11
	RPKM																							
Skeletal muscle	TPM																				0.9		0.1	2
	RPKM																				0.7			
Skin	TPM	2.2	0.3		0.1	13.1	4.1			1.1		0.3		0.1	0.2				0.2		0.1	12	2.6	13
	RPKM	0.3				0.3	0.2					0.1										5.1	0.3	
Small intestine	TPM	81.9	5.6	2.5	6.8	2.7	0.1	4.9		59		70		0.5	20.6	522			0.2	86	7		26	16
	RPKM	3.1		0.7		0.3				12.3		7.3			20	115				31	0.2		5.7	
Smooth muscle*	TPM	0.2						1.6		3.1		0.5			0.7	15.9		0.1		1.3			2.4	9
	RPKM																							
Spleen	TPM	0.4													0.1	0.8						1.5	2	5
	RPKM															0.4						0.3	0.8	
Stomach	TPM	2.6	0.1			8.9			0.2	31		0.7			9.4	0.7				0.8			14	10
	RPKM	0.4				1				1.3					10.7	0.1				0.2			3.5	
Testis	TPM	1.2				0.6	0.1			0.5	0.5	0.3	0.5	0.2	0.7	0.1		0.1			9	4.6	8.5	14
	RPKM	0.1				0.5							0.5		0.3						4.8	2.8	1.6	
Thyroid gland	TPM	0.3				0.1								0.8							0.2		19	5
	RPKM													0.3									1.8	
Tonsil	TPM	0.3				5.2	4.6	0.4		0.2		0.2			0.1	2.2						0.1	4.2	10
	RPKM																							
Urinary bladder	TPM	4.2	0.4	0.2		107	0.2	22.7	0.1	1.1		3.4		0.2	0.1	1	0.8					0.1	0.8	15
	RPKM	4.5				4.6	0.5			0.6		0.3			0.7		0.9						0.1	
Vagina#	TPM	0.3				0.2	0.4																0.5	4
	RPKM																							
No of tissues expressed $		36	11	5	5	27	17	13	10	20	9	28	6	19	24	29	5	6	6	15	15	17	42	

The expression levels of *UGT* genes in human tissues were obtained from RNA-seq datasets from the HPA (Fagerberg L et al, 2014) and GTEx (**GTEx** Consortium, 2013) that is available at The Human Protein Altas (https://www.proteinatlas.org). Data for three UGT genes (1A3, 1A5, 1A8, 1A9) are not available in the GTEx dataset.

TPM: Transcript Per Million; RPKM: Reads Per Kilobase gene model and Million mapped reads. * Tissues available from HPA RNA-seq dataset only; # Tissues available from GTEs only. $ Total number of tissues expressing each UGT gene; % Total number of UGT genes expressed in each tissue.

Table 2
Alternatively spliced transcripts of UGT genes classified by different types of splicing events.

													NCBI Reference Sequences (RefSeq, May 2019)	
	Cassette exons	Wild-type Transcript	Total alternative transcripts	Cassette alternative exons	Alternative 5' splice sites	Alternative 3' splice sites	Intron retention	Internal exonic intronization (Exitron)	Mutually exclusive exons	Alternative promoter and/or first exon	Alternative poly A site and/or terminal exon	Chaotic transcripts	transcript	protein
UGT1A gene	17	9	56							13	?			
UGT1A1	5	v1	5	n2	n1	n3					v2, v3		v1 (NM_000463.2)	NP_000454 (533 aa)
UGT1A3	5	v1	4		n2			n1			v2, v3		v1 (NM_019093.2)	NP_061966 (534 aa)
UGT1A4	5	v1	5	n2	n1						v2, v3	n3	v1 (NM_007120.2)	NP_009051 (534 aa)
UGT1A5	5	v1	2								v2, v3		v1 (NM_019078.1)	NP_061951 (534 aa)
UGT1A6	5	v1	6	n4							v2, v3	n1, n2, n3	v1 (NM_001072.4)	NP_001063 (532 aa)
UGT1A7	5	v1	3								v2, v3	n1	v1 (NM_019077.2)	NP_061950 (530 aa)
UGT1A8	5	v1	3		n1						v2, v3		v1 (NM_019076.4)	NP_061949 (530 aa)
UGT1A9	5	v1	3	n1									v1 (NM_021027.3)	NP_066307 (530 aa)
UGT1A10	5	v1	8	n1,n2, n3								n4, n5, n6	v1 (NM_019075.2)	NP_061948 (530 aa)
Other UGT1A			8							n1, n2, n3, n4, n5, n6, n7		n5, n8		
UGT2B4	6	v1	16	v5, v6,	v11	v8		v9, v10		v7, v12, n1	v2, v3, v4, n4		v1 (NM_021139.3) v5 (NM_001297615.1) v8 (NM_001297616.1)	NP_066962 (528 aa)_i1 NP_001284544 (369aa)_i5 NP_001284545 (392 aa)_i8
UGT2B7	6	v1	44	v6, v7, v22, n3, n4, n7, n10,	n1			n2	E1a, 1b, 1c, 1d (mutually exclusive with E1)	v2, v3, v4, v8, v9, v10, v11, v12, v13, v14, v15, v16, v17, v18, v19, v20, v21, n12, n13, n14, n15,	v5,	n5, n6, n11, n16, n17, n18, n19, n20, n21, n22	v1 (NM_001074.4) v7 (NM_001330719.1) v4 (NM_001349568.1) v6 (NM_011532229.2)	NP_001065 (529aa)_i1 NP_001317648 (369aa)_i4 NP_001336497 (280aa)_i5 XP_011530531 (450aa)_i3
UGT2B10	6	v1	10	n6, n7	n1, n2, n3, n4	n5					n8	n9, n10	v1 (NM_001075.6) n4 (NM-001144767.2) n3 (NM_001290091.1)	NP_001066.1 (528 aa)_i1 NP_001138339.1 (444aa)_i2 NP_001277020.1 (280 aa)_i3
UGT2B11	6	v1	8	n2	n4, n5, n6, n7			n8			n1, n3	n9	v1 (NM_001073.2)	NP_001064.1 (528 aa)_i1
UGT2B15	6	v1	8	n3, n4, n5, n6				n8			n1, n2,	n7	v1 (NM_001076.4)	NP_001067.2 (530 aa)_i1
UGT2B17	6	v1	9	n6, n7				n5		n1, n2, n3, n4		n8, n9	v1 (NM_001077.3)	NP-001068.1 (530 aa)_i1

(continued on next page)

Table 2 (*continued*)

	Cassette exons	Wild-type Transcript	Total alternative transcripts	Cassette alternative exons	Alternative 5' splice sites	Alternative 3' splice sites	Intron retention	Internal exonic intronization (Exitron)	Mutually exclusive exons	Alternative promoter and/or first exon	Alternative poly A site and/or terminal exon	Chaotic transcripts	NCBI Reference Sequences (RefSeq, May 2019)	
													transcript	protein
UGT2B28	6	v1	6	v2	n1			v3			n3, n4	n2	v1 (NM_053039.1) v2 (NM_001207004.1)	NP_444267.1 (529 aa)_i1 NP_001193933.1 (335 aa)_i2
UGT2A1	6	v1	3	v2						n2		n1,, n3	v1 (NM_006798)	NP_006789 (527 aa)_i1
UGT2A2	6	v1	0	v2									v1 (NM_001105677) v2 (NM_001301233.1)	NP_001099147.2 (536 aa)_i1 NP_001288162.1 (492aa)_i2
UGT2A3	6	v1	20	n4, n8, n10, n11, n12, n13, n14, n15, n16, n19	n1, n2							n5, n6, n7, n9, n17, n18, n20	v1 (NM_024743) v2 (NM_001168316.1)	NP_777574.2 (523 aa)_i1 NP_00161788.1 (489 aa)_i2
UGT3A1	7	v1	1									v2	v1 (NM_152404.3) v2 (NM_001171873.1)	NP_689617.3 (523 aa)_i1 NP_00165344.1 (252 aa)_i2
UGT3A2	7	v1	1	v2									v1 (NM_174914.4) v1 (NM_001168316.1)	NP_777574.2 (523 aa)_i1 NP_00161788.1 (489 aa)_i2
UGT8	7	v1, v3, v4, v5	5							v2			v1 (NM_001128174)	NP_1121646 (541 aa) _i1

In this table we have used the isoform nomenclature defined by Tourancheau et al. (2016) including "v", "n", and "i" designations where "v" referred to classical and previously published variant transcripts and "n" referred to novel putative transcripts that they reported for the first time in that study. The term "i" referred to protein isforms encoded by the relevant transcripts.

deregulated in cancerous tissues. An analysis of RNA-seq data from a human liver panel revealed high interindividual variability in the expression levels of variant UGT transcripts in human liver (Tourancheau et al., 2018).

Of the 10 *UGT* genes analysed by these studies, *UGT1* and *UGT2B7* genes undergo the most alternative splicing events. The *UGT1* locus has 33 exons in total (17 constitutive exons and 16 alternative exons) and generates at least 65 potential alternatively spliced transcripts. The *UGT2B7* locus has 19 exons (6 constitutive exons and 13 alternative exons) and generates potentially 44 alternatively spliced transcripts (Rouleau et al., 2016). The alternatively spliced transcripts of all *UGT* genes that have been reported and/or annotated as NCBI Reference Sequences (RefSeq) (accessed, June 2019) are summarized in Table 2. These variant transcripts may involve 1) cassette alternative exons, 2) alternative 5' splice sites, 3) alternative 3' splice sites, 4) alternative promoter and first exons, 5) alternative terminal exon, 6) exitrons, and 7) chaotic transcripts involving at least two different types of alternative splicing events. Variant transcripts containing retained introns have not yet been reported for *UGT* genes. Over 90 % of variant UGT transcripts contain an open reading frame (ORF) potentially encoding variant UGT proteins (Tourancheau et al., 2018). Alternatively spliced transcripts generated by the same type of splicing events tend to have similar biological consequences. In this section, we subclassify alternatively spliced UGT transcripts according to different types of alternative splicing events (Table 2) and discuss their potential biological functions.

2.5.1. Cassette alternative exons

Alternative splicing that involves cassette alternative exons refers to the skipping of one or more internal constitutive exons and/or the inclusion of one or more alternative exons (exons located within an intronic region). This is the most prevalent type of alternative splicing event, accounting for approximately 30 % of all alternative splicing (Roy et al., 2013). As listed in Table 2, almost all *UGT* genes produce this type of alternatively spliced transcript, which primarily occurs via skipping of constitutive exons.

For *UGT1*, the skipping of exon 2 was found in five transcripts (1A1_n2, 1A4_n2, 1A6_n4, 1A9_n1, 1A10_n3) The inclusion of one (1A10-n1) or two (1A10-n2) alternative exons was also reported (Tourancheau et al., 2016).

For the *UGT2A* genes, the skipping of exon 2 (2A3_n13), exon 3 (2A1_v2, 2A2_v2, 2A3_n12), exon 5 (2A3_n14), and exons 2/3/4 (2A3_n15) were reported (Bushey & Lazarus, 2012; Bushey, Dluzen, & Lazarus, 2013). The inclusion of one (2A3_n4, 2A3_n8, 2A3_n9) or two (2A3_n10, 2A3_n11) alternative exons were also found (Tourancheau et al., 2016).

Within the *UGT2B* genes, skipping of exon 3 (2B15_n6, 2B17_n6), exon 4 (2B15_n5), exon 5 (2B4_v5, 2B7_v7, 2B10_n7), exons 2/3 (2B15_n6), exons 4/5 (2B4_v6, 2B17_n7, 2B15_n3, 2B10_n6, 2B28_v2), and exons 2/3/4 (2B15_n4), and the inclusion of one alternative exon (2B7_v6, 2B7_n4, 2B7_n10, 2B11_n2) was reported (Tourancheau et al., 2016).

Of the above-mentioned alternative transcripts, only a few have been studied functionally. The consequences of exon skipping or alternative exon inclusion depends largely on whether it changes the reading frame of the coding sequence. As the majority of *UGT* gene exons are not comprised of a multiple of 3 nt, the skipping of a constitutive exon or inclusion of an alternative exon usually alters the reading frame such that a premature stop codon is introduced, thus generating C-terminally truncated proteins. The biological function of such truncated UGT enzymes has been best studied using the exon 5-skipped transcripts UGT2B4_v5 (Levesque et al., 2010) and UGT2B7_v7 (Menard et al., 2011; Menard, Collin, Margaillan, & Guillemette, 2013). Wild-type UGT2B4 (528 aa) and UGT2B7 (529 aa) enzymes are encoded by six constitutive exons; variant transcripts UGT2B4_v5 and UGT2B7_v7 lack exon 5 and both encode a C-terminally truncated 369 aa protein (UGT2B4_i5, UGT2B7_i4, respectively). Both variant enzymes are enzymatically inactive but can inhibit their wild-type counterparts as discussed in Section 4 (Levesque et al., 2010; Menard et al., 2011; 2013). Wild-type UGT2B7 (2B7_v1) and variant (2B7_v7) transcripts are coexpressed *in vivo* in human liver, kidney, and small intestine. Furthermore, wild-type (2B7_i1) and variant (2B7_i4) proteins colocalize in the microsomal/Golgi fraction, although UGT2B7_i4 protein has also been identified in the cytosol (Levesque et al., 2010).

A few constitutive *UGT* exons comprise a multiple of 3 nt, thus their skipping does not alter the reading frame but can lead to loss of an internal protein region. Such UGT exons include exon 2 (132 bp) of *UGT1*, and exon 3 (132 bp) of *UGT2Bs* and *UGT2As*. Skipping of exon 2 was found in five *UGT1* transcripts (1A1_n2, 1A4_n2, 1A6_n4, 1A9_n1, 1A10_n3) and skipping of exon 3 was also reported in three *UGT2A* transcripts (2A1_v2, 2A2_v2, 2A3_n12) and two *UGT2B* transcripts (2B15_n6, 2B17_n6) (Bushey & Lazarus, 2012; Bushey et al., 2013; Tourancheau et al., 2016). These exons are highly conserved among the *UGT* family members and encode a 44-amino acid region between substrate and co-substrate binding domains. The function of this type of variant protein has been studied using UGT2A1Δex3 (2A1_i2) and UGT2A2Δex3 (2A2_i2) (Bushey & Lazarus, 2012). Briefly, UGT2A1_i2 and UGT2A2_i2 are enzymatically inactive and can inhibit wild-type UGTs (see Section 4). The UGT2A1_v2 transcript is expressed at varying levels in many tissues (lung, trachea, larynx, tonsil, and colon) with the expression ratio of UGT2A1_v2/UGT2A1_v1 being the highest in colon (0.79 ± 0.08) and lung (0.42 ± 0.12). The expression of UGT2A1_i2 was also detected in many tissues with the expression ratio of UGT2A1_i2/UGT2A1_i1 in the range of 0.5 to 0.9 (Bushey & Lazarus, 2012). As UGT2A1 can glucuronidate tobacco-related carcinogens, the abundant expression of UGT2A1_i2 in lung and other relevant tissues may impair this detoxification capacity (Bushey & Lazarus, 2012).

Some alternatively used *UGT* exons also comprise a multiple of 3 nt and hence their inclusion could produce variant proteins with novel in-frame internal regions. Such exons include the *UGT2A1/2* exon 1b (69 bp), *UGT2A1/2* exon 2b (132 bp), *UGT2A3* exon 3b (63 bp), *UGT2B7* exon 1b (132 bp), *UGT2B7* exon 2b (96 bp), *UGT2B7* exon 5b (48 bp), and *UGT2B11* exon 5b (57 bp) (Tourancheau et al., 2016). This possibility is best exemplified by UGT2B7_n4 which has alternative exon 2b inserted between exon 1 and exon 2. This leads to the insertion of a novel 32-aa segment between the substrate- and co-substrate-binding domains, resulting in the UGT2B7_i8 variant protein (Rouleau et al., 2016; Tourancheau et al., 2016). UGT2B7_n4 is abundant in liver, where its expression can reach up to 75 % of wild-type UGT2B7_v1 levels, as well as in kidney and other tissues. Targeted mass spectrometry (MS) analysis has detected peptides specific to the sequence encoded by exon 2b in human livers (Tourancheau et al., 2016). Functions of UGT2B7_i8 are discussed in Section 4.

2.5.2. Alternative terminal exons

Another common type of alternative splicing involves the substitution of a constitutive terminal exon with an alternative terminal exon, which produces a transcript that typically encodes a C-terminally truncated protein. This alternative splicing has been reported for six *UGT* genes: *UGT1* (Girard et al., 2007; Levesque, Girard, Journault, Lepine, & Guillemette, 2007), *UGT2B4* (Levesque et al., 2010), *UGT2B7* (Menard et al., 2011, 2013), *UGT2B10* (Labriet et al., 2018), *UGT2B15* (Tourancheau et al., 2016), and *UGT2B28* (Tourancheau et al., 2016). Studies have investigated potential functions of this type of variant transcript as summarized below.

The *UGT1* gene has an alternative terminal exon 5b (2085 bp). The replacement of the constitutive exon 5a with exon 5b generates a set of nine variant transcripts (UGT1A_v2s). A 134-nt segment at the 5'-end of exon 5b can also be separately spliced between exon 4 and exon 5a, generating another set of nine variant transcripts (UGT1A_v3s). The splicing of exon 5b after exon 4 creates a short in-frame coding sequence with a premature stop codon that is present in both UGT1A_v2s and _v3s. Therefore, UGT1A_vs2 and _v3s have the

same open reading frame and encode the same nine variant proteins (termed UGT1A_i2s) with a novel 10-aa C-terminal peptide (RKKQQSGRQM) replacing the wildtype 99-aa C-terminal region encoded by exon 5a (Girard et al., 2007; Levesque et al., 2007). Wildtype (_v1 s) and variant (_v2s, _v3s) transcripts are co-expressed in many human tissues including liver, kidney, colon, small intestine, esophagus, testis, prostate, bladder and cervix (Girard et al., 2007). However, RT-qPCR using isoform-specific primers showed that the _v1 isoform expression of UGTs 1A1, 1A4, 1A6, and 1A9 is 16-, 17-, 57- and 29-fold higher than the levels of their respective variant _v2/_v3 transcripts in normal human liver specimens (Jones, Sun, Freeman, & Lazarus, 2012). Using isoform-specific antibodies, immunohistochemistry analysis has shown that UGT1A_i1 and _i2 proteins are co-expressed in major drug-metabolizing tissues, including liver, kidney, stomach, intestine and colon (Bellemare et al., 2011). UGT1A_i2 proteins lack the transmembrane domain and the ER retention domain and their presence in the ER is thought to be likely mediated by the novel 10-amino acid C-terminal peptide that has a dilysine motif KKXX and four positively charged residues (Girard et al., 2010). These proteins can interact with and inhibit other UGTs as discussed in Section 4.

The *UGT2B4* gene has four alternative terminal exons 6b/AG_1 (142 bp), 6b/AG_2 (140 bp), 6c (28 bp), and 6d (216 bp). Exons 6b/AG_1 and 6b/AG_2 end at the same nucleotide but 6b/AG_1 begins two nucleotides (AG) upstream of 6b/AG_2. The replacement of the constitutive terminal exon 6a with these alternative terminal exons generates four variant transcripts 2B4_v2 (6b/AG_1), 2B4_v3 (6b/AG_2), 2B4_v4 (6c) and 2B4_n4 (6d) that encode the C-terminally truncated UGT2B4 proteins 2B4_i2 (485 aa), 2B4_i3 (489 aa), 2B4_i4 (445 aa), and a fourth unnamed variant, respectively (Levesque et al., 2010; Tourancheau et al., 2016). Wild-type UGT2B4 transcript (2B4_v1) is expressed in fetal and adult liver, kidney and colon, but variant 2B4_v2 and v3 are only expressed in fetal and adult liver (Levesque et al., 2010). Compared to wild-type UGT2B4 protein (2B4_i1), these variant proteins lack the transmembrane domain and the ER retention sequence; they are enzymatically inactive and can inhibit wild-type UGT2B4_i1 proteins as described in Section 4 (Levesque et al., 2010).

The *UGT2B7* gene has two alternative terminal exons: 6b (64 bp) and 6c (59 bp). The splicing of exon 6b instead of the constitutive terminal exon 6a generates variant transcript UGT2B7_v5 that encodes the variant UGT2B7 protein, 2B7_i2 (457 aa). Transcript UGT2B7_v5 is co-expressed with its wild-type counterpart (2B7_v1) in HepG2 and C3A cell lines and many human tissues such as liver, kidney, colon, and duodenum (Menard et al., 2011, 2013). UGT2B7_i2 proteins lack the transmembrane domain and the ER retention domain, yet are co-localized in the microsomal/Golgi fraction with UGT2B7_i1 proteins (Menard et al., 2013). Retention in the ER membrane may be mediated by the two positively charged residues lysine and arginine ($_{364}$DIK$^+$R$^+$M$_{369}$L) and the C-terminal sequence $_{444}$IAASCGNCMFMK$_{456}$ (Menard et al., 2013). UGT2B7_i2 lacks glucuronidation activity and can alter wildtype UGT2B7 activity (see Section 4).

2.5.3. Alternative promoter and first exons

At least five *UGT* genes (*UGT1A1/2*, *UGT2B4*, *UGT2B7*, *UGT2B17* and *UGT8*) are reported to have multiple promoters (Levesque et al., 2010; Menard et al., 2011, 2013; Tourancheau et al., 2016). The impact of different *UGT2B4* and *UGT2B7* promoters have been studied and are discussed below.

The *UGT2B4* gene has two promoters: the classical promoter upstream of the canonical exon 1 (1a), and an alternative promoter upstream of alternative exons 1b (790 bp) and 1c (79 bp), which are 29,825 bp and 23,923 bp upstream of exon 1a respectively. The classical promoter generates the wildtype transcript (2B4_v1) and the alternative promoter generates four variant UGT2B4 transcripts (2B4_v7, 2B4_v8, 2B4_v12, 2B4_n1) (Levesque et al., 2010; Tourancheau et al., 2016). UGT2B4_v7 lacks exon 1a which means exon 1b is directly spliced to exons 2–6, whereas UGT2B4_v8 contains part of exon 1a spliced between exon 1b and exons 2–6. UGT2B4_v12 contains only three exons (1b/2/3) and UGT2B4_n1 contains only exons 1b and 1c. UGT2B4_v7 and UGT2B4_v8 are expressed in adult and fetal liver, and in kidney. UGT2B4_v7 and _v8 encode variant proteins UGT2B4_i7 (302 aa) and _i8 (392 aa), which lack N-terminal regions of the protein. In both UGT2B4_i7 and _i8, exon 1b encodes a putative 14-residue signal peptide at the N-terminal end. Both proteins are inactive but may be involved in protein interactions as discussed in Section 4.

The *UGT2B7* gene has two mutually exclusive promoters defined as promoter 1 (canonical) and promoter 1a (alternative) (Menard et al., 2011, 2013). Promoter 1a is situated approximately 44 kb upstream from the promoter 1. Four alternatively spliced exon 1 s (termed 1a, 1b, 1c, and 1d) are located between the two promoters. Exon 1b is the closest to promoter 1, residing 6871 bp upstream. Transcription directed by promoter 1 generates the classic exon 1-containing transcripts including the wild-type transcript (2B7_v1) as well as 15 exon 1-containing variant transcripts. Transcription driven by promoter 1a generates 28 alternatively spliced transcripts that lack the classical exon 1 but contain at least one of the four alternatively spliced exon 1 s (Table 2) (Tourancheau et al., 2016). Most of the exon 1b-containing alternative transcripts have a predicted ORF that starts with the first in-frame ATG in exon 2. Of these variant transcripts, those with an alternative exon 1 spliced to the constitutive exons 2–6a (_v4, _v12, _v14, _v16, _v17, _v20) encode UGT2B7_i5 proteins (280 aa). Transcripts lacking exon 5 at the 3′ end (_v9, _v13, _v15, _v19, _v21) encode variant UGT2B7_i6 proteins (208 aa). Transcripts with alternative terminal exon 6b relacing the classical terminal exon 6a (_v8, _v18) encode variant UGT2B7_i7 proteins (120 aa) (Menard et al., 2011, 2013). Therefore, these three variant UGT2B7 proteins are N-terminally truncated and have a common N-terminus encoded by exons 2–4 but unique C-termini. *in vitro* translation experiments showed stable expression of UGT2B7_i5 and _i6 in HEK293 cells (Menard et al., 2011). Consistent with the loss of the N-terminal exon 1-encoded substrate-binding domain, UGT2B7_i5 and _i6 proteins are enzymatically inactive towards UGT2B7 substrates (Menard et al., 2011).

Promoters 1 and 1a regulate the transcription of the *UGT2B7* gene in a tissue-, developmental-stage-, and disease-specific manner (Menard et al., 2013). Transcripts containing exon 1 are abundant in the liver and gastrointestinal tract, whereas transcripts containing exon 1b are predominantly expressed in other extrahepatic tissues (e.g. breast, bladder, testis, trachea, and thymus) as well as in fetal tissues (kidney, lung) and kidney tumor samples. Similar differential expression is seen in liver and kidney cell lines (Menard et al., 2013).

UGT8 has multiple exons of which only 5 are protein coding. Like many developmentally important genes, the first constitutive exon (s) are non-coding and the initiation codon is contained within the designated constitutive exon 2. Currently, *UGT8* is known to generate five different transcripts with different combinations of non-coding 5' exons. However, all of these transcripts contain the coding exons 2–6 (Fig. 4) and hence encode the same 541-aa wildtype UGT8 protein (NP_00112646.2). The constitutive promoter (used in tissues where UGT8 is most abundantly expressed such as the central nervous system) is located upstream of exon 1a (Fig. 4). Three alternatively used exons designated 1a, 1b and 1c located downstream of this promoter are responsible for generating different 5'UTRs. The majority of transcripts (v1, v3 and v4) initiate at the start of exon 1a. The length of exon 1a is variable in transcripts v1 (168 nts), v3 (573 nts), and v4 (565 nts), as transcription of this exon can start at two different positions [v1 (114,598,807), v3/v4 (114,598,402). Exon 1a length is also modified by variable usage of two different donor splice sites during splicing to exon 2: v1/v3 at 114,598,974 and v4 at 114,599,057 (GRC38/hg38). The transcript designated v5 initiates at the beginning of exon 1b. Thus altogether, the canonical *UGT8* promoter initiates transcription in at least three different positions, generating four transcripts (v1, v3, v4, v5). The transcript v2 is likely controlled by an alternative promoter

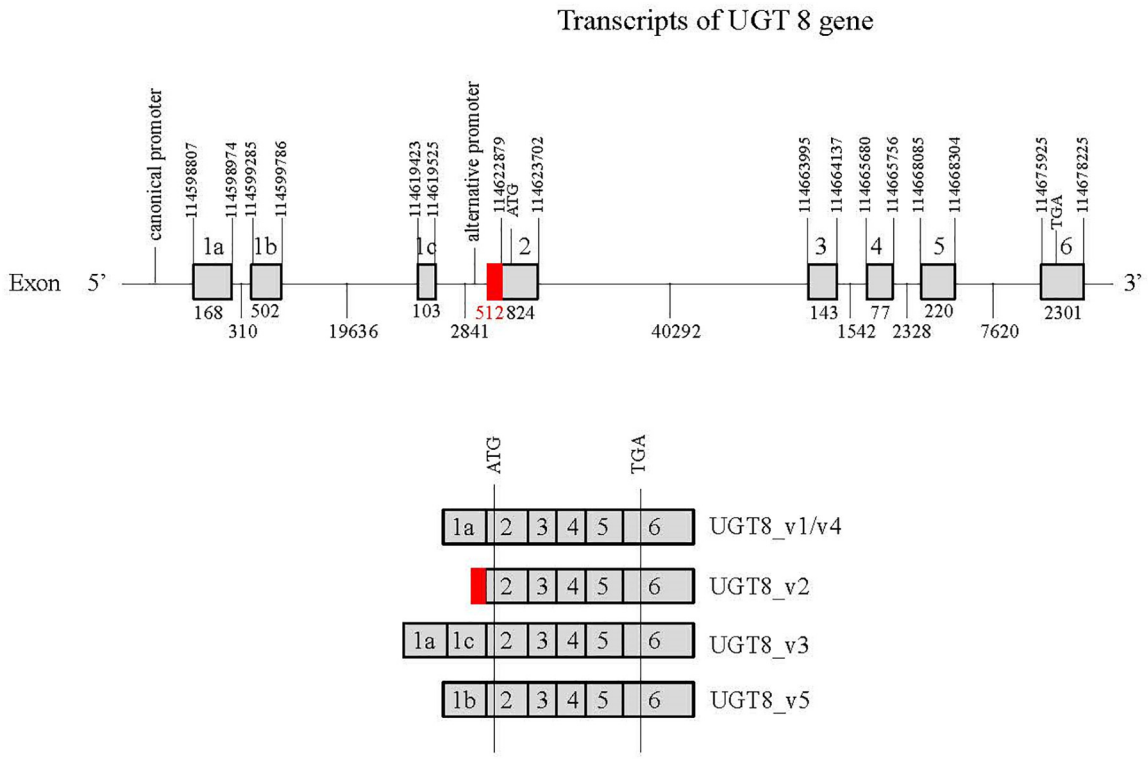


Fig. 4. *UGT8* gene transcription is controlled by two promoters that generate five different transcripts via alternative splicing. A promoter located upstream of exon 1a appears to drive transcription initiation 5′ of exons 1a and 1b; this promoter is constitutively used in most UGT8-expressing tissues. A rarely used alternative promoter is present upstream of the first coding exon (exon 2). The transcripts generated by these different promoters have different 5′UTR sequences but all encode identical full length UGT8 proteins.

that initiates transcription 512 nucleotides upstream of exon 2, thus extending exon 2 and creating yet another novel 5'UTR. It is notable that unlike other UGTs, skipping or inclusion of alternative internal exons that could lead to production of C-terminally truncated, inactive forms of UGT8 has not been reported; this is likely due to the essential role of the UGT8 protein in neural development.

2.5.4. Alternative 5' splice sites and alternative 3' splice sites

Constitutive exons of human multi-exon genes are flanked by highly conserved 5' and 3' splice sites that ensure their splicing into mature mRNAs coding for wild-type functional proteins (Consortium, 2012). However, in many genes there are weak 5'and 3' splice motifs present within exonic and intronic sequences, and these may act as alternative 5' and 3' splice sites (Blencowe, 2006; Roy et al., 2013; Sibley et al., 2015). The use of alternative intronic splice sites adds intronic sequence to the 5' (3' weak splice site) or 3' (5' weak splice site) ends of the exons, whereas the use of exonic alternative splice sites removes exonic sequence from the 5' (3' weak splice site) or 3' (5' weak splice site) ends of the exons. Nearly one-quarter of known alternative spliced transcripts are generated through alternative 5' and 3' splice sites (Blencowe, 2006). Exonic alternative 5' and 3' splice sites are normally present in long exons. Tourancheau A et al. reported alternative UGT transcripts generated using alternative 5' and 3' splice sites that are mainly located in the largest exons (exon 1, exon 6) of multiple *UGT* genes as summarized below (Tourancheau et al., 2016).

Using the classical 5' splice site and an alternative exon 1-exonic 3' splice site, one variant transcript is found for UGT2B4 (2B4_v8). Using the classical 5' splice site and an alternative exon 2-exonic 3' splice site, one variant transcript is found for UGT2B10 (2B10_n5). Using the classical 3' splice site and an alternative exon 1-exonic 5' splice site, variant transcripts are found for *UGT1* (1A1_n1, 1A3_n2, 1A4_n1, 1A8_n1), *UGT2B4* (2B4_v11), *UGT2B7* (2B7_n1), and *UGT2B28* (2B28_n1). Using the classical 3' splice site and an alternative exon 1-exonic 5' splice site, four variant transcripts are found for *UGT2B10* (_n1, _n2, _n3, _n4) and *UGT2B11* (_n4, _n5, _n6, _n7) genes. Because most of these variant transcripts lack part of an exon, they usually have disrupted ORFs. No functional analysis of this type of alternative UGT transcripts has been reported (Tourancheau et al., 2016).

2.5.5. Exitrons (Internal exonic intronization)

Internal exonic intronization refers to the removal of an internal part of an exon from a protein-coding mRNA. With respect to the wild-type mature mRNA, the spliced-out exonic region can be considered as a retained intron. These introns are termed exitrons (exonic introns) (Marquez et al., 2012, 2015). A genome-wide screening identified 923 exitrons in the human genome; half of these with sizes of multiples of 3 nt. Exitrons are normally embedded within long exons and have weak flanking splice sites and thus tend to be retained under normal conditions. Splicing of exitrons out from mature mRNAs thus could generate internally deleted proteins and boost the diversity of the human transcriptome and proteome. Multiple *UGT* genes have variant transcripts containing an internal deletion within their exon 1 s, including UGT1A2P_n1, UGT1A3_n1, UGT2B4_v9, UGT2B4_v10, UGT2B7_n2, UGT2B11_n8, UGT2B15_n8, and UGT2B28_v3 (Table 2) (Tourancheau et al., 2016). In addition, exon 6 is also internally deleted in two variant UGT2B10 transcripts (2B10_n9, 2B10_n10). In general, the coding potential of these variant transcripts and their biological function have not been well investigated. However, the truncated proteins corresponding to the UGT2B10_n9 and _n10 forms (called _i4 and _i5) have been studied *in vitro*. These proteins have shorter half-lives than full length UGT2B10 and are inactive with known UGT2B10 substrates amitriptyline and levomedetomidine (Labriet et al., 2018). Taken together, these observations indicate that *UGT* genes have exitrons in their longest exons, namely exon 1 and exon 6.

2.5.6. Chaotic transcripts

Chaotic transcripts are generated by at least two different types of alternative splicing events (Park et al., 2018; Vaquero-Garcia et al., 2016).

Many chaotic UGT transcripts have been reported as listed in Table 2 (Labriet et al., 2018; Levesque et al., 2010; Rouleau et al., 2016; Tourancheau et al., 2016). Some of these transcripts have predicted ORFs encoding variant proteins. For example, UGT1A4_n3, UGT1A7_n1, and three UGT1A10 transcripts (_n4, _n5, _n6), contain only the isoform specific exon 1 and a novel 3' terminal exon. UGT2B7_n5 has three exons (1/2/2b), UGT2B11_n9 has three exons (1/2/2b), and UGT2B17_n9 has three exons (1/2/2b). These variant transcripts could encode proteins containing only the N-terminal half of the protein. These predicted proteins might compete for substrate binding, and/or interact with full length UGTs; however, this remains to be investigated.

2.5.7. Chimeric splicing

All alternative UGT transcripts as reviewed above contain exons that are located within the boundaries of single *UGT* gene loci. In addition to these canonical alternative transcripts, chimeric transcripts that contain exons from more than one *UGT* gene have been recently reported (Hu et al., 2018). Chimeric transcripts can be generated by trans-splicing, cis-splicing, or rarely chromosomal translocation (Chwalenia, Facemire, & Li, 2017). Trans-splicing allows exons from more than one primary transcript (often derived from different genes) to be spliced in any order into a single mature mRNA (Lei et al., 2016). Cis-splicing is the classical process where exons of a primary transcript derived from a single gene are spliced to form mRNA. Cis-splicing of adjacent genes (cis-SAGe) involves read-through transcription of adjacent genes generating multi-gene primary transcripts, followed by intergenic splicing to form chimeric mRNA (Kannan et al., 2011; Parra et al., 2006). Via cis-SAGe, it is reported that approximately 4 % of tandem genes (two adjacent genes in the same orientation) and 14 % of protein-coding genes in the human genome are involved in formation of chimeric transcripts encoding putative chimeric proteins (Kannan et al., 2011; Parra et al., 2006). Within a 165-kb region of the *UGT2B* cluster, *UGT2B15*, *UGT2B29P2*, *UGT2B17*, and *UGT2B29P1* are located adjacent to each other in a head-to-tail orientation (*UGT2B15-UGT2B29P2-UGT2B17-UGT2B29P1*) (Menard, Eap, Harvey, Guillemette, & Levesque, 2009; Turgeon et al., 2000). The *UGT2B15* and *UGT2B17* genes are controlled by their own promoters to generate single-gene mature mRNAs. As shown in Fig. 5, we recently found that the transcription of *UGT2B15* and *UGT2B17* can continue through the downstream genes

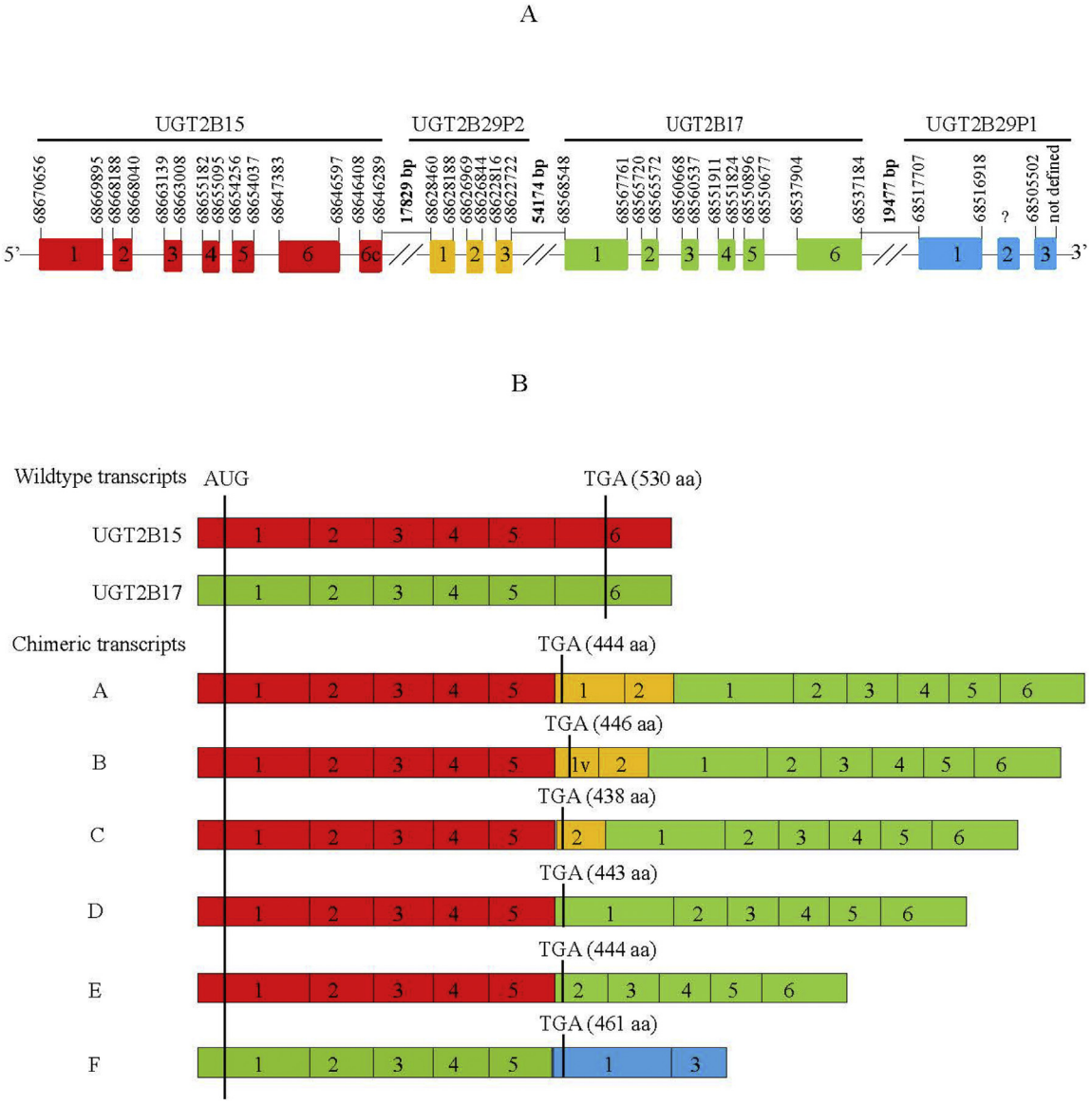


Fig. 5. Chimeric UGT transcripts produced by read-through transcription of four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) and subsequent intergenic splicing. (A) Genomic locations and structures of four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) in accordance with the human genome assembly GRCh38/hg38 (2013). (B) Exon [2B15 (red), 2B29P2 (orange), 2B17 (green), 2B29P1 (blue)] structures of wild-type and chimeric UGT2B15 and UGT2B17 transcripts. Initiation codon: AUG; stop codon: TGA; aa, amino acid.

to generate two-gene (*2B17-2B29P1*) or three-gene (*2B15-2B29P2-2B17*) primary transcripts. Via cis-SAGe, six chimeric transcripts are generated from these multi-gene primary transcripts. The splicing of three exons from *UGT2B29P2* or two exons from *UGT2B29P1* in the chimeric transcripts represents the first evidence for the transcription of these two pseudogenes. The chimeric transcripts have exons from 2 (chimeras D, E, F) or 3 (Chimeras A, B, C) genes (see Fig. 5). As the first exon always lacks an acceptor splice site and the last exon lacks a donor splice site, cis-SAGe often occurs between the second-to-the-last exon of the 5' gene and the second exon of the 3' gene (Chwalenia et al., 2017; Jia, Xie, & Li, 2016; Parra et al., 2006). This 2'-2' rule explains the generation of the chimera D and B as well as the lack of *UGT2B15* exon 6 in chimeras A, B, C, D and E, and the lack of *UGT2B17* exon 6 in chimera F. The exception to this 2'-2' rule is the splicing of *UGT2B29P1* exon 1 in chimera F and the *UGT2B29P1* exon 1 in chimeras C and E. This is because these exon 1 s are spliced into the chimeric transcripts using a cryptic acceptor splice site. All the canonical and variant exons of the four *UGT* genes in the chimeras conform to the "GT-AG" rule suggesting that they are generated by splicing and not genomic rearrangement or cloning artefacts.

All six chimeric transcripts contain the *UGT2B15* (or *UGT2B17*) exons 1-5 with exon 6 (3'-terminal exon) replaced by exons from one or two downstream *UGT* gene(s). The predicted ORFs have the canonical *UGT2B15* or *UGT2B17* ATG and terminate at a premature stop codon in the exons that replace the canonical exon 6; thus they each encode similar C-terminally truncated UGTs. The chimeric UGT2B15 protein A and chimeric UGT2B17 protein chimera F were stably expressed in HEK cells and as discussed in section 4, were found to interact with and inhibit other UGTs.

In addition to *UGT2B29P1* and *UGT2B29P2*, the *UGT2* gene cluster contains 5 other pseudogenes (*2B24P*, *2B25P*, *2B26P*, *2B27P*, and *2B28P*) (Turgeon et al., 2000). Whether these pseudogenes are involved in formation of similar chimeric transcripts with their neighbouring functional *UGT* genes remains to be investigated. In conclusion, the formation of chimeric transcripts among *UGT* genes further increases the diversity of the human UGT transcriptome and proteome.

3. Expanding the diversity of UGT functions by posttranslational modification of UGT proteins

Newly synthetized proteins are subject to a wide range of chemical modifications that begin during transport to their final cellular destinations and continue throughout the lifetime of the protein. Some of the more studied of these modifications include N- and O-linked glycosylation, palmitoylation, sumoylation, ubiquitination, biotinylation, phosphorylation, acetylation and methylation (Omenn et al., 2016). These modifications quite often alter the dynamics of protein folding and stability, as well as directly modulate protein function and signaling. To date, only N-linked glycosylation and phosphorylation have been reported as posttranslational modification processes for UGT proteins.

3.1. N-linked glycosylation

Potential sites for N-linked glycosylation were first described based on primary amino acid sequences of UGTs (Mackenzie, 1986). Subsequently, endoglycosidase H digestions of purified and recombinant proteins demonstrated the presence of oligosaccharide chains attached to asparagine residues in UGT proteins (Green & Tephly, 1989; Mackenzie, 1990). The glycan chain was not essential for catalytic activity and appeared not to alter the substrate preferences of each UGT (Mackenzie, 1990). However, it may be important for correct folding and maintenance of fully-active enzymes, as described for UGT1A9 (Nakajima et al., 2010) and UGT2B7 (Girard-Bock, Benoit-Biancamano, Villeneuve, Desjardins, & Guillemette, 2016). It may also modify UGT interaction with other proteins of the endoplasmic reticulum, such as cytochromes P450 (Nakamura et al., 2016). Although changes in substrate choice as a consequence of altered glycosylation have not been observed, N-linked glycosylation does appear to alter the kinetics of catalysis. Studies of UGT2B7 demonstrated that removal of N-linked glycans increased the affinity of the enzyme for morphine but decreased that for zidovudine, without affecting reaction velocities. The affinity for UDPGA as a cofactor was decreased irrespective of the substrate used (Nagaoka, Hanioka, Ikushiro, Yamano, & Narimatsu, 2012). In summary, there is little evidence to suggest that N-linked glycosylation contributes to the capacity of the UGT superfamily to conjugate thousands of structurally diverse chemicals with sugar moieties to render them less toxic and more amenable to excretion.

3.2. Phosphorylation

Another posttranslational modification that may impact UGT substrate selectivity is phosphorylation. Evidence that phosphorylation alters UGT activity was first shown for UGT1A10, where mutations of predicted protein kinase C phosphorylation sites greatly reduced UGT1A10 activity towards ß-estradiol (Basu et al., 2004). Subsequent studies with classic protein kinase C agonists and antagonists confirmed that at least 2 protein kinase C isozymes were involved in phosphorylating at least 10 UGTs, and that the effects on catalytic activity were dependent on the site phosphorylated (Basu et al., 2005). Phosphorylation of some protein kinase C sites abolished activity whereas phosphorylation of other sites caused a major change in pH optima and activity towards various substrates, which was UGT-specific (Basu et al., 2005, 2008). An example of the latter effect is shown for UGT1A7, where phosphorylation of serine-432 of the enzyme enhanced its activity towards SN38 but reduced its activity towards ß-estradiol about 10-fold (Basu et al., 2005). Whether the capacity to glucuronidate new substrates is induced by phosphorylation remains to be tested. It was suggested that the specificity of these effects on pH optima and activity may depend on the architecture of the active site and the positioning of the charged phospho-tagged serine or threonine residues within this site (Basu et al., 2005). However, most phosphorylation sites are on the periphery of the UGT molecule (as assessed by homology modelling) as would be expected to enable access to the protein kinase, and hence may affect substrate access rather than substrate catalysis.

In addition to phosphorylation of serine/threonine amino acids in UGTs by protein kinase C, phosphorylation of tyrosine has also been reported, as shown for UGT2B7 (Mitra, Basu, & Owens, 2009). In this case, mutation of the three protein kinase C sites in UGT2B7 had little effect on activity, whereas mutation of the two tyrosine kinase sites greatly decreased the activity of the enzyme. Of particular interest was the observation that two tyrosine phosphorylases were involved in phosphorylating UGT2B7 and that their actions had differential effects on activity. Phosphorylation by Src of UGT2B7 allows glucuronidation of 4-hydroxyestrone but not 17ß-estradiol, whereas phosphorylation by an unidentified, non-Src kinase allows the glucuronidation of both chemicals (Mitra et al., 2011). Hence, UGT2B7 substrate selection is altered, depending upon the tyrosine kinases that phosphorylate UGT2B7. UGT2B15 also has three predicted protein kinase C sites and two tyrosine kinase sites. As with UGT2B7, mutation of the tyrosine kinase sites resulted in a 80–90% loss of activity. In contrast to UGT2B7, mutation of two of the PKC sites resulted in 70–100% loss of activity, whereas mutation of the third site had little effect (Chakraborty et al., 2012).

In summary, phosphorylation appears to be necessary for optimal activity of a UGT and may differentially alter its activity towards a set of known substrates. Whether UGT phosphorylation leads to the glucuronidation of novel substrates remains to be determined.

4. Expanding the diversity of UGT functions through protein-protein interactions

4.1. Homo- and hetero-oligomerization of UGTs

The ability of UGTs to function as dimers or higher order oligomers has been long suggested and supported by multiple lines of evidence. This interaction could take the form of homo-oligomerization (proteins of the same isoform interacting) or hetero-oligomerization (different isoforms interacting). The former may be considered to include variants of the same isoform derived by alternative splicing, or by polymorphism. Of note, the subject of UGT dimerization was reviewed relatively recently by Fujiwara, Yokoi, and Nakajima, 2016), and hence this section will include a summary or reanalysis of some previously reviewed studies, as well as an update on new work in this area. Given that it may be argued that there are conceptual differences between studies performed using endogenous proteins, and those performed using heterologous overexpression systems, this section will be divided accordingly. In addition, interactions between variant UGT proteins resulting from genetic differences or alterative splicing, are discussed separately in Section 4.1.2.4.

4.1.1. Evidence for homo- and/or hetero-oligomerization of endogenous UGT proteins

4.1.1.1. Gel filtration and radiation inactivation methods. Gel filtration chromatography of purified proteins provided the earliest evidence that mammalian UGTs exist as oligomers *in vivo*, prior to the molecular cloning of any *UGT* gene. Tukey and Tephly purified UGT proteins that exhibited either estrone or p-nitrophenol glucuronidation activities from rabbit liver. These forms showed apparent molecular weights of 230 kDa based on gel filtration (Tukey & Tephly, 1981). Given that the proteins were found by SDS-PAGE to have monomeric molecular weights of ~53 kDa (consistent with later molecular cloning studies), the data suggested that UGTs might form tetramers. Other studies around this time purified a rat liver microsomal UGT capable of conjugating chenodeoxycholic acid and testosterone, and based on non-denaturing gradient gel electrophoresis predicted an apparent molecular weight of 316 kDa, also suggesting oligomerization (Matern, Matern, & Gerok, 1982).

The other main biochemical method used in early studies to interrogate oligomeric forms of native UGTs was radiation inactivation. The direct relationship between radiation-induced inactivation of the enzymes and both radiation dose and the size of the functional enzyme complex was used to determine molecular weight (Kempner & Schlegel, 1979). Using this approach, the molecular weights of UGTs catalyzing 1-naphthol, 6-hydroxychrysene, 3,6-dihydroxybenzo[*a*]pyrene, and 3,6-dihydroxychrysene were calculated to be 91–218 kDa (Gschaidmeier & Bock, 1994), suggesting dimers and/or tetramers. The radiation-inactivation studies of Peters *et al* indicated that a 41.5 kDa protein catalyzed bilirubin mono-glucuronidation and a 175 kDa protein (presumably a UGT complex) catalyzed di-glucuronidation (Peters, Jansen, & Nauta, 1984). The same study found that UGT complexes with molecular weights of 142 and 159 kDa catalyzed glucuronidation of testosterone and phenolphthalein, respectively. Overall, these data supported a model that UGTs can function in dimeric or higher order complexes *in vivo*

4.1.1.2. Immunoblotting, crosslinking, and immunoprecipitation methods. The development of UGT-specific antibodies allowed SDS-PAGE and immunoblotting analysis to become a mainstay of UGT-interaction studies. In general, UGT oligomers are dissociated by SDS-PAGE and not visible via immunoblotting. However, some studies have reported immunoreactive bands that migrate at the expected size of UGT dimers on SDS-PAGE; albeit at a small fraction of the monomer form, and variably sensitive to reducing agents. To facilitate analysis, a variety of chemical crosslinkers have been used to stabilize interactions. Using the homobifunctional cross-linker 1,6-bis(maleimido)hexane to treat rat hepatic microsomes, Ikushiro *et al* detected bands migrating at 50–60 kDa and at 120–130 kDa on SDS-PAGE that were immunoreactive with antibodies recognizing UGT1 or UGT2B1 (Ikushiro, Emi, & Iyanagi, 1997). Because the crosslinker interacts with sulfhydryl-groups, they postulated that amino acids containing these groups located on the outside of the proteins were important for dimerization.

Analysis of UGT heterodimers has been achieved largely by co-immunoprecipitation, with SDS-PAGE and immunoblotting generally used to identity interacting proteins, although other approaches have been used. For example, in the study by Ikushiro *et al* described above, proteins that co-immunoprecipitated with UGT1 isoforms from rat liver microsomes were identified by amino-terminal peptide sequencing. The major protein co-eluting with UGT1As was identified as UGT2B1 (a testosterone conjugating isoform in the rat), with UGT2B3 and UGT2B6 as minor interactors. The interaction between UGT2B1 and UGT1As was disrupted above pH 8.0, suggesting pH sensitive heterodimerization of UGTs from different families.

The first study showing endogenous interactions of human UGTs used isoform specific antibodies to immunoprecipitate UGT2B7, UGT1A1 or UGT1A6 from human liver microsomes (Fremont, Wang, & King, 2005). In each case, immunoblotting analysis showed that the precipitated UGT isoform robustly co-immunoprecipitated each of the other isoforms.

4.1.1.3. Functional assays of endogenous UGTs. Recently, Konopnicki *et al* studied functional interactions of endogenous UGTs in primary human hepatocytes using isoform-selective siRNA and probe substrates. Down-regulation of UGT2B7 using siRNA reduced UGT1A9 activity with its specific substrate propofol; the UGT1A9 mRNA level was not altered. Inhibition of UGT1A9 expression also produced minor changes in UGT1A4, UGT2B4 and UGT2B7 activity (Konopnicki et al., 2013).

Crigler–Najjar syndrome (CN) patient data has also provided support for functional interaction *in vivo*. CN is usually inherited as an autosomal recessive trait requiring mutations on both alleles. However, Koiwai *et al* (Koiwai et al., 1996) identified a CN type II (CN-II) patient with one wild-type allele and one allele carrying a nonsense mutation resulting in premature truncation. The mutant mRNA was not subject to nonsense mediated decay. The dominant-negative effect of the truncated UGT1A form on the wild-type protein was later confirmed *in vitro* (Suzuki et al., 2014) as discussed in Section 4.1.2. It has also been observed that patients with low activity UGT1A1 variants or hyperbilirubinemia have greater clearance of sorafenib which is metabolized by UGT1A9, suggesting a functional interaction of UGT1A1 and UGT1A9 *in vivo* (Peer et al., 2012).

4.1.2. Evidence for homo- and/or hetero-oligomerization of heterologously-expressed UGT proteins

4.1.2.1. Immunoblotting, immunoprecipitation methods and other affinity methods. The advent of molecular cloning allowed interactions between specific UGT isoforms to be interrogated using heterologous expression systems. Table 3 summarizes all physical interactions between heterologously expressed UGTs reported in literature. We first used this approach to study the dimerization of rat UGT2B1 in COS-1 cells (Meech & Mackenzie, 1997a). When the overexpressed protein was resolved by SDS-PAGE and immunoblotted, we observed bands at ~50 kDa and at ~110 kDa; we proposed that the latter were unusually stable UGT2B1 dimers. In support of the larger band representing a dimer, its size was reduced proportionally in a series of C-terminally-truncated mutants, and by deglycosylation of the UGT. Removal of the transmembrane domain and the C-terminal cytosolic tail (residues after amino acid 493 in UGT2B1) did not affect the appearance of this larger band on immunoblots, suggesting that these regions were not required for dimerization. In the same study, we probed the role of the amino terminal

Table 3
Physical Homo/heterooligomerization of UGTs shown using heterologous expression systems.

Isoform 1	Isoform 2	Interaction	Method	System	Reference
UGT1 (unspecified)	UGT2B1 UGT2B3, UGT2B6	Yes (major) yes (minor) Yes (minor)	Crosslinking with SDS-PAGE	Rat liver microsomes	Ikushiro, 1997
UGT2B7	UGT1A1	Yes	Co-IP and SDS-PAG	human liver microsomes	Fremont, 2005
UGT2B7	UGT1A6	Yes	Co-IP and SDS-PAG	human liver microsomes	Fremont, 2005
UGT1A1	UGT1A6	yes	Co-IP and SDS-PAG	human liver microsomes	Fremont, 2005
UGT1A1	UGT1A1	yes	Crosslinking with SDS PAGE 2-hybrid	Gunn rat fibrosblasts, COS-7	Ghosh, 2001
UGT1A1	UGT1A6	no	2-hybrid*	COS-7	Ghosh, 2001
UGT1A1	UGT1A1Δ152-180	no	2-hybrid*	COS-7	Ghosh, 2001
	UGT1A1:L175E UGT1A1:C223Y	No no	2-hybrid*	COS-7	Ghosh, 2001
UGT1A1	UGT1A1 -pQ331X	yes	Co-affinity purification and SDS-PAGE; AFM	HEK293T	Suzuki, 2014,
UGT1A6	UGT1A9	yes	Co-IP and SDS-PAGE	HEK293T	Fujiwara, 2007
UGT1A1	UGT1A1_i2	Yes	Co-IP and SDS-PAGE	HEK293T	Leveseque, 2007
UGT1A1_i2	UGT1A1_i2	yes	Co-IP and SDS-PAGE	HEK293T	Bellemare, 2010
UGT1A1_i1	UGT1A1_i2	Yes	Co-IP and SDS-PAGE	HEK293T	Bellemare, 2010
UGT1A6_i1	UGT1A6_i1	Yes			
UGT1A7_i1	UGT1A7_i1	yes			
UGT1A9	UGT2B7	yes	Affinity pulldown and SDS-PAGE	Insect cells	Kurkela, 2003
UGT1A9	UGT1A9	yes	Affinity pulldown and SDS-PAGE	Insect cells	Kurkela, 2003
UGT2B7	UGT2B7	yes	Co-IP and SDS-PAGE	HEK293T	Lewis, 2007
UGT2B7	UGT2B7	yes	FRET	Insect cells	Yuan, 2015
UGT1A1	UGT1A1	yes	FRET; CoIP and SDS-PAGE	Insect cells	Liu, 2016
UGT1A9	UGT1A9	yes	FRET; CoIP and SDS-PAGE	Insect cells	Liu, 2016
UGT1A9 SNP variants	UGT1A9 SNP variants	yes	FRET; CoIP and SDS-PAGE	Insect cells	Liu, 2016
UGT1A1	UGT1A1_i2 (UGT1A1*1b)	yes	FRET; CoIP and SDS-PAGE	Insect cells	Liu, 2016
UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9, UGT1A10	Self Self Self Self Self Self Self Self	Yes Yes Yes Yes Yes Yes Yes yes	FRET	COS7	Operana and Tukey, 2007
UGT1A1	UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9, UGT1A10	Yes Yes Yes Yes Yes Yes yes	FRET	COS7	Operana and Tukey, 2007
UGT2B7	UGT2B7 UGT1A1 UGT1A4 UGT1A6 UGT1A9	Yes Yes Yes Yes yes	Native PAGE; Co-IP and SDS-PAGE	HEK293T	Fujiwara, 2010
UGT2B10	UGT2B10_i4 UGT2B10_i5	Yes yes	Co-IP and SDS-PAGE	HEK293T	Labriet, 2018
UGT2B7	UGT2B7_i8	Yes	Co-IP and SDS-PAGE	HEK293T	
UGT2B15chim (truncated)	UGT2B15 UGT2B7	Yes yes	Co-IP and SDS-PAGE	HEK293T	Hu, 2018
UGT2B17chim (truncated)	UGT2B17 UGT2B7	Yes yes	Co-IP and SDS-PAGE	HEK293T	Hu, 2018
UGT1A10	UGT1A10	yes	Non-reducing SDS-PAGE	HEK293T CHO	Kim, 2018

domain in dimerization by generating chimeras between rat UGT2B1 and the baculoviral ecdysteroid glucosyltransferase (EGT). We found that proteins containing the N-terminal domain of UGT2B1 and the C-terminal domain of EGT (UGT/EGT) formed stable dimers on SDS-PAGE; however, proteins containing the N-terminal domain of EGT and the C-terminal domain of UGT2B1 (EGT/UGT) did not. This provided the first evidence that the amino terminal domain may mediate dimerization, which we went on to support using dominant-negative and functional rescue studies as discussed in Section 4.1.2.3 (Meech & Mackenzie, 1997b). Subsequently, several other groups used overexpression systems to study homo- and hetero-dimerization

using non-reducing SDS-PAGE, native PAGE, crosslinking, and immunoprecipitation/affinity purification.

Ghosh *et al* studied interactions of heterologously-expressed human UGT1A1 in Gunn rat fibroblasts that largely lack endogenous UGT1A proteins (Ghosh et al., 2001). Gel-filtration studies supported a dimeric form for the human UGT1A1 protein, and weak bands corresponding to the expected size of a dimer were seen on non-reducing SDS-PAGE. Both a disulfide cross-linker and an amino group cross-linker were able to stabilize human UGT1A1 homodimers such that around half of the protein migrated at the expected dimer size on SDS-PAGE; dimerization was inhibited in alkaline conditions (pH > 9.0). This study also used mammalian two hybrid analyses to study interactions between wild-type and/or mutant forms of UGT1A1 and UGT1A6. They found no interaction between UGT1A1 and UGT1A6 (which contrasts with several later studies). Deletion of an amino-terminal region between amino acids 152–180 prevented homodimerization of UGT1A1; this region includes a predicted membrane-embedded helix (Ciotti, Cho, George, & Owens, 1998). In addition, two point-mutations in UGT1A1: L175E and C223Y, appeared to disrupt homodimerization based on the two-hybrid reporter assay. An important caveat to this approach is that only dimer conformations that allow the fusion proteins (in this case yeast gal4 and mouse NFκb) to reconstitute a functional transcription factor, will result in a positive readout. This may explain why Ghosh *et al* did not detect UGT1A1-UGT1A6 interaction in the 2-hybrid assay, even though this has been consistently reported by multiple studies as discussed below. Nonetheless, the failure of the C223Y mutant to interact with wild-type UGT1A1 in two-hybrid analysis did seem to be supported by functional studies; specifically, this inactive mutant did not have a dominant-negative effect on wild-type UGT1A1 activity.

Fujiwara et al used isoform specific antibodies to show that UGT1A6 co-immunoprecipitated with UGT1A9 in the HEK293 T co-expression system (Fujiwara et al. 2007). They subsequently went on to study interactions between human UGT2B7 and various UGT1As in this system (Fujiwara et al., 2010). Native-PAGE analysis showed that UGT2B7 forms homo-oligomers, and both native PAGE and immunoprecipitation assays indicated that UGT2B7 could form hetero-oligomers with UGT1A1, UGT1A4, UGT1A6, and UGT1A9. Interestingly, co-expression of UGT1As increased the thermal stability of UGT2B7. The functional consequences of these interactions were supported by kinetic analyses as discussed in Section 4.1.2.3 (Fujiwara et al., 2010).

Tagging UGT proteins with various epitope tags has been a useful tool for validating homo-oligomerization. Homo-oligomerization of UGT1A9 was demonstrated by co-expressing UGT1A9 proteins appended with either a poly-Histidine (His) or a hemagglutinin (HA) tag, and then capturing the His-tagged form on Ni-affinity resin. Immunoblotting showed that the affinity-captured complex also contained the HA-tagged form (Kurkela et al., 2003). Liu *et al* also showed homodimerization of UGT1A9 and UGT1A1 by co-immunoprecipitation of tagged proteins (Liu et al., 2016). The interaction only occurred when proteins were co-expressed and not when the protein lysates were mixed, suggesting that UGTs may dimerize during co-translational insertion into the membrane. Co-expressed UGT2B7 proteins tagged with either a HA or myc-epitope co-immunoprecipitated as active homodimers that catalyzed 4-methylumbelliferone glucuronidation; again, when lysates containing each tagged form alone were mixed there was no association (Lewis, Mackenzie, & Miners, 2011).

Variant UGTs that result from genetic variation (including SNPs) or alterative splicing have also been tested for physical interactions using the methods described here; these are discussed in detail in Section 4.1.2.4.

4.1.2.2. FRET with tagged proteins. The use of fluorescent proteins as tags has allowed co-immunoprecipitation, co-localization, and fluorescence resonance energy transfer (FRET) studies to be performed using the same heterologous expression systems. Operana and Tukey (2007) were the first to use this approach with UGTs. They expressed UGT proteins that were fused to different fluorescent proteins (YFP or CFP) at their C-termini and used FRET to demonstrate that UGT1A proteins interact in live COS cells. Each of the full length UGT1A proteins (UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9, and UGT1A10) was able to interact with its own cognate isoform (homodimerize). Moreover, each UGT1A homodimer complex showed a broadly similar FRET efficiency suggesting that their abilities to dimerize do not vary greatly. UGT1A1 was able to dimerize with UGT1A3, UGT1A4, UGT1A6, UGT1A7, UGT1A8, UGT1A9, and UGT1A10. Interactions between other UGT1A family members were not tested. The FRET-derived data were complemented with co-immunoprecipitation analysis using HA-tagged and CFP-tagged UGT proteins. These data confirmed that each of the UGT1A proteins homodimerize and that UGT1A1 can heterodimerize with other UGT1A proteins.

Human UGT1A1 and UGT1A9 were shown to homodimerize in SF9 cells using FRET and this was confirmed using co-immunoprecipitation assays (Section 4.1.2.1) (Liu et al., 2016). This study also used semi-quantitative FRET to study homodimerization of multiple UGT1A9 SNP variants that have different activity profiles as discussed in Section 4.1.2.4.

4.1.2.3. Functional analysis: dominant-negative, rescue, and kinetic studies. We provided some of the earliest evidence for functional interaction of UGTs when we co-expressed two inactive mutant forms of the rat UGT2B1 proteins and restored some testosterone conjugating activity to the complex. The proteins were a truncated form lacking the transmembrane and cytosolic domains, and a form with a point mutation in the sugar binding domain (Meech & Mackenzie, 1997a). We also showed that co-expression of inactive chimeric UGT/EGT proteins with wild-type UGT2B1 protein led to a reduction in activity of the latter, indicating a dominant-negative effect. Subsequently, studies of guinea pig UGTs showed that co-expression of UGT2B22 with UGT2B21 in COS-7 cells increased UGT2B22-mediated morphine glucuronidation (Ishii et al., 2001; Ishii, Miyoshi, Maji, Yamada, & Oguri, 2004). UGT2B22 has no activity with morphine itself. These studies indicate that oligomerisation can provide functional complementation, cooperation, or inhibition depending on the nature of the proteins expressed.

Several studies have now shown that UGT co-expression can alter kinetic properties in an isoform- and substrate-dependent manner. Fujiwara *et al* studied the effects of co-expressing UGT1A1, UGT1A4 and UGT1A6 using probe substrates in HEK293 T cells (Fujiwara, Nakajima, Yamanaka, Katoh et al., 2007, 2007b). The effects of UGT1A4 or UGT1A6 on UGT1A1 activity was substrate dependent (increasing Vmax with bilirubin and decreasing Vmax with estradiol). Similarly, UGT1A1 decreased activity of UGT1A4 with imipramine but not trifluoperazine, while UGT1A6 affected UGT1A4-catalyzed trifluoperazine glucuronidation (increasing Km and Vmax) but not imipramine glucuronidation (Fujiwara, Nakajima, Yamanaka, Katoh et al., 2007, 2007b). Both UGT1A1 and UGT1A4 increased the Vmax of UGT1A6-mediated serotonin and diclofenac glucuronidation. These data were also supported by identification of heterodimers by native PAGE. Related work by the same group showed that UGT1A9 decreased estradiol glucuronidation by UGT1A1 (decreased Vmax), and altered the kinetics of imipramine glucuronidation by UGT1A4 (increased Km and Vmax) and serotonin glucuronidation by UGT1A6 (decreased Vmax). UGT1A1 affected propofol glucuronidation by UGT1A9 (increased Km and decreased Vmax), while UGT1A4 and UGT1A6 increased the Vmax of propofol glucuronidation (Fujiwara, Nakajima, Yamanaka, Katoh et al., 2007, 2007b). Subsequent studies by the same group found that UGT1A6 decreased (S)- and (R)-4'-HPPH glucuronide formation by UGT1A1 and UGT1A9; UGT1A4 increased (S)- and (R)-4'-HPPH O-glucuronide formation by UGT1A1 but decreased formation of the same conjugates by UGT1A9 (Nakajima, Yamanaka, Fujiwara, Katoh, & Yokoi, 2007). The broad conclusion from the aggregate of these studies

is that the effects of UGT1A co-expression (and presumed hetero-oligomerization) on enzyme activities depends on both the isoforms and substrates involved.

Functional interactions between UGT1A and UGT2B forms have also been shown. Kurkela *et. al.* demonstrated that co-expression of UGT1A4 affected glucuronidation by UGT1A9, UGT1A6 and UGT2B7, increasing activity of UGT1A6, but decreasing activity of UGT1A9 and UGT2B7 (Kurkela et al., 2007). Co-expression of UGT1As decreased Km and increased the Vmax of zidovudine glucuronidation by UGT2B7. The affinity of UGT1A1 for bilirubin and estradiol was increased when UGT1A1 was co-expressed with UGT1A4, UGT1A6, and UGT2B7. Co-expression of UGT2B7 also affected the kinetics of estradiol 3-O-glucuronidation by UGT1A1, imipramine N-glucuronidation by UGT1A4, serotonin O-glucuronidation by UGT1A6, and propofol O-glucuronidation by UGT1A9 (Fujiwara et al., 2010). Overall, these studies indicate that the effects of different UGT1A and UGT2B pairings can be cooperative or antagonistic and are substrate dependent.

An intriguing question is whether dimerization of UGTs can lead to new specificities and whether this may explain the glucuronidation of chemicals for which glucuronides have been measured *in vivo*, but for which the responsible UGT isoform remains undefined. Recent work studied the role of UGTs in resistance of acute myeloid leukemia (AML) cells to the chemotherapy drugs cytarabine (Ara-C) and ribavirin (Zahreddine & Borden, 2015). Drug resistant AML cells showed elevated levels of UGT1A but not UGT2 proteins, and mass spectrometry showed that the cells glucuronidated Ara-C and ribavirin. Liver microsomes expressing UGT1A1, UGT1A4, UGT1A6, and UGT1A9 produced the ribavirin glucuronide but not the Ara-C glucuronide, even though a small molecule inhibitor of UGT1A4 restored Ara-C sensitivity to the AML cells, supporting a role for UGT1A4 in Ara-C conjugation. Moreover, recent data in AML patients suggests that UGT1A1 polymorphisms alter Ara-C metabolism (Chen et al., 2018). Together these findings suggest that UGT1A1 and UGT1A4 are involved in Ara-C conjugation, however because microsomes containing these forms are insufficient to conjugate Ara-C, it is possible that the activity requires cooperation with other forms present in the AML cells.

4.1.2.4. Oligomerization can mediate the functional consequences of UGT genetic variation and alternative splicing. All *UGT* genes show genetic variation and their transcripts can be alternatively spliced. An important question is how oligomerization contributes to the *in vitro* and *in vivo* functions of such variant UGT forms.

Suzuki *et al* studied the dominant negative effects of a genetic variant of UGT1A1 found in a patient with dominantly inherited CN Type II (Koiwai et al., 1996) (see Section 4.1.1.3). Co-expression of the truncated inactive UGT1A1-p.Q331X protein with wild-type UGT1A1 in COS cells resulted in greatly reduced activity, which was consistent with the previous interpretation that the truncated form behaves dominantly *in vivo*. Co-affinity purification showed interaction of UGT1A1-p.Q331X with wild-type UGT1A1 in HEK293 T cells. This study also used Atomic Force Microscopy to probe the dimers and Fluorescence Recovery After Photobleaching to show that UGT1A1-p.Q331X was stable and retained in the ER (Suzuki et al., 2014).

Another CNII-associated mutation in the common (shared exon) region of the UGT1A gene produces multiple UGT1A proteins carrying a Y485D amino acid change (Kurkela et al., 2007). The scopoletin glucuronidation activity of this mutant form of UGT1A6 was dramatically increased by UGT1A4. Multiple wild-type UGTs (UGT1A1, 1A3, 1A4, 1A7, 1A8, 1A9, 1A10, 2B4, 2B7, 2B10, 2B15, 2B17) increased the serotonin glucuronidation activity of mutant UGT1A6, with UGT2B7 showing the most dramatic effects and UGT2B10 and UGT1A9 showing the weakest effects, suggesting that heterodimerization could mediate functional complementation. Interestingly, only UGT1A5 and UGT2B28, for which no high activity substrates have been identified, failed to increase mutant UGT1A6 activity. Moreover, co-expression of UGT1A4 lowered the Km of mutant UGT1A6 for UDP-glucuronic acid to that of wild-type UGT1A6, further supporting a rescue effect (Kurkela et al., 2007).

As mentioned above, the studies of Liu *et al* focused on the interactions of UGT1A1 and UGT1A9 SNP variants that display different kinetic properties; they concluded that while low activity variants typically reduced the activity of wild-type UGT1A9, the effects were also substrate dependent (Liu et al., 2016). Yuan et al. studied genetic variants of UGT2B7 that encode the amino acid substitutions A71S, H268Y and D398 N. Both H268Y and D398 N variant proteins impaired the activity of the wild-type and A71S forms with Zidovudine. Moreover, quantitative FRET studies suggested that distances between the donor and acceptor proteins varied between these different pairs, and that codon 71 and codon 268 were involved in the interaction (Yuan, Qian, Xiao, Sun, & Zeng, 2015). The same group later examined interactions of the UGT2B7 variants with UGT1A1 and with UGT1A9 allelic variants using FRET and co-immunoprecipitation (Yuan et al., 2016) and concluded that such variation can affect the strength of interaction.

In addition to genetic variants, alternatively spliced UGT forms can mediate dominant effects and several studies have tested both physical and functional interactions between wild-type and alternatively spliced UGTs as discussed below.

The 45 kDa truncated UGT1A_i2 proteins, described in Section 4.1.2.1, are catalytically inactive. Co-expression of UGT1A1_i2 with UGT1A1_i1 (full length) in HEK293 T cells showed that the i2 form dramatically reduced the activity of the full-length form (Levesque et al., 2007). Similarly, Bellemare et al. co-expressed UGT1A1_i2, UGT1A7_i2, and UGT1A8_i2 with their cognate UGT1A_i1 proteins using an inducible system that attempted to mimic the relative abundance of these isoforms in human tissues. Under these conditions, the truncated forms significantly inhibited the activity of UGT1A_i1 forms with different substrates, without affecting the Km for either the UDPGA or the substrate (Bellemare, Rouleau, Harvey, & Guillemette, 2010). In both studies dimerization was supported by co-immunoprecipitation assays. In another study, knockdown of transcripts encoding these UGT1A_i2 proteins using siRNA in HepG2 cells significantly increased the activity of UGT1A enzymes by 83 % as assayed using raloxifene, a common substrate for multiple UGT1A enzymes (Jones et al., 2012). The studies by Yuan et al. mentioned above tested the ability of UGT1A1_i2 (which they called UGT1A1*1b) to interact with multiple allelic variants of UGT1A9 and UGT2B7 using FRET and co-immunoprecipitation (Yuan et al., 2016). They found positive interactions with almost all combinations of proteins, but FRET efficiencies varied considerably suggesting varying strengths of interaction. Notably, interaction of UGT1A1_i2 with UGT2B7 D398 N (2B7*5 SNP variant) was not detected by FRET but was detected by co-immunoprecipitation, suggesting that FRET may be less sensitive.

Levesque et al. studied multiple truncated UGT2B4 forms generated by alternative splicing and found that all were inactive towards known UGT2B4 substrates such as eugenol and hyodeoxycholic acid (Levesque et al., 2010). However, they did not all have the same capacity to repress glucuronidation by wild-type UGT2B4 (2B4_i1). C-terminally truncated UGT2B4_i2, UGT2B4_i3, and UGT2B4_i5 were inhibitory (up to 40–60% reduction in activity); however, some other C-terminally truncated proteins showed little or no inhibition, which might relate to subtle differences in stability or localization. Physical interactions between these proteins were not tested. The alternatively spliced truncated inactive UGT2B7_i2 and UGT2B7_i4 variants reduced the activity of wild-type UGT2B7_i1 proteins with Zidovudine; this was also supported by physical interaction studies and evidence that wild-type and truncated forms of UGT2B7 colocalize in the ER (Menard et al., 2013).

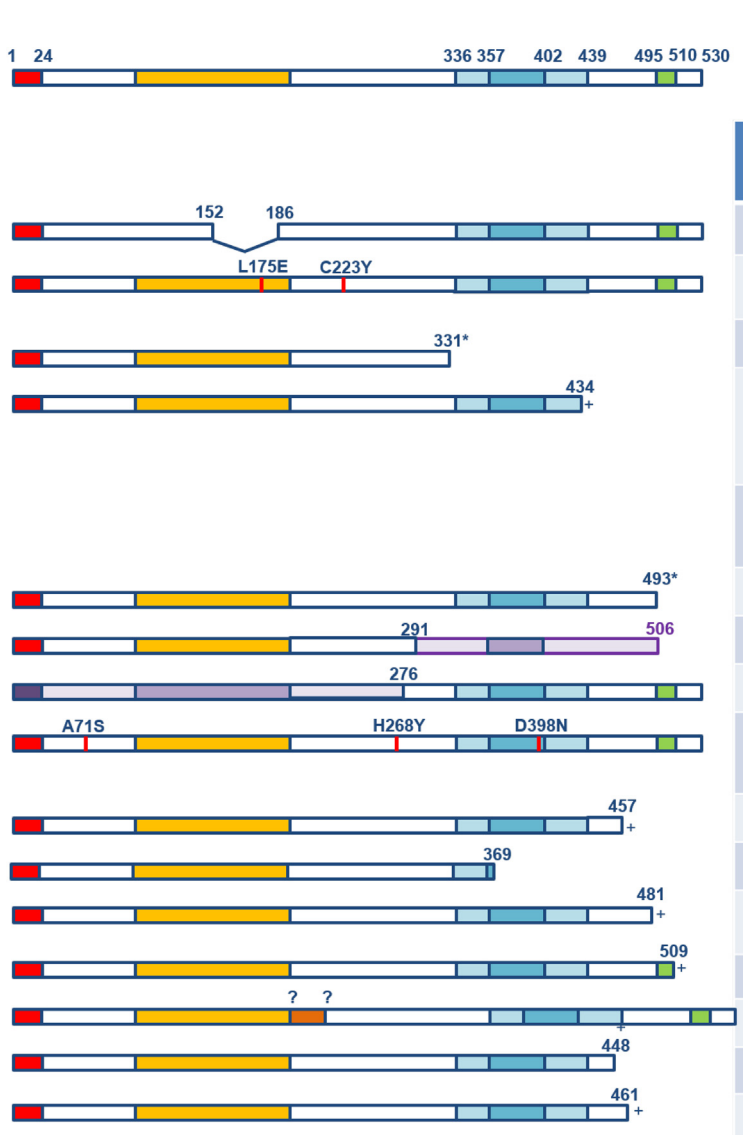
We recently identified chimeric UGT transcripts that produce inactive forms of UGT2B15 and UGT2B17 that are truncated and have novel C-terminal peptides (described in Section 2). These truncated forms inhibited activity of their cognate wild-type forms as well as that of UGT2B7. Interaction between truncated and full-length proteins was shown by co-immunoprecipitation assays (Hu et al., 2018). Other

recent work studied alternatively spliced forms UGT2B10_i4 and UGT2B10_i5 (see Section 2) with novel C-terminal peptides. Both inactive forms co-localized with full length UGT2B10 in the ER and inhibited its activity when co-expressed in HEK293 T cells (Labriet et al., 2018). Interaction between truncated and full-length forms was demonstrated using co-immunoprecipitation. Of interest, this study showed that expression of the inactive UGT2B10_i4 form in HepG2 cells enhanced endogenous UGT2B10 activity. Mechanistically this is not understood but may involve modulation of interactions already occurring between different UGT isoforms in HepG2 cells. To date there is only one clear example of an alternative spliced UGT variant showing a cooperative effect with other UGTs. The UGT2B7_i8 protein contains a 32-aa in-frame insertion between the substrate-domain and co-substrate- binding domain and is inactive (see Section 2). UGT2B7_i8 can interact with UGT2B7_i1 and co-expression of these forms in HEK293 T cells significantly enhances the activity of the latter (Tourancheau et al., 2016). Intriguingly, expression of UGT2B7_i8 also reduced proliferation and enhanced adhesion of HEK293 T cells. The mechanism of this effect is unknown although it was presumed to relate to the metabolic alterations observed in these cells including accumulation of multiple amino acids and reduction in purines and pyridines.

Overall, these data suggest the potential for alterative splicing to control UGT activity in a tissue specific, or signal/stimulus mediated manner, mainly through the production of inactive truncated proteins that can oligomerize with their cognate full-length forms, or other UGT isoforms.

4.1.3. Identifying domains and residues involved in UGT interactions

As summarized in Fig. 6, much of what we know about the interaction domains of UGTs has been derived from analysis of natural protein variants (e.g. splice variants) as well as a small number of synthetic mutations, deletions and chimeras. Because most natural variants are C-terminally truncated, the data is redundant; in particular, loss of the transmembrane and cytoplasmic domains has been studied numerous times and physical interactions, as evidenced by methods such as co-immunoprecipitation, are always retained. The shortest UGT form yet tested for interaction is 331aa and lacks the entire C-terminal co-



Form 1 (shown)	Form 2	Interact	Assay	Ref
1A1Δ 152-186	1A1wt	no	2-hybrid	Ghosh 2001
1A1 L175E C223Y	1A1wt	no	2-hybrid	Ghosh 2001
1A1 331*	1A1wt	yes	coIP	Suzuki 2014
1A1_i2 1A7_i2 1A8_i2	1A1_1(wt) 1A7_1(wt) 1A8_1(wt)	yes	coIP	Levesque 2007; Bellemare 2010; Rouleau 2016
1A1_i2 (1A1*1b)	1A9*1, *2, *3, *5 2B7*1, *2, *5, A71S	yes	FRET and coIP	Yuan 2016
2B1 494*	2B1 494*	yes	SDS-PAGE	Meech 1997
2B1/EGT	2B1/EGT	yes	SDS-PAGE	Meech 1997
EGT/2B1	EGT/2B1	no	SDS-PAGE	Meech 1997
2B7 A71S, H268Y D398N	2B7wt	yes	FRET	Yuan 2015
2B7_i2	2B7_i1 (wt)	yes	coIP	Menard 2013
2B7_i4	2B7_i1 (wt)	Yes		Menard 2013
2B10_i4	2B10_i1 (wt)	yes	coIP	Labriet 2018
2B10_i5	2B10_i1 (wt)	yes	coIP	Labriet 2018
2B7_i8	2B7_i1 (wt)	Yes	coIP	Rouleau 2016
2B15chim	2B15wt 2B7wt	Yes	coIP	Hu 2018
2B17chim	2B17wt 2B7wt	Yes	coIP	Hu 2018

Fig. 6. Summary of experimentally truncated and mutated forms of UGT proteins that have been tested for physical interactions in heterologous expression systems. Studies that have only shown the ability of truncated or mutated UGT forms to alter the activities/kinetics of full length UGTs are not included. Left, schematics of the modified proteins tested for interaction; right, table providing details of the physical interaction assays and outcomes. Note that the point mutations (red bars) shown in UGT1A1 and in UGT2B7 were tested individually and not in the same protein. EGT protein domains are shown in purple.

substrate binding domain; however, it has only been tested for self-dimerization.

While most data suggest that the N-terminal half of the protein is minimally required for UGT-UGT interactions, it is important to note that most physical interaction assays have been qualitative. Yuan et al. used FRET to quantify interactions and their finding that truncated UGT1A1_i2 protein bound less well to UGT1A9 and UGT2B7 than wild type UGT1A1 suggested that the C-terminal region also contributes to dimerization efficacy (Yuan et al.). Based on UGT1A homology models, they predicted that the absence of two C-terminal envelope α-helixes (env1 and env2) and the transmembrane helix in UGT1A1_i2 reduced the ability of the truncated protein to interact.

The two-hybrid studies of Ghosh et al. offer some molecular insights into the putative N-terminal dimerization region in UGT1A1 (Ghosh et al., 2001). Substitution of residues L175E or C233Y significantly reduced the ability to form UGT1A1 homodimers. Further, dimerization was almost abolished when residues 152–180 (spanning a predicted membrane-embedded helical region) were deleted. However, the observation that the two-hybrid assay failed to show association between UGT1A1 and UGT1A6, which was demonstrated by several others via co-immunoprecipitation and FRET, raises concern about the overall validity of the two-hybrid method for faithfully reporting on UGT interactions. Hence caveats will remain until these modifications are tested using more direct methods such as co-immunoprecipitation.

In addition to the interaction domain(s) being poorly delineated, structural insights into UGT-UGT binding domain(s) and its role in small molecule (e.g. drug substrate) recognition are currently lacking. One of the main issues impeding the rational prediction of dimerization/oligomerization domain(s) is the absence of full-length three-dimensional (3D) structure(s) of human UGTs (Nair, Meech, Mackenzie, McKinnon, & Miners, 2015). The N-terminal half is the region of greatest sequence divergence between UGT isoforms (e.g. the N-terminal domains of the UGT1 family are only ~38 % similar); there are currently no experimental structures available for this region. It is uncertain if this variable region with anticipated structural differences may offer binding interactions that are unique with regard to each UGT isoform, allowing preferential recognition of different dimer partners. Moreover, the altered kinetics of substrates with co-expressed UGTs (Section 4.1.2.3) is indicative that dimerization domains may preferentially regulate the effective binding of one substrate over others (e.g. control ligand accessibility and the specificity of active sites).

In the absence of a defined structure, one useful approach is the application of homology models to the investigation of UGT-UGT structure–function; however, this remains at an early stage. Lewis et al. (2011) developed a homology model of UGT2B7 from plant and human template structures to elucidate a putative UGT2B7 dimerization domain. In this model, the proposed dimerization domain was comprised of eighteen amino acids with a putative dimerization signature motif (FPPSYVPVVMS), which is conserved in all UGT2B enzymes. The composition of the dimerization site was mainly hydrophobic and showed the presence of 'proline brackets'. The dimer interface was predicted to form stable salt bridges, aromatic π-π stacking interactions, and S-aromatic (face) interactions between the two UGT2B7 monomers. This proposed dimerization domain was located between residues 183 and 200, consistent with the reported involvement of the N-terminal domain. However, the requirement for this region for dimerization has not been validated by experimental deletion, nor has the role of individual residues within the domain been tested by site-directed mutagenesis. Notably, these UGT2B7 residues lie in a different location to that proposed to mediate dimerization of UGT1A1 and exhibit a different secondary structure at the respective sites: a loop (Lewis et al., 2011) versus a helix (Ghosh et al., 2001). These studies raise the possibility that UGTs from different families have distinct dimerization sites or may interact via multiple binding domains.

Whether disulfide bonding is involved in UGT interactions remains poorly defined. Ghosh et al. found that UGT1A1 dimers were visible on SDS-PAGE only in non-reducing conditions; however because most of the protein was monomeric with or without reducing agent, they concluded that disulfide bonding did not play an important role in dimerization (Ghosh et al., 2001). However, recently Kim et al. (2018) demonstrated that over-expressed UGT1A10 can form covalently crosslinked higher-order oligomers via intermolecular disulfide bonds. The oligomerization of UGT1A10 protein was only observed in HEK293 cells and not in CHO cells. The authors proposed the involvement of three cysteine residues (C72, C183, and C277) that lie on the protein surface of the modelled UGT1A10 structure, where C183 of one monomer might form a disulfide bond with C72 or C277 of another monomer; however, this was not tested experimentally. Moreover, given that these cysteine residues are not conserved, it also remains possible that disulfide bonding is involved in the interactions of only a subset of UGTs.

Overall, there is a paucity of experimental data to define dimerization domains or critical residues, and the current limited attempts at modelling provide very little insight that can be generalized across different known UGT-UGT interactions. Moreover, some studies report contradictory dimerization requirements. A parsimonious model may be proposed in which multiple domains of the UGT protein are involved in interactions, however only those mediated by the N-terminal domain are essential (with C-terminal interactions perhaps stabilizing the complex). The involvement of the N-terminal domain may allow different UGT partners to be bound with different affinities and regulate the substrate specificity of the oligomeric complex.

4.2. Interactions of UGTs with non-UGT proteins

The cytochrome P450 (CYP) enzymes constitute a major family of Phase I drug metabolising enzymes, which are most abundantly expressed in the liver. Like UGTs, eukaryotic CYPs are primarily anchored within the ER membrane. A handful of studies have investigated protein-protein interactions between UGT and CYP proteins, which would be indicative of facilitation of a series of multi-step drug metabolic conversions.

Using affinity chromatography with a CYP1A1-conjugated Sepharose 4B column, Taura and colleagues observed the co-elution of multiple UGT isoforms (although the identity of these isoforms was unknown), from rat liver microsomes (Taura et al., 2000). Similarly, co-immunoprecipitation studies with rat liver microsomes detected UGT proteins following immunoprecipitation with specific antibodies towards CYP3A2, CYP2B2, CYP2C11/13 and CYP1A2 (Ishii et al., 2007). Importantly, all co-precipitated rat UGT proteins were active exhibiting activity toward 4-MU. The affinity of each CYP isoform for UGTs varied, suggesting isoform specificity in interactions between rat UGTs and CYPs.

Interaction between UGT2B7 and CYP3A4 has also been observed by co-immunoprecipitation of human liver microsomes (Fremont et al., 2005; Takeda, Ishii, Iwanaga et al., 2005, 2005b; Takeda et al., 2009), with confirmation provided by overlay assays using glutathione S-transferase-CYP3A4 fusion protein (Takeda, Ishii, Iwanaga et al., 2005, 2005b, Takeda et al., 2009). Freemont et al. found that CYP3A4 also co-immunoprecipitated UGT1A1 and UGT1A6 from human liver microsomes (Fremont et al., 2005). To date, there are no reports on the identification of UGT protein domains required for interaction with CYPs. Furthermore, whether these reported UGT-CYP interactions occur *in vivo* remains to be investigated.

Only three studies have assessed the impact of specific UGT-CYP interactions on the corresponding UGT activities. Whilst co-expression of CYP1A2 or CYP2C9 with UGT2B7 in mammalian COS1 cells had no effect on the kinetic parameters of UGT2B7-catalysed metabolism of morphine to morphine-3-glucuronide, co-expression of CYP3A4 with UGT2B7 increased the Km of morphine-3-glucuronide formation around 10-fold without affecting Vmax (Takeda, Ishii, Iwanaga et al., 2005, 2005b). Assessment of glucuronidation in a Sf-9 insect cell

expression system demonstrated increased Km and Vmax of UGT1A6-mediated serotonin glucuronidation when co-expressed with CYP3A4 (Ishii et al., 2014). Simultaneous expression of CYP3A4 with UGT1A1 or UGT1A7 resulted in increased Vmax values for UGT1A1-mediated and UGT1A7-mediated formation of 4-MU and SN-38 glucuronides, UGT1A1-mediated estradiol 3-O-glucuronidation and UGT1A7-mediated 4-hydroxybiphenol-glucuronidation (Ishii et al., 2014). However, co-expression of CYP3A4 had negligible impact on Km values for all UGT1A1- and UGT1A7-catalysed reactions tested.

Glucuronidation of 4-MU has also been assessed following co-expression of rat UGT2B3 and CYP3A1. Whilst expression of CYP3A1 lowered the Vmax of UGT2B3-mediated conjugation, expression of CYP3A1 increased the Vmax of 4-MU glucuronidation catalysed by a variant of UGT2B3 to which a glycosylation site had been engineered (Nakamura et al., 2016). The fact that N-glycosylation of UGT2B3 altered the sensitivity of UGT2B3 to CYP3A1-dependent modulation suggests a crosstalk between post-translation modification and UGT interactions that remains to be further investigated.

Overall, this small body of literature suggests that the ability of CYP proteins to modulate UGT glucuronidation activity is dependent not only upon the CYP isoform but also on the specific UGT substrate and post-translational modifications of the UGT enzyme.

It is plausible that CYP enzymes can act as substrate suppliers for UGTs in addition to functionally modulating UGT activity. This idea is supported by a study that examined glucuronidation of 3-hydroxybenzo(*a*)pyrene (3-OH-B(*a*)P), a major metabolite of B(a)P produced by CYPs, in rat liver microsomes (Taura et al., 2004). In this study, UGT activity in non-permeabilised microsomes was significantly reduced by the classical CYP inhibitor alpha-naphthoflavone (α-NF), but was not reduced by α-NF under conditions of increased membrane permeability (Taura et al., 2004). This observation suggests that α-NF does not act directly on UGTs, but functions via CYP inhibition within a UGT-CYP complex.

In addition to UGT-CYP protein interactions, interactions between UGTs and other microsomal proteins has also been investigated. Large-scale shotgun analysis using LC–MS/MS of human liver microsomes immunoprecipitated with anti-UGT2B7 antibody identified, in addition to CYP3A4, CYP1A2 and UGT1A, an extensive array of other microsomal/ER proteins. These included the enzymes mannosyl-oligosaccharide glucosidase (MOGS), bile acyl-CoA synthetase (SLC27A5), microsomal triglyceride transfer protein (MTTP), ATP synthase (ATP5MG), NADH dehydrogenase (NDUV1), monoamine oxidase (MAO) and Peroxiredoxin-4 (PRDX4) (Fujiwara & Itoh, 2014). Notably, many interacting proteins are located predominantly in non-ER locations including peroxisomes: e.g. 3-ketoacyl-CoA thiolase (ACAA1) and peroxisomal membrane protein 2 (PXMP2), and in mitochondria: e.g. voltage-dependent anion-selective channel protein 1 (VDAC1), 2-oxoglutarate/malate carrier protein and pyruvate carrier 2 (SLC25A11) and mitochondrial 3-ketoacyl-CoA thiolase (ACAA2) (Fujiwara & Itoh, 2014). It should be noted however that peptide counts for most of these putative partners were very low and they were not experimentally validated.

More recently, an unbiased proteomic investigation of the endogenous protein interactome of human UGT1A enzymes in liver, kidney and intestinal tissues revealed several interacting proteins of pharmacological and metabolic importance (Rouleau et al., 2017). These investigations were performed using affinity purification of UGT1A enzymes with an antibody that recognises all UGT1A isoforms (Bellemare et al., 2011). Observed interacting proteins included the UGT2B enzymes UGT2B4, UGT2B7 and UGT2B17, and metabolic proteins that function in detoxification and bioenergetic pathways such as glutathione S-transferase (GSTA1), alcohol dehydrogenase (ALDH2) and antioxidant enzymes Peroxiredoxin-1 and -2 (PRDX1 and PRDX2). In addition, several proteins involved in fatty acid degradation (ACOT8, ECH1, CPT1A, ACAA2), the glucagon signalling pathway (PHKG2, PHKA2, ITPR2, CALM1) glycolysis/gluconeogenesis (IDH2, GAPDH, GFPT1) and vesicular trafficking (RALGAPA1, GBF1, SH3KBP1) were significantly represented (Rouleau et al., 2017). Of note, ACAA2 was also previously reported as a potential interaction partner of UGT2B7 (Fujiwara & Itoh, 2014). The interaction of UGT1A9 with ACOT8, PHKA2 and SH3KBP1 were experimentally validated by co-immunoprecipitation of transiently expressed proteins. Consistent with the notion that UGTs interact functionally with lipid related proteins, UGT1A9 overexpression increased lipid droplet formation in HEK293 T cells. Collectively, the interaction networks identified in these studies are in line with a model which favours UGT enzymes as part of a "metabolosome" complex with other drug-metabolising enzymes and associated transport proteins to coordinate detoxification and elimination. It also suggests significant coordination between UGTs and other components of endogenous cellular metabolism. Again, many of these interacting proteins are not predominantly ER-located but can be found in mitochondria, peroxisomes, or cytoplasm.

Rouleau et al. studied the interactome of C-terminally truncated UGT1A_i2 isoforms (Rouleau, Roberge, Bellemare, & Guillemette, 2014). Although largely localized to the ER, some expression of UGT1A_i2 isoforms has been observed in cytoplasm (Audet-Delage et al., 2017; Levesque et al., 2007). Consistent with this, global peptide analysis of human kidney and intestine homogenates with an antibody specific for UGT1A_i2 variants found interactions between UGT1A_i2 proteins and cytoplasmic peroxiredoxin 1 (PRDX1) and catalase (CAT) (Rouleau et al., 2014). They reported that these interactions were not observed when a UGT1A_i1 antibody was used, although a later study precipitated PRDX1 with an antibody that recognizes full length UGT1A (Rouleau et al., 2017). A more recent proteomic analysis of UGT1A_i2 interactions identified several high-confidence candidate endogenous interactors in human liver and kidney tissues (Audet-Delage et al., 2017). These included various proteins involved in the citrate cycle, glucagon signalling, glycolysis pathway, tryptophan metabolism and pyruvate metabolism. Co-immunoprecipitation and immunofluorescence studies validated an interaction with pyruvate kinase M2 (PKM), a key enzyme of the glycolysis pathway primarily localised in the cytoplasm, in HT115 colon cancer cells (Audet-Delage et al., 2017).

Relative to the UGT1A family, UGT2B family members other than UGT2B7 have been less studied in terms of interactions with non-UGT proteins. However, UGT2B17 was reported to interact with kinase c-Src in a prostate cancer cell line after prolonged androgen deprivation; moreover, this interaction correlated with activation of c-Src. c-Src is located in the cytoplasm and nucleus where it phosphorylates and activates multiple downstream effectors including androgen receptor (AR) and is associated with androgen-independent AR function. This suggests a novel mechanistic aspect to the role of this steroid-conjugating UGT in steroid signalling in cancer (Li et al., 2016).

Two other interesting interactions that are somewhat buried in the literature of UGT biology relate to the UGT8 protein. The first is the interaction of UGT8 with SLC35A2 (Sprong et al., 2003). UGT8 is the ceramide galactosyl transferase and SLC35A2 is a UDP-galactose transporter that facilitates galactose transport into both the ER and Golgi complex for galactosidation of lipids and proteins. The interaction of UGT8 and SLC35A3 (confusingly referred to in the study as CerGalT and UDP-Galactose Transporter 1 [UGT1] respectively) was shown by coimmunoprecipitation in a heterologous expression system. When the proteins were co-expressed, galactosylceramide production was increased indicating that the transporter increases supply of galactose to UGT8. Moreover, SLC35A2 was localized in the ER when co-expressed with UGT8; whereas it localized exclusively to the Golgi in cells lacking UGT8. Thus, it appears that UGT8 retains the transporter in the ER and could hence influence the balance of ER-localised lipid galactosidation and Golgi-localised protein galactosidation reactions. The second known interaction of UGT8 is with the sigma-1 receptor (Sig-1R) chaperone (Hayashi, Fujimoto, Sprong, & Su, 2012). Sig-1R was shown to bind to UGT8 using immunoprecipitation, and Sig-1R knockdown prolonged the lifetime of UGT8 in the ER and increased its activity.

Because Sig-1R associates with insulin induced gene (Insig) in a partly sterol-dependent manner, it was suggested that interaction of UGT8 with this complex allows its level/activity to be controlled by sterol-sensitive ER-associated degradation (ERAD), possibly allowing crosstalk between sterol and sphingolipid metabolism.

As previously discussed, UGTs are typically resident within the ER; hence it may appear difficult to reconcile the numerous reports of their interactions with proteins that are canonically located in other cellular compartments. However, the ER is known to extend from the nuclear envelope to the cell periphery and to make direct contacts with structures such as the nucleus (via the outer nuclear envelope) and the mitochondria (via mitochondrial-associated membranes) (Nixon-Abell et al., 2016; Vance, 1990). It is also the source of secretory vesicles that deliver proteins to the plasma membrane and potentially other cellular locations. Moreover, the ER plays a central role in peroxisome biogenesis and ER-peroxisome tethers have recently been identified (Hua et al., 2017). Thus, as is discussed in Section 5, it is possible that at least a subpopulation of UGTs and partner proteins could co-locate in almost any region of the cell where they may engage in either canonical activities or atypical 'moonlighting' functions (see Fig. 7)

5. Summary: functional impact of expanding UGT diversity

The 22 UGT enzymes that have been identified in the human genome and functionally characterized in single enzyme expression systems would appear to represent only the tip of the iceberg in terms of UGT functional diversity. Given that a major role of UGTs is to provide a protective interface between the organism and its chemical-rich environment, it is likely that selective pressure has driven the evolution of multiple mechanisms that increase metabolic capacity. In particular, the extensive evidence that UGTs hetero-oligomerize provides vast opportunity for UGTs to modulate one another's functions. Such interactions may alter kinetic properties of UGTs, inhibit their function, or even provide new substrate specificities. Ample studies have shown kinetic effects of UGT co-expression; however, we have little idea of how such changes occur as we lack models for the interacting N-terminal regions of the UGTs. In the absence of such models, we can only speculate on the nature of the interaction interfaces, and hence the mechanism (s) by which interaction alters substrate recognition and catalysis.

A critical question is whether UGT hetero-oligomerization occurs to a sufficient degree as to produce significant deviation of *in vivo* pharmacokinetic values from *in vitro* data generated using heterologous single enzyme systems. To date this question is not readily answered as most studies of UGT interactions are also performed in heterologous over-expression systems, which may artificially enhance interactions. Future work to ablate individual UGT isoforms in cells or in animals using siRNA or CRISPR, as well as assays with probe substrates and isoform specific inhibitors, may help in this regard. It is plausible however that failure to consider UGT interactions may affect our ability to accurately predict drug metabolism and clearance *in vivo*. It is also possible that hetero-oligomers explain the production of some glucuronides for which a single isoform has not been identified as responsible; again, purpose-designed studies will be required to test this hypothesis.

UGT hetero-oligomerization may also provide an alternative function for poorly active (orphan) UGTs. A parallel may be drawn with the orphan SULT4 protein, which is inactive and may dimerize with and inhibit other SULTs (Minchin, Lewis, Mitchell, Kadlubar, & McManus, 2008; Petrotchenko, Pedersen, Borchers, Tomer, & Negishi, 2001; Sidharthan, Butcher, Mitchell, & Minchin, 2014). Unlike SULT4, no UGT to date is considered entirely catalytically incompetent; however, UGT1A5 and UGT2B28 have no significant activity with any known substrate and previous work showed that these were the only wildtype UGTs that could not functionally complement mutant UGT1A6 (Kurkela et al., 2007). In our studies, we have found that UGT2B28 can inhibit the activity of other wildtype UGT enzymes (not shown). Thus, it remains formally possible that some UGTs function less through their own catalytic activity, and more through inhibition/modulation of others.

In contrast to full length UGTs, multiple truncated UGTs produced by alternative splicing are inactive and inhibit the functions of their

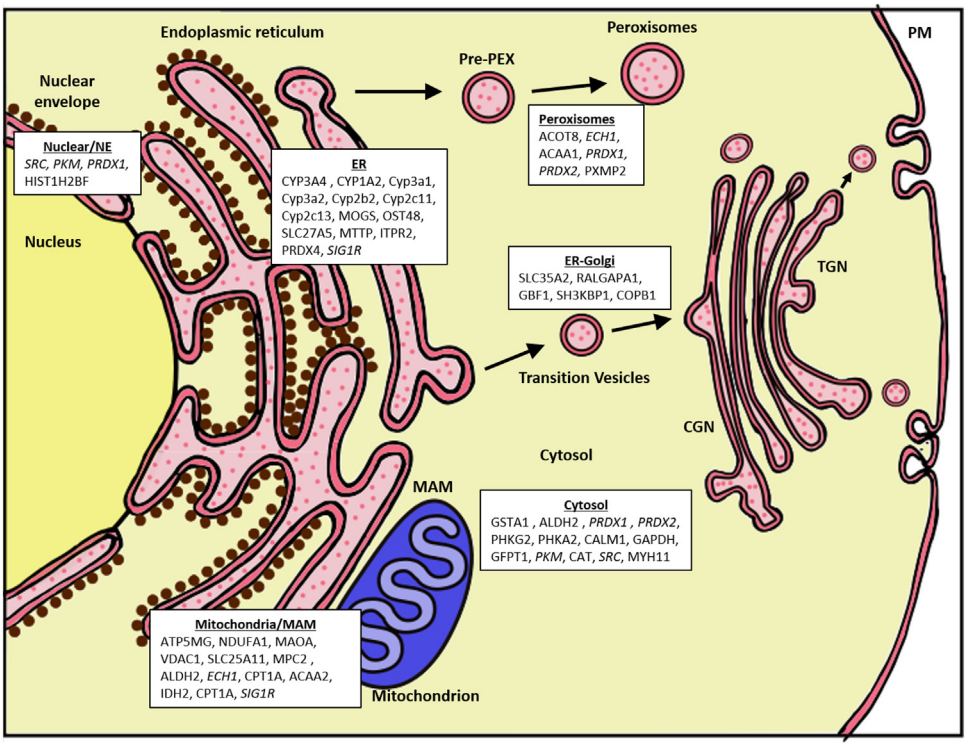


Fig. 7. A subcellular location-oriented summary of reported interactions between UGTs and non-UGT proteins. Protein-partners are listed in their most commonly reported locations; proteins highlighted in italics are commonly found in more than one location. Lower case gene names indicate non-human proteins. ER, endoplasmic reticulum; MAM, mitochondrial associated membranes; NE, nuclear envelope; CGN and TGC, cis and trans Golgi network; PEX, peroxisomes.

cognate full-length forms as well as other UGTs. Alternative splicing provides exceptional capacity to diversify gene function. The fact that *UGT* genes generate such a diverse array of protein variants by alternative splicing suggests that this a fundamental aspect of UGT biology. Like gene regulation, splicing is often tissue specific and can be dynamically regulated so that it is rapidly responsive to cellular demands. However, it may more dramatically change the cellular output by not just altering levels of the target protein, but fundamentally altering its function, and the functions of its multiple interaction partners through dominant effects.

In most normal tissues, alternatively spliced UGT transcripts represent a small fraction of the gene output. However, dysregulation of alternative splicing is a hallmark of cancer and many examples of variant transcripts/proteins with pathogenic effect in cancer are known. It is likely that dysregulated UGT splicing can modulate the regulation of small signalling molecules that promote cancer risk or progression. As an example, the alternatively spliced forms of the androgen conjugating UGT2B15 and UGT2B17 genes that we identified in normal and cancer tissues (Hu et al., 2018) displayed widely varying relative abundance in breast cancers. In addition, the widespread use of alternative 3′ exons encoding different 3′UTR regions can modulate post-transcriptional regulation by miRNAs and hence transcript and protein abundance. Unlike the canonical transcripts, the alternative UGT2B15 and UGT2B17 transcripts mentioned above cannot be regulated by miR-376c and hence may accumulate to higher levels in tissues where this miRNA is abundant. We have also found other variant UGTs with alternate 3′ exons that are overexpressed in different cancer models (unpublished data), which may involve differential miRNA recognition, suggesting that this could be an important aspect of UGT regulation in cancer. An important future direction is to determine the potential pathogenicity of such variants, and whether they act as biomarkers of cancer progression.

Increasing evidence that UGTs interact with multiple non-UGT proteins is greatly expanding our understanding of UGT function. Proteins shown to interact with UGTs are found in a wide variety of subcellular locations and are involved in a wide variety of cellular functions and depicted in Figs. 7 and 8. Interactions with microsomal CYPs and other enzymes such as GSTA and ALDH is suggestive of a 'drug/xenobiotic metabolosome' (Fig. 8) in which various steps in functionalization and conjugation are coordinated to more effectively detoxify xenobiotics. Other interactions may relate to the function of the UGT directly, for example, proteins implicated in vesicular trafficking may be involved in retention of UGTs in the ER by retrograde transport. In support of this idea, we previously found that UGT2B1 immunoprecipitates with the β-COP subunit of the COPI complex which binds to the dilysine ER retention motif (unpublished data). The interaction of UGT8 with SLC35A3 may be indicative of widespread interaction of UGTs with the sugar transporters responsible for supplying their co-substrates, a notion that could be readily tested given that ER-resident transporters for most UDP-sugars are well known.

Several UGT-interacting proteins identified via proteomics are involved in fundamental aspects of cellular metabolism such as fatty acid degradation and glycolysis/gluconeogenesis and are located in cytoplasm, mitochondria or peroxisomes. Some of these proteins were found as part of the UGT1_i2-specific interactome (Audet-Delage et al., 2017; Rouleau et al., 2017). It is possible that these represent novel interactions mediated by the unique C-terminal peptides of the i2 proteins or conformational changes that create new interaction surfaces. However, given that a subpopulation of full length UGT proteins can also be located outside of the ER, it is possible that both full length (i1) and i2 forms interact with these proteins, but the increased localization of i2 protein outside of the ER increases the sensitivity for detecting the interactions. Structures are that make direct contact with the ER such as the nuclear envelope, the mitochondrial-associated membranes (MAM) (Vance, 1990) and peroxisomes (Costello et al., 2017; Hua et al.,

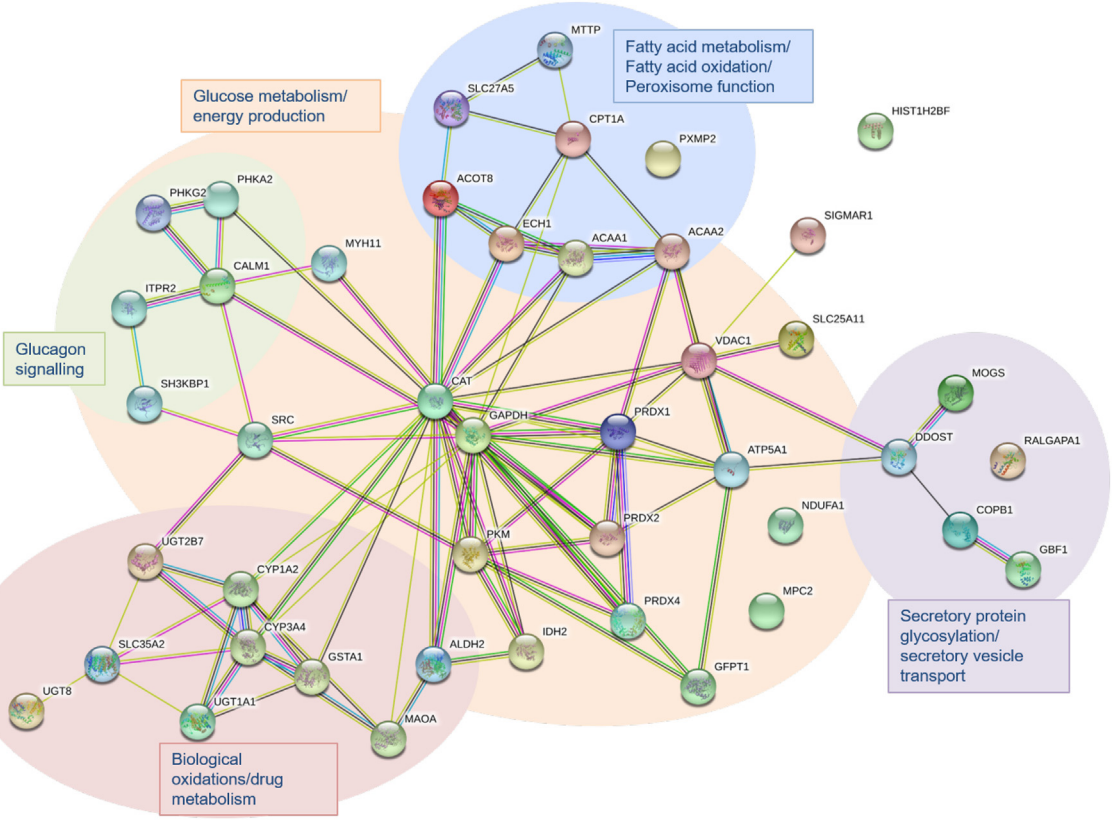


Fig. 8. Network analysis of proteins reported to physically interact with UGTs (UGT2B7, various UGT1A isoforms, or UGT8). Functional interactions between proteins were predicted using STRING and pathway analysis was performed using Toppcluster and Network Analyst with KEGG pathway annotations. Only human proteins are shown.

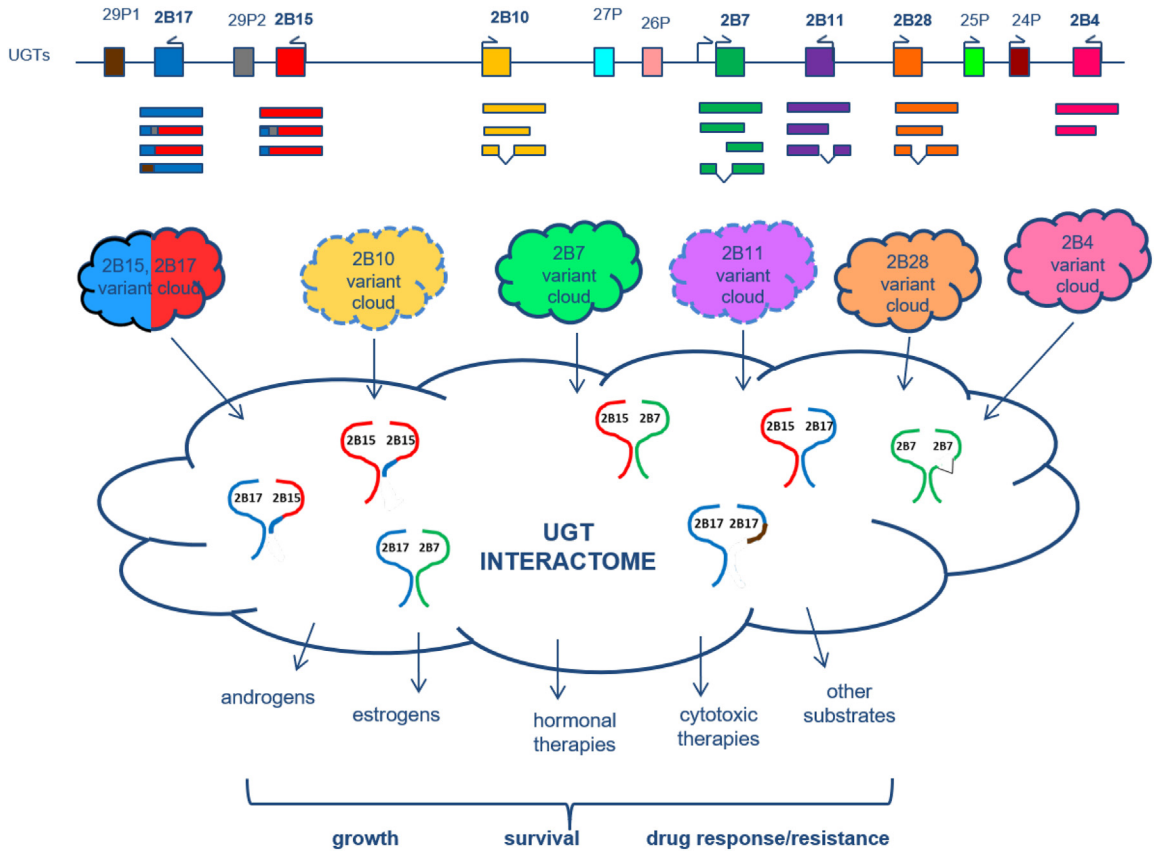


Fig. 9. Model of the functional UGTome generated by a combination of alternative splicing and oligomerization events. As depicted for the UGT2B locus, a variant transcript 'cloud' encodes canonical and variant UGTs that can interact in heterotypic complexes forming the UGT interactome (note that for simplicity, only a subset of known variants and interactions are shown, and non-UGT proteins are omitted). The activities of these complexes with endobiotic and drug substrates affects cell growth and survival (e.g. via altering nuclear receptor signalling), as well as drug response. Not shown: UGT-mediated control of steroid signalling could feed back to alter UGT expression and potentially alternative splicing.

2017) might be expected to contain UGTs, although likely at lower abundance than the ER.

The reports that UGTs interact with lipid metabolic proteins suggests that they may play broad roles in fundamental lipid metabolism; this would be consistent with observations that several UGTs collectively conjugate diverse lipids such as fatty acids, eicosanoids, prostaglandins, farnesol, ceramide, diacylgycerol, and cholesterol (Wade, 1970; Hara & Taketomi, 1982; van der Bijl, Strous, Lopes-Cardozo, Thomas-Oates, & van Meer, 1996; Jude et al., 2001; Little, Williams, Xu, & Radominska-Pandya, 2002; Turgeon et al., 2003; Staines, Sindelar, Coughtrie, & Burchell, 2004; Allain et al., 2019). UGTs might also modulate other pathways involved in the production and use of UDP-sugars through protein-protein interactions, as part of a homeostatic mechanism. The ability of proteins, often enzymes, to carry out non-canonical functions has been termed 'moonlighting' (Jeffery, 2019). Moonlighting is often facilitated by alterative splicing, which provides novel capacity for interactions and/or altered subcellular localization. Despite extensive interactions of UGTs with non-UGT proteins, there is currently no study describing any direct 'moonlighting' (i.e. non-conjugation) functions that UGTs perform in cooperation with these proteins; the possible exceptions being the ability of UGT2B17 to promote c-Src phosphorylation, and the ability of UGT1A9 to modulate lipid droplet formation; however the mechanisms involved remain unknown (Li et al., 2016; Rouleau et al., 2017).

Recent work presented a model for the CYP-ome as a cloud of protein variants produced by alterative splicing that may interact to further diversify functions (Reed & Backes, 2012; Annalora, Marcus, & Iversen, 2017); we suggest that this is also a useful model to apply to the UGTome. In the model of the UGT2B-ome shown in Fig. 9, alternative splicing produces a variant transcript 'cloud' encoding proteins with altered functions. Canonical and variant UGT proteins can interact in heterotypic complexes, and also interact with non-UGT proteins, forming the UGT interactome. The diverse effects of these complexes on conjugation activities with endobiotic and drug substrates, as well as other possible 'moonlighting' functions (Jeffery, 2019), could coordinately affect cell growth and survival, as well as drug responses. Moreover, given that known UGT2B substrates, such as steroids, can modulate both UGT transcription and alternative splicing (Auboeuf et al., 2004), there is the capacity to generate a feedback loop that maintains the UGTome homeostatically, and provides finely tuned responses to dynamic metabolic demands.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

Acknowledgements

We acknowledge funding support from sponsor-id="http://dx.doi.org/10.13039/501100000925">National Health and Medical Research Council (NHMRC) of Australia [Grant IDs GNT1143175 and GNT1085410 (R.M, R.A.M, P.I.M, D.G.H)]. The project was also supported by funding from the sponsor-id="http://dx.doi.org/10.13039/100008017">Flinders Medical Centre Foundation. R.A.M. is a Cancer Council/SA Health Beat Cancer Professorial Chair.

References

Allain, E. P., Rouleau, M., Le, T., Vanura, K., Villeneuve, L., Caron, P., Turcotte, V., Lévesque, E., & Guillemette, C. (2019). Inactivation of prostaglandin E2 as a mechanism for

UGT2B17-Mediated adverse effects in chronic lymphocytic leukemia. *Frontiers in Oncology 9*(606).
Annalora, A. J., Marcus, C. B., & Iversen, P. L. (2017). Alternative splicing in the cytochrome P450 superfamily expands protein diversity to augment gene function and redirect human drug metabolism. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 45*(4), 375–389.
Auboeuf, D., Dowhan, D. H., Li, X., Larkin, K., Ko, L., Berget, S. M., et al. (2004). CoAA, a nuclear receptor coactivator protein at the interface of transcriptional coactivation and RNA splicing. *Molecular and Cellular Biology 24*(1), 442–453.
Audet-Delage, Y., Rouleau, M., Rouleau, M., Roberge, J., Miard, S., Picard, F., Tetu, B., & Guillemette, C. (2017). Cross-talk between alternatively spliced UGT1A isoforms and Colon Cancer cell metabolism. *Molecular Pharmacology 91*(3), 167–177.
Baralle, F. E., & Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nature Reviews Molecular Cell Biology 18*(7), 437–451.
Barré, L., Magdalou, J., Netter, P., Fournel-Gigleux, S., & Ouzzine, M. (2005). 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 272*(4), 1063–1071.
Basu, N. K., Kole, L., Basu, M., Chakraborty, K., Mitra, P. S., & Owens, I. S. (2008). The major chemical-detoxifying system of UDP-glucuronosyltransferases requires regulated phosphorylation supported by protein kinase C. *The Journal of Biological Chemistry 283*(34), 23048–23061.
Basu, N. K., Kovarova, M., Garza, A., Kubota, S., Saha, T., Mitra, P. S., Banerjee, R., Rivera, J., & Owens, I. S. (2005). Phosphorylation of a UDP-glucuronosyltransferase regulates substrate specificity. *Proceedings of the National Academy of Sciences of the United States of America 102*(18), 6285–6290.
Basu, N. K., Kubota, S., Meselhy, M. R., Ciotti, M., Chowdhury, B., Hartori, M., et al. (2004). Gastrointestinally distributed UDP-glucuronosyltransferase 1A10, which metabolizes estrogens and nonsteroidal anti-inflammatory drugs, depends upon phosphorylation. *The Journal of Biological Chemistry 279*(27), 28320–28329.
Bates, D. O., Morris, J. C., Oltean, S., & Donaldson, L. F. (2017). Pharmacology of modulators of alternative splicing. *Pharmacological Reviews 69*(1), 63–79.
Bellemare, J., Rouleau, M., Harvey, M., & Guillemette, C. (2010). Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *The Journal of Biological Chemistry 285*(6), 3600–3607.
Bellemare, J., Rouleau, M., Harvey, M., Popa, I., Pelletier, G., Tetu, B., et al. (2011). Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *The Journal of Pathology 223*(3), 425–435.
Blencowe, B. J. (2006). Alternative splicing: New insights from global analyses. *Cell 126*(1), 37–47.
Bosma, P. J., Chowdhury, J. R., Bakker, C., Gantla, S., de Boer, A., Oostra, B. A., et al. (1995). The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome [see comments]. *The New England Journal of Medicine 333*(18), 1171–1175.
Bosma, P. J., Seppen, J., Goldhoorn, B., Bakker, C., Oude Elferink, R. P., Chowdhury, J. R., Chowdhury, N. R., & Jansen, P. L. (1994). Bilirubin UDP-glucuronosyltransferase 1 is the only relevant bilirubin glucuronidating isoform in man.[Erratum appears in J Biol Chem 1994 Oct 14;269(41):2542]. *The Journal of Biological Chemistry 269*(27), 17960–17964.
Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., & Chambon, P. (1978). Ovalbumin gene: Evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proceedings of the National Academy of Sciences of the United States of America 75*(10), 4853–4857.
Bushey, R. T., Dluzen, D. F., & Lazarus, P. (2013). Importance of UDP-glucuronosyltransferases 2A2 and 2A3 in tobacco carcinogen metabolism. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 41*(1), 170–179.
Bushey, R. T., & Lazarus, P. (2012). Identification and functional characterization of a novel UDP-glucuronosyltransferase 2A1 splice variant: Potential importance in tobacco-related cancer susceptibility. *The Journal of Pharmacology and Experimental Therapeutics 343*(3), 712–724.
Butler, L. M., Duguay, Y., Millikan, R. C., Sinha, R., Gagne, J. -F., Sandler, R. S., et al. (2005). Joint effects between UDP-glucuronosyltransferase 1A7 genotype and dietary carcinogen exposure on risk of colon cancer. *Cancer Epidemiology Biomarkers & Prevention 14*(7), 1626–1632.
Campbell, J. A., Davies, G. J., Bulone, V., & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *The Biochemical Journal 326*(Pt 3), 929–939.
Chakraborty, S. K., Basu, N. K., Jana, S., Basu, M., Raychoudhuri, A., & Owens, I. S. (2012). Protein kinase Calpha and Src kinase support human prostate-distributed dihydrotestosterone-metabolizing UDP-glucuronosyltransferase 2B15 activity. *The Journal of Biological Chemistry 287*(29), 24387–24396.
Chen, P., Zhu, K. -W., Zhang, D. -Y., Yan, H., Liu, H., Liu, Y. -L., Cao, S., Zhou, G., Zeng, H., Chen, S. -P., Zhao, X. -L., Yang, J., & Chen, X. -P. (2018). Influence of UGT1A1 polymorphisms on the outcome of acute myeloid leukemia patients treated with cytarabine-base regimens. *Journal of Translational Medicine 16*(1), 197.
Chouinard, S., Yueh, M. F., Tukey, R. H., Giton, F., Fiet, J., Pelletier, G., Barbier, O., & Belanger, A. (2008). Inactivation by UDP-glucuronosyltransferase enzymes: The end of androgen signaling. *The Journal of Steroid Biochemistry and Molecular Biology 109*(3-5), 247–253.
Chwalenia, K., Facemire, L., & Li, H. (2017). Chimeric RNAs in cancer and normal physiology. *Wiley Interdisciplinary Reviews RNA 8*(6).
Ciotti, M., Cho, J. W., George, J., & Owens, I. S. (1998). Required buried alpha-helical structure in the bilirubin UDP-glucuronosyltransferase, UGT1A1, contains a nonreplaceable phenylalanine. *Biochemistry 37*(31), 11018–11025.
Climente-Gonzalez, H., Porta-Pardo, E., Godzik, A., & Eyras, E. (2017). The functional impact of alternative splicing in cancer. *Cell Reports 20*(9), 2215–2226.
Coetzee, T., Fujita, N., Dupree, J., Shi, R., Blight, A., Suzuki, K., Suzuki, K., & Popko, B. (1996). Myelination in the absence of galactocerebroside and sulfatide: normal structure with abnormal function and regional instability. *Cell 86*(2), 209–219.
Consortium, E. P. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature 489*(7414), 57–74.
Consortium, G. T. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics 45*(6), 580–585.
Costello, J. L., Castro, I. G., Hacker, C., Schrader, T. A., Metz, J., Zeuschner, D., Azadi, A. S., Godinho, L. F., Costina, V., Findeisen, P., Manner, A., Islinger, M., & Schrader, M. (2017). ACBD5 and VAPB mediate membrane associations between peroxisomes and the ER. *The Journal of Cell Biology 216*(2), 331–342.
Coughtrie, M. W., Blair, J. N., Hume, R., & Burchell, A. (1991). Improved preparation of hepatic microsomes for in vitro diagnosis of inherited disorders of the glucose-6-phosphatase system. *Clinical Chemistry 37*(5), 739–742.
Court, M. H. (2010). Interindividual variability in hepatic drug glucuronidation: Studies into the role of age, sex, enzyme inducers, and genetic polymorphism using the human liver bank as a model system. *Drug Metabolism Reviews 42*(1), 209–224.
Court, M. H., Hazarika, S., Krishnaswamy, S., Finel, M., & Williams, J. A. (2008). Novel polymorphic human UDP-glucuronosyltransferase 2A3: Cloning, functional characterization of enzyme variants, comparative tissue expression, and gene induction. *Molecular Pharmacology 74*(3), 744–754.
Court, M. H., Zhang, X., Ding, X., Yee, K. K., Hesse, L. M., & Finel, M. (2012). Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica 42*(3), 266–277.
Dvinge, H., Kim, E., Abdel-Wahab, O., & Bradley, R. K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews Cancer 16*(7), 413–430.
El Marabti, E., & Younis, I. (2018). The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. *Frontiers in Molecular Biosciences 5*, 80.
Fremont, J. J., Wang, R. W., & King, C. D. (2005). Coimmunoprecipitation of UDP-glucuronosyltransferase isoforms and cytochrome P450 3A4. *Molecular Pharmacology 67*(1), 260–262.
Fujiwara, R., & Itoh, T. (2014). Extensive protein-protein interactions involving UDP-glucuronosyltransferase (UGT) 2B7 in human liver microsomes. *Drug Metabolism and Pharmacokinetics 29*(3), 259–265.
Fujiwara, R., Nakajima, M., Oda, S., Yamanaka, H., Ikushiro, S., Sakaki, T., et al. (2010). Interactions between human UDP-glucuronosyltransferase (UGT) 2B7 and UGT1A enzymes. *Journal of Pharmaceutical Sciences 99*(1), 442–454.
Fujiwara, R., Nakajima, M., Yamanaka, H., Katoh, M., & Yokoi, T. (2007). Interactions between human UGT1A1, UGT1A4, and UGT1A6 affect their enzymatic activities. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 35*(10), 1781–1787.
Fujiwara, R., Nakajima, M., Yamanaka, H., Nakamura, A., Katoh, M., Ikushiro, S., Sakaki, T., & Yokoi, T. (2007). Effects of coexpression of UGT1A9 on enzymatic activities of human UGT1A isoforms. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 35*(5), 747–757.
Fujiwara, R., Yokoi, T., & Nakajima, M. (2016). Structure and protein-Protein interactions of human UDP-Glucuronosyltransferases. *Frontiers in Pharmacology 7*, 388.
Gauthier-Landry, L., Belanger, A., & Barbier, O. (2015). Multiple roles for UDP-glucuronosyltransferase (UGT)2B15 and UGT2B17 enzymes in androgen metabolism and prostate cancer evolution. *The Journal of Steroid Biochemistry and Molecular Biology 145*, 187–192.
Ghosh, S. S., Sappal, B. S., Kalpana, G. V., Lee, S. W., Chowdhury, J. R., & Chowdhury, N. R. (2001). Homodimerization of human bilirubin-uridine-diphosphoglucuronate glucuronosyltransferase-1 (UGT1A1) and its functional implications. *The Journal of Biological Chemistry 276*(45), 42108–42115.
Girard-Bock, C., Benoit-Biancamano, M. O., Villeneuve, L., Desjardins, S., & Guillemette, C. (2016). A rare UGT2B7 variant creates a novel N-Glycosylation site at codon 121 with impaired enzyme activity. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 44*(12), 1867–1871.
Girard, H., Levesque, E., Bellemare, J., Journault, K., Caillier, B., & Guillemette, C. (2007). 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. *Pharmacogenetics and Genomics 17*(12), 1077–1089.
Gong, Q. H., Cho, J. W., Huang, T., Potter, C., Gholami, N., Basu, N. K., Kubota, S., Carvalho, S., Pennington, M. W., Owens, I. S., & Popescu, N. C. (2001). Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics 11*(4), 357–368.
Green, M. D., & Tephly, T. R. (1989). N-glycosylation of purified rat and rabbit hepatic UDP-glucuronosyltransferases. *Archives of Biochemistry and Biophysics 273*(1), 72–78.
Gschaidmeier, H., & Bock, K. W. (1994). Radiation inactivation analysis of microsomal UDP-glucuronosyltransferases catalysing mono- and diglucuronide formation of 3,6-dihydroxybenzo(a)pyrene and 3,6-dihydroxychrysene. *Biochemical Pharmacology 48*(8), 1545–1549.
Guillemette, C., Levesque, E., Harvey, M., Bellemare, J., & Menard, V. (2010). UGT genomic diversity: Beyond gene duplication. *Drug Metabolism Reviews 42*(1), 24–44.
Hara, A., & Taketomi, T. (1982). Isolation and determination of cholesterol glucuronide in human liver. *Lipids 17*(8), 515–518.
Hayashi, T., Hayashi, E., Fujimoto, M., Sprong, H., & Su, T. -P. (2012). The lifetime of UDP-galactose: Ceramide galactosyltransferase is controlled by a distinct endoplasmic reticulum-associated degradation (ERAD) regulated by sigma-1 receptor chaperones. *The Journal of Biological Chemistry 287*(51), 43156–43169.
Hong, W., Guoxiu, C., Guangji, W., & Haiping, H. (2018). Regulation of mammalian UDP-Glucuronosyltransferases. *Current Drug Metabolism 19*(6), 490–501.
Hu, D. G., Hulin, J. -A., Wijayakumara, D. D., McKinnon, R. A., Mackenzie, P. I., & Meech, R. (2018). Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17,

2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions. *Mol Pharmacol: mol. 118*, 111773.
Hu, D. G., Mackenzie, P. I., McKinnon, R. A., & Meech, R. (2016). Genetic polymorphisms of human UDP-glucuronosyltransferase (UGT) genes and cancer risk. *Drug Metabolism Reviews 48*(1), 47–69.
Hu, D. G., Meech, R., McKinnon, R. A., & Mackenzie, P. I. (2014). Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metabolism Reviews 46*(4), 421–458.
Hu, D. G., Selth, L. A., Tarulli, G. A., Meech, R., Wijayakumara, D., Chanawong, A., Russell, R., Caldas, C., Robinson, J. L., Carroll, J. S., Tilley, W. D., Mackenzie, P. I., & Hickey, T. E. (2016). Androgen and estrogen receptors in breast cancer coregulate human UDP-Glucuronosyltransferases 2B15 and 2B17. *Cancer Research 76*(19), 5881–5893.
Hua, R., Cheng, D., Coyaud, É., Freeman, S., Di Pietro, E., Wang, Y., Vissa, A., Yip, C. M., Fairn, G. D., Braverman, N., Brumell, J. H., Trimble, W. S., Raught, B., & Kim, P. K. (2017). VAPs and ACBD5 tether peroxisomes to the ER for peroxisome maintenance and lipid homeostasis. *The Journal of Cell Biology 216*(2), 367–377.
Ichikawa, W., Araki, K., Fujita, K., Yamamoto, W., Endo, H., Nagashima, F., Tanaka, R., Miya, T., Kodama, K., Sunakawa, Y., Narabayashi, M., Ando, Y., Akiyama, Y., Kawara, K., & Sasaki, Y. (2008). Re: UGT1A1*28 genotype and irinotecan-induced neutropenia: Dose matters. *Journal of the National Cancer Institute 100*(3), 224–225 author reply 225.
Ikushiro, S., Emi, Y., & Iyanagi, T. (1997). Protein-protein interactions between UDP-glucuronosyltransferase isozymes in rat hepatic microsomes. *Biochemistry 36*(23), 7154–7161.
Ishii, Y., Iwanaga, M., Nishimura, Y., Takeda, S., Ikushiro, S. -I., Nagata, K., Yamazoe, Y., Mackenzie, P. I., & Yamada, H. (2007). Protein-protein interactions between rat hepatic cytochromes P450 (P450s) and UDP-glucuronosyltransferases (UGTs): Evidence for the functionally active UGT in P450-UGT complex. *Drug Metabolism and Pharmacokinetics 22*(5), 367–376.
Ishii, Y., Koba, H., Kinoshita, K., Oizaki, T., Iwamoto, Y., Takeda, S., Miyauchi, Y., Nishimura, Y., Egoshi, N., Taura, F., Morimoto, S., Ikushiro, S., Nagata, K., Yamazoe, Y., Mackenzie, P. I., & Yamada, H. (2014). Alteration of the function of the UDP-glucuronosyltransferase 1A subfamily by cytochrome P450 3A4: Different susceptibility for UGT isoforms and UGT1A1/7 variants. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 42*(2), 229–238.
Ishii, Y., Miyoshi, A., Maji, D., Yamada, H., & Oguri, K. (2004). Simultaneous expression of guinea pig UDP-glucuronosyltransferase 2B21 (UGT2B21) and 2B22 in COS-7 cells enhances UGT2B21-catalyzed chloramphenicol glucuronidation. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 32*(10), 1057–1060.
Ishii, Y., Miyoshi, A., Watanabe, R., Tsuruda, K., Tsuda, M., Yamaguchi-Nagamatsu, Y., Yoshisue, K., Tanaka, M., Maji, D., Ohgiya, S., & Oguri, K. (2001). Simultaneous expression of guinea pig UDP-glucuronosyltransferase 2B21 and 2B22 in COS-7 cells enhances UDP-glucuronosyltransferase 2B21-catalyzed morphine-6-glucuronide formation. *Molecular Pharmacology 60*(5), 1040–1048.
Iyanagi, T., Haniu, M., Sogawa, K., Fujii-Kuriyama, Y., Watanabe, S., Shively, J. E., et al. (1986). Cloning and characterization of cDNA encoding 3-methylcholanthrene inducible rat mRNA for UDP-glucuronosyltransferase. *The Journal of Biological Chemistry 261*(33), 15607–15614.
Izukawa, T., Nakajima, M., Fujiwara, R., Yamanaka, H., Fukami, T., Takamiya, M., Aoki, Y., Ikushiro, S., Sakaki, T., & Yokoi, T. (2009). Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 37*(8), 1759–1768.
Jedlitschky, G., Cassidy, A. J., Sales, M., Pratt, N., & Burchell, B. (1999). Cloning and characterization of a novel human olfactory UDP-glucuronosyltransferase. *The Biochemical Journal 340*(Pt 3), 837–843.
Jeffery, C. J. (2019). An enzyme in the test tube, and a transcription factor in the cell: Moonlighting proteins and cellular factors that affect their behavior. *Protein Science: A Publication of the Protein Society 28*(7), 1233–1238.
Jia, Y., Xie, Z., & Li, H. (2016). Intergenically spliced chimeric RNAs in cancer. *Trends in Cancer 2*(9), 475–484.
Jones, N. R., Sun, D., Freeman, W. M., & Lazarus, P. (2012). Quantification of Hepatic UDP glucuronosyltransferase 1A splice variant expression and correlation of UDP glucuronosyltransferase 1A1 variant expression with glucuronidation activity. *The Journal of Pharmacology and Experimental Therapeutics 342*(3), 720–729.
Jude, A. R., Little, J. M., Czernik, P. J., Tephly, T. R., Grant, D. F., & Radominska-Pandya, A. (2001). Glucuronidation of linoleic acid diols by human microsomal and recombinant UDP-Glucuronosyltransferases: Identification of UGT2B7 as the major isoform involved. *Archives of Biochemistry and Biophysics 389*(2), 176–186.
Jung, H., Lee, D., Lee, J., Park, D., Kim, Y. J., Park, W. Y., Hong, D., Park, P. J., & Lee, E. (2015). Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nature Genetics 47*(11), 1242–1248.
Kannan, K., Wang, L., Wang, J., Ittmann, M. M., Li, W., & Yen, L. (2011). Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proceedings of the National Academy of Sciences of the United States of America 108*(22), 9172–9177.
Kempner, E. S., & Schlegel, W. (1979). Size determination of enzymes by radiation inactivation. *Analytical Biochemistry 92*(1), 2–10.
Kim, M. S., Pinto, S. M., Getnet, D., Nirujogi, R. S., Manda, S. S., Chaerkady, R., Madugundu, A. K., Kelkar, D. S., Isserlin, R., Jain, S., Thomas, J. K., Muthusamy, B., Leal-Rojas, P., Kumar, P., Sahasrabuddhe, N. A., Balakrishnan, L., Advani, J., George, B., Renuse, S., Selvan, L. D., Patil, A. H., Nanjappa, V., Radhakrishnan, A., Prasad, S., Subbannayya, T. , Raju, R., Kumar, M., Sreenivasamurthy, S. K., Marimuthu, A., Sathe, G. J., Chavan, S., Datta, K. K., Subbannayya, Y., Sahu, A., Yelamanchi, S. D., Jayaram, S., Rajagopalan, P. , Sharma, J., Murthy, K. R., Syed, N., Goel, R., Khan, A. A., Ahmad, S., Dey, G., Mudgal, K., Chatterjee, A., Huang, T. C., Zhong, J., Wu, X., Shaw, P. G., Freed, D., Zahari, M. S., Mukherjee, K. K., Shankar, S., Mahadevan, A., Lam, H., Mitchell, C. J., Shankar, S. K., Satishchandra, P., Schroeder, J. T., Sirdeshmukh, R., Maitra, A., Leach, S. D., Drake, C. G., Halushka, M. K., Prasad, T. S., Hruban, R. H., Kerr, C. L., Bader, G. D., Iacobuzio-Donahue, C. A., Gowda, H., & Pandey, A. (2014). A draft map of the human proteome. *Nature 509*(7502), 575–581.
Koiwai, O., Aono, S., Adachi, Y., Kamisako, T., Yasui, Y., Nishizawa, M., et al. (1996). Crigler-Najjar syndrome type II is inherited both as a dominant and as a recessive trait. *Human Molecular Genetics 5*(5), 645–647.
Konopnicki, C. M., Dickmann, L. J., Tracy, J. M., Tukey, R. H., Wienkers, L. C., & Foti, R. S. (2013). Evaluation of UGT protein interactions in human hepatocytes: Effect of siRNA down regulation of UGT1A9 and UGT2B7 on propofol glucuronidation in human hepatocytes. *Archives of Biochemistry and Biophysics 535*(2), 143–149.
Kua, L. -F., Ross, S., Lee, S. -C., Mimura, K., Kono, K., Goh, B. -C., et al. (2012). UGT1A6 polymorphisms modulated lung cancer risk in a Chinese population. *PloS One 7*(8), e42873.
Kurkela, M., Garcia-Horsman, J. A., Luukkanen, L., Morsky, S., Taskinen, J., Baumann, M., Kostiainen, R., Hirvonen, J., & Finel, M. (2003). Expression and characterization of recombinant human UDP-glucuronosyltransferases (UGTs). UGT1A9 is more resistant to detergent inhibition than other UGTs and was purified as an active dimeric enzyme. *The Journal of Biological Chemistry 278*(6), 3536–3544.
Kurkela, M., Patana, A. -S., Mackenzie, P. I., Court, M. H., Tate, C. G., Hirvonen, J., Goldman, A., & Finel, M. (2007). Interactions with other human UDP-glucuronosyltransferases attenuate the consequences of the Y485D mutation on the activity and substrate affinity of UGT1A6. *Pharmacogenetics and Genomics 17*(2), 115–126.
Labriet, A., Allain, E. P., Rouleau, M., Audet-Delage, Y., Villeneuve, L., & Guillemette, C. (2018). Post-transcriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 46*(5), 514–524.
Lee, Y., & Rio, D. C. (2015). Mechanisms and regulation of alternative Pre-mRNA splicing. *Annual Review of Biochemistry 84*, 291–323.
Lei, Q., Li, C., Zuo, Z., Huang, C., Cheng, H., & Zhou, R. (2016). Evolutionary insights into RNA trans-splicing in vertebrates. *Genome Biology and Evolution 8*(3), 562–577.
Levesque, E., Girard, H., Journault, K., Lepine, J., & Guillemette, C. (2007). Regulation of the UGT1A1 bilirubin-conjugating pathway: Role of a new splicing event at the UGT1A locus. *Hepatology 45*(1), 128–138.
Levesque, E., Menard, V., Laverdiere, I., Bellemare, J., Barbier, O., Girard, H., et al. (2010). Extensive splicing of transcripts encoding the bile acid-conjugating enzyme UGT2B4 modulates glucuronidation. *Pharmacogenetics and Genomics 20*(3), 195–210.
Lewis, B. C., Mackenzie, P. I., & Miners, J. O. (2011). Homodimerization of UDP-glucuronosyltransferase 2B7 (UGT2B7) and identification of a putative dimerization domain by protein homology modeling. *Biochemical Pharmacology 82*(12), 2016–2023.
Li, H., Xie, N., Chen, R., Verreault, M., Fazli, L., Gleave, M. E., Barbier, O., & Dong, X. (2016). UGT2B17 expedites progression of castration-resistant prostate cancers by promoting ligand-independent AR signaling. *Cancer Research 76*(22), 6701–6711.
Little, J. M., Williams, L., Xu, J., & Radominska-Pandya, A. (2002). Glucuronidation of the dietary fatty acids, phytanic acid and docosahexaenoic acid, by human UDP-glucuronosyltransferases. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 30*(5), 531–533.
Liu, Y., Gonzalez-Porta, M., Santos, S., Brazma, A., Marioni, J. C., Aebersold, R., Venkitaraman, A. R., & Wickramasinghe, V. O. (2017). Impact of alternative splicing on the human proteome. *Cell Reports 20*(5), 1229–1241.
Liu, Y. Q., Yuan, L. M., Gao, Z. Z., Xiao, Y. S., Sun, H. Y., Yu, L. S., et al. (2016). Dimerization of human uridine diphosphate glucuronosyltransferase allozymes 1A1 and 1A9 alters their quercetin glucuronidation activities. *Scientific Reports 6*, 23763.
Mackenzie, P. I. (1986). Rat liver UDP-glucuronosyltransferase. cDNA sequence and expression of a form glucuronidating 3-hydroxyandrogens. *The Journal of Biological Chemistry 261*(30), 14112–14117.
Mackenzie, P. I. (1990). The effect of N-linked glycosylation on the substrate preferences of UDP glucuronosyltransferases. *Biochemical and Biophysical Research Communications 166*(3), 1293–1299.
Mackenzie, P. I., Bock, K. W., Burchell, B., Guillemette, C., Ikushiro, S. -i., Iyanagi, T., Miners, J. O., Owens, I. S., & Nebert, D. W. (2005a). Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenetics and Genomics 15*(10), 677–685.
Mackenzie, P. I., Bock, K. W., Burchell, B., Guillemette, C., Ikushiro, S., Iyanagi, T., Miners, J. O., Owens, I. S., & Nebert, D. W. (2005b). Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenetics and Genomics 15*(10), 677–685.
Mackenzie, P. I., & Owens, I. S. (1984). Cleavage of nascent UDP glucuronosyltransferase from rat liver by dog pancreatic microsomes. *Biochemical and Biophysical Research Communications 122*(3), 1441–1449.
Mackenzie, P. I., Owens, I. S., Burchell, B., Bock, K. W., Bairoch, A., Belanger, A., Fournel-Gigleux, S., Green, M., Hum, D. W., Iyanagi, T., Lancet, D., Louisot, P., Magdalou, J., Chowdhury, J. R., Ritter, J. K., Schachter, H., Tephly, T. R., Tipton, K. F., & Nebert, D. W. (1997). The UDP glycosyltransferase gene superfamily: Recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics 7*(4), 255–269.
Mackenzie, P. I., Rogers, A., Elliot, D. J., Chau, N., Hulin, J. A., Miners, J. O., et al. (2011). The novel UDP glycosyltransferase 3A2: Cloning, catalytic properties, and tissue distribution. *Molecular Pharmacology 79*(3), 472–478.
Mackenzie, P. I., Rogers, A., Treloar, J., Jorgensen, B. R., Miners, J. O., & Meech, R. (2008). Identification of UDP glycosyltransferase 3A1 as a UDP N-acetylglucosaminyltransferase. *The Journal of Biological Chemistry 283*(52), 36205–36210.
Marquez, Y., Brown, J. W., Simpson, C., Barta, A., & Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Research 22*(6), 1184–1195.

Marquez, Y., Hopfler, M., Ayatollahi, Z., Barta, A., & Kalyna, M. (2015). Unmasking alternative splicing inside protein-coding exons defines exitrons and their role in proteome plasticity. *Genome Research 25*(7), 995–1007.

Matern, H., Matern, S., & Gerok, W. (1982). Isolation and characterization of rat liver microsomal UDP-glucuronosyltransferase activity toward chenodeoxycholic acid and testosterone as a single form of enzyme. *The Journal of Biological Chemistry 257*(13), 7422–7429.

Meech, R., Hu, D. G., McKinnon, R. A., Mubarokah, S. N., Haines, A. Z., Nair, P. C., Rowland, A., & Mackenzie, P. I. (2019). The UDP-Glycosyltransferase (UGT) superfamily: New members, new functions, and novel paradigms. *Physiological Reviews 99*(2), 1153–1222.

Meech, R., & Mackenzie, P. (1998). Determinants of UDP glucuronosyltransferase membrane association and residency in the endoplasmic reticulum. *Archives of Biochemistry and Biophysics 356*(1), 77–85.

Meech, R., & Mackenzie, P. I. (1997a). Structure and function of uridine diphosphate glucuronosyltransferases. *Clinical and Experimental Pharmacology & Physiology 24*(12), 907–915.

Meech, R., & Mackenzie, P. I. (1997b). UDP-glucuronosyltransferase, the role of the amino terminus in dimerization. *The Journal of Biological Chemistry 272*(43), 26913–26917.

Meech, R., & Mackenzie, P. I. (2010). UGT3A: Novel UDP-glycosyltransferases of the UGT superfamily. *Drug Metabolism Reviews 42*(1), 43–52.

Meech, R., Miners, J. O., Lewis, B. C., & Mackenzie, P. I. (2012). The glycosidation of xenobiotics and endogenous compounds: Versatility and redundancy in the UDP glycosyltransferase superfamily. *Pharmacology & Therapeutics 134*(2), 200–218.

Meech, R., Mubaroka, N., Shivasami, A., Rogers, A., Nair, P. C., Hu, D. G., McKinnon, R. A., & Mackenzie, P. I. (2014). A novel function for UDP glycosyltransferase 8 (UGT8): Galactosidation of bile acids. *Molecular pharmacology: mol. 114*, 093823.

Meech, R., Mubarokah, N., Shivasami, A., Rogers, A., Nair, P. C., Hu, D. G., McKinnon, R. A., & Mackenzie, P. I. (2015). A novel function for UDP glycosyltransferase 8: Galactosidation of bile acids. *Molecular Pharmacology 87*(3), 442–450.

Menard, V., Collin, P., Margaillan, G., & Guillemette, C. (2013). Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 41*(12), 2197–2205.

Menard, V., Eap, O., Harvey, M., Guillemette, C., & Levesque, E. (2009). Copy-number variations (CNVs) of the human sex steroid metabolizing genes UGT2B17 and UGT2B28 and their associations with a UGT2B15 functional polymorphism. *Human Mutation 30*(9), 1310–1319.

Menard, V., Eap, O., Roberge, J., Harvey, M., Levesque, E., & Guillemette, C. (2011). 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 21*(10), 631–641.

Miley, M. J., Zielinska, A. K., Keenan, J. E., Bratton, S. M., Radominska-Pandya, A., & Redinbo, M. R. (2007). Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *Journal of Molecular Biology 369*(2), 498–511.

Minchin, R. F., Lewis, A., Mitchell, D., Kadlubar, F. F., & McManus, M. E. (2008). Sulfotransferase 4A1. *The International Journal of Biochemistry & Cell Biology 40*(12), 2686–2691.

Mitra, P. S., Basu, N. K., Basu, M., Chakraborty, S., Saha, T., & Owens, I. S. (2011). Regulated phosphorylation of a major UDP-glucuronosyltransferase isozyme by tyrosine kinases dictates endogenous substrate selection for detoxification. *The Journal of Biological Chemistry 286*(2), 1639–1648.

Mitra, P. S., Basu, N. K., & Owens, I. S. (2009). Src supports UDP-glucuronosyltransferase-2B7 detoxification of catechol estrogens associated with breast cancer. *Biochemical and Biophysical Research Communications 382*(4), 651–656.

Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Research 10*(2), 459–472.

Nagaoka, K., Hanioka, N., Ikushiro, S., Yamano, S., & Narimatsu, S. (2012). The effects of N-glycosylation on the glucuronidation of zidovudine and morphine by UGT2B7 expressed in HEK293 cells. *Drug Metabolism and Pharmacokinetics 27*(4), 388–397.

Nair, P. C., Meech, R., Mackenzie, P. I., McKinnon, R. A., & Miners, J. O. (2015). Insights into the UDP-sugar selectivities of human UDP-glycosyltransferases (UGT): A molecular modeling perspective. *Drug Metabolism Reviews 47*(3), 335–345.

Nakajima, M., Koga, T., Sakai, H., Yamanaka, H., Fujiwara, R., & Yokoi, T. (2010). N-Glycosylation plays a role in protein folding of human UGT1A9. *Biochemical Pharmacology 79*(8), 1165–1172.

Nakajima, M., Yamanaka, H., Fujiwara, R., Katoh, M., & Yokoi, T. (2007). Stereoselective glucuronidation of 5-(4′-hydroxyphenyl)-5-phenylhydantoin by human UDP-glucuronosyltransferase (UGT) 1A1, UGT1A9, and UGT2B15: effects of UGT-UGT interactions. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 35*(9), 1679–1686.

Nakamura, A., Nakajima, M., Yamanaka, H., Fujiwara, R., & Yokoi, T. (2008). Expression of UGT1A and UGT2B mRNA in human normal tissues and various cell lines. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 36*(8), 1461–1464.

Nakamura, T., Yamaguchi, N., Miyauchi, Y., Takeda, T., Yamazoe, Y., Nagata, K., Mackenzie, P. I., Yamada, H., & Ishii, Y. (2016). Introduction of an N-Glycosylation site into UDP-Glucuronosyltransferase 2B3 alters its sensitivity to cytochrome P450 3A1-Dependent modulation. *Frontiers in Pharmacology 7*, 427.

Nakano, M., & Nakajima, M. (2018). Current knowledge of microRNA-mediated regulation of drug metabolism in humans. *Expert Opinion on Drug Metabolism & Toxicology 14*(5), 493–504.

Nishimura, M., & Naito, S. (2006). Tissue-specific mRNA expression profiles of human phase I metabolizing enzymes except for cytochrome P450 and phase II metabolizing enzymes. *Drug Metabolism and Pharmacokinetics 21*(5), 357–374.

Nixon-Abell, J., Obara, C. J., Weigel, A. V., Li, D., Legant, W. R., Xu, C. S., Pasolli, H. A., Harvey, K., Hess, H. F., Betzig, E., Blackstone, C., & Lippincott-Schwartz, J. (2016). Increased spatiotemporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science 354*(6311), aaf3928.

O'Dwyer, P. J., & Catalano, R. B. (2006). Uridine diphosphate glucuronosyltransferase (UGT) 1A1 and irinotecan: Practical pharmacogenomics arrives in cancer therapy. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology 24*(28), 4534–4538.

Ohno, S., & Nakajin, S. (2009). Determination of mRNA expression of human UDP-glucuronosyltransferases and application for localization in various human tissues by real-time reverse transcriptase-polymerase chain reaction. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 37*(1), 32–40.

Omenn, G. S., Lane, L., Lundberg, E. K., Beavis, R. C., Overall, C. M., & Deutsch, E. W. (2016). Metrics for the human proteome project 2016: Progress on identifying and characterizing the human proteome, including post-translational modifications. *Journal of Proteome Research 15*(11), 3951–3960.

Operana, T. N., & Tukey, R. H. (2007). Oligomerization of the UDP-glucuronosyltransferase 1A proteins: Homo- and heterodimerization analysis by fluorescence resonance energy transfer and co-immunoprecipitation. *The Journal of Biological Chemistry 282*(7), 4821–4829.

Ouzzine, M., Magdalou, J., Burchell, B., & Fournel-Gigleux, S. (1999). An internal signal sequence mediates the targeting and retention of the human UDP-Glucuronosyltransferase 1A6 to the endoplasmic reticulum. *The Journal of Biological Chemistry 274*(44), 31401–31409.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J., & Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics 40*(12), 1413–1415.

Park, E., Pan, Z., Zhang, Z., Lin, L., & Xing, Y. (2018). The expanding landscape of alternative splicing variation in human populations. *American Journal of Human Genetics 102*(1), 11–26.

Parra, G., Reymond, A., Dabbouseh, N., Dermitzakis, E. T., Castelo, R., Thomson, T. M., Antonarakis, S. E., & Guigo, R. (2006). Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Research 16*(1), 37–44.

Peer, C. J., Sissung, T. M., Kim, A., Jain, L., Woo, S., Gardner, E. R., Kirkland, C. T., Troutman, S. M., English, B. C., Richardson, E. D., Federspiel, J., Venzon, D., Dahut, W., Kohn, E., Kummar, S., Yarchoan, R., Giaccone, G., Widemann, B., & Figg, W. D. (2012). Sorafenib is an inhibitor of UGT1A1 but is metabolized by UGT1A9: Implications of genetic variants on pharmacokinetics and hyperbilirubinemia. *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research 18*(7), 2099–2107.

Peters, W. H., Jansen, P. L., & Nauta, H. (1984). The molecular weights of UDP-glucuronyltransferase determined with radiation-inactivation analysis. A molecular model of bilirubin UDP-glucuronyltransferase. *The Journal of Biological Chemistry 259*(19), 11701–11705.

Petrotchenko, E. V., Pedersen, L. C., Borchers, C. H., Tomer, K. B., & Negishi, M. (2001). The dimerization motif of cytosolic sulfotransferases. *FEBS Letters 490*(1-2), 39–43.

Radominska-Pandya, A., Pokrovskaya, I. D., Xu, J., Little, J. M., Jude, A. R., Kurten, R. C., et al. (2002). Nuclear UDP-glucuronosyltransferases: Identification of UGT2B7 and UGT1A6 in human liver nuclear membranes. *Archives of Biochemistry and Biophysics 399*(1), 37–48.

Reed, J. R., & Backes, W. L. (2012). Formation of P450 · P450 complexes and their effect on P450 function. *Pharmacology & Therapeutics 133*(3), 299–310.

Ritter, J. K., Chen, F., Sheen, Y. Y., Tran, H. M., Kimura, S., Yeatman, M. T., et al. (1992). A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *The Journal of Biological Chemistry 267*(5), 3257–3261.

Rouleau, M., Audet-Delage, Y., Desjardins, S., Rouleau, M., Girard-Bock, C., & Guillemette, C. (2017). endogenous protein interactome of human UDP-glucuronosyltransferases exposed by untargeted proteomics. *Frontiers in Pharmacology 8*, 23.

Rouleau, M., Roberge, J., Bellemare, J., & Guillemette, C. (2014). Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Molecular Pharmacology 85*(1), 29–36.

Rouleau, M., Tourancheau, A., Girard-Bock, C., Villeneuve, L., Vaucher, J., Duperré, A. -M., Audet-Delage, Y., Gilbert, I., Popa, I., Droit, A., & Guillemette, C. (2016). Divergent expression and metabolic functions of human glucuronosyltransferases through alternative splicing. *Cell Reports 17*(1), 114–124.

Roy, B., Haupt, L. M., & Griffiths, L. R. (2013). Review: Alternative splicing (AS) of genes as an approach for generating protein complexity. *Current Genomics 14*(3), 182–194.

Schaefer, O., Ohtsuki, S., Kawakami, H., Inoue, T., Liehner, S., Saito, A., Sakamoto, A., Ishiguro, N., Matsumaru, T., Terasaki, T., & Ebner, T. (2012). Absolute quantification and differential expression of drug transporters, cytochrome P450 enzymes, and UDP-glucuronosyltransferases in cultured primary human hepatocytes. *Drug Metabolism and Disposition: the Biological Fate of Chemicals 40*(1), 93–103.

Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nature Reviews Genetics 17*(1), 19–32.

Servedio, V., d'Apolito, M., Maiorano, N., Minuti, B., Torricelli, F., Ronchi, F., Zancan, L., Perrotta, S., Vajro, P., Boschetto, L., & Iolascon, A. (2005). Spectrum of UGT1A1 mutations in Crigler-Najjar (CN) syndrome patients: Identification of twelve novel alleles and genotype-phenotype correlation. *Human Mutation 25*(3), 325.

Shepherd, S. R., Baird, S. J., Hallinan, T., & Burchell, B. (1989). An investigation of the transverse topology of bilirubin UDP-glucuronosyltransferase in rat hepatic endoplasmic reticulum. *The Biochemical Journal 259*(2), 617–620.

Sibley, C. R., Emmett, W., Blazquez, L., Faro, A., Haberman, N., Briese, M., Trabzuni, D., Ryten, M., Weale, M. E., Hardy, J., Modic, M., Curk, T., Wilson, S. W., Plagnol, V., & Ule, J. (2015). Recursive splicing in long vertebrate genes. *Nature 521*(7552), 371–375.

Sidharthan, N. P., Butcher, N. J., Mitchell, D. J., & Minchin, R. F. (2014). Expression of the orphan cytosolic sulfotransferase SULT4A1 and its major splice variant in human

tissues and cells: Dimerization, degradation and polyubiquitination. *PloS One 9*(7), e101520.
Sneitz, N., Court, M. H., Zhang, X., Laajanen, K., Yee, K. K., Dalton, P., Ding, X., & Finel, M. (2009). Human UDP-glucuronosyltransferase UGT2A2: cDNA construction, expression, and functional characterization in comparison with UGT2A1 and UGT2A3. *Pharmacogenetics and Genomics 19*(12), 923–934.
Sprong, H., Degroote, S., Nilsson, T., Kawakita, M., Ishida, N., van der Sluijs, P., et al. (2003). Association of the Golgi UDP-galactose transporter with UDP-galactose: Ceramide galactosyltransferase allows UDP-galactose import in the endoplasmic reticulum. *Molecular Biology of the Cell 14*(8), 3482–3493.
Staines, A. G., Sindelar, P., Coughtrie, M. W. H., & Burchell, B. (2004). Farnesol is glucuronidated in human liver, kidney and intestine in vitro, and is a novel substrate for UGT2B7 and UGT1A1. *The Biochemical Journal 384*(Pt 3), 637–645.
Sun, D., Liu, Y., McCloskey, D. E., & Lazarus, P. (2011). Abstract 712: Characterization of exemestane metabolic pathways. *Cancer Research 71*(8 Supplement), 712.
Supek, F., Minana, B., Valcarcel, J., Gabaldon, T., & Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. *Cell 156*(6), 1324–1335.
Suzuki, M., Hirata, M., Takagi, M., Watanabe, T., Iguchi, T., Koiwai, K., Maezawa, S., & Koiwai, O. (2014). Truncated UDP-glucuronosyltransferase (UGT) from a Crigler-Najjar syndrome type II patient colocalizes with intact UGT in the endoplasmic reticulum. *Journal of Human Genetics 59*(3), 158–162.
Takeda, S., Ishii, Y., Iwanaga, M., Mackenzie, P. I., Nagata, K., Yamazoe, Y., Oguri, K., & Yamada, H. (2005). Modulation of UDP-glucuronosyltransferase function by cytochrome P450: Evidence for the alteration of UGT2B7-catalyzed glucuronidation of morphine by CYP3A4. *Molecular Pharmacology 67*(3), 665–672.
Takeda, S., Ishii, Y., Iwanaga, M., Nurrochmad, A., Ito, Y., Mackenzie, P. I., Nagata, K., Yamazoe, Y., Oguri, K., & Yamada, H. (2009). Interaction of cytochrome P450 3A4 and UDP-glucuronosyltransferase 2B7: Evidence for protein-protein association and possible involvement of CYP3A4 J-helix in the interaction. *Molecular Pharmacology 75*(4), 956–964.
Takeda, S., Ishii, Y., Mackenzie, P. I., Nagata, K., Yamazoe, Y., Oguri, K., et al. (2005). Modulation of UDP-glucuronosyltransferase 2B7 function by cytochrome P450s in vitro: Differential effects of CYP1A2, CYP2C9 and CYP3A4. *Biological & Pharmaceutical Bulletin 28*(10), 2026–2027.
Taura, K., Naito, E., Ishii, Y., Mori, M. A., Oguri, K., & Yamada, H. (2004). Cytochrome P450 1A1 (CYP1A1) inhibitor alpha-naphthoflavone interferes with UDP-glucuronosyltransferase (UGT) activity in intact but not in permeabilized hepatic microsomes from 3-methylcholanthrene-treated rats: Possible involvement of UGT-P450 interactions. *Biological & Pharmaceutical Bulletin 27*(1), 56–60.
Taura, K. I., Yamada, H., Hagino, Y., Ishii, Y., Mori, M. A., & Oguri, K. (2000). Interaction between cytochrome P450 and other drug-metabolizing enzymes: evidence for an association of CYP1A1 with microsomal epoxide hydrolase and UDP-glucuronosyltransferase. *Biochemical and Biophysical Research Communications 273*(3), 1048–1052.
Tourancheau, A., Margaillan, G., Rouleau, M., Gilbert, I., Villeneuve, L., Levesque, E., Droit, A., & Guillemette, C. (2016). Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *The Pharmacogenomics Journal 16*(1), 60–70.
Tourancheau, A., Rouleau, M., Guauque-Olarte, S., Villeneuve, L., Gilbert, I., Droit, A., et al. (2018). Quantitative profiling of the UGT transcriptome in human drug-metabolizing tissues. *The Pharmacogenomics Journal 18*(2), 251–261.
Tukey, R. H., & Tephly, T. R. (1981). Purification of properties of rabbit liver estrone and p-nitrophenol UDP-glucuronyltransferases. *Archives of Biochemistry and Biophysics 209*(2), 565–578.
Turgeon, D., Carrier, J. S., Levesque, E., Beatty, B. G., Belanger, A., & Hum, D. W. (2000). Isolation and characterization of the human UGT2B15 gene, localized within a cluster of UGT2B genes and pseudogenes on chromosome 4. *Journal of Molecular Biology 295*(3), 489–504.
Turgeon, D., Chouinard, S., Belanger, P., Picard, S., Labbe, J. F., Borgeat, P., et al. (2003). Glucuronidation of arachidonic and linoleic acid metabolites by human UDP-glucuronosyltransferases. *Journal of Lipid Research 44*(6), 1182–1191.
van der Bijl, P., Strous, G. J., Lopes-Cardozo, M., Thomas-Oates, J., & van Meer, G. (1996). Synthesis of non-hydroxy-galactosylceramides and galactosyldiglycerides by hydroxy-ceramide galactosyltransferase. *The Biochemical Journal 317*, 589–597.
Vance, J. E. (1990). Phospholipid synthesis in a membrane fraction associated with mitochondria. *The Journal of Biological Chemistry 265*(13), 7248–7256.
Vaquero-Garcia, J., Barrera, A., Gazzara, M. R., Gonzalez-Vallinas, J., Lahens, N. F., Hogenesch, J. B., Lynch, K. W., & Barash, Y. (2016). A new view of transcriptome complexity and regulation through the lens of local splicing variations. *Elife 5*, e11752.
Wade, A. P. (1970). The estimation of cholesterol sulphate and cholesterol glucuronide in plasma. *Clinica Chimica Acta 27*(1), 109–116.
Wahl, M. C., Will, C. L., & Luhrmann, R. (2009). The spliceosome: Design principles of a dynamic RNP machine. *Cell 136*(4), 701–718.
Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., & Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature 456*(7221), 470–476.
Wishart, G. J., & Fry, D. J. (1980). Evidence from rat liver nuclear preparations that latency of microsomal UDP-glucuronosyltransferase is associated with vesiculation. *The Biochemical Journal 186*(3), 687–691.
Yao, L., Qiu, L. X., Yang, Z., Yu, X. J., Zhong, Y., et al. (2010). The association between TA-repeat polymorphism in the promoter region of UGT1A1 and breast cancer risk: A meta-analysis. *Breast Cancer Research and Treatment 122*(3), 879–882.
Yuan, L. -M., Gao, Z. -Z., Sun, H. -Y., Qian, S. -N., Xiao, Y. -S., Sun, L. -L., et al. (2016). Inter-isoform Hetero-dimerization of Human UDP-Glucuronosyltransferases (UGTs) 1A1, 1A9, and 2B7 and Impacts on Glucuronidation Activity. *Scientific Reports 6*, 34450.
Yuan, L., Qian, S., Xiao, Y., Sun, H., & Zeng, S. (2015). Homo- and hetero-dimerization of human UDP-glucuronosyltransferase 2B7 (UGT2B7) wild type and its allelic variants affect zidovudine glucuronidation activity. *Biochemical Pharmacology 95*(1), 58–70.
Zahreddine, H. A., & Borden, K. L. (2015). Molecular pathways: GLI1-Induced drug glucuronidation in resistant Cancer cells. *Clinical Cancer Research : an Official Journal of the American Association for Cancer Research 21*(10), 2207–2210.
Zhu, Z., Chung, Y. M., Sergeeva, O., Kepe, V., Berk, M., Li, J., Ko, H. K., Li, Z., Petro, M., DiFilippo, F. P., Lee, Z., & Sharifi, N. (2018). Loss of dihydrotestosterone-inactivation activity promotes prostate cancer castration resistance detectable by functional imaging. *The Journal of Biological Chemistry 293*(46), 17829–17837.