Pharmacogenomics of human uridine diphospho-glucuronosyltransferases (UGTs) and clinical implications

Chantal Guillemette[1,3], Éric Lévesque[2] and Michèle Rouleau[1]

Pharmacogenomics Laboratory, Centre Hospitalier de l'Université Laval (CHU) de Québec Research Center and [1]Faculty of Pharmacy and [2]Faculty of Medicine of Laval University, G1V 4G2, Quebec, Canada; [3]Canada Research Chair in Pharmacogenomics.

Short Title: UGT pharmacogenomics in drug metabolism and diseases

Key words: Glucuronidation, *UGT*, Pharmacogenomics, Drug metabolism, Drug toxicity, Polymorphisms, Cancer susceptibility and progression, Alternative splicing, Epigenetic.

Author to whom correspondence should be sent:
Chantal Guillemette, Ph.D.
Canada Research Chair in Pharmacogenomics
Pharmacogenomics Laboratory
CHU de Québec Research Center
2705 Boul. Laurier, R4720
Québec, Canada, G1V 4G2
Tel. (418)654-2296; Fax.(418)654-2761
E-mail: Chantal.Guillemette@crchul.ulaval.ca

ABSTRACT

Glucuronidation, mediated by UDP-glucuronosyltransferase enzymes (UGTs), is a major phase II biotransformation pathway and, complementary to phase I metabolism and membrane transport, one of the most important cellular defense mechanism responsible for the inactivation of therapeutic drugs, other xenobiotics and numerous endogenous molecules. Individual variability in UGT enzymatic pathways is significant and may have profound pharmacological and toxicological implications. Several genetic and genomic processes are underlying this variability and are discussed in the context of drug metabolism and diseases such as cancer.

AN OVERVIEW OF THE GLUCURONIDATION PATHWAY, UGT ENZYMES AND THEIR CONTRIBUTION TO DRUG METABOLISM

The glucuronidation system, discovered more than 60 years ago by the pioneering work of Prof. Geoffrey J. Dutton, is a major phase II biotransformation reaction taking place in most human tissues. Glucuronic acid conjugation by uridine diphospho-glucuronosyltransferase enzymes (UGTs; EC 2.4.1.17) offers protection against dermal, airway and ingestion exposures to foreign chemicals.[1] UGTs are key specific enzymes governing drug metabolism, inactivating nearly 35% of current drugs (**Table 1**).[2] UGTs also participate in the homeostasis of a number of endogenous compounds, thereby contributing significantly to constitutive cellular metabolic pathways.

Glucuronidation occurs for a wide variety of lipophilic endogenous molecules, environmental compounds and synthetic drugs. The enzymatic reaction consists in the transfer of the sugar

moiety of the widely abundant co-substrate uridine 5′-diphosphoglucuronic acid (UDPGA) at multiple sites, including carbonyl, carboxyl, sulfuryl, hydroxyl, and amine groups (**Figure 1**). The most common glucuronidation reactions consist of *O*- and *N*-glucuronidation, occurring through conjugation of aliphatic alcohols, phenols, carboxylic acids, thiols and amines (primary, secondary and tertiary).

A large diversity of structurally unrelated endogenous substrates and medications commonly used in clinical practice are subjected to glucuronidation. Endogenous substrates include bilirubin, sex-steroids, thyroid hormones, bile acids and fat-soluble vitamins. Synthetic drugs include medications from most therapeutic classes (**Table 1**). Conjugation of these lipophilic molecules with the polar sugar glucuronic acid typically leads to their inactivation and increased water solubility. It enables the recognition of the resulting metabolites by hepatic and renal anion transporters such as hepatocellular organic anion-transporting polypeptides (OATPs) and multidrug resistance-associated protein 2 (MRP2) which is mandatory for their excretion from the body, either through urine (generally high molecular weight compounds) or bile (generally low molecular weight compounds). Some notable exceptions exist, in particular the conjugation of morphine into the active and more potent analgesic morphine-6-*O*-glucuronide metabolite, acting similarly to morphine on the μ-opioid receptor in the central nervous system. This conjugation is associated with the clinical effects of morphine therapy and the development of central side effects.[3, 4]

In clinical settings, the activity of UGT enzymes has a critical impact on first-pass metabolism of drugs and their bioavailability. In addition, glucuronidated drugs excreted in bile are susceptible

to deconjugation by bacterial β-glucuronidases. This process, which releases free drug in the intestine that may be reabsorbed and undergo entero-hepatic cycling, has significant implications for some glucuronidated drugs, such as the anticancer agent irinotecan and mycophenolic acid (MPA) discussed below. Glucuronidation also results in the detoxification of a high number of other xenobiotics including numerous carcinogenic chemicals and endocrine disruptors making this system one of the most important cellular defense mechanism.

In recent years, a large body of work has highlighted numerous factors considerably influencing drug, carcinogen and hormone glucuronidation reactions, including age, sex, nutrition, environmental exposure, pathological conditions as well as genetic determinants. These aspects have profound pharmacological and toxicological implications and influence susceptibility to multiple diseases and their clinical evolution. We herein discuss current and emerging molecular mechanisms having an influence on glucuronidation pathways in humans, focussing on their clinical impact on drug metabolism and diseases such as cancer.

Human UGTs

To conjugate this structurally diverse and broad range of lipophilic molecules, human *UGT* loci have evolved to encode several distinct proteins arising from exon or gene duplications, but also from alternative promoter usage and various splicing events. Four subfamilies of UGT proteins (UGT1A, UGT2, UGT3 and UGT8) share a common UGT sequence signature of about 50 amino acids (**Figure 2**).[5] Glucuronidating enzymes have been classified in two major families, namely UGT1A and UGT2, the latter being further subdivided in UGT2A and UGT2B (http://www.flinders.edu.au/medicine/sites/clinical-pharmacology/ugt-homepage.cfm, accessed

April 15, 2014). Glucuronidation by UGT3 and UGT8 subfamily members does not appear to contribute significantly to drug inactivation and will therefore not be discussed herein.

The genomic structure of human UGTs is organized to optimize and coordinate the broad expression of 19 enzymatically active UGT proteins with distinct but sometimes overlapping substrate specificity. The substantial contribution of UGT1A and UGT2B enzymes in drug metabolism predominates whereas for UGT2A it is still uncertain. The human liver certainly expresses by far the widest variety of UGT enzymes, 13 being co-expressed in this organ at different levels (**Figure 3**). UGT enzymes are also significantly expressed in several other metabolizing tissues relevant to first-pass metabolism, enterohepatic cycling and the bioavailability of many drugs, such as the intestine and kidneys. An overview of the isoenzymes expressed in these drug metabolizing tissues as well as their relative expression levels is provided in **Figure 3**. UGTs display tissue-specific distribution in many other peripheral tissues, namely lungs, skin and those related to steroid hormones, also contributing to local drug metabolism in these tissues. Expression levels of most UGT enzymes are under the control of several specific transcription factors (TFs), including the hepatic nuclear factors HNF1α and HNF4α. Ligand-activated TFs/nuclear receptors (NRs), including the aryl hydrocarbon receptor AhR, the peroxisome proliferator-activated receptor PPARγ, the pregnane X receptor PXR and the constitutive androstane receptor CAR, further modulate UGT expression in response to endo- and xenobiotics. For instance, NR-mediated autoregulation of UGT expression by endobiotics such as steroid hormones and bile acids greatly contribute to cellular homeostasis while NRs are also involved in the first line modulation of responses towards many xenobiotics.[6,7] Transcriptional regulation has been particularly well documented for hepatic expression.

Because of the high level of sequence identity among UGT enzymes (up to 95% amino acid identity), the lack of a characterized tertiary structure for full-length proteins and a clear definition of the active site, significant limitations associated with *in vitro-in vivo* extrapolation, and limited availability of computational models, it remains highly challenging to predict drug glucuronidation *in vivo* and foresee which UGT (and potentially multiple UGTs) participates in the metabolism of a given compound partly or primarily cleared by glucuronidation in humans.

UGT1A family members

The UGT1A family members are all encoded by a single gene locus on chromosome 2q37.1. The human *UGT1A* gene well exemplifies the usage of alternative promoters and first exons to increase protein diversity. Indeed, this complex locus consists of tandem arrayed variable first exons (n=13) that are alternatively spliced to four common exons (**Figure 2**). From the 13 possible mRNA isoforms, nine conduct to functionally active enzymes (UGT1A1, 1A3, 1A4, 1A5, 1A6, 1A7, 1A8, 1A9 and 1A10) and four are pseudogenes (UGT1A2p, UGT1A11p, UGT1A12p, and UGT1A13p).[8] More recently, one additional terminal exon, named exon 5b, was identified within the common region of the gene in the intron 4 (**Figure 2**).[9] Inclusion of this novel exon 5b in place of the classical exon 5a generates nine shorter additional protein products (referred to as UGT1A isoforms 2 or i2s) with negative regulatory functions upon glucuronidation rates through protein-protein interactions with UGT1A enzymes.[10] The discovery of these alternative splicing events has shed light on one of few emerging mechanisms likely implicated in the control UGT-mediated glucuronidation and drug glucuronidation (as discussed below).

The first exons and their associated promoter regions are distinct for each of the *UGT1A* mRNA variants, and provide the basis for substrate specificity and functional diversity of UGT1A enzymes (**Figure 2**). It allows a precise expression of each UGT1A protein by selective activation of their associated promoter in a tissue-specific manner. This unique genomic organization, observed for only few human genes such as the human leukocyte antigen loci (HLA), has not been observed for other drug metabolizing enzymes (DMEs). While UGT1As are mostly produced in the liver and numerous extrahepatic organs, some are exclusively encountered in extrahepatic tissues, namely UGT1A7, UGT1A8 and UGT1A10 (**Figure 3; Table 1**).

This family of enzymes is well recognized to conjugate several xenobiotics, such as drugs and environmental carcinogens but are also involved in the metabolism of several endogenous compounds, namely bilirubin, bile acids, serotonin and sex-steroid hormones. The most studied UGT1A enzyme in the context of drug metabolism is certainly the bilirubin-conjugating enzyme UGT1A1 that also conjugates a significant number of pharmacological compounds. UGT1A3, UGT1A4, UGT1A6 and UGT1A9 are also among chief UGTs involved in the metabolism of numerous drugs. Collectively, the latter UGT1A enzymes conjugate approximately 55% of known glucuronidated drugs. An overview of common drugs conjugated by specific UGT enzymes is provided in **Table 1**.

UGT2 family members

The human UGT2 family members are encoded by multiple duplicated genes clustered on

chromosome 4q13. In contrast to UGT1A proteins, most UGT2 enzymes are encoded by distinct genes. UGT2A1 and UGT2A2 mRNAs are transcribed from the same *UGT2A* locus by the use of alternative first exons, as demonstrated for the *UGT1* locus, while UGT2A3 is encoded by a distinct gene.[11, 12] All seven enzymes of the UGT2B subfamily (UGT2B4, UGT2B7, UGT2B10, UGT2B11, UGT2B15, UGT2B17 and UGT2B28) are encoded by distinct genes that share high sequence identity due to multiple gene duplication events (**Figure 2**). Genomic duplication is highlighted by the fact that UGT2B15 and UGT2B17 share more than 97% nucleotide sequence identity and the presence of at least five pseudogenes interspersed among the *UGT2B* genes within 750 kb of highly similar genetic sequences (**Figure 2**). The genomic structure of *UGT2* genes is composed of 6 distinct exons, with the first two exons contributing to the putative substrate-binding domain (corresponding to exon 1 of UGT1A) and thus displaying the lowest degree of sequence similarity among UGT2B enzymes. The other coding exons are translated into the conserved UDPGA-binding and transmembrane domains (**Figure 2**). In the last years, this common structural gene architecture has been revised especially for *UGT2B4* and *UGT2B7*, revealing numerous additional coding exons and several novel UGT proteins (see section below).[13, 14]

UGT2B enzymes are potentially major determinants of responses to several xenobiotics and in the metabolism of common drugs used in the clinic, most particularly UGT2B7 (discussed below) that is involved in the clearance of approximately 25% of common medications. Other UGT2Bs involved in the inactivation and excretion of pharmacological compounds are UGT2B4 as well as UGT2B15 and UGT2B17 (**Table 1**). Interestingly, two of the *UGT2B* genes, UGT2B17 and UGT2B28, are among the most commonly deleted genes of the human genome,

inducing significant changes in the pharmacokinetics of drug substrates and subsequent therapeutic consequences.[15,16] In addition, UGT2B enzymes play a critical role in steroid inactivation and alteration of these pathways affect numerous endocrine-related diseases such as cancer, as discussed below.

PHARMACOGENOMICS OF UGTS AND DRUG RESPONSE

Like many other DMEs, UGTs are highly polymorphic in humans and affected by common inherited polymorphisms and copy number variations (CNVs), in addition to rare genetic alterations at the origin of single-gene disorders. The genetic alterations present divergent frequencies among ethnic populations and surely contribute as a source of phenotypic differences with the potential to translate into variable drug clearance and response. A portray of genetic polymorphisms across most human UGTs can be found at http://www.pharmacogenomics.pha.ulaval.ca/cms/ugt_alleles/. Of those, a number of non-synonymous polymorphisms and regulatory variants have been reported at the *UGT1* locus and *UGT2* genes and are summarized in **Figure 4**.

Several reports of distinct drug-reaction profiles associated with UGT common germline variations have highlighted the relevance of studying UGT polymorphisms to predict drug exposure, toxicity and response. In patients, germline information in UGTs may play a role in optimizing the dose and selection of therapy. Thus far, studies have addressed whether inherited differences in drug metabolism caused by genetic polymorphisms in *UGT* genes contribute to variable clearance of drugs and revealed a broad spectrum of effect sizes. A recent publication

provides a systematic review of pharmacokinetic parameters of many drugs biotransformed by polymorphic UGT enzymes[2], and two examples for which a wealth of data is available are discussed below. However, much less studies have clearly addressed how the genetic effect within a UGT pathway affects outcomes of drug therapy (both in terms of efficacy and toxicities). Because of the paucity of data, a lot of work remains to be done to provide strong evidence of specific effects of UGT polymorphisms on altered drug responses that would require a dosage adjustment or a change in therapy, and subsequent use of UGT polymorphisms for the individualization of drug therapy. Few examples are emerging and are discussed in the following sections to illustrate that genetic determinants in *UGT* genes meaningfully modify drug response, disease risk and progression.

The widely used chemotherapeutic drug irinotecan

In clinical settings, the most discussed example relating *UGT* germline genetic variations with drug-induced toxicities pertains to *UGT1A1* polymorphisms associated with the metabolism of the anticancer agent irinotecan (CPT-11 or 7-ethyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothecin). Irinotecan is a chemotherapeutic drug clinically used in combination with 5-fluorouracil and leucovorin (FOLFIRI) in the treatment of metastatic colorectal cancer (CRC) and also administered to treat other solid tumors. One drawback of irinotecan treatment is its unpredictable associated severe diarrhea and hematologic toxicities, which vary greatly among patients. Irinotecan is a prodrug biotransformed into the pharmacologically active metabolite, SN-38, which undergoes extensive glucuronidation (>70%) to form the corresponding inactive glucuronide (SN-38G). Several UGT1As are involved in this reaction,

predominantly the UGT1A1 enzyme highly expressed in the liver (**Figure 5**). SN-38 is responsible for the anticancer effect but also severe toxicity and as such, factors modulating negatively its inactivation by the UGT pathway and its elimination from plasma will predispose to drug-induced hematological toxicity and potentially influence treatment efficacy. Conversely, enhanced UGT1A activity will favour delivery of SN-38G to the gastrointestinal tract via biliary excretion and may predispose to diarrhea, as resident intestinal bacteria β-glucuronidase activity will release active SN-38 in these tissues. Accordingly, UGT1A activity in the liver but also in extrahepatic tissue such as the intestine may be relevant to the development of irinotecan-associated toxicities.

In the general population, the interindividual variation in UGT1A1 expression is partly explained by a common polymorphic alteration in the atypical TATA-box region of the *UGT1A1* gene (**Figure 4**). The most common and reference allele *UGT1A1*1* contains six TA repeats, whereas the principal common variant allele *UGT1A1*28* (rs8175347) contains seven TA repeats and less frequent variants comprise five (*UGT1A1*33*) or eight (*UGT1A1*34*) TA repeats. Increasing number of TA repeats leads to a reduced expression of the *UGT1A1* gene. [17] Most studies have investigated the *UGT1A1*28* allele and revealed an association with reduced SN-38 glucuronidation rates and an approximately 2-fold higher risk of severe neutropenia induced by irinotecan, helping the identification of those patients who would benefit from reduced doses of the drug (recently reviewed by Barbarino and collaborators).[18] Notably, some regulatory bodies, including the U.S. Food and Drug Administration, have recommended that reduced initial dose of irinotecan should be considered in patients homozygous for *UGT1A1*28* to minimize drug exposure and risk of toxicities. Overall, there is a strong biological rationale and reproducible

evidence linking *UGT1A1*28* allele and dose-limiting toxicities, and a number of studies supporting that genotyping of the *UGT1A1*28* allele prior to irinotecan administration in CRC patients would be cost-effective.[19] This marker thus falls into the category of pharmacogenetic markers associated with chemotherapeutic drug-related toxicity. However, the evidence of an impact on drug efficacy is still lacking.

It is anticipated that for some drugs, the efficacy and adverse drug reactions will not only differ between individuals but also across different populations since some *UGT1A1* polymorphisms are exclusively or mostly observed in particular ethnic populations, supporting an environmental pressure in sequence selection. For instance, the frequency of the low promoter activity *UGT1A1*28* allele varies significantly with a frequency of approximately 0–3% of the Asian population, 2–13% of the Caucasian population and up to 16–19% of Africans whereas the *UGT1A1*33* and *UGT1A1*34* alleles are predominant in Africans. Other variants such as the *UGT1A1*6* variant affecting the enzyme sequence (Arg71Gly; rs4148323) predominates in Asians and is linked to an increased incidence of severe neutropenia in Asian cancer patients in both high/medium and low doses of irinotecan (**Figure 4**).[20, 21] Considering the large degree of diversity in *UGT* genes across human populations, a careful evaluation of the haplotype structure is also highly relevant to drug metabolism and the evaluation of drug safety and efficacy. Other genotypes both in UGT1A1 and in additional UGT1A enzymes are likely to improve prediction of risk and exposure to SN-38, especially when observed in combination. This is sustained by the fact that SN-38 is metabolized by several other UGT1A enzymes, namely UGT1A7 and UGT1A9 expressed respectively in extrahepatic (GI tract) and hepatic tissues (**Figure 5B**). However, few variations in these other enzymes were consistently associated with severe

toxicities across populations while still very few studies examined their associations with severe toxicity in combination with the *UGT1A1* variant alleles. For instance, the low activity coding variant *UGT1A7*3* comprising three non-synonymous variations (N129K, R131K and W208R) was associated with lower incidence of severe diarrhea and an increased risk of neutropenia alone and in combination with the *UGT1A1*28* allele.[22] Similarly, a recent prospective study suggested that haplotype analyses of *UGT1A7* and *UGT1A9* markers along with a *UGT1A* SNP in the 3' untranslated regions (3'UTR; common to all UGT1A enzymes) allowed the identification of a protective haplotype associated with a 2-fold lower risk of severe neutropenia as well as higher risk characterized by 2 and 3 other unfavorable *UGT1A* alleles revealing a dosage effect (odd ratios from 2.15 to 5.28).[23] It is thus likely that the consideration of additional *UGT* alleles will exceed the predictive value of *UGT1A1*28* itself in the context of irinotecan-related chemotherapy. It remains however challenging to pinpoint the causal variant(s) because *UGT1A* variants, although likely functionally relevant to SN-38 glucuronidation, are to some degree frequently in close linkage disequilibrium and further display variable haplotypic structures among ethnic groups.[24, 25] There is thus a requirement to pursue comprehensive haplotype studies including variants across the *UGT1A* loci to refine pharmacogenetic testing, to provide functional data for novel markers in order to better predict metabolizer phenotype and also consider ethnic specific and *UGT1A* haplotype structures for future clinical applications.

Pretreatment selection according to genetic analysis is conceivable to achieve better clinical benefits while minimizing toxicity. However, clinically, little is known about the optimal strategy for managing patients with *UGT1A* variance and definitive therapeutic recommendations based on *UGT1A* actionable genotypes are still awaited. As for many other pharmacogenes, the

cumulative level of evidence rather than randomized clinical trials will likely provide the basis for the therapeutic recommendations associated with UGT genotypes.[26] Recent research efforts have been targeted to UGT genotype-guided clinical studies. At least five studies have been published with three examining Asian patients and two focusing on European patients. Findings support that doses as high as 370 mg/m^2 and 310 mg/m^2 could be tolerated for homozygous *1/*1 and*1/*28 heterozygous carriers, respectively and a dosage lower than 150 mg/m^2 (between 100 to 130 depending on the study) could be administered for *UGT1A1*28* homozygous carriers.[27, 28] Besides, significant challenges remain regarding whether response is maintained with dosage adjustment in *UGT1A1*28* homozygous patients. Likewise, patients with favourable *UGT1A* 3'UTR alleles associated with a significant lower risk of severe toxicity might benefit from increased irinotecan dosing potentially improving response rates and survival. Irinotecan is therefore one of the most documented examples of a meaningful UGT effect on drug clearance and drug-related toxicity. It is expected that findings in the context of UGTs and irinotecan will serve as a basis for other pharmacogenomics studies for drugs predominantly metabolized by *UGT1A* gene products.

The frequently used immunosuppressant MPA

Another example of clinical relevance pertains to one of the most frequently used immunosuppressive drugs for the prevention of acute and chronic rejections among hematopoietic stem cell and solid organ transplant patients, mycophenolic acid (MPA). Derived from the inactive mycophenolate mofetyl (MMF) prodrug of MPA and the enteric-coated mycophenolate sodium salt, MPA presents a narrow therapeutic index and significant inter-

individual variability in blood concentrations. This drug, given at a fixed dose, is associated with adverse effects including gastrointestinal and bone marrow toxicity and infection. Most studies support a relationship between MPA pharmacokinetic parameters and clinical outcome, namely in relation to acute rejection and drug-related toxicity. Therapeutic monitoring has been applied particularly in renal transplantation, but how to best measure MPA exposure is still a subject of debate.

The glucuronic acid conjugation reaction by UGT1A and UGT2B7 enzymes mediates a significant proportion of the total clearance of MPA. It occurs in the liver, intestine, and kidney to yield two glucuronide metabolites excreted predominantly in bile. A major inactive and more water-soluble phenolic 7-O-glucuronide (MPAG) accounts for 95% of the total metabolic elimination pathway and undergo significant enterohepatic recirculation along with MPA. A less abundant (5%) acyl-glucuronide (AcMPAG) has been shown to have some pharmacologic and toxicologic effects, seemingly through direct proinflammatory and proliferative effects as well as binding covalently to proteins and macromolecules, which might explain the immunotoxicity conferred by this immunosuppressive drug.

Mainly, three UGT isoforms are involved in the metabolism of MPAG and AcMPAG, namely the hepatic and extrahepatic enzyme UGT1A9 is responsible for MPAG formation, along with a contribution of the extrahepatic UGT1A8 isoenzyme. The UGT2B7 enzyme, expressed in liver and extrahepatic tissues, and to a minor extent UGT1A8, mediates AcMPAG formation. One of the mechanisms underlying differences in MPA exposure related to germline variations of the UGT1A and UGT2B7 pathways were recently reviewed. [29]The potential clinical significance of UGT genetic variability is particularly exemplified by several studies supporting a relationship

between polymorphisms in the *UGT1A9* promoter, namely variations at positions -275T>A and -2152C>T associated with increased gene expression, and significantly lower MPA plasma concentrations in healthy subjects and renal transplant recipients. Consistent with the immunosuppressant effects of MPA exposure, some studies also reported an increased risk of acute rejection for -275/-2152 carriers. Opposite effect was observed for the low activity *UGT1A9*3* (98T>C; M33T) allele, such that carriers may benefit receiving a lower MPA dose.[2]Two coding variants of UGT1A8, corresponding to the *2 (A173G) and *3 (C277Y) alleles, were also evaluated but inconclusive and conflicting results were reported on the relationship with MPA and MPA glucuronide exposure. As for UGT2B7, a few studies have examined *UGT2B7* polymorphisms focusing on the tightly linked promoter variants (-842G>A and -79G>A) and the nonsynonymous variation at codon H268Y (*UGT2B7*2* allele), An increased AcMPAG exposure but not to MPA was associated with these variations in some of theses studies.

Overall, not all studies consistently observed an association between UGT variations and MPA pharmacokinetic parameters and little evidence is available regarding a meaningful link with transplantation outcome such as acute rejection and drug-related toxicities. However, most studies were small, underpowered and tested a short list of variations whereas these associations may vary depending on SNP frequency, , time after transplant and the prescribed concomitant drug that affects MPA pharmacokinetics. Indeed, co-administration of other immunosuppressive agents, such as cyclosporine, tacrolimus or corticosteroids, influences MPA exposure through an interaction or modulation of drug metabolism and transporters, affecting enterohepatic circulation of MPA and its glucuronides and hence modifying drug exposure. More detailed

genetic profiling of the different UGT enzymes, along with drug transporters and other biologically relevant candidate genes, as well as their influence on MPA, MPAG and AcMPAG pharmacokinetics, need to be better defined prospectively to translate findings into clinical recommendations, and in different transplant populations for most clinically used immunosuppressive regimens involving MPA.

UGTs AS DISEASE SUSCEPTIBILITY GENES AND PROGNOSTIC MARKERS

Several diseases or conditions are linked to variations in the *UGT* genes independently of drug metabolism. Because the UGT pathway is integral to the biotransformation of numerous carcinogens, industrial chemicals, dietary components as well as hormones, several factors affecting the *UGT* genes including genetic polymorphisms and CNVs, have been shown to influence disease predisposition and even prognosis (**Table 2**). Relevant to cancer prevention, studies have identified specific *UGT* variants as low-penetrance susceptibility genes for various types of cancers. Unrelated to drug treatment, the evidence is also emerging regarding genetic factors in *UGT* genes or altered UGT expression levels being linked to cancer aggressiveness or overall outcome of the disease, independently of clinico-pathological characteristics. Prognostic information related to UGTs may be clinically relevant since it can be used to determine a particular or individualized treatment plan.

Genetic variants in UGTs may increase or protect from cancer by respectively decreasing or increasing the inactivation of procarcinogens and carcinogenic molecules. At present, cancers most studied in the context of UGT-linked risks relate to hormonal exposure such as prostate, breast, uterine and ovarian cancers, as well as cancers linked to carcinogen exposure including

head and neck, bladder, oesophageal, lung and colorectal cancers (**Table 2**). Most data arise from candidate gene studies with few more recent genome-wide association studies also identifying *UGT1A* and *UGT2* genes as novel cancer susceptibility loci involved in tumour development.[30-37] Overall, most studies were from case-control study design and have confirmed a significant link between *UGT* gene polymorphisms and the risk of cancer. Relative risks reached 0.2 for protective genotypes and over 2-fold for risk genotypes, while other studies have failed to find significant associations. Disparities have been most particularly observed between ethnic groups and are likely explained by differences in genetic background (as for drug metabolism), and environmental exposures. Environmental factors broadly are defined as endogenous or exogenous risk factors such as levels of sex steroid hormones, cigarette smoking and meat-related mutagen exposure. Sometimes, the lack of detailed and valid assessments of genes and environmental interactions has limited the impact of some studies on UGTs and cancer predisposition. Frequently, limited sample size has also precluded the detection of any small effect expected to be caused by *UGT* polymorphisms and resulted in negative studies or positive associations that were not subsequently replicated by others. Overall, validation of the association of *UGT* polymorphisms, individually or in combination, with cancer risk for different types of cancers is clearly warranted by using larger, well-designed studies from different ethnic populations. Overall, findings may have implications for cancer treatment and prevention.

UGT2B17 gene as an example: clinical implications of copy number variation and its expression level

The occurrence of a large deletion in the entirety of the *UGT2B17* and *UGT2B28* genes has been

demonstrated as two of the most commonly deleted genes in the genome. These deletions occur at variable frequencies among ethnic groups, with over 50% of Caucasians and 75% of Asians having a reduced copy number in at least one of these genes. This inactivating genomic mechanism definitely provides a unique opportunity to study the impact of the UGT2B17 and UGT2B28 pathways in various clinical contexts. For instance, the *UGT2B17* gene deletion has a significant impact on steroid metabolism leading to altered circulating and tissular steroid hormone levels, doping in sports, osteoporosis and increased risk of fracture, in addition to be associated with higher risk of prostate cancer and recurrence after radical prostatectomy.[38-43] These observations are consistent with a meaningful role of the UGT pathway to the maintenance of intracellular levels of sex-steroid hormones in target cells such as in the prostate. Also, the link between *UGT2B17* CNV and prostate cancer is in line with the fact that prostate carcinogenesis is androgen-dependent, supporting the hypothesis that reduced androgen glucuronidation by UGT-mediated inactivation would lead to an increased exposure to active hormone and higher risk of cancer and relapse.[42, 44] Similarly, the *UGT2B28* gene deletion was associated with amenorrhea and hyperandrogenism, Addison disease and prostate cancer recurrence after prostatectomy.[40, 42, 44-46] Also, UGT2B17 is highly expressed in GI tract and a recent genetic association study supports its role in modifying colon cancer risk,[44] especially in men, consistent with the observation that men have higher UGT2B17 expression and activity than women.

Besides, the uniqueness of UGT2B17 is related to its emerging biological pleiotropic effects in minor HLA disparities and immune recognition in stem cell transplantation. In contrast to previously defined human minor H antigens, the immunogenicity of UGT2B17 lies in its differential expression in donor and recipient cells as a consequence of a homozygous gene

deletion in the donor.[47, 48] UGT2B17 thus possesses immunogenic properties (**Figure 6**) that predispose to a serious complication in transplant recipients, termed graft-vs-host disease, affecting mainly the liver and GI tract expressing UGT2B17. In addition to this seemingly non-enzymatic immune function, UGT2B17 is overexpressed in high-risk chronic lymphocytic leukemia (CLL).[49] UGT2B17 gene overexpression rather than CNV was prognostic of CLL outcomes and is suspected as a disease accelerator and a potential drug target. Moreover, the induction of *UGT2B17* gene expression was associated with resistance to fludarabine-containing regimens in a subset of CLL patients, which may further modify drug response. Consistent with this notion, UGT2B17 is a key enzyme for the inactivation of anticancer agents, namely suberoylanilide hydroxamic acid (SAHA), an oral histone deacetylase inhibitor used in the treatment of cutaneous T-cell lymphoma and in clinical trials for treatment of multiple other cancers. Individuals with the null UGT2B17 genotype present lower SAHA glucuronidation rates that may have clinical implications.[50] Overall, UGT2B17 is somehow unique since it has emerging implications in several detoxification, endocrine and immune processes whereas some of its biological function also appears independent of its enzymatic function. Whether these properties apply to other UGTs remain unknown.

EMERGING MECHANISMS AS A SOURCE OF VARIABILITY IN THE GLUCURONIDATION PATHWAY

Numerous recent investigations have uncovered a wide array of molecular mechanisms beyond genetic variations and NRs-mediated regulation of UGTs likely governing the heterogeneity in expression and glucuronidation activity of UGT enzymes. In this section, we discuss novel molecular aspects, including epigenetic and pre-mRNA alternative splicing mechanisms that

appear determinant to UGT expression levels, protein diversity and glucuronidation phenotypes (**Figure 6**).

Epigenetic-induced variability: DNA methylation and microRNAs

Epigenetics modifications, which refer to inheritable aspects of gene functions that are not specified by the genomic DNA sequence, include DNA methylation, histone modifications, chromatin architecture, and non-coding RNAs such as microRNAs (miRNAs). Epigenetic marks are reversible and modulated during development, by tissue-specific factors, age, sex, environmental factors and drug treatment.[51] In parallel, deregulation of epigenetic patterns is associated with induction and propagation of disease states.[52] Although epigenetic mechanisms are evolving as key processes contributing to wide interindividual variations in DMEs, transport and subsequent drug responses,[53] their contribution to variability in drug glucuronidation is only starting to be recognized.

DNA methylation, occurring predominantly on cytosines in the context of CpG dinucleotides, is a generally stable repressive mark associated with chromatin compaction, interfering with transcription factor (TF) binding. Relevant to UGTs, one of the first study assessing epigenetic marks on human *UGT* genes pinpointed a relationship between SN-38 inactivation by UGT1A1, hypermethylation levels of the *UGT1A1* gene in CRC patients and increased tumor cell sensitivity to SN-38.[54] Hypermethylated CpG sites at the *UGT1A1* promoter correlated with reduced binding of transcription factors such as HNF1α, which regulates *UGT* expression, and transcriptional repression of the *UGT1A1* gene. This highlights the potential influence of

UGT1A1 gene methylation status on clinical responses to drug therapy. Hepatic drug glucuronidation is likely also affected by methylation status in healthy individuals, as suggested by a recent analysis of 46 normal human livers.[55] UGT epigenetic profiles may thus serve as predictors of glucuronidation potential and explain interindividual differences in drug metabolism.

Besides, the mechanisms for the observed tissue-specific expression of certain UGTs are still unclear and DNA methylation and histone posttranslational modifications may certainly play a role. UGT1A1 is abundant in liver but undetectable in kidney while other UGTs are absent in liver but abundant in the GI tract such as UGT1A10. DNA methylation of *UGT1A1*, hypoacetylation of histone H3, and decreased binding of HNF1α are involved in differential tissue expression between liver and kidney cell lines.[56, 57] For instance, the *UGT1A10* promoter appears hypermethylated in hepatocytes and this process would interfere with the binding of the transcription factors HNF1α and Cdx2, resulting in the defective expression of UGT1A10 in human liver. In turns, the *UGT1A10* promoter is hypomethylated in the epithelium of the small intestine consistent with the reported high expression in this tissue. The *UGT* transcriptional reprogramming through changes in methylation and histone acetylation may lead to reactivation or repression of specific UGTs in a tissue- and cell-type specific manner and subsequently modulate local drug metabolism and potentially drug response.

MicroRNAs consist in short noncoding RNAs of approximately 22 nucleotides that regulate protein expression by binding to UTRs in corresponding mRNAs, resulting in translational repression or mRNA degradation. miRNAs contribute to many kinds of human processes and

diseases, including cancer, and their overexpression in certain tumors suggests candidate disease biomarkers and drug targets.[58] There has been increasing efforts to define miRNA-mediated regulation to DMEs, transporters and nuclear receptors, but our current knowledge is limited to disease-specific phenomena and a lack of understanding of its role in regulating human UGTs. While the assessment of the impact of miRNAs on UGT expression is in its infancy, current data support their significant contribution to the regulation of specific UGT expression, likely in a tissue-specific manner. Since an miRNA may target more than one gene, potential co-regulation of DMEs is possible, whereas a UGT gene might be regulated by multiple miRNAs that may act synergistically. A first experimental evidence of a specific miRNA interfering with UGT expression was provided recently.[59] The miRNA miR-491-3p regulates the expression of some UGT1A enzymes in hepatic *in vitro* cellular models, by targeting a region located in the 3'UTR common to all UGT1A enzymes. However, the expression of the hepatic UGT1A4 and UGT1A9 was not altered by this miRNA, highlighting that factors other than target sequence will influence regulation by miRNA, such as mRNA secondary structure or SNPs in the target sequence and illustrating the complexity of this regulatory mechanism. Interestingly, miR-491-3p, while expressed in the liver, is much more abundant in extrahepatic tissues such as the colon, suggesting that it might play a prevalent role in regulating metabolism in these tissues. Another recent study identified 56 hepatic miRNAs predicted to target the 3' UTR of DMEs and assessed their relationship with putative target gene expression.[60] For instance, the expression of 7 miRNAs inversely correlated with that of UGT2B7 and UGT2B17 in a collection of 92 human livers, suggesting these miRNAs as potential regulators, but they remain to be experimentally validated. miRNA-mediated regulation of UGTs is therefore a potential factor affecting their gene expression and subsequent drug metabolism, likely operating in a tissue-specific manner.

Moreover, indirect regulation of UGT expression by miRNAs may also occur, through modulation of transcription factors involved in UGT expression such as Nrf2, HNF4, HIF1α and PPARα.[58] It is therefore expected that future research will establish a role for microRNAs in altered drug glucuronidation. The clinical impact of this knowledge may be substantial in view of the modulation in miRNA expression patterns induced by drug treatments that may further affect individual drug responses.

Alternative splicing of UGT transcripts and expansion of the UGT transcriptome and proteome

Alternative splicing is a genetic process for controlling gene expression, increasing expressional flexibility and the complexity at the transcriptome and proteome level while diversifying cellular functions of alternate proteins. With recent genome-wide analyses based on next generation sequencing, questioning the extent of alternative splicing becomes achievable and is revealed as an important mechanism explaining transcriptome diversity. A comprehensive assessment of the UGT transcriptome is crucial for interpreting the functional elements of each *UGT* genomic regions and for considering their role in drug metabolism, diseases, and potentially in other cellular functions. Several examples have illustrated that mechanistic understanding of splice variants and their role in therapeutic resistance can lead to novel treatment ideas.[61] A number of alternative mechanisms have been reported for human *UGT* genes, in particular specific alternative promoters and exon usage.[9, 14] Considering the wide occurrence of gene-associated mRNA diversity, it is highly probable that the *UGT* loci have additional undescribed mRNA species. Current observations sustain that alternative promoter and/or exon usage may play a

substantive role in modulating the expression/activity of most human UGTs (e.g. UGT1A, UGT2A, UGT2B4, UGT2B17 and UGT2B28)[10] and that this process may influence drug metabolism and potentially explain some of the variability in the glucuronidation pathway observed in patients.

Alternative splicing of UGT1A mRNAs and modulation of metabolic pathways

Through the extensive analysis of human *UGT1A* genomic sequences, a novel class of human UGT proteins named isoforms 2 (or i2s) produced by alternative splicing was uncovered and involve a new exon in the common intron 4 of the gene, exon 5b (exon 5a being the canonical exon 5).[9, 62] The inclusion of exon 5b in the mature transcript leads to the production of nine shorter alternate UGT1A_i2 proteins. Exon 5b causes a premature end of translation, loss of the transmembrane domain and the inclusion of ten novel C-terminal amino acids (R_{435}KKQQSGRQM_{444}). The 45-kDa i2 proteins and the classical functional 55-kDa i1 enzymes are co-produced in the same tissue structures of the liver, kidney, stomach, intestine and colon. Both in tissues and model cell lines, they are co-expressed and co-localized within the ER. Functional analysis demonstrate that the shorter i2s lack glucuronidation activity and rather have a dominant-negative regulatory role, possibly by forming inactive heteromeric complexes with i1 enzymes with reduced glucuronidation rates. The extent of i2-mediated inhibition varies depending on drug substrates and isoenzymes, ranging from 41–82% in the context of i2s expression levels being below those of i1s, as predominantly observed in human tissues. Noteworthy, this degree of inhibition by i2s is similar to the impact of common genetic polymorphisms associated with significantly altered responses to endogenous compounds and xenobiotics, supporting a potential meaningful effect of i2s on drug metabolism *in vivo*. A recent

investigation demonstrated the impact of partially depleting endogenous i2 on drug bioavailability and cellular response to SN-38 in colon cancer cells.[63] A first evidence for the *in vivo* clinical relevance of these alternative i2 proteins was also reported. Enhanced acetaminophen glucuronidation and decreased risk of unintentional acetaminophen-induced acute liver failure is associated with a SNP (rs8330) in the *UGT1A* 3'UTR. This SNP, possibly by altering the binding site for Srp (splicing regulatory protein), is associated with increased exon 5a/5b expression ratios that would lead to an enhanced abundance of UGT1A enzymes compared to i2 regulators.[64] An assessment of i1/i2 expression in normal and cancer tissues further indicates a differential expression of these isoforms and suggests altered glucuronidation rates between normal and disease tissues.[65] An intriguing observation is the fact that alternate UGT1A_i2 proteins not only localize to ER in human tissue samples but also to the cytoplasm suggesting potential additional function(s) in this cellular compartment (**Figure 6**). Evidence that i2s would influence oxidative pathways, relevant to drug response were exposed by recent proteomics and cellular investigations. In particular, i2s interact and modulate the biological activity of the ROS scavengers and enzymatic partners' catalase and peroxiredoxin 1.[63] The broader cellular functions of UGT1A proteins bring to light a potentially wider interconnection of UGTs with other metabolic pathways, which may affect drug metabolism and response.

Alternative splicing of UGT2B7

The UGT2B7 enzyme plays a prominent role in drug metabolism and is involved in the metabolism of non-steroidal anti-inflammatory drugs, morphine, epirubicin and MPA (**Table 2**). UGT2B7 glucuronidation capacity varies up to 7-fold among individuals while no frequent (>5%) genetic component has been consistently recognized as contributing factor to this

variability. Most studies on *UGT2B7* genetic variants focused on the *2 (802C>T; His268Tyr; rs7439366) allele highly prevalent in the Caucasians population (~0.50), that has however little impact on substrate selectivity and activity. Two other major haplotypes comprised in the proximal promoter of the *UGT2B7* gene (-138A>G, rs7668258 and -900A>G, rs7438135) that may be further modulated by the presence of rare polymorphisms,[66] however, none of these genomic variants appear to explain a significant part of the high interindividual variability in *UGT2B7* gene expression.

Further experimental evidence clearly indicates inconsistencies in the metabolic fate of UGT2B7 substrates in relation with *UGT2B7* mRNA expression, while liver *UGT2B7* mRNA expression correlates modestly with UGT2B7-mediated activity, suggesting that posttranscriptional mechanisms might be involved.[67, 68] Indeed, the occurrence of extensive pre-mRNA alternative splicing events and 4 novel exons in the *UGT2B7* locus were recently uncovered and may contribute to the high interindividual UGT2B7-dependent glucuronidation activity.[14, 69] These novel splicing events clearly display tissue-specific patterns with a predominance in several extra-hepatic tissues, namely in the kidney. Two main categories of novel 2B7 products were observed, with variable N- and C-terminal ends, influencing *in vitro* zidovudine (AZT) drug glucuronidation activity. [70] Similar observations were described for UGT2B4 and novel alternate isoforms.[13] As reported for alternate UGT1A_i2 proteins, shorter UGT2B7 isoforms with different C-terminal ends repress UGT2B7 glucuronidation activity via protein-protein interactions with the UGT2B7 enzyme.[69] Conversely, a second group of UGT2B7 variants possess alternate 5' ends, owing to the use of alternative promoters and first exons.[70, 71] Their expression occurs at the expense of transcripts encoding the active UGT2B7 enzyme thereby

drastically affecting drug glucuronidation capacity.[70] Furthermore, expression of variants with different N-terminal ends is also associated with impaired UGT2B7 function in human tissues during prenatal life and in cancer. Regardless of the fact that these mRNAs may conduct to proteins, their expression clearly affect functional UGT expression and activity in cell models and tissues studied. These processes may thus considerably modify the glucuronidation potential across tissues and cells, and disease state. Overall, these observations clearly unveil another level of complexity in the regulation of intracellular glucuronidation activity that may have important physiological and pharmacological implications. Understanding the regulation of this complex transcriptional system will be of great interest in the near future and may underlie part of the interindividual glucuronidation variability currently unexplained by UGT single nucleotide polymorphisms or other non-genetic causes.

PERSPECTIVES AND CONCLUDING REMARKS

Glucuronidation of drugs and other xenobiotics by UGT enzymes is undoubtedly a major and essential component of drug metabolic pathways, conjugating a large fraction of clinically relevant drugs. Currently, UGT2B7, UGT1A1, UGT1A3, UGT1A4 and UGT1A9 are the hepatic UGT enzymes most actively involved in drug detoxification and elimination, conjugating nearly 70% of common drugs known to be glucuronidated (**Table 1**). With the evolving chemical properties of new drugs tending towards being larger, more lipophilic and more aromatic molecules, an increasing contribution of the glucuronidation pathway is clearly anticipated. [72, 73] Exploring the highly complex mechanisms underlying the regulation of UGT expression and activity is essential to better define key aspects and predict those most likely to have significant

impact on drug efficacy and side effects.

In the last decades, a wealth of research reports have greatly contributed to decipher important sources of variability in the glucuronidation rates of many drugs, and have uncovered a significant impact of several genetic variations (SNPs and CNV) at loci encoding these enzymes. The translation of these genetic polymorphisms into meaningful pharmacokinetics consequences, toxicities and bioavailability is starting to emerge but remains under evaluated for most drugs. However, it still remains largely speculative to infer *in vivo* relevance of specific genetic variations in one or multiple UGTs from *in vitro* data.[2] In addition, a better understanding of noncoding variations (located in the 5' and 3' UTR regions as well as intronic regions) is required but involves complex and labor-intensive functional studies to assess mechanism of action. Being able to quantify UGT enzymes and the activity of a specific UGT is also required to establish the relationship between genotypes and phenotypes. At this stage, a challenging task is to pursue the identification of specific probe substrates and biomarkers of specific UGT activity. Mass spectrometry-based quantification methods are emerging as novel tools to catalog expressed UGT proteins in a given human tissue.[74, 75] However, the high degree of shared sequence among the UGT enzymes, even more in cases of genetic or splicing variants of a single UGT, exposes some limits to these powerful approaches. Evidently, the high interindividual variability in glucuronidation activity remains unexplained in part by known genetic polymorphisms (as illustrated above for UGT2B7) supporting the prevalence of other contributing mechanisms and the possible existence of rare variants, as observed recently for drug target and transporter genes.[76, 77] These rare variants are predicted to have a larger effect than the common variants identified thus far, and may collectively explain at least in part, the

wide interindividual variation observed in glucuronidation phenotypes. On a population basis, this genetic information may have limited impact, but on a patient perspective, these data could be of primary importance if efficacy can be enhanced and toxicity significantly abrogated.

More recent research efforts have focused on a wide array of alternatively spliced variants expanding the UGT transcriptome and proteome. Whether extensive alternative splicing applies to all other UGTs remains to be demonstrated but is likely given that all *UGT* genes arise from gene/exon duplication clusters and share high sequence similarity. At present, researchers currently face to clearly target the appropriate transcripts to be quantified by better selecting amplification strategy to avoid further conflicting results. The occurrence of these unsuspected variants may have escaped detection and contribute to the lack of correlation between mRNA levels and glucuronidation activity previously reported.[67, 68] Besides, the functional impact of such diversity requires extensive characterization, especially to demonstrate their *in vivo* relevance, but already suggests multiple regulatory avenues, modulation of glucuronidation activity by dominant negative interactions being a new one (**Figure 6**). Several additional studies are required to capture genetic/genomics variables associated with interindividual glucuronidation capacity. Additional answers may come from RNA-seq experiments and/or trough resequencing outliers to pinpoint unidentified variants associated with major changes in specific glucuronidation phenotypes. These tasks are convoluted by the fact that UGT loci are duplicated regions with impressive sequence similarity among members of the same family complicating molecular studies. Globally, it seems clear that several genetic and genomic processes including common SNPs, rarer potential unidentified variants, epigenetic, alternative splicing and post-transcriptional regulatory mechanisms (such as protein-protein interaction)

affect glucuronidation phenotypes (**Figure 6**). Further understanding of the molecular processes is required to help reduce toxicities, enhance drug efficacy and help individualized therapy.

Connexions between genetic and epigenetic mechanisms occurring simultaneously are relevant in determining glucuronidation phenotypes. For instance, the functional mechanism of SNPs in the *UGT* 3'UTR region may be in some cases related to miRNA regulation by altering an miRNA-binding site in the 3'UTR region of a *UGT*.[78] Also, SNPs might affect splice site integrity,[79] whereas the interplay between alternative splicing and UGT SNPs has been reported.[64] Recent studies also point to a key function of chromatin structure and histone modifications in alternative splicing regulation. This reinforces the need to better study each of these individual mechanisms and to assess their interconnection and contribution to variable drug glucuronidation and their relevance to improve modeling of drug response and individualized therapy. These processes will help address issues of SNP function, unbalanced mRNA and protein levels, clarify some of the interindividual variability in glucuronidation and may lead to the improvement of prediction of drug metabolism and responses. Furthermore, the complex pharmacological profile of most drugs is not only dependent on UGTs but also on a host of enzymes and transporters involved in metabolic transformation, active transport proteins, intestinal absorption, and hepatobiliary secretion mechanisms. Modulation of any of these detoxification pathways through inhibition, induction, saturation, genetic polymorphisms and epigenetics modifications are likely to influence the concentration of glucuronides, affecting systemic drug exposure and efficacy. Such information is required to reach the ultimate goal of better-tailored pharmacological interventions and personalized medicine.

ACKNOWLEDGMENTS

The authors apologize for citing only a few of numerous highly relevant publications, due to space limitations. They would like to thank Véronique Turcotte for collecting the information regarding UGT drug substrates and the graduate students Mélanie Rouleau and Guillaume Margaillan in the Guillemette's lab for discussions. The excellent artwork by France Couture is also acknowledged. The Canadian Institutes of Health Research (CIHR grants MOP-42392 and MOP-84223), The Natural Sciences and Engineering Research Council of Canada, Prostate Cancer Canada and The Canada Research Chair Program (C.G.) supported some of the work discussed. C.G. holds a CIHR Canada Research Chair in Pharmacogenomics (Tier I). E.L. is a recipient of a CIHR clinician-scientist phase II award and Prostate Cancer Canada rising star award (RS2013-55).

Conflict of interest

The authors declared no conflict of interest

References

1. Tukey, R.H. & Strassburg, C.P. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu Rev Pharmacol Toxicol* **40**, 581-616 (2000).
2. Stingl, J.C., Bartels, H., Viviani, R., Lehmann, M.L. & Brockmoller, J. Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther* **141**, 92-116 (2014).
3. Ritter, J.K. Roles of glucuronidation and UDP-glucuronosyltransferases in xenobiotic bioactivation reactions. *Chem Biol Interact* **129**, 171-193 (2000).
4. Gretton, S.K. *et al.* Plasma morphine and metabolite concentrations are associated with clinical effects of morphine in cancer patients. *J Pain Symptom Manage* **45**, 670-680 (2013).
5. Mackenzie, P.I. *et al.* The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics* **7**, 255-269 (1997).
6. Bigo, C., Caron, S., Dallaire-Theroux, A. & Barbier, O. Nuclear receptors and endobiotics glucuronidation: the good, the bad, and the UGT. *Drug Metab Rev* **45**, 34-47 (2013).
7. Bock, K.W. Functions and transcriptional regulation of adult human hepatic UDP-glucuronosyl-transferases (UGTs): mechanisms responsible for interindividual variation of UGT levels. *Biochem Pharmacol* **80**, 771-777 (2010).
8. Gong, Q.H. *et al.* Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics* **11**, 357-368 (2001).

9. Levesque, E., Girard, H., Journault, K., Lepine, J. & Guillemette, C. Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus. *Hepatology* **45**, 128-138 (2007).

10. Guillemette, C., Levesque, E., Harvey, M., Bellemare, J. & Menard, V. UGT genomic diversity: beyond gene duplication. *Drug Metab Rev* **42**, 24-44 (2010).

11. Bushey, R.T., Dluzen, D.F. & Lazarus, P. Importance of UDP-glucuronosyltransferases 2A2 and 2A3 in tobacco carcinogen metabolism. *Drug Metab Dispos* **41**, 170-179 (2013).

12. Sneitz, N. *et al.* Human UDP-glucuronosyltransferase UGT2A2: cDNA construction, expression, and functional characterization in comparison with UGT2A1 and UGT2A3. *Pharmacogenet Genomics* **19**, 923-934 (2009).

13. Levesque, E. *et al.* Extensive splicing of transcripts encoding the bile acid-conjugating enzyme UGT2B4 modulates glucuronidation. *Pharmacogenet Genomics* **20**, 195-210 (2010).

14. Menard, V., Eap, O., Roberge, J., Harvey, M., Levesque, E. & Guillemette, C. 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 (2011).

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

16. McCarroll, S.A. *et al.* Common deletion polymorphisms in the human genome. *Nat Genet* **38**, 86-92 (2006).

17. Beutler, E., Gelbart, T. & Demina, A. Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? *Proc Natl Acad Sci U S A* **95**, 8170-8174 (1998).

18. Barbarino, J.M., Haidar, C.E., Klein, T.E. & Altman, R.B. PharmGKB summary: very important pharmacogene information for UGT1A1. *Pharmacogenet Genomics* **24**, 177-183 (2014).

19. Pichereau, S., Le Louarn, A., Lecomte, T., Blasco, H., Le Guellec, C. & Bourgoin, H. Cost-effectiveness of UGT1A1*28 genotyping in preventing severe neutropenia following FOLFIRI therapy in colorectal cancer. *J Pharm Pharm Sci* **13**, 615-625 (2010).

20. Cheng, L. *et al.* UGT1A1*6 polymorphisms are correlated with irinotecan-induced toxicity: a system review and meta-analysis in Asians. *Cancer Chemother Pharmacol* **73**, 551-560 (2014).

21. Han, F.F. *et al.* Associations between UGT1A1*6 or UGT1A1*6/*28 polymorphisms and irinotecan-induced neutropenia in Asian cancer patients. *Cancer Chemother Pharmacol* **73**, 779-788 (2014).

22. Sai, K. & Saito, Y. Ethnic differences in the metabolism, toxicology and efficacy of three anticancer drugs. *Expert Opin Drug Metab Toxicol* **7**, 967-988 (2011).

23. Levesque, E. *et al.* Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J Pharmacol Exp Ther* **345**, 95-101 (2013).

24. Menard, V., Girard, H., Harvey, M., Perusse, L. & Guillemette, C. Analysis of inherited genetic variations at the UGT1 locus in the French-Canadian population. *Hum Mutat* **30**, 677-687 (2009).

25. Yang, J., Cai, L., Huang, H., Liu, B. & Wu, Q. Genetic variations and haplotype diversity of the UGT1 gene cluster in the Chinese population. *PLoS One* **7**, e33988 (2012).

26. Wheeler, H.E., Maitland, M.L., Dolan, M.E., Cox, N.J. & Ratain, M.J. Cancer pharmacogenomics: strategies and challenges. *Nat Rev Genet* **14**, 23-34 (2013).

27. Hertz, D.L. & McLeod, H.L. Use of pharmacogenetics for predicting cancer prognosis and treatment exposure, response and toxicity. *J Hum Genet* **58**, 346-352 (2013).

28. Toffoli, G. *et al.* Genotype-driven phase I study of irinotecan administered in combination with fluorouracil/leucovorin in patients with metastatic colorectal cancer. *J Clin Oncol* **28**, 866-871 (2010).

29. Dupuis, R., Yuen, A. & Innocenti, F. The influence of UGT polymorphisms as biomarkers in solid organ transplantation. *Clin Chim Acta* **413**, 1318-1325 (2012).

30. Garcia-Closas, M. *et al.* Common genetic polymorphisms modify the effect of smoking on absolute risk of bladder cancer. *Cancer Res* **73**, 2211-2220 (2013).

31. Kpoghomou, M.A., Soatiana, J.E., Kalembo, F.W., Bishwajit, G. & Sheng, W. UGT2B17 Polymorphism and Risk of Prostate Cancer: A Meta-Analysis. *ISRN Oncol* **2013**, 465916 (2013).

32. Lu, P.H., Chen, M.B., Wu, X.Y., Gu, J.H., Liu, Y. & Gu, R.M. Genetic polymorphisms of UGT1A7 and cancer risk: evidence from 21 case-control studies. *Cancer Invest* **29**, 645-654 (2011).

33. Nagar, S. & Remmel, R.P. Uridine diphosphoglucuronosyltransferase pharmacogenetics and cancer. *Oncogene* **25**, 1659-1672 (2006).

34. Scherer, D. *et al.* Genetic variation in UGT genes modify the associations of NSAIDs with risk of colorectal cancer: Colon cancer family registry. *Genes Chromosomes Cancer*, (2014).

35. Selinski, S. *et al.* Rs11892031[A] on chromosome 2q37 in an intronic region of the UGT1A locus is associated with urinary bladder cancer risk. *Arch Toxicol* **86**, 1369-1378 (2012).

36. Tang, W. *et al.* Mapping of the UGT1A locus identifies an uncommon coding variant that affects mRNA expression and protects from bladder cancer. *Hum Mol Genet* **21**, 1918-1930 (2012).

37. Elahi, A. *et al.* Detection of UGT1A10 polymorphisms and their association with orolaryngeal carcinoma risk. *Cancer* **98**, 872-880 (2003).

38. Anielski, P., Simmchen, J., Wassill, L., Ganghofner, D. & Thieme, D. Epidemiological investigation of the UGT2B17 polymorphism in doping control urine samples and its correlation to T/E ratios. *Drug Test Anal* **3**, 645-651 (2011).

39. Giroux, S., Bussieres, J., Bureau, A. & Rousseau, F. UGT2B17 gene deletion associated with an increase in bone mineral density similar to the effect of hormone replacement in postmenopausal women. *Osteoporos Int* **23**, 1163-1170 (2012).

40. Grant, D.J. *et al.* Association of uridine diphosphate-glucuronosyltransferase 2B gene variants with serum glucuronide levels and prostate cancer risk. *Genet Test Mol Biomarkers* **17**, 3-9 (2013).

41. Jakobsson, J. *et al.* Large differences in testosterone excretion in Korean and Swedish men are strongly associated with a UDP-glucuronosyl transferase 2B17 polymorphism. *J Clin Endocrinol Metab* **91**, 687-693 (2006).

42. Nadeau, G. *et al.* Deletions of the androgen-metabolizing UGT2B genes have an effect on circulating steroid levels and biochemical recurrence after radical prostatectomy in localized prostate cancer. *J Clin Endocrinol Metab* **96**, E1550-1557 (2011).

43. Yang, T.L. *et al.* Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. *Am J Hum Genet* **83**, 663-674 (2008).

44. Angstadt, A.Y. *et al.* The effect of copy number variation in the phase II detoxification genes UGT2B17 and UGT2B28 on colorectal cancer risk. *Cancer* **119**, 2477-2485 (2013).

45. Bronstad, I., Wolff, A.S., Lovas, K., Knappskog, P.M. & Husebye, E.S. Genome-wide copy number variation (CNV) in patients with autoimmune Addison's disease. *BMC Med Genet* **12**, 111 (2011).

46. Zhang, Y., Dai, Y., Tu, Z., Li, Q., Wang, L. & Zhang, L. Array-CGH detection of UGT2B28 gene deletion in a girl with primary amenorrhea and hyperandrogenism. *Int J Gynaecol Obstet* **109**, 164-166 (2010).

47. McCarroll, S.A. *et al.* Donor-recipient mismatch for common gene deletion polymorphisms in graft-versus-host disease. *Nat Genet* **41**, 1341-1344 (2009).

48. Murata, M., Warren, E.H. & Riddell, S.R. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. *J Exp Med* **197**, 1279-1289 (2003).

49. Gruber, M. *et al.* Overexpression of uridine diphospho glucuronosyltransferase 2B17 in high-risk chronic lymphocytic leukemia. *Blood* **121**, 1175-1183 (2013).

50. Balliet, R.M., Chen, G., Gallagher, C.J., Dellinger, R.W., Sun, D. & Lazarus, P. Characterization of UGTs active against SAHA and association between SAHA glucuronidation activity phenotype with UGT genotype. *Cancer Res* **69**, 2981-2989 (2009).

51. Margueron, R. & Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**, 285-296 (2010).

52. Baylin, S.B. & Jones, P.A. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer* **11**, 726-734 (2011).

53. Zanger, U.M. & Schwab, M. Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* **138**, 103-141 (2013).

54. Belanger, A.S., Tojcic, J., Harvey, M. & Guillemette, C. Regulation of UGT1A1 and HNF1 transcription factor gene expression by DNA methylation in colon cancer cells. *BMC Mol Biol* **11**, 9 (2010).

55. Yasar, U., Greenblatt, D.J., Guillemette, C. & Court, M.H. Evidence for regulation of UDP-glucuronosyltransferase (UGT) 1A1 protein expression and activity via DNA methylation in healthy human livers. *J Pharm Pharmacol* **65**, 874-883 (2013).

56. Oda, S., Fukami, T., Yokoi, T. & Nakajima, M. Epigenetic regulation is a crucial factor in the repression of UGT1A1 expression in the human kidney. *Drug Metab Dispos* **41**, 1738-1743 (2013).

57. Oda, S., Fukami, T., Yokoi, T. & Nakajima, M. Epigenetic regulation of the tissue-specific expression of human UDP-glucuronosyltransferase (UGT) 1A10. *Biochem Pharmacol* **87**, 660-667 (2014).

58. Yokoi, T. & Nakajima, M. microRNAs as mediators of drug toxicity. *Annu Rev Pharmacol Toxicol* **53**, 377-400 (2013).

59. Dluzen, D.F. *et al.* Regulation of UDP-glucuronosyltransferase 1A1 expression and activity by microRNA 491-3p. *J Pharmacol Exp Ther* **348**, 465-477 (2014).

60. Rieger, J.K., Klein, K., Winter, S. & Zanger, U.M. Expression variability of absorption, distribution, metabolism, excretion-related microRNAs in human liver: influence of nongenetic factors and association with gene expression. *Drug Metab Dispos* **41**, 1752-1762 (2013).

61. Dehm, S.M. mRNA splicing variants: exploiting modularity to outwit cancer therapy. *Cancer Res* **73**, 5309-5314 (2013).

62. Girard, H., Levesque, E., Bellemare, J., Journault, K., Caillier, B. & Guillemette, C. 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 (2007).

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

64. Court, M.H. *et al.* The UDP-glucuronosyltransferase (UGT) 1A polymorphism c.2042C>G (rs8330) is associated with increased human liver acetaminophen glucuronidation, increased UGT1A exon 5a/5b splice variant mRNA ratio, and decreased

risk of unintentional acetaminophen-induced acute liver failure. *J Pharmacol Exp Ther* **345**, 297-307 (2013).

65. Bellemare, J. *et al.* Immunohistochemical expression of conjugating UGT1A-derived isoforms in normal and tumoral drug-metabolizing tissues in humans. *J Pathol* **223**, 425-435 (2011).

66. Hu, D.G., Meech, R., Lu, L., McKinnon, R.A. & Mackenzie, P.I. Polymorphisms and Haplotypes of the UDP-Glucuronosyltransferase 2B7 Gene Promoter. *Drug Metab Dispos* **42**, 854-862 (2014).

67. Izukawa, T. *et al.* Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers. *Drug Metab Dispos* **37**, 1759-1768 (2009).

68. Peterkin, V.C. *et al.* 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 (2007).

69. Menard, V., Collin, P., Margaillan, G. & Guillemette, C. Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos* **41**, 2197-2205 (2013).

70. Menard, V. *et al.* 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 (2013).

71. Innocenti, F. *et al.* Single nucleotide polymorphism discovery and functional assessment of variation in the UDP-glucuronosyltransferase 2B7 gene. *Pharmacogenet Genomics* **18**, 683-697 (2008).

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

73. Walters, W.P., Green, J., Weiss, J.R. & Murcko, M.A. What do medicinal chemists actually make? A 50-year retrospective. *J Med Chem* **54**, 6405-6416 (2011).

74. Fallon, J.K., Neubert, H., Hyland, R., Goosen, T.C. & Smith, P.C. 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 (2013).

75. Sato, Y. *et al.* Optimized methods for targeted Peptide-based quantification of human uridine 5'-diphosphate-glucuronosyltransferases in biological specimens using liquid chromatography-tandem mass spectrometry. *Drug Metab Dispos* **42**, 885-889 (2014).

76. Nelson, M.R. *et al.* An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. *Science* **337**, 100-104 (2012).

77. Ramsey, L.B. *et al.* Rare versus common variants in pharmacogenetics: SLCO1B1 variation and methotrexate disposition. *Genome Res* **22**, 1-8 (2012).

78. Sun, C., Southard, C., Olopade, O.I. & Di Rienzo, A. Differential allelic expression of c.1568C > A at UGT2B15 is due to variation in a novel cis-regulatory element in the 3'UTR. *Gene* **481**, 24-28 (2011).

79. Gantla, S. *et al.* Splice-site mutations: a novel genetic mechanism of Crigler-Najjar syndrome type 1. *Am J Hum Genet* **62**, 585-592 (1998).

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

81. Strassburg, C.P., Strassburg, A., Nguyen, N., Li, Q., Manns, M.P. & Tukey, R.H. Regulation and function of family 1 and family 2 UDP-glucuronosyltransferase genes (UGT1A, UGT2B) in human oesophagus. *Biochem J* **338 (Pt 2)**, 489-498 (1999).

82. Ohno, S. & Nakajin, S. Determination of mRNA expression of human UDP-glucuronosyltransferases and application for localization in various human tissues by real-time reverse transcriptase-polymerase chain reaction. *Drug Metab Dispos* **37**, 32-40 (2009).

Figure Legends

Figure 1. The glucuronidation reaction. ***A.*** Schematic illustration of the glucuronidation reaction by a UGT enzyme located in the endoplasmic reticulum (ER). ***A.*** The UGT enzyme lies predominantly on the luminal side of the ER. A short transmembrane domain directs the C-terminal tail and a dilysine motif (++) on the cytosolic side. The UDP-glucuronic acid (UDPGA) co-substrate is actively transported into the ER while lipophilic substrates (illustrated by R and R-OH) are directly conjugated by UGTs or following functionalization by other drug metabolism enzymes such as cytochrome P450 (P450). The hydrophilic glucuronide conjugate is actively released for elimination. ***B.*** Schematic representation of the enzymatic reaction. The conjugation reaction involves the conversion of the UDP-glucuronic acid (co-substrate) α-bond into a β-bond between the nucleophilic substrate and the sugar via an S_N2 mechanism. ***C.*** Functional groups targeted by UGT enzymes. Glucuronidated moieties are highlighted. *O*-linked and *N*-linked glucuronidation are the most frequently encountered.

Figure 2. Schematic representation of human *UGT* loci and encoded UGT proteins. ***Top*** *UGT1* locus, located on chromosome 2, comprises 13 first exons which are alternatively spliced to 3 common exons and alternative 3' exons 5a or 5b. Alternative first exons shown in grey do not encode functional proteins due sequence defects (e.g. pseudogenes). The *UGT1A9* mature mRNA is shown as an example. ***Bottom*** *UGT2* locus, located on chromosome 4, comprises distinct but highly similar genes translated into 10 mRNAs encoding UGT2A and UGT2B enzymes. Each gene comprises at least six exons represented by colored boxes. The direction of transcription is indicated by an arrow over each gene. Pseudogenes are represented by grey boxes. The mature mRNA encoding UGT2B7 is shown as an example. ***Center*** The primary

structure of UGT1A, UGT2A and UGT2B enzymes are highly related. The N-terminal variable domain, encoded by exon 1 (UGT1A) or exons 1 and 2 (UGT2), is responsible for substrate specificity. The variable domain also comprises a signal peptide cleaved to generate the mature protein after proper localization in the ER, and a putative membrane-anchoring domain. The C-terminal half of UGT enzymes is highly similar among the UGT1A and UGT2 subfamilies. It includes the co-substrate UDPGA binding domain, a UGT signature sequence, the transmembrane region and a cytosolic dilysine motif common to all glucuronidating enzymes.

Figure 3. Wide expression profiles of UGTs in major drug metabolizing tissues. The specific UGTs detected/undetected in each human tissue are summarized according to quantitative reverse transcription-PCR analysis of UGT mRNA levels.[72, 80, 81] Detection of UGT1A variants are based on exon 1 amplification startegies and do not discriminate between 3' variants. The relative expression of UGT mRNAs in the small intesting is given, based on published qPCR quantifications.[80, 82] The relative abundance of hepatic UGT proteins is also represented according to recent targeted quantitative proteomics data.[74, 75] The distribution of UGT2A enzymes is not depicted their contribution to drug metabolism is uncertain.

Figure 4. Schematic representation of most common known non-synonymous and regulatory single nucleotide polymorphisms (SNPs) and copy number variations (CNVs). Non-synonymous SNPs are shown in purple boxes below the UGT loci. Regulatory SNPs are represented by yellow boxes above the loci. Pink boxes represent copy number variations of UGT2B17 and UGT2B28 genes due to whole gene deletions. An exhaustive list of SNPs with

corresponding rs numbers and allele names is available at http://www.pharmacogenomics.pha.ulaval.ca/cms/ugt_alleles.

Figure 5. Pharmacogenomics of UGTs and drug response: irinotecan as an example. *A.* Upon administration, irinotecan is predominantly converted to its active metabolite SN-38 by carboxyesterases (hCE2) with a little fraction subjected to CYP3A4-dependent oxidation. Both hepatic and extrahepatic glucuronidation of SN-38 (SN-38G) by multiple UGT enzymes, namely UGT1A1, UGT1A7 and UGT1A9, dictates exposure levels and subsequent antitumor activity and toxicity. ***B.*** The expression and glucuronidation activity of UGT enzymes are collectively affected by SNPs and other genetic alterations. Specific UGT1A haplotypes conferring either increased protection or increased risk for severe neutropenia have been identified. Combination of variants across the *UGT1A* loci is likely to refine pharmacogenetic testing compared to the predictive value of the current predictive marker *UGT1A1*28*.

Figure 6. Schematic representation of genetic and genomics mechanisms contributing to interindividual glucuronidation phenotypes and expansion of the UGT transcriptome and proteome. Genomic, transcriptional and posttranscriptional elements determine the resulting glucuronidation activity in a tissue and cell type-specific manner. The tissue-specific expression of UGT family members is orchestrated by cooperative interactions between transcription factors (TFs), ligand-activated nuclear receptors (NRs) and the transcription-initiation complex. Single nucleotide polymorphisms (SNPs), copy number variation (CNV), chromatin state (such as histone acetylation and DNA methylation) and TFs/NRs collectively regulate UGT gene expression and alternative splicing events. MicroRNAs (miRNAs) also regulate UGT expression

by directly modulating mRNA stability and translation of UGTs or indirectly by modulating UGT-dependent TFs. Mature mRNAs encode UGT enzymes and numerous alternative isoforms that mediate several protein-protein interactions with UGT enzymes modulating glucuronidation activity. In addition, data support interactions with drug metabolizing enzymes (DMEs) and other metabolic enzymes (cytochrome P450, glutathione S-transferases, catalase, peroxiredoxin, etc), likely influencing their activity. Besides their glucuronidating functions, some UGT members are reported to play other roles including interactions with other metabolic enzymes in the cytoplasm and a function as minor antigens for UGT2B17.

Table 1. The contribution of UGT enzymes to the metabolism of clinically relevant drugs.

UGTs	Tissue distribution		Estimated % of drugs		Typical drug substrates[4]	Probe substrates
	Liver	Extra hepatic[1]	200 most prescribed[2]	known glucuronidated drugs[3]		
1A1	✓	✓	18	11	SN-38, simvastatin, etoposide, ezetimibe, ethinyltestradiol, atorvastatin, codeine	bilirubin
1A3	✓	✓	24	13	atorvastatin, simvastatin, ezetimibe, telmisartan	
1A4	✓	✓	4	11	tamoxifen, lamotrigine, olanzapine, amitryptyline midazolam, tacrolimus	trifluoperazine tacrolimus
1A5		✓			unknown	
1A6	✓	✓	2	6	Deferiprone, paracetamol	serotonin deferiprone
1A7		✓	11	3	SN-38	
1A8		✓	3	4	mycophenolic acid	
1A9	✓	✓	6	14	SN-38, paracetamol, flavopiridol, propofol, entacapone, R-oxazepam, mycophenolic acid, edaravone, sorafenib, tolcapone	propofol
1A10		✓	13	6	SN-38	
2B4	✓	✓	5	3	acetaminophen, deferiprone	
2B7	✓	✓	25	19	zidovudine, epirubicin, fenofibrate, morphine, codeine, NSAIDs, mycophenolic acid, chloramphenicol, efavirenz, naproxen, naloxone	zidovudine
2B10	✓	✓	1	2	diphenhydramine, olanzapine,	

					levomedetomidine	
2B11	✓	✓			unknown	unknown
2B15	✓	✓	4	5	tamoxifen, S-oxazepam, lorazepam, dabigatran, R-methadone, tolcapone	S-oxazepam
2B17	✓	✓	4	3	vorinostat	
2B28	✓	✓			unknown	unknown

[1]Extrahepatic tissues consist in drug metabolizing tissues other than liver such as intestine, colon, kidney.
[2]According to www.rxlist.com (*list of 200 most prescribed drugs in 2012; accessed April 11th, 2014*).
UGTs collectively conjugate 111 (55%) of the 200 most prescribed drugs. Listed % represents the proportion of glucuronidated drugs metabolized by each UGT.
[3]Reviewed by Stingl et al. 2014;[2] % represents the contribution of each UGT to the conjugation of all known glucuronidated drugs.
[4]The list of substrates is not exhaustive.

Table 2. UGTs as susceptibility and prognostic genes for numerous cancers.

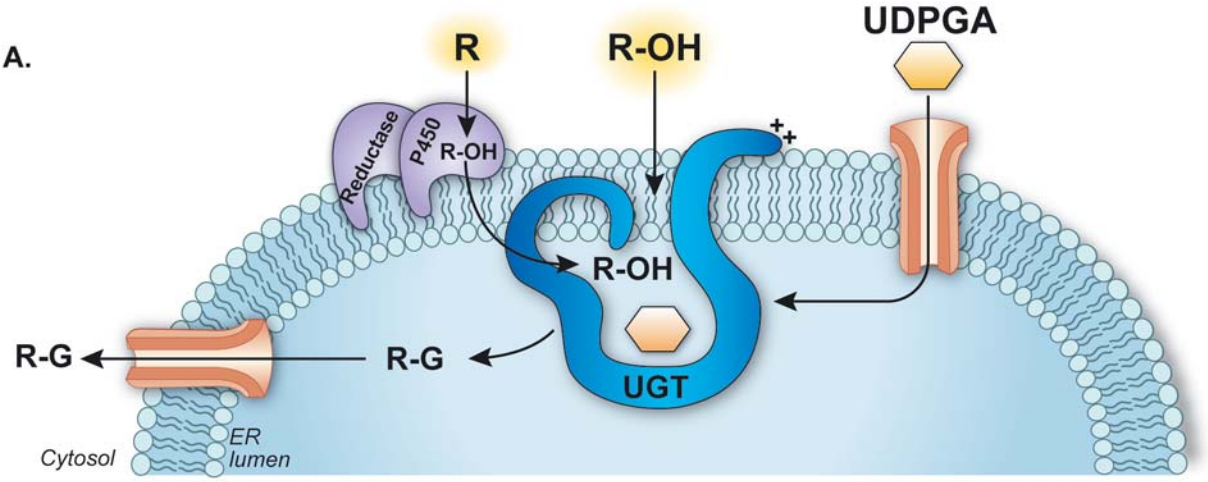
		Clinical impact	
Gene[1]	**Tumor site**	**Risk**	**Prognosis[2]**
UGT1A	Bladder	✓	✓
UGT1A1	Chronic lymphocytic leukemia	✓	✓
	Colorectal	✓	
	Head and neck	✓	
UGT1A6	Bladder	✓	
	Breast	✓	
	Colorectal	✓	
UGT1A7	Bladder	✓	
	Colorectal	✓	
	Head and neck	✓	
	Hepatocellular	✓	
	Orolaryngeal	✓	
UGT1A8	Esophageal	✓	
UGT1A10	Orolaryngeal	✓	
UGT2B4	Breast	✓	
	Colon	✓	
	Esophageal	✓	
UGT2B15	Prostate	✓	✓
	Colon	✓	
UGT2B17	Prostate	✓	✓
	Lung	✓	
	Chronic lymphocytic leukemia		✓
UGT2B28	Prostate		✓
	Colorectal	✓	

Available data are mostly from case-control studies, including a few from GWAS, as well as case series studies.

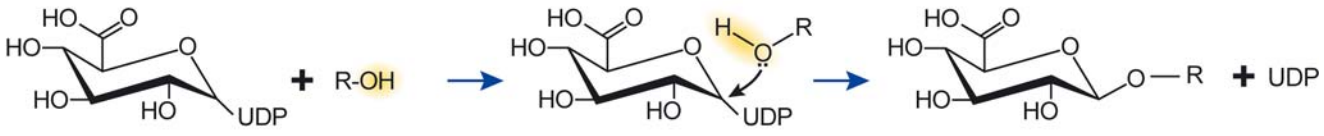
[1]Germline DNA variations and/or altered gene expression levels.

[2]Cancer relapse, aggressiveness and/or survival; unrelated to drug treatment.

A.

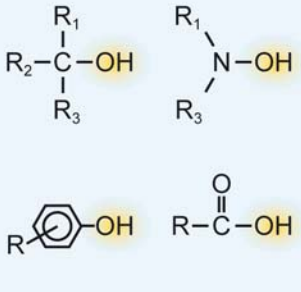
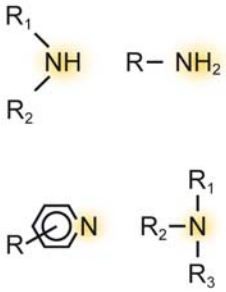
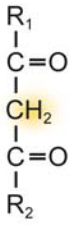


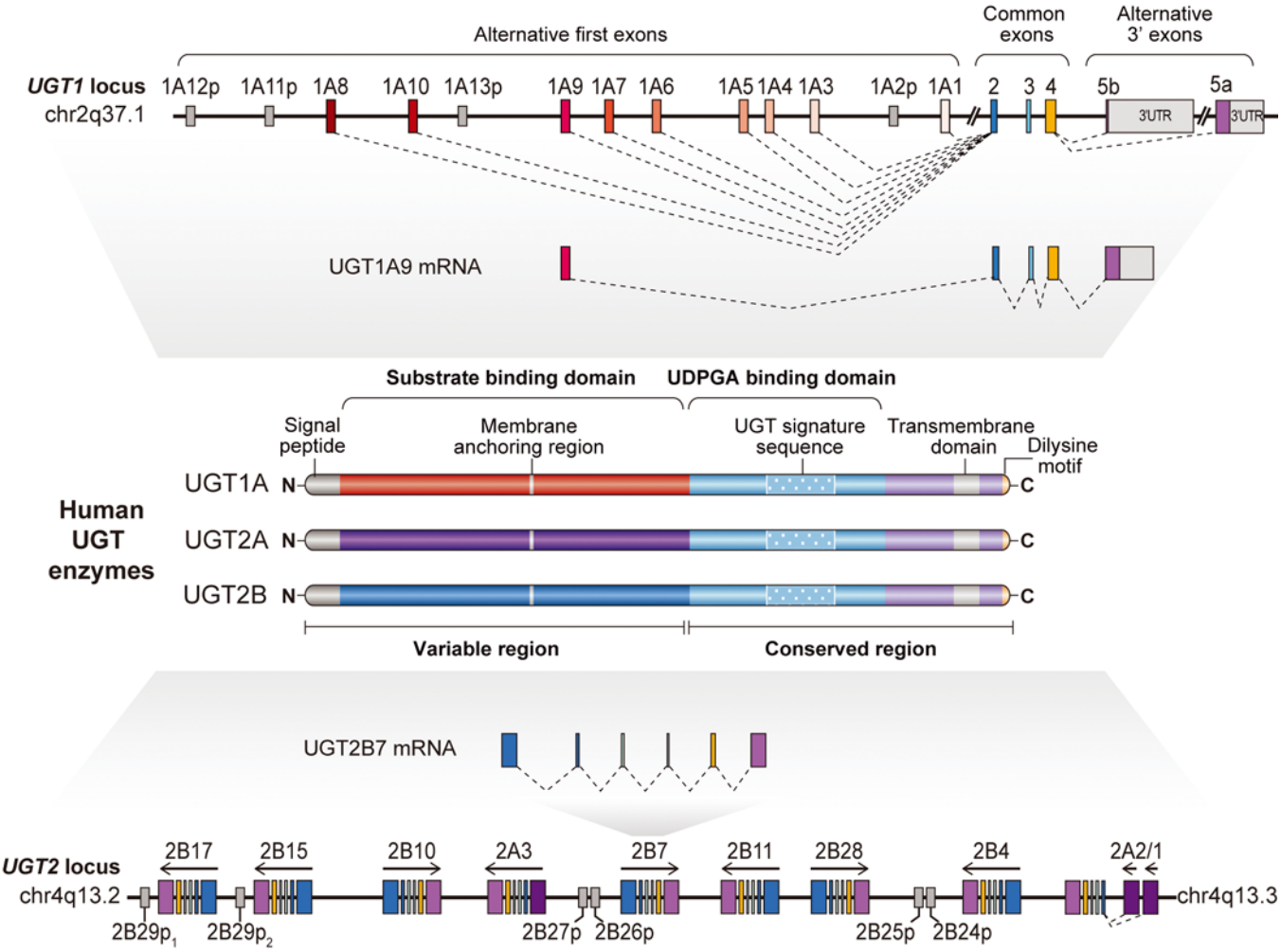
B.

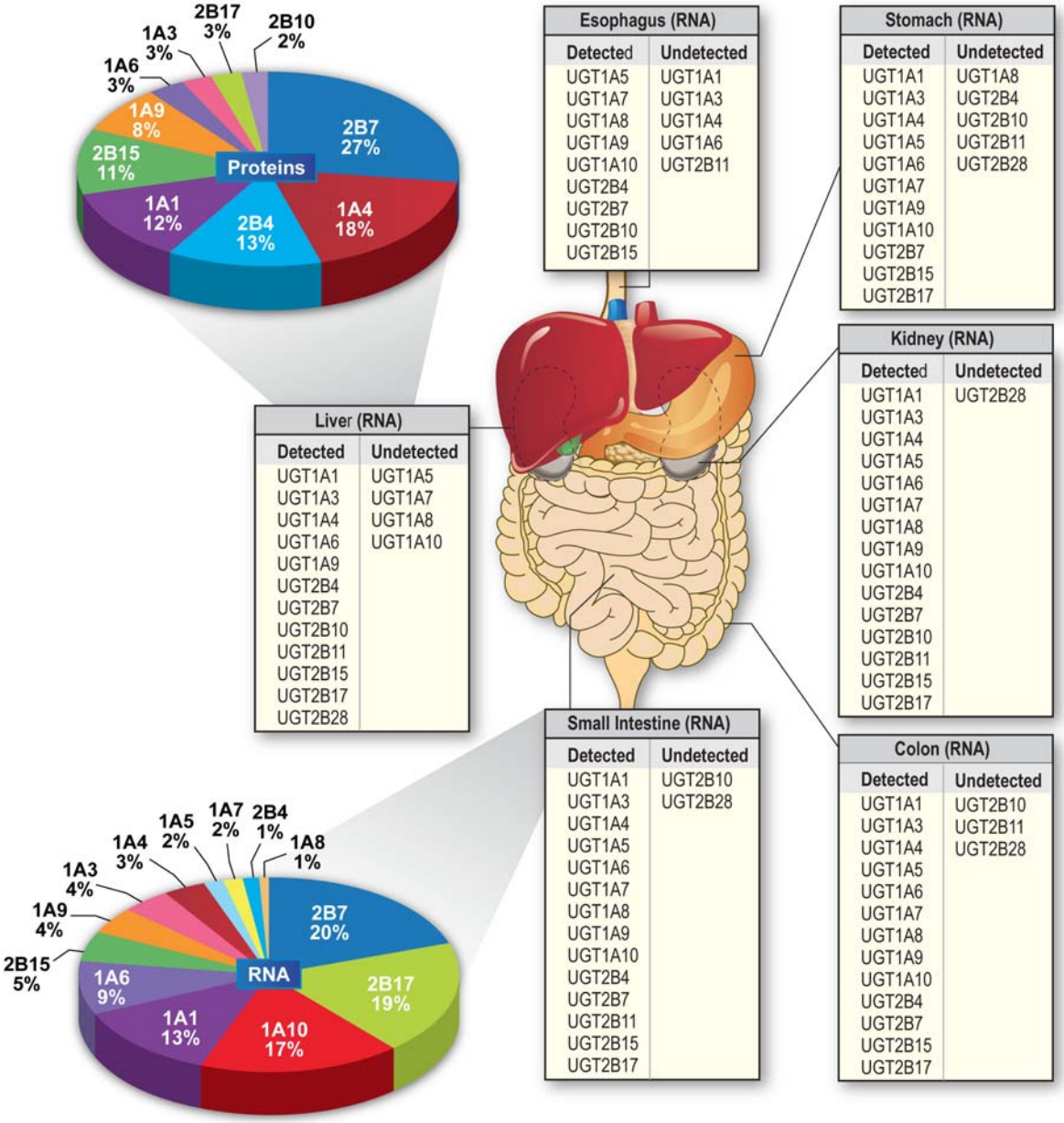


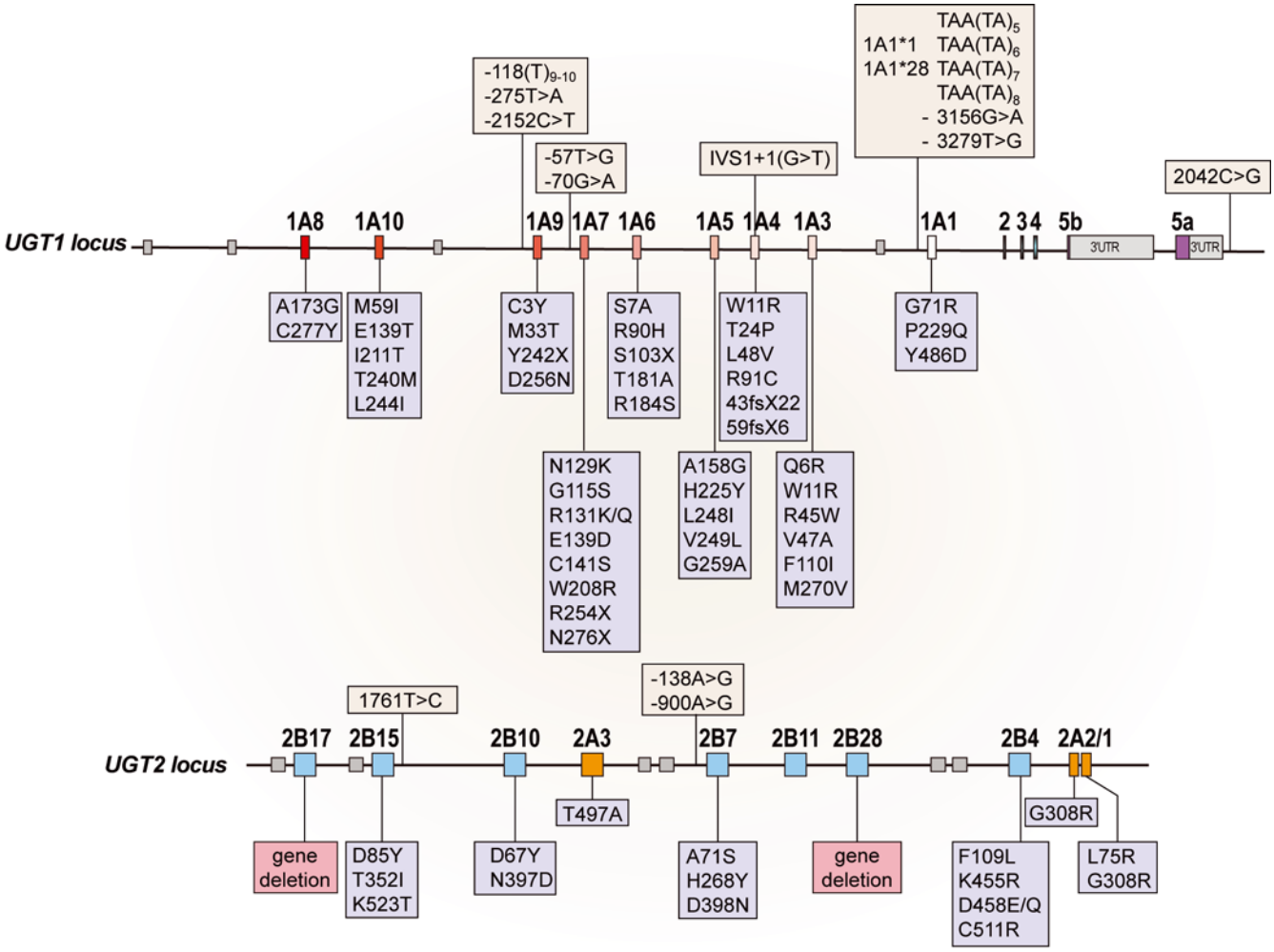
UDPGA co-substrate | **Nucleophilic substrate** | **Nucleophilic attack** | **Substrate glucuronide (-G)**

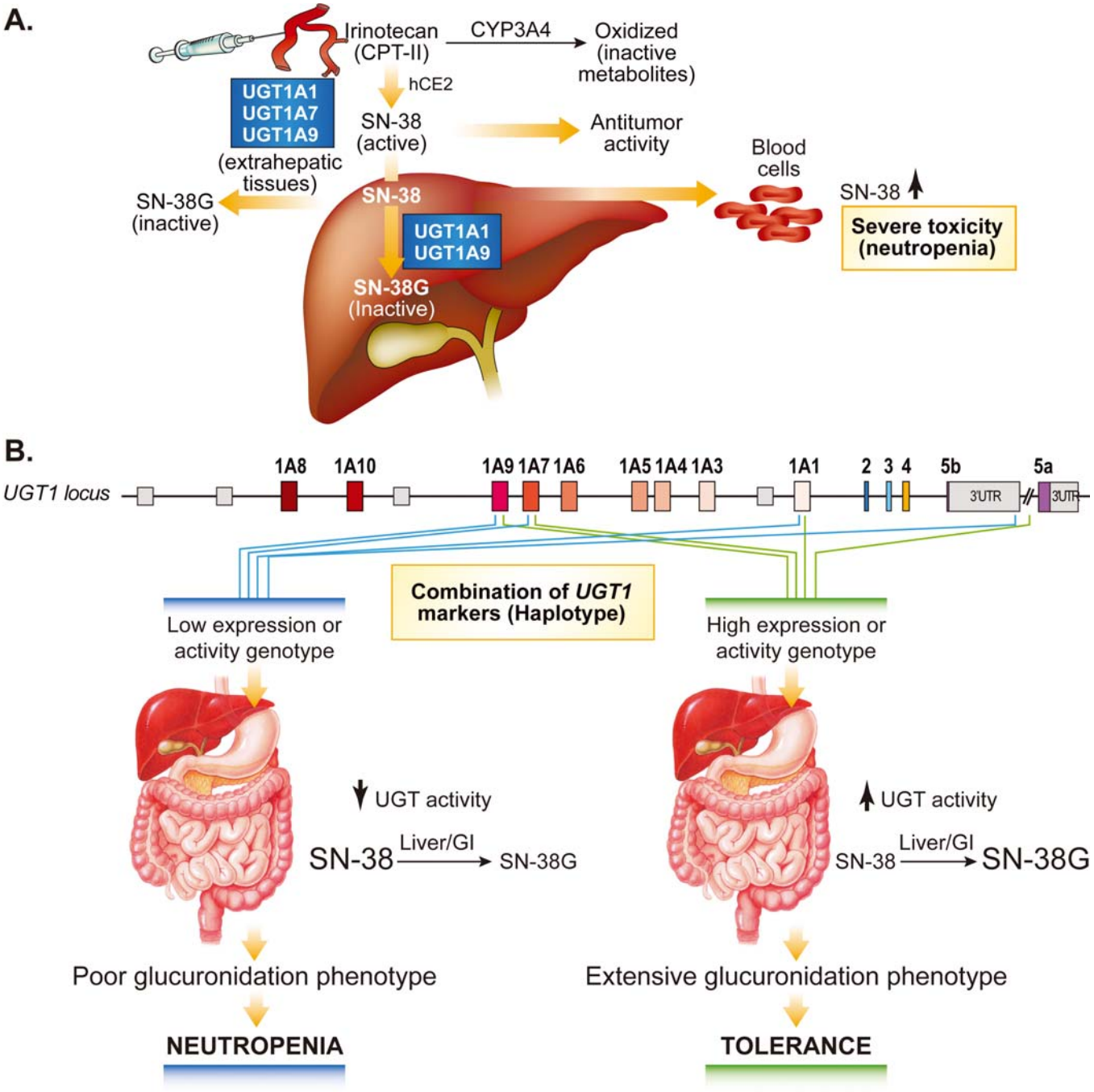
C.

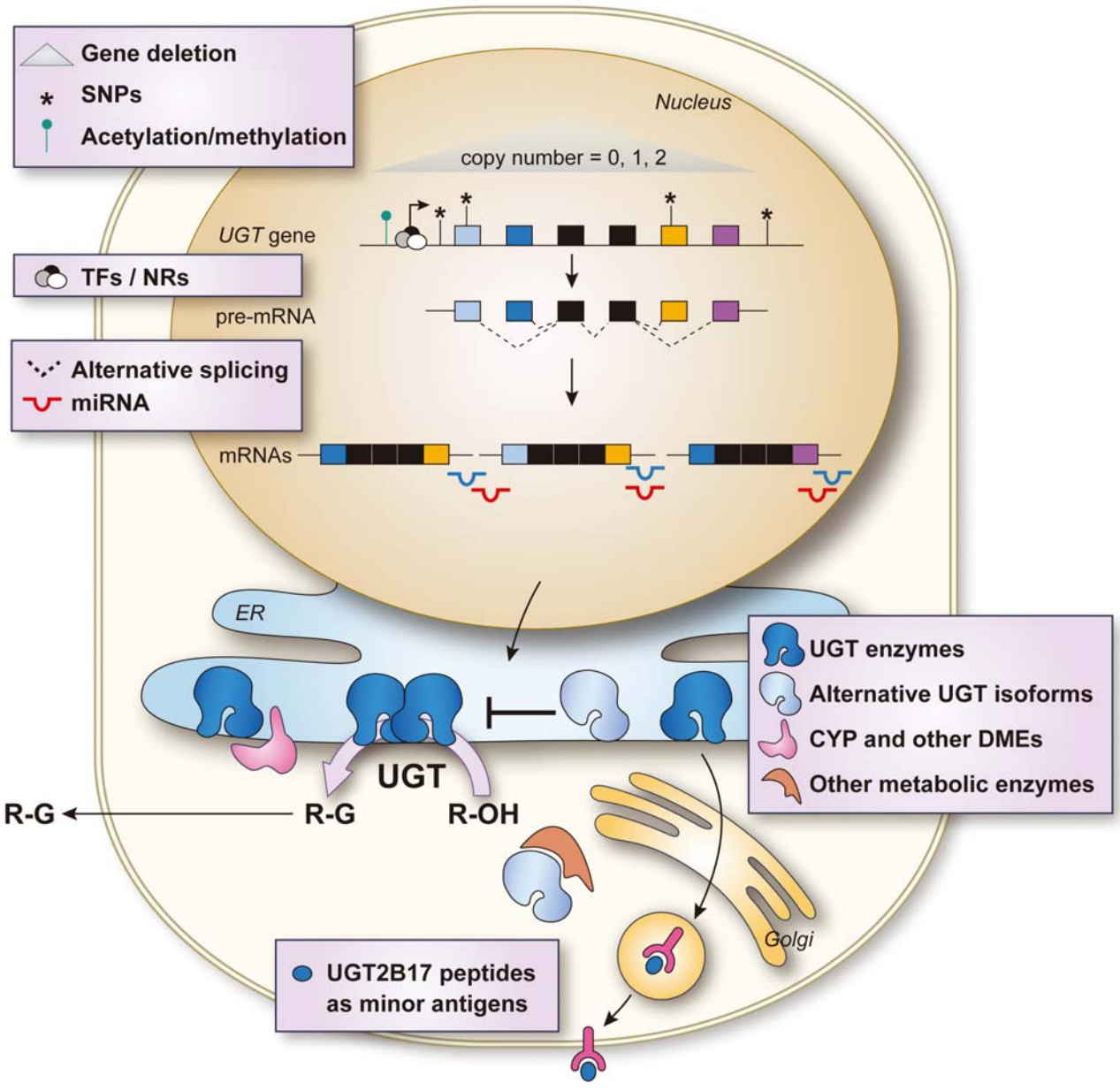
	Hydroxyl/Carboxyl	Amines	Sulfhydryl	Enolic acids
Functional groups				
Glucuronides	*O*-linked	*N*-linked	*S*-linked	*C*-linked
UGTs	1A, 2A, 2B	1A, 2B7, 2B10	1A, 2A, 2B	1A9


Alternative first exons
Common exons
Alternative 3' exons
UGT1 locus
chr2q37.1
1A12p 1A11p 1A8 1A10 1A13p 1A9 1A7 1A6 1A5 1A4 1A3 1A2p 1A1 2 3 4 5b 5a
3'UTR
3'UTR
UGT1A9 mRNA
Substrate binding domain
UDPGA binding domain
Signal peptide
Membrane anchoring region
UGT signature sequence
Transmembrane domain
Dilysine motif
Human UGT enzymes
UGT1A
UGT2A
UGT2B
N
C
Variable region
Conserved region
UGT2B7 mRNA
UGT2 locus
chr4q13.2
chr4q13.3
2B17 2B15 2B10 2A3 2B7 2B11 2B28 2B4 2A2/1
2B29p1 2B29p2 2B27p 2B26p 2B25p 2B24p


2B17 3%
2B10 2%
1A3 3%
1A6 3%
1A9 8%
2B15 11%
Proteins
2B7 27%
1A1 12%
2B4 13%
1A4 18%
Esophagus (RNA)
Detected
Undetected
UGT1A5
UGT1A7
UGT1A8
UGT1A9
UGT1A10
UGT2B4
UGT2B7
UGT2B10
UGT2B15
UGT1A1
UGT1A3
UGT1A4
UGT1A6
UGT2B11
Stomach (RNA)
Detected
Undetected
UGT1A1
UGT1A3
UGT1A4
UGT1A5
UGT1A6
UGT1A7
UGT1A9
UGT1A10
UGT2B7
UGT2B15
UGT2B17
UGT1A8
UGT2B4
UGT2B10
UGT2B11
UGT2B28
Kidney (RNA)
Detected
Undetected
UGT1A1
UGT1A3
UGT1A4
UGT1A5
UGT1A6
UGT1A7
UGT1A8
UGT1A9
UGT1A10
UGT2B4
UGT2B7
UGT2B10
UGT2B11
UGT2B15
UGT2B17
UGT2B28
Liver (RNA)
Detected
Undetected
UGT1A1
UGT1A3
UGT1A4
UGT1A6
UGT1A9
UGT2B4
UGT2B7
UGT2B10
UGT2B11
UGT2B15
UGT2B17
UGT2B28
UGT1A5
UGT1A7
UGT1A8
UGT1A10
Small Intestine (RNA)
Detected
Undetected
UGT1A1
UGT1A3
UGT1A4
UGT1A5
UGT1A6
UGT1A7
UGT1A8
UGT1A9
UGT1A10
UGT2B4
UGT2B7
UGT2B11
UGT2B15
UGT2B17
UGT2B10
UGT2B28
Colon (RNA)
Detected
Undetected
UGT1A1
UGT1A3
UGT1A4
UGT1A5
UGT1A6
UGT1A7
UGT1A8
UGT1A9
UGT1A10
UGT2B4
UGT2B7
UGT2B15
UGT2B17
UGT2B10
UGT2B11
UGT2B28
1A5 2%
1A7 2%
2B4 1%
1A8 1%
1A4 3%
1A3 4%
1A9 4%
2B15 5%
RNA
2B7 20%
1A6 9%
2B17 19%
1A1 13%
1A10 17%


UGT1 locus
1A8
1A10
1A9
1A7
1A6
1A5
1A4
1A3
1A1
2
3
4
5b
5a
3'UTR
3'UTR
A173G
C277Y
M59I
E139T
I211T
T240M
L244I
-118(T)9-10
-275T>A
-2152C>T
-57T>G
-70G>A
C3Y
M33T
Y242X
D256N
S7A
R90H
S103X
T181A
R184S
IVS1+1(G>T)
W11R
T24P
L48V
R91C
43fsX22
59fsX6
TAA(TA)5
1A1*1 TAA(TA)6
1A1*28 TAA(TA)7
TAA(TA)8
- 3156G>A
- 3279T>G
G71R
P229Q
Y486D
2042C>G
N129K
G115S
R131K/Q
E139D
C141S
W208R
R254X
N276X
A158G
H225Y
L248I
V249L
G259A
Q6R
W11R
R45W
V47A
F110I
M270V
UGT2 locus
2B17
2B15
2B10
2A3
2B7
2B11
2B28
2B4
2A2/1
1761T>C
-138A>G
-900A>G
gene deletion
D85Y
T352I
K523T
D67Y
N397D
T497A
A71S
H268Y
D398N
gene deletion
F109L
K455R
D458E/Q
C511R
G308R
L75R
G308R


A.
Irinotecan (CPT-II)
CYP3A4
Oxidized (inactive metabolites)
hCE2
UGT1A1
UGT1A7
UGT1A9
(extrahepatic tissues)
SN-38 (active)
Antitumor activity
Blood cells
SN-38
Severe toxicity (neutropenia)
SN-38G (inactive)
SN-38
UGT1A1
UGT1A9
SN-38G (Inactive)
B.
UGT1 locus
1A8
1A10
1A9
1A7
1A6
1A5
1A4
1A3
1A1
2
3
4
5b
5a
3'UTR
3'UTR
Low expression or activity genotype
Combination of UGT1 markers (Haplotype)
High expression or activity genotype
UGT activity
SN-38
Liver/GI
SN-38G
UGT activity
SN-38
Liver/GI
SN-38G
Poor glucuronidation phenotype
Extensive glucuronidation phenotype
NEUTROPENIA
TOLERANCE


Gene deletion
SNPs
Acetylation/methylation
Nucleus
copy number = 0, 1, 2
UGT gene
TFs / NRs
pre-mRNA
Alternative splicing
miRNA
mRNAs
ER
UGT
R-G
R-G
R-OH
UGT enzymes
Alternative UGT isoforms
CYP and other DMEs
Other metabolic enzymes
Golgi
UGT2B17 peptides
as minor antigens