Supplemental material to this article can be found at:
http://molpharm.aspetjournals.org/content/suppl/2018/06/29/mol.118.111773.DC1

1521-0111/94/3/938–952$35.00
MOLECULAR PHARMACOLOGY
Copyright © 2018 by The American Society for Pharmacology and Experimental Therapeutics
https://doi.org/10.1124/mol.118.111773
Mol Pharmacol 94:938–952, September 2018

Intergenic Splicing between Four Adjacent *UGT* Genes (*2B15, 2B29P2, 2B17, 2B29P1*) Gives Rise to Variant UGT Proteins That Inhibit Glucuronidation via Protein-Protein Interactions[S]

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

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

Received January 15, 2018; accepted June 8, 2018

ABSTRACT

Recent studies have investigated alternative splicing profiles of UDP-glucuronosyltransferase *(UGT)* genes and identified over 130 different alternatively spliced UGT transcripts. Although *UGT* genes are highly clustered, the formation of chimeric transcripts by intergenic splicing between two or more *UGT* genes has not yet been reported. This study identified 12 chimeric transcripts (chimeras A–L) containing exons from two or three genes of the four neighboring *UGT* genes (*UGT2B15*, *UGT2B29P2*, *UGT2B17*, and *UGT2B29P1*) in human liver and prostate cancer cells. These chimeras typically contain the first five exons of *UGT2B15* or *UGT2B17* (exons 1–5) spliced to a terminal exon (exon 6) from a downstream *UGT* gene. Hence they encode truncated UGTs with novel C-terminal peptides. Functional assays of representative chimeric UGT proteins (termed chimeric UGT2B15 and chimeric UGT2B17) showed that they are inactive and can repress the activity of wild-type UGTs. Coimmunoprecipitation assays demonstrated heterotypic interactions between chimeric UGT2B15 (or chimeric UGT2B17) and the UGT2B7 protein. Thus oligomerization of the chimeric UGTs with wild-type UGTs may explain their inhibitory activity. Studies in breast and prostate cancer cells showed that both wild-type and chimeric UGT2B15 and UGT2B17 transcripts are regulated in a similar way at the transcriptional level by sex hormones through their canonical promoters but are differentially regulated at the post-transcriptional level by micro-RNA 376c via their unique 3′-untranslated regions. In conclusion, the formation of chimeric transcripts by intergenic splicing among *UGT* genes represents a novel mechanism contributing to the diversity of the human UGT transcriptome and proteome. The differential post-transcriptional regulation of wild-type and variant transcripts by micro-RNAs may contribute to their deregulated expression in cancer.

Introduction

Glucuronidation is the process of conjugating glucuronic acid to small lipophilic molecules of both exogenous and endogenous origin, rendering them generally inactive, more water soluble, and more readily excreted (Mackenzie et al., 2005). Glucuronidation plays an important role in detoxification, therapeutic drug clearance, and maintenance of homeostasis of bioactive endogenous molecules such as bile acids and steroid hormones. Glucuronidation is primarily catalyzed by the UDP-glucuronosyltransferase (UGT) 1A and 2B subfamilies (Hu et al., 2014c). The nine functional UGT1A enzymes are encoded by a single *UGT1* locus at 2q37 through the splicing of an isoform-specific exon 1 to a set of shared exons 2–5. The seven functional UGT2B enzymes (2B4, 2B7, 2B10, 2B11, 2B15, 2B17, and 2B28) are encoded by individual genes that form a gene cluster at 4q13 (Mackenzie et al., 2005).

In addition to the 16 wild-type transcripts that encode the functional UGT1A and UGT2B enzymes, variant UGT transcripts generated from alternative splicing of canonical exons and/or alternative exons have been reported (Girard et al., 2007; Lévesque et al., 2010; Ménard et al., 2011). A recent comprehensive search of transcripts from *UGT1* and *UGT2* loci using targeted next-generation RNA-sequencing identified over 130 variant transcripts in normal and cancerous drug-metabolizing tissues (liver, kidney, intestine, and colon) and hormone-dependent tissues (prostate, breast, and uterus) (Tourancheau et al., 2016). The *UGT1* and *UGT2B7* genes have the most extensive splicing complexity reported to date (Tourancheau et al., 2016). *UGT1* variants are generated by the splicing of one or both of two alternative exons (exon 5b and exons 5b/5a) after the canonical exon 4. This generates two different variants for each of the nine UGT1A transcripts,

This study was supported by the National Health and Medical Research Council (NHMRC) of Australia [Grants ID1020931 and ID1085410] and funding from the Flinders Medical Centre Foundation. The investigators were supported by a Cancer Council/SA Health Beat Cancer Professorial Chair and an Australian Research Council Future Fellowship.
https://doi.org/10.1124/mol.118.111773.
[S] This article has supplemental material available at molpharm.aspetjournals.org.

ABBREVIATIONS: AR, androgen receptor; bp, base pair; cis-SAGe, cis-splicing of adjacent gene; FBS, fetal bovine serum; HEK, human embryonic kidney; mi, micro; miR-376c, microRNA 376c; 4-MU, methylumbelliferone; PCR, polymerase chain reaction; RT, reverse transcription; RT-qPCR, reverse transcriptase–real-time quantitative polymerase chain reaction; qPCR, quantitative real-time polymerase chain reaction; UGT, UDP-glucuronosyltransferase; UTR, untranslated region.

MOLECULAR PHARMACOLOGY

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

UGT1A_v2 (1/2/3/4/5b) and UGT1A_v3 (1/2/3/4/5b/5a), respectively. The UGT1A_v2 and v3 transcripts have identical open reading frames and thus encode identical proteins (termed UGT1A_i2), which have a novel 10-residue peptide replacing the 99-residue C-terminal domain of the wild-type UGT1A proteins (UGT1A_i1) (Lévesque et al., 2007). The *UGT2B7* gene has six canonical exons and 13 alternative exons and generates at least 44 different transcripts through a complex splicing pattern (Tourancheau et al., 2016).

Studies of UGT transcript diversity have revealed the common occurrence of variants that have the canonical 3′-terminal exon replaced by an alternative exon with a short open reading frame (e.g., UGT1A_v2, UGT2B7_v5). These variant mRNAs encode C-terminally truncated proteins (e.g., UGT1A_i2, UGT2B7_i2) that lack the transmembrane and cytosolic domains and possess no UGT activity; however, they negatively regulate glucuronidation activity when coexpressed with their wild-type counterparts via protein-protein interactions (Girard et al., 2007; Ménard et al., 2011).

All alternatively spliced UGT transcripts reported so far contain exons that are located within the boundaries of single *UGT* gene locus. Transcripts that contain exons from more than one gene locus are termed chimeric transcripts (Parra et al., 2006) and have been found in both normal and cancer tissues (Chwalenia et al., 2017). Several mechanisms contribute to the formation of chimeric transcripts, including trans-splicing, cis-splicing of adjacent genes (termed cis-SAGe), and chromosomal translocation (Lei et al., 2016). It is estimated that about 4% of all tandemly arranged genes (two adjacent genes in the same orientation) and 14% of protein-coding genes in the human genome are involved in formation of chimeric transcripts via cis-SAGe (Parra et al., 2006; Kannan et al., 2011). Chimeric transcripts of drug/xenobiotic-metabolizing cytochrome P450 genes have been identified in human liver (Zaphiropoulos, 1999; Finta and Zaphiropoulos, 2000b). As an example, the four-gene *CYP2C* cluster (*CYP2C18-CYP2C19-CYP2C9-CYP2C8)* gives rise to chimeric transcripts containing exons from up to three of these genes (Zaphiropoulos, 1999; Finta and Zaphiropoulos, 2000a).

Despite the extensive clustering of *UGT* genes, chimeric *UGT* transcripts have not yet been reported. *UGT2B15*, *UGT2B29P2*, *UGT2B17*, and *UGT2B29P1* are neighboring genes tandemly arranged within a 165-kb region (*UGT2B15-UGT2B29P2-UGT2B17-UGT2B29P1*) (Turgeon et al., 2000; Ménard et al., 2009). *UGT2B15* and *UGT2B17* are both protein coding genes comprising six canonical exons each; *UGT2B29P1* and *UGT2B29P2* are pseudogenes that do not encode known proteins. We report here the discovery of chimeric transcripts that contain exons from two or three of these four *UGT* genes in prostate cancer cell lines as well as human liver tissues. Most of these chimeric transcripts contain the canonical exons 1–5 of either UGT2B15 or UGT2B17 linked to a 3′-terminal exon from a downstream gene (i.e., from *UGT2B29P2*, *UGT2B17,* or *UGT2B29P1*). These transcripts encode C-terminally truncated UGT2B15 and UGT2B17 proteins with novel C-terminal peptides. The chimeric UGT proteins have no glucuronosyltransferase activity but they negatively regulate glucuronidation activity when coexpressed with wild-type UGTs, suggesting both homotypic and heterotypic interactions.

Materials and Methods

Human Tissues and Cancer Cell Lines. Normal human liver tissues were obtained from the liver bank of the Department of Clinical Pharmacology of Flinders Medical Centre, Flinders University of South Australia, Australia, as previously reported (McKinnon et al., 1991; Bhasker et al., 2000; Hu et al., 2014b). Ethics approval was granted by the Flinders Medical Centre Research Ethics Committee. An RNA panel totaling 20 human tissues (prostate, liver, colon, spleen, lung, testis, kidney, placenta, bladder, brain, adipose, ovary, cervix, heart, skeletal muscle, small intestine, thyroid, thymus, esophagus, trachea) was purchased from Thermo Fisher Scientific (Ambion brand; Waltham, MA). The prostate cancer VCaP and LNCaP cell lines, the breast cancer MCF7 cell line, and the human embryonic kidney (HEK)293 cell line were purchased from American Type Culture Collection (ATCC, Manassas, VA). VCaP and HEK293 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (FBS). MCF7 and LNCaP cells were maintained in RPMI-1640 medium containing 5% FBS.

RNA Preparation and Reverse Transcription. Total RNA was extracted from liver tissues and cancer cell lines using TRIzol reagent according to the manufacturer's protocol (Invitrogen/Thermo Fisher Scientific, Carlsbad, CA). Reverse transcription was carried out using Invitrogen reagents as previously reported (Hu and Mackenzie, 2009; Hu et al., 2015). In brief, total RNA (~1 μg) was treated with DNase I at room temperature for 15 minutes and then reverse-transcribed using Superscript III (50 units) and random hexamer primers (50 ng) at 50°C for 50 minutes in a 20-μl reaction containing 50 mM Tris-HCl, pH 8.0, 75 mM KCl, and 3 mM $MgCl_2$. The resulting cDNAs were diluted five times in RNase-free H_2O and used for polymerase chain reaction (PCR) or quantitative PCR as described below.

PCR Amplification and Cloning of Chimeric UGT cDNA into pCR Blunt Vector. PCRs were conducted to amplify chimeric cDNAs that contained exons from at least two genes of the four adjacent *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*) (Fig. 1A) using cDNA from human tissues or cancer cell lines and Phusion High-Fidelity DNA polymerase according to the manufacturers recommendations (New England Biolabs Ltd., Hitchin, UK). Specifically, chimeric cDNA (Fig. 1B) that contained exons located between *UGT2B15* (either exon 2 or exon 4) and *UGT2B17* (either exon 2 or exon 4) were amplified using primer pairs 2F and 2R, or 4F and 4R, from VCaP and human liver cDNAs. Chimeric cDNA (Fig. 1C) that included exons located between *UGT2B15* exon 1 and *UGT2B17* exon 6 were amplified using primers 15E1-F (+348/+375) and 17E6-R (−257/−281) from VCaP and human liver cDNAs. Chimeric cDNA (Fig. 1D) that contained exons located between *UGT2B17* and *UGT2B29P1* were amplified using the forward primer 17E5F specific to *UGT2B17* exon 5 and the reverse primer specific to *UGT2B29P1* exon 1 (29P1E1R), exon 2 (29P1E2R), or exon 3 (29P1E3R). The sequences of all PCR primers are provided in Table 1. Amplified products were purified using QIAquick PCR purification kit according to the manufacturer's protocol (QIAGEN, Hilden, Germany) and subsequently cloned into the pCR Blunt vector using the Zero Blunt PCR cloning kit (Thermo Fisher Scientific) according to the manufacturer's protocol. Inserts were sequenced using primers M13F and M13R (Thermo Fisher Scientific).

Cloning of PCR Products Encoding Full-Length Wild-Type and Chimeric cDNAs of *UGT2B15* and *UGT2B17* into the pEF-IRESpuro6 Expression Vector. PCRs were carried out to amplify both wild-type and chimeric UGT2B15 or UGT2B17 full-length coding sequences from human tissue and cell line cDNAs using Phusion High-Fidelity DNA Polymerase (New England Biolabs Ltd.) with specific primers (Fig. 1A; Table 1), and PCR cycling conditions of an initial activation step of 98°C for 30 seconds, and 35 cycles of 98°C for 5 seconds, 60°C for 20 seconds, and 72°C for 1 minute. The 1593-base pair (bp) coding sequence of the wild-type UGT2B15 was amplified from VCaP cDNA using primers 15F and 15R6; the 1593-bp coding sequence of the wild-type UGT2B17 was amplified from colon cDNA

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

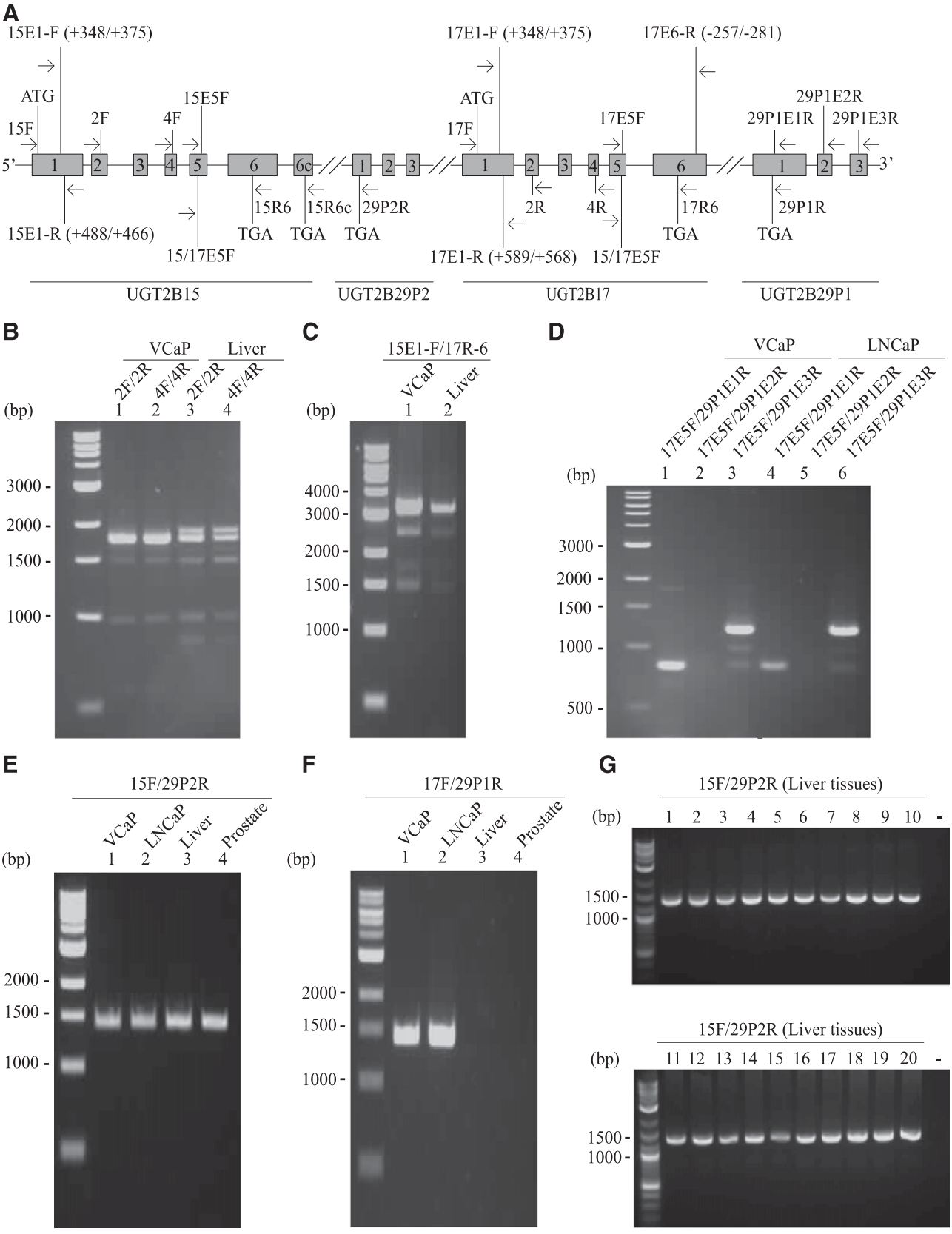


Fig. 1. RT-PCRs using primers that are specific to four neighboring *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*). (A) Schematic representation of four neighboring *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*) showing the positions of exon-specific RT-PCR primers. The primers 15E1-F and 15E1-R (or 17E1-F and 17E1-R) are numbered relative to the "A" of the start codon of UGT2B15 (NM_001076) (or UGT2B17, NM_001077), which is designated +1. The 17E6-R primer sequence is numbered relative to the "T" of the TGA stop codon of UGT2B17 (NM_001077), which is designated −1. (B and C) Agarose gels showing multiple RT-PCR products that were amplified from VCaP cells or human liver using (B) primer pairs 2F/2R or 4F/4R or (C) primer pair 15E1-F/17E6-R. (D) Agarose gels showing RT-PCR products that were amplified from VCaP and LNCaP cells using forward primer 17E5F (specific to

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

TABLE 1

Sequences of primers for PCR and cloning

PCR primers are numbered relative to the "A" of the initiator codon (ATG), which is assigned as "+1", or relative to the "T" of the stop codon (TGA), which is assigned as "−1." Incorporated restriction sites of both MluI (ACGCGT) and XhoI (CTCGAG) are underlined.

Primers	Nucleotide Sequences (5′ to 3′)
RT-PCR 5′ to 3′	
2F	TTTGAATTTCCTCGCCCATT
2R	ATCCCAATAGGTTCGAATGA
4F	TAGGTTCCAATACTCGACTG
4R	AAGTATTTGGCTTCTTGCCA
17E5F	CCAATGGCATCTATGAGGCA
29P1E1R	TATGCGGGAAATGGGGAATG
29P1E2R	TTCAAAATCCCAGTACGTTT
29P1E3R	ACCAGAGCTGTTCCTATTTG
17E6-R (−257/−281)	AACTTTTGTGTGGGGTAACTTTTGG
15E1-F (+348/+375)	GTGTTGGGAATATTATGACTACAGTAAC
15E1-R (+488/+466)	TCAGCCAGTAGCTCACCACAGGG
17E1-F (+348/+375)	GTGTTGGGAATATTCTGACTATAATATA
17E1-R (+589/+568)	CAGGTACATAGGAAGGAGGGAA
15E5-qPCR F	AATGGCATCTATGAGGCG
15/17E5F	TGATAACATTGCTCACATGA
15E6C-qPCR R	AATCACCCGTCTTCTGTGTCf
29P2E1-qPCR R	GGACATGTATTTAACATGCC
29P1E1-qPCR R	ATAGTCCTCCTGGTGCA
Cloning	
15F	ACTGCTCGAGATGTCTCTGAAATGGAC
15R6	GACCACGCGTGGCTTTTGATATAACTAATCT
15R6C	GTATACGCGTAATCACCCGTCTTCTGTGTC
29P2R	CGATACGCGTGGACATGTATTTAACATGCC
17F	ACTGCTCGAGATGTCTCTGAAATGGAT
17R6	GTATACGCGTATAAGGAAGCTCAGTAACT
29P1R	CTAGACGCGTTGTATCAGGAGAATATT

using primers 17F and 17R6. The 1335-bp chimeric full-length coding cDNA (termed chimeric UGT2B15 in this study) containing *UGT2B15* exons 1–5 spliced to *UGT2B29P2* exon 1 was amplified from VCaP cDNA using primers 15F and 29P2R. Chimeric UGT2B15 was also amplified from cDNAs of LNCaP cells, and prostate and liver tissue (Fig. 1, E and G). The 1386-bp chimeric full-length coding cDNA (termed chimeric UGT2B17 in this study) containing *UGT2B17* exons 1–5 spliced to *UGT2B29P1* exon 1 was amplified from VCaP cDNA using primers 17F and 29P1R. Chimeric UGT2B17 was also amplified from LNCaP cDNA but not liver and prostate cDNAs (Fig. 1F). The amplification products were confirmed on ethidium bromide–stained 1% agarose gels, digested with XhoI and MluI (New England BioLabs Ltd.), purified using QIAquick PCR purification kit (QIAGEN), and then cloned into the XhoI/MluI sites of the pEF_IRESpuro6 expression vector. The identities of the resultant constructs were confirmed by DNA sequencing. The vector that expressed wild-type UGT2B7 (pEF_IRESpuro6/wild-type UGT2B7) or C-terminal cMYC-tagged UGT2B7 (pEF_IRESpuro6/wild-type UGT2B7/cMYC) was previously reported (Jin et al., 1993; Stone et al., 2003; Lewis et al., 2011).

Quantitative Real-Time Polymerase Chain Reaction. Isoform specific quantitative real-time (qPCR) assays were designed for wild-type and chimeric UGT2B15 and UGT2B17 transcripts and the UGT2B15/6C splice variant transcript (containing UGT2B15 exons 1–5 and exon 6C). The isoform-specific primers (Fig. 1A; Table 1), amplified:

1. wild-type UGT2B15 [15E1-F (+348/+375) and 15E1-R (+488/+466)],
2. wild-type UGT2B17 [17E1-F (+348/+375) and 17E1-R (+589/+568)],
3. chimeric UGT2B15 [15E5-qPCR F (specific to *UGT2B15* exon 5) and 29P2E1-qPCR R (specific to *UGT2B29P2* exon 1)],
4. chimeric UGT2B17 [15/17E5F (specific to both *UGT2B15* and *UGT2B17*exon 5s) and 29P1E1-qPCR R (specific to *UGT2B29P1* exon 1), and
5. chimeric UGT2B15/6C [15/17E5F and 15E6c-qPCR R (specific to *UGT2B15* exon 6C)].

Reactions were conducted as previously reported (Hu and Mackenzie, 2009; Hu et al., 2014a) using cDNAs of human tissues and cancer cells (VCaP, LNCaP, and MCF7). To generate the standard calibration curves for the qPCR analyses, we prepared four serial 10-fold dilutions containing known copy numbers (e.g., 6000, 600, 60, and 6) of a pEF_IRESpuro6 expression vector carrying: 1) chimeric UGT2B15 for quantifying both wild-type and chimeric UGT2B15, 2) chimeric UGT2B17 for measuring both wild-type and chimeric UGT2B17 transcripts, or 3) UGT2B15/6C for quantifying variant UGT2B15/6C transcripts. The standard curves allowed copy numbers of each transcript to be calculated in the experimental samples as previously reported (Hu and Mackenzie, 2009; Hu et al., 2014a).

Expression of Wild-Type and Chimeric UGT2B15 and UGT2B17 Enzymes and Western Blotting. HEK293 cells lacking UGT expression were used as cellular models for overexpression of wild-type and chimeric UGTs. HEK293 cells were plated into T25 flasks at 50% confluence and cultured overnight. Cells were transfected with 1 μg of a pEF_IRESpuro6 vector expressing wild-type or

UGT2B17 exon 5) and a reverse primer specific to exon 1 (29P1E1R), exon 2 (29P1E2R), or exon 3 (29P1E3R) of the *UGT2B29P1* gene. (E and G) Agarose gels showing the 1335-bp full coding region of the UGT2B15 chimeric transcript amplified from 1) VCaP, LNCaP, liver, and prostate in (E) or 2) a panel of 20 different human livers in (G) using the forward primer 15F (specific to *UGT2B15* exon 1) and the reverse primer 29P2R (specific to *UGT2B29P2* exon 1). (F) Agarose gel showing the 1386-bp full coding sequence of the UGT2B17 chimeric transcript amplified from VCaP and LNCaP cells using the forward primer 17F (specific to *UGT2B17* exon 1) and the reverse primer 29P1R (specific to *UGT2B29P1* exon 1). The primer sequences are given in Table 1. ATG, initiator codon; F, forward primer; R, reverse primer; TGA, stop codon.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

chimeric UGT2B15 (or UGT2B17) proteins using Lipofectamine 2000. Forty-eight hours post-transfection, whole-cell lysates were prepared in radioimmunoprecipitation assay buffer (50 mM Tris-HCl, pH 8.0, 1% NP-40, 150 mM sodium chloride, 0.5% sodium deoxycholate, and 0.1% sodium dodecyl sulfate). Protein concentrations were determined using the Bradford Protein Assay (Bio-Rad, Gladesville, NSW, Australia). For Western blotting assays, 20 μg of protein of each whole-cell lysate was separated on SDS-polyacrylamide gels (10%) and transferred to nitrocellulose membranes. Membranes were incubated first with the rabbit anti-UGT2B15/UGT2B17 antibody that recognizes both wild-type UGT2B15 and UGT2B17 as previously reported (Wijayakumara et al., 2015) and then with a horseradish peroxidase-conjugated donkey anti-rabbit secondary antibody (NeoMarkers; Fremont, CA). Immunosignals were detected with the SuperSignalWest Pico Chemiluminescent kit (Thermo Fisher Scientific) using an ImageQuant LAS 4000 luminescent image analyzer (GE Healthcare, Chalfont St. Giles, United Kingdom).

Glucuronidation Assays and Western Blotting Assays. To test the effect of chimeric UGT2B15 on UGT2B7-glucuronidation activity, six T25 flasks of HEK293 cells at 50% confluence were cotransfected with pEF_IRESpuro6/UGT2B7 vector and pEF_IRESpuro6/chimeric UGT2B15 in different ratios; the empty pEF_IRESpuro6 vector was added to make up the total amount of DNA in each flask to 9 μg. The exact ratios of pEF_IRESpuro6/UGT2B7 to pEF_IRESpuro6/chimeric UGT2B15 to pEF_IRESpuro6 plasmids in each flask were as follows: 1) 3:0:6; 2) 3:1.5:4.5; 3) 3:3:3; 4) 3:6:0; and 5) 0:3:6. The sixth flask was transfected with 9 μg of empty pEF_IRESpuro6 DNA only. To test the effect of chimeric UGT2B17 on UGT2B7-morphine glucuronidation activity, an identical set of transfections was performed using the pEF_IRESpuro6/chimeric UGT2B17 vector. Forty-eight hours post-transfection, cells were harvested, washed with 1× phosphate-buffered saline twice, and then lysed in 250 μl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.6). Cell lysates were used for morphine glucuronidation assays and Western blotting assays. Morphine glucuronidation assays were carried out using high-performance liquid chromatography using an Agilent 1100 series instrument (Agilent Technologies, Sydney, NSW, Australia) as previously reported (Hu et al., 2014d; Wijayakumara et al., 2017). Western blotting assays were conducted as described above. UGT2B7 protein was detected using a rabbit anti-UGT2B7 antibody as reported elsewhere (Hu et al., 2014b). Immunoblot band densitometry was carried out using Multi Gauge Ver3.0 software (FUJIFILM, Tokyo, Japan).

To test the effect of chimeric UGT2B15 on wild-type UGT2B15-4-MU (methylumbelliferone) glucuronidation activity, four T25 flasks of HEK293 cells were cotransfected with pEF_IRESpuro6/wild-type UGT2B15 vector and pEF_IRESpuro6/chimeric UGT2B15 in different ratios; the empty pEF_IRESpuro6 vector was added to make up the total amount of DNA in each flask to 3 μg. The exact microgram ratios of pEF_IRESpuro6/wild-type UGT2B15 to pEF_IRESpuro6/chimeric UGT2B15 to pEF_IRESpuro6 plasmids in each flask were as follows: 1) 1:0:2; 2) 0:2:1; 3) 1:2:0, and 4) 0:0:3. The fourth flask was transfected with 3 μg of empty pEF_IRESpuro6 DNA only. To test the effect of chimeric UGT2B17 on UGT2B17-4MU glucuronidation activity, an identical set of transfections was performed using the pEF_IRESpuro6/wild-type UGT2B17 vector and the pEF_IRESpuro6/chimeric UGT2B17 vector. Forty-eight hours post-transfection, whole-cell lysates were prepared as described above and used for 4-MU glucuronidation assays as previously reported (Wijayakumara et al., 2015), and Western blotting assays using the anti-UGT2B15/UGT2B17 antibody for detection of both wild-type and chimeric UGT2B15 and UGT2B17.

Coimmunoprecipitation. The pEF_IRESpuro6/wild-type UGT2B7-cMYC vector expressing cMYC-tagged wild-type UGT2B7 was previously reported (Lewis et al., 2011). HEK293 cells in T25 flasks were cotransfected with 5 μg of pEF_IRESpuro6/wild-type UGT2B7-cMYC vector and 5 μg of pEF_IRESpuro6/chimeric UGT2B15 (or pEF_IRESpuro6/chimeric UGT2B17) and cultured for 48 hours. Cells were lysed on ice with 200 μl hypotonic lysis buffer (20 mM Tris pH 7.4, 10 mM $MgCl_2$, 10 mM KCl, 1 mM EDTA, 10% glycerol, 1% Triton-X and 1 mM dithiothreitol) as described previously (Klenova et al., 2002). Lysates from two T25 flasks were pooled together, supplemented to 350 mM NaCl, centrifuged at 14,000*g* for 10 minutes at 4°C to remove cellular debris and precleared for 1 hour with either protein A Sepharose (GE Healthcare) or protein G magnetic beads (Cell Signaling Technology, Danvers, MA). Precleared lysates were incubated overnight at 4°C with 3 μg of mouse anti-cMYC antibody (Santa Cruz Biotechnology, Dallas, TX) or rabbit anti-UGT2B15/17 antibody. Sepharose or magnetic beads were added and incubated for 3 hours. Beads were washed three times with phosphate-buffered saline and bound proteins were eluted in SDS buffer (for Sepharose) or 0.1 M glycine pH 2 (for magnetic beads). Samples precipitated with anti-UGT2B15/2B17 antibody were analyzed by Western blotting with anti-cMYC antibody; samples precipitated with anti-cMYC antibody were analyzed by Western blotting with anti-UGT2B15/2B17 antibody.

Treatment of Breast and Prostate Cancer Cells with Sex Steroids. The prostate cancer VCaP and LNCaP cells were cultured in six-well plates for 2 days in phenol red-free RPMI-medium supplemented with 5% dextran-coated charcoal-stripped FBS and then treated with 0.1% ethanol (vehicle control) or 1 nM dihydrotestosterone for 24 hours. MCF7 cells were cultured in six-well plates for 2 days in phenol red-free RPMI-medium supplemented with 5% dextran-coated charcoal-stripped FBS and then treated with 0.1% ethanol (vehicle control) or 1 nM 17β-estradiol for 24 hours. Total RNA was prepared from treated cells and then subjected to RT-qPCR for measuring both the wild-type and chimeric UGT2B15 and UGT2B17 transcript levels as described above.

Transfection of Cells with MicroRNA. Cells were cultured for 32 hours in phenol red-free RPMI 1640 medium (Invitrogen) supplemented with 5% dextran-coated charcoal-stripped FBS and then replated into six-well plates at approximately 1×10^6 cells per well. After overnight culture, cells were transfected in triplicate with microRNA 376c (miR-376c) mimic or a scrambled control mimic (miR-neg) at 30 nM using 8 μl Lipofectamine 2000 (Invitrogen) per well. Twenty-four hours post-transfection, total RNA was prepared from the cells and RT-qPCR was performed to quantify both the wild-type and chimeric UGT2B15 and UGT2B17 transcript levels as described above.

Results

Discovery of Chimeric Transcripts Containing Exons of *UGT2B15*, *UGT2B29P2*, and *UGT2B17*. UGT2B15 and UGT2B17 mRNAs are highly expressed in prostate cancer VCaP cells (Hu et al., 2010; Wijayakumara et al., 2015). In contrast, there is no evidence for the expression of mRNAs corresponding the *UGT2B29P2* pseudogene that is located 18 kb downstream of *UGT2B15* and 54 kb upstream of *UGT2B17* (LOC728807, NG_022012.1, 2265 bp) (Turgeon et al., 2000; Ménard et al., 2009) (see Fig. 2A). To determine whether transcripts exist that contain exons from both *UGT2B15* and *UGT2B17*, we performed RT-PCR using primer pairs that spanned both genes. Specifically, forward primers 2F or 4F that bind *UGT2B15* exon 2 or exon 4, respectively, were paired with reverse primers 2R or 4R that bind *UGT2B17* exon 2 or exon 4, respectively (Fig. 1A). Primer pairs 2F/2R or 4F/4R generated multiple amplicons (ranging from 1 to 2 kb) from VCaP cDNA suggesting that chimeric (intergene) transcripts indeed exist (Fig. 1B). Cloning and sequencing of these amplicons identified eight different chimeric transcripts that contained exons from two or more of the three adjacent genes *UGT2B15*, *UGT2B17*, and *UGT2B29P2*. These transcripts (Fig. 2B) included four 2F/2R-derived chimeras [designated Ca $(E2\text{-}E5)_{2B15}(E1\text{-}E2)_{2B29P2}(E1\text{-}E2)_{2B17}$, Da $(E2\text{-}E5)_{2B15}(E1\text{-}E2)_{2B29P2}(E2)_{2B17}$, Fa $(E2\text{-}E5)_{2B15}(E2)_{2B29P2}(E1\text{-}E2)_{2B17}$, Ga $(E2\text{-}E5)_{2B15}(E1\text{-}E2)_{2B17}$], and four 4F/4R-derived chimeras [designated

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

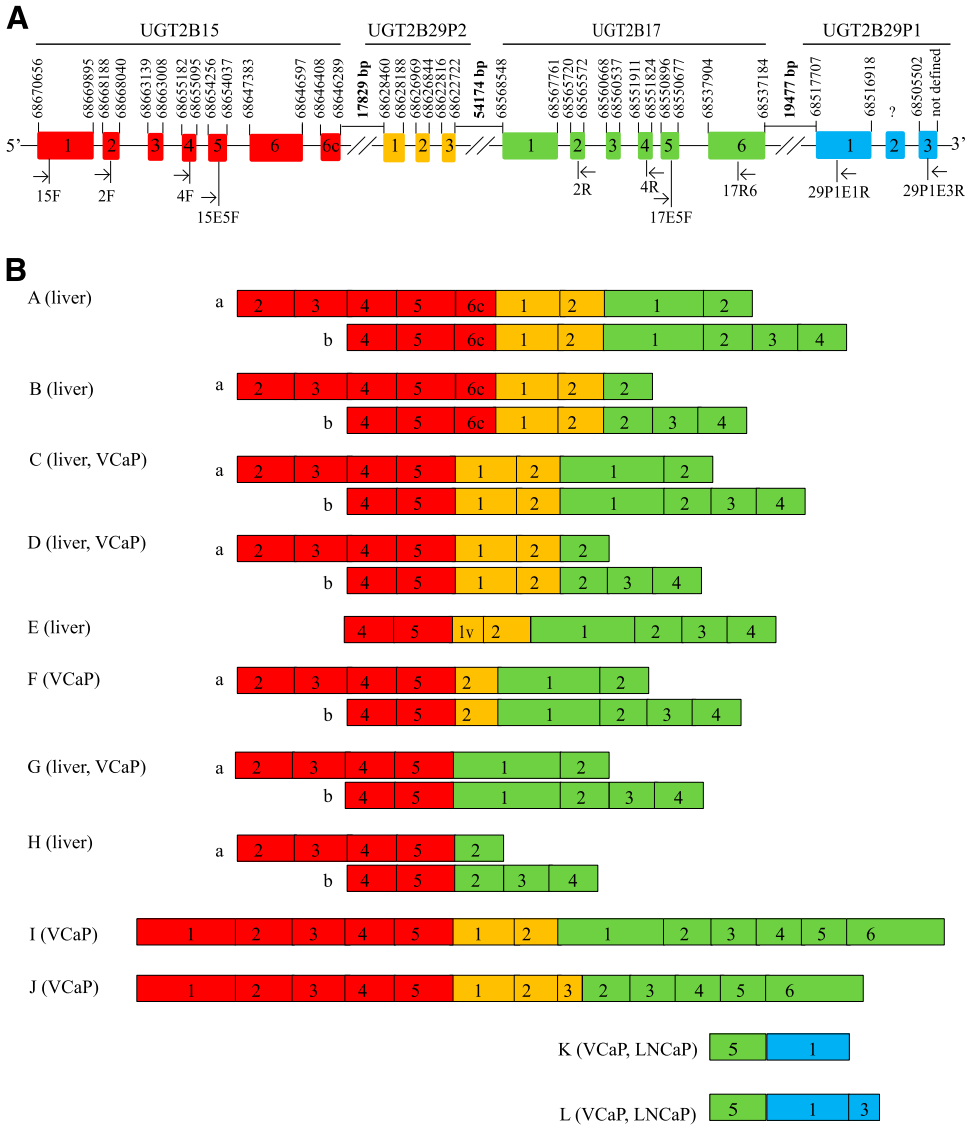


Fig. 2. Exons from four neighboring *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*) are spliced to generate a variety of chimeric transcripts. (A) A schematic diagram showing the gene structures of four neighboring *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*). The genomic coordinates that define the start and end of exons for four neighboring *UGT* genes [*2B15* (red), *2B29P2* (orange), *2B17* (green), *2B29P1* (blue)] are in accordance with the human genome assembly GRCh38/hg38 (2013). Previously identified alternatively spliced exons of *UGT2B15* (1b and 6b) and *UGT2B17* (1b, 1c, 1d, 2b) are omitted for simplicity. (B) Exon structures of chimeric transcripts amplified from different cDNA sources using various primer pairs, including: 1) eight chimeras (Aa, Ba, Ca, Da, Ea, Fa, Ga, Ha) generated using primers 2F/2R, 2) eight chimeras (Ab, Bb, Cb, Db, Eb, Fb, Gb, and Hb) using primers 4F/4R, 3) two chimeras (I and J) using primers 15E1-F/17E6-R, 4) chimera K generated using primers 17E5F/29P1E1R, and 5) chimera L using primers 17E5F/29P1E3R. F, forward primer. R, reverse primer.

Cb (E4-E5)$_{2B15}$(E1-E2)$_{2B29P2}$(E1-E4)$_{2B17}$, Db (E4-E5)$_{2B15}$(E1-E2)$_{2B29P2}$(E2-E4)$_{2B17}$, Fb (E4-E5)$_{2B15}$(E2)$_{2B29P2}$(E1-E4)$_{2B17}$, Gb (E4-E5)$_{2B15}$(E1-E4)$_{2B17}$].

UGT2B15 and *UGT2B17* are highly expressed in human liver, the major organ for drug metabolism and clearance (Court, 2010). Using liver cDNA (a pool from six human liver tissues) as templates, RT-PCR with the primer pairs 2F/2R and 4F/4R produced multiple amplicons that were similar in size to those amplified from VCaP cDNA (Fig. 1B), suggesting the expression of these chimeras in normal liver. Indeed, cloning and sequencing of these liver cDNA-derived amplicons detected the same eight VCaP cDNA-derived chimeras, as well as seven new chimeras including three 2F/2R-derived chimeras [Aa (E2-E5-E6c)$_{2B15}$(E1-E2)$_{2B29P2}$(E1-E2)$_{2B17}$, Ba (E2-E5-E6c)$_{2B15}$(E1-E2)$_{2B29P2}$(E2)$_{2B17}$, Ha (E2-E5)$_{2B15}$(E2)$_{2B17}$] and four 4F/4R-derived chimeras [Ab (E4-E5-E6c)$_{2B15}$(E1-E2)$_{2B29P2}$(E1-E4)$_{2B17}$, Bb (E4-E5-E6c)$_{2B15}$(E1-E2)$_{2B29P2}$(E2-E4)$_{2B17}$, E (E4-E5)$_{2B15}$(E1v-E2)$_{2B29P2}$(E1-E4)$_{2B17}$, and Hb (E4-E5)$_{2B15}$(E2-E4)$_{2B17}$].

Analysis of the exonic structures of the 15 chimeric sequences listed above suggested that they could be parsed into eight distinct chimeric transcripts designated A, B, C, D, E, F, G, H (Fig. 2B). Because our initial PCR amplification experiments used forward primers in UGT2B15 exon 2 or 4, the cloned chimeric amplicons do not contain the 5′ end of the native UGT2B15 transcript that includes the ATG codon (i.e., exon 1). To test whether transcripts exist that include the 5′ end of UGT2B15 (and the 3′ end of UGT2B17), we repeated RT-PCR using a primer pair (15E1-F/17E6-R) that spans from *UGT2B15* exon 1 to *UGT2B17* exon 6. These primers generated multiple amplicons ranging from 2 to 3.5 kb in length from VCaP or liver cDNA (Fig. 1C). Cloning and sequencing the 15E1-F/17E6-R-derived amplicons from VCaP cDNA identified two chimeras that spanned from *UGT2B15* exon 1 to *UGT2B17* exon 6 and also included exons from *UGT2B29P2* (Fig. 2B), namely, chimera I [(E1-E5)$_{2B15}$(E1-E2)$_{2B29P2}$(E1-E6)$_{2B17}$] and J [(E1-E5)$_{2B15}$(E1-E3)$_{2B29P2}$(E2-E6)$_{2B17}$].

In summary, multiple chimeric transcripts (A to J) containing exons from three *UGT* genes (*2B15*, *2B29P2*, and *2B17*) were discovered in prostate cancer VCaP cells and human liver. Supplemental Figs. 1–9 show the sequencing

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

chromatograms corresponding to the chimeric parts of these chimeric transcripts.

Discovery of Chimeric Transcripts Containing Exons of *UGT2B17* and *UGT2B29P1*. The *UGT2B29P1* pseudogene is located downstream of *UGT2B17* (AF179880, 1021 bp) and is highly conserved in structure and sequence with *UGT2B29P2* (95% identity) (Ménard et al., 2009). Given that *UGT2B29P2* exons were spliced to *UGT2B15* exon 5 in numerous chimeras (A to F, see Fig. 2), we hypothesized that a similar splicing of *UGT2B29P1* exons to *UGT2B17* exon 5 might occur. To test this hypothesis, RT-PCR was carried out using the forward primer 17E5F (specific to *UGT2B17* exon 5) and a reverse primer that was specific to *UGT2B29P1* exon 1 (29P1E1R), exon 2 (29P1E2R), or exon 3 (29P1E3R). RT-PCR with the primer pairs 17E5F/29P1E1R and 17E5F/29P1E3R generated products from VCaP and LNCaP cDNA (Fig. 1D) but not from liver cDNA (data not shown). No products were generated using reverse primer 29P1E2R in *UGT2B29P1* exon 2 (Fig. 1D). Cloning and sequencing the VCaP cDNA-derived amplicons identified two chimeras, namely chimera K $(E5)_{2B17}(E1)_{2B29P1}$ with *UGT2B29P1* exon 1 spliced to *UGT2B17* exon 5, and chimera L $(E5)_{2B17}(E1\text{-}E3)_{2B29P1}$ with *UGT2B29P1* exons 1 and 3 spliced to *UGT2B17* exon 5. Supplemental Fig. 10 shows the sequencing chromatogram spanning the chimeric part of chimera L.

Discovery of Novel Exons from Pseudogenes *UGT2B29P1* and *UGT2B29P2* and Variant Exons from *UGT2B15* and *UGT2B17*. Novel exons from *UGT2B29P1* and *UGT2B29P2* and variant exons from *UGT2B15* and *UGT2B17* found in the chimeras (Fig. 2) are described further in this section. The pseudogenes *UGT2B29P1* and *UGT2B29P2* were originally identified on the basis of their sequence similarity with *UGT2B* genes (Turgeon et al., 2000; Ménard et al., 2009); however, the expression of their predicted mRNAs has not been reported in tissues or cell lines. Six exons found in the new chimeric transcripts were transcribed from either *UGT2B29P1* (E1, E3) or *UGT2B29P2* (E1, E1v, E2, E3). Among the two *UGT2B29P1* exons, E1 (790 bp) is contained within the previously predicted 1021-bp *UGT2B29P1* sequence (AF179880) (Supplemental Fig. 11A), and E3 (95 bp) is novel and is located approximately 11 kb downstream of exon 1 (Supplemental Fig. 11B). Among the four *UGT2B29P2* exons, E1 (273 bp) and E2 (126 bp) are contained within the predicted 2265-bp *UGT2B29P2* seqeunce (LOC728807) (Supplemental Fig. 11C); E3 (95 bp) is novel and is located approximately 4 kb downstream of E2 (Supplemental Fig. 11D). In addition, exon 1v (E1v) in chimera E contains only the last 18 nucleotides of *UGT2B29P2* E1, suggesting an alternative splice acceptor site for this exon.

When aligned to the UGT2B17 mRNA reference sequence (NM_001077), the *UGT2B17* exon 1 (E1) seen in five chimeras (A, C, E, G, I) begins 21 nucleotides upstream of the reference sequence E1 (Supplemental Fig. 12A), again suggesting a different splice acceptor site. Sequence alignment showed that the E1 sequences of *UGT2B15*, *UGT2B29P2*, *UGT2B17*, and *UGT2B29P1* have a high similarity at their 3′-ends; however, the *UGT2B29P2* E1 is shorter and lacks approximately 60% of the 5′-end sequences of the other three genes (Supplemental Fig. 12A). The exon 2 (E2) sequences of the four genes are also conserved at their 3′-ends but divergent at their 5′-ends (Supplemental Fig. 12B). The exon 3 (E3) sequences of the four genes are similar but the E3 sequences from the two pseudogenes are shorter and divergent from *UGT2B15* and *UGT2B17* at both the 5′- and 3′-ends (Supplemental Fig. 12C).

UGT2B15 exon 6C (178 bp) is a recently reported alternative exon located immediately downstream of the canonical exon 6 (Tourancheau et al., 2016). We found two variant forms of exon 6C in our chimeras, namely exons 6Cv1 (76 bp) in chimera A and exon 6Cv2 (120 bp) in chimera B. All three exon 6Cs start at the same nucleotide C (68646408) but exons 6Cv1 and 6Cv2 end 111 and 67 nucleotides earlier than exon 6C, respectively (Supplemental Fig. 11E).

Most canonical exons of human genes are all flanked by the dinucleotide "AG" at the 5′-end and the dinucleotide "GT" at the 3′-end (Breathnach et al., 1978; Mount, 1982). As shown in Supplemental Fig. 11, all of the novel and variant exons from the four *UGT* genes that were found in the chimeras in this study conform to this "GT-AG" rule.

Chimeric Transcripts Encode Novel Variant UGT2B15 and UGT2B17 Proteins. Although only the I and J chimeric amplicons were generated using a primer in *UGT2B15* exon 1, we proposed that all the chimeras generated using internal primers (A, B, C, D, E, F G) may also be derived from longer transcripts that include the *UGT2B15* exon 1. With this assumption as a basis, we predicted the existence of six variant UGT2B15 proteins that include the previously defined type 1 variant (with exon 6C in place of canonical exon 6) and five new variants termed types 2–6. All of the variant UGT2B15 proteins would have a common 438-aa N-terminus encoded by *UGT2B15* exons 1–5 and a unique short C-terminal peptide that is encoded by exon 6C, or an exon from *UGT2B29P2* or *UGT2B17* (Fig. 3). Transcripts with *UGT2B15* exon 6C (chimeras A and B), *UGT2B29P2* exon 1 (chimeras C and D), *UGT2B29P2* exon 1v (chimera E), *UGT2B29P2* exon 2 (chimera F), *UGT2B17* exon 1 (chimera G), and *UGT2B17* exon 2 (chimera H) spliced to *UGT2B15* exon 5 would encode proteins type 1 (439 aa), type 2 (444 aa), type 3 (446 aa), type 4 (438 aa), type 5 (443 aa), and type 6 (444 aa), respectively (Fig. 3). Chimeras K and L contain *UGT2B29P1* exon 1 spliced to the *UGT2B17* exon 5 and are both predicted to encode the same 461 aa variant UGT2B17 protein.

To assess the overall prevalence of protein-encoding chimeric UGT2B15 and UGT2B17 transcripts in prostate cancer cell lines and in prostate and liver tissues by RT-PCR, we used primer pair 15F/29P2R that amplifies the full chimeric UGT2B15 coding sequence (1335 bp), and primer pair 17F/29P1R that amplifies the full chimeric UGT2B17 coding sequence (1386 bp), respectively (see Fig. 1, E–G). The full chimeric UGT2B15 sequence was efficiently amplified from VCaP, LNCaP, liver tissue, and prostate tissue samples, whereas the full chimeric UGT2B17 sequence was only amplified from the cell lines. In addition, the 1320-bp full-length coding sequence of the type 1 UGT2B15 variant protein (containing exon 6C) was successfully amplified from VCaP cells and liver tissues (data not shown). Quantification of mRNA levels using RT-qPCR revealed low expression of the UGT2B15 chimeric transcript, ranging from 0.12% to 2.04% of the levels of wild-type UGT2B15 transcript in a panel of 20 liver tissues (Fig. 4A). The expression level of the previously identified type 1 UGT2B15 variant was slightly higher, ranging from 0.49% to 6.02% of wild-type transcript levels in the same samples (Fig. 4B). Western blotting assays

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

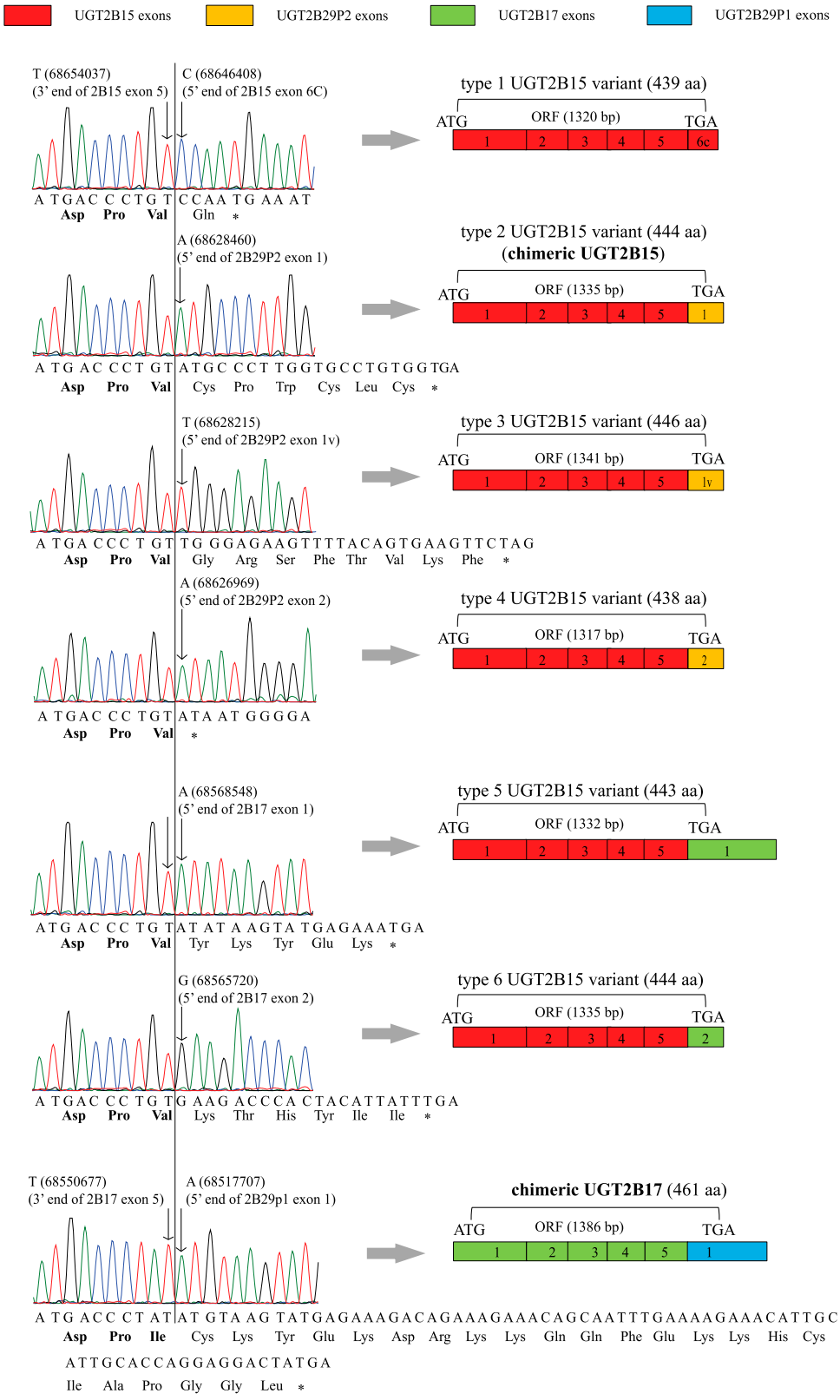


Fig. 3. Chimeric transcripts predict a group of C-terminally truncated UGT2B15 or UGT2B17 proteins. Sequencing chromatograms (left) show the novel splicing junctions (genomic coordinates according to GRCh38/hg38). The novel splicing junctions are indicated by a vertical line. The last three amino acids encoded by *UGT2B15* (or *UGT2B17*) exon 5 are in bold. (Right) Schematic showing the exon structures of the seven predicted transcripts that encode six C-terminally truncated UGT2B15 proteins (types 1–6) and one C-terminally truncated UGT2B17 protein (termed chimeric UGT2B17 in subsequent functional studies). Type 2 UGT2B15 protein (termed chimeric UGT2B15 in subsequent functional studies) contains a novel 6-residue C-terminal peptide resulting from the replacement of *UGT2B15* exon 6 by UGT2B29P2 exon 1. Likewise, chimeric UGT2B17 protein contains a novel 23-residue peptide that is generated by the replacement of the *UGT2B17* exon 6 by *UGT2B29P1* exon 1. ATG, initiator codon; ORF, open reading frame; TGA, stop codon; *, the end of translation.

with an anti-UGT2B15 antibody showed abundant expression of wild-type UGT2B15 protein in all 11 human liver microsome samples. In most liver samples, we also detected a lower immunoreactive protein band with the same mobility as the chimeric UGT2B15 protein that was heterologously expressed in HEK239T cells (Supplemental Fig. 13). Lysates from HEK293T cells overexpressing wild-type UGT2B15 did not show the lower immunoreactive band, suggesting that it is not a proteolytic breakdown product. In the absence of an isoform-specific antibody, this lower band cannot be unequivocally

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

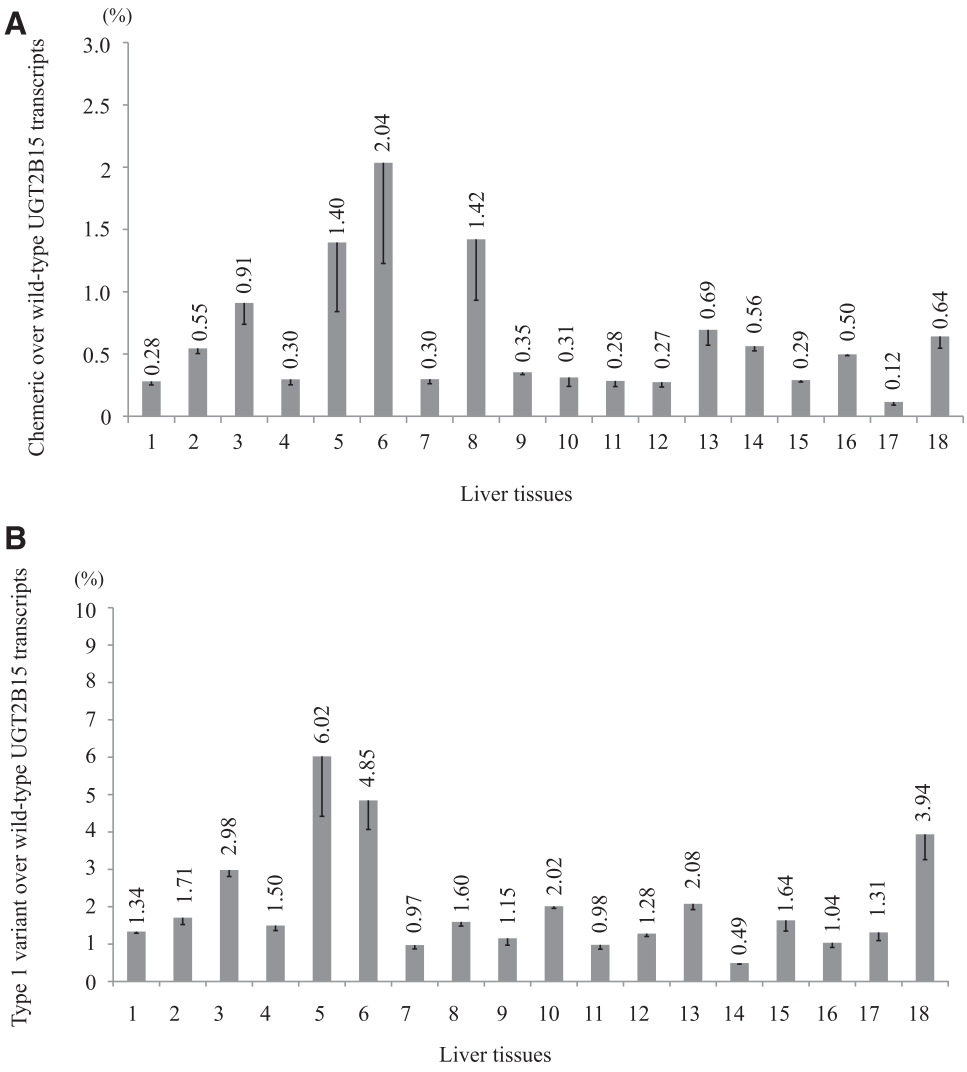


Fig. 4. Quantitative measurement of type 1 and type 2 UGT2B15 variant transcripts in normal human liver tissues. The type 1 variant transcript has exon 6C replacing the canonical exon 6, whereas the type 2 variant (chimeric) transcript has UGT2B29P2 exon 1 replacing the canonical exon 6. RT-qPCR analysis of cDNA samples from 18 human normal liver tissues using primers specific to wild-type, type 1, or type 2 variant UGT2B15 transcripts. The expression levels of variant UGT2B15 transcripts are presented relative to those of the wild-type UGT2B15 transcripts (set as a value of 100%). Data shown are the means ± 1 S.D. from three independent measurements of the ratios of chimeric over wild-type UGT2B15 transcripts (A) or type 1 variant over wild-type UGT2B15 transcripts (B).

identified; however, the result suggests that the chimeric UGT2B15 protein is present in the human liver at low levels.

Inhibition of UGT2B7-Mediated Morphine Glucuronidation Activity by Chimeric UGT2B15 and UGT2B17 Proteins Via Direct Protein-Protein Interaction. The chimeric UGT proteins identified in this study lack the exon 6-encoded transmembrane and cytoplasmic domains and thus are predicted to be inactive (Mackenzie, 1986; Radominska-Pandya et al., 2010). Recent studies have shown that similar C-terminally truncated UGT variants repress the glucuronidation activity of wild-type UGTs via direct protein-protein interaction (Bellemare et al., 2010; Ménard et al., 2013). To test whether chimeric UGT2B15 and UGT2B17 proteins could repress glucuronidation, we constructed four pEF_IRESpuro6 expression vectors that expressed wild-type UGT2B15 (pEF_IRESpuro6/wild-type UGT2B15), wild-type UGT2B17 (pEF_IRESpuro6/wild-type UGT2B17), chimeric (type 2) UGT2B15 (pEF_IRESpuro6/chimeric UGT2B15), or chimeric UGT2B17 (pEF_IRESpuro6/chimeric UGT2B17). Western blotting analysis of HEK293 cells transfected with these pEF_IRESpuro6 expression vectors showed production of these four proteins at their expected sizes (Fig. 5). The calculated apparent molecular weight of wild-type UGT2B15 (530 aa), wild-type UGT2B17 (530 aa), chimeric UGT2B15 (444 aa), and chimeric UGT2B17 (461 aa) were approximately 50, 50, 45, 47 kDa, respectively.

To test whether chimeric UGT2B15 and UGT2B17 proteins could repress glucuronidation activity, pEF_IRESpuro6/wild-type UGT2B7 was cotransfected into HEK293 cells with pEF_IRESpuro6/chimeric UGT2B15 or with pEF_IRESpuro6/chimeric UGT2B17 at various ratios (0.5, 1, or 2). As shown in Fig. 6A, chimeric UGT2B15 alone had no morphine glucuronidation activity but it inhibited the morphine glucuronidation activity of the coexpressed UGT2B7 by 40%–60% ($P < 0.05$). Likewise, chimeric UGT2B17 had no activity toward morphine but it inhibited UGT2B7-mediated morphine glucuronidation by 70%–85% ($P < 0.001$) (Fig. 6B).

Analysis of the human liver RNA panel indicated that the chimeric UGT2B15 is expressed at low levels relative to wild-type UGT2B15 (Fig. 4). Chimeric UGT2B15 transcript levels were also less than 2% of wild-type UGT2B7 transcript levels (not shown). These low expression ratios suggest that endogenously produced chimeric UGT2B15 probably does not significantly inhibit UGT2B7 activity in liver. A wider screen of

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

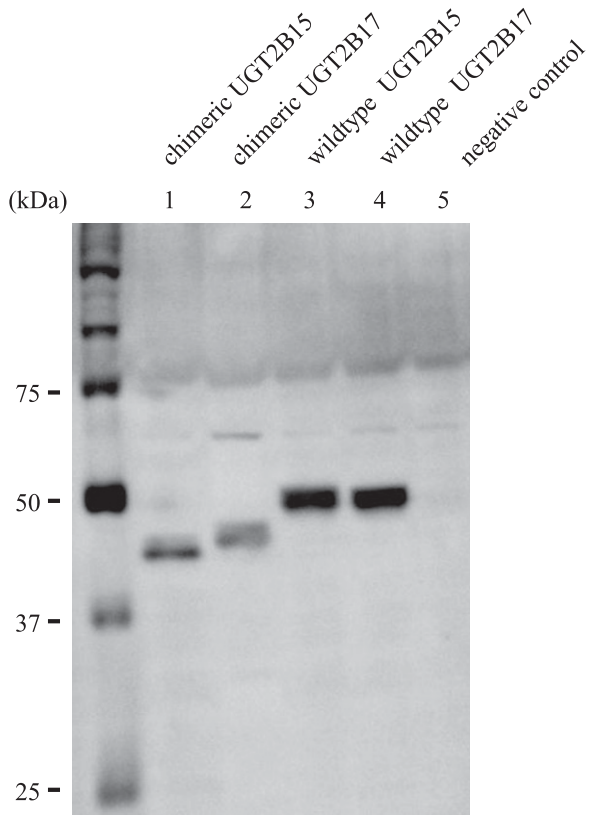


Fig. 5. Expression of chimeric UGT2B15 and UGT2B17 variant proteins in HEK293 cells assessed by Western blotting. HEK293 cells were transfected with either empty pEF_IRESpuro6 vector or pEF_IRESpuro6 vectors expressing wild-type UGT2B15, wild-type UGT2B17, chimeric UGT2B15, or chimeric UGT2B17; 48 hours post-transfection whole-cell lysates were analyzed by Western blotting using the rabbit anti-UGT2B15/UGT2B17 antibody that recognizes both wild-type and chimeric UGT2B15 and UGT2B17 proteins. Molecular weight markers are indicated on the left of the figure.

20 different human tissues revealed that chimeric UGT2B15 transcripts are also present at low levels in prostate, testis, lung, kidney, esophagus, small intestine, and colon (not shown). In most tissues, the ratio of chimeric UGT2B15 transcripts to wild-type UGT2B15, UGT2B17, or UGT2B7 transcripts was also low (<2%); however, in testis we found that the relative ratio of chimeric UGT2B15 to UGT2B7 was >60% (Supplemental Fig. 14). This ratio would be predicted to be inhibitory on the basis of our in vitro studies (Fig. 6).

To investigate the mechanisms underlying the inhibitory effects of chimeric UGT2B15 and UGT2B17 on wild-type UGTB7, coimmunoprecipitation experiments were performed using HEK293 cells that coexpressed cMYC-tagged UGT2B7 and chimeric UGT2B15 or chimeric UGT2B17. Western blotting analysis of the anti-cMYC-precipitated samples using the anti-UGT2B15/2B17 antibody showed that chimeric UGT2B15 (or chimeric UGT2B17) proteins were coprecipitated (Fig. 7A). Conversely, Western blotting of the anti-UGT2B15/2B17-precipitated samples using the anti-cMYC antibody revealed that UGT2B7 protein had been coprecipitated (Fig. 7B). None of the three proteins were detected in the control preimmune sera-immunoprecipitated (IP) samples (Fig. 7). These data suggest that the chimeric UGT2B15 and UGT2B17 proteins negatively regulate UGT2B7 activity via heterodimerization.

Chimeric UGT2B15 and UGT2B17 Variants Inhibit the Activity of Their Wild-Type Counterparts as Assessed by 4-Methylumbelliferone Glucuronidation Assays. To test whether chimeric UGT2B15 and UGT2B17 proteins could repress the glucuronidation activity of their wild-type counterparts, both wild-type and chimeric UGT2B15 (or UGT2B17) were coexpressed at a chimeric to wild-type ratio of 2:1 in HEK293 cells; the cells were harvested 48 hours post-transfection for glucuronidation assays using 4-MU. As shown in Fig. 6C, chimeric UGT2B15 did not glucuronidate 4-MU but it reduced the 4-MU-glucuronidation activity of wild-type UGT2B15 by 24% ($P < 0.01$). Likewise, chimeric UGT2B17 did not conjugate 4-MU but it repressed the 4-MU glucuronidation activity of wild-type UGT2B17 by 36% ($P < 0.001$) (Fig. 6D). The potential interaction between wild-type and chimeric UGT2B15 (or UGT2B17) was not assessed owing to the lack of an antibody specific for chimeric UGT2B15 or UGT2B17.

Regulation of Wild-Type and Chimeric UGT2B15 and UGT2B17 Transcripts by Steroids in Prostate and Breast Cancer Cells. It is well known that androgens down-regulate both the *UGT2B15* and *UGT2B17* genes in prostate cancer cell lines through recruitment of the androgen receptor (AR) at their proximal promoters (Chouinard et al., 2007; Bao et al., 2008). Using RT-qPCRs assays with transcript-specific primers, we showed here that dihydrotestosterone at 1 nM significantly reduced both the wild-type and chimeric UGT2B15 (Fig. 8A) and UGT2B17 (Fig. 8B) transcript levels in LNCaP or VCaP cells. Estrogens are known to upregulate both wild-type *UGT2B15* and *UGT2B17* in breast cancer cell lines via estrogen receptor α binding to an estrogen response unit at their proximal promoters (Harrington et al., 2006; Hu and Mackenzie, 2009; Hu et al., 2016). We showed here that 17β-estradiol (E2) at 1 nM significantly increased both the wild-type and chimeric UGT2B15 transcript levels in breast cancer MCF7 cells (Fig. 8C). The chimeric UGT2B17 transcripts were not detected by RT-qPCR in MCF7 cells (data not shown). Overall, these results indicate that wild-type and chimeric UGT2B15 and UGT2B17 transcripts are regulated at the transcriptional level by sex hormones in prostate and breast cancer cells in a similar manner.

miR-376c Reduces Wild-Type UGT2B15 and UGT2B17 Transcript Levels but Not Chimeric UGT2B15 and UGT2B17 Transcript Levels in Prostate Cancer Cells. We and others have shown that miR-376c regulates UGT2B15 and UGT2B17 via a conserved miR-376c target site located in their 3′-untranslated regions (3′-UTR) (Wijayakumara et al., 2015; Margaillan et al., 2016). Both chimeric UGT2B15 and UGT2B17 transcripts lack the terminal exon (exon 6) that contains the wild-type 3′-UTR and the miR-376c target site. Hence the chimeric transcripts are not expected to be regulated by miR-376c. To test this hypothesis, we transfected either miR-376c mimics or a scrambled control (miR-neg) into prostate cancer LNCaP and VCaP cells and measured the expression levels of wild-type and chimeric UGT2B15 and UGT2B17 transcripts using RT-qPCR. As expected, miR-376c significantly reduced the levels of wild-type UGT2B15 and UGT2B17 transcripts; however, it did not significantly alter the levels of chimeric UGT2B15 and UGT2B17 transcripts (Fig. 9).

Discussion

Recent studies have identified over 130 different alternatively spliced UGT transcripts (Girard et al., 2007; Lévesque

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

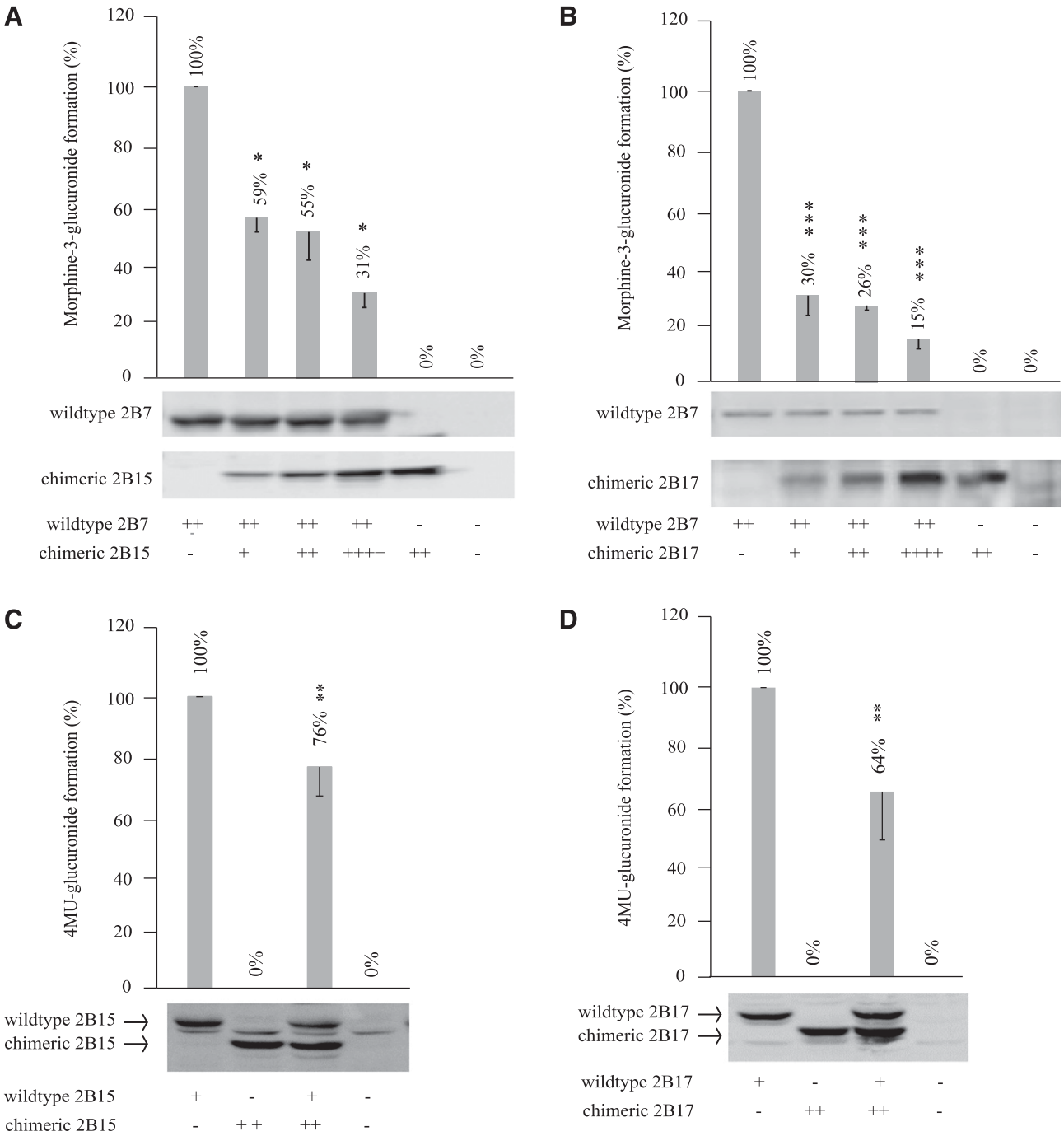


Fig. 6. Chimeric UGT2B15 and UGT2B17 proteins repress the glucuronidation activity of wild-type UGTs. (A and B) HEK293 cells were cotransfected with pEF_IRESpuro6/wild-type UGT2B7 vector and the pEF_IRESpuro6/chimeric UGT2B15 (A) or the pEF_IRESpuro6/chimeric UGT2B17 vector (B). Forty-eight hours post-transfection, whole-cell lysates were subjected to morphine glucuronidation assays. (C and D) HEK293 cells were cotransfected with pEF_IRESpuro6/wild-type UGT2B15 (C) or UGT2B17 (D) vector and the pEF_IRESpuro6/chimeric UGT2B15 or UGT2B17 at the varying ratios described in the *Materials and Methods*. Forty-eight hours post-transfection, whole-cell lysates were subjected to 4-MU glucuronidation assays. For Western blotting assays, the anti-UGT2B15/UGT2B17 antibody was used for detection of both wild-type and chimeric UGT2B15 and UGT2B17, and the anti-cMYC antibody was used to detect the cMYC-tagged UGT2B7. Morphine-glucuronidation activities (formation of morphine-3-glucuronide) were normalized to the UGT2B7 band intensities in each sample and then presented relative to that of the cells transfected with UGT2B7 alone (set as a value of 100%). Likewise, the 4-MU-glucuronidation activities were normalized to the wild-type UGT2B15 (or UGT2B17) band intensities in each sample and then presented relative to that of the cells transfected with wild-type UGT2B15 or UGT2B17 only (set as a value of 100%). Data shown are means ± 1 S.D. from three independent experiments. Statistical analysis was conducted by one-way analysis of variance followed by Tukey's multiple comparison test using GraphPad Prism version 7.03. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

et al., 2010; Ménard et al., 2011; Bushey and Lazarus, 2012; Tourancheau et al., 2016). By contrast, chimeric transcripts containing exons from two or more different *UGT* genes have not yet been reported. We report here the discovery of 12 chimeric transcripts containing exons from two or three genes of the four neighboring *UGT* genes (*UGT2B15*,

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

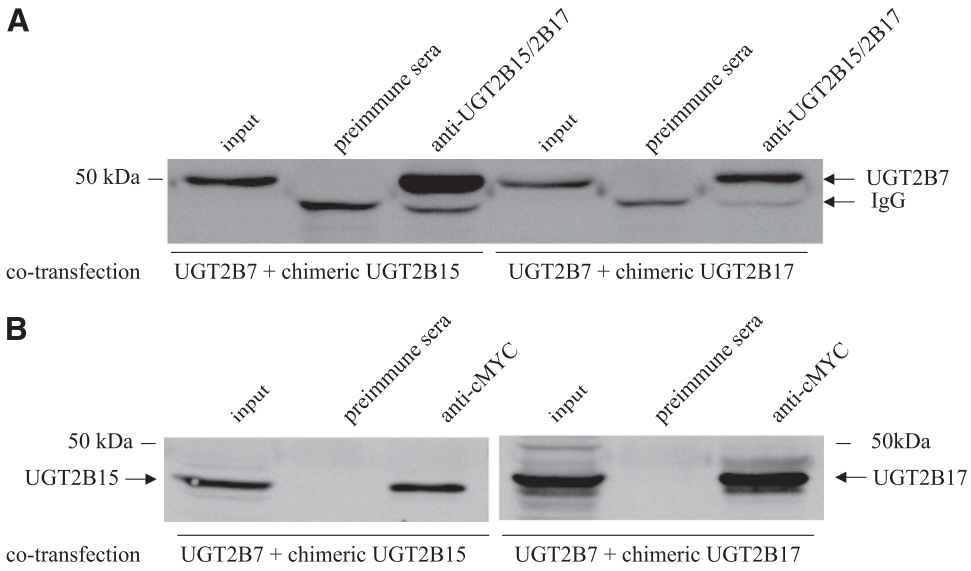


Fig. 7. Coimmunoprecipitation of cMYC-tagged UGT2B7 with chimeric UGT2B15 and UGT2B17 proteins. HEK293 cells were cotransfected with a pEF_IRESpuro6 vector expressing cMYC-tagged UGT2B7 and a pEF_IRESpuro6 vector expressing chimeric UGT2B15 or chimeric UGT2B17 protein. Forty-eight hours post-transfection, cell lysates were immunoprecipitated with either a mouse anti-cMYC antibody or a rabbit anti-UGT2B15/UGT2B17 antibody. (A) Anti-UGT2B15/2B17 antibody-precipitated samples analyzed by Western blotting using the anti-cMYC antibody for detection of cMYC-tagged UGT2B7. (B) Anti-cMYC-precipitated samples analyzed by Western blotting using the anti-UGT2B15/UGT2B17 antibody for detection of chimeric UGT2B15 or UGT2B17 proteins.

UGT2B29P2, *UGT2B17*, and *UGT2B29P1*). All of the splice junctions seen in the chimeras, including those involving newly discovered exons from *UGT2B29P1* and *UGT2B29P2* and the variant exons from *UGT2B15* and *UGT2B17*, conform to the "GT-AG" rule (Burset et al., 2000; Al-Balool et al., 2011), suggesting that the chimeras are genuine splicing products and not the result of translocations or other artifacts (Zaphiropoulos, 1998). A recent study estimated that, when combined, alternative transcripts contribute 5%–61% of the total expression level of UGT1 isoforms and 8%–100% of the total expression level of UGT2 isoforms in normal human liver, kidney, and colon tissues (Tourancheau et al., 2018). In particular, alternative UGT2B15 variants were found to represent up to 6% of UGT2B15 hepatic expression (Tourancheau et al., 2018). We found that the chimeric UGT2B15 transcript levels represented up to about 2% of wild-type UGT2B15 transcript levels in human livers, somewhat less than the previously characterized type I variant containing exon 6C (at up to 6%).

Chimeric transcripts can be generated by trans-splicing or cis-splicing (Chwalenia et al., 2017). Trans-splicing is a non-canonical splicing process by which exons of two different primary transcripts are spliced into a single mRNA (Lei et al., 2016). Cis-SAGe involves splicing exons of adjacent genes into chimeric transcripts; the RNA substrate is generally a read-through multigene primary transcript (Parra et al., 2006; Kannan et al., 2011). *UGT2B15*, *UGT2B29P2*, *UGT2B17*, and *UGT2B29P1* are located adjacent to each other within a 165-kb region in a head-to-tail orientation (*UGT2B15-UGT2B29P2-UGT2B17-UGT2B29P1*) (Turgeon et al., 2000; Ménard et al., 2009). Cis-SAGe is thus the probable mechanism underlying the formation of the chimeric transcripts described herein. Chimeras I and J contained exons from three genes (*UGT2B15*, *UGT2B29P2*, *UGT2B17*), thus indicating that the transcription machinery is able to read through these three neighboring genes, generating a three-gene primary RNA transcript. Chimeras A–J had different combinations of exons, most probably derived from alternative splicing of this three-gene primary RNA transcript. The most frequent cis-SAGe splicing pattern (termed the 2'-2' rule) occurs between the second-to-last exon of the 5′ gene and the second exon of the 3′ gene (Parra et al., 2006; Jia et al., 2016; Chwalenia et al., 2017). Among the 12 chimeras we identified, two were formed according to this 2'-2' rule: Chimera F had *UGT2B15* exon 5 spliced to *UGT2B29P2* exon 2 and chimera H had *UGT2B15* exon 5 spliced to *UGT2B17* exon 2. The last exon of the upstream gene lacks a splicing donor splice site and hence it is always excluded from chimeric transcripts. This explains the lack of *UGT2B15* exon 6 in most of the chimeras (C–J), and *UGT2B17* exon 6 in chimeras K and L. The first exon of the downstream gene in a multigene primary transcript lacks a splicing acceptor site and is typically excluded from chimeric transcripts, unless it contains a cryptic acceptor site. In five of our chimeras (A, C, E, G, I) *UGT2B17* exon 1 is incorporated via a cryptic acceptor site located 21 nucleotides upstream of the *UGT2B17* transcription start site.

Pseudogenes are found throughout the *UGT1* and *UGT2* loci (Mackenzie et al., 2005). The six novel exons contained within the chimeric transcripts described here were transcribed from the pseudogenes *UGT2B29P1* and *UGT2B29P2*. These pseudogenes were predicted on the basis of their genomic sequence similarities to functional *UGT2B* genes; however, their transcription had not been previously reported (Turgeon et al., 2000; Ménard et al., 2009). RNA transcripts from pseudogenes, it has been suggested, serve as miRNA decoys and antisense RNAs that may affect the parental gene expression (Pei et al., 2012). Our work suggests another way that pseudogene sequences contribute to the control of the transcriptional output of the parent genes, that is, by incorporation into chimeric transcripts. Interestingly, similar chimeric transcripts containing exons from pseudogenes and neighboring functional genes have been reported for the *CYP2C* and *CYP3A* genes (Finta and Zaphiropoulos, 2000b; Warner et al., 2001). Finta and Zaphiropoulos (2000b) reported a 15-exon variant CYP3A7 transcript containing two exons from the *CYP3AP1* pseudogene spliced to the *CYP3A7* terminal exon. This CYP3A7/CYP3AP1 chimeric transcript encodes a C-terminally variant CYP3A7 protein with the last four amino acids replaced by a novel 36-amino acid sequence encoded by the *CYP3AP1* exons. On the basis of our studies, four

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

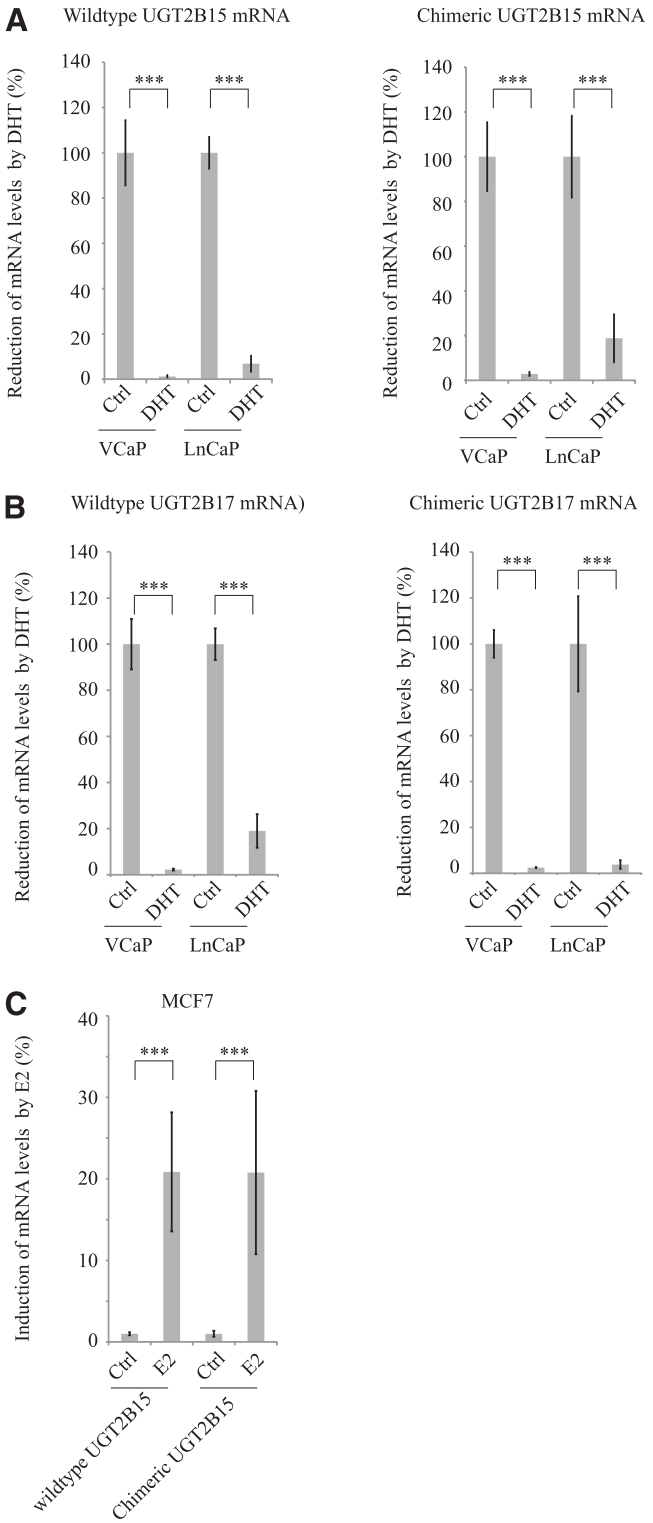


Fig. 8. Regulation of wild-type and chimeric UGT2B15 and UGT2B17 transcripts by sex steroid hormones in prostate and breast cancer cells. (A and B) Prostate cancer VCaP and LNCaP cells were cultured for 2 days in steroid-stripped media and then treated with vehicle control (0.1% ethanol) or 1 nM dihydrotestosterone for 24 hours. (C) MCF7 cells were cultured for 2 days in steroid-stripped media and then treated with vehicle control (0.1% ethanol) or 1 nM 17β-estradiol for 24 hours. Total RNA was prepared and subjected to RT-qPCR to measure UGT2B15 (A) and UGT2B17 (B) transcripts. The expression levels of wild-type and chimeric transcripts in steroid-treated cells were first normalized to that of 18S rRNA and then presented relative to the levels in the vehicle-treated cells (set as a value of 100%). Data shown are means ± S.D. from three independent experiments. Statistical analysis was conducted by two-way analysis of variance followed by Bonferroni's multiple comparison test using GraphPad Prism version 7.03. ***$P < 0.001$.

UGT2B29P2 exons and two *UGT2B29P1* exons may now be considered alternatively spliced exons of *UGT2B15* and *UGT2B17*, respectively. The *UGT2* gene cluster contains at least five other pseudogenes (Turgeon et al., 2000); whether these are involved in formation of chimeric transcripts with neighboring *UGT* genes remains to be investigated.

Previously reported C-terminally truncated UGT variants possess no activity but can interact with and inhibit functional UGTs (Girard et al., 2007; Ménard et al., 2011). Most of the chimeras reported here encoded similar C-terminally truncated UGTs that were able to inhibit the activity of wild-type UGT2B15, UGT2B17, or UGT2B7 proteins, most probably by direct interaction. However, native contexts in which levels of chimeric UGTs are sufficient to inhibit wild-type enzymes require further exploration. Our studies suggest that chimeric UGT2B15 or UGT2B17 variants are probably not sufficiently abundant to inhibit their full-length counterparts or UGT2B7 in normal liver. However, it is worth noting that levels of each of these wild-type UGTs vary extremely in human livers (by over 50-fold) (Izukawa et al., 2009), and we only assessed a relatively small number of liver samples. We did find that in testis, ratios of chimeric UGT2B15 to wild-type UGT2B7 were over 60%, which is well within the range of inhibitory activity. Moreover, our preliminary analysis indicates that ratios of chimeric to wild-type UGT2B15 transcripts ranges very widely (from <1% to >80%) in breast cancer specimens (data not shown). However, the extent to which cancers may display elevated ratios of chimeric UGT2B15 or UGT2B17 to their wild-type counterparts (or other UGTs) is yet to be comprehensively assessed. Splicing dysregulation is frequently seen in cancer, including prostate and breast cancer, and there is considerable interest in determining whether alternatively spliced gene products contribute to cancer progression and/or may be useful as cancer biomarkers. As an example, the alternatively spliced variant of the androgen receptor called ARv7 plays an important role in androgen-independence in prostate cancer, and may be both a prognostic and predictive biomarker (Bastos and Antonarakis, 2018). Given the important role of UGT2B15 and UGT2B17 in steroid (mainly androgen) metabolism in prostate and breast, future work should examine whether the expression of chimeric (or other variant) UGT2B15 and UGT2B17 forms correlate with features such as cancer grade, progression, and outcomes.

Both wild-type and chimeric UGT2B15 and UGT2B17 transcripts were presumed to be expressed from the canonical promoters, and, wild-type and chimeric transcripts were indeed regulated by androgens and estrogens in an equivalent manner. However the presence of different 3′-terminal exons in wild-type and chimeric UGT2B15 and UGT2B17 transcripts suggested differential regulation by miRNAs (An et al., 2013). Indeed, the chimeric UGT2B15 and UGT2B17 transcripts were not regulated by miR-376c, a miRNA that binds to a conserved site in the 3′-UTRs of the wild-type UGT2B15 and UGT2B17 transcripts and mediates their downregulation (Wijayakumara et al., 2015). In contexts in which miRNA-mediated regulation is particularly important

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

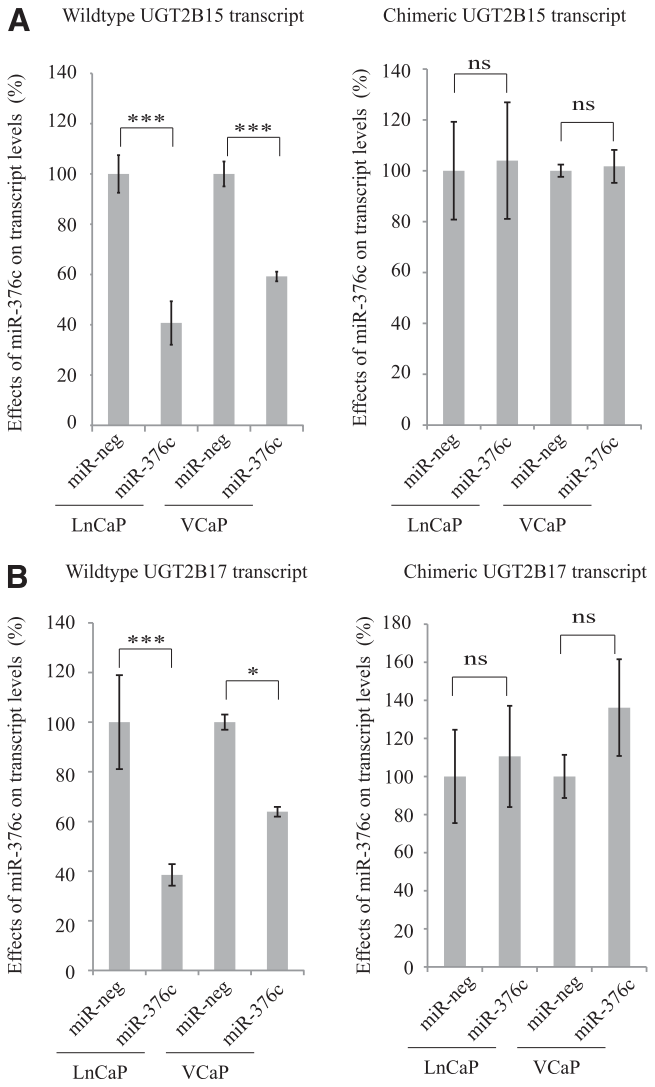


Fig. 9. miR-376c reduces the levels of wild-type UGT2B15 and UGT2B17 transcripts but not the levels of chimeric UGT2B15 and UGT2B17 transcripts in prostate cancer cells. VCaP and LNCaP cells were cultured for 32 hours in steroid-stripped media and then transfected in triplicate with either miR-376c or miR-neg at 30 nM per well. Twenty-four hours post-transfection, total RNA was extracted and subjected to RT-qPCR to measure the levels of wild-type and chimeric UGT2B15 (A) and UGT2B17 (B) transcripts. The expression levels of wild-type and chimeric transcripts in miR-376c-transfected cells were first normalized to that of 18S rRNA and then presented relative to the levels in the miR-neg-transfected cells (set as a value of 100%). Data shown are means ± S.D. from three independent experiments. Statistical analysis was conducted by two-way analysis of variance followed by Bonferroni's multiple comparison test using GraphPad Prism version 7.03. $^{*}P < 0.05$; $^{***}P < 0.001$. ns, not significant.

for maintaining optimal levels of wild-type UGT2B15 and UGT2B17, the insensitivity of the chimeric transcripts to these miRNAs could result in significant increases in the relative ratio of chimeric to wild-type UGT expression.

In conclusion, this study reports the discovery and functional characterization of chimeric transcripts comprising exons from two or three of the four neighboring *UGT* genes *2B15*, *2B29P2*, *2B17*, and *2B29P1*. Our results indicate that the formation of chimeric transcripts among adjacent *UGT* genes (including pseudogenes) via cis-SAGe represents a novel mechanism that further increases the complexity of the human UGT transcriptome and proteome.

Authorship Contributions

Participated in research design: Hu, Hulin, McKinnon, Mackenzie, Meech.

Conducted experiments: Hu, Hulin, Wijayakumara.

Performed data analysis: Hu, Hulin, McKinnon, Mackenzie, Meech.

Wrote or contributed to the writing of the manuscript: Hu, Hulin, McKinnon, Mackenzie, Meech.

References

Al-Balool HH, Weber D, Liu Y, Wade M, Guleria K, Nam PL, Clayton J, Rowe W, Coxhead J, Irving J, et al. (2011) Post-transcriptional exon shuffling events in humans can be evolutionarily conserved and abundant. *Genome Res* **21**:1788–1799.

An J, Zhu X, Wang H, and Jin X (2013) A dynamic interplay between alternative polyadenylation and microRNA regulation: implications for cancer (Review). *Int J Oncol* **43**:995–1001.

Bao BY, Chuang BF, Wang Q, Sartor O, Balk SP, Brown M, Kantoff PW, and Lee GS (2008) Androgen receptor mediates the expression of UDP-glucuronosyltransferase 2 B15 and B17 genes. *Prostate* **68**:839–848.

Bastos DA and Antonarakis ES (2018) CTC-derived AR-V7 detection as a prognostic and predictive biomarker in advanced prostate cancer. *Expert Rev Mol Diagn* **18**: 155–163.

Bellemare J, Rouleau M, Girard H, Harvey M, and Guillemette C (2010) Alternatively spliced products of the UGT1A gene interact with the enzymatically active proteins to inhibit glucuronosyltransferase activity in vitro. *Drug Metab Dispos* **38**: 1785–1789.

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.

Breathnach R, Benoist C, O'Hare K, Gannon F, and Chambon P (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *Proc Natl Acad Sci USA* **75**:4853–4857.

Burset M, Seledtsov IA, and Solovyev VV (2000) Analysis of canonical and non-canonical splice sites in mammalian genomes. *Nucleic Acids Res* **28**:4364–4375.

Bushey RT and Lazarus P (2012) Identification and functional characterization of a novel UDP-glucuronosyltransferase 2A1 splice variant: potential importance in tobacco-related cancer susceptibility. *J Pharmacol Exp Ther* **343**:712–724.

Chouinard S, Barbier O, and Bélanger A (2007) UDP-glucuronosyltransferase 2B15 (UGT2B15) and UGT2B17 enzymes are major determinants of the androgen response in prostate cancer LNCaP cells. *J Biol Chem* **282**:33466–33474.

Chwalenia K, Facemire L, and Li H (2017) Chimeric RNAs in cancer and normal physiology. *Wiley Interdiscip Rev RNA* **8**.

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.

Finta C and Zaphiropoulos PG (2000a) The human CYP2C locus: a prototype for intergenic and exon repetition splicing events. *Genomics* **63**:433–438.

Finta C and Zaphiropoulos PG (2000b) The human cytochrome P450 3A locus. Gene evolution by capture of downstream exons. *Gene* **260**:13–23.

Girard H, Lévesque 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.

Harrington WR, Sengupta S, and Katzenellenbogen BS (2006) Estrogen regulation of the glucuronidation enzyme UGT2B15 in estrogen receptor-positive breast cancer cells. *Endocrinology* **147**:3843–3850.

Hu DG, Gardner-Stephen D, Severi G, Gregory PA, Treloar J, Giles GG, English DR, Hopper JL, Tilley WD, and Mackenzie PI (2010) A novel polymorphism in a forkhead box A1 (FOXA1) binding site of the human UDP glucuronosyltransferase 2B17 gene modulates promoter activity and is associated with altered levels of circulating androstane-3α,17β-diol glucuronide. *Mol Pharmacol* **78**:714–722.

Hu DG, Hickey TE, Irvine C, Wijayakumara DD, Lu L, Tilley WD, Selth LA, and Mackenzie PI (2014a) Identification of androgen receptor splice variant transcripts in breast cancer cell lines and human tissues. *Horm Cancer* **5**:61–71.

Hu DG and Mackenzie PI (2009) Estrogen receptor alpha, fos-related antigen-2, and c-Jun coordinately regulate human UDP glucuronosyltransferase 2B15 and 2B17 expression in response to 17beta-estradiol in MCF-7 cells. *Mol Pharmacol* **76**:425–439.

Hu DG, Mackenzie PI, Lu L, Meech R, and McKinnon RA (2015) Induction of human UDP-Glucuronosyltransferase 2B7 gene expression by cytotoxic anticancer drugs in liver cancer HepG2 cells. *Drug Metab Dispos* **43**:660–668.

Hu DG, Meech R, Lu L, McKinnon RA, and Mackenzie PI (2014b) Polymorphisms and haplotypes of the UDP-glucuronosyltransferase 2B7 gene promoter. *Drug Metab Dispos* **42**:854–862.

Hu DG, Meech R, McKinnon RA, and Mackenzie PI (2014c) Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab Rev* **46**:421–458.

Hu DG, Rogers A, and Mackenzie PI (2014d) Epirubicin upregulates UDP glucuronosyltransferase 2B7 expression in liver cancer cells via the p53 pathway. *Mol Pharmacol* **85**:887–897.

Hu DG, Selth LA, Tarulli GA, Meech R, Wijayakumara D, Chanawong A, Russell R, Caldas C, Robinson JL, Carroll JS, et al. (2016) Androgen and estrogen receptors in breast cancer coregulate human UDP-glucuronosyltransferases 2B15 and 2B17. *Cancer Res* **76**:5881–5893.

Izukawa T, Nakajima M, Fujiwara R, Yamanaka H, Fukami T, Takamiya M, Aoki Y, Ikushiro S, Sakaki T, and Yokoi T (2009) Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers. *Drug Metab Dispos* **37**:1759–1768.

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

Jia Y, Xie Z, and Li H (2016) Intergenically spliced chimeric RNAs in cancer. *Trends Cancer* **2**:475–484.

Jin C, Miners JO, Lillywhite KJ, and Mackenzie PI (1993) Complementary deoxyribonucleic acid cloning and expression of a human liver uridine diphosphate-glucuronosyltransferase glucuronidating carboxylic acid-containing drugs. *J Pharmacol Exp Ther* **264**:475–479.

Kannan K, Wang L, Wang J, Ittmann MM, Li W, and Yen L (2011) Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc Natl Acad Sci USA* **108**:9172–9177.

Klenova E, Chernukhin I, Inoue T, Shamsuddin S, and Norton J (2002) Immunoprecipitation techniques for the analysis of transcription factor complexes. *Methods* **26**:254–259.

Lei Q, Li C, Zuo Z, Huang C, Cheng H, and Zhou R (2016) Evolutionary insights into RNA trans-splicing in vertebrates. *Genome Biol Evol* **8**:562–577.

Lévesque E, Girard H, Journault K, Lépine 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.

Lévesque E, Ménard V, Laverdière I, Bellemare J, Barbier O, Girard H, and Guillemette C (2010) Extensive splicing of transcripts encoding the bile acid-conjugating enzyme UGT2B4 modulates glucuronidation. *Pharmacogenet Genomics* **20**:195–210.

Lewis BC, Mackenzie PI, and Miners JO (2011) Homodimerization of UDP-glucuronosyltransferase 2B7 (UGT2B7) and identification of a putative dimerization domain by protein homology modeling. *Biochem Pharmacol* **82**: 2016–2023.

Mackenzie PI (1986) Rat liver UDP-glucuronosyltransferase. Sequence and expression of a cDNA encoding a phenobarbital-inducible form. *J Biol Chem* **261**: 6119–6125.

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.

Margaillan G, Lévesque É, and Guillemette C (2016) Epigenetic regulation of steroid inactivating UDP-glucuronosyltransferases by microRNAs in prostate cancer. *J Steroid Biochem Mol Biol* **155** (Pt A):85–93.

McKinnon RA, Hall PD, Quattrochi LC, Tukey RH, and McManus ME (1991) Localization of CYP1A1 and CYP1A2 messenger RNA in normal human liver and in hepatocellular carcinoma by in situ hybridization. *Hepatology* **14**:848–856.

Ménard 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.

Ménard V, Eap O, Harvey M, Guillemette C, and Lévesque E (2009) 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.

Ménard V, Eap O, Roberge J, Harvey M, Lévesque 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.

Mount SM (1982) A catalogue of splice junction sequences. *Nucleic Acids Res* **10**: 459–472.

Parra G, Reymond A, Dabbouseh N, Dermitzakis ET, Castelo R, Thomson TM, Antonarakis SE, and Guigó R (2006) Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res* **16**:37–44.

Pei B, Sisu C, Frankish A, Howald C, Habegger L, Mu XJ, Harte R, Balasubramanian S, Tanzer A, Diekhans M, et al. (2012) The GENCODE pseudogene resource. *Genome Biol* **13**:R51.

Radominska-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.

Stone AN, Mackenzie PI, Galetin A, Houston JB, and Miners JO (2003) Isoform selectivity and kinetics of morphine 3- and 6-glucuronidation by human udp-glucuronosyltransferases: evidence for atypical glucuronidation kinetics by UGT2B7. *Drug Metab Dispos* **31**:1086–1089.

Tourancheau A, Margaillan G, Rouleau M, Gilbert I, Villeneuve L, Lévesque 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.

Tourancheau A, Rouleau M, Guauque-Olarte S, Villeneuve L, Gilbert I, Droit A, and Guillemette C (2018) Quantitative profiling of the UGT transcriptome in human drug-metabolizing tissues. *Pharmacogenomics J* **18**:251–261.PubMed

Turgeon D, Carrier JS, Lévesque E, Beatty BG, Bélanger A, and Hum DW (2000) Isolation and characterization of the human UGT2B15 gene, localized within a cluster of UGT2B genes and pseudogenes on chromosome 4. *J Mol Biol* **295**:489–504.

Warner SC, Finta C, and Zaphiropoulos PG (2001) Intergenic transcripts containing a novel human cytochrome P450 2C exon 1 spliced to sequences from the CYP2C9 gene. *Mol Biol Evol* **18**:1841–1848.

Wijayakumara DD, Hu DG, Meech R, McKinnon RA, and Mackenzie PI (2015) Regulation of human UGT2B15 and UGT2B17 by miR-376c in prostate cancer cell lines. *J Pharmacol Exp Ther* **354**:417–425.

Wijayakumara DD, Mackenzie PI, McKinnon RA, Hu DG, and Meech R (2017) Regulation of UDP-glucuronosyltransferases UGT2B4 and UGT2B7 by MicroRNAs in liver cancer cells. *J Pharmacol Exp Ther* **361**:386–397.

Zaphiropoulos PG (1998) Non-homologous recombination mediated by Thermus aquaticus DNA polymerase I. Evidence supporting a copy choice mechanism. *Nucleic Acids Res* **26**:2843–2848.

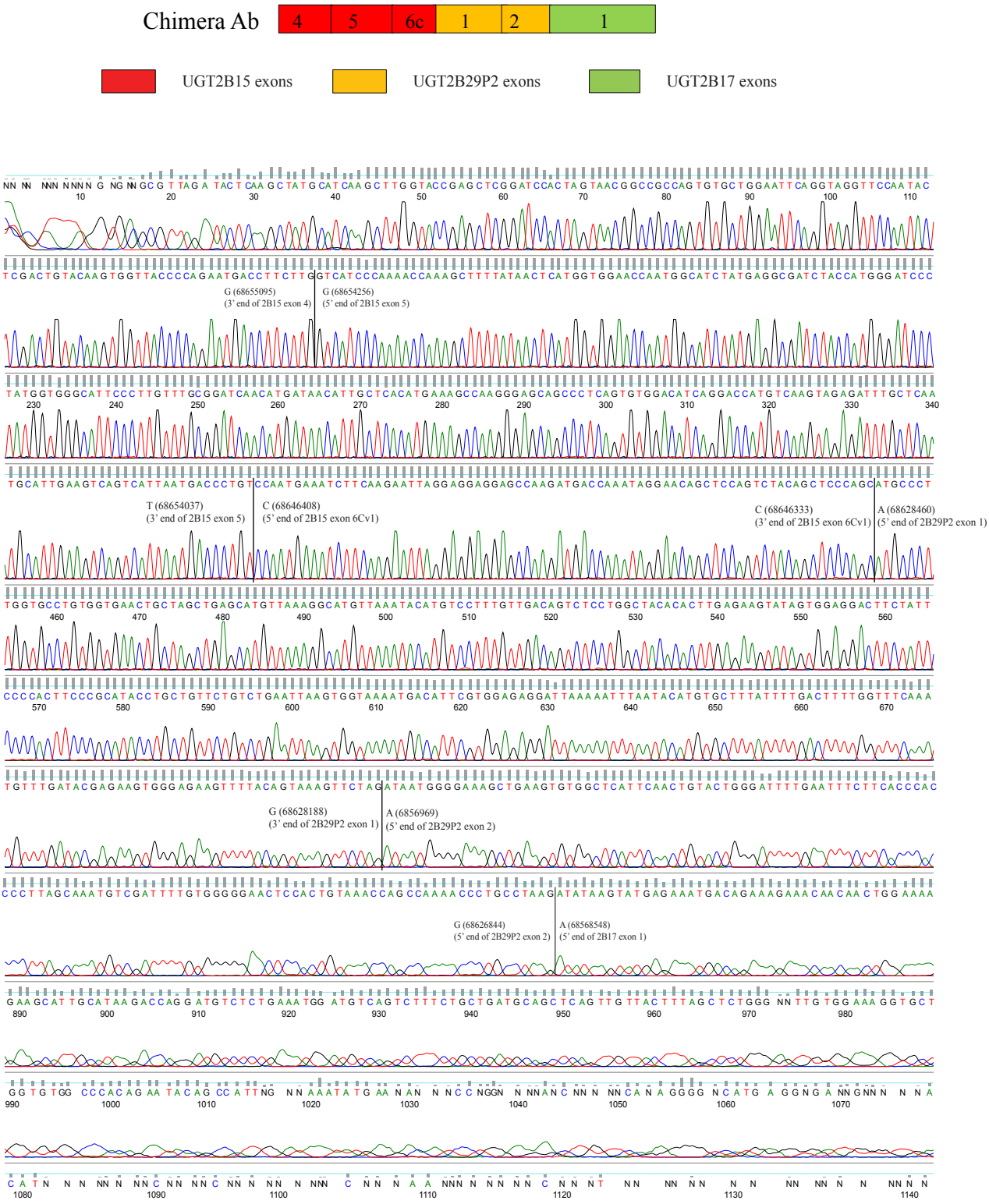
Zaphiropoulos PG (1999) RNA molecules containing exons originating from different members of the cytochrome P450 2C gene subfamily (CYP2C) in human epidermis and liver. *Nucleic Acids Res* **27**:2585–2590.

Address correspondence to: Dr. Robyn Meech, Department of Clinical Pharmacology and Flinders Centre for Innovation in Cancer, Flinders University College of Medicine and Public Health, Flinders Medical Centre, Bedford Park, South Australia, Australia. E-mail: robyn.meech@flinders.edu.au

Downloaded from molpharm.aspetjournals.org at ASPET Journals on April 23, 2024

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

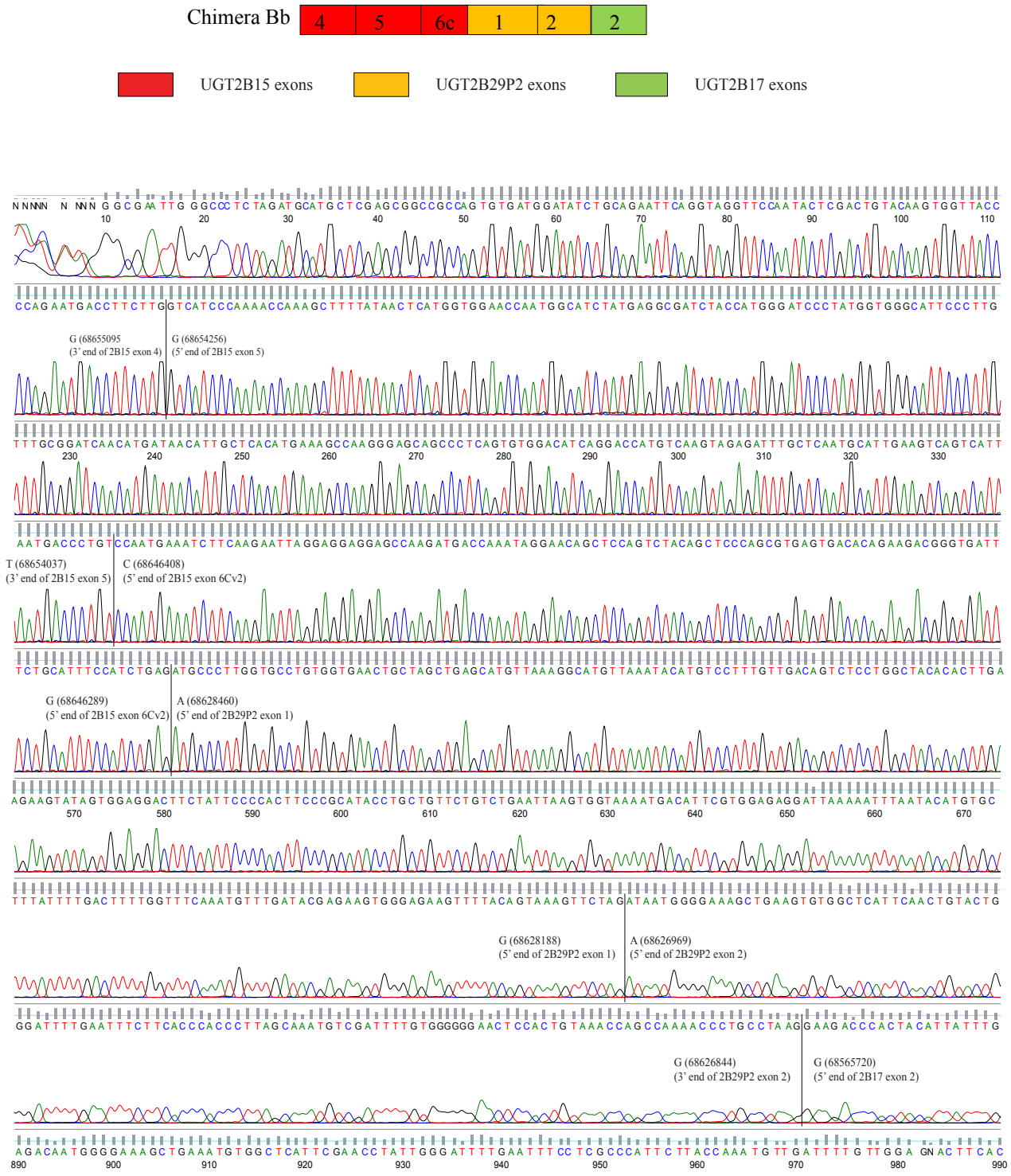
Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 1

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 2

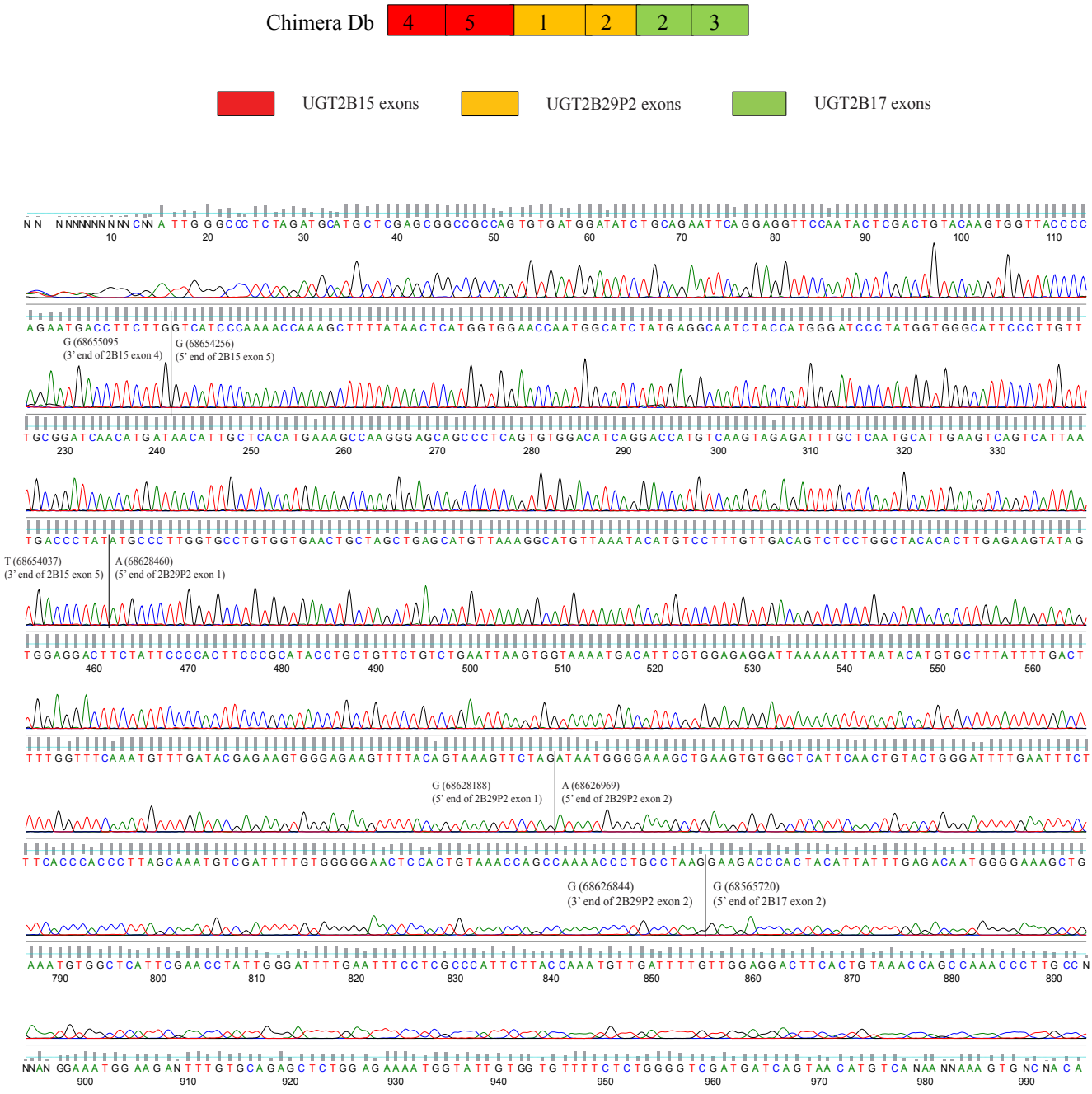
Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

Supplemental Figure 3

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

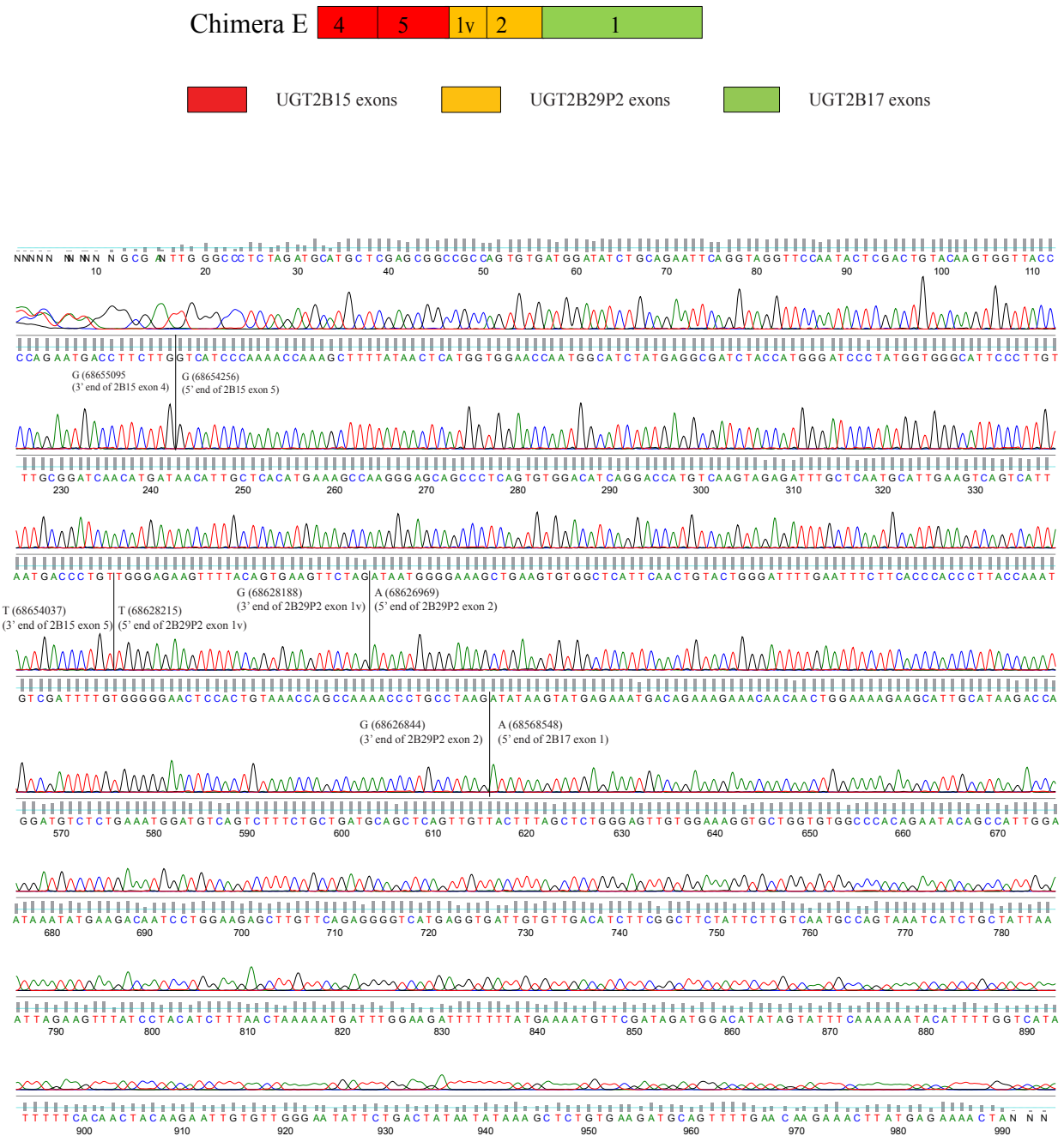
Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 4

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

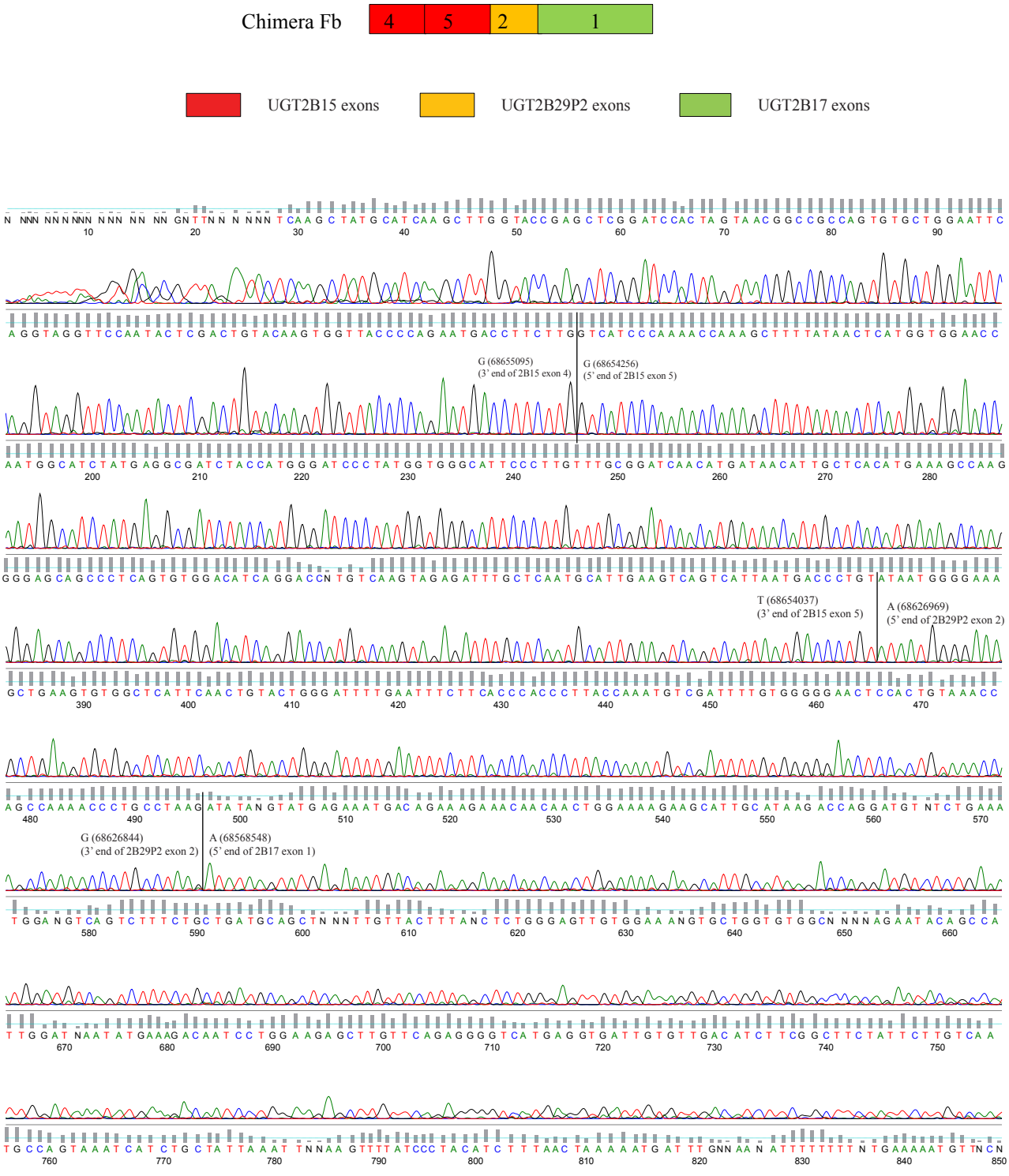
Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 5

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 6

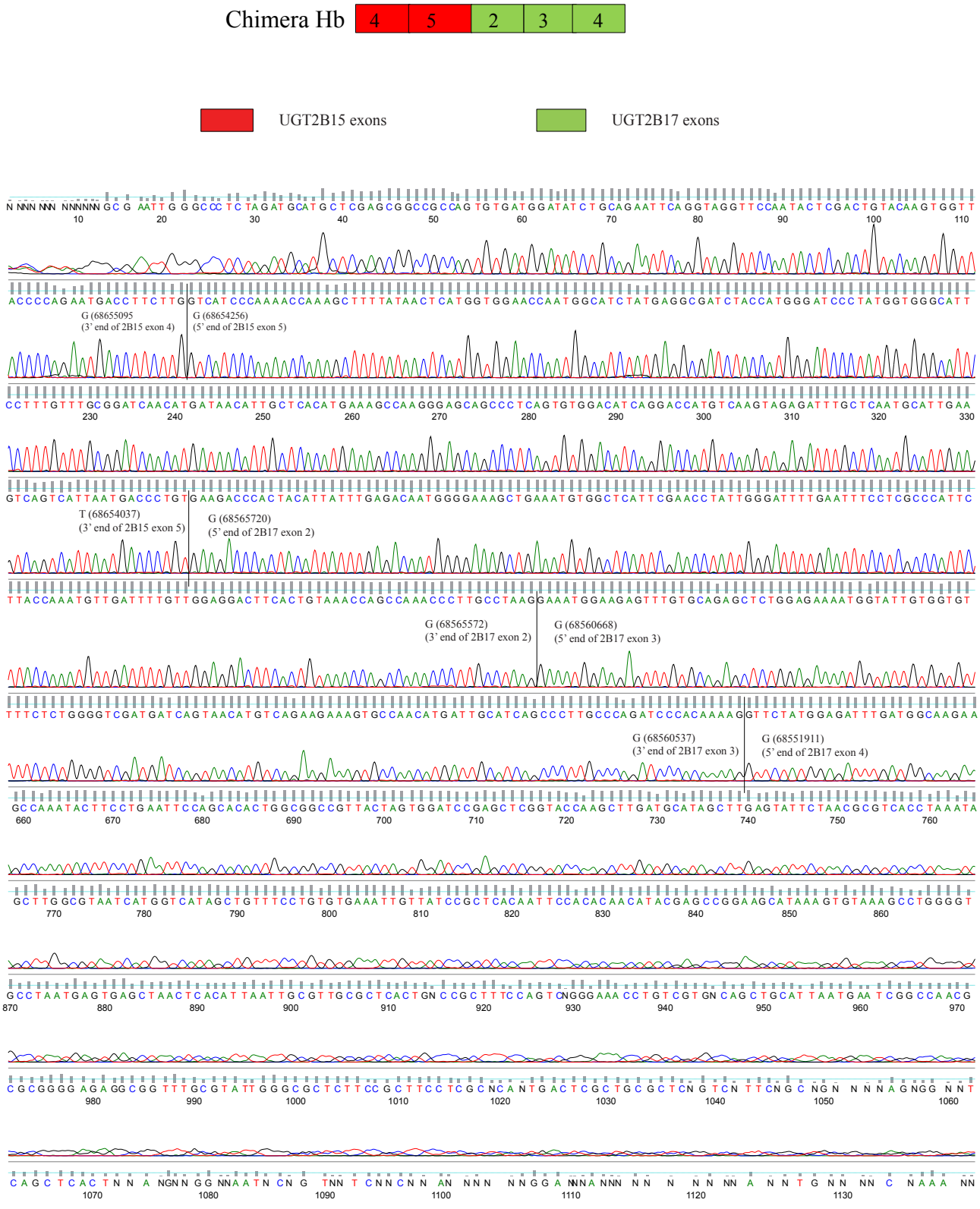
Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

Supplemental Figure 7

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

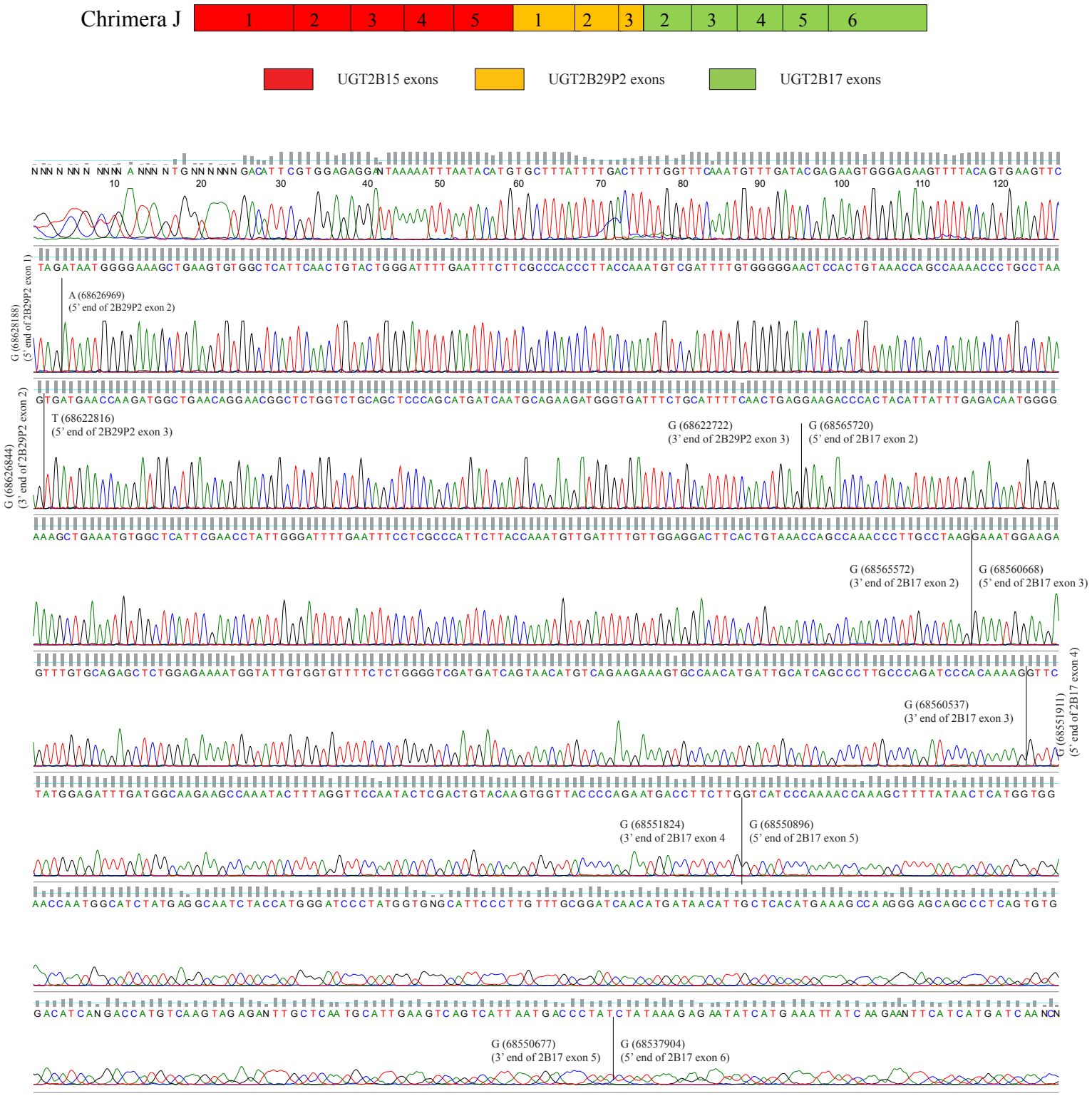
Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 8

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 9

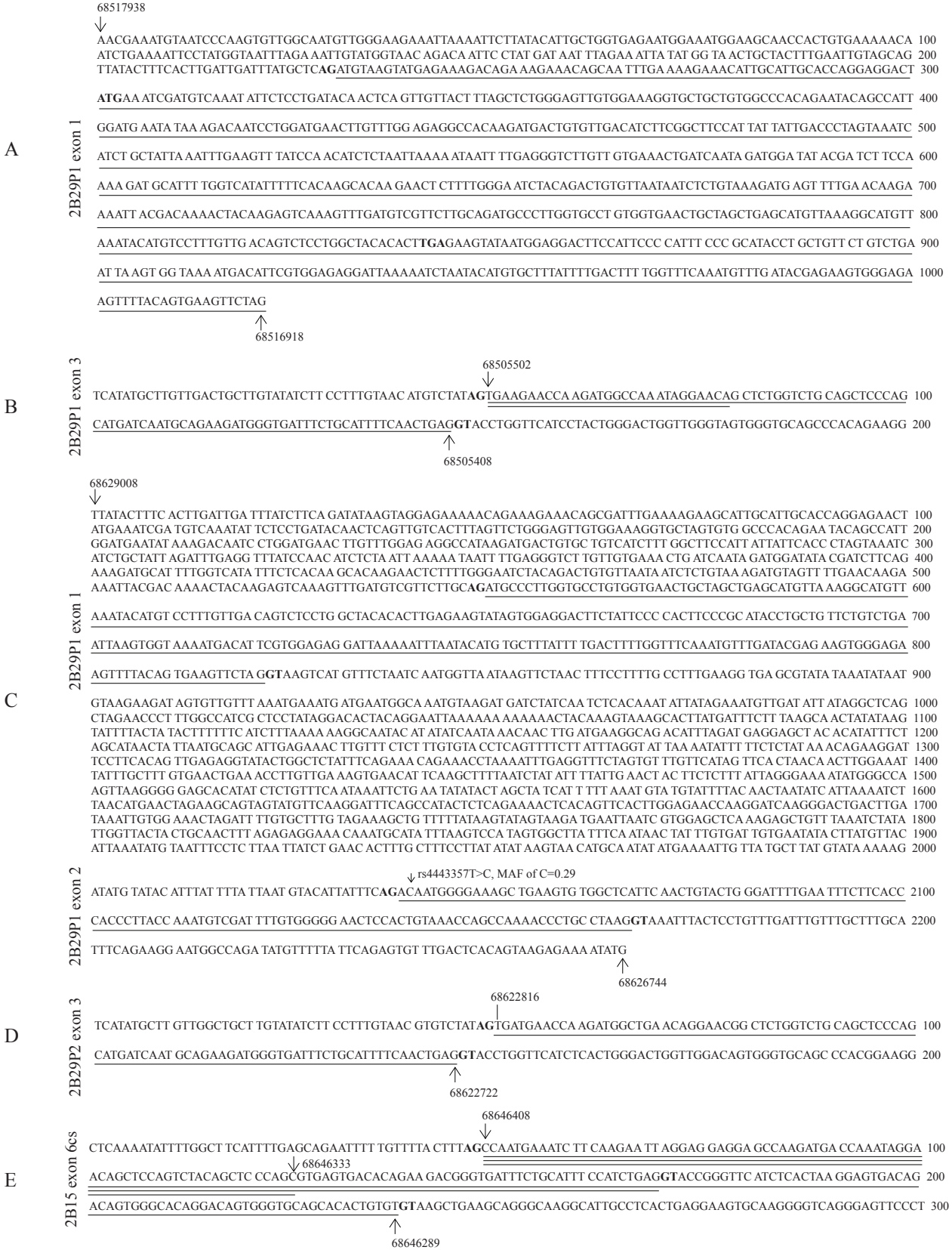
Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

Supplemental Figure 10

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 11

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

A

Sequence alignment of exon 1s of *UGT2B15*, *UGT2B17*, *UGT2B29P1*, *UGT2B29P2*

```
UGT2B29P2
UGT2B29P1 ATGTAAGTATGAG AAAGACAGA AAGAAACAGCAATTTGAA AAG AAACATTGCATTGCACCAGGA GGACATTGA AATCGATGTCAAATAT TCT CCTGATAC 100
UGT2B15                                    AAACAACAACTGGAA AAG AAGCATTGCATAAGACCAGGATGTCTCTGA AATGGACGTCAGTCTT TCTGCTGATAC  75
UGT2B17   ATATAAGTATGAGAAATGACAGAAAGAAACAACAACTGGAA AAG AAGCACTGCATAAGACCAGGATGTCTCTGA AATGGATGTCAGTCTT TCTGCTGATGC 80
                                                 ↑ start of UGT2B17 exon 1 (NM_001076)
UGT2B29P2
UGT2B29P1 AACTCAGT TGTTACTT TAGCTCT GGGAGTTGTG GAAAGGTG CTGCTGTGGC CCACAGA ATACAG CCATTG GATGAATATAAAGA CAATCC TGGATGAACT 200
UGT2B15   AGCTCAGT TGTTACTT TAGCTCT GGAAGCTGTG GAAAGGTG CTAGTGTGGC CCACAGA ATACAG CCATTG GATAAATATGAAGA CAATCCTGGAAGAGCT 175
UGT2B17   AGCTCAGT TGTTACTT TAGCTCT GGGAGTTGTG GAAAGGTG CTGGTGTGGC CCACAGA ATACAG CCATTG GATAAATATGAAGA CAATCCTGGAAGAGCT 180

UGT2B29P2
UGT2B29P1 TGTTTGGAGA GGC CACAAGA TGACTGTGTT GACAT CTTCG GCTTCCATTA TTATTGACCCTAG TA AATCATCT GCT ATTA AATTTGAAGT TTATCCAACA 300
UGT2B15   TGTTCAGAGG GGC CATGAGGTGACTGTGTT GACAT CTTCG GCTTCTACTC TTGTCAATGCCAG TA AATCATCT GCT ATTA AATTAGAAGT TTATCCTACA 275
UGT2B17   TGTTCAGAGG GGC CATGAGGTGATTGTGTT GACAT CTT CG GCTTCTATTC TTGTCAATGCCAG TA AATCATCT GCT ATTA AATTAGAAGT TTATCCTACA 280

UGT2B29P2
UGT2B29P1 TCTCTAATTA AAA ATAATTT TGAGGGTCTTGTTGTGA AACTGATCAATAG ATGGATA TACGATCT TCCAA A AGATGCATTT TGGTCATAT T TT TCAC AAG 400
UGT2B15   TCTTTAACTA AAA ATTATTT GGAAGATTCTCTTCTGA AA ATTCTCAATAG ATGGATA TATGGTGT TTCAA A  AAATACATTT TGGTCATAT T TT TCAC AAT 375
UGT2B17   TCTTTAACTA AAA ATGATTTGGAAGATTTT TTTATGA AA ATGTTCAATAG ATGGACA TATAGTAT TTCAA A AAATACATTT TGGTCATAT T TT TCAC AAC 380

UGT2B29P2
UGT2B29P1 CACAAGAA CTCTTTTGGGA ATCTACAGACTGTGTTAATA A TCTCTGTAAAGATGTAGTTTT GAACAAGAAAATTACGACAAAACTACAAG AGTC AAAGTT 500
UGT2B15   TACAAGAA TTGTGTTGGGA ATATTATGACTACAGTAACA AGCTCTGTAAAGATGCAGTTTTGAATAAGAAACTT ATGATGAAACTACAAG AGTC AAAGTT 475
UGT2B17   TACAAGAA TTGTGTTGGGA ATATTCTGACTATA ATA TAA AGCTCTGTGAAGATGCAGTTTTGAACAAGAAACTTATGAGAAAACTACAAG AGTC AAA ATT 480

UGT2B29P2                                  ATGCCCTTGGTGCCTGTGGTGAACTGCTAGCTG AGCATGTTAA - AGGCATGTTAAATACA T GTCCTTTGTTGACAGTCTCCTGG  84
UGT2B29P1 TGATGTCGT TCTTGCAGATGCCCTTGGTGCCTGTGGTGAACTGCTAGCTG AGCATGTTAA - AGGCATGTTAAATACA T GTCCTTTGTTGACAGTCTCCTGG 600
UGT2B15   TGATGTCATTCTGGCAGATGCCCTTAATCCCTGTGGTGAGCTACTGGCTG A ACTATT TAA CATACCCTT TCTG TACA - GTCTTCGATTC -- -- TCTGTTGG 575
UGT2B17   TGATGTCCTTCTGGCAGATGCCGTTAATCCCTGTGGTGAGCTGCTGGCTGA GCTACT TAA CATACCCTT TCTG TACA - GTCTCCGCTTC -- -- TCTGTTGG 580

UGT2B29P2 CTACACACTTGAGA AGTATAGTGG AGGACTTCTA TTCCCCACT TCCCGCATACCTGCTGT TCTGTCTGAATTAAGTGGTA AAATGACAT TCG TGGAGAGG 184
UGT2B29P1 CTACACACTTGAGA AGTATAATGG AGG ACTTCCATTCCCCATT TCCCGCATACCTGCTGT TCTGTCTGAATTAAGTGGTA AAATGACAT TCG TGGAGAGG 700
UGT2B15   CTACACATTTGAGA AGAATGGTGG AGGATTTCTGTTCCCTCCT TCCTATG TACCTG TTGT TATGTCAGAATTAAGTGATC AAATGATTT TCA TGGAGAGG 675
UGT2B17   CTACACAGTTGAGA AGAATGGTGG AGGATTTCTGTTCCCTCCT TCCTATG TACCTGTTGT TATGTCAGAATTAAGTG ATC AAATGATTT TCA TGGAGAGG 680

UGT2B29P2 ATTAAAAATTTAATACATGTGCTTTATTTT GACTTTTGGT TTCAAATGTTTGATACGA - GA AGTGGGAGAAGTTTTACAGTGAAGTTCTAG  273
UGT2B29P1 ATTAAAAATCTAATACATGTGCTTTATTTT GACTTTTGGT TTCAAATGTTTGATACGA - GA AGTGGGAGAAGTTTTACAGTGAAGTTCTAG  790
UGT2B15   ATAAAAAATATGATACATATGCTTTATTTT GACTTTTGGT TTCAAATTTATGATCTGAAGA AGTGGGACCAGTTTTA TAGTGAAGTTCTAG  762
UGT2B17   ATAAAAAATATGATATATATGCTTTATTTT GACTTTTGGT TTCAAGCATATGATCTGAAGAAGTGGGACCAGTTTTA TAGTGAAGTTCTAG  767
```

B

Sequence alignment of exon 2s of *UGT2B15*, *UGT2B17*, *UGT2B29P1*, *UGT2B29P2*

```
UGT2B15   GAAGACCCACTACATTATTTGAGACAATGGGGA AAGCTGA AA TGTGGCTCATTCGAACCT ATTGGGAT TT TGAATT TCCT CGC CCAT TCTTACCAAATGT 100
UGT2B17   GAAGACCCACTACATTATTTGAGACAATGGGGA AAGCTGA AA TGTGGCTCATTCGAACCT ATTGGGAT TT TGAATT TCCT CGC CCAT TCTTACCAAATGT 100
UGT2B29P1                    ACAATGGGGA AAGCTGA AG TGTGGCTCATTCAAACGT ACTGGGAT TT TGAATT TCTT CAC CCACCCT TACCAAATGT  83
UGT2B29P2                    ATAATGGGGA AAGCTGA AG TGTGGCTCATTCAACTGT ACTGGGAT TT TGAATT TCTT CAC CCACCCT TACCAAATGT  77

UGT2B15   TGATTTTGT T GGAGGACTTC ACTGTAAACC AGCCAAACCC CTGCCTAAG  149
UGT2B17   TGATTTTGT T GGAGGACTTC ACTGTAAACC AGCCAAACCC TTGCCTAAG  149
UGT2B29P1 CGATTTTGTG GGGGAACTCC ACTGTAAACC AGCCAAAACC CTGCCTAAG  126
UGT2B29P2 CGATTTTGTGGGGG AACTCC ACTGTAAACC AGCCAAAACC CTGCCTAAG  126
```

C

Sequence alignment of exon 3s of *UGT2B15*, *UGT2B17*, *UGT2B29P1*, *UGT2B29P2*

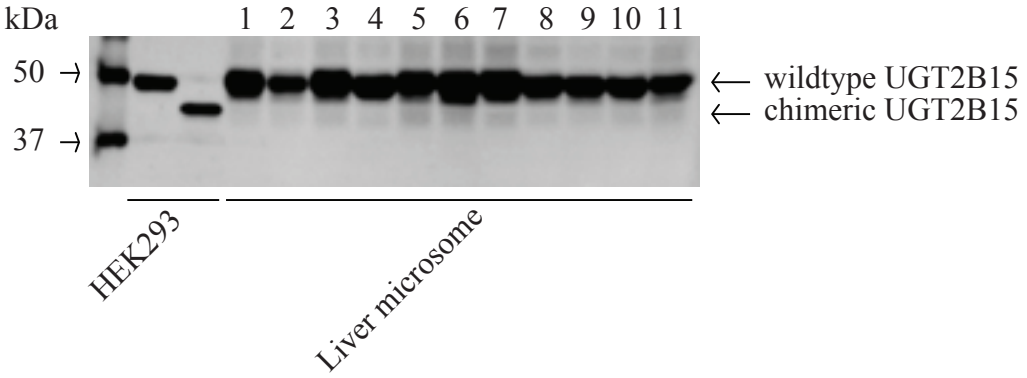
```
UGT2B15   GAAATGGAAGAGTTTGTGCAGAGCTCTGGAGAAAATGGTA TTGTGGTGTTTTCTCTGGGGTCGATG ATCAGTAACATGTCAGAAGAAAGTGCCAACATGA 100
UGT2B17   GAAATGGAAGAGTTTGTGCAGAGCTCTGGAGAAAATGGTA TTGTGGTGTTTTCTCTGGGGTCGATG ATCAGTAACATGTCAGAAGAAAGTGCCAACATGA 100
UGT2B29P1                  TGAAGAACCAAGATGGCCAAATAGGAACAGCTCTGGT - - - - CTGCAG - CTCCCAGCATGA TCAATGCAGAAGATGGGTG - - - - ATTT  78
UGT2B29P2                  TGATGAACCAAGATGGCTGAACAGGAACGGCTCTGGT - - - - CTGCAG - CTCCCAGCATGA TCAATGCAGAAGATGGGTG - - - - ATTT  78

UGT2B15   TTGCATCAGCCCTTGCCCAG ATCCCACAAA AG  132
UGT2B17   TTGCATCAGCCCTTGCCCAG ATCCCACAAA AG  132
UGT2B29P1 CTGCATTTTCAACTGAG  95
UGT2B29P2 CTGCATTTTCAACTGAG  95
```

Supplemental Figure 12

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

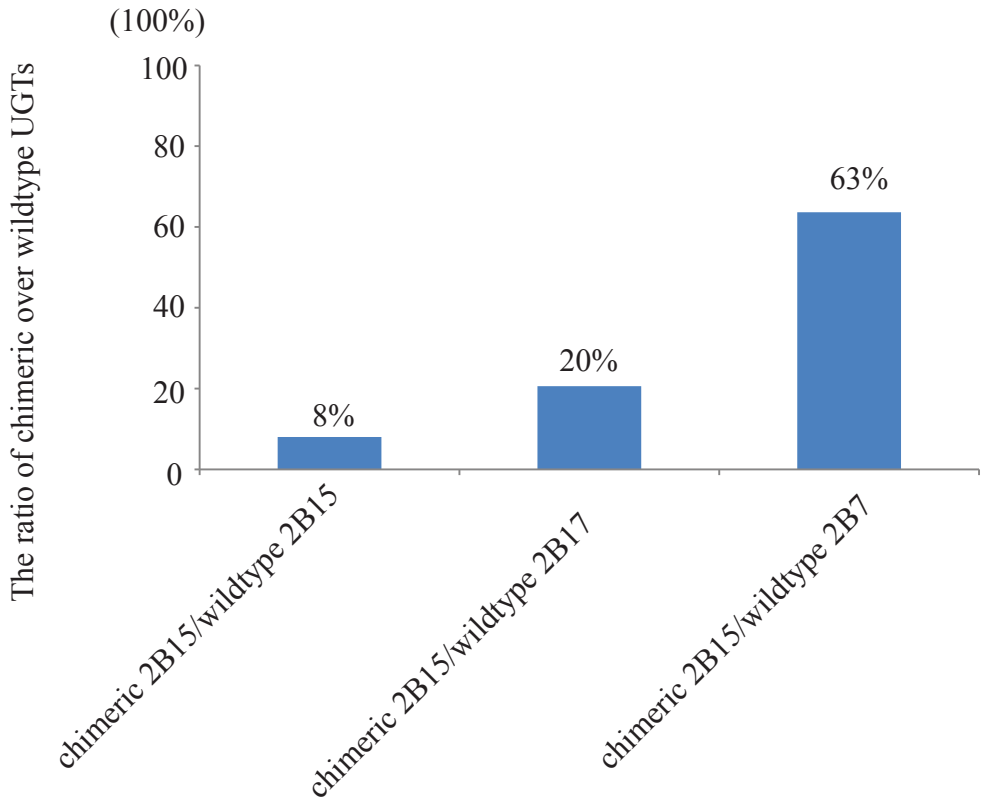
Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 13

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech



Supplemental Figure 14

Molecular Pharmacology
Supplemental Materials

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

Supplemental Figure Legends

Supplemental Figs 1-10. The sequencing chromatograms of RT-PCR products identified novel chimeric transcripts of four neighbouring *UGT* genes (*2B15, 2B29P2, 2B17,* and *2B29P1*). RT-PCR was carried out to amplify chimeric transcripts using cDNA from normal liver tissues or prostate cancer VCaP and LNCaP cells. Shown are the sequencing chromatograms that contain the chimeric parts of ten chimeric transcripts. The exon/exon junctions are indicated by a vertical line and the starting and ending nucleotides of all exons are positioned according to the human genome assembly GRC38/hg38 (2013).

Supplemental Fig. 11. The sequences of novel exons from pseudogenes *UGT2B29P1 and UGT2B29P1*, and variant exons from functional genes *UGT2B15* and *UGT2B17*. (A) The *UGT2B29P1* exon 1 (underlined) in the chimeras is contained within the predicted *UGT2B29P1* gene sequence (AF179880). (B) The predicted *UGT2B29P1* exon 3 (single-underlined) and the *UGT2B29P1* exon 3 (double-underlined) detected in the chimeras. (C) The *UGT2B29P2* exon 1 and exon 2 (underlined) seen in the chimeras are contained within the predicted *UGT2B29P2* gene sequence (NG_022012.1). (D) The *UGT2B29P2* exon 3 (underlined) seen in the chimeras. (E) The reported *UGT2B15* exon 6C (187 bp, single-underlined) and the variant exon 6Cv1 (76 bp, triple-underlined) seen in Chimera A and the variant exon 6Cv2 (120 bp, double-underlined) seen in Chimera B. The genomic coorinates that indicate the start and end of each exon are given according to the human genome assembly GRC38/hg38 (2013). All of the exons are flanked by AG (bold) at the 5'-end and GT (bold) at the 3'-end and thus conform to the "GT-AG" rule.

Molecular Pharmacology
Supplemental Materials

Intergenic splicing between four adjacent UGT genes (2B15, 2B29P2, 2B17, 2B29P1) gives rise to variant UGT proteins that inhibit glucuronidation via protein-protein interactions

Dong Gui Hu, Julie-Ann Hulin, Dhilushi D. Wijayakumara, Ross A. McKinnon, Peter I. Mackenzie, and Robyn Meech

Supplemental Figure 12. Sequence alignment of exons 1-3 of four neighbouring *UGT* genes (*2B15*, *2B29P2*, *2B17*, *2B29P1*). Shown are the sequence alignment of exons 1 (A), exons 2 (B) and exons 3 (C) of *UGT2B15* (NM_001076), *UGT2B17* (NM_001077), *UGT2B29P1*, and *UGT2B29P2*. Five novel exons, including *UGT2B29P1* exons 1 and 3 and *UGT2B29P2* exons 1, 2 and 3, are found in chimeric transcripts in this study. The *UGT2B17* exon 1 in the chimeric transcripts starts 21 nucleotides earlier compared to exon 1 in wildtype *UGT2B17* exon 1 (indicated by an arrow) as annotated in the reference sequence NM_001077.

Supplemental Figure 13. Western Blotting assays show the expression of wild-type and putative chimeric UGT2B15 proteins in human liver. Eleven human liver microsomal samples and wildtype and chimeric UGT2B15 proteins that were expressed in HEK293 cells, were subjected to standard western blotting assays using the rabbit anti-UGT2B15/UGT2B17 antibody that recognizes both wildtype and chimeric UGT2B15 proteins. Molecular weight markers are indicated on the left of the figure.

Supplemental Figure 14. Analysis of chimeric and wildtype UGTs in human testis. RT-qPCR quantification of chimeric UGT2B15 and three wildtype UGTs (2B7, 2B15, 2B17) were conducted using total RNA from human testis as described in Materials and Methods. Data shown are the levels (plotted as the percentage) of the chimeric UGT2B15 transcript relative to the levels of the respective wildtype UGTs.