



Un nouveau variant rare entraînant la modification  
post-traductionnelle de l'enzyme UGT2B7 et affectant  
son activité

**Mémoire**

**Camille Girard-Bock**

**Maîtrise en sciences pharmaceutiques**  
Maître ès sciences (M. Sc.)

Québec, Canada

© Camille Girard-Bock, 2016

Un nouveau variant rare entraînant la modification  
post-traductionnelle de l'enzyme UGT2B7 et affectant  
son activité

**Mémoire**

**Camille Girard-Bock**

Sous la direction de :

Chantal Guillemette, directrice de recherche

## Résumés

La superfamille des UDP-glucuronosyltransférases (UGT) est constituée de glycoprotéines résidant au réticulum endoplasmique et sujettes aux modifications post-traductionnelles (PTM, *post-translational modifications*). L'enzyme UGT2B7 est d'un intérêt particulier vu son action sur une grande variété de médicaments. La plupart des études actuelles n'ont pour sujet que les variants communs de cette enzyme et n'examinent donc qu'une fraction de la diversité génétique de celle-ci. En effet, les variants rares (fréquence allélique en deçà de 1%) peuvent potentiellement avoir un effet considérable puisqu'ils sont prédits comme étant bien plus nombreux que les variants communs au sein du génome humain. La présente étude fait état de la découverte d'un variant rare d'*UGT2B7* possédant un intérêt pharmacogénétique potentiel et encodant une substitution d'acide aminé au codon 121. Cette variation peu fréquente, retrouvée chez deux individus au sein d'une population de 305 sujets sains, mène à la traduction d'une asparagine (Asn) plutôt qu'un acide aspartique (Asp) au codon 121 (UGT2B7 p.D121N). Cette substitution est prédite comme créant un motif de N-glycosylation NX(S/T) subséquent validé par traitement à l'endoglycosidase de fractions microsomales issues de surexpressions dans des cellules HEK293 et par inhibition à la tunicamycine de la N-glycosylation d'UGT2B7 produites de façon endogène dans des HEK293. De plus, la présence d'un oligosaccharide additionnel sur l'enzyme UGT2B7, affectant potentiellement son repliement, résulte en la diminution, respectivement par 49 et 40%, de la formation de glucuronides à partir de la zidovudine et de l'acide mycophénolique. Une analyse de la base de données dbSNP a permis la découverte de 32 variants rares pouvant potentiellement créer ou abolir des motifs de N-glycosylation au sein d'enzymes UGT2B. Ensemble, ces variants ont le potentiel d'augmenter la proportion de la variance de la voie des UGT qui est expliquée par des modifications post-traductionnelles telles la N-glycosylation, affectant ainsi le métabolisme des médicaments.

The UDP-glucuronosyltransferase (UGT) superfamily consists of glycoproteins resident of the endoplasmic reticulum membranes that undergo post-translational modifications (PTM). UGT2B7 is of particular interest because of its action on a wide variety of drugs. Most studies currently survey common variants and are only examining a small fraction of the genetic diversity. However, rare variants (frequency <1%) might have significant effect as they are predicted to greatly outnumber common variants in the human genome. Here, we discovered a rare single nucleotide *UGT2B7* variant of potential pharmacogenetic relevance that encodes a nonconservative amino acid substitution at codon 121. This low-frequency variation, found in two individuals of a population of 305 healthy volunteers, leads to the translation of an asparagine (Asn) instead of an aspartic acid (Asp) (*UGT2B7* p.D121N). This amino acid change was predicted to create a putative N-glycosylation motif NX(S/T) subsequently validated upon endoglycosidase H treatment of microsomal fractions and inhibition of N-glycosylation of endogenously produced *UGT2B7* with tunicamycin from HEK293 cells. The presence of an additional N-linked glycan on the *UGT2B7* enzyme, likely affecting proper protein folding, resulted in a significant decrease, respectively by 49 and 40%, in the formation of zidovudine and mycophenolic acid glucuronides. A systematic survey of the dbSNP database uncovered 32 rare and naturally occurring missense variations predicted to create or disrupt N-glycosylation sequence motifs in the other *UGT2B* enzymes. Collectively, these variants have the potential to increase the proportion of variance explained in the UGT pathway due to changes in PTM such as N-linked glycosylation with consequences on drug metabolism.

# Table des matières

Résumés.....	iii
Liste des figures.....	vii
Liste des abréviations .....	viii
Remerciements .....	ix
Avant-propos .....	x
Introduction .....	1
1. La voie enzymatique de glucuronidation .....	1
1.1. Structure et fonction des enzymes UDP-glucuronosyltransférases .....	1
1.2. Organisation génomique des gènes codant pour les enzymes UGT .....	3
1.3. UGT2B7.....	3
2. Déterminants de la variabilité de glucuronidation.....	6
2.1. Facteurs environnementaux .....	6
2.2. Facteurs physiologiques.....	8
2.3. Facteurs génétiques et génomiques.....	9
2.3.1. Les polymorphismes .....	10
2.3.2. L'épissage alternatif.....	12
3. Les modifications post-traductionnelles.....	13
3.1. Incidence et effet sur le protéome.....	13
3.2. La N-glycosylation .....	14
3.3. La phosphorylation .....	17
3.4. L'ubiquitination .....	18
3.5. L'acétylation .....	18
Hypothèses et objectifs.....	19
Résultats : Un variant rare d'UGT2B7 crée un nouveau site de N-glycosylation et affecte l'activité de l'enzyme .....	21

Discussion.....	38
Conclusion.....	43
Annexe : Divergences dans l'expression et la fonction métabolique des glucuronosyl transférase humaines via l'épissage alternatif.....	45
Bibliographie .....	91

## Liste des figures

Figure 1. Organisation génomique des UGT humaines .....	2
Figure 2. Diagramme de la contribution de chaque UGT .....	4
Figure 3. Structure secondaire de la région C-terminale d'UGT2B7 .....	5
Figure 4. Variation dans l'activité de glucuronidation d'enzymes UGT au foie .....	7
Figure 5. Occurrence des modifications post-traductionnelles expérimentalement validées dans la base de données dbPTM.....	14

## Liste des abréviations

Acide mycophénolique (MPA)

Acide mycophénolique glucuronide (MPAG)

Acyl-MPA-glucuronide (AcMPAG)

Endoglycosidase H (Endo H)

Isoformes 2 (i2)

Réticulum endoplasmique (RE)

Single nucleotide polymorphism (SNP)

UDP-glucuronosyltransférases (UGT)

Uridine di-phosphate acide glucuronique (UDPGA)



## Remerciements

Je tiens d'abord à remercier le jury pour l'attention qu'il porte à ce mémoire.

Je veux également remercier ma directrice de recherche Dre Chantal Guillemette de m'avoir donné l'occasion de m'impliquer dans une variété de projets au sein de son laboratoire, j'y ai appris énormément. Je suis heureuse d'avoir eu la chance de débiter mes études graduées dans un tel environnement et une telle équipe, dans lesquels se reflète votre soif d'excellence et votre passion pour la recherche.

J'aimerais aussi remercier tous les gens qui m'ont côtoyée au laboratoire au cours de ma maîtrise, toute l'équipe *labo CG*. Merci à Yannick, Anaïs et surtout Lyne d'avoir toujours été présents pour me montrer quoi faire et répondre à mes questions dans le labo. Merci à Alan qui a été présent de la même façon mais à l'ordinateur plutôt qu'à la *bench*. Merci à nos chimistes, Pat et Véro, et aux professionnelles de recherche, Lyne, Andréa, Michèle et Sylvie, sans qui ni le projet au cœur de ce mémoire ni bien d'autres projets ne seraient possibles. Merci à tous mes collègues, ceux que j'ai vu terminer, Alan, Anaïs, Guillaume, Sandra, Sylvia, et ceux qui sont encore là, Adrien, Anne-Marie, Éric, Luciana, Yannick, pour toutes les heures de dîner et autres discussions sur des sujets tous plus improbables les uns que les autres, mais qui font de nous une si belle équipe. Un merci tout spécial à Michèle qui m'a guidée et écoutée.

Je souhaite aussi remercier ma famille. Merci à mes frères et sœurs, Simon, Louis, Pascale et Michelle, de toujours savoir me faire rire autant et d'avoir appris le mot *pharmacogénomique* pour pouvoir répondre à ceux qui vous demandent ce que fait votre grande sœur à Québec. Merci grand-maman Claire de ton appui et ta confiance. Merci à mon papa, Jean-François, pour ton enthousiasme face à la recherche et tes encouragements. Un énorme merci à ma maman, Marcelle, je ne pourrai jamais assez te remercier, je ne serais pas rendue où je suis sans toi, merci de croire en moi et de me pousser à me dépasser.

Enfin, merci à mon amoureux, Nico, pour ton soutien et tes encouragements dans les moments difficiles, pour ta patience face à la distance et surtout pour ton optimisme infini.

## Avant-propos

Ce mémoire, intitulé « Un nouveau variant rare entraînant la modification post-traductionnelle de l'enzyme UGT2B7 et affectant son activité », est construit sous la forme d'une insertion d'article et est présenté à la faculté des études supérieures de l'Université Laval en vue de l'obtention du grade de Maîtres ès Sciences (M. Sc.).

L'article qui constitue le cœur de ce mémoire s'intitule « A rare UGT2B7 variant creates a novel N-glycosylation site at codon 121 with impaired enzyme activity », j'en suis le premier auteur et il a été publié dans le journal *Drug Metabolism and Disposition* en octobre 2016. Ma directrice de recherche, Dre Guillemette, et Dre Benoit-Biancamano, alors candidate au doctorat, ont conceptualisé l'étude et contribué à l'écriture et à la révision du manuscrit. Dre Benoit-Biancamano a établi la lignée cellulaire surexprimant la protéine variante d'UGT2B7 au codon 121 en collaboration avec Lyne Villeneuve, professionnelle de recherche au laboratoire. Dre Benoit-Biancamano a aussi contribué aux essais enzymatiques utilisant le MPA comme substrat. La professionnelle de recherche Sylvie Desjardins a effectué les essais sur les microsomes avec l'endoglycosidase H et l'O-glycosidase ainsi que les traitements sur les cellules en expression transitoire avec la tunicamycine et la thapsigargine. Pour ma part, j'ai effectué les essais enzymatiques avec l'AZT et les immunobuvardages de type Western permettant de normaliser les activités enzymatiques et l'ensemble des analyses de cinétiques enzymatiques. J'ai fait l'analyse des variants ponctuels à partir de bases de données publiques et j'ai produit la figure s'y rattachant. Enfin, j'ai rédigé l'ébauche du manuscrit. Tous les auteurs ont contribué à l'analyse des données présentées dans l'article et à la révision du manuscrit. J'estime à 60% ma contribution à ce manuscrit.

Au cours de ma maîtrise, j'ai également travaillé sur des variants d'épissage des gènes UGT, particulièrement le variant UGT2B7\_n4, qui possède un exon 2 supplémentaire. Mes travaux ont été inclus dans l'article intitulé « Divergent expression and metabolic functions of human glucuronosyltransferases through alternative splicing » publié dans le journal *Cell Reports* en septembre 2016 et présenté en annexe de ce mémoire. Ma contribution porte spécifiquement sur la réalisation des expériences de mise au point de nouveaux anticorps, d'essais d'immunoprécipitation pour détecter la protéine variante par spectrométrie de

masse (figure S2) ainsi que par l'analyse des données de métabolomique non ciblée dans les modèles cellulaires HepG2 et HEK293 surexprimant ou non le variant. J'ai conceptualisé et produit les figures et tableaux associés, soit la figure 5F et le tableau S1. J'estime ma contribution à ce manuscrit à 20% et j'en suis 3<sup>ième</sup> auteur.

# Introduction

## 1. La voie enzymatique de glucuronidation

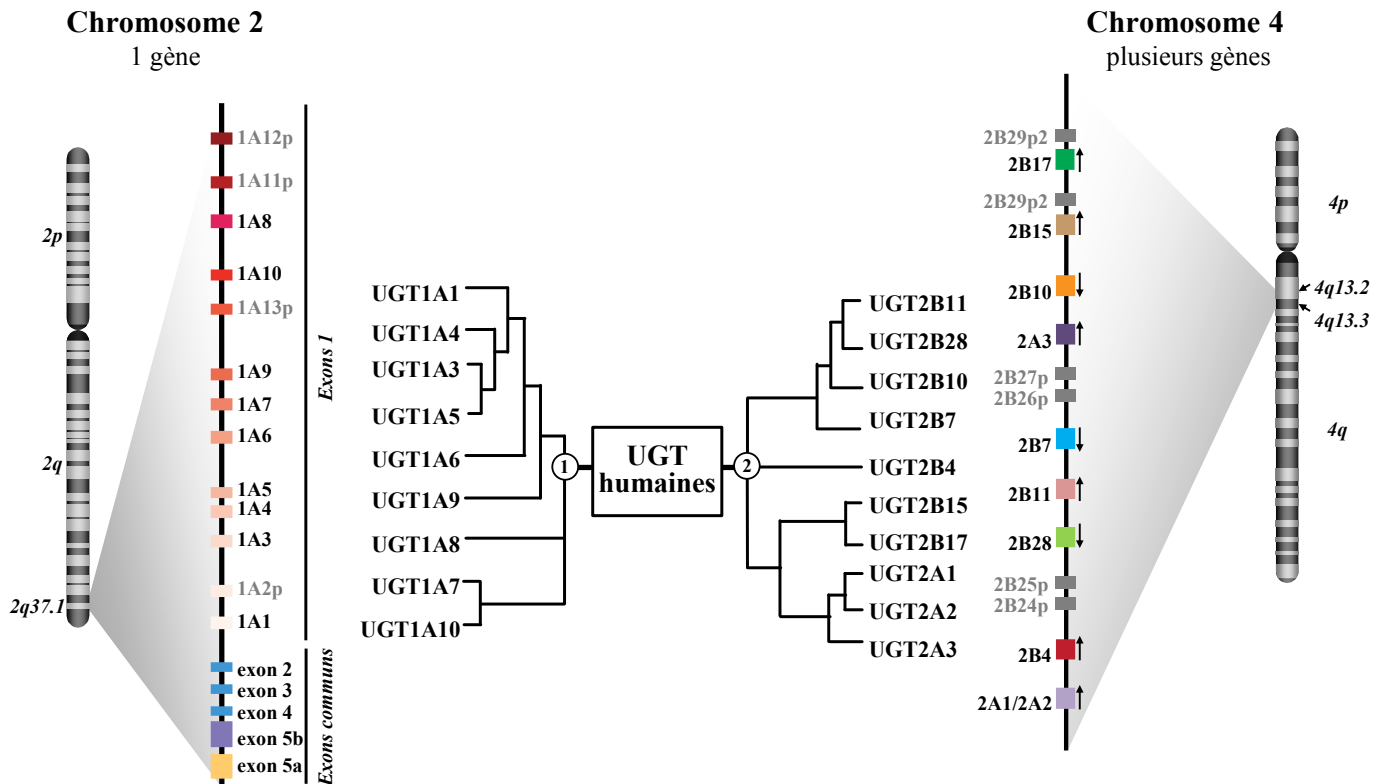
L'homéostasie est la capacité des êtres vivants à stabiliser leur environnement interne suite à une perturbation, qu'elle soit de nature endogène, telle la sécrétion d'une hormone, ou exogène, telle l'ingestion d'une toxine ou d'un médicament. Des réactions métaboliques interviennent alors dans le but de rendre biologiquement inactif le composé causant le déséquilibre et de lui permettre d'être excrété par l'organisme. Parmi ces réactions, la réaction de glucuronidation, qui d'un point de vue pharmacologique s'inscrit dans le métabolisme de phase II des médicaments. Cette réaction est également impliquée dans le maintien des concentrations endogènes de plusieurs composés (Bock 2015). La glucuronidation s'effectue en présence du cofacteur uridine di-phosphate acide glucuronique (UDPGA), un sucre provenant de l'UDP-glucose. La réaction consiste en le transfert de l'acide glucuronique de l'UDPGA à un groupement nucléophile du substrat. Les métabolites générés sont généralement inactifs, plus polaires et plus facilement excrétés de l'organisme via la bile et l'urine. Il existe cependant quelques exceptions. C'est notamment le cas de la morphine, dont le métabolite glucuronide sur son groupement hydroxyle alicyclique (morphine-6-glucuronide) conserve des propriétés analgésiques s'apparentant à celles de son composé parent, sans toutefois entraîner d'effets secondaires aussi importants (Coller et al. 2009).

### 1.1. Structure et fonction des enzymes UDP-glucuronosyltransférases

Enzymes du métabolisme de phase II, les UDP-glucuronosyltransférases (UGT) exercent des fonctions de détoxification au sein d'une panoplie d'organismes : mammifères, insectes, plantes, bactéries (Bock 2016). Chez l'humain, les UGT sont responsables d'une réaction de conjugaison produisant des métabolites glucurono-conjugués. Cette réaction, dite de glucuronidation, intervient dans le métabolisme de plus de 55% des médicaments les plus prescrits (Guillemette et al. 2014). Elle est également impliquée dans le métabolisme de molécules endogènes, telles la bilirubine ou les hormones stéroïdiennes. L'activité enzymatique de l'UGT envers un substrat donné est influencée par son profil d'expression tissulaire et cellulaire ainsi que par la nature de l'UGT. Chaque UGT possède pour chacun de ses substrats une affinité et même une

régiosélectivité qui lui est propre (Wu et al. 2012). Certains substrats peuvent être glucuronidés par plusieurs UGT alors que d'autres sont spécifiques à une seule enzyme. Ces derniers s'avèrent utiles afin d'établir quelles UGT sont actives dans un tissu donné ou envers un substrat donné, par essais de compétition par exemple (Seo et al. 2014, Margailan et al. 2015).

Cette activité est influencée par le positionnement du groupement nucléophile à l'intérieur de la poche catalytique de l'enzyme, lequel dépend de la position du groupement au sein de la molécule et de la nature des autres groupements qui s'y trouvent (Sorich et al. 2004, Wu et al. 2012). La structure tridimensionnelle exacte de cette poche catalytique est méconnue puisque la nature transmembranaire (au sein de la membrane du réticulum endoplasmique) de ces enzymes les rend difficile à cristalliser à des fins de cristallographie par rayon-X. En effet, le domaine de liaison au substrat des enzymes UGT se situe dans la portion N-terminale de la protéine, laquelle contient aussi le peptide signal qui guide l'insertion de la protéine dans le réticulum endoplasmique (RE) suite à sa traduction (Ouzzine et al. 1999). Ainsi, le site actif des UGT se retrouve dans la lumière du RE. Dans la portion C-terminale de la protéine se situent deux



**Figure 1. Organisation génomique des UGT humaines**  
Figure adaptée de (Guillemette et al. 2010).

domaines connus, le domaine de liaison au co-substrat UDPGA et un domaine transmembranaire (Meech and Mackenzie 1997).

## 1.2. Organisation génomique des gènes codant pour les enzymes UGT

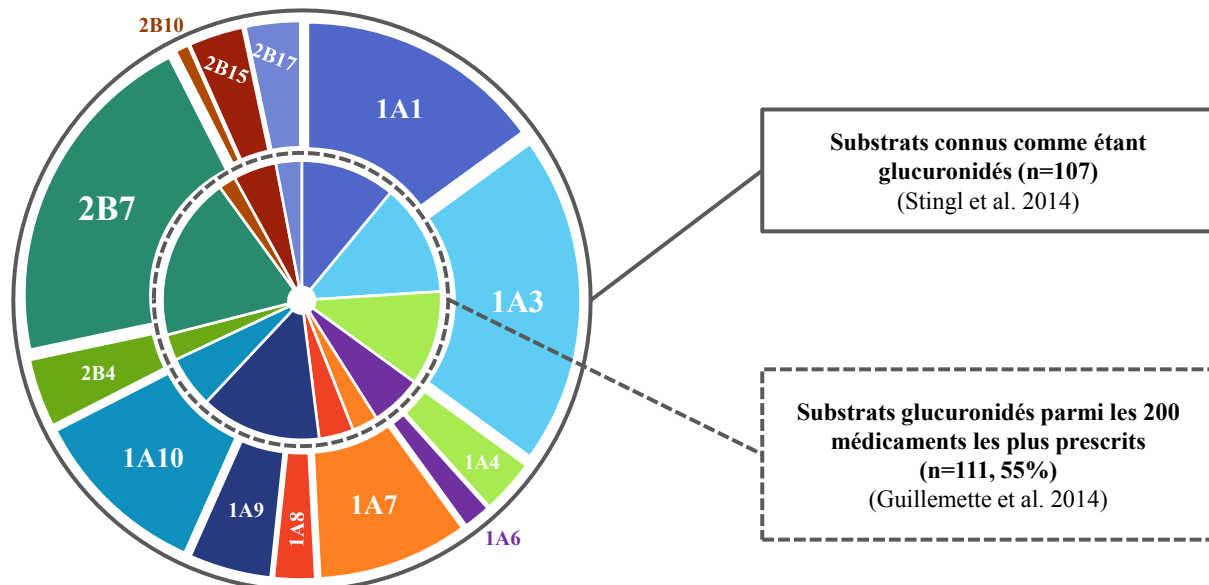
Le gène *UGT1* humain situé sur le chromosome 2 permet, grâce à des événements d'épissage alternatif impliquant des promoteurs et exons 1 alternatifs, l'expression de plus de neuf enzymes UGT1A (Gong et al. 2001). L'épissage alternatif d'exons 5 terminaux mène de surcroît à l'expressions d'autant de protéines tronquées, dites isoformes 2 ou i2, lesquelles régulent l'activité enzymatique des UGT1A\_i1 (Girard et al. 2007, Levesque et al. 2007). Située sur le chromosome 4, la sous-famille des *UGT2B* est composée de sept gènes distincts, chacun d'entre eux codant pour une enzyme UGT2B : UGT2B4, UGT2B7, UGT2B10, UGT2B11, UGT2B15, UGT2B17 et UGT2B28. La sous-famille des *UGT2A* est quant à elle composée de deux gènes qui codent pour 3 enzymes (Mackenzie et al. 2005, Owens et al. 2005). L'organisation génomique ainsi que la phylogénie de ces gènes sont présentées dans la figure 1.

## 1.3. UGT2B7

L'enzyme UGT2B7 est responsable de la glucuronidation de pratiquement un cinquième de tous les substrats connus comme étant glucuronidés (Stingl et al. (2014); Figure 2). Parmi les 200 médicaments les plus prescrits, 111 sont conjugués par les enzymes UGT. De ce nombre, 25% sont des substrats d'UGT2B7 (Guillemette et al. (2014); Figure 2). Ces statistiques laissent entrevoir quelques raisons qui font de cette enzyme la plus étudiée des UGT de la famille 2B. Ainsi, cette protéine est l'isoforme UGT majeure métabolisant plusieurs substrats exogènes, entre autres, les médicaments antirétroviraux efavirenz et zidovudine (AZT) (Belanger et al. 2009), le composé anti-cancéreux épirubicine (Innocenti et al. 2001), les anti-inflammatoires non-stéroïdiens diclofénac, clonixine et ibuprofène (Sakaguchi et al. 2004), la morphine (Gradinaru et al. 2015), le bêta-bloquant propranolol (Yu et al. 2010) ainsi que l'acide valproïque, un médicament antiépileptique (Wang et al. 2016). Elle est également impliquée dans le métabolisme de plusieurs substrats endogènes comme l'acide arachidonique, les prostaglandines E2 et B1 (Little et al. 2004) et plusieurs estrogènes (Lepine et al. 2004) et androgènes (Turgeon et al. 2001).

L'enzyme UGT2B7 est abondante au foie mais se retrouve aussi dans d'autres tissus du métabolisme des médicaments tels les reins mais également dans tous le système gastro-intestinal (à l'exception du rectum) ainsi qu'au pancréas (Court et al. 2012, Margaillan et al. 2015). Longue de 529 acides aminés, UGT2B7 apparaît à une hauteur d'environ 52kDa lors d'immunobuvardages de type Western.

Les UGT partagent une forte homologie de séquence, l'enzyme UGT2B7 possède donc les mêmes domaines protéiques que les autres membres de cette famille. Ainsi se trouvent dans la portion N-terminale de l'enzyme le peptide signal la guidant à la membrane du RE et le domaine de liaison au substrat. Bien que la conformation et les résidus importants de ce domaine de liaison restent peu connus, une étude de cette région via la création de protéines chimériques UGT2B7 et UGT2B15 met en lumière son rôle sur l'activité enzymatique et la spécificité du substrat (Lewis et al. 2007). En effet, l'UGT2B7 chimérique, pour laquelle les résidus 61 à 194 ont été remplacés par ceux d'UGT2B15, montrait une activité envers la testostérone et la phénolphtaléine alors qu'UGT2B7 (séquence de référence) ne présentait aucune activité pour ces substrats. De plus, l'étude a montré que les enzymes UGT2B15 chimériques avaient un profil enzymatique sigmoïde pour la glucuronidation du 4-méthylumbelliférone s'apparentant plus à celui d'UGT2B7 que d'UGT2B15. Enfin, ce type de profil sigmoïdal laisse croire aux auteurs



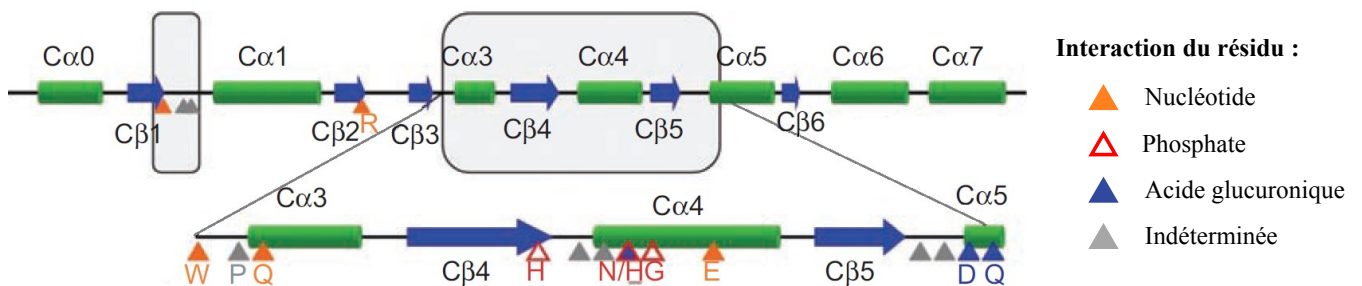
**Figure 2. Diagramme de la contribution de chaque UGT**  
 Contribution de chaque UGT à la glucuronidation de leurs substrats (cercle extérieur) et de ceux compris dans la liste des 200 médicaments les plus prescrits (cercle intérieur).  
 Données issues de (Guillemette et al. 2014, Stingl et al. 2014).

que cette région d'UGT2B7 posséderait des propriétés de dimérisation.

De telles propriétés ont été validées dans d'autres études, montrant également pour UGT2B7 une capacité de dimérisation avec d'autres UGT, telles les UGT1A, augmentant ainsi son affinité et sa vitesse de formation du glucuronide pour la AZT (Fujiwara et al. 2010). Une étude récente montre l'implication des codons 71 et 268 du domaine N-terminal dans la formation des dimères ainsi que l'effet de polymorphismes tels UGT2B7\*2 (H268Y), \*71S (A71S) et UGT2B7\*5 (D398N) sur cette dimérisation (Yuan et al. 2015).

La portion N-terminale de l'UGT2B7 comprend également deux motifs de N-glycosylation aux positions 67 et 68. La position 68 fut expérimentalement validée comme N-glycosylée et comme affectant l'activité enzymatique (Nagaoka et al. 2012). Également situé dans la portion N-terminale de l'enzyme, le variant ponctuel affectant le codon 121, dont il est question dans ce mémoire, présente par sa proximité au site de liaison au substrat un intérêt particulier puisqu'il crée à cet endroit un nouveau site de N-glycosylation, ce qui lui confère le potentiel de modifier l'activité enzymatique d'UGT2B7.

De son côté, la portion C-terminale de la protéine est mieux caractérisée, en effet il a été possible de déterminer la structure tridimensionnelle de la portion située entre les résidus 285 et 451 par cristallographie (Miley et al. 2007). Ces travaux ont permis d'identifier des résidus importants du site de liaison au co-substrat UDPGA en superposant la structure de cette région avec celle d'une glucosyl transférase végétale, VvGT1. Bien que l'homologie entre ces deux protéines soit faible (19%), les domaines de liaison au co-substrats se superposaient aisément, laissant deviner un



**Figure 3. Structure secondaire de la région C-terminale d'UGT2B7**

Les deux régions identifiées comme domaines de liaison au co-substrat sont encadrées en gris et les résidus importants sont indiqués par un triangle. Figure adaptée de (Radomska-Pandya et al. 2010).



mécanisme de liaison similaire. En combinant cette information et celle obtenue en alignant les séquences de plusieurs UGT pour observer les résidus conservés, les résidus soupçonnés d'être impliqués dans la liaison au co-substrat ont été identifiés. Les interactions de ces résidus ont été confirmés par des mutations ponctuelles et des essais enzymatiques et sont présentés dans la figure 3. Également, il se trouve dans cette portion de la protéine un site de N-glycosylation expérimentalement validé à la position 315 qui affecte également l'activité de l'enzyme (Nagaoka et al. 2012).

## 2. Déterminants de la variabilité de glucuronidation

Plusieurs facteurs entrent en jeu dans l'effet potentiel d'un xénobiotique sur le corps humain, tant du point de vue de son innocuité que de son efficacité. L'expression et l'activité des voies du métabolisme des médicaments font partie de ces facteurs (Evans 2004). Parmi ces voies, la glucuronidation, une réaction de conjugaison prise en charge par les enzymes UGT. Tel qu'illustré à la figure 4, où l'on peut observer une grande variabilité dans l'activité de plusieurs UGT retrouvées au foie, il existe une différence non négligeable entre les individus en ce qui a trait à leur capacité de glucuronidation (Court 2010).

### 2.1. Facteurs environnementaux

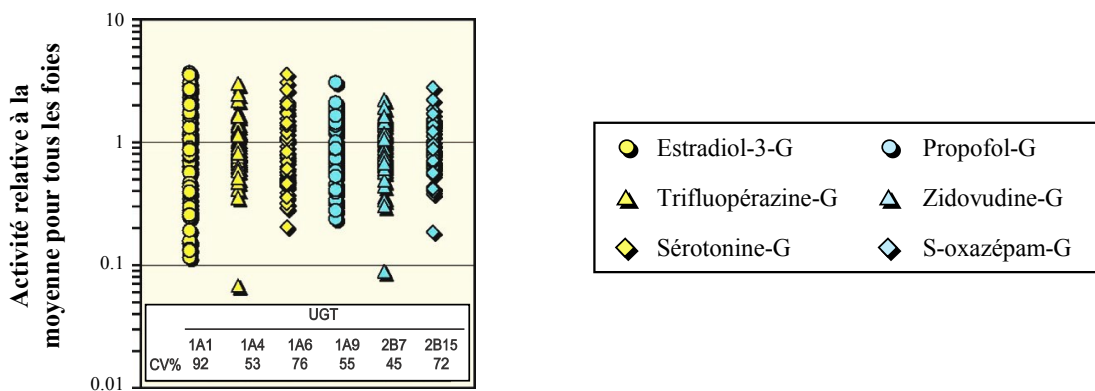
Une partie de la variabilité de glucuronidation observée entre individus est causée par des différences de nature environnementales. En effet, des variations dans l'alimentation, la prise de tabac, d'alcool, de produits naturels ou médicamenteux sont autant de différences pouvant entraîner cette variabilité.

Parmi les composés retrouvés dans l'alimentation pouvant influencer la pharmacocinétique de substrats des UGT se retrouve la vitamine A qui agit comme un co-régulateur (avec les hormones thyroïdiennes) de l'expression de gènes de la famille *Ugt1a* chez les rats (Haberkorn et al. 2002). L'effet que peut avoir l'alimentation s'étend également au concept de jeûne et de sous-nutrition. Ainsi, un jeûne de 36 heures induit chez le rat l'expression d'*Ugt1a1* (de Vries et al. 2015). Chez l'homme, la prise de nourriture, comparativement au jeûne, entraîne une

augmentation de la concentration plasmatique maximale de l'acide mycophénolique glucuronide (MPAG), le métabolite actif du médicament antirejet mycophénolate mofétil (Bullingham et al. 1996).

La consommation d'alcool a elle aussi la capacité de modifier la conjugaison par les enzymes UGT. Une étude portant sur la glucuronidation du lorazépam, une benzodiazépine principalement utilisée pour ses propriétés anxiolytiques, a en effet noté chez l'homme une diminution de 18% de la clairance du médicament suite à une administration d'éthanol sur une courte période (Hoyumpa et al. 1981). Une autre étude sur le médicament antiépileptique ézogabine, dont la conjugaison dépend elle aussi des UGT, dénote une augmentation de l'exposition et de la concentration plasmatique maximale lorsque le composé est ingéré avec de l'alcool (Crean and Tompson 2013).

Également, la prise de certains produits naturels, telles des herbes médicinales, peut influencer la glucuronidation. Par exemple, le millepertuis perforé est une herbe utilisée pour ses vertus antidépressives mais dont le composé actif, l'hyperforine, entraîne l'activation du récepteur nucléaire PXR (*pregnane X receptor*) qui à son tour est responsable d'une augmentation de l'expression d'UGT1A1 (Moore et al. 2000, Sugatani et al. 2012). Une étude *in vitro* montre de surcroît que de hautes concentrations de composés retrouvés dans le cannabis ont un impact sur la production de l'éthanol-glucuronide, dont 50% de la glucuronidation est faite par l'UGT1A9 et l'UGT2B7 (Al Saabi et al. 2013).



**Figure 4. Variation dans l'activité de glucuronidation d'enzymes UGT au foie**

Les substrats utilisés sont des substrats spécifiques aux UGT testées, les tests ont été effectués dans une même banque de microsomes de foie. Figure adaptée de (Court 2010).

L'utilisation de produits du tabac est aussi reconnue comme ayant un impact sur l'expression de plusieurs gènes des voies du métabolisme des médicaments. C'est le cas d'UGT2A1 : des cellules provenant de lavages broncho-alvéolaires de poumons de fumeurs ont une expression significativement plus faible de ce gène comparativement à celles de poumons de non-fumeurs (Thum et al. 2006).

Finalement, la prise de médicaments peut aussi contribuer à une variation dans la glucuronidation. De nombreuses occurrences d'interactions entre des médicaments et les enzymes UGT sont répertoriées. Par exemple, le diclophénac et l'acide méfénamique, deux anti-inflammatoires nonstéroïdiens, ont récemment été identifiés comme ayant un potentiel inhibiteur sur les enzymes UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A9 et UGT2B7 *in vitro* (Joo et al. 2015). Également, les médicaments antiépileptiques oxcarbazépine et lamotrigine sont éliminés par glucuronidation mais lorsqu'utilisés en combinaison le premier induirait l'expression d'UGT1A4 menant ainsi à une diminution de 29% de la concentration plasmatique du second (Johannessen and Landmark 2010).

## 2.2. Facteurs physiologiques

De nombreux facteurs physiologiques doivent être considérés lorsqu'on compare la capacité de glucuronidation entre individus, entre autres la masse corporelle, l'âge, le sexe, la fonction rénale et hépatique.

L'expression des UGT dans les tissus humains change en effet au cours du développement. D'une part, l'on observe une différence marquée de l'activité de glucuronidation entre la naissance et l'âge adulte. Plusieurs études suggèrent que cette différence s'explique en grande partie par le changement de masse corporelle au cours du développement et ces modèles s'appliquent chez les sujets d'âge pédiatrique en général. Cependant, le changement de masse corporelle seul ne permet pas d'expliquer la différence de l'activité de glucuronidation observée chez les nouveaux-nés (Wang et al. 2013, Mahmood 2015). Il y a effectivement un écart dans la capacité de glucuronidation des nouveaux-nés âgés de moins de 8 jours lorsque comparés aux nouveaux-nés plus âgés, lequel ne peut être attribué qu'à la différence de poids (Allegaert et al. 2008). Plusieurs UGT dont UGT2B7 et UGT1A9 ont en effet une expression et une activité qui croissent avec l'âge du nouveau-né (Zaya et al. 2006, Miyagi et al. 2012). D'autre part, au cours

de l'âge adulte, des différences, quoique plus modestes, liées à l'âge ont été observées dans la clairance de substrats des UGT. Une étude a récemment établi un lien entre l'âge et la clairance des antiépileptiques lamotrigine et oxcarbazépine, montrant dans une cohorte de plus de 1500 patients une augmentation de celle-ci jusqu'à la mi-trentaine, suivie d'une diminution (Wegner et al. 2013).

La grossesse est un autre facteur physiologique influençant la glucuronidation. Une étude antérieure du même groupe portant sur deux femmes prenant en combinaison les médicaments antiépileptiques mentionnés ci-haut au cours de leur grossesse a observé une augmentation des clairances de ces composés (Wegner et al. 2010). Cet effet s'explique, du moins en partie, par la hausse des concentrations des hormones estradiol et progestérone lesquelles activent respectivement les récepteurs nucléaires ER $\alpha$  et PXR qui à leur tour entraînent une augmentation de la transcription des gènes UGT (Jeong 2010).

En outre, des différences de conjugaison par les enzymes UGT sont observées entre les sexes. Une étude sur le propofol, un anesthésique général de courte durée administré par voie intraveineuse, a mis au jour des différences significatives entre les quantités de trois métabolites glucuronidés de ce composé dans le sang d'hommes et de femmes, ces dernières ayant des quantités de 1,7 à 2,1 fois plus élevées de glucuronides (Loryan et al. 2012).

Enfin, la fonction des tissus du métabolisme des médicaments, tels le foie et les reins, peut aussi expliquer une différence d'élimination d'un composé. Une étude comparant l'élimination du médicament antiépileptique rétigabine entre des sujets jeune (18-40 ans) et âgés (66-81 ans) émet l'hypothèse selon laquelle la diminution observée dans l'élimination chez les sujets âgés serait due à un déclin de la fonction rénale et plutôt qu'à une différence dans la conjugaison par les enzymes de phase II telles les UGT (Hermann et al. 2003).

### 2.3. Facteurs génétiques et génomiques

La génétique et la génomique jouent un grand rôle dans la variabilité de réponse aux médicaments, tel que le démontre l'importance qui est accordée à l'origine ethnique des patients autant la recherche que lors d'essais cliniques. La réaction de glucuronidation n'y fait pas

exception et plusieurs mécanismes, tels l'épissage alternatif et les polymorphismes, expliquent en partie la variabilité interindividuelle de cette voie.

### 2.3.1. Les polymorphismes

Les polymorphismes sont les variations génétiques les plus étudiées dans le cadre de la variabilité du métabolisme des xénobiotiques entre les individus. En ce qui concerne les UGT, de nombreux polymorphismes ont en effet un impact important sur la conjugaison des substrats de ces enzymes, tel que le démontrent plusieurs études *in vivo* et *in vitro* (Stingl et al. 2014). Ces variations incluent les insertions/délétions, les variations du nombre de copies d'un gène ainsi que les polymorphismes ponctuels. Une attention particulière sera apportée à décrire ces derniers puisqu'ils sont un élément important de l'étude présentée dans la discussion de ce mémoire.

Premièrement, les insertions et délétions sont des polymorphismes caractérisés par l'ajout ou le retrait d'un ou plusieurs nucléotides au sein de la séquence d'ADN. Ils peuvent ainsi entraîner la modification d'un amplificateur ou d'un promoteur, ou encore, lorsque situé dans la région codante, un décalage du cadre de lecture, un codon stop prématuré ou une modification de la séquence peptidique. Par exemple, une publication récente relate les résultats d'une étude de phase I ayant pour but d'évaluer la dose maximale d'irinotécan tolérée par des patients classés en groupes selon leur génotype. L'étude montre que les patients ne possédant pas le polymorphisme UGT1A1\*28, une insertion affectant la séquence TATA du promoteur du gène, tolèrent une dose plus de deux fois plus élevée de cet agent anticancéreux que celle que tolèrent les patients hétérozygotes pour cet allèle (Innocenti et al. 2014).

Deuxièmement, les variations du nombre de copies d'un gène consistent en un nombre de copies du gène différent d'une copie par allèle. Ainsi il peut s'agir de l'absence d'un gène sur un ou deux allèles d'un individu ou encore de la duplication du gène menant à la présence de plus d'une copie du gène par allèle. Notamment, les gènes UGT2B17 et UGT2B28 comptent parmi les gènes humains dont la délétion est la plus fréquente (McCarroll et al. 2006, Stranger et al. 2007). Des tissus de foie d'individus possédants 0 ou 2 copies du gène UGT2B17 ont été comparés en ce qui a trait à leur activité enzymatique envers le médicament exemestane, utilisé dans le traitement du cancer du sein. Les résultats montrent une diminution de 14 fois de la formation du métabolite glucuronidé pour les tissus d'individus sans copie d'UGT2B17 en

comparaison avec les tissus de ceux possédant 2 copies. Également, pour la même comparaison, le  $V_{max}/K_m$  était 36 fois moins élevé (Sun et al. 2010).

Finalement, les polymorphismes ponctuels, abrégés SNP pour *single nucleotide polymorphism*, sont des variations modifiant une seule paire de base du génome. Ils peuvent y modifier un nucléotide situé dans une région codante ou non. Même dans la région non-codante, un SNP a le potentiel de modifier l'expression d'un gène s'il se trouve au sein d'un amplificateur ou encore au promoteur du gène. Un de ces polymorphismes, situé au promoteur d'*UGT1A9*, diminue significativement chez les patients porteurs l'exposition au MPA lors d'un traitement au mycophénolate mofétil (Levesque et al. 2007, Levesque et al. 2008, Mazidi et al. 2013). Un autre polymorphisme, par sa position au promoteur d'*UGT2B7* (-161C>T), diminue la clairance de la lamotrigine dans la population thaïlandaise où il a été étudié (Singkham et al. 2013). Ce même SNP a récemment été associé à une meilleure survie sans récurrence chez des patientes atteintes de cancer du sein traitées avec l'épirubicine. L'étude montre en effet une meilleure survie des patientes hétérozygotes pour cette variation (Sawyer et al. 2016). Par contre, bien que ces résultats soient prometteurs l'article ne décrit pas par quel mécanisme le SNP modifie la pharmacocinétique du composé.

Les SNP qui modifient un nucléotide de la séquence codante d'un gène et entraînent un changement d'acide aminé dans la protéine se nomment SNP non-synonymes. Le changement d'acide aminé a le potentiel d'affecter l'activité de l'enzyme de plusieurs façons. Un exemple est le variant *UGT2B7\*4* qui, bien qu'ayant une fréquence allélique de 0,4% dans la population japonaise où il a été étudié, a une importance certaine puisqu'il réduit l'activité de conjugaison de l'enzyme envers l'acide désoxycholique et le tétrachlorocathécol à respectivement 2 et 11% de l'activité qu'a l'enzyme d'*UGT2B7\*1* pour ces substrats (Saeki et al. 2004, Radominska-Pandya et al. 2010). Son influence est due au fait qu'il modifie en une asparagine l'acide aspartique à la position 398 de l'enzyme, résidu clé du site de liaison de la protéine avec son co-substrat. D'autres mécanismes permettent à un SNP de modifier l'activité d'une enzyme, notamment la modification d'un site d'interaction protéique ou encore d'un site de modification post-traductionnelle.

D'autres variants, malgré une fréquence allélique élevée dans la population, peuvent avoir un effet modeste ou dépendant du substrat sur l'activité enzymatique. C'est le cas du polymorphisme UGT2B7\*2, le plus fréquent de ce gène, dont la fréquence avoisine les 50% dans la population caucasienne. La substitution pour une tyrosine de l'histidine en position 268 de l'enzyme entraînée par ce SNP ne permet pas d'expliquer la variabilité de glucuronidation de cette enzyme pour plusieurs médicaments tels la AZT ou le mycophénolate mofétil (Djebli et al. 2007, Levesque et al. 2007, Peterkin et al. 2007, Levesque et al. 2008, Kwara et al. 2009). Ainsi, la fréquence allélique d'un polymorphisme d'un gène UGT n'est pas systématiquement représentative de sa contribution à la variabilité de glucuronidation.

### 2.3.2. L'épissage alternatif

L'épissage alternatif consiste en l'inclusion ou l'exclusion d'exons lors de la transcription des ARN messagers. Il touche l'ensemble des gènes à multiples exons (Pan et al. 2008, Wang et al. 2008, Carstens 2014). Une expansion importante du transcriptome, soit l'entièreté des ARN messagers produits par un seul gène, découle des événements d'épissage alternatif. Cette hausse du nombre de transcrits via l'épissage alternatif est fréquemment associée avec une diversification du protéome, soit l'ensemble des protéines produites par ce gène (Nilsen and Graveley 2010, Shabalina et al. 2014). Ce phénomène permet l'expression différentielle de gènes dans diverses situations, lors du développement ou en présence de néoplasie par exemple (Eswaran et al. 2013, Biamonti et al. 2014, Zhang et al. 2014). Tel que mentionné plus tôt, ce sont des événements d'épissage alternatif qui permettent au locus *UGT1A* de produire 9 enzymes mais aussi 9 protéines enzymatiquement inactives (Gong et al. 2001, Levesque et al. 2007). Ces dernières influencent la glucuronidation puisqu'elles ont un effet répresseur sur l'activité des enzymes UGT1A (Girard et al. 2007). Notre groupe a également mis au jour plusieurs événements d'épissage alternatif des gènes *UGT2B*. D'abord au niveau d'*UGT2B7*, 6 nouveaux exons ont été identifiés menant à 22 variants d'épissage alternatif. Les protéines enzymatiquement inactives codées par ces variants ont été mises en lien avec une régulation négative des enzymes provenant des transcrits canoniques d'*UGT2B7*. Cette régulation s'effectue via des interactions protéine-protéine et une expression différentielle des protéines variantes est observée dans des tissus normaux, fœtaux et tumoraux (Menard et al. 2011, Menard et al. 2013, Menard et al. 2013). De plus, une étude de transcriptomique a permis d'identifier

plus de 132 nouveaux transcrits UGT issus d'événements d'épissage alternatif de tous les gènes UGT (Tourancheau et al. 2016). Telles que présentées dans l'annexe de ce mémoire, l'identification et la caractérisation de protéines issues de certains de ces transcrits permet d'entrevoir pour ces protéines des fonctions autres que la glucuronidation (Rouleau et al. 2016).

### 3. Les modifications post-traductionnelles

#### 3.1. Incidence et effet sur le protéome

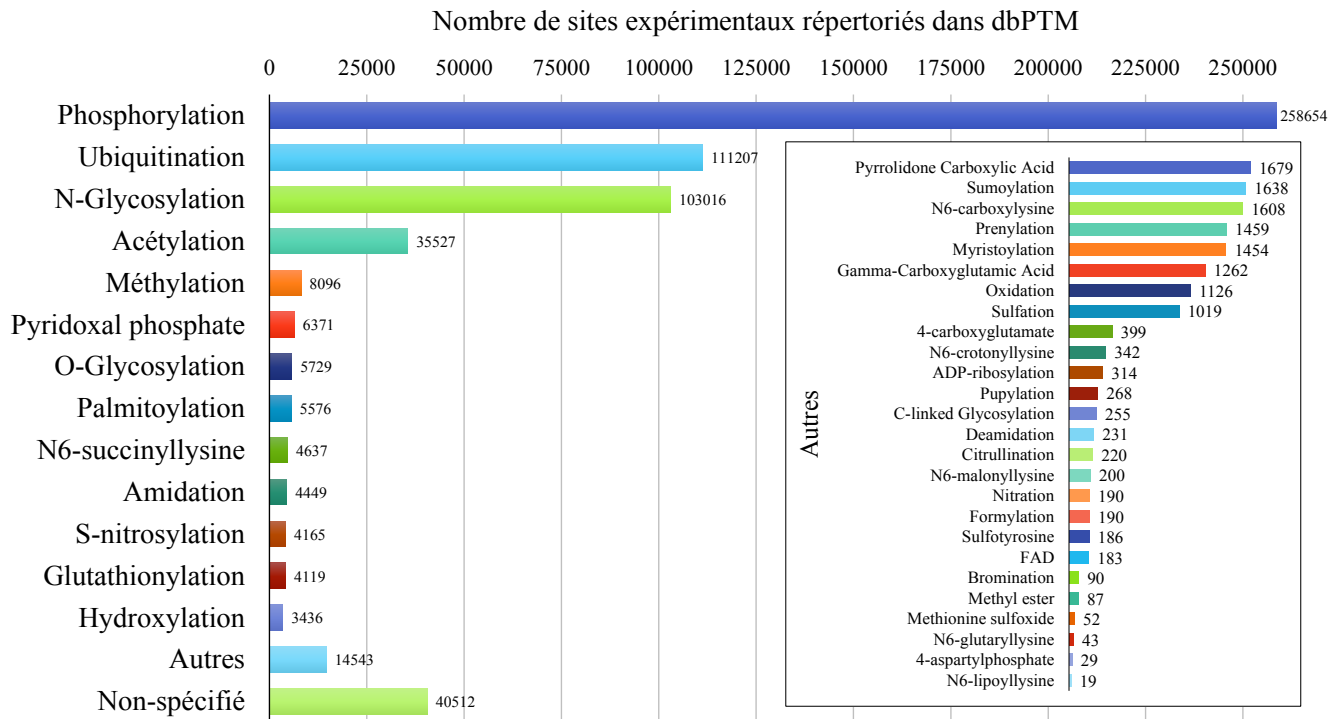
Les modifications post-traductionnelles consistent en l'altération de la structure primaire d'une protéine suite à sa traduction dans le but d'obtenir une protéine dite mature. Plusieurs de ces modifications se font via l'attachement de manière covalente d'un groupement à la protéine. Ledit groupement fonctionnel peut être de nature variée. Les clivages de sous-unités peptidiques, comme lors du processus de maturation de l'insuline (Steiner and Oyer 1967), ou la dégradation protéique font aussi partie des altérations possibles. Au niveau du protéome, ces modifications sont ce qui permet aux quelques 100 000 transcrits issus de presque 25 000 gènes humains de produire plus d'un million de protéines distinctes. Grâce aux modifications post-traductionnelles, ces protéines diffèrent sur le plan de leur repliement, leurs propriétés physiques et chimiques, leur conformation, leur localisation au sein de la cellule, leur stabilité et leur activité (Mann and Jensen 2003). La base de données dbSNP répertorie entre autres les modifications post-traductionnelles validées expérimentalement. Les données qu'elle contient peuvent provenir de d'autres bases de données plus spécifiques à un type de modification (p. ex., la phosphorylation) ou d'articles scientifiques analysés par une méthode de *text-mining* (Lu et al. 2013). Les statistiques provenant de cette base de données permettent de constater que parmi les sites de modifications post-traductionnelles expérimentalement validés 42% sont des sites de phosphorylation, 18% d'ubiquitination, 17% de N-glycosylation et 6% d'acétylation (Figure 5). Chacune de ces modifications sera brièvement présentée puis mise en contexte avec les enzymes UGT lorsqu'approprié, en portant une attention particulière à la N-glycosylation. Aucune des autres modifications post-traductionnelles présentées à la figure 5 ne sont connues comme affectant les enzymes UGT.



### 3.2. La N-glycosylation

Il existe plusieurs types de glycosylation qui consistent tous en l'attachement par une enzyme de type glycosyl-transférase d'un oligosaccharide à un acide aminé d'une protéine. Cette dernière se nomme alors glycoprotéine. On parle de N-glycosylation lorsque l'oligosaccharide est attaché à l'azote d'une asparagine. En contraste, on parle plutôt d'O-glycosylation lorsque cet oligosaccharide est attaché à une sérine ou une thréonine et de C-glycosylation lorsqu'il est attaché à l'indole d'un tryptophane. Ces deux derniers types de glycosylation sont toutefois moins courants que le premier, lequel correspond en effet à près de 95% des événements de glycosylation (Figure 5).

La N-glycosylation se produit spécifiquement sur les résidus asparagines se trouvant dans un motif Asn-Xaa-(Ser/Thr) où Xaa représente n'importe quel acide aminé à l'exception de la proline (Thomsen and Nielsen 2012). De surcroit, il semble que la position du motif au sein de la séquence protéique ait également une influence sur la probabilité qu'il soit N-glycosylé, les positions à proximité des extrémités N- et C-terminales seraient défavorisées. La probabilité de



**Figure 5. Occurrence des modifications post-traductionnelles expérimentalement validées dans la base de données dbPTM**

Données acquises sur le site web de dbPTM (<http://dbptm.mbc.nctu.edu.tw/statistics.php>) consulté le 5 mai 2016.

N-glycosylation selon la position dans la séquence suivrait une distribution uni modale dont la valeur maximale se situerait au tiers de la protéine à partir de l'extrémité N-terminale (Nilsson and von Heijne 2000, Gupta and Brunak 2002). La N-glycosylation est un processus essentiel, l'absence complète de N-glycosylation est létale alors qu'un fonctionnement altéré des enzymes impliquées via des mutations entraîne plusieurs problèmes chez l'humain entre autres un retard mental et des dysfonctions au niveau du foie et des intestins (Freeze and Westphal 2001). Également, les mutations supprimant le motif de N-glycosylation Asn-Xaa-(Ser/Thr) sont significativement plus abondantes dans une banque de mutations liées à des maladies que dans une banque de mutations neutres (Li et al. 2010).

Ce processus joue en effet divers rôles dans la maturation et le bon fonctionnement des protéines qu'il affecte. Pour certaines protéines, la présence d'oligosaccharides augmente leur stabilité en augmentant leur résistance à la protéolyse, comme dans le cas de la fibronectine (Bernard et al. 1982), ou à la dénaturation due à la chaleur, comme c'est le cas pour l'interféron-bêta et l'interleukine-5 (Kodama et al. 1993, Runkel et al. 1998). La solubilité de la protéine peut également être augmentée selon le contenu en acide sialique des N-glycans qui y sont ajoutés (Stowers et al. 2002). L'ajout de tels oligosaccharides est ainsi utilisée en glyco-ingénierie afin d'augmenter la demi-vie des composés dans le sérum (Sinclair and Elliott 2005). La N-glycosylation est aussi impliquée dans la sécrétion, la fonction de récepteurs, la spécificité d'enzymes, la modulation de l'activité enzymatique (Skropeta 2009) ainsi que dans le repliement et l'interaction protéique (Imperiali and O'Connor 1999).

La technique généralement utilisée pour confirmer la N-glycosylation est l'analyse d'immunobuvardages de type Western de l'enzyme native et mutante (dont les asparagines des sites de N-glycosylations à confirmer ont été mutées) que l'on soumet ou non à un traitement à l'endoglycosidase H (Endo H). Cette enzyme clive l'oligosaccharide à proximité de son point d'attache à l'asparagine du motif de N-glycosylation. Si le site est bien un site de N-glycosylation, on observe alors une différence due à la présence de l'oligosaccharide entre la hauteur à laquelle migrent les protéines mutantes et non mutantes sur le gel pour les échantillons non-traités alors que ceux traités à l'Endo H migreront à la même hauteur. Également, un traitement à la tunicamycine de cellules dans lesquelles on surexprime la protéine investiguée dans des cellules (dans le cas des UGT les HEK293 sont utilisées car elles n'expriment pas

constitutivement ces enzymes) peut être utilisé. Ce traitement empêche la N-glycosylation et est comparé au traitement avec l'Endo H dans le cadre d'essais enzymatiques. Dans le cas où le traitement à l'Endo H n'affecte pas l'activité enzymatique, mais que celle-ci est affectée par la tunicamycine cela signifie que la N-glycosylation de l'enzyme n'est essentielle qu'au stade du repliement et ne l'est plus par la suite. De plus, la N-glycosylation peut aussi être validée via une technique d'analyse du glycoprotéome combinant la digestion multiple des enzymes et l'extraction en phase solide par chimie hydrazide des peptides résultants (Chen et al. 2009). Enfin, des logiciels de prédiction tel NetGly1.0, qui analyse le contexte dans lequel se retrouvent les motifs de N-glycosylation au sein d'une séquence donnée, peuvent prédire le potentiel de N-glycosylation d'un motif (Gupta 2015).

Plusieurs enzymes UGT humaines sont sujettes à la N-glycosylation. La séquence de l'UGT1A6 contient deux motifs de N-glycosylation mais seule la position 345 est N-glycosylée (Ouzzine et al. 2006). Également, les UGT1A4 et UGT1A9 sont aussi N-glycosylées en diverses positions et l'on sait de surcroît grâce à l'utilisation de la tunicamycine que le repliement adéquat d'UGT1A9 dépend de cette glycosylation (Nakajima et al. 2010). Les membres de la famille UGT2B possèdent de nombreux motifs de N-glycosylation, dont plusieurs ont été validés ou infirmés expérimentalement. Ainsi, il a été démontré que malgré la présence d'un motif en position 315, l'UGT2B4 n'est pas N-glycosylée (Barbier et al. 2000). Cependant, une étude utilisant plutôt l'analyse du glycoprotéome de foies humains a identifié ce site comme étant glycosylé dans l'UGT2B4 (Chen et al. 2009). Le logiciel de prédiction NetGly1.0 n'identifie cependant pas le site d'UGT2B4 comme potentiellement glycosylé, ce qui peut s'expliquer par le fait que le logiciel possède une précision globale de 76% (Gupta 2015). D'autres sites de N-glycosylation au sein d'autres UGT2B voient leur statut différer selon la méthode utilisée pour la validation du site. Ainsi, Barbier et al. ont déterminé qu'UGT2B17 était glycosylée, sans toutefois préciser sur lequel ou lesquels de ses trois sites potentiels, alors qu'aucun glyco-peptide n'a été observé par Chen et al. pour cette enzyme (Barbier et al. 2000, Chen et al. 2009). Cependant, les deux études sont en accord quant au fait que seul le site en position 65 d'UGT2B15 serait N-glycosylée alors que les sites aux positions 316 et 483 ne le seraient pas. Ces résultats divergents selon la technique utilisée pour la validation s'expliquent possiblement par le fait que l'environnement cellulaire, et non la seule présence ou absence d'un motif de N-glycosylation, influence la glycosylation d'un site.

L'enzyme UGT2B7 possède elle aussi trois sites potentiels de N-glycosylation aux positions 67, 68 et 315. Une étude utilisant des endoglycosidases a non seulement permis de valider la N-glycosylation des sites aux positions 68 et 315 et d'infirmer celle de la position 67, mais aussi d'explorer l'influence distincte de la N-glycosylation sur deux substrats d'UGT2B7 (Nagaoka et al. 2012). En effet, des essais enzymatiques ont montré que suite à une déglycosylation on observait une diminution de l'affinité pour l'AZT et le co-substrat UDPGA (indépendamment du substrat utilisé) ainsi qu'une augmentation de l'affinité pour la morphine, le tout en n'affectant pas la vitesse de formation du glucuronide ou d'utilisation du co-substrat.

### 3.3. La phosphorylation

La phosphorylation consiste en l'ajout d'un groupement phosphate sur une protéine, elle est la modification post-traductionnelle la plus courante, touchant près du tiers des protéines cellulaires (Cohen 2000). Le groupement phosphate est attaché à une sérine, une thréonine ou, plus rarement, une tyrosine. Le contexte de ces acides aminés au sein de la séquence protéique joue également un rôle, les enzymes responsables de la phosphorylation, les kinases, reconnaissent en effet divers motifs de phosphorylation (Wong et al. 2007). Son caractère réversible fait de cette modification une réaction abondamment utilisée dans la signalisation cellulaire, la phosphorylation pouvant activer ou inactiver une protéine. Quelques exemples probants sont les protéines de la voie de signalisation Ras-Raf-MEK-MAPK avec la kinase Raf qui active MEK via sa phosphorylation qui à son tour active MAPK de la même façon. Cette dernière ira par la suite phosphoryler de nombreuses autres protéines, contrôlant ainsi plusieurs facteurs de transcription. Des mutations de cette voie peuvent empêcher le changement de statut de phosphorylation de ces protéines, laissant la voie activée, ce qui se produit dans environ 30% des cancers humains (Mandal et al. 2016).

Plusieurs séquences UGT contiennent des sites de phosphorylations potentiels, mais seuls certains d'entre eux ont été validés expérimentalement comme étant phosphorylés, entraînant des conséquences diverses selon le cas. Ainsi, les UGT1A7 et UGT1A10 sont phosphorylées par la sérine-thréonine kinase PKC. La mutation des thréonines d'UGT1A7 aux positions 73 et 202 cause une perte complète de l'activité alors que la mutation de sa sérine en position 432 modifie son pH optimal ainsi que sa spécificité de substrats. Quant à elle, UGT1A10, lorsque mutée à

cette position, ne subit qu'un faible changement de pH optimal et aucun effet sur sa spécificité (Basu et al. 2005, Basu et al. 2008). Dans le cas d'UGT2B7, la mutation de sites reconnus par la kinase PKC n'entraîne aucun effet sur l'activité alors que la mutation de sites reconnus par la tyrosine-kinase Src, aux positions 236 et 438, abolit pratiquement son activité envers les catéchols estrogènes (Mitra et al. 2009). Également, la nature des tyrosines-kinases impliquées dans cette phosphorylation influence la spécificité du substrat (Mitra et al. 2011). Enfin, l'UGT2B15 est régulée via sa phosphorylation à la fois par la kinase PKC et par la Src. Cette régulation complexe permet entre autres aux cellules prostatiques d'effectuer un contrôle fin des niveaux de dihydrotestostérone (Chakraborty et al. 2012).

### 3.4. L'ubiquitination

L'ubiquitination est l'ajout d'un groupement ubiquitine à une protéine, sur les résidus lysines. Ce groupement est lui-même un polypeptide, long de 76 acides aminés, environ 8,5kDa, qui peut également suite à sa liaison à la protéine s'assembler à d'autres ubiquitines afin de former un polymère (Komander 2009). Cette modification, telle la phosphorylation, est réversible et joue un grand rôle dans la signalisation cellulaire, mais également dans la dégradation via le protéasome, influençant donc la demi-vie des protéines. Les UGT ne sont pas connues comme étant la cible de cette modification post-traductionnelle.

### 3.5. L'acétylation

L'acétylation, l'ajout d'un groupement acétyle sur une lysine, est bien connue comme affectant les histones, qui sont des protéines impliquées dans le repliement de la chromatine. Cette modification a ainsi un impact sur l'expression génique et par le fait même sur le protéome. Par exemple, le promoteur d'*UGT1A1* se retrouve à proximité d'une histone H3 hypoacétylée au rein mais hyperacétylée au foie ce qui influence l'attachement d'un facteur de transcription et donc l'expression différentielle du gène entre ces tissus (Oda et al. 2013). L'acétylation de protéines autres que les histones est également possible, jouant tel le fait l'ubiquitination un rôle dans la stabilité et la dégradation protéique (Sadoul et al. 2008). Aucune littérature n'indique l'acétylation d'enzymes UGT.

## Hypothèses et objectifs

L'UGT2B7 est impliquée dans le métabolisme de nombreux médicaments. Ainsi, la grande variabilité interindividuelle observée dans son activité de glucuronidation constitue une problématique dans une optique de mise au point et d'optimisation de traitements pharmaceutiques. À l'aube de l'ère de la médecine personnalisée, il devient en effet crucial de comprendre les divers acteurs influençant cette variabilité. Celle-ci n'est en outre pas entièrement expliquée par les facteurs environnementaux, physiologiques ou génomiques connus. De l'étude de ces derniers facteurs, on retient le lien parfois faible entre les polymorphismes ponctuels fréquents et l'étendue de la variation observée dans la glucuronidation. Ainsi, l'hypothèse de ce travail était que la présence de variations ponctuelles peu fréquentes pourrait expliquer une partie de la variabilité de la voie de glucuronidation. Ces variations pourraient entraîner notamment des changements au niveau post-traductionnel.

Le premier objectif du projet a donc été la recherche de nouvelles variations ponctuelles du gène *UGT2B7* par le biais de séquençage. Nous avons d'abord étudié une cohorte de 305 volontaires sains précédemment recrutés dans le cadre d'une étude portant sur la pharmacocinétique du mycophénolate mofétil (Levesque et al. 2007, Levesque et al. 2008). Ensuite, la région exonique contenant le nucléotide affecté par le génotype *UGT2B7\*2* a été séquencée pour l'ensemble de la cohorte afin de sélectionner les sujets pour l'étude pharmacocinétique. Le séquençage de cette région a révélé que deux individus étaient hétérozygotes au nucléotide en position 361 du gène. Comme chez un individu hétérozygote une seule des deux allèles du gène possède la mutation, la nouvelle variation possède une fréquence allélique de 0.328% dans cette population (2 allèles mutantes sur un total de 610 allèles). Sa position au sein de la séquence du gène *UGT2B7* fait en sorte que le codon 121 est traduit en une asparagine plutôt qu'en l'acide aspartique retrouvé dans la protéine *UGT2B7* de référence.

Puisque l'analyse de cette nouvelle variation à l'aide du logiciel de prédiction de N-glycosylation NetNGlyc 1.0 (Gupta 2015) conclut que cette position possède une haute probabilité d'être glycosylée, le second objectif du projet était de valider expérimentalement cette prédiction. Procédant tel que l'ont fait Barbier et Nagaoka afin de valider les sites de N-glycosylation d'UGT2B (Barbier et al. 2000, Nagaoka et al. 2012), la surexpression des protéines mutantes et

non-mutantes dans des cellules de la lignée HEK293 suivie d'un traitement à l'Endo H, analysé par immunobuvardage de type Western, a permis la validation du nouveau site de N-glycosylation.

Suite à cette validation, le troisième objectif était d'évaluer l'impact de cette glycosylation additionnelle sur l'activité enzymatique d'UGT2B7. Cet objectif est atteint en procédant à des essais enzymatiques afin de déterminer les constantes cinétiques (vélocité, affinité et clairance) associées à la glucuronidation de l'AZT ou de l'acide mycophénolique par des microsomes de cellules HEK293 surexprimant l'enzyme mutante ou non-mutante.

Enfin, au vu de l'impact important de la nouvelle variation identifiée sur l'activité enzymatique d'UGT2B7 via la création d'un site de N-glycosylation, un quatrième objectif du projet était d'inventorier les variants ponctuels provenant de bases de données publiques pouvant affecter de façon similaire la N-glycosylation d'UGT de la famille 2B. Cet objectif est réalisé en acquérant d'abord l'information associée aux SNP non-synonymes des gènes *UGT2B* répertoriés dans la base de données dbSNP (Sherry et al. 2001) puis en couplant cette information (c.-à-d. la position de l'acide aminé affecté au sein de la protéine, sa nature et sa fréquence) aux séquences nucléotidiques et peptidiques des *UGT2B*. L'inspection visuelle des acides aminés avoisinants une mutation permet de reconnaître si un site de N-glycosylation est créé ou aboli par celle-ci.

## Résultats : Un variant rare d'UGT2B7 crée un nouveau site de N-glycosylation et affecte l'activité de l'enzyme

Camille Girard-Bock, Marie-Odile Benoit-Biancamano, Lyne Villeneuve, Sylvie Desjardins, Chantal Guillemette

La superfamille des UDP-glucuronosyltransférases (UGT) est constituée de glycoprotéines résidant au réticulum endoplasmique et sujettes aux modifications post-traductionnelles (PTM, post-translational modifications). L'enzyme UGT2B7 est d'un intérêt particulier vu son action sur une grande variété de médicaments. La plupart des études actuelles n'ont pour sujet que les variants communs de cette enzyme et n'examinent donc qu'une fraction de la diversité génétique de celle-ci. En effet, les variants rares (fréquence allélique en deçà de 1%) peuvent potentiellement avoir un effet considérable puisqu'ils sont prédits comme étant bien plus nombreux que les variants communs au sein du génome humain. La présente étude fait état de la découverte d'un variant rare d'UGT2B7 possédant un intérêt pharmacogénétique potentiel et encodant une substitution d'acide aminé au codon 121. Cette variation peu fréquente, retrouvée chez deux individus au sein d'une population de 305 sujets sains, mène à la traduction d'une asparagine (Asn) plutôt qu'un acide aspartique (Asp) au codon 121 (UGT2B7 p.D121N). Cette substitution est prédite comme créant un motif de N-glycosylation NX(S/T) subséquent validé par traitement à l'endoglycosidase de fractions microsomales issues de surexpressions dans des HEK293 et par inhibition à la tunicamycine de la N-glycosylation d'UGT2B7 produites de façon endogène dans des HEK293. De plus, la présence d'un oligosaccharide additionnel sur l'enzyme UGT2B7, affectant potentiellement son repliement, résulte en la diminution, respectivement par 49 et 40%, de la formation de glucuronides à partir de la zidovudine et de l'acide mycophénolique. Une analyse de la base de données dbSNP a permis la découverte de 32 variants rares pouvant potentiellement créer ou abolir des motifs de N-glycosylation au sein d'enzymes UGT2B. Ensemble, ces variants ont le potentiel d'augmenter la proportion de la variance de la voie des UGT qui est expliquée, par le biais de modifications post-traductionnelles telles la N-glycosylation qui affectent le métabolisme des médicaments.

**Article publié dans le journal *Drug Metabolism and Disposition***



## **Title Page**

### **SHORT COMMUNICATION**

#### **A rare UGT2B7 variant creates a novel N-glycosylation site at codon 121 with impaired enzyme activity**

**Authors:** Camille Girard-Bock, Marie-Odile Benoit-Biancamano, Lyne Villeneuve, Sylvie Desjardins, Chantal Guillemette

Pharmacogenomics Laboratory, Canada Research Chair in Pharmacogenomics, Centre Hospitalier Universitaire de Québec Research Center and Faculty of Pharmacy, Laval University, Québec City, Québec, Canada (CGB, LV, SD, CG)

Département de Pathologie et de Microbiologie, Faculté de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, QC J2S 2M2, Canada (MOBB)

#### **Non-standard abbreviations:**

AcMPAG, mycophenolic acid acyl-glucuronide; AZT, zidovudine (3'-azido-3'-deoxythymidine); Endo H: Endoglycosidase H; HEK293, human embryo kidney 293 cells; MPA, mycophenolic acid; MPAG, mycophenolic acid phenolic-glucuronide; UDPGA, UDP-glucuronic acid; UGT, UDP-glucuronosyltransferase.

## Abstract

UDP-glucuronosyltransferase (UGT) superfamily are glycoproteins resident of the endoplasmic reticulum membranes that undergo post-translational modifications (PTM). UGT2B7 is of particular interest because of its action on a wide variety of drugs. Most studies currently survey common variants and are only examining a small fraction of the genetic diversity. However, rare variants (frequency <1%) might have significant effect as they are predicted to greatly outnumber common variants in the human genome. Here, we discovered a rare single nucleotide *UGT2B7* variant of potential pharmacogenetic relevance that encodes a nonconservative amino acid substitution at codon 121. This low-frequency variation, found in two individuals of a population of 305 healthy volunteers, leads to the translation of an asparagine (Asn) instead of an aspartic acid (Asp) (*UGT2B7* p.D<sup>121</sup>N). This amino acid change was predicted to create a putative N-glycosylation motif NX(S/T) subsequently validated upon endoglycosidase H treatment of microsomal fractions and inhibition of N-glycosylation of endogenously produced UGT2B7 with tunicamycin from HEK293 cells. The presence of an additional N-linked glycan on the UGT2B7 enzyme, likely affecting proper protein folding, resulted in a significant decrease, respectively by 49 and 40%, in the formation of zidovudine and mycophenolic acid glucuronides. A systematic survey of the dbSNP database uncovered 32 rare and naturally occurring missense variations predicted to create or disrupt N-glycosylation sequence motifs in the other UGT2B enzymes. Collectively, these variants have the potential to increase the proportion of variance explained in the UGT pathway due to changes in PTM such as N-linked glycosylation with consequences on drug metabolism.

## INTRODUCTION

Metabolic enzymes of the UDP-glucuronosyltransferase (UGT) superfamily catalyze glucuronidation reactions involved in the disposition of endogenous molecules, drugs and other xenobiotics. UGTs are glycoproteins resident of the endoplasmic reticulum (ER) membranes that undergo post-translational modifications (PTM) such as glycosylation and phosphorylation (Mackenzie 1990, Chakraborty et al. 2012, Riches and Collier 2015). This family of proteins are involved in drug metabolism, accounting for approximately 55% of the 200 most prescribed drugs (Guillemette et al. 2014). UGT2B7 is of particular interest as it is the most prevalent member of this enzyme family, conjugates almost one fifth of all drugs known to be conjugated to glucuronic acid (Stingl et al. 2014). For instance, UGT2B7 is involved in the inactivation and elimination of carboxylic acid-containing drugs including the nonsteroidal anti-inflammatory drugs (NSAIDs), and of a large variety of other drugs including opioids, the anticonvulsant valproic acid, morphine, codeine, efavirenz, fenofibric acid, mycophenolic acid (MPA) and zidovudine (AZT) (Stingl et al. 2014).

Glucuronide formation varies greatly between individuals and may be explained in part by the presence of single nucleotide polymorphisms (SNPs) (Guillemette et al. 2014). In the case of UGT2B7, there is a well-established connection to variation in drug efficacy and toxicity, especially for the most common polymorphism *UGT2B7*\*2 allele that encodes a nonconservative amino acid substitution His<sup>268</sup>Tyr in the substrate-binding domain linked to altered pharmacokinetics of several drugs but not all substrates examined (Bhasker et al. 2000, Stingl et al. 2014). Most pharmacogenomics studies currently survey common variants and thus are only observing a small fraction of the genetic diversity in any gene. Different forms of genetic variations within a *UGT* locus, including common and rare coding and regulatory variants, can exist and have separate and yet cumulative effects. Rare variants (allele frequency <1%) may also cause interindividual differences in therapeutic effects and adverse reactions to drugs but have been much less studied (Nelson et al. 2012). This is highly relevant since rare variants are predicted to greatly outnumber common variants in the human genome (Marth et al. 2011) and they may be very important in the genomic contribution to treatment response and toxicity (Ramsey et al. 2012, Gillis et al. 2014). For example, it was recently established that as much as 92% of known variants in the cytochrome P450 superfamily have an allelic frequency below 1%,

and 83% below 0.1% (Fujikura et al. 2015). Similar distribution of variants by allelic frequency might be observed in UGT2B7 and other UGTs, underscoring the importance of studying rare variants.

Here we report the identification of a novel rare missense *UGT2B7* variant (NM\_001074.2:c.361G>A) in a Caucasian cohort of 305 healthy volunteers. This rare variation leads to the translation of an asparagine (Asn) instead of an aspartic acid (Asp) at position 121 of the UGT2B7 protein (NP\_001065.2:p.Asp121Asn; D<sup>121</sup>N). Investigation of its functional impact using bioinformatics tools predicted that this variation creates a putative N-glycosylation site, which was experimentally validated. We then studied whether this additional PTM site affects glucuronidation of AZT and MPA, demonstrating a drastic reduction in UGT2B7 catalytic activity. Lastly, we established a comprehensive dataset of 32 rare variants potentially affecting gain or loss of N-glycosylation sites in other family members, with some located in regions highly conserved across *UGT* genes, by integrating data from the dbSNP database along with their respective allelic frequency. Collectively, these low-frequency variants have the potential to increase the proportion of variance explained in the glucuronidation pathway.

## MATERIALS AND METHODS

**Genotyping of healthy Caucasian volunteers.** The variation at position 361 in the *UGT2B7* gene causing amino acid change Asp<sup>121</sup>Asn was initially observed in one of 52 participants of a previous pharmacokinetic study (Levesque et al. 2007, Levesque et al. 2008). These healthy volunteers were selected in a population of 305 Caucasian subjects that were subsequently genotyped for the variation at codon 121 using specific primers, as described (Levesque et al. 2007, Levesque et al. 2008), and a second subject was found to carry this variation.

**Heterologous expression of the variant UGT2B7 enzyme and enzymatic assays.** A HEK293 cell line expressing the variant UGT2B7 protein at codon 121 (UGT2B7Asn<sup>121</sup>) was established by mutagenesis using primers 5'-CCCAACAACATCCTCTCTTAAATGAAA-3' (forward) and 5'-TTTCAATTTAAGAGAGGATGAGTTGTTGGG-3' (reverse) based on the reference UGT2B7 cDNA (Menard et al. 2011). Relative quantification of UGT protein content in microsomal proteins was performed by immunoblot analysis using an anti-UGT2B antibody

EL-93 (dilution 1:2000) (Lepine et al. 2004) and with an anti-calnexin antibody (dilution 1:5000; Stressgen Biotechnologies, Victoria, Canada) to normalize for sample loading. Kinetic parameters were assessed for both cell lines in the presence of increasing concentrations of substrates, MPA (25 to 1500  $\mu$ M; MP Biomedicals, Santa Ana, CA) or AZT (100 to 5000  $\mu$ M; Sigma-Aldrich, St. Louis, MO) and fixed UDPGA (5 mM), or co-substrate UDPGA (from Sigma-Aldrich, St. Louis, MO) (50 to 5000  $\mu$ M; fixed AZT (500  $\mu$ M or 1 mM)), for 1-hour incubation at 37°C. Enzymatic assays and liquid chromatography-mass spectrometry were performed to assess glucuronide (G) formation as described (Benoit-Biancamano et al. 2007). Absolute velocities ( $V_{\max}$ : pmol/min/mg protein) were normalized for UGT protein content assessed by western blotting and expressed as relative  $V_{\max}$  (pmol/min/mg protein/UGT content). Kinetic parameters, according to the Michaelis-Menten model, were calculated with Sigma Plot 11 using the Enzyme Kinetics 1.3 module (SYSTAT Software Inc., San Jose, CA). Data are derived from at least two independent experiments performed in triplicates. *P*-value calculations using Student's t-test were performed with Excel 2016 (Microsoft Corporation, Redmond, WA).

**Experimental confirmation of N-glycosylation using endoglycosidase digestion.** Enzymes were obtained from New England BioLabs Inc. (Ipswich, MA). Microsomes (20  $\mu$ g) from HEK293-UGT2B7 cell lines were incubated in Glycoprotein Denaturing Buffer (0.5% SDS, 40 mM DTT) for 10 minutes at 100°C to ensure denaturation of the protein content. For endoglycosidase H (Endo H) treatment, GlycoBuffer 3 (50 mM Sodium Citrate (pH 6 @ 25°C) and Endo H (500 U) were added to microsomes whereas for O-glycosidase treatment, microsomes were supplemented with GlycoBuffer 2 (50 mM Sodium Phosphate (pH 7.5 @ 25°C), NP-40 (1 %), neuraminidase (100 U) and O-glycosidase (40 000 U), in a 20  $\mu$ l final volume. Samples were incubated for 1h at 37°C and subsequently resolved by 10% SDS-polyacrylamide gel electrophoresis using standard procedure. Immunodetection of UGT2B7 was conducted with a polyclonal UGT2B7 antibody (1:5000; 16661-1-AP, ProteinTech Group, Rosemont, IL).

Inhibition of endogenous glycosylation of UGT2B7 was conducted as described with some modifications (Nakajuma et al. 2010). Briefly, HEK293 cells plated in 10 cm-dishes were transiently transfected with 4  $\mu$ g of pcDNA3 constructs driving expression of reference UGT2B7 or the c121 variant (Menard et al. 2011) using Lipofectamine 2000 (Invitrogen). Inhibition of

glycosylation was achieved with tunicamycin (Sigma-Aldrich) added at time of transfection (final concentration 0.1 and 1  $\mu\text{g/ml}$ ). Controls consisted of vehicle only (0.1% ethanol). Cells were harvested 24h/40h post-transfection by washing twice in PBS then homogenates were prepared by harvesting cells in phosphate-buffered saline containing 0.5 mM DTT. To control for ER stress, cells were transfected as above, thapsigargin (0.5  $\mu\text{M}$ ; Sigma-Aldrich) was added at time of transfection and cells were harvested 24 h post-treatment as above. Protein concentration in homogenates was determined by a bicinchoninic acid assay. Immunodetection of UGT2B7 was as described above.

**Analysis of UGT2B genetic variations.** Single nucleotide variations (SNPs) for UGT2B coding sequences were retrieved from the dbSNP database using the NCBI browser (U.S. National Library of Medicine, Bethesda, MD). Alignments of UGT2Bs were obtained using the Clustal O (1.2.1) multiple sequence alignment tool. Coding variations were analyzed to establish whether they affect the sequence context Asn-X-Ser/Thr (create or disrupt NX(S/T) motif), where X is any amino acid except proline.

## RESULTS AND DISCUSSION

Our analysis of *UGT2B7* gene sequences in a population of 305 healthy volunteers uncovered a novel missense variation observed in two individuals and corresponding to a nonsynonymous coding variation at codon 121 (Asp<sup>121</sup>Asn). A Sanger sequencing chromatogram is presented in **Fig.1A** designating a double peak of an adenine (A) and a guanine (G) at position 361 (allelic frequency of 0.328% for the variant A allele). This population was previously used to identify candidates carrying specific *UGT1A* and *UGT2B7* genetic variations for a pharmacokinetic study of mycophenolate mofetyl (Levesque et al. 2007, Levesque et al. 2008). An *in silico* analysis for prediction of PTM sites using the NetNGlyc 1.0 tool (<http://www.cbs.dtu.dk/services/NetNGlyc/>) also confirmed with the GlycoEP software (Chauhan et al. 2013), revealed that this amino acid change creates a putative asparagine-linked N-glycosylation motif NX(S/T) consisting of an asparagine followed 2 positions downstream by a threonine as pictured in **Fig.1D**. Three such motifs are typically found in the UGT2B7 protein, of which glycosylation at positions 68 and 315 was experimentally validated (Nagaoka et al. 2012). The western blot analysis of HEK293 microsomal preparations showed that the variant UGT2B7Asn<sup>121</sup> protein has a higher molecular weight than the UGT2B7Asp<sup>121</sup> enzyme, which is

consistent with the addition of an oligosaccharide by N-glycosylation. In contrast to O-glycosidase treatment, when treated with Endo H, oligosaccharides are removed leaving both UGT2B7 proteins at the same lower molecular weight, thus confirming the presence of an additional N-glycosylation on the variant UGT2B7Asn<sup>121</sup> protein (**Fig.1B**). As further evidence of N-glycosylation causing the enhanced mobility shift of the UGT2B7Asn<sup>121</sup> protein, N-glycosylation of endogenously produced UGT2B7 was inhibited with tunicamycin (Nakajima et al. 2010). Both the reference and variant UGT2B7 were detected as multiple protein bands with low drug concentration, suggesting a partially perturbed glycosylation, whereas the high drug concentration prevented the formation of slower migrating proteins, thus confirming that the mobility shift is caused by N-glycosylation (**Fig.1C**). Inhibition was not due to a general ER stress impairing glycosylation, given that thapsigargin did not perturb the mobility of either UGT2B7 (**Fig.1C**). We also used mass spectrometry analysis with the goal to detect the glycosylated peptide containing the sequence at codon 121. This approach, using cellular fractions from HEK293-UGT2B7 cell models enriched for UGT2B7 by affinity purification and treated or not with endoglycosidase PNGase F, was inconclusive. It permitted the detection of multiple UGT2B7 tryptic peptides (protein coverage up to 29%; data not shown) but not the codon 121-bearing peptide.

Previous data support that N-glycosylation plays a significant role in the enzymatic activity of UGT2B7 (Barbier et al. 2000, Nagaoka et al. 2012), suggesting that the novel UGT2B7Asn<sup>121</sup> variant may affect enzyme activity. A significant alteration in the conjugation of AZT and MPA was observed, with a decreased activity by 49% and 40%, respectively, associated with the variant Asn<sup>121</sup> protein compared to the reference Asp<sup>121</sup> enzyme, suggesting an altered protein folding. No significant differences were noted in the affinity ( $K_m$ ) of the enzyme (**Table 1, Fig.1E**). Nagaoka and colleagues showed that disruption of N-glycosylation sites at position 68 and 315 by mutagenesis leads to significant changes in the activity of the UGT2B7 enzyme (Nagaoka et al. 2012). Similarly, abolition of the N-glycosylation of UGT2B15 resulted in decreased enzyme activity without changing  $K_m$  (Barbier et al. 2000). In line, a rare *UGT1A1* variant creating a glycosylation site (K<sup>402</sup>T) was reported to cause a drastic decrease in UGT activity associated with severe hyperbilirubinemia (Crigler-Najjar type 1) (Sneitz et al. 2010). These observations underscore the need for further research on the impact of the N-glyco variants on UGT2B7 protein function as well as on other UGTs.

Our goal was next to build a framework for better understanding the effects of non-synonymous variations on the N-glycosylation of UGT2B enzymes. Through a systematic survey of publicly available data (dbSNP database), we revealed numerous naturally occurring missense variations, most with low allelic frequency (below 0.01%), predicted to affect N-glycosylation of UGT2B enzymes (**Table 2, Fig.1F**). Very little information is still available regarding the glycosylation profile of UGT enzymes, precluding us from establishing the potential functional relevance of these rare variants. Clustering of the variants according to their position in the UGT sequences helped visualize the fact that many of them localize in the same regions of UGT1, UGT2A or UGT2B enzymes. These potentially correspond to conserved and functional NX(S/T) motifs across several UGT proteins (**Fig.1F**), hence making them interesting candidates for further *in vitro* validation. For example, the first cluster is noteworthy because N-glycosylation in this region of UGT2B7 and UGT2B15, on residues 68 and 65 respectively, was experimentally validated in liver tissue by a method combining multiple digestion and hydrazide chemistry (Chen et al. 2009). Likewise, codon 69 of UGT2B4, UGT2B11 and UGT2B28 is affected by naturally occurring rare variants creating at this position the final serine or threonine of an NX(S/T) glycosylation motif, whereas a similar motif is present in this region for UGT1A and UGT2A enzymes. A comprehensive survey of the impact of these additional rare variants on the N-glycosylation of UGT enzymes will be required to evaluate if variations that lead to changes in the glycosylation pattern of a UGT protein can be damaging.

In conclusion, we discovered a rare UGT2B7 variant resulting in a NX(S/T) glycosylation gain that significantly affects rates of drug glucuronidation. A more thorough understanding of the significance of this variant in the context of drug treatment would require genotyping a larger population, initially to acquire a better appreciation of the allelic frequency across populations and to perform pharmacokinetic analyses in individuals carrying this variation. Likewise, our exhaustive analysis of variations data in additional drug conjugating UGT enzymes exposes numerous missense variations with low allelic frequency potentially creating or disrupting N-glycosylation sites, with some in regions conserved across all UGT enzymes. Accordingly, a large proportion of variability in the UGT pathway may be due to rare variants of significant effect size with a profound impact on their biological function due to changes in PTM such as N-linked glycosylation.



**Acknowledgments:** We thank Patrick Caron, Véronique Turcotte for technical assistance and Michèle Rouleau for helpful discussion.

**Authorship Contributions**

*Participated in research design:* Benoit-Biancamano, Guillemette

*Conducted experiments:* Girard-Bock, Benoit-Biancamano, Villeneuve, Desjardins

*Performed data analysis:* Girard-Bock, Benoit-Biancamano, Villeneuve, Desjardins, Guillemette

*Wrote or contributed to the writing of the manuscript:* Girard-Bock, Benoit-Biancamano, Guillemette

## REFERENCES

- Barbier O, Girard C, Breton R, Belanger A and Hum DW (2000). N-glycosylation and residue 96 are involved in the functional properties of UDP-glucuronosyltransferase enzymes. *Biochemistry* **39**: 11540-11552.
- Benoit-Biancamano MO, Caron P, Levesque E, Delage R, Couture F and Guillemette C (2007). Sensitive high-performance liquid chromatography-tandem mass spectrometry method for quantitative analysis of mycophenolic acid and its glucuronide metabolites in human plasma and urine. *J Chromatogr B Analyt Technol Biomed Life Sci* **858**: 159-167.
- Bhasker CR, McKinnon W, Stone A, Lo AC, Kubota T, Ishizaki T and Miners JO (2000). Genetic polymorphism of UDP-glucuronosyltransferase 2B7 (UGT2B7) at amino acid 268: ethnic diversity of alleles and potential clinical significance. *Pharmacogenetics* **10**: 679-685.
- Chakraborty SK, Basu NK, Jana S, Basu M, Raychoudhuri A and Owens IS (2012). Protein kinase Calpha and Src kinase support human prostate-distributed dihydrotestosterone-metabolizing UDP-glucuronosyltransferase 2B15 activity. *J Biol Chem* **287**: 24387-24396.
- Chauhan JS, Rao A and Raghava GP (2013). In silico platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PLoS One* **8**: e67008.
- Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L and Zou H (2009). Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res* **8**: 651-661.
- Fujikura K, Ingelman-Sundberg M and Lauschke VM (2015). Genetic variation in the human cytochrome P450 supergene family. *Pharmacogenet Genomics* **25**: 584-594.
- Gillis NK, Patel JN and Innocenti F (2014). Clinical implementation of germ line cancer pharmacogenetic variants during the next-generation sequencing era. *Clin Pharmacol Ther* **95**: 269-280.
- Guillemette C, Levesque E and Rouleau M (2014). Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* **96**: 324-339.
- Lepine J, Bernard O, Plante M, Tetu B, Pelletier G, Labrie F, Belanger A and Guillemette C (2004). Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* **89**: 5222-5232.

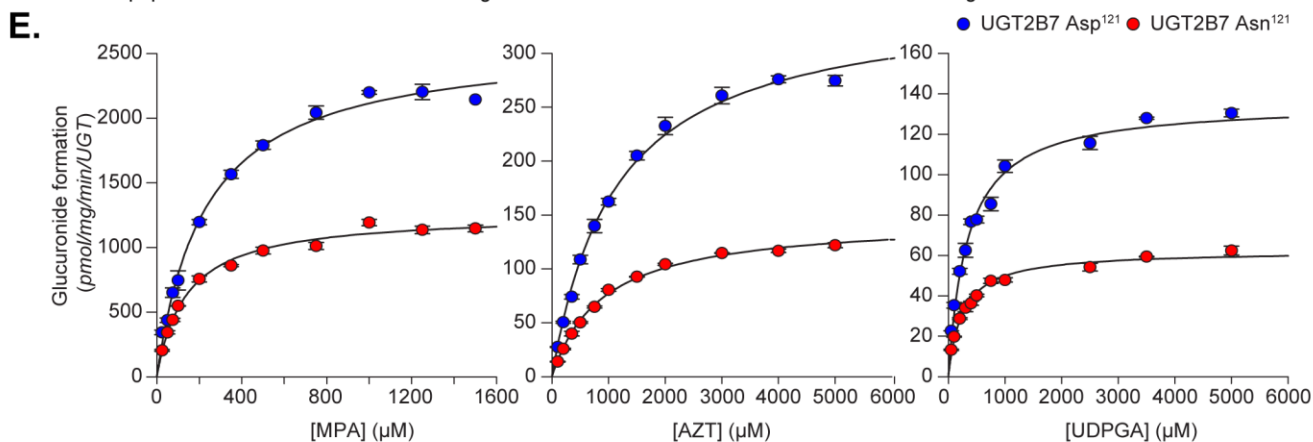
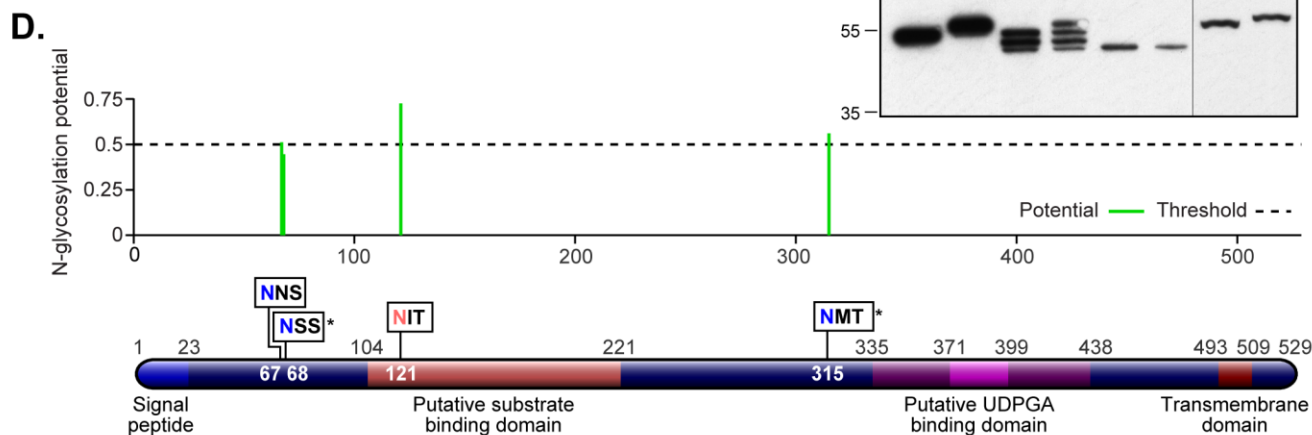
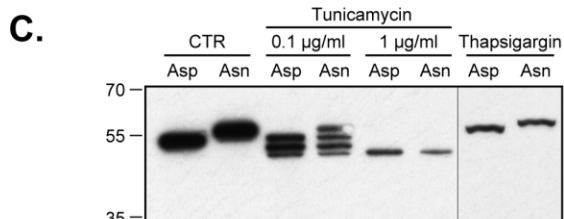
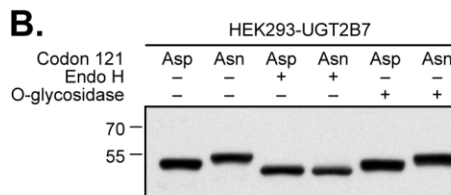
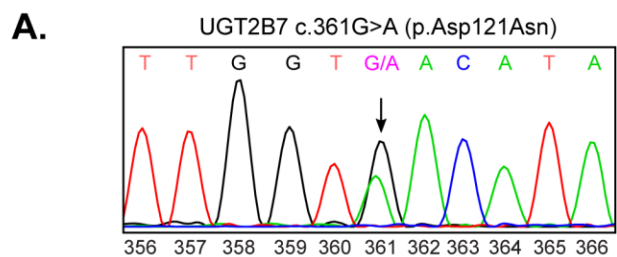
- Levesque E, Benoit-Biancamano MO, Delage R, Couture F and Guillemette C (2008). Pharmacokinetics of mycophenolate mofetil and its glucuronide metabolites in healthy volunteers. *Pharmacogenomics* **9**: 869-879.
- Levesque E, Delage R, Benoit-Biancamano MO, Caron P, Bernard O, Couture F and Guillemette C (2007). The impact of UGT1A8, UGT1A9, and UGT2B7 genetic polymorphisms on the pharmacokinetic profile of mycophenolic acid after a single oral dose in healthy volunteers. *Clin Pharmacol Ther* **81**: 392-400.
- Mackenzie PI (1990). The effect of N-linked glycosylation on the substrate preferences of UDP glucuronosyltransferases. *Biochem Biophys Res Commun* **166**: 1293-1299.
- Marth GT, Yu F, Indap AR, Garimella K, Gravel S, Leong WF, Tyler-Smith C, Bainbridge M, Blackwell T, Zheng-Bradley X, Chen Y, Challis D, Clarke L, Ball EV, Cibulskis K, Cooper DN, Fulton B, Hartl C, Koboldt D, Muzny D, Smith R, Sougnez C, Stewart C, Ward A, Yu J, Xue Y, Altshuler D, Bustamante CD, Clark AG, Daly M, DePristo M, Flicek P, Gabriel S, Mardis E, Palotie A, Gibbs R and Genomes P (2011). The functional spectrum of low-frequency coding variation. *Genome Biol* **12**: R84.
- Menard V, Eap O, Roberge J, Harvey M, Levesque E and 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. *Pharmacogenet Genomics* **21**: 631-641.
- Nakajima M, Koga T, Sakai H, Yamanaka H, Fujiwara R, Yokoi T (2010). N-Glycosylation plays a role in protein folding of human UGT1A9 *Biochem Pharmacol.* **79**:1165-72.
- Nagaoka K, Hanioka N, Ikushiro S, Yamano S and Narimatsu S (2012). The effects of N-glycosylation on the glucuronidation of zidovudine and morphine by UGT2B7 expressed in HEK293 cells. *Drug Metab Pharmacokinet* **27**: 388-397.
- Nelson MR, Wegmann D, Ehm MG, Kessner D, St Jean P, Verzilli C, Shen J, Tang Z, Bacanu SA, Fraser D, Warren L, Aponte J, Zawistowski M, Liu X, Zhang H, Zhang Y, Li J, Li Y, Li L, Woollard P, Topp S, Hall MD, Nangle K, Wang J, Abecasis G, Cardon LR, Zollner S, Whittaker JC, Chissoe SL, Novembre J and Mooser V (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**: 100-104.
- Ramsey LB, Bruun GH, Yang W, Trevino LR, Vattathil S, Scheet P, Cheng C, Rosner GL, Giacomini KM, Fan Y, Sparreboom A, Mikkelsen TS, Corydon TJ, Pui CH, Evans WE and

- Relling MV (2012). Rare versus common variants in pharmacogenetics: SLCO1B1 variation and methotrexate disposition. *Genome Res* **22**: 1-8.
- Riches Z and Collier AC (2015). Posttranscriptional regulation of uridine diphosphate glucuronosyltransferases. *Expert Opin Drug Metab Toxicol* **11**: 949-965.
- Sneitz N, Bakker CT, de Knegt RJ, Halley DJ, Finel M and Bosma PJ (2010). Crigler-Najjar syndrome in The Netherlands: identification of four novel UGT1A1 alleles, genotype-phenotype correlation, and functional analysis of 10 missense mutants. *Hum Mutat* **31**: 52-59.
- Stingl JC, Bartels H, Viviani R, Lehmann ML and Brockmoller J (2014). Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther* **141**: 92-116.

**Footnotes:** This work is supported by the Canadian Institutes of Health (CIHR MOP-142318). CGB received studentships from the Fonds d'enseignement et de recherche (FER) of Laval University's Faculty of Pharmacy and from the Fondation Desjardins and Fondation du CHU de Québec. CG holds a Canada Research Chair in Pharmacogenomics (Canadian Research Chair Program).

## Legends for Figures

**Figure 1.** **A)** Sanger sequencing chromatogram revealing an individual heterozygote for the novel nonsense variation at position 361 ((NC\_000004.12:g.69096881G>A, NM\_001074.2:c.361G>A, NP\_001065.2:p.Asp121Asn)) of the *UGT2B7* gene. The arrow indicates the position of the nucleotide variation. **B)** N-glycosylation profile of the UGT2B7 enzyme (Asp<sup>121</sup>) and variant protein (Asn<sup>121</sup>) assessed by Western blot analysis. Microsomal preparations (20 ug) were untreated or treated with endoglycosidases to cleave attached oligosaccharides. **C)** Inhibition of endogenous N-glycosylation. HEK293 cells transiently transfected to express UGT2B7 (Asp<sup>121</sup>) or variant UGT2B7 (Asn<sup>121</sup>) were treated with either vehicle (CTR), tunicamycin (0.1 or 1.0 µg/ml) or thapsigargin (0.5 µM) at time of transfection. UGT2B7 were immunodetected in cell homogenates (10 µg). **D)** Graphical representation of the UGT2B7 protein sequence with putative domains, N-glycosylation sites (blue), those experimentally validated (marked with a star, (Nagaoka et al. 2012)), and the new site created in the UGT2B7Asn<sup>121</sup> variant protein (red) as predicted with the NetNGlyc 1.0 tool. **E)** Glucuronide formation by the UGT2B7Asp<sup>121</sup> reference enzyme (blue) and the variant UGT2B7Asn<sup>121</sup> (red) using varying concentrations of MPA (25 to 1500 µM), AZT (100 to 5000 µM) or UDPGA (50 to 5000 µM in the presence of 500 µM of AZT). **F)** Partial sequence alignment of the 19 UGT1 and UGT2 enzymes in which putative N-glycosylation motifs (grey) and positions of rare variants (listed in Table 1) either creating (green) or disrupting (red) such motifs are indicated. Amino acid positions given above sequence alignments correspond to the UGT2B7 protein and letters represent clusters of putative N-glycosylation sites among UGTs (see **Table 2**)



**F.** Sequence alignment of UGT2B7 variants

	67(a)	121	142	165(b)	189	213(c)	315(d)	322	342(e)	349	402(f)	434	482(g)
UGT2B4	DPNS <b>E</b> STLK	TFNDILRKF	QES	LLKIP	FPP	NMI	SNTS	NVI	NKPD	LNT	NIA	NDPL	DLT
UGT2B7	DPN <b>S</b> SALK	IFG <b>D</b> TRKF	QES	LFN <b>E</b>	FPP	NMI	SNMT	NVI	NKPD	LNT	NIA	NDPS	DLT
UGT2B10	DPND <b>S</b> STLK	AINDIIRNF	QES	LFNIP	FPP	NML	SNMT	NVI	NKPD	LNT	NIA	NDPS	NLT
UGT2B11	DPND <b>A</b> STLK	ELYDIFRNF	QES	LLNIR	FPP	NMI	SNMT	NVI	NKPD	LNT	NIA	NDPL	DLT
UGT2B15	NASKSSAIK	EYYDYSNKL	QES	LFN <b>E</b>	FPP	NMI	SNMS	NMI	KKPN	SNT	NIA	NDPV	NLT
UGT2B17	NASKSSAIK	EYSDYNIKL	QES	LLNIP	FPP	NMI	SNMS	NMI	KKPN	SNT	NIA	NDPI	NLT
UGT2B28	DPND <b>A</b> FTLK	EFHDIFRNF	QES	LLNIP	FPP	NMI	SNMT	NVI	NKPD	LNT	NIA	NDPS	DLT
UGT1A1	RDGAF--YT	KDSAMLLSG	AES	YLSLP	NPF	NML	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A3	KEENF--FT	NMSLVY <b>RS</b>	NAT	YLSIP	NPS	NML	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A4	KEEF--FT	NVSLALHRC	NAT	YLSIP	NPS	NML	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A5	KEENF--FT	NMSLIHRS	HAT	YLSIP	NPS	NML	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A6	KESKY--YT	IVIGLYFIN	KES	YGLGP	DPV	NFL	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A7	GRSLN--CT	GIFDLFFSN	KES	YFSLP	APL	NHI	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A8	GKSLN--CT	GFFNLFFSH	KES	YFSLP	APL	NHI	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A9	GRSLN--CT	DIFDLFFSN	KES	YFSLP	APL	NHI	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT1A10	ERSLN--CT	GFLDLFFSH	KES	YFSLP	APL	NHI	SEIP	MAI	TRPS	NNT	NAK	NDKS	DLT
UGT2A1	TPTS <b>N</b> PSLT	DFHMVSQEI	KKS	KLGIY	YPP	NFI	KNLT	NLI	KKPA	NNT	NIA	NEPS	DLT
UGT2A2	NSNPDSV <b>N</b>	TFFQINQL	QKG	KLGIY	APV	NTI	KNLT	NLI	KKPA	NNT	NIA	NEPS	DLT
UGT2A3	DYRKPSALK	EIRGTLKMM	QET	LLAVP	APL	NSM	QNV <b>T</b>	NI <b>I</b>	KK <b>P</b> S	AN <b>T</b>	NIA	TD <b>S</b> S	DLT

Legend:  
 ■ NX(S/T) motifs  
 ■ Sites disrupted  
 ■ Sites created

**TABLE 1.**

*Kinetic parameters for UGT2B7 substrates (AZT and MPA) using microsomal protein preparations isolated from HEK293 cells stably expressing the UGT2B7 enzyme (Asp<sup>121</sup>) and the novel variant UGT2B7 protein (Asn<sup>121</sup>)<sup>1</sup>.*

	Substrate			Substrate			Co-substrate		
	MPA			AZT			UDPGA <sup>2</sup>		
	K <sub>m</sub>	Relative V <sub>max</sub>	CL <sub>int</sub>	K <sub>m</sub>	Relative V <sub>max</sub>	CL <sub>int</sub>	K <sub>m</sub>	Relative V <sub>max</sub>	CL <sub>int</sub>
UGT2B7	$\mu\text{M}$	$\text{pmol/min/mg/UGT}$ content	$\mu\text{L/min/mg}$	$\mu\text{M}$	$\text{pmol/min/mg/UGT}$ content	$\mu\text{L/min/mg}$	$\mu\text{M}$	$\text{pmol/min/mg/UGT}$ content	$\mu\text{L/min/mg}$
Asp <sup>121</sup>	174 ± 46	2194 ± 65	13	1018 ± 152	267 ± 13	0.26	314 ± 37	139 ± 12	0.45
Asn <sup>121</sup>	152 ± 29	1323 ± 121**	8	968 ± 100	137 ± 21*	0.14	231 ± 24	58 ± 6*	0.25

<sup>1</sup>Rates of glucuronide formation were adjusted relative to UGT2B7 protein content determined by Western blot. Values correspond to the mean of at least two independent experiments performed in triplicate. Relative Acyl-MPA-G, AZT-G formation rates and co-substrate utilization rates were significantly different between UGT2B7Asp<sup>121</sup> and UGT2B7Asn<sup>121</sup>; CL<sub>int</sub> = clearance (V<sub>max</sub>/K<sub>m</sub>); \*P < 0.05 and \*\*P < 0.01 vs. Asp<sup>121</sup>. <sup>2</sup>In the presence of fixed concentration of AZT (500 μM). Similar results were observed using AZT at 1 mM (K<sub>m</sub> = 400 ± 26 μM and V<sub>max</sub> = 156 ± 3 pmol/min/mg/UGT content for UGT2B7Asp<sup>121</sup> and K<sub>m</sub> = 240 ± 31 μM and V<sub>max</sub> = 56 ± 2 pmol/min/mg/UGT content for UGT2B7Asn<sup>121</sup>).

**TABLE 2.** Naturally occurring rare variants in UGT2B family members predicted to create (gain) or disrupt (loss) N-glycosylation sites (NX(S/T) motif).

Clusters	UGT	Variant	Position chr.4	Amino acid	NX(S/T) motif		MAF (%)
					sequence	effect	
	UGT2B4	rs767028471	69495657	Pro69Ser	NSP	gain	< 0.01
	UGT2B7	rs748435925	69096719	Asn67His	NNS	loss	< 0.01
	UGT2B7	rs373587868	69096729	Ser70Phe	NSS	loss	< 0.01
a	UGT2B10	rs61750900	68816218	Asp67Asn	DSS	gain	<0.01
	UGT2B11	rs756939149	69214518	Ala69Thr	NDA	gain	< 0.01
	UGT2B15	rs770648228	68670425	Asn65Ser	NAS	loss	< 0.01
	UGT2B28	rs150261084	69280705	Ala69Thr	NDA	gain	0.14
-	UGT2B15	rs368012995	68670362	Leu86Ser	NYL	gain	-
-	UGT2B7	This report	69096881	Asp121Asn	DIT	gain	0.33
b	UGT2B7	rs747704916	69097019	Pro167Ser	NIP	gain	< 0.01
	UGT2B15	rs747378153	68670117	Pro168Thr	NIP	gain	< 0.01
-	UGT2B10	rs764895973	68816534	Ser172Asn	SFS	gain	< 0.01
c	UGT2B4	rs760430619	69495218	Ile215Thr	NMI	gain	< 0.01
	UGT2B17	rs747935682	68567838	Ile216Thr	NMI	gain	< 0.01
-	UGT2B17	rs748669369	68560628	Ile305Thr	NGI	gain	< 0.01
	UGT2B4	rs756195714	69489500	Ser314Asn	SNT	gain	-
	UGT2B7	rs755080081	69102879	Asn315Tyr	NMT	loss	< 0.01
d	UGT2B7	rs752908839	69102885	Thr317Ala	NMT	loss	< 0.01
	UGT2B28	rs768879480	69286824	Asn315Asp	NMT	loss	< 0.01
	UGT2B28	rs144043239	69286831	Thr317Ile	NMT	loss	< 0.01
-	UGT2B7	rs745439209	69102907	Ile324Thr	NVI	gain	< 0.01
e	UGT2B10	rs767451406	68826437	Pro343Ser	NKP	gain	< 0.01
	UGT2B11	rs201637982	69205540	Pro344Thr	NKP	gain	< 0.01
f	UGT2B7	rs201964275	69108222	Ala404Thr	NIA	gain	< 0.01
	UGT2B10	rs750605168	68827448	Ala403Ser	NIA	gain	< 0.01
-	UGT2B10	rs545266191	68827497	Met419Thr	NTM	gain	0.02
-	UGT2B4	rs375220784	69480893	Met443Thr	NAM	gain	< 0.01
-	UGT2B28	rs781334372	69294551	Lys444Asn	KLS	gain	< 0.01
	UGT2B4	rs767837637	69480777	Asp482Asn	DLT	gain	< 0.01
g	UGT2B10	rs112561475	68830733	Asn481Asp	NLT	loss	1.30
	UGT2B10	rs567502684	68830734	Asn481Ser	NLT	loss	< 0.01
	UGT2B10	rs536035975	68830740	Thr483Asn	NLT	loss	0.10
-	UGT2B4	rs770843089	69480747	Asp492Asn	DVT	gain	< 0.01

Horizontal lines group variants according to their positions. MAF: Minor allele frequency.



## Discussion

La voie de glucuronidation, impliquée dans le métabolisme de plus de la moitié des médicaments les plus prescrits (Guillemette et al. 2014), est médiée par plusieurs enzymes UGT. Il existe diverses sources à l'origine des différences observées entre les individus en ce qui a trait au métabolisme par les enzymes UGT. Ces sources de variation sont plus ou moins bien définies et étudiées selon les composés et les enzymes impliquées. Des variations génomiques des gènes UGT contribuent à la variabilité interindividuelle de l'activité de cette voie. Par exemple, l'allèle UGT1A1\*28, qui en opposition à l'allèle plus courant UGT1A1\*1 possède une répétition TA supplémentaire au sein de la boîte TATA de son promoteur, est associée à une réduction de la transcription du gène et une réduction de la glucuronidation par l'enzyme UGT1A1 (Hall et al. 1999). Ainsi, la monographie du médicament irinotécan, utilisé dans le traitement des cancers colorectaux, suggère que la dose administrée aux individus homozygotes pour cette allèle, représentant 10% de la population nord-américaine, soit réduite de 20% afin d'éviter la manifestation d'effets secondaires indésirables liés à une trop faible élimination du composé actif (Etienne-Grimaldi et al. 2015). Une étude allemande a récemment établi que le génotypage des individus atteints du cancer du côlon traités à l'irinotécan était non seulement avantageux du point de vue médical, mais également monétaire (Butzke et al. 2016). Cet exemple dénote bien l'importance d'identifier les sources de variation de la voie de glucuronidation afin d'optimiser la posologie tout en minimisant le risque d'effets indésirables et ainsi potentiellement améliorer le traitement des patients.

Plusieurs variations génétiques fréquentes, notamment des polymorphismes ponctuels, ont été investiguées afin d'expliquer la variabilité observée dans l'activité des enzymes UGT envers divers substrats xénobiotiques. Toutefois, cette variabilité pourrait également être du moins partiellement prédite par des variants dont la fréquence allélique serait plus faible. En effet, bien que chacune d'entre elle soit rare, ces variations se trouvent collectivement en plus grand nombre au sein du génome que celles plus courantes dans la population (Nelson et al. 2012). Les enzymes du métabolisme de phase I des médicaments cytochromes P450 illustrent ce fait puisqu'elles sont issues de gènes dans lesquels 83% des

variations ponctuelles ont une fréquence allélique en deçà de 1% (Fujikura et al. 2015). De tels variants à faible fréquence allélique sont non seulement d'intéressants acteurs potentiels dans la variabilité de la réponse aux médicaments de par leur abondance dans le génome humain, mais aussi à cause de l'impact qu'ils peuvent avoir sur l'activité enzymatique, en affectant par exemple le promoteur d'un gène ou encore en entraînant la modification d'un acide aminé. En effet, le changement d'un acide aminé au sein d'une protéine peut se traduire en une panoplie d'effets selon la position, la nature et le rôle du résidu impliqué. Parmi ces effets d'une substitution d'acide aminé, on retrouve entre autres la possibilité qu'un motif de modification post-traductionnelle soit altéré, ce qui peut en retour causer un changement de l'activité enzymatique.

Au laboratoire, la recherche de variants susceptibles d'expliquer la variabilité de glucuronidation a été faite dans le cadre d'études portant sur le rôle de divers membres de la famille des UGT dans la pharmacocinétique du mycophénolate mofétil, un agent immunosuppresseur (Benoit-Biancamano et al. 2007, Levesque et al. 2007, Levesque et al. 2008). Ces études démontrent une variabilité de plus de 50% entre les individus dans la formation l'acyl-MPA-glucuronide (AcMPAG) par l'enzyme UGT2B7. Ce métabolite est formé en moindre quantité en comparaison avec le métabolite inactif phénol-MPA-glucuronide (issu de la glucuronidation du MPA par les UGT1A8 et UGT1A9), mais possède une importance clinique puisqu'il est actif et est lié à divers effets indésirables de nature inflammatoire (Maes et al. 2002).

Dans le cadre de l'analyse des séquences UGT de 305 sujets sains, deux individus porteurs d'une variation rare non-synonyme affectant le codon 121 d'UGT2B7 ont été identifiés, établissant la fréquence allélique de ce variant dans cette population à 0,328%. Le changement du nucléotide en position 361 de la région codante du gène, une guanine, en une adénine mène à la traduction d'une asparagine plutôt qu'un acide aspartique en position 121 de la protéine. Cette substitution crée à cette position un motif de N-glycosylation validé par plusieurs outils de prédiction bio-informatique. La validation expérimentale de ce nouveau site a été effectuée suite à la surexpression de l'enzyme variante dans des cellules HEK293 (UGT négatives), laquelle est ensuite comparée à la protéine UGT2B7 de

référence. Brièvement, les microsomes issus de ces deux surexpressions sont analysés par immunobuvardage de type Western, analyse qui permet d'observer que la protéine mutante possède un poids moléculaire plus élevé que la protéine de référence, le poids supplémentaire correspondant à celui entraîné par l'ajout d'un oligosaccharide sur la protéine mutante. Ceci est de surcroit vérifié par l'analyse par immunobuvardage de microsomes traités à l'Endo H, enzyme qui clive les oligosaccharides propres à la N-glycosylation en leur point d'attache avec l'asparagine de la protéine. Suite au traitement à l'Endo H, les bandes correspondant à chaque surexpression migrent plus bas sur la membrane, conséquence de la perte des oligosaccharides. Également, on n'observe plus de différence de hauteur entre les bandes, permettant de conclure que la différence observée entre les échantillons non-traités était bel et bien due à un site de N-glycosylation additionnel. L'O-glycosidase n'a quant à elle aucun effet sur la migration des échantillons, montrant que ce type de glycosylation (l'O-glycosylation) n'entre pas en jeu dans la différence de hauteur observée entre les bandes. De surcroit, la présence d'un site de N-glycosylation additionnel modifiant la migration sur gel de la protéine variante est vérifiée par un traitement à la tunicamycine dans des cellules exprimant de façon transitoire ou le variant ou l'UGT2B7 de référence. Ce traitement inhibe toute N-glycosylation et suite à celui-ci on observe des bandes à des hauteurs égales pour les deux protéines, corroborant à nouveau l'idée que la différence de hauteur soit due à un site de N-glycosylation additionnel sur l'UGT2B7 variante au codon 121. Enfin, l'inhibition observée n'est pas due à un stress du réticulum endoplasmique perturbant la glycosylation puisque la thapsigargine n'a affecté la mobilité d'aucun des deux échantillons.

L'effet de ce nouveau site de N-glycosylation sur l'activité de l'enzyme a été évalué à l'aide d'essais enzymatiques. En effet, à partir de ces essais sont calculées des constantes cinétiques, constante d'affinité ( $K_m$ ), vitesse ( $V_{max}$ ) et clairance, lesquelles nous informent sur les paramètres affectés par la mutation. Les essais sont effectués avec les microsomes de cellules HEK293 sur exprimant ou l'enzyme UGT2B7 de référence ou l'enzyme mutante au codon 121. Deux substrats ont été testés en des concentrations variables, le MPA et l'AZT. Ce dernier a aussi été testé en concentrations fixes avec variation de la concentration du co-substrat UDPGA afin d'établir également les constantes

liées à son utilisation par l'enzyme. Bien que le codon 121 se situe au domaine de liaison au substrat de l'UGT2B7, l'affinité ( $K_m$ ) des enzymes pour les substrats testés ne variait pas selon s'il s'agissait de l'enzyme mutante ou non. L'impact fonctionnel de variations dans ce domaine est difficile à prédire puisque ce dernier se situe dans une portion de la protéine dont la structure tridimensionnelle n'est pas connue, rendant ardue l'identification des résidus impliqués dans la réaction enzymatique. D'autre part, la vitesse ( $V_{max}$ , courbe de saturation) de formation de l'AZT-glucuronide, de l'AcMPAG et d'utilisation du co-substrat UDPGA étaient toutes diminuées pour la protéine mutante au codon 121 en comparaison avec l'UGT2B7 classique.

La diminution de la vitesse de formation des substrats et d'utilisation du co-substrat couplée à une absence de modification de l'affinité malgré la position de la N-glycosylation laisse croire à une implication de l'oligosaccharide dans le repliement adéquat de la protéine plutôt qu'au site catalytique. Cette hypothèse sera vérifiée d'abord en effectuant des essais enzymatiques sur des microsomes traité ou non avec l'Endo H. Ces essais nous permettraient d'observer si d'enlever l'oligosaccharide de la protéine alors qu'elle est déjà repliée cause un changement à son activité. En utilisant ce test Nagaoka et al ont établi que les sites de glycosylations usuels d'UGT2B7 ne semblaient pas essentiels à son activité pour l'AZT une fois la protéine mature et repliée (Nagaoka et al. 2012). En effet, ils n'ont pas observé de différence dans l'activité des microsomes traités ou non à l'endoglycosidase. Ils concluent donc que l'effet sur l'activité enzymatique qu'ils observent lors de la mutation des sites de N-glycosylation est dû à un mauvais repliement de la protéine, tel que validé pour l'UGT1A9 par l'équipe de Nakajima (Nakajima et al. 2010). Cette équipe avait utilisé la tunicamycine afin d'inhiber dans des cellules toute N-glycosylation, ce qui contrairement au traitement à l'endoglycosidase avait significativement réduit l'activité enzymatique. Dans le cas du variant au codon 121, un tel traitement permettra de déterminer si la glycosylation de ce résidu influence l'activité de l'enzyme via un effet sur son repliement.

L'impact de la variation au codon 121 pourrait également être étudiée *in vivo*, en procédant à une étude pharmacocinétique concernant des sujets porteurs. En effet, il se pourrait que

cette variation soit à même d'expliquer une fraction du phénotype de métabolisateur lent observé chez certains sujets. Toutefois, ceci requerra le séquençage d'une cohorte plus nombreuse. Aussi, tant *in vitro* qu'*in vivo*, il serait intéressant de tester l'effet du variant sur l'activité d'UGT2B7 pour d'autres substrats exogènes de l'enzyme tels la morphine, l'épirubicine, l'efavirenz, ainsi que des substrats endogènes tels les prostaglandines E2 et l'acide arachidonique.

D'autres variations affectant la N-glycosylation pourraient perturber l'activité d'enzymes de la famille des UGT2B. En effet, des variants ponctuels peuvent mener au gain ou à la perte de sites de N-glycosylation en modifiant le motif NX(S/T). Une recherche systématique de tels variants pour l'ensemble du protéome a été réalisée récemment par Mazumder et al. sans toutefois identifier de variations affectant les enzymes UGT (Mazumder et al. 2012). Cette recherche n'a en effet été faite que sur les variations qui étaient associées à des séquences protéiques dans la base de données Uniprot, laquelle ne répertorie pas tous les variants. C'est pourquoi dans le cadre du projet présenté ici tous les variants non-synonymes répertoriés dans la base de données dbSNP pour chaque *UGT2B* ont été manuellement associées aux séquences des enzymes. Une fois les variations triées selon si elles modifiaient les asparagines, sérines ou thréonines ou menaient à la traduction d'un de ces trois résidus, une inspection visuelle des acides aminés avoisinants permettait de déterminer si un motif de N-glycosylation était aboli ou créé. Ce faisant, 32 variations, pour la plupart rares (fréquence allélique < 1%), ont été identifiées comme influençant potentiellement la N-glycosylation d'enzymes de la famille des UGT2B.

Étant donnée la forte homologie de séquence entre les gènes de la famille UGT2B, plusieurs des 32 variations identifiées se trouvent à être regroupées selon leur localisation lorsqu'on effectue l'alignement des séquences protéiques des UGT2B. Ces variations sont particulièrement intéressantes car certaines créent des sites de N-glycosylation à une position d'une UGT2B qui équivaut dans une autre à un site de glycosylation validé expérimentalement. C'est le cas de variants d'*UGT2B4*, *UGT2B11* et *UGT2B28* qui créent des motifs de N-glycosylation à des positions avoisinants celles des premiers sites de N-glycosylation validés de la séquence des UGT2B7 et UGT2B15 (Chen et al. 2009). C'est

aussi le cas d'un variant affectant le résidu 482 d'UGT2B4 créant à cette position l'asparagine d'un site de N-glycosylation qui s'aligne à celui en position 481 d'UGT2B10, lequel a aussi été validé expérimentalement (Chen et al. 2009).

La base de données dbSNP, bien que réunissant les données de plusieurs autres bases, ne contient que les variations auxquelles ont été assignées des identifiants rsID. Ainsi, l'inclusion dans cette étude de variations provenant d'autres bases de données telles le *Exome Aggregation Consortium* (ExAC), qui contient également des variations qui n'ont pas encore de rsID, permettrait de broser un tableau plus complet des variations affectant potentiellement la N-glycosylation des UGT2B. De plus, les outils de prédiction du potentiel de N-glycosylation utilisés dans l'étude du variant au codon 121 pourraient servir à établir, à partir des variations répertoriées dans l'étude comme créant un nouveau motif de N-glycosylation, une liste de celles dont le potentiel est le plus élevé, guidant ainsi le choix de variations à investiguer davantage.

## Conclusion

En conclusion, le séquençage de patients d'études précédentes sur la pharmacocinétique du mycophénolate mofétil a permis d'identifier un nouveau variant ponctuel dont la position au sein de la séquence mène à la création d'un motif de N-glycosylation. Ce site est prédit comme ayant un haut potentiel de glycosylation *in silico* et a été validé expérimentalement par l'analyse par immunobuvardage de type Western de traitements à l'Endo H. Il a également été établi que le variant entraînait une diminution de la vitesse de formation des glucuronides à partir des substrats AZT et MPA et de la vitesse d'utilisation du co-substrats UDPGA en présence stable d'AZT. Aucun effet n'a été décelé sur l'affinité de l'enzyme variante pour ces substrats. Ensemble, ces résultats laissent croire à un effet de l'oligosaccharide supplémentaire sur le repliement de la protéine, ce qui pourrait être vérifié par des traitements à l'endoglycosidase et la tunicamycine. L'effet du variant pourrait se traduire *in vivo* en un phénotype de métabolisateur lent, ce qui serait à confirmer grâce à des études de pharmacocinétiques sur des sujets porteurs. Le séquençage d'un plus grand nombre de sujet permettrait également une meilleure estimation de la fréquence allélique. En outre, l'analyse de bases de données de variants ponctuels a permis

d'identifier 32 variations qui pourraient affecter la N-glycosylation d'enzymes UGT2B et potentiellement leur activité enzymatique. Des études subséquentes seront nécessaires afin de vérifier cet effet potentiel et afin d'établir si des variants ponctuels pourraient affecter de la même façon l'activité enzymatique en altérant des motifs associés à d'autres modifications post-traductionnelles.

## Annexe : Divergences dans l'expression et la fonction métabolique des glucuronosyl transférases humaines via l'épissage alternatif

### Résumé

Le maintien de l'homéostasie cellulaire et la détoxification de xénobiotiques sont médiés par 19 UDP-glucuronosyl transférases (UGT) humaines, encodées par 10 gènes, qui composent la voie de glucuronidation. Le séquençage en profondeur d'organes majeurs du métabolisme met au jour une augmentation substantielle du transcriptome UGT via l'épissage alternatif, avec des variants représentant de 20 à 60% de l'expression des transcrits canoniques. Près du cinquième des variants exprimés comprennent des séquences conservant le cadre de lecture, leur permettant de créer des caractéristiques structurelles et fonctionnelles distinctes. Des fonctions biologiques propres à ces protéines UGT alternatives ont été subséquemment révélées lors d'essais dans des cellules. Certaines isoformes ont été identifiées comme inhibant ou induisant l'inactivation de médicaments et stéroïdes ainsi que comme perturbant le métabolisme global (énergie, acides aminés, nucléotides), l'adhésion et la prolifération cellulaire. Les travaux présentés dans cet article mettent en lumière l'importance biologique de l'expression alternative des UGT. Nous proposons que cette dernière augmente la diversité protéique à travers l'évolution de régulateurs métaboliques provenant d'enzymes spécifiques.

**Article publié dans le journal *Cell Reports***



## **Divergent expression and metabolic functions of human glucuronosyltransferases through alternative splicing**

Michèle Rouleau<sup>1</sup>, Alan Tourancheau<sup>1</sup>, Camille Girard-Bock<sup>1</sup>, Lyne Villeneuve<sup>1</sup>, Jonathan Vaucher<sup>2</sup>, Anne-Marie Duperré<sup>1</sup>, Yannick Audet-Delage<sup>1</sup>, Isabelle Gilbert<sup>1</sup>, Ion Popa<sup>2</sup>, Arnaud Droit<sup>2</sup>, Chantal Guillemette<sup>1</sup>‡

<sup>1</sup>Pharmacogenomics Laboratory, Centre Hospitalier Universitaire (CHU) de Québec Research Center, Québec, QC, Canada and Faculty of Pharmacy, Laval University, Québec, Canada, G1V 4G2

<sup>2</sup>Faculty of Medicine, Laval University, Québec, Canada, G1V 0A6

‡Canada Research Chair in Pharmacogenomics

Corresponding Author/Lead Contact:

Chantal Guillemette, Ph.D.

Canada Research Chair in Pharmacogenomics

Pharmacogenomics Laboratory

CHU de Québec Research Center

2705 Blvd. Laurier, R4720

Québec, Canada, G1V 4G2

Tel. (418) 654-2296

E-mail: [Chantal.Guillemette@crchudequebec.ulaval.ca](mailto:Chantal.Guillemette@crchudequebec.ulaval.ca)

## **Summary**

Maintenance of cellular homeostasis and xenobiotic detoxification is mediated by 19 human UDP-glucuronosyltransferase enzymes (UGTs) encoded by 10 genes that comprise the glucuronidation pathway. Deep RNA sequencing of major metabolic organs exposes a substantial expansion of the UGT transcriptome by alternative splicing, with variants representing 20 to 60% of canonical transcript expression. Nearly a fifth of expressed variants comprise in-frame sequences that may create distinct structural and functional features. Follow up cell-based assays reveal biological functions for these alternative UGT proteins. Some isoforms were found to inhibit or induce inactivation of drugs and steroids in addition to perturbing global cell metabolism (energy, amino acids, nucleotides), cell adhesion and proliferation. This work highlights the biological relevance of alternative UGT expression, which we propose increases protein diversity through evolution of metabolic regulators from specific enzymes.

## Introduction

Human glycosyltransferases utilize different sugar nucleotide donors to regulate a wide variety of cellular processes (Guillemette et al. 2004, Little et al. 2004, Breton et al. 2012, Rowland et al. 2013, De Bruyn et al. 2015). One large family of membrane-associated uridine diphospho-glucuronosyltransferases (UGTs) employ uridine diphospho-glucuronic acid (UDPGA) to conjugate a variety of small-molecule acceptors and catalyze the most important metabolic pathway for the human body's elimination of both endogenous and exogenous lipophilic molecules. These endoplasmic reticulum (ER) resident type I membrane proteins transfer the sugar moiety of the co-substrate UDPGA to the nucleophilic functional group of their substrates to promote, in most cases, their inactivation and elimination. Glucuronidation affects bioactivity and bioavailability of harmful xenobiotics, chemicals, and drugs from food, the environment, or pharmacological treatments. In parallel, glucuronidation maintains the homeostasis of the heme-breakdown product bilirubin, multiple endogenous hormones, secondary metabolites, and other endobiotics (Radomska-Pandya et al. 1999, Guillemette 2003, Wells et al. 2004, Bock 2015).

In humans, 10 genes encode 19 canonical UGT enzymes (isoform 1s or 11s) that display remarkable plasticity. For example, a single *UGT1* gene on chromosome 2q37 encodes nine UGT1A enzymes through usage of individual alternative promoters and first exons, whereas 10 UGT2 enzymes (three UGT2A and seven UGT2B) are synthesized from nine independent genes clustered on chromosome 4q13 (Mackenzie et al. 2005). Although amino acid sequences among UGTs are highly similar, substrate specificity is dictated by the slightly more divergent first (UGT1) or first/second (UGT2) exons that encode the N-terminal half of each enzyme. UGTs are found in nearly all human tissues and display tissue- and cell type-specific expression, but they are most abundant in metabolically highly active organs such as the liver, kidney, and intestine (Court et al. 2012, Rowland et al. 2013, Margaillan et al. 2015, Margaillan et al. 2015).

Alternative splicing (AS) is a key mechanism in the control of gene expression, transcriptomes, and protein diversity, with over 90% of human multi-exon genes undergoing AS regulation before forming mature transcript isoforms (Wang et al. 2008). The important contributions of AS to an individual's response to endogenous and exogenous molecules, including drugs, and the link to various human diseases are emerging but remain largely unknown (Garcia-Blanco et al. 2004, Gamazon and Stranger 2014, Oltean and Bates 2014). Our recent work established that AS expands the coding capacity of human *UGTs*, allowing over 180 canonical and alternative *UGT* transcripts in various tissues involved in drug metabolism as well as hormone-dependent tissues (Tourancheau et al. 2016). Thus, a large genetic diversity characterizes the transcriptome landscape of human *UGTs* and implies a functional diversification of the UGT proteome that remains unexplored.

In this study, we applied targeted next-generation RNA sequencing (CaptureSeq) (Mercer et al. 2012, Clark et al. 2015) to quantify the UGT transcriptome and expression of alternative transcripts in the liver, kidney, and gastrointestinal tissues, which are most relevant to the metabolic functions of this crucial enzymatic pathway. We further addressed tissue-specific protein expression as well as the function of selected UGT isoforms where in-frame sequences have been introduced using cell-based assays combined with untargeted metabolomics (**Fig. 1**). Our findings reveal a complex expression pattern of UGT alternative variants and distinct functional properties, including antagonizing and inducing UGT transferase function, likely through protein–protein interaction, while inducing substantial rewiring of cell metabolism that affects cellular behavior.

## Results

### **AS contributes to quantitative profiles of the UGT transcriptome in human metabolic tissues**

A quantitative transcriptome analysis of all 10 human *UGT1* and *UGT2* genes was conducted in pooled samples of normal liver, kidney, and intestine/colon tissues as well as in tumors originating from the kidney and the intestine/colon. A CaptureSeq approach was used to achieve sufficient sequencing depth and ensure the fullest coverage of alternative UGT variants. Reads were mapped to the human genome sequence (hg19) complemented with the recently published comprehensive human UGT transcriptome (Tourancheau et al. 2016) (**Fig. 1**). The data revealed that AS contributes substantially to the UGT transcriptome landscape, affecting each expressed *UGT1* and *UGT2* gene in a tissue-specific manner, and the AS products are significantly altered in neoplastic tissues (**Fig. 2A, B**). Alternatively spliced species were abundant in normal tissues, representing an average of 19, 24, and 34% of the total UGT expression in liver, intestine/colon, and kidney, respectively (**Fig. 2C**). In intestine/colon and kidney tumor tissues, these levels were strikingly elevated to an average of 33 and 61%, respectively. Changes in liver cancer could not be assessed owing to the lack of availability of hepatic tumors. The abundance and tissue-specific expression of several alternative transcripts supports a physiologically relevant role in the surveyed tissues.

### **Diverse structural features of alternative UGT isoforms**

A common structural organization is shared by the 19 known human *UGT1* and *UGT2* enzymes, each encoded by a canonical mRNA variant v1. An N-terminal signal peptide and a C-terminal transmembrane region direct the substrate and UDPGA co-substrate-binding domains and catalytic site of each mature protein to the luminal side of the ER (**Fig. 3A**). Based on these structural features, 164 expressed non-v1 alternative transcripts were classified into five categories (**Fig. 3B**). Nearly half of expressed alternative UGT variants lacked a sequence encoding the substrate-binding domain (N-terminal), frequently due to truncation or skipping of exon 1, and this alteration was especially frequent among *UGT2* transcripts. In the absence of the N-terminal substrate-binding domain, the encoded proteins

would be expected to lack transferase activity. Another prominently expressed class of variants that constituted over one-third of alternative variants lacked part or all of the exons encoding the C-terminal co-substrate domain. Remarkably, over 20% of the expressed variants comprised introduction of in-frame sequences. The domain organization and sequences of the putative isoforms imply that several may have different subcellular distribution, substrate specificity, or catalytic activity and thus could have altered biological functions.

### **Functional diversity of selected alternate UGTs containing in-frame sequences**

*Shorter UGT1 isoforms with an alternative C-terminal sequence have antagonistic functions.*

A single *UGT1* locus encodes half of the human UGT enzymes, i.e., 9 UGT1A enzymes. It is well established that mature transcripts encoding each UGT1A enzyme include only one of the different *UGT1* exon-1 sequences associated with four downstream exons common to all UGT1A enzymes (i1s). A splicing event involving the use of an alternative 3'-terminal exon (exon 5b) generates three categories of *UGT1* transcripts, namely the canonical *v1* (exon 5a) and alternative *v2* (exon 5b) and *v3* (exons 5b and 5a) (**Fig. 4A**). Our CaptureSeq data and those of an independent RNA sequencing (RNA-Seq) study revealed significant expression of *v2/v3*, with these transcripts representing between 9 and 20% of canonical *UGT1A* transcripts in normal tissues and high interindividual variability (**Fig. S1A, B**). In kidney and intestine/colon tumor tissues, *v2/v3* constituted between 19 and 23% of *UGT1A\_v1* expression. Because of a stop codon in exon 5b, the alternative *v2/v3* variants both encode shorter, 45-kDa isoform 2 proteins (i2s) lacking the C-terminal 99 amino acid residues encoded by exon 5a that comprise the membrane-spanning domain and the short cytosolic charged tail. In alternate UGT1A\_i2 proteins, this sequence is replaced by a charged 10-residue C-terminal sequence not found in any other human proteins. The functions of i2s were studied in the HEK293 human embryonic kidney cell line, in which endogenous UGTs are not detected. UGT1A1\_i2 had a remarkably long half-life compared with that of UGT1A1\_i1 (11.6 h versus 1.3 h), whereas co-expression of the two isoforms did not alter their respective half-lives (**Fig. 4B**). *In situ* enzymatic assays in

intact cells supported the notion that UGT1A1\_i2 lacks transferase activity; rather, it has an antagonist role, leading to a significant reduction in UGT1A1\_i1-mediated glucuronidation of the anti-cancer agent SN-38 (the active metabolite of irinotecan;  $-73.5\%$ ;  $P < 0.001$ ) and the endogenous substrate estradiol ( $E_2$ ;  $-75\%$ ;  $P < 0.001$ ) (**Fig. 4C**). These antagonistic effects of i2 are consistent with the co-immunoprecipitation (IP) of UGT1A\_i2 with UGT1A\_i1 (**Fig. 4D**) as well as with the co-localization of both isoforms in the ER as assessed by immunofluorescence (IF; **Fig. 4E**) and in cells of human tissues by immunohistochemistry (IHC; **Fig. 4F**). The data also indicated that UGT1A proteins are distributed in other subcellular compartments (**Fig. S2**). In colon tissues, i1 and i2 co-localized at surface epithelial cells and in intestinal gland (crypt) cells, whereas in liver tissues the two isoforms co-localized in hepatocytes.

*A UGT2B7 isoform with a spacer domain enhances drug inactivation and provokes remodeling of cellular metabolism and phenotypic changes.*

An alternative full-length *UGT2B7\_n4* transcript, confirmed by PCR analysis as containing an introduced sequence, exon 2b (**Fig. S3A**), encodes an alternate UGT2B7 protein termed isoform 8 (UGT2B7\_i8) that has a unique 32-residue in-frame internal region residing at the interface between the N-terminal substrate-binding domain and the C-terminal co-substrate-binding domain (**Fig. 5A**). CaptureSeq data analysis revealed that expression of exon 2b-containing transcripts in three pools of three liver samples was low compared with the canonical transcript (**Fig. S1C**). However, the analysis of an independent RNA-Seq dataset derived from 18 different individuals indicated that the UGT2B7\_i8-encoding transcript represents up to 75% of the canonical UGT2B7 transcripts, with an average coefficient of variation of 350% (**Fig. S1D**). Similar patterns of expression were noted in the kidney, with a greater proportion of UGT2B7\_i8-encoding transcript in kidney tumors relative to the canonical transcript. The encoded UGT2B7\_i8 had a significant half-life of  $\sim 6$  h that is not affected when co-expressed with the i1 protein, which has a half-life of 12 h (**Fig. S3B**). Expression of the encoded i8 protein along with UGT2B7 in human liver was further established in hepatocytes using an i8-specific antibody, supporting the expression data of the alternative isoform and its partial co-localization with UGT2B7\_i1 (**Fig. 5B**). Staining was also observed in smooth muscle cells of hepatic arteries as well as in the same

structures of esophageal, breast, uterine, testicular, and skeletal muscle tissues (data not shown), suggesting a potential unique cell type-specific expression of UGT2B7\_i8 that deserves further investigation. Endogenous alternate protein expression was also corroborated by the detection of peptides specific to the sequence encoded by exon 2b in human liver samples using a targeted mass spectrometry (MS) approach (**Fig. S4**).

UGT-negative (HEK293) and UGT-positive (HepG2) cell models were established to study cellular functions of this UGT2B7 isoform and to reproduce the expression observed in tissues ( $i8 < i1$ ). We could not detect the formation of glucuronide from zidovudine (AZT), the probe substrate of the UGT2B7 enzyme, in UGT2B7\_i8-expressing HEK293 cells. However, the co-expression of UGT2B7\_i8 and UGT2B7\_i1 induced 1.3- and 2.1-fold increases in the formation of the glucuronide in HEK293 and HepG2 cells, respectively, implying a UGT-activating function (**Fig. 5C**). The potential of UGT2B7\_i8 and UGT2B7\_i1 proteins to interact was further supported by co-IP experiments and by their co-localization in the ER and may underlie the induction in glucuronidation activity (**Fig. 5D-E**). Untargeted metabolic profiling of the HEK293 and HepG2 cell models further revealed that the levels of a vast array of cellular metabolites were significantly altered in cells overexpressing UGT2B7\_i8 compared with control cells (**Fig. 5F; Table S1**). Despite the distinctive basal metabolomes of the kidney and liver cell models, expression of UGT2B7\_i8 induced remarkable accumulation of multiple amino acids in both cell types, with glutamine level being most altered. Most purines and pyrimidines were reduced upon i8 expression in HEK cells, whereas several glycolytic and TCA cycle intermediates were significantly reduced in HepG2 cells (**Fig. 5F, Table S1**). These alterations in numerous metabolites essential for cell growth were associated with 3.5-fold enhanced adhesion but 3.7-fold slower proliferation in HEK293 cells overexpressing UGT2B7\_i8 compared with control cells (**Fig. 5G**). In support of a specific effect of the UGT2B7\_i8 protein, adhesion and proliferation of HEK293 cells overexpressing UGT2B7\_i1 were similar to the control cells despite higher expression of i1 than i8.



## Discussion

Our findings reveal large differences in the levels of naturally occurring alternative mRNAs transcribed from *UGT* loci in major metabolic tissues and demonstrate that these alternate transcripts are expressed in a tissue-specific manner, with remodeling of the UGT transcriptome in cancer tissues. Among this large collection of alternative UGT variants, expression of selected alternately spliced isoforms with in-frame sequences was confirmed at the protein level in human tissues such as the liver and colon. In cell-based assays, alternate UGT proteins were shown to differentially regulate the glucuronidation pathway by inhibiting or promoting the inactivation of drugs and hormones, likely through protein–protein interactions. We also uncovered that the alternate UGT isoforms have functions distinct from the canonical glucuronidation function, and these appear to broadly affect global metabolism, which results in changes in cell behavior. These findings imply that this transcriptome diversity expands the proteome and biological functions of human UGTs, likely playing crucial roles in multiple cellular processes.

Our targeted CaptureSeq approach allowed high-resolution, quantitative evaluation of *UGT* expression in the liver, kidney, and intestine/colon. A strength of our study is the mapping of RNA sequencing short reads on the exhaustive human *UGT* transcriptome recently established (Tourancheau et al. 2016), leading to precise assignment of reads to the appropriate *UGT* loci and transcript reconstruction than previously possible. This work revealed that alternative *UGT* mRNAs constitute an appreciable proportion of the total UGT transcriptome in the surveyed tissues, representing 19 to 60% of all *UGT* mRNAs. In line with this, there are many examples in the literature of protein functional expansion created by AS, among which the human tRNA synthetases and brain neurexins are remarkable recent examples (Lo et al. 2014, Treutlein et al. 2014). Our data further indicate that UGT AS programs are constitutive in normal tissues, implying that these events are associated with normal biological processes. This observation was also validated in an independent RNA-Seq dataset not captured for UGT sequences (Chhibber et al. 2016). These findings support a role for AS in controlling coordinated cellular responses induced

by small molecules and exposure to xenobiotics, including drugs of various classes. Given the key role played by UGTs in cellular homeostasis by inactivation of a variety of endogenous lipophilic cellular constituents, one can envision that expression of alternative UGTs may be triggered by the cellular environment and/or stimuli such as hormones and lipids (Hu et al. 2014, Dates et al. 2015). Expression of alternative UGTs is also highly likely to be triggered by exogenous stimuli and toxic xenobiotics that are common substrates of UGTs (Hu et al. 2014). It remains unknown if other enzymes involved in parallel cellular and drug-metabolism functions diversify their expression profiles and functional complexity through splice variants. A recent study indicated that this might be the case for a vast majority (>70%) of factors in drug-related pathways, such as cytochromes P450 and other transferases (Chhibber et al. 2016). However, the regulation and functional consequences of AS events for these clinically important pharmacogenes remain undefined.

We estimate that only a small fraction of the UGT variants we quantified may undergo nonsense-mediated mRNA decay according to the 50-bp rule (Popp and Maquat 2013), and therefore most observed *UGT* transcripts likely have the potential to be translated. Multiple distinct features characterize putative UGT isoforms, including truncated substrate or co-substrate-binding domains and distinctive in-frame sequences. We provide evidence that AS may represent a potential mechanism to modify canonical UGT function and to produce alternative UGT isoforms with introduced domains and divergent functions. Our findings with the *UGT1A* and *UGT2B7* transcripts, which encode, respectively, a shorter protein with an in-frame C-terminal sequence and a longer protein with an internal peptide sequence introduced, clearly support this notion. At the transcript level, *UGT1A* and *UGT2B7* expression shows high interindividual variability in the liver and kidney. Specific antibodies raised against the peptide sequences not found in canonical UGTs or any other known proteins labeled several human tissues (IHC). Co-localization of alternate isoforms with canonical UGTs in the same cell types, such as hepatocytes, was evident, even though relative abundance could not be deduced. According to functional assays in intact proliferating cells, both *UGT1A1\_i2* and *UGT2B7\_i8* lack the ability to transfer a

glucuronic acid moiety to the typical substrates of the canonical enzymes and therefore would be enzymatically null. Alternate UGT1A\_i2s contain complete binding sites for both the substrate and co-substrate and a C-terminal dilysine motif encoded by the new terminal exon, but they lack the transmembrane-binding domain, which likely affects membrane topology and conformation and thus UGT activity. In contrast, UGT2B7\_i8 contains all UGT domains but also an internal 32-residue sequence that potentially disrupts the coordination of co-substrate-to-substrate  $S_N2$  transfer of the glucuronic acid for typical UGT2B7 substrates, such as AZT, as observed in this study. Alternatively, this region may also potentially adapt to a different set of substrates, changing substrate specificity compared with the canonical UGT2B7 enzyme that conjugates a wide variety of lipophilic metabolites and is involved in the clearance of ~25% of common medications (Guillemette et al. 2014, Stingl et al. 2014). A larger set of substrates, as well as co-substrates such as UDP-glucose, that are metabolized by UGT2B7 in addition to UDPGA (Chau et al. 2014), needs to be tested to determine whether the alternative UGT isoforms are truly enzymatically null.

When co-expressed with UGTs in human tissues and cells, alternate UGTs either inhibited or activated cellular glucuronide formation, significantly affecting inactivation of drugs and hormones. To our knowledge, a UGT acting as an inducer/activator of the glucuronidation pathway has not been reported before. Previous investigations by our group and others indicate that UGTs form oligomeric complexes (Operana and Tukey 2007, Ishii et al. 2010, Rouleau et al. 2013). The immunofluorescence evidence of close proximity in the ER membrane and co-IP data strongly suggest that the modulatory effects of the UGT1A and UGT2B7 alternative isoforms are caused by direct protein-protein interactions with their cognate UGT enzymes, and perhaps other UGTs, leading to a mixture of active and inactive complexes (Operana and Tukey 2007, Ishii et al. 2010, Rouleau et al. 2013). Given the enhanced stability of at least some alternative isoforms relative to the canonical UGTs, as shown for UGT1A1\_i2, the impact of alternative isoforms on cellular glucuronidation activity is likely greater than anticipated from mRNA expression levels. We suggest that AS hampers or potentiates canonical UGT functions. This likely helps coordinate and fine-

tune cellular responses to numerous endogenous and exogenous stimuli, with the potential to affect drug responses and disease development and progression, especially tumorigenesis, given the altered ratio of alternative/canonical variant expression and the large interindividual variability in their expression.

Our functional investigations of the UGT2B7\_i8 variant provide further evidence that alternative isoforms have extended metabolic functions that affect global cell metabolism, suggesting that they may be part of multi-enzyme complexes of a core metabolic network. The biological relevance of a metabolic shift induced by expression of UGT2B7\_i8 alone warrants further characterization, but our data also suggest that the glucuronidation pathway and primary cell metabolism (e.g., amino acids and nucleotides) are interrelated enzymatic processes. Whether dimerization/oligomerization and/or glucuronidation of alternative endogenous substrates link the biological activity of these isoforms with cell metabolism remains to be addressed. Furthermore, UGT2B7\_i8 expression is functionally linked to modulation of cell adhesion and proliferation, supporting the conclusion that UGTs substantially impact multiple cellular processes.

## **Conclusions**

This study provides a comprehensive quantitative portrait of hepatic, renal, and gastrointestinal UGT transcriptomes and reveals the abundance of alternate UGT variants, which clearly expand the UGT proteome. A first biological consequence of altering the UGT transcriptome was emphasized by the functional characterization of alternative isoforms with in-frame sequences introduced, which were found to antagonize or induce cellular glucuronidation, thus affecting inactivation of drugs and sex hormones. This study also raises the exciting possibility that different UGTs may exhibit biological functions independent of glucuronic acid transferase activity. Alternate isoform expression induced drastic shifts in multiple metabolic pathways and modified cellular phenotypes, suggesting crosstalk between various enzymatic pathways and UGTs that could serve to maintain cellular homeostasis and increase cell fitness. The broader regulatory and functional relevance of such an extensive splicing program remains to be fully characterized. We

expect that this work will enable the exploration of human UGT-mediated metabolism and its potentially extended role in health and disease.

## **Experimental Procedures**

### **RNA samples, sequencing, and mRNA expression**

UGT expression was characterized in normal tissue samples from 9 liver, 9 kidney, and 15 intestine and colon samples as well as in samples from 9 kidney, and 9 intestine and colon tumor tissues of mixed-gender origin. For each tissue, equivalent amounts of total RNA from at least three samples were pooled, and three pools per tissue type were prepared and used for production of sequencing libraries. UGT sequences were enriched from paired-end bar-coded cDNA libraries with UGT capture (Tourancheau et al. 2016). Libraries were sequenced on an Illumina HiSeq 2500 system (McGill University and Génome Québec Innovation Center, Montreal, QC, Canada). Trimmed reads were aligned with the splice-aware software TopHat2 on the UCSC hg19 reference genome with annotation provided by Illumina iGenome combined with the complete annotated human UGT loci described recently by our group (Tourancheau et al. 2016). The average normalized expression of replicates is provided throughout the text. Detailed library construction, sequence analysis, and validation of *UGT2B7\_n4* expression by reverse transcription-PCR and Sanger sequencing are described in Supplemental Experimental Procedures.

Expression of *UGT1A* and *UGT2B7* variants was examined in 18 normal liver and 18 normal kidney samples in the RNA-Seq dataset from the Pharmacogenomics Research Network (Chhibber et al. 2016) accessed from the GEO accession number GSE70503. Alignment of reads on the complete human UGT transcriptome and quantification using Cufflinks are described in Supplemental Experimental Procedures.

### **Cell models**

HT115 cells expressing endogenous UGT1A isoforms and HEK293 cell models expressing myc- or V5-tagged UGT1A1 isoforms have been described previously (Bellemare et al. 2010, Rouleau et al. 2014). The pool of HEK293 cells stably expressing *UGT2B7\_v1* was

established by supplementing cell culture media with G418 (Invitrogen, 1 mg/ml). The pool of HEK293 cells co-expressing *UGT2B7\_v1* and *UGT2B7\_n4* (corresponding to the exon 2b encoding transcript) was established by subsequent transfection of HEK-2B7\_v1 with the *UGT2B7\_n4* construct (with or without a V5 tag) and selection with blasticidin (Wisent, 10 µg/ml). The HEK293 cell model expressing *UGT2B7\_i1* and *UGT2B7\_i8-V5* was used only to demonstrate protein-protein interactions by IP. HepG2 cells, which express *UGT2B7\_v1* endogenously but not *UGT2B7\_n4*, were transfected with the *UGT2B7\_n4* construct (without tag), and the pool was established with blasticidin selection. Control HEK293 and HepG2 cells were produced by transfection with the parental vectors and selection as above.

### **Analysis of protein expression**

*Antibodies.* *UGT2B7* polyclonal antibody was from ProteinTech Group (16661-1-AP; western blotting (WB): 1:5000; IHC: 1:800; IF: 1:500) and monoclonal antibody was from Abcam (ab57685; IP: 1:100; Abcam, Toronto, ON, Canada; note that ab57685 is sold as an anti-*UGT2B10* antibody but it also recognizes *UGT2B7* (**Fig. S5**)). The purified pan-*UGT1A* polyclonal antibody (RC-71; WB: 1:1000), purified *UGT1A\_i1* polyclonal antibody (#9348; IHC and IF: 1:1000; IP: 1:200), and purified *UGT1A\_i2* antibody (#4863; IHC: 1:1000) have been described (Albert et al. 1999, Bellemare et al. 2011). *UGT1A\_i2* (#4C5E7; IF: 1:100) and *UGT2B7\_i8* (IF and IHC: 1:1000) antibodies were custom made (Genscript, Piscataway, NJ, USA). Anti-V5 (WB: 1:5000) and anti-myc (clone 4A6, WB: 1:5000) were from Invitrogen (Life Technologies, Burlington, Canada) and EMD Millipore (Etobicoke, ON, Canada), respectively. Cell compartment-specific antibodies used in IF were: anti-58K Golgi protein (1:100; ab27043; Abcam), anti-protein disulfide isomerase (PDI: 1:100; ab2792; Abcam) for the ER, anti-cytochrome c (1:300; cat #12963S; Cell Signaling Technology) for the mitochondria; DNA was stained with Hoechst 33342 (1:1500; Sigma) or DRAQ5 (1:2000; Life Technologies Inc.). Production of antibodies against variants is detailed in Supplemental Experimental Procedures.

*mRNA and protein stability of UGT1A and UGT2B7.* Cells were treated with 20 µg/ml cycloheximide for 0–16 h and collected at various time points as described (Turgeon et al.

2001). Cell homogenates (20 µg protein) prepared in phosphate-buffered saline containing 0.5 mM dithiothreitol were analyzed by WB with RC-71 (UGT1A) or anti-UGT2B7 from ProteinTech.

*IF.* Detection of endogenous UGT1A\_i1 and \_i2 in HT115 cells was as described (Bellemare et al. 2010). Stably expressed UGT2B7\_i1 and UGT2B7\_i8 in the HEK cell models were detected with anti-2B7 16661-1-AP.

*IHC.* Paraffin-embedded tissue blocks were available for individual patients. The institutional board approved the study, and written consent was given by all patients concerning the use of their tissues for research purposes. Serial sections (5 µm) were deparaffinized, rehydrated, and processed using the IDetect SuperStain HRP polymer kit (Empire Genomics, Buffalo, NY, USA) with an overnight incubation of tissues at 4°C with primary antibodies. Control sections were incubated with the anti-UGT2B7\_i8 preadsorbed for 3 h at room temperature with an excess of the immunogenic peptide (2 µM) to ascertain labeling specificity.

*IP.* HEK293 UGT1A-tagged cell models were grown to confluency in 10-cm culture dishes, washed, scraped into 800 µl lysis buffer/dish, and immunoprecipitated as described in Supplemental Experimental Procedures using anti-UGT1A\_i1 or control rabbit IgG. For UGT2B7\_i8 IP, HEK293 cells grown in two 15-cm dishes were scraped and cross-linked with 0.125% formaldehyde for 10 min at 37°C then quenched with 0.125 M glycine. Cells were collected by centrifugation prior to lysis in 1 ml lysis buffer and IP as described in Supplemental Experimental Procedures.

*MS-coupled multiple reaction monitoring.* Tryptic digests of UGT2B7 immunoprecipitated from human liver S9 fractions were analyzed by MS-coupled multiple reaction monitoring on a 6500QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer (Sciex, Concord, ON, Canada). The UGT2B7 signature peptide ADVWLIR and the UGT2B7\_i8-specific peptide LDSFNTGWINK were detected in tryptic digests of the immunoprecipitated UGT2B7 samples, and peptide identity was confirmed by co-injection of isotopically labeled [<sup>13</sup>C6,<sup>15</sup>N2]Lys and [<sup>13</sup>C6,<sup>15</sup>N4]Arg synthetic peptides.

*Cell adhesion and proliferation.* HEK293 and HepG2 cells (pool of control cells or cells stably expressing *UGT2B7\_n4*) plated in E-view PET 16-well plates (10,000 cells/well, four replicates/experiment, n = 3) were monitored in real time on an Xcelligence DP system (ACEA Biosciences Inc., San Diego, CA, USA). Doubling time was determined with the RTCA software 2.0 (ACEA Biosciences Inc.) using normalized cell index values between the 20 and 45 h time points. Alternatively, HEK293 and HepG2 cells were plated in 6-well plates (125,000 HEK cells/well; 100,000 HepG2 cells/well), incubated for specified times, and counted with a TC-10 automated cell counter (Bio-Rad) at 48, 72 and 96 h after plating. Cell media were replaced after the 48 h time point.

### **Metabolic functions**

*In situ glucuronidation assays.* HEK293 and HepG2 cells (control cells or cells stably expressing *UGT2B7\_v1*, *UGT2B7\_n4* or both, as stated in the figure legends) were seeded in 24-well plates (HEK293: 80,000 cells/well, HepG2: 125,000 cells/well). Assays were initiated 72 h after seeding by replacing the culture medium with fresh medium (1 ml/well) containing a labeled UGT substrate probe (for UGT1A1: 5  $\mu$ M SN-38 prepared by hydrolysis of irinotecan-HCl (McKesson, ON, Canada) and/or 25  $\mu$ M estradiol (Steraloids, Newport, RI, USA); for UGT2B7: AZT (Sigma-Aldrich, St. Louis, MO, USA)). Cells were incubated for 4 h, and the media were collected and stored at  $-20^{\circ}\text{C}$  until glucuronide assessment by MS-based analysis as described (Lepine et al. 2004, Belanger et al. 2009). Briefly, for AZT-G, the culture media were diluted 1:3 with 25% aqueous methanol solution containing AZT-d3-G standard (Toronto Research Chemicals, Toronto, ON, Canada) prior to analysis as described (Belanger et al. 2009).

*Untargeted metabolomics analysis.* HEK293 and HepG2 cells (stable pool of control cells or cells expressing *UGT2B7\_n4*) were plated in 10-cm culture dishes (1.5 million HEK cells/dish; 3.5 million HepG2 cells/dish) and grown for 96 h, with a change to fresh medium after the first 48 h. Cells were harvested by trypsinization, counted, rinsed twice in ice-cold phosphate-buffered saline, snap-frozen on dry ice, and stored at  $-80^{\circ}\text{C}$  until extraction for metabolomics analysis. Untargeted global metabolite profiling was conducted



on triplicate samples at the West Coast Metabolomics Center (University of California at Davis) as described (Fiehn et al. 2008). Relative quantification data provided as normalized peak heights were further normalized for cell counts in each sample and expressed as fold change ratios of mean metabolite levels in the cells expressing variant UGT2B7 versus control cells. Details about the data analysis are provided in Supplemental Experimental Procedures.

## **Competing Interests**

The authors have no conflicts of interest to disclose.

## **Author contributions**

Conceptualization: MR, CG; Methodology: MR, CG; Investigation: MR, AT, CGB, LV, AMD, YAD, IG, AD, CG; Validation: MR, AT, CGB, LV, JV, IP, AD, CG; Writing – Original Draft: MR, CG; Writing – Review & Editing: MR, AT, CGB, LV, JV, AMD, YAD, IG, IP, AD, CG; Supervision: MR, CG; Project administration: CG; Funding Acquisition: CG

## **Accession numbers**

RNA sequencing data are available in the GEO database (GSE80463)

## **Acknowledgements**

We acknowledge Patrick Caron and Véronique Turcotte for their contribution to the mass spectrometry analysis of glucuronide metabolites, Sylvie Desjardins for contributing to IP experiments, Andréa Fournier for contributing to the adhesion and proliferation

experiments, Joannie Roberge and Hugo Girard for contributing to the protein stability experiments, Sandra Guauque-Olarte for GEO submission of RNA sequencing data, Johanne Ouellet and Michèle Orain for contributing to the IHC analysis, Isabelle Kelly from the Proteomics platform at the CHU de Québec Research Center for support with the multiple reaction monitoring analysis, France Couture for artwork, and Dr. Éric Lévesque for helpful discussion.

This work was financially supported by grants from the Canadian Institutes for Health Research (MOP-42392 and MOP-142318), the Natural Sciences and Engineering Research Council of Canada (342176-2012), and a Canadian Research Chair in Pharmacogenomics to CG (Tier I). CGB holds an FER studentship award from the Faculty of Pharmacy of Laval University. YAD received studentship funding from the Centre de Recherche en Endocrinologie Moléculaire et Oncologique et Génomique Humaine, Laval University and Fonds de Recherche du Québec – Santé.

## References

Albert, C., Vallee, M., Beaudry, G., Belanger, A., and Hum, D.W. (1999). The monkey and human uridine diphosphate-glucuronosyltransferase UGT1A9, expressed in steroid target tissues, are estrogen-conjugating enzymes. *Endocrinology* 140, 3292-3302.

Belanger, A.S., Caron, P., Harvey, M., Zimmerman, P.A., Mehlotra, R.K., and Guillemette, C. (2009). Glucuronidation of the antiretroviral drug efavirenz by UGT2B7 and an in vitro investigation of drug-drug interaction with zidovudine. *Drug Metab Dispos* 37, 1793-1796.

Bellemare, J., Rouleau, M., Harvey, M., and Guillemette, C. (2010a). Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *J Biol Chem* 285, 3600-3607.

Bellemare, J., Rouleau, M., Harvey, M., Popa, I., Pelletier, G., Tetu, B., and Guillemette, C. (2011). Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *J Pathol* 223, 425-435.

Bellemare, J., Rouleau, M., Harvey, M., Tetu, B., and Guillemette, C. (2010b). Alternative-splicing forms of the major phase II conjugating UGT1A gene negatively regulate glucuronidation in human carcinoma cell lines. *Pharmacogenomics J* 10, 431-441.

Bock, K.W. (2015). Roles of human UDP-glucuronosyltransferases in clearance and homeostasis of endogenous substrates, and functional implications. *Biochem Pharmacol* 96, 77-82.

Breton, C., Fournel-Gigleux, S., and Palcic, M.M. (2012). Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol* 22, 540-549.

Chau, N., Elliot, D.J., Lewis, B.C., Burns, K., Johnston, M.R., Mackenzie, P.I., and Miners, J.O. (2014). Morphine glucuronidation and glucosidation represent complementary metabolic pathways that are both catalyzed by UDP-glucuronosyltransferase 2B7: kinetic, inhibition, and molecular modeling studies. *J Pharmacol Exp Ther* 349, 126-137.

Chhibber, A., French, C.E., Yee, S.W., Gamazon, E.R., Theusch, E., Qin, X., Webb, A., Papp, A.C., Wang, A., Simmons, C.Q., *et al.* (2016). Transcriptomic variation of

pharmacogenes in multiple human tissues and lymphoblastoid cell lines. *Pharmacogenomics J. On line*: February 9, 2016; doi:10.1038/tpj.2015.93

Clark, M.B., Mercer, T.R., Bussotti, G., Leonardi, T., Haynes, K.R., Crawford, J., Brunck, M.E., Cao, K.A., Thomas, G.P., Chen, W.Y., *et al.* (2015). Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. *Nat Methods* 12, 339-342.

Court, M.H., Zhang, X., Ding, X., Yee, K.K., Hesse, L.M., and Finel, M. (2012). Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica* 42, 266-277.

Dates, C.R., Fahmi, T., Pyrek, S.J., Yao-Borengasser, A., Borowa-Mazgaj, B., Bratton, S.M., Kadlubar, S.A., Mackenzie, P.I., Haun, R.S., and Radominska-Pandya, A. (2015). Human UDP-Glucuronosyltransferases: Effects of altered expression in breast and pancreatic cancer cell lines. *Cancer Biol Ther* 16, 714-723.

De Bruyn, F., Maertens, J., Beauprez, J., Soetaert, W., and De Mey, M. (2015). Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnol Adv* 33, 288-302.

Fiehn, O., Wohlgemuth, G., Scholz, M., Kind, T., Lee do, Y., Lu, Y., Moon, S., and Nikolau, B. (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* 53, 691-704.

Gamazon, E.R., and Stranger, B.E. (2014). Genomics of alternative splicing: evolution, development and pathophysiology. *Hum Genet* 133, 679-687.

Garcia-Blanco, M.A., Baraniak, A.P., and Lasda, E.L. (2004). Alternative splicing in disease and therapy. *Nat Biotechnol* 22, 535-546.

Guillemette, C. (2003). Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *Pharmacogenomics J* 3, 136-158.

Guillemette, C., Belanger, A., and Lepine, J. (2004). Metabolic inactivation of estrogens in breast tissue by UDP-glucuronosyltransferase enzymes: an overview. *Breast Cancer Res* 6, 246-254.

- Guillemette, C., Levesque, E., and Rouleau, M. (2014). Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* 96, 324-339.
- Hu, D.G., Meech, R., McKinnon, R.A., and Mackenzie, P.I. (2014). Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab Rev* 46, 421-458.
- Ishii, Y., Takeda, S., and Yamada, H. (2010). Modulation of UDP-glucuronosyltransferase activity by protein-protein association. *Drug Metab Rev* 42, 145-158.
- Lepine, J., Bernard, O., Plante, M., Tetu, B., Pelletier, G., Labrie, F., Belanger, A., and Guillemette, C. (2004). Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* 89, 5222-5232.
- Little, J.M., Kurkela, M., Sonka, J., Jantti, S., Ketola, R., Bratton, S., Finel, M., and Radomska-Pandya, A. (2004). Glucuronidation of oxidized fatty acids and prostaglandins B1 and E2 by human hepatic and recombinant UDP-glucuronosyltransferases. *J Lipid Res* 45, 1694-1703.
- Lo, W.S., Gardiner, E., Xu, Z., Lau, C.F., Wang, F., Zhou, J.J., Mendlein, J.D., Nangle, L.A., Chiang, K.P., Yang, X.L., *et al.* (2014). Human tRNA synthetase catalytic nulls with diverse functions. *Science* 345, 328-332.
- Mackenzie, P.I., Bock, K.W., Burchell, B., Guillemette, C., Ikushiro, S., Iyanagi, T., Miners, J.O., Owens, I.S., and Nebert, D.W. (2005). Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics* 15, 677-685.
- Margaillan, G., Rouleau, M., Fallon, J.K., Caron, P., Villeneuve, L., Turcotte, V., Smith, P.C., Joy, M.S., and Guillemette, C. (2015a). Quantitative profiling of human renal UDP-glucuronosyltransferases and glucuronidation activity: a comparison of normal and tumoral kidney tissues. *Drug Metab Dispos* 43, 611-619.

Margaillan, G., Rouleau, M., Klein, K., Fallon, J.K., Caron, P., Villeneuve, L., Smith, P.C., Zanger, U.M., and Guillemette, C. (2015b). Multiplexed Targeted Quantitative Proteomics Predicts Hepatic Glucuronidation Potential. *Drug Metab Dispos* 43, 1331-1335.

Mercer, T.R., Gerhardt, D.J., Dinger, M.E., Crawford, J., Trapnell, C., Jeddloh, J.A., Mattick, J.S., and Rinn, J.L. (2012). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol* 30, 99-104.

Oltean, S., and Bates, D.O. (2014). Hallmarks of alternative splicing in cancer. *Oncogene* 33, 5311-5318.

Operana, T.N., and Tukey, R.H. (2007). Oligomerization of the UDP-glucuronosyltransferase 1A proteins: homo- and heterodimerization analysis by fluorescence resonance energy transfer and co-immunoprecipitation. *J Biol Chem* 282, 4821-4829.

Popp, M.W., and Maquat, L.E. (2013). Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu Rev Genet* 47, 139-165.

Radomska-Pandya, A., Czernik, P.J., Little, J.M., Battaglia, E., and Mackenzie, P.I. (1999). Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* 31, 817-899.

Rouleau, M., Collin, P., Bellemare, J., Harvey, M., and Guillemette, C. (2013). Protein-protein interactions between the bilirubin-conjugating UDP-glucuronosyltransferase UGT1A1 and its shorter isoform 2 regulatory partner derived from alternative splicing. *Biochem J* 450, 107-114.

Rouleau, M., Roberge, J., Bellemare, J., and Guillemette, C. (2014). Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Mol Pharmacol* 85, 29-36.

Rowland, A., Miners, J.O., and Mackenzie, P.I. (2013). The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol* 45, 1121-1132.

- Stingl, J.C., Bartels, H., Viviani, R., Lehmann, M.L., and Brockmoller, J. (2014). Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther* *141*, 92-116.
- Tourancheau, A., Margaillan, G., Rouleau, M., Gilbert, I., Villeneuve, L., Levesque, E., Droit, A., and Guillemette, C. (2016). Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J* *16*, 60-70.
- Treutlein, B., Gokce, O., Quake, S.R., and Sudhof, T.C. (2014). Cartography of neurexin alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc Natl Acad Sci U S A* *111*, E1291-1299.
- Turgeon, D., Carrier, J.S., Levesque, E., Hum, D.W., and Belanger, A. (2001). Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. *Endocrinology* *142*, 778-787.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* *456*, 470-476.
- Wells, P.G., Mackenzie, P.I., Chowdhury, J.R., Guillemette, C., Gregory, P.A., Ishii, Y., Hansen, A.J., Kessler, F.K., Kim, P.M., Chowdhury, N.R., *et al.* (2004). Glucuronidation and the UDP-glucuronosyltransferases in health and disease. *Drug Metab Dispos* *32*, 281-290.

## Figure Legends

**Figure 1: Profiling the UGT transcriptome and characterization of UGT isoforms: an overview of the experimental design.** The UGT transcriptome in the main drug metabolizing organs ( $\geq 9$  samples per tissue) was quantified by RNA-Seq after a targeted capture step. The complete human UGT sequence annotation (Tourancheau et al. 2016) enabled precise mapping of reads to canonical and alternative transcripts. Expression and functional analysis of mRNA variants and predicted protein isoforms were conducted by multiple approaches. MS: mass spectrometry; IF: immunofluorescence; IHC: immunohistochemistry.

**Figure 2: Tissue-specific expression of canonical and alternative *UGT* transcripts.** Venn diagrams showing the large diversity in expression patterns between tissues (**A**) and between healthy and tumor tissues (**B**). Total number of canonical (v1) UGTs and alternative variants (alt) are given for each tissue. Values in the Venn diagrams represent number of alternative transcripts (alt) expressed above 600 fragments per kilobase of transcript per million mapped reads in CaptureSeq. **C**. Average quantitative expression of alternative transcripts in healthy and tumor tissues. Proportion of alternative transcripts represents total expression of alternative transcripts relative to total v1 expression in each tissue. See also Fig. S1.

**Figure 3: Structural organization of UGT enzymes and alternative isoforms.** **A**. Schematic of functional domains and structural organization of UGT enzymes in the endoplasmic reticulum (ER). The bulk of the substrate-binding (red) and co-substrate-binding (blue) domains lie on the luminal side of the ER. The transmembrane domain (grey segment) positions the positively charged tail (orange, ++ ) on the cytoplasmic side of the ER. **B**. Quantification of alternative UGT levels classified on the basis of truncated regions and inclusion of sequences predicted from alternative transcript expression levels. N-term: N-terminal; C-term: C-terminal.



**Figure 4. Functional characterization of alternate UGT1A<sub>i2</sub>.** **A.** The nine well-characterized UGT1A<sub>v1</sub> transcripts are produced from splicing of alternative exon 1s to the common exons 2–5a (blue lines) at the single *UGT1* locus and encode the nine UGT1A<sub>i1</sub> enzymes. The *UGT1A<sub>v2</sub>* and *UGT1A<sub>v3</sub>* transcripts (nine of each type) arise from the use of alternative exon 5b, without or with exon 5a, respectively (green lines). *UGT1A<sub>v2</sub>* and *\_v3* transcripts encode the same nine UGT1A<sub>i2</sub> proteins, because exon 5b includes a stop codon. UGT1A<sub>i2</sub> and UGT1A<sub>i1</sub> are distinguishable by their divergent C-termini. The UGT1A<sub>i2</sub>-specific antibody is directed towards the unique i2 amino acid sequence (green box). Domains are organized as in Fig. 3A. **B.** UGT1A<sub>i2</sub>s have a longer protein half-life compared with UGT1A<sub>i1</sub>. *Top panel:* Protein levels of UGT1A<sub>i1</sub>, UGT1A<sub>i2</sub> or both in the HEK293 cell models were measured at different times after treatment with cycloheximide using a pan-UGT1A antibody recognizing both isoforms. *Lower panel:* Densitometric quantification of UGT1A protein level is expressed as a function of time after cycloheximide treatment. Data are presented as means + SD. **C.** The HEK293 cell models (UGT negative) were developed to express UGT1A<sub>i1</sub>, UGT1A<sub>i2</sub>, or both isoforms. *In situ* cell assays demonstrate that glucuronidation of SN-38 and estradiol (E<sub>2</sub>) by UGT1A<sub>i1</sub> is impaired by the co-expression of UGT1A<sub>i2</sub> (n = 2 independent assays in triplicate). UGT1A<sub>i2</sub> did not catalyze glucuronidation of these substrates. Activity was normalized to UGT1A<sub>i1</sub> level. Data are presented as means + SD, \*\*\**P*<0.001, Student's *t*-test. **D.** Interactions between UGT1A<sub>i1</sub> and i2. IP of i1 from HEK cell lysates with the specific anti-i1 co-purifies UGT1A<sub>i2</sub>. Lanes “ctr” served as markers for i1 and i2 protein mobility and detection. **E.** Endogenous UGT1A<sub>i1</sub> and i2 co-localize in the endoplasmic reticulum (ER) of colorectal cancer HT115 cells. Protein disulfide isomerase (PDI) is an ER marker, whereas Hoechst (H) labels nuclei. Fluorescence intensity profiles are given for cross-sections (dashed lines). See also Fig. S2 for a comparison of co-localization with multiple specific subcellular markers including the ER. Scale bars: 10 μm. **F.** Immunohistochemical investigation of UGT1A<sub>i1</sub> and I2 in consecutive sections of normal human liver and colon tissues. H: hepatocytes; PV: portal

vein; IG: intestinal gland; L: lumen; LP: lamina propria; SE: surface epithelium. Scale bars: 200  $\mu\text{m}$ .

**Figure 5. Functional characterization of UGT2B7\_i8.** **A.** Alternative splicing at the *UGT2B7* locus. The canonical *UGT2B7\_v1* transcript excludes the exon 2b (blue lines) and encodes the UGT2B7\_i1 enzyme, whereas *UGT2B7\_n4* includes exon 2b (green region, line and boxed sequence) that extends the coding sequence to produce UGT2B7\_i8. Anti-2B7\_i8 was raised against the underlined immunogenic peptide. **B.** Immunohistochemical investigation of i8 expression in normal liver tissues. *Left:* UGT2B7\_i1, labeled by the commercial antibody against UGT2B7, is expressed predominantly in hepatocytes but not in hepatic arteries or bile ducts. *Center:* The purified anti-i8 specifically and strongly labels smooth muscle cells of hepatic arteries whereas hepatocytes are weakly labeled and bile ducts are unlabeled. *Right:* Labeling is impaired by preadsorption of anti-2B7\_i8 with the immunogenic peptide prior to tissue labeling, demonstrating specificity of staining with anti-2B7\_i8. Representative images from consecutive sections of two liver tissues are shown. BD: bile ducts; H: hepatocytes; HA: hepatic arteries; PV: portal vein. Scale bars: 200  $\mu\text{m}$ . **C.** Metabolic alterations associated with the expression of UGT2B7\_i8. HEK293 kidney and HepG2 liver cell models (expressing endogenous UGTs, including UGT2B7) were developed. UGT2B7\_i8 has an apparent molecular mass of 62 kDa, slightly larger than that the UGT2B7\_i1 enzyme (55 kDa), as expected, upon analysis of the microsomal fractions (20  $\mu\text{g}$ ) from each cell model. L: Human liver microsomes, as a positive control. *In situ* glucuronidation activity towards the UGT2B7-specific substrate zidovudine (AZT) was enhanced by co-expression of UGT2B7\_i1 and the i8 isoform in both cell models (HEK293, n = 4 assays; HepG2, n = 3 assays, in triplicate). UGT2B7\_i8 did not have glucuronidation activity with AZT as a substrate. Activity was normalized to UGT2B7\_i1 level. Data are presented as means + SD, \*\*\* $P < 0.001$ , Student's *t*-test. **D.** Interactions between UGT2B7\_i8 and i1. IP of i8 from cell lysates of HEK293 cells co-expressing UGT2B7\_i1 and tagged UGT2B7\_i8-V5 with anti-V5 co-purifies UGT2B7\_i1. **E.** UGT2B7\_i1 and i8 co-localize in the endoplasmic reticulum (ER). Each UGT2B7 variant protein stably expressed in HEK293 cells co-localizes with the ER marker protein disulfide

isomerase (PDI). Fluorescence intensity profiles are given for cross-sections (dashed lines). Scale bars: 10  $\mu\text{m}$ . **F.** Cellular metabolic profiles reveal an important modulation of levels of amino acids and nucleotides induced by UGT2B7\_i8 expression in HEK293 and HepG2 cell models. Data points for which the coefficient of variation exceeded 50% were omitted for clarity. Detailed metabolomics data are given in **Table S1**. **G.** UGT2B7\_i8 enhanced adhesion but reduced the proliferation rate of HEK293 cells. *Left* Real-time cell adhesion was measured by electrical impedance (cell index) (n=2 independent assays in quadruplicate; significantly different for all time points). *Right* Proliferation was assessed by cell counts (n=2 independent assays in triplicate) and doubling time was determined from real time cell index data. Data are presented as means + SD, \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , Student's *t*-test. See also Fig. S3 and S4.

# Figures

Figure 1.

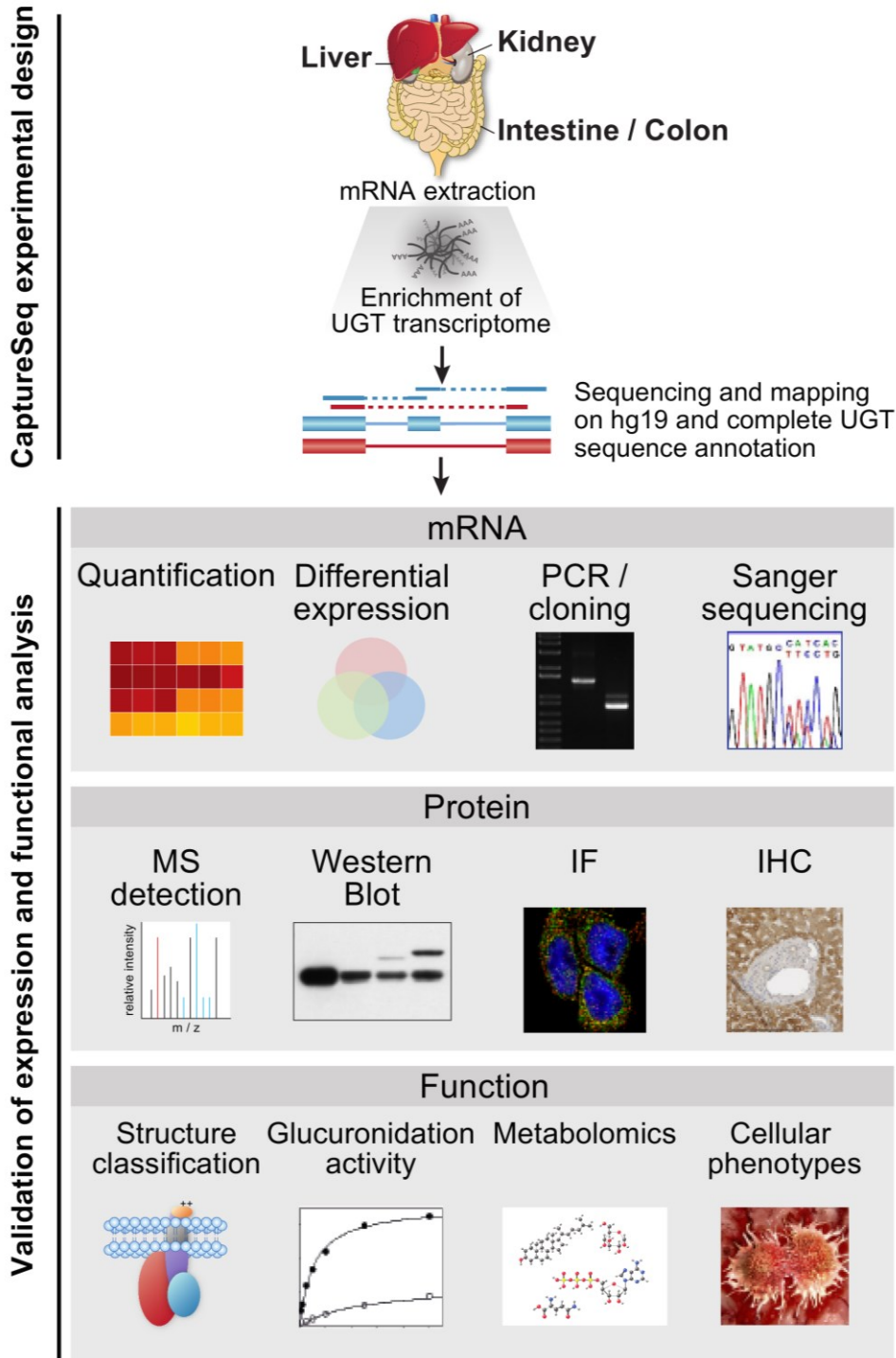
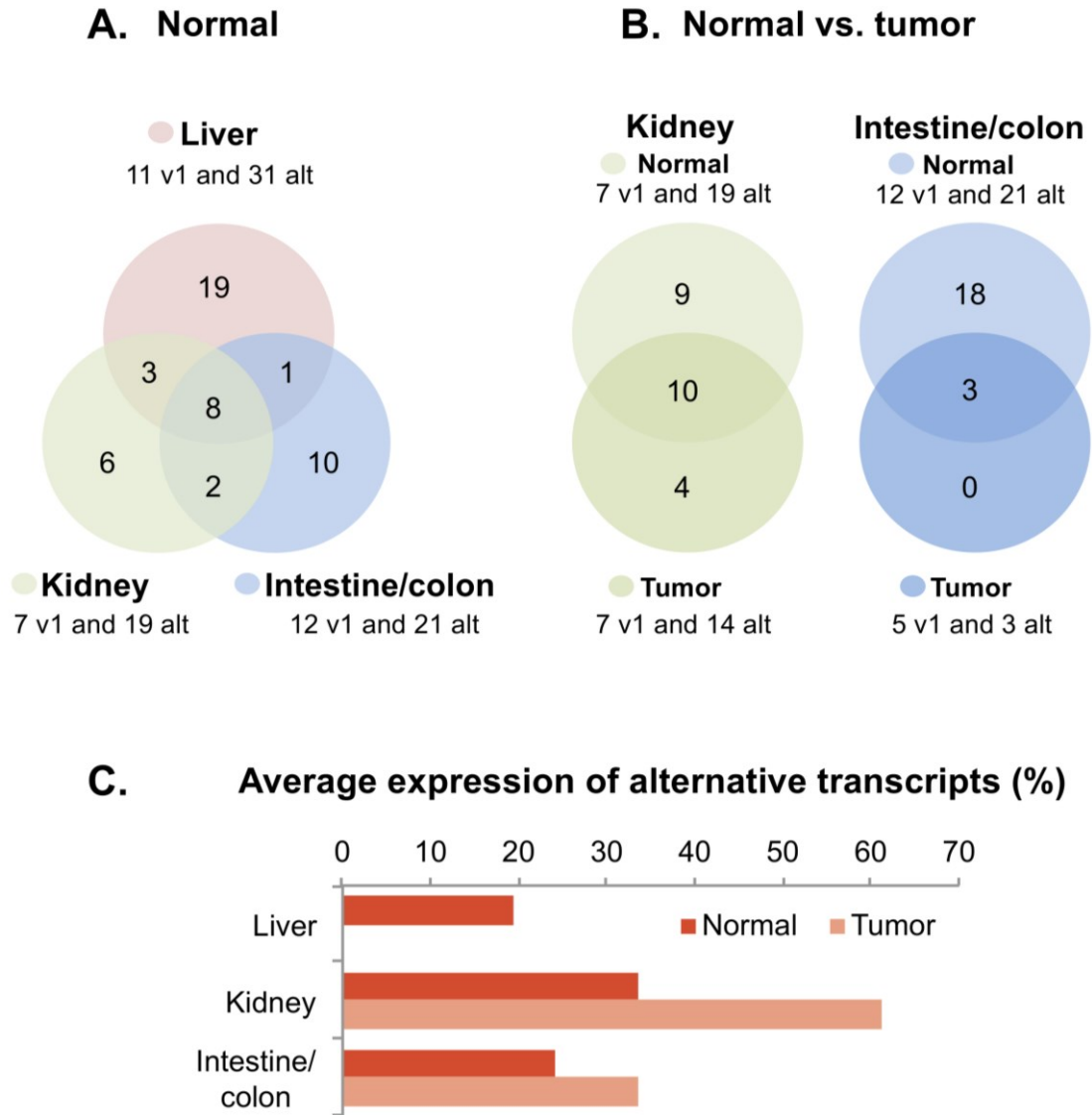
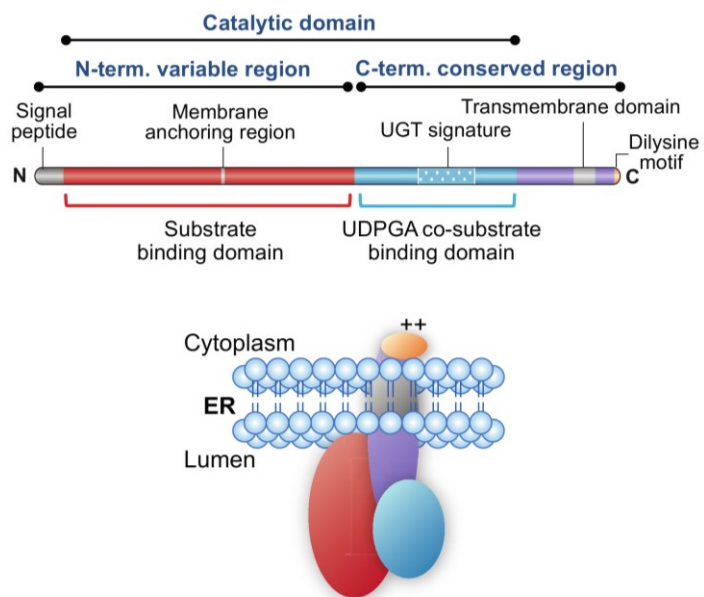


Figure 2.

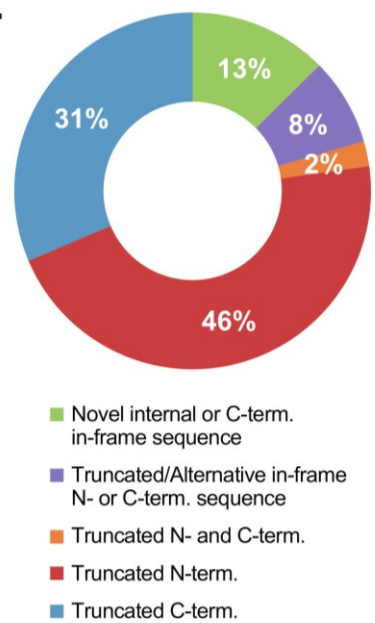


**Figure 3.**

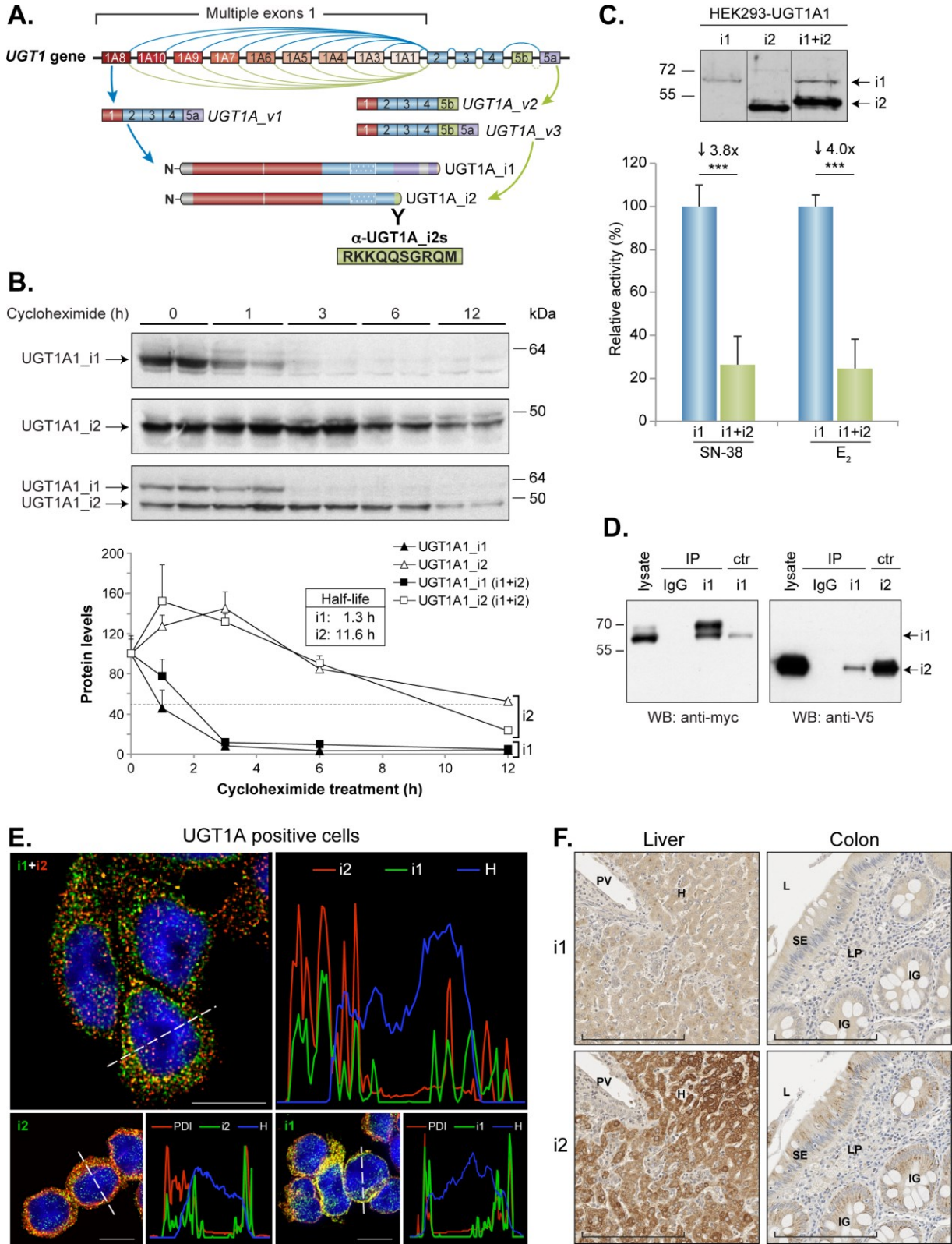
**A.**



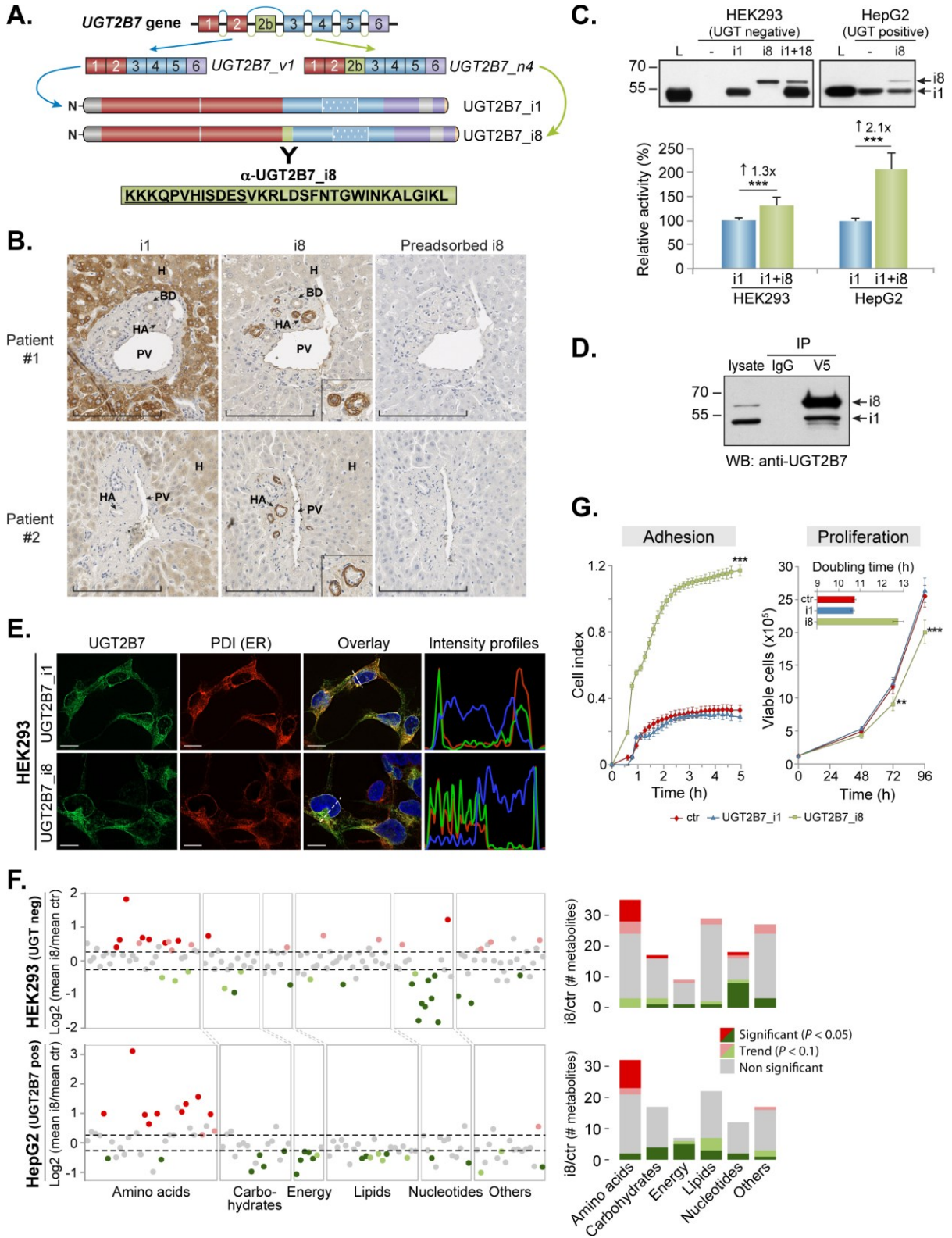
**B.**



**Figure 4.**



**Figure 5.**





## Supplemental Figure Legends

**Figure S1:** Expression of *UGT* alternative variants in human liver and kidney tissues established by RNA-Seq, Related to Figure 2. A. Average expression of *UGT1A\_v2/v3* variants established by CaptureSeq from three pools of at least three tissue samples (current study). Abundance of *v2/v3* relative to the canonical variant is indicated within bars. B. Expression of *UGT1A\_v2/v3* variants relative to canonical *v1* in liver and kidney tissues from 18 individuals in the PGRN dataset. One liver sample has a relative expression >1000% and is not shown. C. Average expression of *UGT2B7\_n4-n6* variants established by CaptureSeq from three pools of at least three tissue samples (current study). Abundances of *n4-n6* relative to the canonical variant are indicated within bars. D. Expression of *UGT2B7\_n4-n6* variants relative to canonical *v1* in liver and kidney tissues from 18 individuals in the PGRN dataset. One kidney sample has a relative expression >1000% and is not shown.

**Figure S2:** Subcellular distribution of endogenous *UGT1A\_i1* and *UGT1A\_i2*, Related to Figure 4. Endogenous *UGT1A\_i1* and *i2* were labeled with their respective specific antibody, and cells were co-labeled with antibodies specific to subcellular compartment protein markers described in Experimental Procedures. Fluorescence intensity profiles are given for cross sections (dashed lines). Scale bars: 10  $\mu\text{m}$ .

**Figure S3:** Validation of *UGT2B7\_i8* encoding transcript expression, Related to Figure 5. A. PCR detection of *UGT2B7\_n4* transcripts (encoding *UGT2B7\_i8*) in kidney tissues with schematics of PCR strategy and products. B. The high stability of *UGT2B7\_i1* enzyme is not influenced by *UGT2B7\_i8*. *UGT2B7* protein levels in each HEK293 cell model treated for the indicated times with cycloheximide were measured with anti-*UGT2B7* that detects both isoforms. Calnexin was used as a loading control. Protein levels were assessed by densitometric scanning and normalized to calnexin levels.

**Figure S4:** Detection of UGT2B7\_i8 in human liver samples, Related to Figure 5. A. Experimental scheme for the detection of the UGT2B7\_i8 unique peptide by multiple reaction monitoring (MRM). B. The common UGT2B7 peptide ADVWLIR (left) and the i8-specific peptide LDSFNTGWINK (right) were detected in tryptic digests of UGT2B7 immunoprecipitated from human liver samples (upper chromatograms). Identification of peptides by MS was confirmed by the perfect matching of retention times and selected MRM transitions with those of stable isotope-labeled peptides mixed with the immunopurified UGT2B7 (lower chromatograms). Representative chromatograms are shown (n = 2).

**Figure S5:** Specificity of the commercial antibody ab57685 for human UGT2B proteins, Related to Experimental Procedures and Figure S4. Microsomal fractions (20 µg) from HEK293 cell models expressing each of the specified UGT2B proteins were analyzed by western blotting with ab57685 (1:5000). Ab57685 is not specific to UGT2B10, as it also efficiently recognizes UGT2B4, UGT2B7, UGT2B11, and UGT2B28. HEK293 control cells transfected with an empty vector (-) and a human liver microsomal fraction (L; 1 µg) were used as negative and positive controls for immunodetection, respectively. The line within the immunoblot denotes removal of an empty lane.

## Supplemental Tables

**Table S1:** Detailed metabolomics data related to Figure 5F.

# Supplemental Figures

Figure S1:

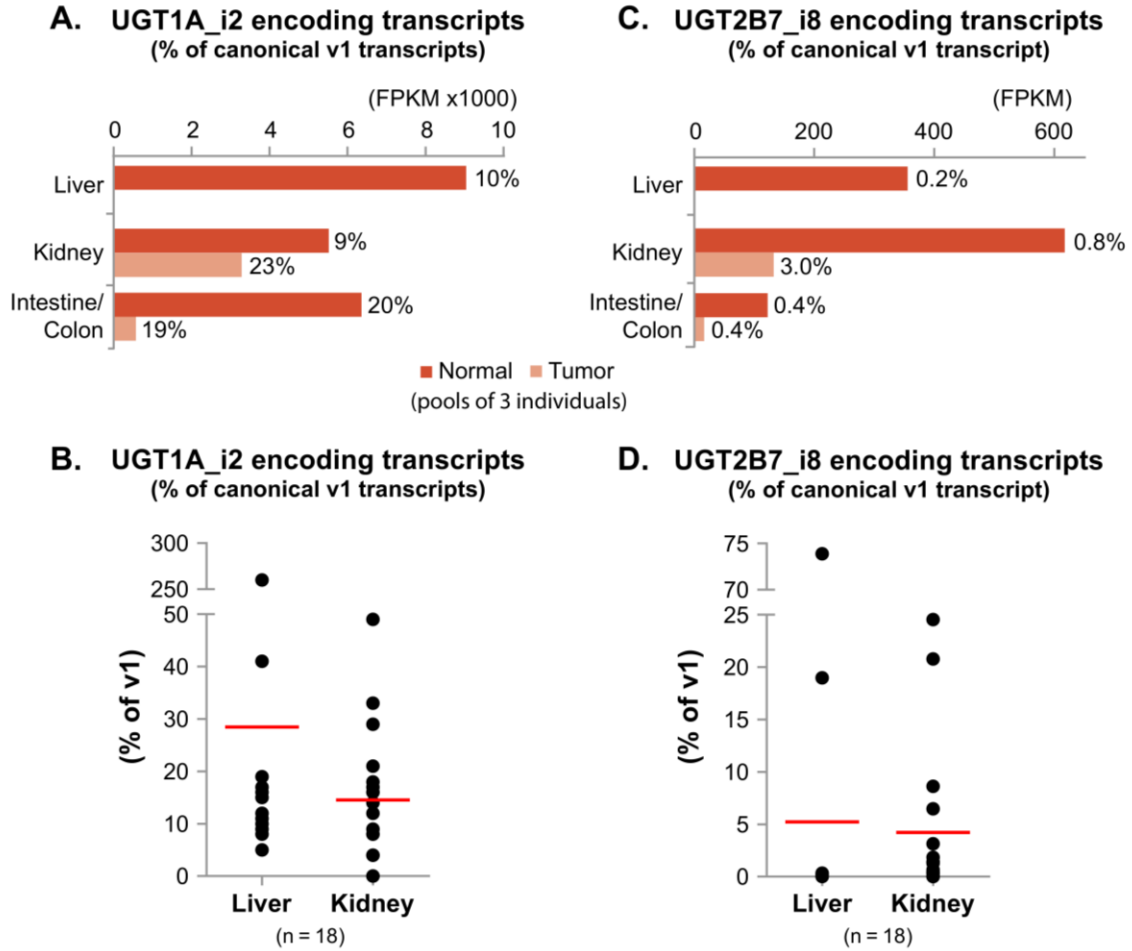
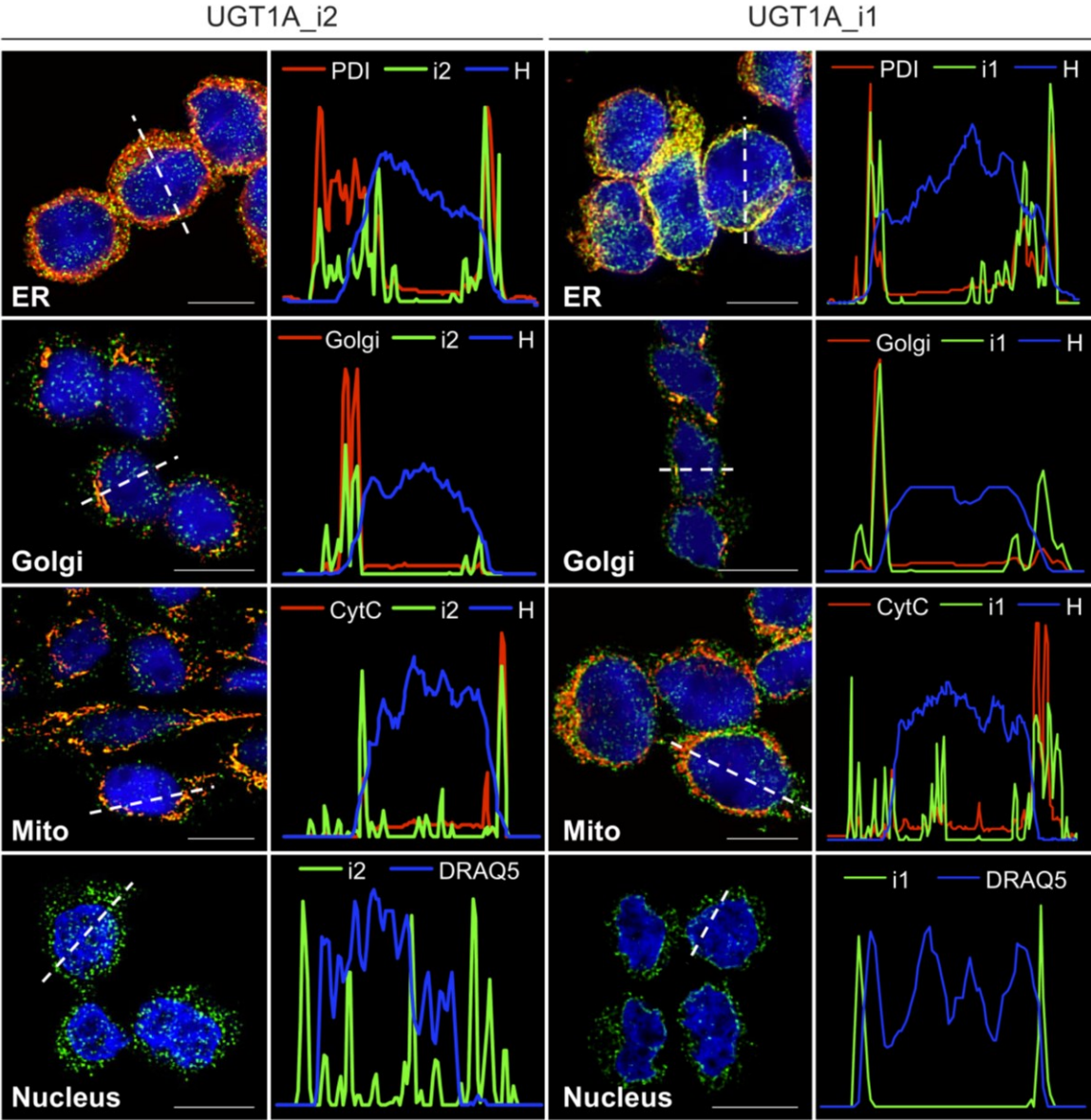
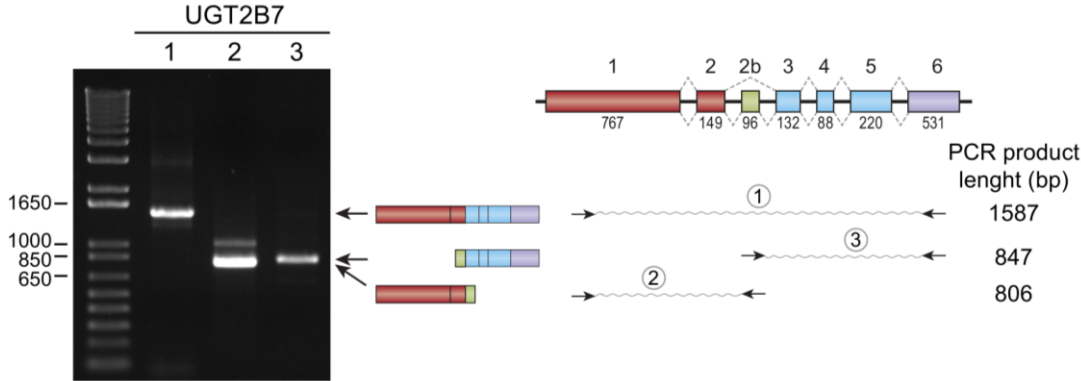


Figure S2:

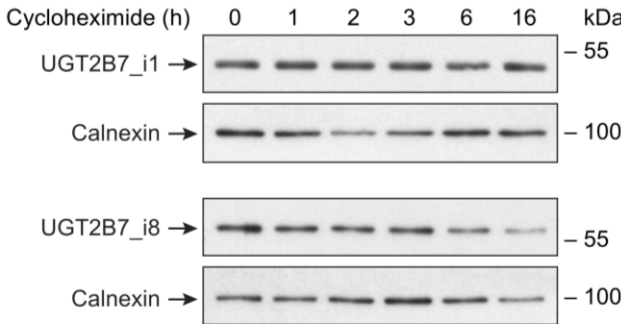


**Figure S3:**

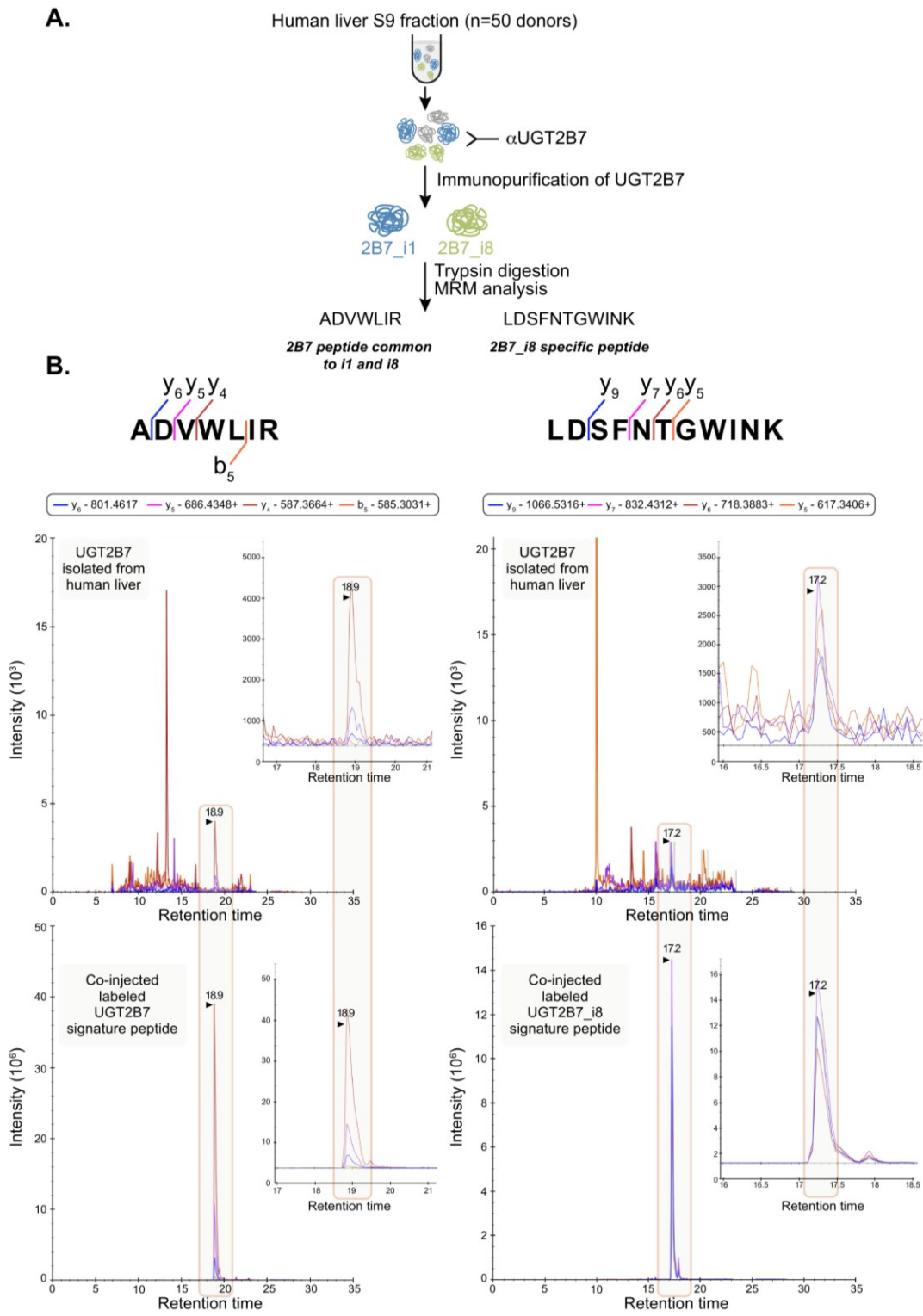
**A.**



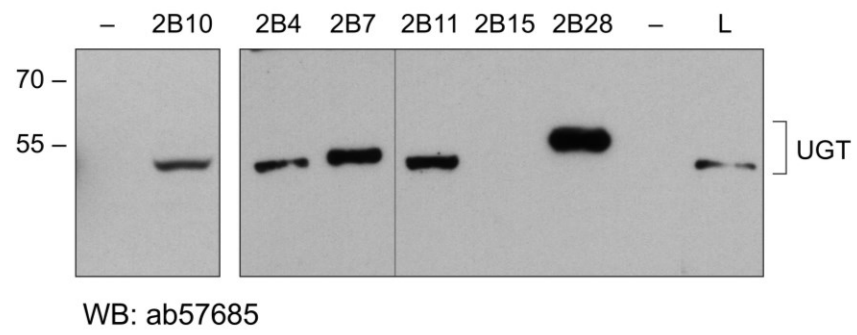
**B.**



**Figure S4:**



**Figure S5:**



## Supplemental Experimental Procedures

### RNA samples, sequencing libraries and processing of sequencing data

Total RNA samples were purchased from Agilent Technologies (Santa Clara, CA, USA), Clontech Laboratories (Mountain View, CA, USA), and Origene (Rockville, MD, USA), or they were available through previous studies (Menard et al. 2013, Tourancheau et al. 2016). Participants provided written consent, and the local research ethics committee approved the research. Total RNA from patient tissues was prepared using Tri Reagent (Sigma-Aldrich, Oakville, ON, Canada) and treated with the RNeasy MinElute Cleanup kit with DNase I (Qiagen, Toronto, ON, Canada). RNA concentrations were measured with a Nanodrop ND-1000 (ThermoFisher Scientific, Ottawa, ON, Canada), and RNA quality was assessed using a 2100 Bioanalyzer (Agilent Technologies). Samples used had an RNA integrity number (RIN) between 7.2 and 9.7. Pooled RNA samples were spiked with the External RNA Controls Consortium (ERCC) standard RNA Spike-In Mix 1 (ThermoFisher Scientific). Poly(A) and cDNA synthesis steps were conducted on each RNA pool (1  $\mu$ g) with the TruSeq RNA sample Preparation kit v2 (Illumina, San Diego, CA, USA) according to manufacturer's protocol (Institut de Biologie Intégrative et des systèmes, Laval University sequencing platform, Québec, QC, Canada). Prior to sequencing, enrichment of UGT sequences with capture probes spanning genomic sequences of the *UGT1* and *UGT2* loci, including introns, intergenic sequences, and 50,000 bp on both side of each UGT locus, was conducted as described previously (Tourancheau et al. 2016).

Quality and basic metrics of paired-end datasets (2  $\times$  100 bp) were assessed with fastQC (v0.11.2), and mapping was conducted with Bowtie2 (v2.2.3; default parameters). Reads were trimmed with Trimmomatic (v0.36) (Bolger et al. 2014). Trimmed reads were aligned with the splice-aware software TopHat2 (v2.0.13; parameters: --no-novel-juncs) on the UCSC hg19 reference genome with annotation provided by Illumina iGenome combined with the complete annotated human UGT loci described recently by our group (Tourancheau et al. 2016). Normalized expression was reported in fragments per kilobase of transcript per million mapped reads from mapped read files (.bam) with the Cufflinks suite of tools (v2.2.1, parameters: --max-bundle-frags 10000000). Cuffquant and Cuffdiff



were used with the annotation used for the mapping step with the default library normalization and cross-replicate dispersion estimation methods. Extraction and treatment of normalized expression levels were performed using R (v3.1.2) and cummeRbund (v2.8.2) (Team" 2008, Goff et al. 2013).

### **Validation of *UGT2B7\_n4* expression**

*UGT2B7\_n4* variant expression was validated by reverse transcription-PCR and Sanger sequencing using a nested PCR strategy. Total RNA from normal human kidney (50 µg; Clontech Laboratories) was reverse transcribed with 125 pmol pDN6 random hexamers using 200 U Superscript III (Invitrogen, Carlsbad, CA, USA). Full-length *UGT2B7* mRNA was first amplified with 10 pmol each of the primers 289-F (5'-GACTGTACTGGCATCTTC) and 2702-R (5'-TGGGAATAAACTGAAGTAGTCT), 50 ng cDNA, 3 mM MgCl<sub>2</sub>, 200 mM dNTPs, and 1.25 U AmpliTaq Gold (Applied Biosystems Inc., Foster City, CA, USA.) in a final volume of 50 µl 1× PCR reaction buffer. The cycling conditions were: 10 min at 95°C, followed by 40 cycles of melting at 95°C for 30 s, annealing at 58°C for 40 s, and elongation at 72°C for 60–90 s, and ending with a final elongation step at 72°C for 7 min. PCR reactions were carried out in a 9700 thermal cycler (Applied Biosystems Inc.). From exon 1 to exon 2b of *UGT2B7\_n4* was subsequently amplified with primers 289-F and 4136-R (5'-GCTTGATGCCAGAGCCTTAT) and from exon 2b to the 3'UTR with the primers 4191-F (5'-ATGAATCTGTAAAGAGGCTCG) and 2702-R using 1 µl of a 1:100 dilution of the first PCR reaction in conditions as above. Sanger sequencing of the PCR amplicons confirmed their identity.

### **Plasmid constructs and cell models**

The *UGT2B7\_n4* plasmid construct was generated by inserting the exon 2b sequence into the *UGT2B7\_v1* pcDNA6 construct (Menard et al. 2011) with the Q5 Site-Directed Mutagenesis kit (New England Biolabs Ltd., Whitby, ON, Canada) and purified primers (exon2b-F: 5'-GATTCCTTCAACACTGGGTGGATAAATAAGGCTCTGGGCATCAAG-

CTAGAAATGGAAGACTTTGTACAGAGC; exon2b-R: 5'-GAGCCTCTTTACAGATTCATCACTGATATGGACTGGTTGTTTCTTCTTCTTAGGCAGGGGTTTGGC) and was verified by Sanger sequencing.

HEK293 and HepG2 cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA) and grown as described (Levesque et al. 1997, Menard et al. 2013). HEK293 cells were transfected with Effectene (Qiagen), and HepG2 cells were transfected by electroporation with the Neon Transfection System (Invitrogen). HT115 cells were from the European Collection of Cell Cultures (Salisbury, UK).

### **Antibodies specific to novel isoforms**

A mouse monoclonal antibody specific to UGT1A\_i2 isoforms was produced against the synthetic peptide C-I<sub>432</sub>NDRKKQSGRQM<sub>444</sub> by Genscript (Piscataway, NJ, USA). The specificity of this antibody was verified by western blotting.

Antibodies specific to UGT2B7\_i8 were raised in rabbits against the synthetic peptide K<sub>290</sub>KKKQPVHISDES<sub>303</sub>-C and affinity purified with the immunogenic peptide by Genscript according to the proprietary Polyexpress protocol. The immunogenic peptide was specific to the unique sequence encoded by exon 2b, and the generated antibody specifically recognized UGT2B7\_i8 and not the canonical UGT2B7\_i1 as determined by immunoprecipitation and immunoblotting.

### **Multiple reaction monitoring mass spectrometry (MRM-MS)**

Human liver S9 fraction (8 mg protein; Xenotech LLC, Lenexa, KS, USA) was lysed for 45 min on ice in a total volume of 4 ml lysis buffer containing 0.05 M Tris-HCl pH 7.4, 0.15 M NaCl, 1% (w/v) Igepal CA-630 (Sigma-Aldrich), 1 mM dithiothreitol, and Complete protease inhibitor cocktail (Roche, Laval, QC, Canada). Lysates were homogenized by

pipetting up and down through fine needles (18G and 20G, 10 times each) on ice and cleared by centrifugation for 15 min at 13,000 g. UGT2B7 was performed immunoprecipitated with 10 µg of purified antibody (ab57685) for 1 h at 4°C on an orbital shaker. Protein G-coated magnetic beads (200 µl Dynabeads, Life Technologies, Burlington, ON, Canada) were added and incubated overnight at 4°C. Beads were washed three times with 4 ml lysis buffer and five times with 50 mM ammonium bicarbonate and stored at -20°C until analysis. Beads were resuspended in 10 µl trypsin (10 ng/µl in 50 mM ammonium bicarbonate; ThermoFisher Scientific) and incubated at 37°C for 5 h. Trypsin digestion was stopped by addition of formic acid to 1%. The tryptic digest was recovered, dried, and resuspended in 30 µl sample buffer (3% acetonitrile, 0.1% trifluoroacetic acid, 0.5% acetic acid). Peptides were desalted on a C18 Empore filter (ThermoFisher Scientific), eluted in 80% acetonitrile-0.1% trifluoroacetic acid, and dried in a vacuum centrifuge.

The UGT2B7 tryptic peptide ADVWLIR, present in nearly all UGT2B7 isoforms, was selected as the common signature peptide (Fallon et al. 2013). UGT2B7\_i8-specific peptide selection was conducted with the open source Skyline v2.6 program (MacCoss Lab, Seattle, WA, USA). UGT2B7\_i8 sequence was digested *in silico* to create a list of potential tryptic peptides of 5 to 25 amino acid residues. Theoretical MRM transitions were generated by including y-ions from +2 and +3 charge state precursors with mass >300 Da and <1500 Da. UGT2B7\_i8 immunoprecipitated from the HEK2B7\_i8 cell model described above was analyzed by MRM. The peptide LDSFNTGWINK, producing the most sensitive and specific signal, was selected. Specificity of the selected peptides for the UGT2B7 isoforms was verified by a tblastn search of the reference human RNA sequences (NCBI accessed March 21, 2016). Corresponding isotopically labeled synthetic crude peptides [<sup>13</sup>C6,<sup>15</sup>N2]Lys and [<sup>13</sup>C6,<sup>15</sup>N4]Arg were obtained from Pierce Protein Biotechnology (ThermoFisher Scientific) and solubilized in 0.1% formic acid.

For MRM-MS analysis, each tryptic digest was reconstituted in 10  $\mu$ l of a peptide standard solution containing 5  $\mu$ M of each synthetic peptide and analyzed on a 6500QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer equipped with a nanoLC 400 cHiPLC nanoflex system controlled by Analyst 1.6 (Sciex, Concord, ON, Canada) and a nanospray ionization source. MS analyses were conducted in positive ion mode with an ionspray voltage of 2500 V. Peptides were desalted on a 200  $\mu$ m  $\times$  6 mm chip trap column packed with ChromXP C18, 3  $\mu$ m (Eksigent, Sciex), at 2  $\mu$ l/min solvent A (0.1% formic acid). Then the peptides were eluted at a flow rate of 1  $\mu$ l/min on a 200  $\mu$ m  $\times$  15 cm chip column packed with ChromXP C18, 3  $\mu$ m (Eksigent), over a 30-min linear gradient from 5 to 40% solvent B (acetonitrile with 0.1% formic acid) and a 10-min linear gradient from 40 to 95% solvent B. The nebulizer gas was set at 20[ $^{\circ}$ C] (Gas 1), curtain gas at 35 $^{\circ}$ C, and heater at 100 $^{\circ}$ C. MRM-MS analyses were performed using the four most intense transitions for each of the target peptides for the light and heavy form. A blank solvent injection was run between samples to prevent sample carryover. Samples containing 5 fmol digested bovine serum albumin were injected before and after the analysis to confirm system stability.

### **Statistical analysis of metabolomics data**

Raw metabolomics data were normalized with cell counts in the corresponding sample (Table S1). Statistical significance was determined by Student's *t*-tests on natural logarithm-transformed normalized raw data. Homogeneity of variances was assessed by an F-test and results taken into account in the Student's *t*-test. All calculations were performed with Excel 2016. Graphical representations of the metabolomics data were made using the ggplot2 package in the statistical software R (v3.1.2) (Team" 2008).

## References for Supplemental Experimental Procedures

- Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* *30*, 2114-2120.
- Fallon, J.K., Neubert, H., Hyland, R., Goosen, T.C., and Smith, P.C. (2013). Targeted quantitative proteomics for the analysis of 14 UGT1As and -2Bs in human liver using NanoUPLC-MS/MS with selected reaction monitoring. *J Proteome Res* *12*, 4402-4413.
- Goff, L., Trapnell, C., and Kelley, D. (2013). cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data. R package version 2.8.2.
- Levesque, E., Beaulieu, M., Green, M.D., Tephly, T.R., Belanger, A., and Hum, D.W. (1997). Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene. *Pharmacogenetics* *7*, 317-325.
- Menard, V., Collin, P., Margaillan, G., and Guillemette, C. (2013). Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos* *41*, 2197-2205.
- Menard, V., Eap, O., Roberge, J., Harvey, M., Levesque, E., and 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. *Pharmacogenet Genomics* *21*, 631-641.
- R Development Core Team, R.C. (2008). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; URL: <https://www.R-project.org>.
- Tourancheau, A., Margaillan, G., Rouleau, M., Gilbert, I., Villeneuve, L., Levesque, E., Droit, A., and Guillemette, C. (2016). Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J* *16*, 60-70.

## Bibliographie

- Al Saabi A, Allorge D, Sauvage FL, Tournel G, Gaulier JM, Marquet P and Picard N (2013). Involvement of UDP-glucuronosyltransferases UGT1A9 and UGT2B7 in ethanol glucuronidation, and interactions with common drugs of abuse. *Drug Metab Dispos* **41**: 568-574.
- Albert C, Vallee M, Beaudry G, Belanger A and Hum DW (1999). The monkey and human uridine diphosphate-glucuronosyltransferase UGT1A9, expressed in steroid target tissues, are estrogen-conjugating enzymes. *Endocrinology* **140**: 3292-3302.
- Allegaert K, Vanhole C, Vermeersch S, Rayyan M, Verbesselt R and de Hoon J (2008). Both postnatal and postmenstrual age contribute to the interindividual variability in tramadol glucuronidation in neonates. *Early Hum Dev* **84**: 325-330.
- Barbier O, Girard C, Breton R, Belanger A and Hum DW (2000). N-glycosylation and residue 96 are involved in the functional properties of UDP-glucuronosyltransferase enzymes. *Biochemistry* **39**: 11540-11552.
- Basu NK, Kole L, Basu M, Chakraborty K, Mitra PS and Owens IS (2008). The major chemical-detoxifying system of UDP-glucuronosyltransferases requires regulated phosphorylation supported by protein kinase C. *J Biol Chem* **283**: 23048-23061.
- Basu NK, Kovarova M, Garza A, Kubota S, Saha T, Mitra PS, Banerjee R, Rivera J and Owens IS (2005). Phosphorylation of a UDP-glucuronosyltransferase regulates substrate specificity. *Proc Natl Acad Sci U S A* **102**: 6285-6290.
- Belanger AS, Caron P, Harvey M, Zimmerman PA, Mehlotra RK and Guillemette C (2009). Glucuronidation of the antiretroviral drug efavirenz by UGT2B7 and an in vitro investigation of drug-drug interaction with zidovudine. *Drug Metab Dispos* **37**: 1793-1796.
- Bellemare J, Rouleau M, Harvey M and Guillemette C (2010). Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *J Biol Chem* **285**: 3600-3607.
- Bellemare J, Rouleau M, Harvey M, Popa I, Pelletier G, Tetu B and Guillemette C (2011). Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *J Pathol* **223**: 425-435.
- Bellemare J, Rouleau M, Harvey M, Tetu B and Guillemette C (2010). Alternative-splicing forms of the major phase II conjugating UGT1A gene negatively regulate glucuronidation in human carcinoma cell lines. *Pharmacogenomics J* **10**: 431-441.
- Benoit-Biancamano MO, Caron P, Levesque E, Delage R, Couture F and Guillemette C (2007). Sensitive high-performance liquid chromatography-tandem mass spectrometry method for quantitative analysis of mycophenolic acid and its glucuronide metabolites in human plasma and urine. *J Chromatogr B Analyt Technol Biomed Life Sci* **858**: 159-167.
- Bernard BA, Yamada KM and Olden K (1982). Carbohydrates selectively protect a specific domain of fibronectin against proteases. *J Biol Chem* **257**: 8549-8554.
- Bhasker CR, McKinnon W, Stone A, Lo AC, Kubota T, Ishizaki T and Miners JO (2000). Genetic polymorphism of UDP-glucuronosyltransferase 2B7 (UGT2B7) at amino acid 268: ethnic diversity of alleles and potential clinical significance. *Pharmacogenetics* **10**: 679-685.

- Biamonti G, Catillo M, Pignataro D, Montecucco A and Ghigna C (2014). The alternative splicing side of cancer. *Semin Cell Dev Biol* **32**: 30-36.
- Bock KW (2015). Roles of human UDP-glucuronosyltransferases in clearance and homeostasis of endogenous substrates, and functional implications. *Biochem Pharmacol* **96**: 77-82.
- Bock KW (2016). The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution. *Biochem Pharmacol* **99**: 11-17.
- Bolger AM, Lohse M and Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.
- Breton C, Fournel-Gigleux S and Palcic MM (2012). Recent structures, evolution and mechanisms of glycosyltransferases. *Curr Opin Struct Biol* **22**: 540-549.
- Bullingham R, Shah J, Goldblum R and Schiff M (1996). Effects of food and antacid on the pharmacokinetics of single doses of mycophenolate mofetil in rheumatoid arthritis patients. *Br J Clin Pharmacol* **41**: 513-516.
- Butzke B, Oduncu FS, Severin F, Pfeufer A, Heinemann V, Giessen-Jung C, Stollenwerk B and Rogowski WH (2016). The cost-effectiveness of UGT1A1 genotyping before colorectal cancer treatment with irinotecan from the perspective of the German statutory health insurance. *Acta Oncol* **55**: 318-328.
- Carstens RP (2014). Networking in an alternative splicing world. *Mol Cell* **54**: 903-904.
- Chakraborty SK, Basu NK, Jana S, Basu M, Raychoudhuri A and Owens IS (2012). Protein kinase Calpha and Src kinase support human prostate-distributed dihydrotestosterone-metabolizing UDP-glucuronosyltransferase 2B15 activity. *J Biol Chem* **287**: 24387-24396.
- Chau N, Elliot DJ, Lewis BC, Burns K, Johnston MR, Mackenzie PI and Miners JO (2014). Morphine glucuronidation and glucosidation represent complementary metabolic pathways that are both catalyzed by UDP-glucuronosyltransferase 2B7: kinetic, inhibition, and molecular modeling studies. *J Pharmacol Exp Ther* **349**: 126-137.
- Chauhan JS, Rao A and Raghava GP (2013). In silico platform for prediction of N-, O- and C-glycosites in eukaryotic protein sequences. *PLoS One* **8**: e67008.
- Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L and Zou H (2009). Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res* **8**: 651-661.
- Chhibber A, French CE, Yee SW, Gamazon ER, Theusch E, Qin X, Webb A, Papp AC, Wang A, Simmons CQ, Konkashbaev A, Chaudhry AS, Mitchel K, Stryke D, Ferrin TE, Weiss ST, Kroetz DL, Sadee W, Nickerson DA, Krauss RM, George AL, Schuetz EG, Medina MW, Cox NJ, Scherer SE, Giacomini KM and Brenner SE (2016). Transcriptomic variation of pharmacogenes in multiple human tissues and lymphoblastoid cell lines. *Pharmacogenomics J*.
- Clark MB, Mercer TR, Bussotti G, Leonardi T, Haynes KR, Crawford J, Brunck ME, Cao KA, Thomas GP, Chen WY, Taft RJ, Nielsen LK, Enright AJ, Mattick JS and Dinger ME (2015). Quantitative gene profiling of long noncoding RNAs with targeted RNA sequencing. *Nat Methods* **12**: 339-342.
- Cohen P (2000). The regulation of protein function by multisite phosphorylation--a 25 year update. *Trends Biochem Sci* **25**: 596-601.

- Coller JK, Christrup LL and Somogyi AA (2009). Role of active metabolites in the use of opioids. *Eur J Clin Pharmacol* **65**: 121-139.
- Court MH (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 Metab Rev* **42**: 209-224.
- Court MH, Zhang X, Ding X, Yee KK, Hesse LM and Finel M (2012). Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica* **42**: 266-277.
- Crean CS and Tompson DJ (2013). The effects of ethanol on the pharmacokinetics, pharmacodynamics, safety, and tolerability of ezogabine (retigabine). *Clin Ther* **35**: 87-93.
- Dates CR, Fahmi T, Pyrek SJ, Yao-Borengasser A, Borowa-Mazgaj B, Bratton SM, Kadlubar SA, Mackenzie PI, Haun RS and Radomska-Pandya A (2015). Human UDP-Glucuronosyltransferases: Effects of altered expression in breast and pancreatic cancer cell lines. *Cancer Biol Ther* **16**: 714-723.
- De Bruyn F, Maertens J, Beauprez J, Soetaert W and De Mey M (2015). Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnol Adv* **33**: 288-302.
- de Vries EM, van Beeren HC, Ackermans MT, Kalsbeek A, Fliers E and Boelen A (2015). Differential effects of fasting vs food restriction on liver thyroid hormone metabolism in male rats. *J Endocrinol* **224**: 25-35.
- Djebli N, Picard N, Rerolle JP, Le Meur Y and Marquet P (2007). Influence of the UGT2B7 promoter region and exon 2 polymorphisms and comedications on Acyl-MPAG production in vitro and in adult renal transplant patients. *Pharmacogenet Genomics* **17**: 321-330.
- Eswaran J, Horvath A, Godbole S, Reddy SD, Mudvari P, Ohshiro K, Cyanam D, Nair S, Fuqua SA, Polyak K, Florea LD and Kumar R (2013). RNA sequencing of cancer reveals novel splicing alterations. *Sci Rep* **3**: 1689.
- Etienne-Grimaldi MC, Boyer JC, Thomas F, Quaranta S, Picard N, Loriot MA, Narjoz C, Poncet D, Gagnieu MC, Ged C, Broly F, Le Morvan V, Bouquie R, Gaub MP, Philibert L, Ghiringhelli F, Le Guellec C, Collective work by Groupe de Pharmacologie Clinique O and French Réseau National de Pharmacogenétique H (2015). UGT1A1 genotype and irinotecan therapy: general review and implementation in routine practice. *Fundam Clin Pharmacol* **29**: 219-237.
- Evans WER, M. V. (2004). Moving towards individualized medicine with pharmacogenomics. *Nature* **429**: 464-468.
- Fallon JK, Neubert H, Hyland R, Goosen TC and Smith PC (2013). Targeted quantitative proteomics for the analysis of 14 UGT1As and -2Bs in human liver using NanoUPLC-MS/MS with selected reaction monitoring. *J Proteome Res* **12**: 4402-4413.
- Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y, Moon S and Nikolau B (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* **53**: 691-704.
- Freeze HH and Westphal V (2001). Balancing N-linked glycosylation to avoid disease. *Biochimie* **83**: 791-799.



- Fujikura K, Ingelman-Sundberg M and Lauschke VM (2015). Genetic variation in the human cytochrome P450 supergene family. *Pharmacogenetics and Genomics* **25**: 584-594.
- Fujiwara R, Nakajima M, Oda S, Yamanaka H, Ikushiro S, Sakaki T and Yokoi T (2010). Interactions between human UDP-glucuronosyltransferase (UGT) 2B7 and UGT1A enzymes. *J Pharm Sci* **99**: 442-454.
- Gamazon ER and Stranger BE (2014). Genomics of alternative splicing: evolution, development and pathophysiology. *Hum Genet* **133**: 679-687.
- Garcia-Blanco MA, Baraniak AP and Lasda EL (2004). Alternative splicing in disease and therapy. *Nat Biotechnol* **22**: 535-546.
- Gillis NK, Patel JN and Innocenti F (2014). Clinical implementation of germ line cancer pharmacogenetic variants during the next-generation sequencing era. *Clin Pharmacol Ther* **95**: 269-280.
- Girard H, Levesque E, Bellemare J, Journault K, Caillier B and 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. *Pharmacogenet Genomics* **17**: 1077-1089.
- Goff L, Trapnell C and Kelley D (2013). cummeRbund: Analysis, exploration, manipulation, and visualization of Cufflinks high-throughput sequencing data.. *R package version 2.8.2*.
- Gong QH, Cho JW, Huang T, Potter C, Gholami N, Basu NK, Kubota S, Carvalho S, Pennington MW, Owens IS and Popescu NC (2001). Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* **11**: 357-368.
- Gradinaru J, Romand S, Geiser L, Carrupt PA, Spaggiari D and Rudaz S (2015). Inhibition screening method of microsomal UGTs using the cocktail approach. *Eur J Pharm Sci* **71**: 35-45.
- Guillemette C (2003). Pharmacogenomics of human UDP-glucuronosyltransferase enzymes. *Pharmacogenomics J* **3**: 136-158.
- Guillemette C, Belanger A and Lepine J (2004). Metabolic inactivation of estrogens in breast tissue by UDP-glucuronosyltransferase enzymes: an overview. *Breast Cancer Res* **6**: 246-254.
- Guillemette C, Levesque E, Harvey M, Bellemare J and Menard V (2010). UGT genomic diversity: beyond gene duplication. *Drug Metab Rev* **42**: 24-44.
- Guillemette C, Levesque E and Rouleau M (2014). Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* **96**: 324-339.
- Gupta R. (2015, August 4th 2015). "NetNGlyc 1.0 Server." Retrieved May 19th, 2016, from <http://www.cbs.dtu.dk/services/NetNGlyc/>.
- Gupta R and Brunak S (2002). Prediction of glycosylation across the human proteome and the correlation to protein function. *Pac Symp Biocomput*: 310-322.
- Haberkorn V, Heydel JM, Mounie J, Artur Y and Goudonnet H (2002). Vitamin A modulates the effects of thyroid hormone on UDP-glucuronosyltransferase expression and activity in rat liver. *Mol Cell Endocrinol* **190**: 167-175.

- Hall D, Ybazeta G, Destro-Bisol G, Petzl-Erler ML and Di Rienzo A (1999). Variability at the uridine diphosphate glucuronosyltransferase 1A1 promoter in human populations and primates. *Pharmacogenetics* **9**: 591-599.
- Hermann R, Ferron GM, Erb K, Knebel N, Ruus P, Paul J, Richards L, Cnota HP and Troy S (2003). Effects of age and sex on the disposition of retigabine. *Clin Pharmacol Ther* **73**: 61-70.
- Hoyumpa AM, Patwardhan R, Maples M, Desmond PV, Johnson RF, Sinclair AP and Schenker S (1981). Effect of short-term ethanol administration on lorazepam clearance. *Hepatology* **1**: 47-53.
- Hu DG, Meech R, McKinnon RA and Mackenzie PI (2014). Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab Rev* **46**: 421-458.
- Imperiali B and O'Connor SE (1999). Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr Opin Chem Biol* **3**: 643-649.
- Innocenti F, Iyer L, Ramirez J, Green MD and Ratain MJ (2001). Epirubicin glucuronidation is catalyzed by human UDP-glucuronosyltransferase 2B7. *Drug Metab Dispos* **29**: 686-692.
- Ishii Y, Takeda S and Yamada H (2010). Modulation of UDP-glucuronosyltransferase activity by protein-protein association. *Drug Metab Rev* **42**: 145-158.
- Jeong H (2010). Altered drug metabolism during pregnancy: hormonal regulation of drug-metabolizing enzymes. *Expert Opin Drug Metab Toxicol* **6**: 689-699.
- Johannessen SI and Landmark CJ (2010). Antiepileptic drug interactions - principles and clinical implications. *Curr Neuropharmacol* **8**: 254-267.
- Joo J, Kim YW, Wu Z, Shin JH, Lee B, Shon JC, Lee EY, Phuc NM and Liu KH (2015). Screening of non-steroidal anti-inflammatory drugs for inhibitory effects on the activities of six UDP-glucuronosyltransferases (UGT1A1, 1A3, 1A4, 1A6, 1A9 and 2B7) using LC-MS/MS. *Biopharm Drug Dispos* **36**: 258-264.
- Kodama S, Tsujimoto M, Tsuruoka N, Sugo T, Endo T and Kobata A (1993). Role of sugar chains in the in-vitro activity of recombinant human interleukin 5. *Eur J Biochem* **211**: 903-908.
- Komander D (2009). The emerging complexity of protein ubiquitination. *Biochem Soc Trans* **37**: 937-953.
- Kwara A, Lartey M, Boamah I, Rezk NL, Oliver-Commey J, Kenu E, Kashuba AD and Court MH (2009). Interindividual variability in pharmacokinetics of generic nucleoside reverse transcriptase inhibitors in TB/HIV-coinfected Ghanaian patients: UGT2B7\*1c is associated with faster zidovudine clearance and glucuronidation. *J Clin Pharmacol* **49**: 1079-1090.
- Lepine J, Bernard O, Plante M, Tetu B, Pelletier G, Labrie F, Belanger A and Guillemette C (2004). Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphosphoglucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* **89**: 5222-5232.
- Levesque E, Beaulieu M, Green MD, Tephly TR, Belanger A and Hum DW (1997). Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene. *Pharmacogenetics* **7**: 317-325.

- Levesque E, Benoit-Biancamano MO, Delage R, Couture F and Guillemette C (2008). Pharmacokinetics of mycophenolate mofetil and its glucuronide metabolites in healthy volunteers. *Pharmacogenomics* **9**: 869-879.
- Levesque E, Delage R, Benoit-Biancamano MO, Caron P, Bernard O, Couture F and Guillemette C (2007). The impact of UGT1A8, UGT1A9, and UGT2B7 genetic polymorphisms on the pharmacokinetic profile of mycophenolic acid after a single oral dose in healthy volunteers. *Clin Pharmacol Ther* **81**: 392-400.
- Levesque E, Girard H, Journault K, Lepine J and Guillemette C (2007). Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus. *Hepatology* **45**: 128-138.
- Lewis BC, Mackenzie PI, Elliot DJ, Burchell B, Bhasker CR and Miners JO (2007). Amino terminal domains of human UDP-glucuronosyltransferases (UGT) 2B7 and 2B15 associated with substrate selectivity and autoactivation. *Biochem Pharmacol* **73**: 1463-1473.
- Li S, Iakoucheva LM, Mooney SD and Radivojac P (2010). Loss of post-translational modification sites in disease. *Pac Symp Biocomput*: 337-347.
- Little JM, Kurkela M, Sonka J, Jantti S, Ketola R, Bratton S, Finel M and Radominska-Pandya A (2004). Glucuronidation of oxidized fatty acids and prostaglandins B1 and E2 by human hepatic and recombinant UDP-glucuronosyltransferases. *J Lipid Res* **45**: 1694-1703.
- Lo WS, Gardiner E, Xu Z, Lau CF, Wang F, Zhou JJ, Mendlein JD, Nangle LA, Chiang KP, Yang XL, Au KF, Wong WH, Guo M, Zhang M and Schimmel P (2014). Human tRNA synthetase catalytic nulls with diverse functions. *Science* **345**: 328-332.
- Loryan I, Lindqvist M, Johansson I, Hiratsuka M, van der Heiden I, van Schaik RH, Jakobsson J and Ingelman-Sundberg M (2012). Influence of sex on propofol metabolism, a pilot study: implications for propofol anesthesia. *Eur J Clin Pharmacol* **68**: 397-406.
- Lu CT, Huang KY, Su MG, Lee TY, Bretana NA, Chang WC, Chen YJ, Chen YJ and Huang HD (2013). DbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Res* **41**: D295-305.
- Mackenzie PI (1990). The effect of N-linked glycosylation on the substrate preferences of UDP glucuronosyltransferases. *Biochem Biophys Res Commun* **166**: 1293-1299.
- Mackenzie PI, Bock KW, Burchell B, Guillemette C, Ikushiro S, Iyanagi T, Miners JO, Owens IS and Nebert DW (2005). Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics* **15**: 677-685.
- Maes B, Oellerich M, Ceuppens JL, Armstrong VW, Evenepoel P, Kuypers D, Messiaen T, Shipkova M, Wieland E and Vanrenterghem Y (2002). A new acute inflammatory syndrome related to the introduction of mycophenolate mofetil in patients with Wegener's granulomatosis. *Nephrol Dial Transplant* **17**: 923-926.
- Mahmood I (2015). Prediction of glucuronidated drug clearance in pediatrics ( $\leq 5$  years): An allometric approach. *Eur J Drug Metab Pharmacokinet* **40**: 53-59.
- Mandal R, Becker S and Strebhardt K (2016). Stamping out RAF and MEK1/2 to inhibit the ERK1/2 pathway: an emerging threat to anticancer therapy. *Oncogene* **35**: 2547-2561.

- Mann M and Jensen ON (2003). Proteomic analysis of post-translational modifications. *Nat Biotechnol* **21**: 255-261.
- Margaillan G, Rouleau M, Fallon JK, Caron P, Villeneuve L, Turcotte V, Smith PC, Joy MS and Guillemette C (2015). Quantitative profiling of human renal UDP-glucuronosyltransferases and glucuronidation activity: a comparison of normal and tumoral kidney tissues. *Drug Metab Dispos* **43**: 611-619.
- Margaillan G, Rouleau M, Klein K, Fallon JK, Caron P, Villeneuve L, Smith PC, Zanger UM and Guillemette C (2015). Multiplexed Targeted Quantitative Proteomics Predicts Hepatic Glucuronidation Potential. *Drug Metab Dispos* **43**: 1331-1335.
- Marth GT, Yu F, Indap AR, Garimella K, Gravel S, Leong WF, Tyler-Smith C, Bainbridge M, Blackwell T, Zheng-Bradley X, Chen Y, Challis D, Clarke L, Ball EV, Cibulskis K, Cooper DN, Fulton B, Hartl C, Koboldt D, Muzny D, Smith R, Sougnez C, Stewart C, Ward A, Yu J, Xue Y, Altshuler D, Bustamante CD, Clark AG, Daly M, DePristo M, Flicek P, Gabriel S, Mardis E, Palotie A, Gibbs R and Genomes P (2011). The functional spectrum of low-frequency coding variation. *Genome Biol* **12**: R84.
- Mazidi T, Rouini MR, Ghahremani MH, Dashti-Khavidaki S, Lessan-Pezeshki M, Ahmadi FL, Salam-Zadeh J, Mandegary A and Gholami K (2013). Impact of UGT1A9 Polymorphism on Mycophenolic Acid Pharmacokinetic Parameters in Stable Renal Transplant Patients. *Iran J Pharm Res* **12**: 547-556.
- Mazumder R, Morampudi KS, Motwani M, Vasudevan S and Goldman R (2012). Proteome-wide analysis of single-nucleotide variations in the N-glycosylation sequon of human genes. *PLoS One* **7**: e36212.
- McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, Altshuler DM and International HapMap C (2006). Common deletion polymorphisms in the human genome. *Nat Genet* **38**: 86-92.
- Meech R and Mackenzie PI (1997). Structure and function of uridine diphosphate glucuronosyltransferases. *Clin Exp Pharmacol Physiol* **24**: 907-915.
- Menard V, Collin P, Margaillan G and Guillemette C (2013). Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos* **41**: 2197-2205.
- Menard V, Eap O, Roberge J, Harvey M, Levesque E and 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. *Pharmacogenet Genomics* **21**: 631-641.
- Menard V, Levesque E, Chen S, Eap O, Joy MS, Ekstrom L, Rane A and Guillemette C (2013). Expression of UGT2B7 is driven by two mutually exclusive promoters and alternative splicing in human tissues: changes from prenatal life to adulthood and in kidney cancer. *Pharmacogenet Genomics* **23**: 684-696.
- Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddloh JA, Mattick JS and Rinn JL (2012). Targeted RNA sequencing reveals the deep complexity of the human transcriptome. *Nat Biotechnol* **30**: 99-104.
- Miley MJ, Zielinska AK, Keenan JE, Bratton SM, Radomska-Pandya A and Redinbo MR (2007). Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *J Mol Biol* **369**: 498-511.

- Mitra PS, Basu NK, Basu M, Chakraborty S, Saha T and Owens IS (2011). Regulated phosphorylation of a major UDP-glucuronosyltransferase isozyme by tyrosine kinases dictates endogenous substrate selection for detoxification. *J Biol Chem* **286**: 1639-1648.
- Mitra PS, Basu NK and Owens IS (2009). Src supports UDP-glucuronosyltransferase-2B7 detoxification of catechol estrogens associated with breast cancer. *Biochem Biophys Res Commun* **382**: 651-656.
- Miyagi SJ, Milne AM, Coughtrie MW and Collier AC (2012). Neonatal development of hepatic UGT1A9: implications of pediatric pharmacokinetics. *Drug Metab Dispos* **40**: 1321-1327.
- Moore LB, Goodwin B, Jones SA, Wisely GB, Serabjit-Singh CJ, Willson TM, Collins JL and Kliewer SA (2000). St. John's wort induces hepatic drug metabolism through activation of the pregnane X receptor. *Proc Natl Acad Sci U S A* **97**: 7500-7502.
- Nagaoka K, Hanioka N, Ikushiro S, Yamano S and Narimatsu S (2012). The effects of N-glycosylation on the glucuronidation of zidovudine and morphine by UGT2B7 expressed in HEK293 cells. *Drug Metab Pharmacokinet* **27**: 388-397.
- Nakajima M, Koga T, Sakai H, Yamanaka H, Fujiwara R and Yokoi T (2010). N-Glycosylation plays a role in protein folding of human UGT1A9. *Biochem Pharmacol* **79**: 1165-1172.
- Nelson MR, Wegmann D, Ehm MG, Kessner D, St Jean P, Verzilli C, Shen J, Tang Z, Bacanu SA, Fraser D, Warren L, Aponte J, Zawistowski M, Liu X, Zhang H, Zhang Y, Li J, Li Y, Li L, Woollard P, Topp S, Hall MD, Nangle K, Wang J, Abecasis G, Cardon LR, Zollner S, Whittaker JC, Chissoe SL, Novembre J and Mooser V (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**: 100-104.
- Nilsen TW and Graveley BR (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**: 457-463.
- Nilsson I and von Heijne G (2000). Glycosylation efficiency of Asn-Xaa-Thr sequons depends both on the distance from the C terminus and on the presence of a downstream transmembrane segment. *J Biol Chem* **275**: 17338-17343.
- Oda S, Fukami T, Yokoi T and Nakajima M (2013). Epigenetic regulation is a crucial factor in the repression of UGT1A1 expression in the human kidney. *Drug Metab Dispos* **41**: 1738-1743.
- Oltean S and Bates DO (2014). Hallmarks of alternative splicing in cancer. *Oncogene* **33**: 5311-5318.
- Operana TN and Tukey RH (2007). Oligomerization of the UDP-glucuronosyltransferase 1A proteins: homo- and heterodimerization analysis by fluorescence resonance energy transfer and co-immunoprecipitation. *J Biol Chem* **282**: 4821-4829.
- Ouzzine M, Barre L, Netter P, Magdalou J and Fournel-Gigleux S (2006). Role of the carboxyl terminal stop transfer sequence of UGT1A6 membrane protein in ER targeting and translocation of upstream luminal domain. *FEBS Lett* **580**: 1953-1958.
- Ouzzine M, Magdalou J, Burchell B and Fournel-Gigleux S (1999). An internal signal sequence mediates the targeting and retention of the human UDP-glucuronosyltransferase 1A6 to the endoplasmic reticulum. *J Biol Chem* **274**: 31401-31409.
- Owens IS, Basu NK and Banerjee R (2005). UDP-glucuronosyltransferases: gene structures of UGT1 and UGT2 families. *Methods Enzymol* **400**: 1-22.

- Pan Q, Shai O, Lee LJ, Frey BJ and Blencowe BJ (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40**: 1413-1415.
- Peterkin VC, Bauman JN, Goosen TC, Menning L, Man MZ, Paulauskis JD, Williams JA and Myrand SP (2007). Limited influence of UGT1A1\*28 and no effect of UGT2B7\*2 polymorphisms on UGT1A1 or UGT2B7 activities and protein expression in human liver microsomes. *Br J Clin Pharmacol* **64**: 458-468.
- Popp MW and Maquat LE (2013). Organizing principles of mammalian nonsense-mediated mRNA decay. *Annu Rev Genet* **47**: 139-165.
- Radomska-Pandya A, Bratton SM, Redinbo MR and Miley MJ (2010). The crystal structure of human UDP-glucuronosyltransferase 2B7 C-terminal end is the first mammalian UGT target to be revealed: the significance for human UGTs from both the 1A and 2B families. *Drug Metab Rev* **42**: 133-144.
- Radomska-Pandya A, Czernik PJ, Little JM, Battaglia E and Mackenzie PI (1999). Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev* **31**: 817-899.
- Ramsey LB, Bruun GH, Yang W, Trevino LR, Vattathil S, Scheet P, Cheng C, Rosner GL, Giacomini KM, Fan Y, Sparreboom A, Mikkelsen TS, Corydon TJ, Pui CH, Evans WE and Relling MV (2012). Rare versus common variants in pharmacogenetics: SLCO1B1 variation and methotrexate disposition. *Genome Res* **22**: 1-8.
- Riches Z and Collier AC (2015). Posttranscriptional regulation of uridine diphosphate glucuronosyltransferases. *Expert Opin Drug Metab Toxicol* **11**: 949-965.
- Rouleau M, Collin P, Bellemare J, Harvey M and Guillemette C (2013). Protein-protein interactions between the bilirubin-conjugating UDP-glucuronosyltransferase UGT1A1 and its shorter isoform 2 regulatory partner derived from alternative splicing. *Biochem J* **450**: 107-114.
- Rouleau M, Roberge J, Bellemare J and Guillemette C (2014). Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Mol Pharmacol* **85**: 29-36.
- Rowland A, Miners JO and Mackenzie PI (2013). The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol* **45**: 1121-1132.
- Runkel L, Meier W, Pepinsky RB, Karpusas M, Whitty A, Kimball K, Brickelmaier M, Muldowney C, Jones W and Goelz SE (1998). Structural and functional differences between glycosylated and non-glycosylated forms of human interferon-beta (IFN-beta). *Pharm Res* **15**: 641-649.
- Sadoul K, Boyault C, Pabion M and Khochbin S (2008). Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie* **90**: 306-312.
- Saeki M, Saito Y, Jinno H, Tanaka-Kagawa T, Ohno A, Ozawa S, Ueno K, Kamakura S, Kamatani N, Komamura K, Kitakaze M and Sawada J (2004). Single nucleotide polymorphisms and haplotype frequencies of UGT2B4 and UGT2B7 in a Japanese population. *Drug Metab Dispos* **32**: 1048-1054.
- Sakaguchi K, Green M, Stock N, Reger TS, Zunic J and King C (2004). Glucuronidation of carboxylic acid containing compounds by UDP-glucuronosyltransferase isoforms. *Arch Biochem Biophys* **424**: 219-225.
- Sawyer MB, Pituskin E, Damaraju S, Bies RR, Vos LJ, Prado CM, Kuzma M, Scarfe AG, Clemons M, Tonkin K, Au HJ, Koski S, Joy AA, Smylie M, King K, Carandang D,

- Damaraju VL, Hanson J, Cass CE and Mackey JR (2016). A Uridine Glucuronosyltransferase 2B7 Polymorphism Predicts Epirubicin Clearance and Outcomes in Early-Stage Breast Cancer. *Clin Breast Cancer* **16**: 139-144 e133.
- Seo KA, Kim HJ, Jeong ES, Abdalla N, Choi CS, Kim DH and Shin JG (2014). In vitro assay of six UDP-glucuronosyltransferase isoforms in human liver microsomes, using cocktails of probe substrates and liquid chromatography-tandem mass spectrometry. *Drug Metab Dispos* **42**: 1803-1810.
- Shabalina SA, Ogurtsov AY, Spiridonov NA and Koonin EV (2014). Evolution at protein ends: major contribution of alternative transcription initiation and termination to the transcriptome and proteome diversity in mammals. *Nucleic Acids Res* **42**: 7132-7144.
- Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM and Sirotkin K (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**: 308-311.
- Sinclair AM and Elliott S (2005). Glycoengineering: the effect of glycosylation on the properties of therapeutic proteins. *J Pharm Sci* **94**: 1626-1635.
- Singkham N, Towanabut S, Lertkachatarn S and Punyawudho B (2013). Influence of the UGT2B7 -161C>T polymorphism on the population pharmacokinetics of lamotrigine in Thai patients. *Eur J Clin Pharmacol* **69**: 1285-1291.
- Skropeta D (2009). The effect of individual N-glycans on enzyme activity. *Bioorg Med Chem* **17**: 2645-2653.
- Sneitz N, Bakker CT, de Knecht RJ, Halley DJ, Finel M and Bosma PJ (2010). Crigler-Najjar syndrome in The Netherlands: identification of four novel UGT1A1 alleles, genotype-phenotype correlation, and functional analysis of 10 missense mutants. *Hum Mutat* **31**: 52-59.
- Sorich MJ, Miners JO, McKinnon RA and Smith PA (2004). Multiple pharmacophores for the investigation of human UDP-glucuronosyltransferase isoform substrate selectivity. *Mol Pharmacol* **65**: 301-308.
- Steiner DF and Oyer PE (1967). The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma. *Proc Natl Acad Sci U S A* **57**: 473-480.
- Stingl JC, Bartels H, Viviani R, Lehmann ML and Brockmoller J (2014). Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther* **141**: 92-116.
- Stowers AW, Chen LH, Zhang Y, Kennedy MC, Zou L, Lambert L, Rice TJ, Kaslow DC, Saul A, Long CA, Meade H and Miller LH (2002). A recombinant vaccine expressed in the milk of transgenic mice protects Aotus monkeys from a lethal challenge with Plasmodium falciparum. *Proc Natl Acad Sci U S A* **99**: 339-344.
- Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurles ME and Dermitzakis ET (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* **315**: 848-853.
- Sugatani J, Uchida T, Kurosawa M, Yamaguchi M, Yamazaki Y, Ikari A and Miwa M (2012). Regulation of pregnane X receptor (PXR) function and UGT1A1 gene expression by posttranslational modification of PXR protein. *Drug Metab Dispos* **40**: 2031-2040.
- Sun D, Chen G, Dellinger RW, Sharma AK and Lazarus P (2010). Characterization of 17-dihydroexemestane glucuronidation: potential role of the UGT2B17 deletion in exemestane pharmacogenetics. *Pharmacogenet Genomics* **20**: 575-585.

- Team" RDC (2008). R: A Language and Environment for Statistical Computing. **R Foundation for Statistical Computing**: <https://www.R-project.org>
- Thomsen M and Nielsen M (2012). Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Res* **40**: W281-287.
- Thum T, Erpenbeck VJ, Moeller J, Hohlfeld JM, Krug N and Borlak J (2006). Expression of xenobiotic metabolizing enzymes in different lung compartments of smokers and nonsmokers. *Environ Health Perspect* **114**: 1655-1661.
- Tourancheau A, Margaillan G, Rouleau M, Gilbert I, Villeneuve L, Levesque E, Droit A and Guillemette C (2016). Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J* **16**: 60-70.
- Treutlein B, Gokce O, Quake SR and Sudhof TC (2014). Cartography of neuexin alternative splicing mapped by single-molecule long-read mRNA sequencing. *Proc Natl Acad Sci U S A* **111**: E1291-1299.
- Turgeon D, Carrier JS, Levesque E, Hum DW and Belanger A (2001). Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. *Endocrinology* **142**: 778-787.
- Wang C, Sadhavisvam S, Krekels EH, Dahan A, Tibboel D, Danhof M, Vinks AA and Knibbe CA (2013). Developmental changes in morphine clearance across the entire paediatric age range are best described by a bodyweight-dependent exponent model. *Clin Drug Investig* **33**: 523-534.
- Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP and Burge CB (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* **456**: 470-476.
- Wang Q, Zhao L, Liang M, Dong Y, Yun W, Qiu F, Meng H and Guo Y (2016). Effects of UGT2B7 Genetic Polymorphisms on Serum Concentrations of Valproic Acid in Chinese Children With Epilepsy Comedicated With Lamotrigine. *Ther Drug Monit* **38**: 343-349.
- Wegner I, Edelbroek P, de Haan GJ, Lindhout D and Sander JW (2010). Drug monitoring of lamotrigine and oxcarbazepine combination during pregnancy. *Epilepsia* **51**: 2500-2502.
- Wegner I, Wilhelm AJ, Sander JW and Lindhout D (2013). The impact of age on lamotrigine and oxcarbazepine kinetics: a historical cohort study. *Epilepsy Behav* **29**: 217-221.
- Wells PG, Mackenzie PI, Chowdhury JR, Guillemette C, Gregory PA, Ishii Y, Hansen AJ, Kessler FK, Kim PM, Chowdhury NR and Ritter JK (2004). Glucuronidation and the UDP-glucuronosyltransferases in health and disease. *Drug Metab Dispos* **32**: 281-290.
- Wong YH, Lee TY, Liang HK, Huang CM, Wang TY, Yang YH, Chu CH, Huang HD, Ko MT and Hwang JK (2007). KinasePhos 2.0: a web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res* **35**: W588-594.
- Wu B, Wang X, Zhang S and Hu M (2012). Accurate prediction of glucuronidation of structurally diverse phenolics by human UGT1A9 using combined experimental and in silico approaches. *Pharm Res* **29**: 1544-1561.



- Yu L, Qian M, Liu Y, Yao T and Zeng S (2010). Stereoselective metabolism of propranolol glucuronidation by human UDP-glucuronosyltransferases 2B7 and 1A9. *Chirality* **22**: 456-461.
- Yuan L, Qian S, Xiao Y, Sun H and Zeng S (2015). Homo- and hetero-dimerization of human UDP-glucuronosyltransferase 2B7 (UGT2B7) wild type and its allelic variants affect zidovudine glucuronidation activity. *Biochem Pharmacol* **95**: 58-70.
- Zaya MJ, Hines RN and Stevens JC (2006). Epirubicin glucuronidation and UGT2B7 developmental expression. *Drug Metab Dispos* **34**: 2097-2101.
- Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keefe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, Deng S, Liddelow SA, Zhang C, Daneman R, Maniatis T, Barres BA and Wu JQ (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**: 11929-11947.