



Pharmacogénomique de la voie de glucuronidation : mécanismes moléculaires et impact clinique

Thèse

Adrien Labriet

Doctorat en sciences pharmaceutiques
Philosophiæ doctor (Ph. D.)

Québec, Canada

**Pharmacogénomique
de la voie de glucuronidation :
mécanismes moléculaires et impact clinique**

Thèse

Adrien Labriet

Sous la direction de :

Chantal Guillemette, directrice de recherche

Résumé

La voie de glucuronidation catalyse l'inactivation de nombreux métabolites endogènes, tels que la bilirubine ou les hormones stéroïdiennes, ainsi que des substances exogènes incluant notamment des médicaments et des carcinogènes. Ce processus enzymatique opéré par les enzymes UDP-glucuronosyltransférases (UGT) entraîne généralement l'abolition de l'activité biologique ou pharmacologique des composés et en augmente la solubilité afin de favoriser leur élimination de l'organisme. Une importante variabilité dans l'expression et l'activité de la voie de glucuronidation est observée entre les individus, pouvant affecter l'exposition aux substrats de cette voie enzymatique.

Dans le cadre de ma thèse, je me suis intéressé à l'influence des polymorphismes génétiques et de l'épissage alternatif sur la variabilité de la voie de glucuronidation et la réponse au traitement. La relation entre les variations génétiques et la réponse au traitement de chimiothérapie de première intention à base d'irinotécan a d'abord été étudiée chez des patients atteints du cancer colorectal métastatique (CCRm). De nouveaux marqueurs germinaux ont été associés à la survie des patients de deux cohortes indépendantes, notamment dans les gènes *UGT1*, *CES1* (carboxylestérase 1), *ABCC1* (*multidrug resistance-associated protein 1*), *RPL28* (protéine ribosomique L28) et du facteur de transcription *HNF1A* (*hepatocyte nuclear factor 1-alpha*). D'autre part, nos travaux révèlent que les variants d'épissage d'*UGT2B10* représentent plus de la moitié du transcriptome dérivé de ce gène dans le tissu hépatique. Les protéines alternatives codées par ces nouveaux transcrits affectent la capacité de glucuronidation d'agents pharmacologiques pris en charge par l'enzyme. L'exposition des cellules hépatiques à certains composés exogènes semble remodeler l'épissage du gène *UGT2B10* au détriment de l'expression de l'enzyme.

Les travaux présentés dans cette thèse supportent que la génétique du patient ainsi que les processus d'épissage alternatif au niveau tissulaire aient le potentiel d'affecter la capacité de glucuronidation et la réponse au traitement. Puisque ces deux mécanismes contribuent à la variabilité de la voie des UGT, ils pourraient constituer de nouveaux biomarqueurs de réponse aux composés pris en charge par ces enzymes.

Abstract

The UDP-glucuronosyltransferases enzymes (UGTs) are responsible for the conjugation of their co-substrate UDP-glucuronic acid to numerous endogenous and exogenous substrates, including many drugs. This metabolic reaction, called glucuronidation, generally results in substrate inactivation and leads to increased solubility, leading to its elimination through bile and urine. Along with other enzymatic conversions and transport pathways, glucuronidation is a determinant of drug response. However, an important interindividual variability in its expression and activity is observed in patients and may lead to variable bioavailability and response to drugs.

During my thesis, my interests were to investigate the role of germline genetic variations and alternative splicing events as mechanisms involved in the interindividual variability of the UGT pathway. First, we investigated genetic polymorphisms of irinotecan-related pathways including UGTs, in relation to outcomes in 417 metastatic colorectal cancer patients treated with first-line irinotecan-based chemotherapy. We identified several markers of survival, namely polymorphisms in *UGT1*, *CES1* (carboxylesterase 1), *ABCC1* (multidrug resistance-associated protein 1), *RPL28* (ribosomal protein L28) genes as well as in *HNF1A* (hepatocyte nuclear factor 1-alpha) transcription factor gene. Then, we demonstrated that alternative transcripts represent half of the *UGT2B10* hepatic transcriptome. Alternative proteins arising from *UGT2B10* splicing events alter enzymatic activity for UGT2B10 substrates. We also showed that pharmacological compounds may affect splicing events at the *UGT2B10* locus.

These studies support an impact of common genetic variations and post-transcriptional mechanisms on glucuronidation, drug exposure and possibly drug response. Germline variations and alternative splicing seem to significantly contribute to the high interindividual variability of the UGT pathway. They may be potential biomarkers of the response to drugs metabolized by this pathway.

Table des matières

Résumé	ii
Abstract.....	iii
Table des matières	iv
Liste des figures.....	vi
Liste des tableaux	vii
Liste des abréviations	viii
Remerciements	x
Avant-propos	xii
Introduction	1
1) La pharmacogénomique en clinique.....	4
2) Voies PK/PD des traitements.	7
2.1) Pharmacocinétique : absorption, distribution, métabolisme et excrétion	7
2.2) Pharmacodynamie.....	12
2.3) Voies PK/PD de l'irinotécan.....	13
3) La voie de glucuronidation : fonctions, organisation, régulation et variabilité inter-individuelle	18
3.1) Fonctions de la voie de glucuronidation	18
3.2) Organisation de la voie de glucuronidation	20
3.3) Régulation de l'expression de la voie de glucuronidation	25
3.4) Variabilité interindividuelle d'expression et d'activité de la voie de glucuronidation.....	28
3.5) Sources de variabilité interindividuelle de la voie de glucuronidation et réponse au traitement.....	30
4) Hypothèses et objectifs de recherche	39
5) Justification de la méthodologie retenue	40
5.1) Découverte de marqueurs génétiques prédictifs de la réponse au traitement	40
5.2) Étude des mécanismes moléculaires en lien avec les marqueurs liés à l'issue clinique	43
5.3) Étude de la fonction de nouveaux variants d'épissage	44

Chapitre 1 : Combination of germline variations in irinotecan pharmacogenetic pathways associated with survival in metastatic colorectal cancer treated with FOLFIRI-based chemotherapy.....	46
1.1 Résumé.....	46
1.2 Abstract.....	48
Chapitre 2 : Germline variability and tumor expression level of ribosomal protein gene <i>RPL28</i> are associated with survival of metastatic colorectal cancer patients.....	77
2.1 Résumé.....	77
2.2 Abstract.....	78
Chapitre 3 : Improved progression-free survival in irinotecan-treated metastatic colorectal cancer patients carrying the HNF1A coding variant p.I27L	124
3.1 Résumé.....	124
3.2 Abstract.....	125
Chapitre 4 : Post-transcriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing.....	164
4.1 Résumé.....	164
4.2 Abstract.....	165
Discussion.....	213
Conclusion.....	231
Annexe A : Germline polymorphisms associated to survival of mCRC patients treated with irinotecan-containing regimen.....	233
Bibliographie	238

Liste des figures

Figure 1. Mécanismes impliqués dans l'influx et l'efflux des médicaments.	9
Figure 2. Enzymes du métabolisme de phase I (gauche) et II (droite) des médicaments. ...	12
Figure 3. Voies PK de l'irinotécan.	17
Figure 4. La réaction de glucuronidation.....	19
Figure 5. Organisation génique et protéique des UGT.	22
Figure 6. Quantification de l'expression des transcrits canoniques et alternatifs des UGT dans plusieurs tissus du métabolisme des médicaments sains et cancéreux.....	38

Liste des tableaux

Tableau 1. Exemples de marqueurs de nature somatique et héréditaire pour lesquels un effet sur la réponse aux anticancéreux a été démontré.	2
Tableau 2. Exemples de variants d'épissage alternatif pour lesquels un effet sur la réponse aux anticancéreux a été démontré.....	4
Tableau 3. Réponse variable au régime FOLFIRI pour le traitement de 1 ^{ère} ligne du CCRm.	14
Tableau 4. Les UGT catalysent la glucuronidation de multiples substrats endogènes et exogènes incluant des médicaments.	24
Tableau 5. L'expression des UGT peut être induite par divers récepteurs nucléaires activés par des ligands endogènes et exogènes.....	27
Tableau 6. L'expression protéique des UGT au niveau hépatique est très variable entre individus.	29
Tableau 7. Variations génétiques affectant les gènes <i>UGT</i>	32
Tableau 8. Exemples de polymorphismes des gènes <i>UGT</i> qui affectent l'exposition au médicament.....	33
Tableau 9. Issues cliniques étudiées en lien avec le régime FOLFIRI.....	42

Liste des abréviations

5-FU	5-fluorouracile
ABC	<i>ATP-binding cassette</i>
ADME	absorption, distribution, métabolisme, excrétion
ADN	acide désoxyribonucléique
AhR	récepteur aux hydrocarbures aromatiques
ALG9	alpha-1,2-mannosyltransférase
APC	<i>7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino] carbonyloxycamptothecin</i>
AR	récepteur des androgènes
ARN	acide ribonucléique
AR-V7	récepteur des androgènes – variant 7
ATP	adénosine triphosphate
BCRP	<i>breast cancer resistance protein</i>
BSEP	<i>bile salt export pump</i>
CAR	récepteur constitutif des androstanes (gène <i>NR1I3</i>)
CCRm	cancer colorectal métastatique
CES	carboxylestérase
CLCN3	<i>chloride voltage-gated channel 3</i>
CNV	variation du nombre de copies
CPIC	<i>Clinical Pharmacogenetics Implementation Consortium</i>
CPT-11	irinotécan (<i>7-ethyl-10[4-(1-piperidino)-1-piperidino] carbonyloxycamptothecin</i>)
CTC	cellule tumorale circulante
CV	coefficient de variation
CYP	cytochrome P450
DPWG	<i>Dutch Pharmacogenetics Working Group</i>
<i>DPYD</i>	gène de la dihydropyrimidine déshydrogénase (DPD)
FDA	<i>Food and Drug Administration</i>
FOLFIRI	acide folinique + 5-fluorouracile + irinotécan
GST	glutathion S-transférase
GTEx	<i>the Genotype-Tissue Expression project</i>
HNF	<i>hepatocyte nuclear factors</i>
HR	<i>hazard ratio</i>
htSNP	polymorphismes étiquettes
kb	kilobase
LD	déséquilibre de liaison
MDR	<i>multidrug resistance protein</i>
miRNA	micro-ARN

MRP	<i>multidrug resistance-related protein</i>
NADPH	nicotinamide adénine dinucléotide phosphate, forme réduite
NPC	<i>7-ethyl-10-(4-amino-1-piperidino) carbonyloxycamptothecin</i>
OAT/OCT	<i>organic anion transporters/organic cation transporters</i>
OATP	<i>organic anion transporting polypeptides</i>
OS	survie globale
PBREM	<i>phenobarbital-responsive enhancer module</i>
PCR	réaction en chaîne par polymérase
PD	pharmacodynamie
PFS	survie sans progression
PK	pharmacocinétique
PXR	<i>pregnane X receptor</i> (gène <i>NR1I2</i>)
RPL	protéine ribosomique de la grande sous-unité 60S
RPS	protéine ribosomique de la petite sous-unité 40S
r_s	coefficient de corrélation de Spearman
SLC	<i>solute carrier</i>
SN-38	métabolite actif de l'irinotécan, <i>7-ethyl-10-hydroxycamptothecin</i>
SN-38G	SN-38 glucuronide
SNP	polymorphisme d'un seul nucléotide
SULT	sulfotransférase
TCGA	<i>The Cancer Genome Atlas</i>
TOP1	topoisomérase 1
TS	thymidylate synthase
UDP	uridine diphosphate
UDPGlcA	acide UDP-glucuronique
UGT	UDP-glucuronosyltransférase
U-PGx	<i>Ubiquitous Pharmacogenomics consortium</i>
USF	<i>upstream stimulatory factor</i>
<i>USP40</i>	<i>ubiquitin carboxyl-terminal hydrolase 40</i>
UTR	région non traduite
VDR	récepteur de la vitamine D
<i>VKORC1</i>	<i>vitamin K epoxide reductase complex subunit 1</i>

Remerciements

Je tiens tout d'abord à remercier Dr Chantal Guillemette, ma directrice de thèse, qui m'a permis de réaliser mon doctorat dans son laboratoire et m'a soutenu tout au long de mon parcours. J'ai grandement apprécié d'avoir eu la chance de pouvoir contribuer à différents projets, dans un laboratoire productif et dynamique. Merci pour votre confiance dans mon travail, grâce à votre aide j'ai appris à développer mon sens critique et mon autonomie.

Je veux ensuite remercier les membres du jury Drs Durocher, Drolet et Lorient, qui ont accepté d'évaluer ma thèse. Merci aussi à nos collaborateurs : le Dr Toffoli et toute son équipe ainsi que certains membres de la plateforme de bioinformatique qui m'ont aidé dans l'obtention de données de génotypes et d'expression de cohortes publiques. Un grand merci aussi à David Simonyan, de la plateforme de recherche clinique et évaluative pour sa contribution à mon projet de recherche ainsi que pour ses conseils et explications en statistiques.

Je souhaite également remercier tous les membres du laboratoire pour leur soutien. Merci Michèle pour toute ton aide, ton temps et tes encouragements en toutes circonstances. Merci Lyne pour ta bonne humeur et tes conseils judicieux dans le lab tout au long de ces 4 années. Merci aussi aux autres membres du labo – actuels et anciens – pour leur disponibilité et leur aide : Patrick, Véronique, Sylvie, Andréa, Anne-Marie et Camille.

Yannick et Eric, c'est probablement vous que j'ai le plus vu durant mon doctorat, notamment grâce à notre cubicule-colocation de 2 m³. Un grand merci pour vos conseils et votre aide, et surtout pour les repas, 5 à 7, soirées et autres bons moments. Merci également Joanie, pour ton calme, ta patience et ta gentillesse. Bonne chance à vous tous pour la suite de vos parcours. Bonne chance aussi aux autres étudiants du groupe, Vincent et Jean-Philippe pour la suite de leurs études.

Je veux aussi remercier Dr Éric Lévesque pour son enthousiasme et pour l'opportunité offerte de m'impliquer dans le projet portant sur le cancer de la prostate mené conjointement avec le groupe du Dre Guillemette, et portant spécifiquement sur le rôle de la protéine UGT2B17 (non présenté dans cette thèse). Merci aussi aux Drs Isabelle

Laverdière, Étienne Audet-Walsh et Jean-Philippe Lambert et à leurs équipes respectives pour leurs conseils et leur dynamisme. Je souhaite aussi remercier toutes les personnes du secteur R4 pour leur bonne humeur : Mandy, Annie-Claude, Mélanie, Stéphane et tous les autres.

Merci à tous mes amis, en France, au Canada et ailleurs pour les bons moments passés ensemble. Je souhaite terminer en remerciant ma famille, qui m'a toujours soutenu dans mes projets. Plus particulièrement, merci à mes parents Dominique et Brigitte, ma sœur Lauriane et ma grand-mère Paulette. Enfin, merci Marie pour ton soutien tout au long de cette aventure.

Avant-propos

Cette thèse intitulée « Pharmacogénomique de la voie de glucuronidation : mécanismes moléculaires et impact clinique » est présentée à la Faculté des études supérieures et postdoctorales de l'Université Laval pour l'obtention du grade de *Philosophiae doctor*. Elle a été rédigée sous la forme d'une thèse par articles.

Contributions présentées comme chapitres dans cette thèse.

Le premier article est intitulé « *Combination of germline variations in irinotecan pharmacogenetic pathways associated with survival in metastatic colorectal cancer treated with FOLFIRI-based chemotherapy* ». Je suis premier auteur de cet article et ma contribution s'élève à 80% du manuscrit. J'ai contribué aux analyses des données de survie en lien avec les données génétiques qui étaient disponibles. J'ai réalisé les essais *in vitro*, l'analyse des données d'expression génique des bases de données publiques. J'ai participé à l'interprétation des données, à la rédaction et la révision du manuscrit. Éric Lévesque, Derek Jonker et Félix Couture ont recruté la cohorte canadienne alors que nos collègues les Drs Elena De Mattia, Erika Cecchin, Angela Buonadonna, Mario d'Andrea et Giuseppe Toffoli ont constitué la cohorte italienne. Lyne Villeneuve a effectué des analyses génétiques et validé les analyses statistiques. David Simonyan a contribué aux analyses des données de survie. Chantal Guillemette a conceptualisé et supervisé l'étude, participé à l'analyse des données, a rédigé et révisé le manuscrit. Tous les auteurs ont participé à la révision du manuscrit. L'article inséré dans cette thèse est identique à la version la plus récente.

Le deuxième article est intitulé « *Germline variability and tumor expression level of ribosomal protein gene RPL28 are associated with survival of metastatic colorectal cancer patients* », dont la version révisée a été soumise pour publication le 21 juin 2019. Je suis premier auteur de cet article et ma contribution s'élève à 80% du manuscrit. J'ai effectué les analyses des données de survie et d'expression, et participé au génotypage et aux essais cellulaires. J'ai contribué à la rédaction et la révision du manuscrit. Éric Lévesque, Derek Jonker et Félix Couture ont recruté la cohorte canadienne alors que nos collègues les Drs

Elena De Mattia, Erika Cecchin, Angela Buonadonna, Mario d'Andrea et Giuseppe Toffoli ont constitué la cohorte italienne. Lyne Villeneuve a validé les analyses statistiques et contribué au génotypage. David Simonyan a contribué aux analyses des données de survie. Eric P. Allain a contribué à l'analyse biostatistique des données d'expression. Chantal Guillemette a conceptualisé et supervisé l'étude et participé à l'analyse des données. Chantal Guillemette, Éric Lévesque et Michèle Rouleau ont rédigé le manuscrit auquel j'ai également contribué. Tous les auteurs ont révisé le manuscrit. L'article inséré dans cette thèse est identique à la version la plus récente.

Le troisième article est intitulé « *Improved progression-free survival in irinotecan-treated metastatic colorectal cancer patients carrying the HNF1A coding variant p.I27L* » et a été publié dans le journal *Frontiers in Pharmacology* (accepté le 22 septembre 2017, publié le 10 octobre 2017). Je suis premier co-auteur et ma contribution s'élève à 70% du travail. J'ai effectué les analyses des données et réalisé les essais cellulaires, participé à l'analyse des données et à la rédaction du manuscrit. Elena De Mattia, Erika Cecchin, Angela Buonadonna, Mario d'Andrea et Giuseppe Toffoli ont recruté la cohorte italienne, contribué à l'obtention des génotypes, effectué les analyses sur les données de survie et fourni les données pharmacocinétiques. Éric Lévesque, Derek Jonker et Félix Couture ont constitué la cohorte canadienne. Lyne Villeneuve a validé les analyses statistiques et contribué au génotypage. Chantal Guillemette a conceptualisé et supervisé l'étude, participé à l'analyse des données et à la rédaction du manuscrit. Tous les auteurs ont participé à la révision du manuscrit. L'article inséré dans cette thèse est identique à la version la plus récente.

Le quatrième article est intitulé « *Post-transcriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing* » et a été publié dans le journal *Drug Metabolism & Disposition* (accepté le 31 janvier 2018, publié en ligne le 4 avril 2018). Je suis premier auteur de cet article et ma contribution s'élève à 75% du manuscrit. J'ai fait les analyses de données et statistiques, les immunobuvardages, les essais enzymatiques, les essais de glycosylation, les co-immunoprécipitations et les essais de stabilité des protéines à l'étude. J'ai aussi participé à l'obtention des modèles *in vitro* et à la rédaction. Eric P. Allain a analysé les données de séquençage à haut débit. Yannick Audet-Delage a réalisé les

immunofluorescences. Lyne Villeneuve a contribué à la mise en place des modèles cellulaires, aux essais de stabilité protéique et a réalisé les essais enzymatiques en cellules intactes. Michèle Rouleau a contribué à la conceptualisation et la supervision de l'étude, l'analyse des données, et a participé à la rédaction du manuscrit. Chantal Guillemette a conceptualisé et supervisé l'étude, contribué à l'analyse des données et à la rédaction du manuscrit. Tous les auteurs ont participé à la révision du manuscrit. L'article inséré dans cette thèse est identique à la version la plus récente.

Autres contributions non-présentées dans cette thèse.

J'ai contribué à un cinquième article intitulé « *Factors affecting interindividual variability of hepatic UGT2B17 protein expression examined using a novel specific monoclonal antibody* » qui a été publié dans le journal *Drug Metabolism & Disposition* (accepté le 21 février 2019, publié en ligne le 28 mars 2019). Je suis deuxième auteur de cet article et ma contribution s'élève à 10% du manuscrit. J'ai participé aux analyses statistiques et aux essais cellulaires.

J'ai contribué à un sixième article intitulé « *Association of STAT-3 rs1053004 and VDR rs11574077 with FOLFIRI-related gastrointestinal toxicity in metastatic colorectal cancer patients* » qui a été publié dans le journal *Frontiers in Pharmacology* (accepté le 28 mars 2018, publié le 13 avril 2018). Je suis quatrième auteur de cet article et ma contribution s'élève à 10% du manuscrit. J'ai contribué aux analyses génétiques.

J'ai contribué à un septième article intitulé « *Germline polymorphisms in the nuclear receptors PXR and VDR as novel prognostic markers in metastatic colorectal cancer patients treated with FOLFIRI* » qui a été soumis pour publication le 26 juin 2019. Je suis quatrième auteur de cet article et ma contribution s'élève à 10% du manuscrit. J'ai contribué aux analyses génétiques.

Introduction

La pharmacogénomique est l'étude de l'influence du génome sur la pharmacologie. Elle vise à identifier des variations du génome qui influencent la réponse clinique, notamment dans les gènes de la pharmacocinétique (PK) et de la pharmacodynamie (PD) du médicament. Actuellement, la majorité des traitements sont administrés à tous les patients selon une dose standard, parfois ajustée en fonction de certains facteurs physiologiques ou pathologiques. Cependant, la réponse au traitement, en termes d'innocuité et d'efficacité, est très variable entre les individus, notamment à cause de variations dans les voies PK/PD. En oncologie, les variations génétiques héréditaires et les variations tumorales somatiques constituent les marqueurs pharmacogénomiques les plus étudiés (**Tableau 1**). Les mutations somatiques sont le plus souvent liées à la réponse de la tumeur et donc la survie du patient. Les variations héréditaires (ou germinales) ont davantage été étudiées en lien avec les effets indésirables induits par le traitement, mais il apparaît désormais qu'elles sont également liées à la survie des patients.

De plus en plus de médicaments sont accompagnés de mises en garde, d'informations ou de recommandations pharmacogénomiques dans leur monographie. C'est par exemple le cas de 260 médicaments autorisés par la FDA¹ (*US Food and Drug Administration*). Bien que la personnalisation des traitements soit en plein essor, plusieurs limites ralentissent sa mise en application clinique. Le manque de lignes directrices standardisées sur les actions à prendre par les cliniciens suite au résultat d'un test génétique est notamment une limitation importante. Des regroupements internationaux tels que le CPIC (*Clinical Pharmacogenetics Implementation Consortium*), le DPWG (*Dutch Pharmacogenetics Working Group*) ou le consortium U-PGx (*Ubiquitous Pharmacogenomics*) travaillent dans l'optique d'y remédier (Bank et al., 2018; Caudle et al., 2014; van der Wouden et al., 2017). Par exemple, le CPIC répertorie actuellement 42 médicaments, dont 6 agents anticancéreux, classés dans la catégorie A. Cette catégorie correspond à des paires gènes/médicaments pour lesquelles il existe des preuves solides selon lesquelles la génétique du patient altère sa réponse, et pour lesquelles une adaptation du traitement est

¹ <https://www.fda.gov/Drugs/ScienceResearch/ucm572698.htm>, consulté le 2 avril 2019

possible et recommandée (ajustement de la dose ou thérapie alternative) selon le résultat du test génétique. Dans le cas des antinéoplasiques, qui ont souvent une fenêtre thérapeutique étroite, de faibles variations dans la PK/PD peuvent fortement altérer la réponse du patient, démontrant la nécessité de développer de nouveaux outils et recommandations basés sur la pharmacogénomique pour optimiser l'usage de ces agents.

Tableau 1. Exemples de marqueurs de nature somatique et héréditaire pour lesquels un effet sur la réponse aux anticancéreux a été démontré.

Gène	Médicament	Marqueur	Issue clinique
<i>ABL</i>	Bosutinib, dasatinib, imatinib, nilotinib, ponatinib	Somatique	Efficacité
<i>ALK</i>	Crizotinib	Somatique	Efficacité
<i>BRAF</i>	Vemurafenib	Somatique	Efficacité
<i>EGFR</i>	Afatinib, cétuximab, erlotinib, panitumumab, vandétanib	Somatique	Efficacité
<i>FCγR</i>	Cétuximab, rituximab, trastuzumab	Somatique	Efficacité
<i>HER2</i>	Lapatinib, pertuzumab, trastuzumab, emtansine	Somatique	Efficacité
<i>KRAS</i>	Cétuximab, panitumumab	Somatique	Efficacité
<i>KIT</i>	Imatinib	Somatique	Efficacité
<i>MET</i>	Tramétinib	Somatique	Efficacité
<i>BIM</i>	Imatinib	Germinal	Efficacité
<i>CYP2B6</i>	Cyclophosphamide	Germinal	Toxicité
<i>CYP2D6</i>	Tamoxifène	Germinal	Efficacité
<i>DPYD</i>	Capécitabine, 5-fluorouracile	Germinal	Toxicité
<i>G6PD</i>	Rasburicase	Germinal	Toxicité
<i>MLH1, MSH2, MSH6, PMS2</i>	5-fluorouracile	Germinal	Efficacité
<i>SLCO1B1</i>	Méthotrexate	Germinal	Toxicité
<i>SLC28A3</i>	Anthracyclines	Germinal	Toxicité
<i>TCL1A</i>	Inhibiteurs de l'aromatase	Germinal	Toxicité
<i>TPMT</i>	Mercaptopurines, thioguanine, cisplatine	Germinal	Toxicité
<i>UGT1A1</i>	Irinotécan	Germinal	Toxicité

Adapté de Filipski et al., 2014.

Environ 55% des médicaments les plus prescrits sont inactivés et éliminés, principalement au foie, par les enzymes UDP-glucuronosyltransférases (UGT), ce qui en fait une voie PK primordiale dans la réponse au médicament (Guillemette et al., 2014). C'est le cas de l'agent antinéoplasique irinotécan, dont l'inactivation du métabolite actif SN-38 est principalement médiée par l'enzyme UGT1A1 au niveau hépatique et qui est un des premiers exemples d'utilisation de la pharmacogénomique. Ce médicament est utilisé contre le cancer colorectal métastatique (CCRm) en combinaison avec le 5-fluorouracile (5-FU) et l'acide folinique dans le régime FOLFIRI, qui constitue un des traitements de chimiothérapie de première intention (Modest et al., 2019; Sanchez-Gundin et al., 2018; Sandhu et al., 2019). Le variant génétique *UGT1A1**28, entraînant une réduction de l'activité d'UGT1A1 et donc de l'élimination du SN-38, a été associé à un risque accru de neutropénies de grades 3-4 causées par ce médicament. Au-delà de l'étude des variants génétiques, le domaine s'intéresse aussi à d'autres sources de variabilité interindividuelle affectant les voies PK/PD et donc potentiellement la réponse clinique du patient au traitement. C'est le cas de l'épissage alternatif, qui apparaît comme un mécanisme affectant la réponse à divers traitements en oncologie (**Tableau 2**). L'épissage alternatif est une source majeure de diversité pour la voie de glucuronidation. En effet, selon les récents travaux du laboratoire, plus de 130 variants alternatifs ont été répertoriés en plus des 19 transcrits canoniques codant pour les 19 enzymes UGT (Tourancheau et al., 2016). Parmi les UGT, l'enzyme UGT2B10 est un bon candidat à étudier puisque plus de la moitié de son transcriptome hépatique est constitué de variants alternatifs.

La première section de l'introduction vise à présenter les conditions pour l'utilisation en clinique d'un marqueur pharmacogénomique et des exemples appliqués à l'oncologie. La deuxième section présente ensuite les voies PK/PD impliquées dans la réponse au médicament, qui sont les candidates les plus étudiées pour la découverte de marqueurs prédictifs de la réponse. Parmi ces voies, la voie de glucuronidation est présentée dans la troisième section. Nos travaux s'intéressant aux variations génétiques héréditaires des voies PK/PD en lien avec la réponse au traitement, incluant la voie de glucuronidation, sont présentés dans les chapitres I à III ; ceux investiguant le rôle de l'épissage alternatif de la voie de glucuronidation sont présentés dans le chapitre IV.

Tableau 2. Exemples de variants d'épissage alternatif pour lesquels un effet sur la réponse aux anticancéreux a été démontré.

Gène	Médicament	Cancer	Issue clinique	Référence
<i>AR</i>	Thérapies anti-androgènes	Prostate	Efficacité	(Paschalis et al., 2018)
<i>NCOR2</i>	Tamoxifène	Sein	Efficacité	(Zhang et al., 2013)
<i>SPP1</i>	Tamoxifène, CMF	Sein	Efficacité	(Zduniak et al., 2016)
<i>CASP3</i>	CEF	Sein	Efficacité	(Vegran et al., 2006)
<i>BCR-ABL1</i>	Imatinib	LMC	Efficacité	(Berman et al., 2016)
<i>SLC22A1</i>	Imatinib	LMC	Efficacité	(Grinfeld et al., 2013)
<i>DPYD</i>	Fluoropyrimidines	Plusieurs	Toxicité	(Meulendijks et al., 2015)

CEF, cyclophosphamide + épirubicine + 5-fluorouracile ; CMF, cyclophosphamide + méthotrexate + 5-fluorouracile ; LMC, Leucémie myéloïde chronique.

1) La pharmacogénomique en clinique

Certains facteurs sont reconnus pour impacter la réponse entre les patients et sont pris en compte en clinique lors du traitement. Il s'agit par exemple de l'âge, du sexe, du statut de performance, de la consommation d'alcool ou de tabac, de la fonction hépatique ou rénale. Malgré leur prise en compte, la réponse demeure souvent très différente d'un patient à l'autre. Cette fluctuation qu'on ne peut attribuer aux facteurs physiologiques ou environnementaux peut découler de la variabilité génétique entre les individus. La pharmacogénomique cherche à identifier une ou des variation(s) génétique(s) permettant de prédire la réponse de chaque patient au traitement.

Le polymorphisme *UGT1A1**28 est un exemple portant sur le risque de neutropénies sévères causées par l'irinotécan. C'est le variant le plus étudié du locus *UGT1* et il consiste en la répétition de 7 motifs TA au lieu de 6 dans la boîte TATA du promoteur d'*UGT1A1*. Il a été associé au syndrome de Gilbert, caractérisé par une hyperbilirubinémie asymptomatique résultant d'une baisse de la glucuronidation de la bilirubine (Beutler et al., 1998; Bosma et al., 1995; Monaghan et al., 1996). Sa fréquence change entre les groupes ethniques allant de 0 à 3% chez les Asiatiques, 2 à 13% chez les Caucasiens et 16 à 19% chez les Africains. Cette affection de la région promotrice entraîne une diminution de l'expression et de l'activité d'*UGT1A1* de 30 à 70% et une augmentation d'environ 15% de

l'exposition systémique au SN-38 (Bosma et al., 1995; Guillemette et al., 2000; Innocenti et al., 2009). Ainsi, chez les patients traités avec le régime FOLFIRI et qui sont porteurs homozygotes du polymorphisme *UGT1A1**28, le risque de développer des neutropénies sévères est plus élevé que chez les homozygotes non porteurs, car la diminution de la glucuronidation du SN-38 par *UGT1A1* induit une plus grande exposition au métabolite actif toxique (Cote et al., 2007; Etienne-Grimaldi et al., 2015; Iyer et al., 2002; Liu et al., 2008; Marcuello et al., 2004; Toffoli et al., 2006). Plusieurs agences de réglementation recommandent donc une réduction de la dose d'irinotécan administrée aux patients homozygotes pour le polymorphisme *UGT1A1**28 (Etienne-Grimaldi et al., 2015; Hoskins et al., 2007; Stingl et al., 2014). Ainsi, la dose standard de 180 mg/m² d'irinotécan est inférieure à celle qui pourrait être tolérée par les patients *UGT1A1**1/*1 (370 à 390 mg/m²) et *UGT1A1**1/*28 (310 à 340 mg/m²) alors qu'elle devrait être réduite de 30% pour les patients *UGT1A1**28/*28 atteints du CCRm traités avec le régime FOLFIRI (Marcuello et al., 2011; Toffoli et al., 2010). La diminution de dose pour les porteurs *UGT1A1**28/*28 permet de maintenir un risque de neutropénies sévères similaire à celui des non-porteurs traités avec la dose standard. L'augmentation de la dose pour les non porteurs homozygotes conduit à une augmentation du taux de réponse sans augmentation de la survenue d'effets indésirables. L'ajout de bévacizumab au régime FOLFIRI ne change pas cette observation, les patients *UGT1A1**1/*1 et *UGT1A1**1/*28 pouvant recevoir une augmentation de la dose d'irinotécan afin d'améliorer la réponse thérapeutique sans augmenter le risque d'effets indésirables (Lu et al., 2015; Toffoli et al., 2017; Yeh et al., 2016). Le marqueur *UGT1A1**28 est associé à une variation d'environ 15% de l'exposition au SN-38, mais jusqu'à 30% de la variation dans l'exposition au SN-38 peut être expliquée par les polymorphismes *ABCC1* rs35605, *ABCB1* rs10276036 et *UGT1A1**93 (Innocenti et al., 2009). Cela semble indiquer que l'utilisation de plusieurs marqueurs expliquant une plus grande part de variation dans l'exposition au traitement, comparativement à l'utilisation d'un seul marqueur, pourrait améliorer la prédiction de réponse des patients. Cependant, 70% de la fluctuation dans l'exposition au SN-38 reste à être expliquée, suggérant que d'autres variants génétiques et/ou d'autres sources de variabilité puissent influencer l'exposition au médicament, et donc potentiellement les toxicités et la réponse.

Bien que le domaine de la pharmacogénomique soit loin d'avoir évalué l'impact de tous les variants génétiques héréditaires dans la réponse au médicament, un intérêt existe déjà sur la contribution d'autres sources de variabilité interindividuelle, comme par exemple l'épissage alternatif. Un des exemples les plus intéressants pour lequel plusieurs études ont été publiées est celui du variant d'épissage alternatif 7 du récepteur aux androgènes (AR-V7) (Li et al., 2018; Paschalis et al., 2018). Le cancer de la prostate métastatique résistant à la castration dépend toujours de la signalisation par le récepteur aux androgènes (AR) malgré la chirurgie (Montgomery et al., 2008). L'enzalutamide (inhibiteur du AR) et l'abiratéron (inhibiteur de la synthèse des androgènes) sont deux thérapies ciblées utilisées pour supprimer la signalisation via le AR et limiter le développement du cancer (Beer et al., 2014; Ryan et al., 2013). Or, 20 à 40% des patients ne répondent pas à ces thérapies dès leur initiation. Cette résistance primaire est notamment due à l'expression par ces patients du AR-V7, car ce récepteur variant induit une signalisation constitutive sans être lié par les androgènes. Ainsi, il a été montré que la détection du AR-V7 dans les cellules tumorales circulantes (CTC) des patients était associée à la résistance au traitement à l'enzalutamide et l'abiratéron (Antonarakis et al., 2014; Hu et al., 2009). En revanche, la présence du AR-V7 n'altère pas la réponse à la chimiothérapie à base de taxanes, qui est une ligne de traitement efficace après l'enzalutamide ou l'abiratéron pour les patients résistants à ces thérapies ciblées (Antonarakis et al., 2015; Scher et al., 2016). Un test permettant de détecter l'expression du AR-V7 dans les CTC a été développé et une étude préliminaire a montré que le résultat de ce test influençait la décision clinique, notamment pour les patients AR-V7 positifs (Markowski et al., 2017). Suite à cette étude, une étude prospective sur 120 patients est en cours pour valider ces observations (NCT02269982). Le variant d'épissage AR-V7 est donc un marqueur pharmacogénomique prometteur puisqu'il permet d'identifier les patients qui ne répondront pas à l'enzalutamide et l'abiratéron, et qu'il existe des alternatives de traitement efficaces.

Malgré le développement important du domaine de la pharmacogénomique, son application clinique s'avère parfois longue et difficile à mettre en place. En effet, après la découverte du lien entre une variation génétique et la réponse à un médicament, une étape de validation du marqueur doit être faite. Il existe deux grandes écoles de pensées : la première affirmant qu'un marqueur doit être validé dans une étude randomisée idéalement en double aveugle,

la seconde considérant qu'une validation prospective est suffisante (Filipski et al., 2014; Kapoor et al., 2016; van der Wouden et al., 2016). Cette divergence dans le type de validation nécessaire pour porter le marqueur en clinique, en plus du temps et du coût nécessaire à la réalisation de l'étude peuvent considérablement ralentir l'arrivée du marqueur en clinique. Ensuite, un test génétique dont la validité analytique et clinique a été démontrée doit être disponible. Son utilisation devrait permettre d'améliorer l'issue clinique pour le patient. L'utilité clinique repose sur le fait qu'un test réalisé en lien avec un traitement propose une alternative (par exemple une modulation de la dose ou une autre thérapie efficace) à un patient qui présente un risque de toxicité ou d'absence de réponse au traitement pour lequel le test a été effectué (Relling et al., 2015). La rentabilité pour le système de santé et la prise en charge des coûts associés sont aussi des facteurs à prendre en compte dans la mise en application clinique du marqueur et peuvent limiter son développement. Enfin, le manque de formation et d'information des professionnels de santé a été identifié comme une des causes majeures de la difficulté à implanter la pharmacogénomique en clinique (Relling et al., 2015; van der Wouden et al., 2016). Des consortiums internationaux (CPIC, DPWG, U-PGx, etc.) ont donc été créés pour informer les cliniciens sur les tests génétiques et établir des lignes directrices standardisées afin de les aider à adapter la prescription efficacement selon le profil génétique du patient.

2) Voies PK/PD des traitements.

La PK d'un médicament correspond à son absorption, sa distribution, son métabolisme et son excrétion (ADME) et constitue donc le devenir de la molécule dans l'organisme. Elle est complémentaire à sa PD, qui correspond aux effets produits par le médicament sur l'organisme. La variabilité de ces voies pourrait influencer la réponse au traitement.

2.1) Pharmacocinétique : absorption, distribution, métabolisme et excrétion

Selon le mode d'administration du traitement (oral, intraveineux, aérosol, etc.), l'absorption du médicament va se dérouler au niveau de divers organes comme les intestins, le foie ou les poumons. Cela va affecter la biodisponibilité du médicament, définie comme la fraction de la dose administrée qui parvient dans la circulation systémique sous la même forme après le premier passage hépatique, en tenant compte de la vitesse à laquelle cela se

produit. De la même façon une excrétion via l'urine va impliquer les reins alors qu'une excrétion via la bile va impliquer le foie et les intestins. On peut ainsi estimer la clairance pour chaque organe, c'est-à-dire le volume de sang épuré par unité de temps. La clairance totale correspond à la somme des clairances de chaque organe. La demi-vie du médicament, dépendante du volume de distribution et de la clairance, correspond au temps requis pour diminuer de moitié la concentration plasmatique du médicament.

Il existe trois façons principales pour le médicament et ses métabolites de pénétrer dans la cellule et d'en sortir : par diffusion passive à travers la membrane lipidique, par des mécanismes de trafic membranaire ou par le biais de transporteurs (**Figure 1**) (Cocucci et al., 2017). Les transporteurs impliqués dans l'absorption et l'excrétion des médicaments et leurs métabolites se séparent principalement en deux grandes familles : celle des transporteurs SLC (*solute carriers*) et celle des transporteurs ABC (*ATP-binding cassette*). Les transporteurs SLC sont surtout impliqués dans l'absorption intracellulaire des substrats sur la base de mécanismes passifs, couplés ou d'échanges, ne dépendant pas directement de l'hydrolyse d'ATP (Nigam, 2015; Sprowl et al., 2014). La famille de protéines SLC se décompose en plusieurs sous-familles dont les OATP (*organic anion-transporting polypeptides*) et les OAT/OCT (*organic cation and anion transporters*) sont les plus étudiées en lien avec le métabolisme des médicaments. Au niveau hépatique ce sont OATP1B1, OATP1B3 et OCT1 qui sont les plus exprimés et sont responsables de l'absorption de diverses molécules telles que les agents anticancéreux oxaliplatine, cisplatine, irinotécan ou paclitaxel (Li et al., 2014; Sprowl et al., 2014). Les transporteurs ABC sont quant à eux impliqués dans l'efflux des xénobiotiques d'une façon directement dépendante de l'hydrolyse d'ATP (Chen et al., 2016; Nigam, 2015). Parmi les protéines ABC les plus étudiées pour leur rôle dans l'exposition aux médicaments se trouvent notamment MDR1, BSEP, BCRP et les MRP1-9. L'ensemble des transporteurs sont parfois regroupés sous l'appellation de phase III du métabolisme des médicaments.

L'expression des transporteurs est très différente entre les individus. Ainsi, plusieurs études regroupant de 4 à 51 individus, ont démontré une importante variabilité de l'expression de la protéine MRP2 au foie avec des coefficients de variation (CV) allant de 42 à 87% (Deo et al., 2012; Fallon et al., 2016; Li et al., 2009; Ohtsuki et al., 2012; Tucker et al., 2012).

Deux de ces études comprenant respectivement 51 et 13 sujets ont également montré que, pour les individus ayant un niveau protéique de MRP2 détectable, la différence entre l'individu exprimant le moins et celui exprimant le plus MRP2 variait de 6,6 à 7 fois (Deo et al., 2012; Tucker et al., 2012). L'expression protéique de MDR1 au foie est aussi irrégulière avec des CV de 39% à 76% (13 et 4 individus, respectivement), et une étude quantifie une variation de 4 fois entre l'expression minimale et celle maximale retrouvée chez 13 individus (Fallon et al., 2016; Tucker et al., 2012). Ces différentes études démontrent aussi une variabilité importante dans l'expression hépatique d'autres transporteurs, comme par exemple BCRP ou MRP3.

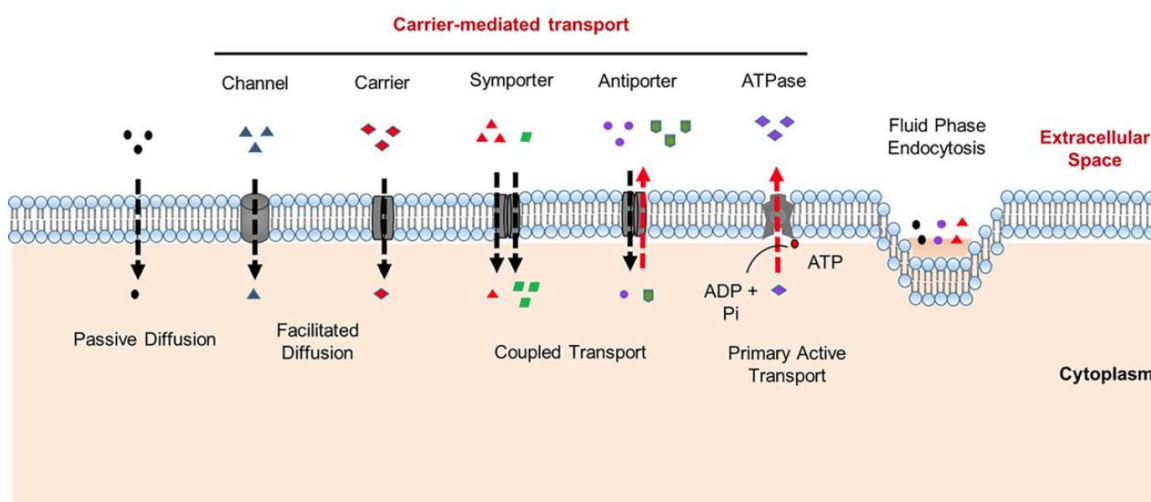


Figure 1. Mécanismes impliqués dans l'influx et l'efflux des médicaments.

Figure issue de Cocucci et al., 2017. Reproduit avec la permission du *Copyright Clearance Center* pour *John Wiley and Sons*.

La distribution du médicament se caractérise tout d'abord par son transport sanguin et sa fixation aux protéines plasmatiques, notamment l'albumine qui est la plus abondante. Cette fixation joue un rôle de réserve de médicament, car seule la fraction libre de médicament est active. Ensuite, la diffusion tissulaire correspond à la répartition du médicament dans les organes et tissus. Elle dépend des transporteurs membranaires d'influx et d'efflux, du débit sanguin, des propriétés physico-chimiques de la molécule et de la fixation aux protéines tissulaires. Le volume de distribution, qui correspond à la quantité de médicament dans l'organisme divisé par sa concentration plasmatique, et la fixation protéique sont les paramètres essentiels de la distribution du médicament. Ces paramètres et leurs variations

ne sont cependant pas l'objet de nos travaux, car ils ne sont pas directement liés à la génétique du patient.

Les réactions de phase I et II du métabolisme des médicaments vont influencer la distribution et la clairance du médicament en modifiant la molécule et donc ses propriétés physico-chimiques. La phase I du métabolisme des médicaments correspond à la fonctionnalisation de molécules lipophiles, afin de les rendre davantage solubles et actives, par la formation ou le clivage de divers groupements via des réactions d'oxydation, de réduction ou d'hydrolyse (**Figure 2**). Bien qu'il existe plusieurs types de réactions catalysées par différentes familles d'enzymes, ce sont principalement les cytochromes P450 (CYP) qui contribuent aux réactions de phase I du métabolisme des médicaments (Evans et al., 1999).

La famille des CYP regroupe chez l'homme 57 enzymes fonctionnelles et 58 pseudogènes (Nelson et al., 2004). Cependant, la majeure partie de l'activité envers les xénobiotiques incluant les médicaments est catalysée par 11 enzymes des familles 1, 2 et 3. Parmi les médicaments actuellement utilisés en clinique, environ 30% sont métabolisés par CYP3A4/3A5, 20% par CYP2D6 et 13% par CYP2C9 qui sont les principales enzymes contributrices (Zanger et al., 2013). Elles sont exprimées au niveau du réticulum endoplasmique principalement dans les cellules hépatiques mais aussi dans de multiples autres tissus (Park et al., 2014; Werck-Reichhart et al., 2000). Les CYP sont des hémoprotéines ayant le plus souvent une activité monooxygénase formant un groupement hydroxyle en utilisant de l'oxygène et du NADPH (McLean et al., 2015). Parmi leurs substrats, on retrouve de nombreux médicaments comme par exemple le tamoxifène, l'abiratéron, le sorafénib ou le paclitaxel qui sont aussi des substrats des enzymes UGT du métabolisme de phase II, en plus de nombreuses molécules endogènes comme les stéroïdes (McDonnell et al., 2013).

CYP3A4 est la plus abondamment exprimée au niveau hépatique, malgré une large variabilité interindividuelle (Achour et al., 2014a). Une étude portant sur 42 tissus hépatiques humains a démontré une fluctuation d'expression protéique de 40 fois entre les individus exprimant le moins CYP3A4 et ceux l'exprimant le plus (Westlind-Johnsson et al., 2003). Au niveau de l'ARN, celle-ci est supérieure à 800 fois chez 32 sujets. Un autre

exemple est celui de CYP3A5, qui n'est exprimé que par environ 20 à 30% des individus (Aoyama et al., 1989; Koch et al., 2002; Westlind-Johnsson et al., 2003; Wrighton et al., 1990). Plusieurs polymorphismes contribuent à cette absence d'expression de CYP3A5, mais le plus fréquent est *CYP3A5*3* qui aboutit à l'introduction d'un codon stop prématuré (Daly, 2006; Kuehl et al., 2001). Parmi les individus exprimant CYP3A5, une étude a rapporté une variabilité d'expression protéique de plus de 7 fois (Westlind-Johnsson et al., 2003). Dans le cas de CYP2D6 on peut classer les individus en quatre grandes catégories à savoir métaboliseur lent, intermédiaire, rapide et ultra-rapide sur la base de l'activité de l'enzyme envers ses substrats (Tracy et al., 2016). L'expression de CYP2D6 est très irrégulière entre et au sein de ces différents groupes, allant d'une quantité indétectable de protéine pour les métaboliseurs lents, à 2,6 pmol/mg de protéine (CV = 104%) pour les métaboliseurs intermédiaires, 7,6 pmol/mg de protéine (CV = 62%) pour les métaboliseurs rapides et jusqu'à 24 pmol/mg de protéine (CV = 32%) pour les métaboliseurs ultra-rapides (Zanger et al., 2001). Le phénotype de métaboliseur lent est associé aux polymorphismes *CYP2D6*3* (décalage du cadre de lecture), *CYP2D6*4* (défaut d'épissage) et *CYP2D6*5* (délétion du gène) (Bertilsson et al., 2002). Cependant, toutes ces différences ne sont pas expliquées entièrement par les polymorphismes, notamment pour les individus qui ne sont pas des métaboliseurs lents mais dont l'expression change de 140 fois entre la plus faible et la plus importante, suggérant une contribution importante d'autres sources de variabilité (Gaedigk et al., 2008; Zanger et al., 2001). L'épissage alternatif pourrait y contribuer car pour les 57 gènes *CYP*, ce sont environ 1000 transcrits qui peuvent être générés (Annalora et al., 2017).

Les réactions de conjugaison constituent le métabolisme de phase II du médicament. Elles consistent au transfert d'un groupement sur une molécule dans le but de réduire son activité et augmenter sa solubilité pour favoriser son élimination de l'organisme par la bile ou l'urine. Cette voie est complémentaire au métabolisme de phase I, et ces deux voies peuvent aussi bien agir simultanément que séquentiellement. Les réactions de conjugaison sont effectuées en majorité par les UGT puis par les SULT, les GST et d'autres enzymes (**Figure 2**) (Evans et al., 1999). Les UGT sont une famille d'enzymes qui transfère le sucre du cosubstrat acide UDP-glucuronique (UDP-GlcA) sur leur substrat, une réaction nommée glucuronidation et qui se déroule principalement au foie mais aussi dans de nombreux

autres tissus. Les UGT conjuguent plus de 55 % des 200 médicaments les plus prescrits en plus de réguler les niveaux de nombreux substrats endogènes comme la bilirubine ou les hormones stéroïdiennes (Guillemette et al., 2014). Des β -glucuronidases d'origine humaine ou bactérienne peuvent cliver le glucuronide et libérer à nouveau la molécule active dans le tractus gastrointestinal. La voie de glucuronidation, ainsi que certaines sources de sa variabilité, sont présentées en détail dans la section 2 de l'introduction car il s'agit du principal sujet de mes travaux.

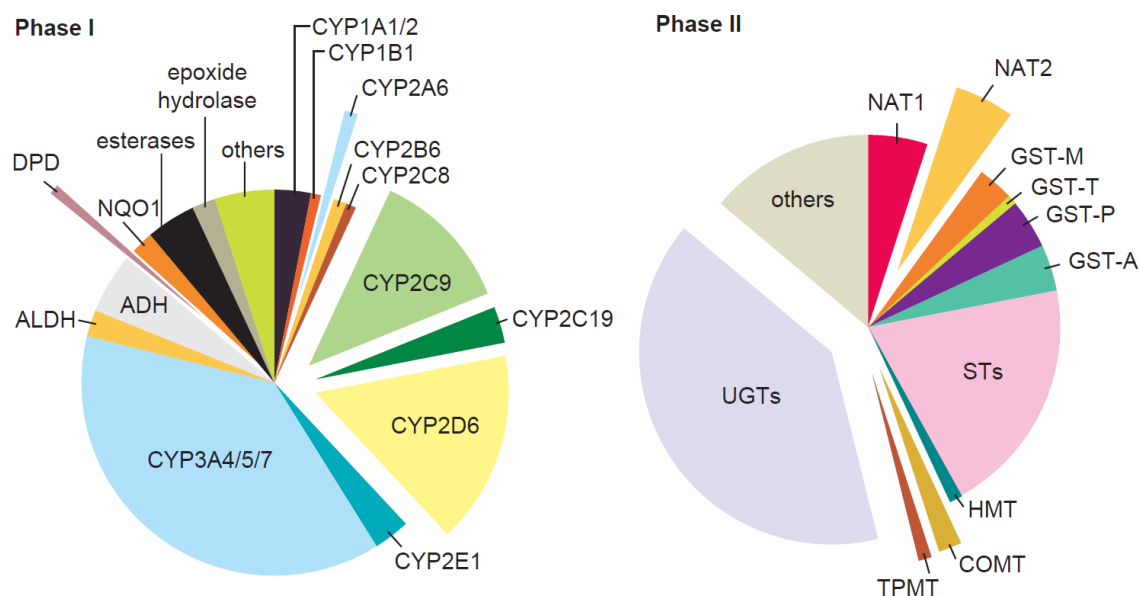


Figure 2. Enzymes du métabolisme de phase I (gauche) et II (droite) des médicaments. La contribution relative est estimée par la taille des sections du graphique. ADH, alcool déshydrogénase ; ALDH, aldéhyde déshydrogénase ; CYP, cytochrome P450 ; COMT, catéchol-O-méthyltransférase ; DPD, dihydropyrimidine déshydrogénase ; GST, glutathion S-transférase ; HMT, histamine méthyltransférase NQO1, NAD(P)H quinone oxydoréductase 1 ; NAT, N-acétyltransférase ; STs, sulfotransférases TPMT, thiopurine méthyltransférase ; UGTs, UDP-glucuronosyltransférases. Reproduit avec la permission de *American Association for the Advancement of Science* de *Evans, W.E. and Relling, M.V., Pharmacogenomics: translating functional genomics into rational therapeutics. Science, 1999. 286(5439): p. 487-91.*

2.2) Pharmacodynamie

Un composé pharmacologique cible généralement plusieurs voies cellulaires afin d'exercer son effet. Cet effet peut se décomposer en une réponse thérapeutique (effet voulu) et/ou la

survenue d'effets indésirables. Le terme de fenêtre thérapeutique se définit comme l'intervalle de concentration d'un traitement à l'intérieur duquel on obtient une réponse thérapeutique sans la survenue d'effet indésirable. Comme démontré dans la première section de l'introduction avec l'exemple du variant AR-V7, la variabilité des voies PD peut affecter la réponse et donc servir de marqueur pharmacogénomique potentiel. Un autre exemple est celui de la réponse au 5-FU. Un polymorphisme de la région 5' non traduite du gène de la thymidylate synthase (TS), cible thérapeutique du 5-FU, consiste en la répétition d'un motif de 28 paires de bases 2, 3, 4, 5 ou 9 fois. Plus le nombre de répétitions est élevé, plus l'expression de la TS est élevée (Deenen et al., 2011; Horie et al., 1995; Kaneda et al., 1987). Chez les individus avec 3 répétitions, un variant génétique peut également diminuer l'expression du gène à un niveau comparable à celui des individus ayant 2 répétitions du motif de 28 paires de bases (Kawakami et al., 2003; Mandola et al., 2003). La plupart des études chez des patients traités avec du 5-FU ont associé cette expression plus élevée à une activité clinique réduite, c'est-à-dire moins d'effets indésirables mais aussi une réponse et une survie réduites (Deenen et al., 2011). Cependant, le manque de reproductibilité entre les études fait que ce marqueur n'est actuellement pas applicable en clinique.

2.3) Voies PK/PD de l'irinotécan

L'irinotécan est un traitement que nous avons étudié dans les chapitres I à III car il est pris en charge par la voie des UGT. Dans le cas du régime FOLFIRI, un des régimes standards de chimiothérapie pour le traitement de première ligne du CCRm, d'importantes différences de réponse au traitement sont observées entre les patients (**Tableau 3**). Par exemple, le taux de réponse varie aux environs de 40%. La survenue d'effets indésirables est aussi hétérogène entre les patients, comme c'est le cas avec les diarrhées et les neutropénies sévères de grades 3-4. Même si l'ajout de thérapies ciblées au régime FOLFIRI a pu améliorer la réponse, celle-ci n'en demeure pas moins variable avec des taux de réponse inférieurs à 60% dans trois essais cliniques de phase III avec les anticorps monoclonaux bévacizumab ou cétuximab (Cremolini et al., 2015; Heinemann et al., 2014; Van Cutsem et al., 2011).

Tableau 3. Réponse variable au régime FOLFIRI pour le traitement de 1^{ère} ligne du CCRm.

n =	Réponse ^b	TR	PFS médiane (mois)	OS médiane (mois)	D	N	Réf.
35	CR = 11%, PR = 20%, SD = 43%, PD = 26%	31%	7	17 (95% CI = 9-25)	9%	9%	(Teufel et al., 2004)
264	CR = 4%, PR = 33%, SD = 40%, PD = 23%	37%	7 (de 1 à 47)	14 (de 1 à 48)	9%	10%	(Maiello et al., 2005)
30 ^a	CR = 3%, PR = 33%, SD = 37%, PD = 27%	36%	7	15	17%	20%	(Souglakos et al., 2005)
122	CR = 6%, PR = 32%, SD = 34%, PD = 24%	38%	7	17	12%	28%	(Falcone et al., 2007)
146	CR = 3%, PR = 30%, SD = 27%, PD = 40%	33%	7 (de 1 à 39)	19 (de 1 à 56)	11%	28%	(Souglakos et al., 2006)
137	CR = 6%	47%	8	23	14%	43%	(Fuchs et al., 2007)
40 ^a	ND	45%	8	17 (95% CI = 12-22)	15%	8%	(Francois et al., 2008)
173	CR = 9%, PR = 48%, SD = 25%, PD = 21%	57%	8 (95% CI = 7-9)	22 (95% CI = 20-25)	12%	16%	(Aranda et al., 2009)
48	SD = 35%, PD = 13%	45%	8 (95% CI = 7-11)	19 (95% CI = 14-29)	2%	23%	(Santoro et al., 2008)
153	ND	ND	25 (de 19 à 39)	ND	14%	23%	(Ychou et al., 2009)
46	CR = 2%, PR = 46%, SD = 22%, PD = 24%	48%	9 (95% CI = 7-12)	17 (95% CI = 8-25)	6%	13%	(Skof et al., 2009)
116	ND	ND	8 (95% CI = 7-9)	24 (95% CI = 18-28)	15%	40%	(Jackson et al., 2009)
28	ND	ND	8 (95% CI = 6-10)	20 (95% CI = 13-30)	11%	54%	
52	CR = 2%, PR = 38%, SD = 52%, PD = 4%	40%	8 (95% CI = 6-9)	22 (95% CI = 18-30)	2%	44%	(Okuyama et al., 2011)
599	CR <1%, PR = 38%, SD = 47%, PD = 9%	39%	8 (95% CI = 8-9)	19 (95% CI = 17-20)	11%	25%	(Van Cutsem et al., 2011)

58	CR = 2%, PR = 35%, SD = 30%, PD = 33%	37%	9 (95% CI = 6-10)	18 (95% CI = 17-21)	ND	ND	(Silvestris et al., 2014)
79	CR = 1%, PR = 42%, SD = 35%, PD = 17%	43%	7 (95% CI = 6-9)	19 (95% CI = 16-22)	ND	ND	(Dong et al., 2015)
70 ^a	CR = 3%, PR = 32%, SD = 29%, PD = 19%	35%	8 (95% CI = 5-9)	11 (95% CI = 9-17)	ND	ND	(Aparicio et al., 2016)
39	CR + PR = 44%, SD = 44%, PD = 12%	44%	8	18	8%	21%	(Paez et al., 2019)

^acohortes de patients âgés (≥ 70 ans) ; ^bCR, réponse complète ; PR, réponse partielle ; SD, maladie stable ; PD, progression de la maladie. D, diarrhées sévères ; N, neutropénies sévères ; ND, non déterminé ou non rapporté ; OS, survie globale ; PFS, survie sans progression ; TR, taux de réponse.

Le métabolisme de l'irinotécan est principalement hépatique (Mathijssen et al., 2001). L'irinotécan est un pro-médicament qui est hydrolysé dans l'organisme par les carboxylestérases en SN-38, son métabolite actif (Gerrits et al., 1997; Kawato et al., 1991; Satoh et al., 1994). La volumineuse chaîne latérale de l'irinotécan est clivée par les carboxylestérases CES1 et CES2, principalement au niveau hépatique, avec une affinité et une vélocité supérieure pour CES2, afin de former le SN-38 (**Figure 3**) (Humerickhouse et al., 2000; Sanghani et al., 2004; Smith et al., 2006). Cependant, CES1 est environ 50 fois plus exprimée que CES2 au foie, et participe donc de façon non-négligeable à la conversion de l'irinotécan en SN-38 (Hatfield et al., 2011; Ross et al., 2012). L'irinotécan peut aussi être oxydé par CYP3A4 dans le foie en APC, un métabolite inactif qui ne peut plus être converti en SN-38 ou en NPC, un autre métabolite inactif, pouvant être converti en SN-38 par les carboxylestérases (**Figure 3**) (Dodds et al., 1998; Haaz et al., 1998; Rivory et al., 1996; Santos et al., 2000). D'autres métabolites (nommés M2, M3 et M4) ont été découverts et sont formés par l'action des CYP3A4 et/ou 3A5 sur l'irinotécan (Lokiec et al., 1996; Sai et al., 2001; Santos et al., 2000). Le SN-38 est principalement éliminé par glucuronidation au niveau hépatique par UGT1A1 mais aussi par UGT1A9 et d'autres UGT1A de façon moins importante (**Figure 3**) (Gagne et al., 2002; Hanioka et al., 2001; Xiao et al., 2018). Dans les tissus extra-hépatiques, UGT1A7 contribue également à l'inactivation du SN-38 (Ciotti et al., 1999; Gagne et al., 2002). En revanche, les β -glucuronidases bactériennes du tractus gastrointestinal sont capables de convertir le SN-38G (métabolite glucuronide inactif) en SN-38, le métabolite actif et toxique. Cette exposition du tractus gastrointestinal au SN-38 peut induire des toxicités sévères, notamment des diarrhées de grades 3-4.

L'irinotécan entre dans la cellule par diffusion alors que le métabolite actif SN-38 entre dans les cellules hépatiques principalement par le biais du transporteur OATP1B1 mais aussi OATP1B3 (Fujita et al., 2014; Nozawa et al., 2005). Le SN-38 est relargué dans la circulation sanguine depuis les cellules du foie par le transporteur MRP1 alors que l'irinotécan, le SN-38 et le SN-38G sont excrétés vers la bile par divers membres de la famille ABC incluant notamment MDR1, MRP2 et BCRP (Chen et al., 1999; de Jong et al., 2007; Luo et al., 2002a; Nakatomi et al., 2001; Tagen et al., 2010; Wang et al., 2016). Ainsi, c'est l'action combinée de plusieurs transporteurs qui régule l'exposition au pro-

médicament, au métabolite actif et aux métabolites inactifs avec probablement une contribution de la diffusion membranaire.

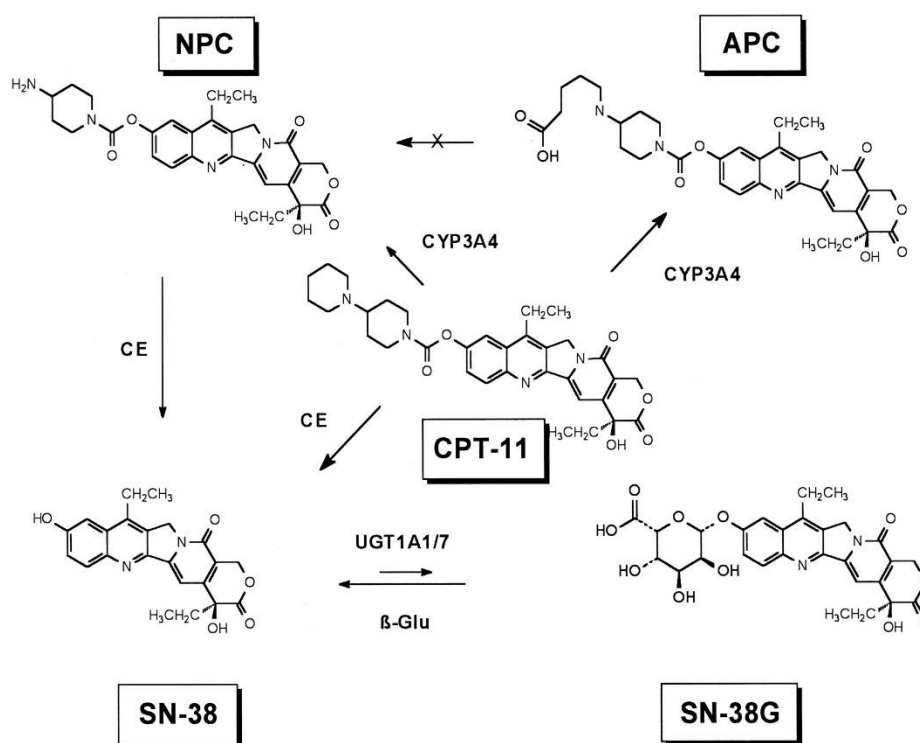


Figure 3. Voies PK de l'irinotécan.

L'irinotécan (CPT-11, 7-éthyl-10-[4-(1-piperidino)-1-piperidino] carbonyloxy camptothécine) est hydrolysé par les carboxylestérases (CE) en SN-38 (7-éthyl-10-hydroxycamptothécine, métabolite actif) qui est par la suite inactivé principalement par les UGT. Les métabolites NPC et APC sont obtenus par action de CYP3A4 sur l'irinotécan. Les β -glucuronidases bactériennes (β -Glu) du tractus gastrointestinal peuvent former du SN-38 à partir de son glucuronide. Figure issue de Mathijssen et al., 2001. Reproduit avec la permission du *Copyright Clearance Center* pour *American Association for Cancer Research*.

Le SN-38 est un inhibiteur de la topoisomérase I, impliquée dans la réplication mais aussi la réparation de l'ADN (Wang, 2002). Cette enzyme effectue une coupure simple brin de l'ADN ce qui entraîne sa relaxation pour permettre la réplication, puis elle répare cette coupure. Le SN-38 se lie au complexe de clivage TOP1-ADN et le stabilise, conduisant à une collision avec la fourche de réplication. Cela produit une coupure double brin irréversible de l'ADN, et l'activation de voies du stress cellulaire et de l'arrêt du cycle cellulaire en phase S, aboutissant à la mort de la cellule (Beretta et al., 2008). C'est un

mécanisme d'action qui est donc particulièrement efficace sur les cellules se répliquant rapidement, comme c'est le cas des cellules cancéreuses.

3) La voie de glucuronidation : fonctions, organisation, régulation et variabilité inter-individuelle

Parmi les voies PK influençant la réponse au traitement, la voie de glucuronidation a le potentiel de jouer un rôle majeur car elle prend en charge un grand nombre de médicaments actuellement sur le marché. Son rôle essentiel est aussi soutenu par l'impact que peut avoir sa variabilité génétique. Par exemple le polymorphisme *UGT1A1*28* est responsable à lui seul de 15% de l'augmentation dans l'exposition au SN-38, avec des conséquences cliniques majeures, telles que décrites précédemment.

3.1) Fonctions de la voie de glucuronidation

Les UGT conjuguent une vaste gamme de substrats lipophiles. Cette réaction, nommée glucuronidation, est une réaction de conjugaison à un sucre qui a lieu au réticulum endoplasmique et contribue à la phase II du métabolisme des médicaments (**Figure 4A**) (Radomska-Pandya et al., 1999). Un groupement polaire est ajouté à une molécule lipophile dans le but d'augmenter sa solubilité dans l'eau et de faciliter son élimination de l'organisme (Rowland et al., 2013). La réaction catalytique consiste au transfert du sucre de l'UDP-GlcA, qui est le cosubstrat ubiquitaire des UGT, sur le substrat. Le mécanisme réactionnel de la glucuronidation est une substitution nucléophile de type 2 (**Figure 4B**). Le substrat est activé par déprotonation grâce à une base et peut donc effectuer une attaque nucléophile sur le carbone 1 de l'UDP-GlcA aboutissant à la formation d'un glucuronide avec libération d'un UDP qui pourra être recyclé en UDP-GlcA (Guillemette et al., 2014). Les substrats sont glucuronidés le plus souvent sur un groupement réactif carboxyle, carbonyle, sulfuryle, hydroxyle ou amine (**Figure 4C**). Puisque l'enzyme est située dans la lumière du réticulum endoplasmique, le substrat et le cosubstrat doivent y parvenir par le biais d'un transporteur ou par diffusion passive (Battaglia et al., 1996; Goto et al., 2001; Kobayashi et al., 2006; Rowland et al., 2015).

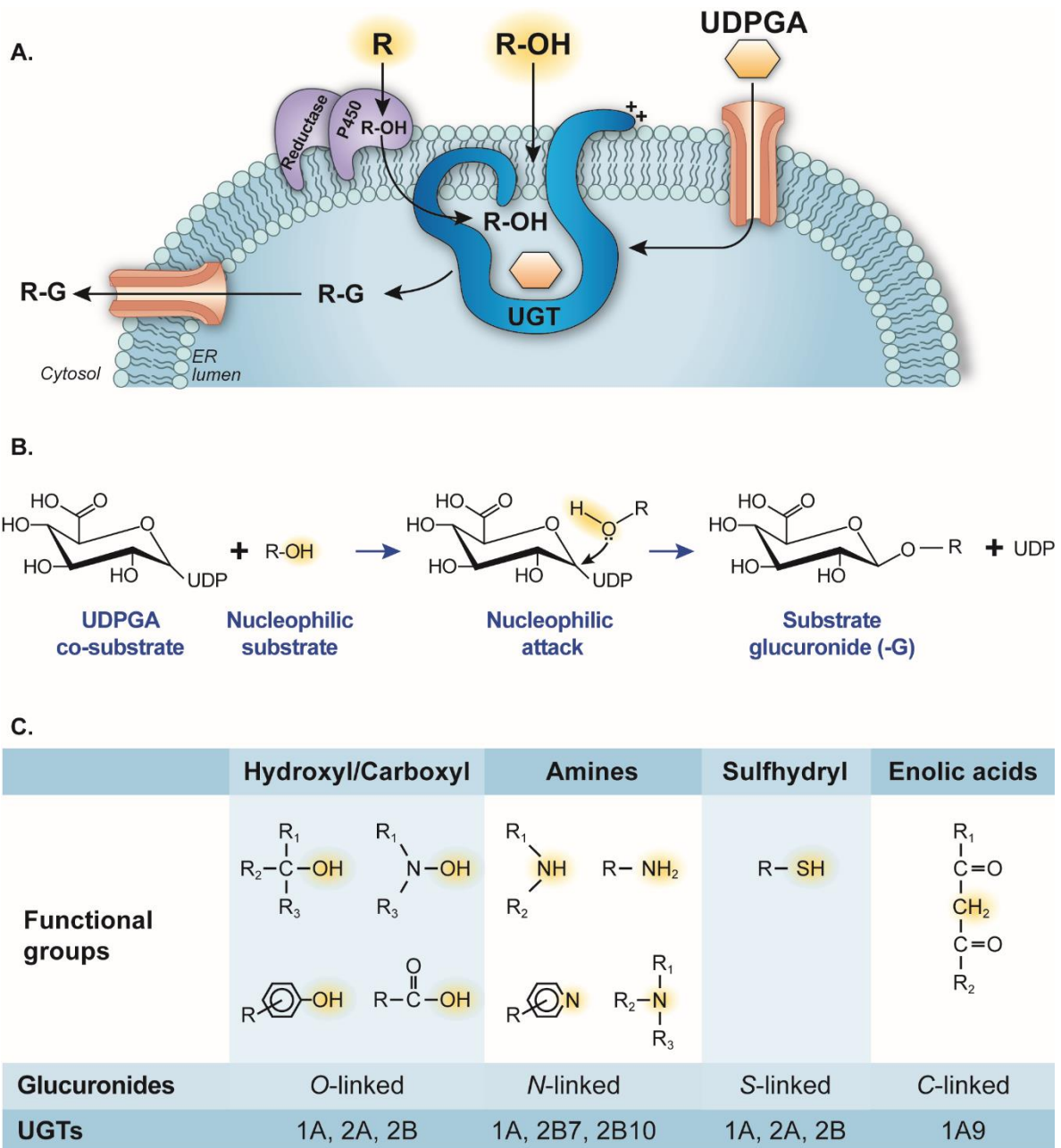


Figure 4. La réaction de glucuronidation.

A. Les enzymes UGT sont localisées au réticulum endoplasmique et catalysent la conjugaison du sucre du cosubstrat UDP-GlcA au substrat. **B.** La glucuronidation est une attaque nucléophile de type 2 résultant en la libération d'un glucuronide et d'UDP. **C.** Différents types de groupements fonctionnels peuvent être glucuronidés. Figure issue de Guillemette et al., 2014. Reproduit avec la permission du *Copyright Clearance Center* pour *John Wiley and Sons*.

Généralement, la réaction prédominante consiste en la O-glucuronidation ou la N-glucuronidation d'un alcool aliphatique, d'un phénol, d'un acide carboxylique, d'un thiol ou d'une amine. L'ajout du sucre rend la molécule plus hydrosoluble facilitant donc son excrétion via la bile et l'urine. Cela augmente aussi son encombrement stérique, diminuant son activité et sa toxicité en l'empêchant d'agir sur un récepteur par exemple. D'une façon générale, la glucuronidation réduit l'activité et l'effet biologique des composés qui y sont soumis, sauf dans quelques cas comme celui de la morphine où un important effet physiologique est attribuable à la morphine-6-glucuronide (Frances et al., 1992; van Dorp et al., 2008). C'est aussi le cas pour l'acide *all-trans* rétinol dont le glucuronide a une activité biologique (Barua et al., 2004; Formelli et al., 1996; Miller et al., 1986).

3.2) Organisation de la voie de glucuronidation

Les 19 enzymes UGT sont réparties dans deux grandes familles (**Figure 5, haut et bas**), UGT1A et UGT2, cette dernière étant subdivisée en deux sous-familles UGT2A et UGT2B (Guillemette et al., 2010; Guillemette et al., 2014). Le locus unique du gène *UGT1* situé sur le chromosome 2q37.1 code pour les neuf enzymes UGT1A (1A1, 1A3-1A10) fonctionnelles ainsi que quatre pseudogènes (1A2p, 1A11p-1A13p) (Gong et al., 2001). Les exons 2 à 5 sont communs (Gong et al., 2001; Ritter et al., 1992). L'utilisation de promoteurs alternatifs permet une expression tissulaire et cellulaire spécifique de chaque isoforme. De la même façon, chacun des exons 1 alternatifs confère une spécificité de substrat à chaque isoforme UGT1A. Les UGT2 sont toutes codées par six exons mais proviennent de gènes distincts tous situés au chromosome 4q13. UGT2A1 et 2A2 sont codées par le même gène selon un mécanisme de promoteurs alternatifs similaire aux UGT1A alors qu'UGT2A3 résulte de l'expression d'un gène distinct (Court et al., 2008; Mackenzie et al., 2005; Sneitz et al., 2009). C'est également le cas pour chacune des 7 enzymes UGT2B4, 2B7, 2B10, 2B11, 2B15, 2B17, 2B28. Ces gènes semblent résulter d'évènements successifs de duplication génique et partagent une grande identité de séquence, ce qui conduit à une similarité pouvant atteindre 97% entre les séquences codantes d'*UGT2B15* et d'*UGT2B17* (Belanger et al., 1998; Turgeon et al., 2000). Plusieurs pseudogènes répartis entre les gènes *UGT2B* témoignent également de ces évènements de duplication. En plus des UGT1A et UGT2, il existe deux UGT3A (3A1 et

3A2) et une UGT8A (8A1) aux rôles moins étudiés et plus modestes quant au métabolisme des médicaments (Rowland et al., 2013). Les UGT3A n'utilisent pas l'UDP-GlcA comme cosubstrat mais l'UDP-N-acétylglucosamine (3A1) et l'UDP-glucose ou l'UDP-xylose (3A2) (MacKenzie et al., 2011; Mackenzie et al., 2008). L'UGT8A1 est une céramide glycosyltransférase impliquée dans la synthèse des glycosphingolipides (Bosio et al., 1996; Ichikawa et al., 1996).

Au niveau protéique, la structure d'une UGT s'articule autour d'une région variable en N-terminal et d'une région conservée en C-terminal (**Figure 5, centre**). A l'extrémité N-terminale, on retrouve un peptide signal qui est clivé une fois la protéine correctement adressée au réticulum endoplasmique (Meech et al., 1998; Teasdale et al., 1996). La région variable correspond au domaine spécifique de liaison au substrat et est codée par l'exon 1 pour les UGT1A et l'exon 1 et 2 pour les UGT2. Au centre de la séquence peptidique de cette zone de spécificité au substrat se trouve une région d'ancrage à la membrane. La partie conservée C-terminale des UGT comprend un site de liaison spécifique du cosubstrat UDP-GlcA, au sein duquel apparaît une séquence signature « glycosyltransférase » (Ciotti et al., 1998). Un domaine transmembranaire et un motif dilysine cytosolique se trouvent à l'extrémité C-terminale et sont communs aux enzymes UGT.

Au niveau subcellulaire, les UGT sont localisées dans la lumière du réticulum endoplasmique et sont fixées à sa membrane grâce au domaine transmembranaire C-terminal et au domaine d'ancrage N-terminal. Seule l'extrémité C-terminale et son motif dilysine sont cytosoliques. La structure complète des UGT n'est pas connue malgré quelques études partielles sur la région C-terminale des UGT1A et d'UGT2B7 (Miley et al., 2007; Osborne et al., 2018; Radomska-Pandya et al., 2010). Seules des études de modélisation permettent de prédire la structure des UGT (Laakkonen et al., 2010; Radomska-Pandya et al., 1999). L'étude des séquences d'acides aminés en lien avec la structure des UGT pourraient permettre de comprendre la spécificité de substrat de chaque UGT et de développer des inhibiteurs spécifiques (Kerdpin et al., 2009; Osborne et al., 2019; Wu et al., 2011).

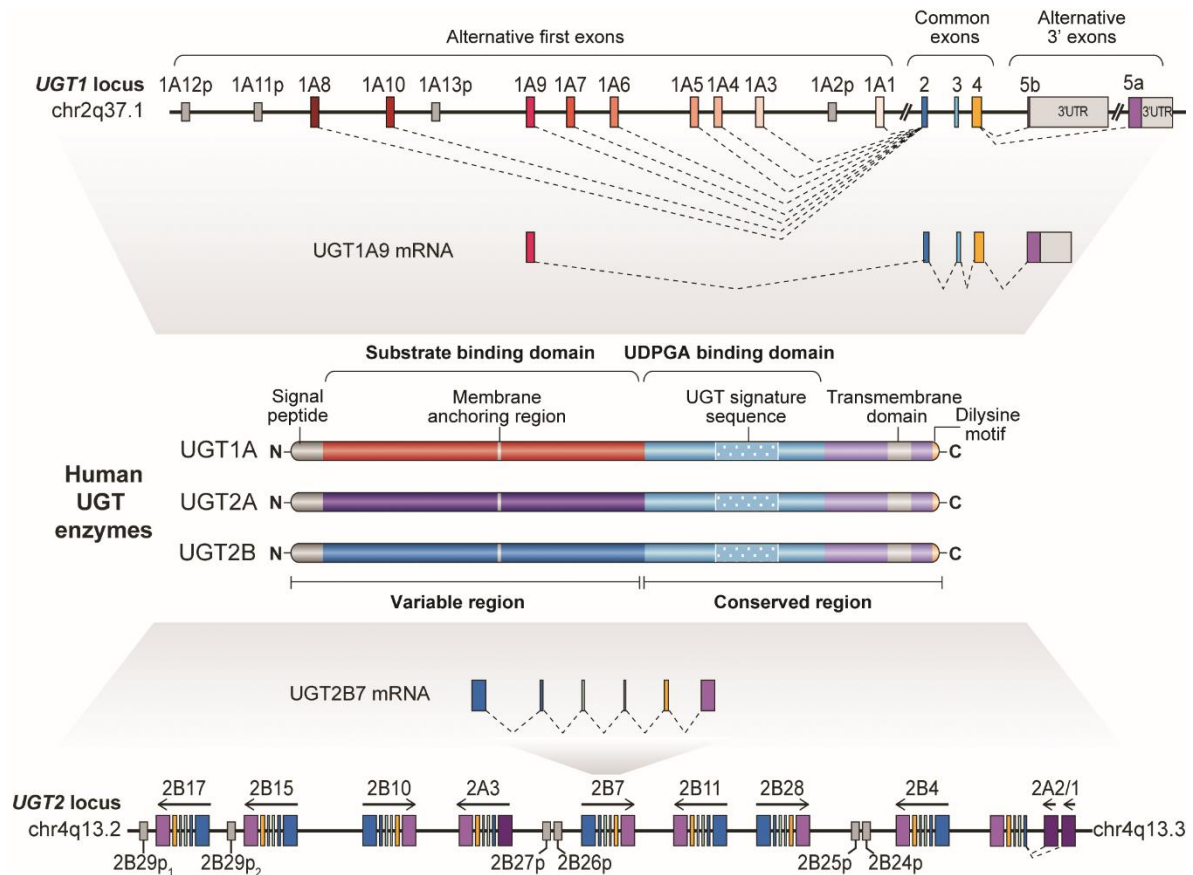


Figure 5. Organisation génique et protéique des UGT.

Le locus *UGT1* code pour les 9 UGT1A grâce à l'usage de promoteurs alternatifs et d'exons 1 spécifiques à chaque isoforme couplés aux exons 2 à 5 communs (haut). Les UGT2A1 et 2A2 sont produites sur le même modèle alors qu'UGT2A3 et toutes les UGT2B proviennent de gènes différents (bas). Les loci *UGT1* et *UGT2* sont parsemés de pseudogènes résultants d'événements de duplication. Les enzymes UGT1A, UGT2A et UGT2B ont une structure similaire avec un domaine variable de liaison au substrat et d'ancrage à la membrane en N-terminal et un domaine constant de liaison au cosubstrat, un domaine transmembranaire et un motif dilysine en C-terminal (centre). Figure issue de Guillemette et al., 2014. Reproduit avec la permission du *Copyright Clearance Center* pour *John Wiley and Sons*.

Les UGT prennent en charge 55% des 200 médicaments les plus prescrits et environ 35% de tous ceux actuellement utilisés, en plus de réguler la concentration de multiples molécules endogènes (**Tableau 4**) (Bock, 2015; Guillemette et al., 2014; Turgeon et al., 2003). Les différentes UGT sont actives envers divers substrats de façon spécifique ou redondante. Par exemple, la bilirubine est exclusivement glucuronidée par UGT1A1, alors que le SN-38 est pris en charge par UGT1A1 et 1A9 au foie et 1A7 et 1A10 dans les tissus extra-hépatiques. Ainsi, il apparaît que les UGT présentent un profil d'expression tissulaire et cellulaire distinct (Court et al., 2012; Ohno et al., 2009). Les UGT se retrouvent dans de nombreux tissus mais c'est dans le foie que leur abondance est la plus importante (Guillemette et al., 2014). Cependant, UGT1A5, 1A7 et 1A10 sont exclusivement extra-hépatiques (Strassburg et al., 1997). Les UGT sont également retrouvées dans d'autres tissus du métabolisme des médicaments tels que l'estomac, l'intestin, le côlon et les reins. Puisqu'elles régulent l'homéostasie hormonale, il n'est pas surprenant de les voir exprimées dans les ovaires, les testicules, le sein et la prostate. Enfin, leur présence dans la peau, les poumons et le tractus gastrointestinal, qui sont les premières barrières de l'organisme lors de l'exposition aux xénobiotiques, a été établie. Il est cependant important de noter que leurs niveaux d'expression en ARN diffèrent selon le tissu (Court et al., 2012; Nakamura et al., 2008; Ohno et al., 2009; Ohno et al., 2011).

Tableau 4. Les UGT catalysent la glucuronidation de multiples substrats endogènes et exogènes incluant des médicaments.

UGT	Substrats endogènes	Substrats exogènes
1A1	Bilirubine*, estradiol, T4	Étoposide, acétaminophène, éthinylestradiol, SN-38, simvastatine, codéine, ézétimibe, atorvastatine
1A3	Acide chénodésoxycholique, estrone, T4	Telmisartan, simvastatine, ézétimibe, atorvastatine
1A4	Androstanédiol, prégnanediol	Amitriptyline, midazolam, imipramine, kétotifène, lamotrigine, tacrolimus*, tamoxifène, trifluopérazine*
1A5	--	1-hydroxypyrene
1A6	Sérotonine*, 5-OH-tryptophol	Défériprone, acétaminophène
1A7	--	Coumarine, thymol, SN-38
1A8	Estrone	Acide mycophénolique
1A9	Estrogènes, T4, 20-HETE	Diclofénac, flavopiridol, acide mycophénolique, acétaminophène, propofol*, sorafénib, entacapone, R-oxazépam, édaravone, tolcapone, SN-38
1A10	Hydroxystéroïdes, dopamine	Apigénine, SN-38
2B4	Acides biliaires	Carvédilol, défériprone, fénofibrate, acétaminophène
2B7	Hormones stéroïdes, acide <i>all-trans</i> rétinoïque, 20-HETE	Carvédilol, naloxone, chloramphénicol, épirubicine, fénofibrate, morphine, codéine, acide mycophénolique, naproxène, zidovudine*, éfavirenz, anti-inflammatoires non stéroïdiens
2B10	12-HETE, 15-HETE	Diphénhydramine, nicotine, olanzapine, amitriptyline, midazolam, kétotifène, tamoxifène, lévométdomidine
2B11	12-HETE, 15-HETE	--
2B15	Testostérone	Lorazépam, S-oxazépam, acétaminophène, tamoxifène, tolcapone, dabigatran, R-méthadone
2B17	Dihydrotestostérone, testostérone	Vorinostat
2B28	Estradiol, testostérone	--

Adapté de Bock 2015, Guillemette et al., 2014, Kato et al., 2013. Ce tableau présente des exemples de substrats endogènes et exogènes pour chaque UGT mais n'est pas exhaustif. HETE, acide hydroxyeicosatétraénoïque ; SN-38, métabolite actif de l'irinotécan ; T4, thyroxine. Les substrats marqués d'un astérisque (*) sont des substrats spécifiques à l'enzyme UGT mentionnée.

3.3) Régulation de l'expression de la voie de glucuronidation

L'expression de chaque isoforme UGT est tissu-spécifique. Elle est régulée pour répondre à la fois au maintien de l'homéostasie cellulaire en modulant les niveaux de molécules endogènes et à la défense de l'organisme contre de multiples xénobiotiques. La régulation transcriptionnelle de ces isoenzymes est spécifique à chacune. Malgré tout, et puisque certaines UGT sont coexprimées au sein d'un même tissu ou en réponse à un même substrat, il existe aussi une redondance dans les mécanismes de régulation pour plusieurs UGT (Hu et al., 2014). Ainsi, l'expression constitutive dans chaque tissu est médiée par des facteurs de transcription spécifiques au tissu considéré alors que l'induction en réponse à l'exposition aux composés endogènes ou exogènes est régulée par des récepteurs nucléaires dont ils sont les ligands (**Tableau 5**). Par exemple, UGT1A1 est l'unique isoforme responsable de la conjugaison de la bilirubine et la principale enzyme contribuant à l'inactivation du SN-38. Elle est exprimée au foie mais aussi dans plusieurs tissus extra-hépatiques incluant l'estomac, les reins, l'intestin et le côlon. L'expression constitutive d'*UGT1A1* est notamment liée à la présence dans la région promotrice de site de liaison pour les facteurs de transcription HNF1A, enrichi au foie, et USF1/2, exprimés de façon plus ubiquitaire, en plus du site de la boîte TATA (Belanger et al., 2010; Bernard et al., 1999). Cependant cette expression constitutive d'*UGT1A1* peut être réprimée par exemple via la méthylation des histones dans les tissus hépatiques fœtaux ou dans le rein par méthylation de l'ADN (Nie et al., 2017; Oda et al., 2013). L'expression d'*UGT1A1* est inductible en réponse à des composés endogènes et exogènes, par le biais d'un site de liaison PBREM (*phenobarbital-responsive enhancer module*) dans le promoteur qui peut être activé par divers récepteurs nucléaires incluant le récepteur AhR (**Tableau 5**). La présence de bilirubine, ligand endogène du récepteur AhR, va donc induire l'expression d'*UGT1A1* via entre autres ce récepteur, dans le but d'éliminer de l'organisme ce déchet métabolique toxique (Togawa et al., 2008). L'activation du récepteur CAR résulte aussi en une augmentation de la clairance de la bilirubine, par le biais d'une hausse de l'expression d'*UGT1A1* et d'autres gènes impliqués dans son élimination (Huang et al., 2003). PXR et CAR sont impliqués dans l'expression de plusieurs CYP, UGT et transporteurs, démontrant une régulation coordonnée des voies PK (Chai et al., 2013; Chen et al., 2012; Wallace et al., 2013). Par exemple, PXR peut être activé par la rifampicine, et induire l'expression

d'enzymes comme CYP3A4 ou UGT1A1 (Bertilsson et al., 1998; Lehmann et al., 1998; Luo et al., 2002b; Moscovitz et al., 2018; Sugatani et al., 2012; Wang et al., 2014). Ils sont aussi capables d'interagir avec des facteurs de transcription comme les HNF. Ainsi, PXR et HNF4 sont impliqués dans la régulation de CYP3A4 et CYP7A1, qui jouent respectivement un rôle dans le métabolisme des médicaments et dans la synthèse des acides biliaires (Li et al., 2005; Lim et al., 2009).

D'autre part, le récepteur PXR est soumis à l'épissage alternatif et l'expression de 3 isoformes (T1-3) a été démontrée (Bertilsson et al., 1998; Blumberg et al., 1998; Dotzlaw et al., 1999; Lehmann et al., 1998). Les isoformes T1 et T2 sont capables d'induire l'expression d'UGT1A1 mais pas l'isoforme T3 (Gardner-Stephen et al., 2004). L'expression des variants PXR est très différente entre les individus, et cela représente donc une source potentielle de modulation de l'expression de voies du métabolisme des médicaments qui sont régulées par PXR, incluant les CYP et UGT (Gardner-Stephen et al., 2004). Une autre étude a démontré une importante variabilité de l'expression de PXR au niveau du côlon chez 14 individus, à la fois dans les tissus sains et tumoraux (Raynal et al., 2010). Des expériences *in vitro* ont aussi démontré qu'une augmentation de l'expression de PXR dans des cellules de cancer du côlon réduisait la sensibilité au SN-38, à cause d'une induction de l'expression des UGT1A responsables de sa glucuronidation. Ces exemples suggèrent que la variabilité génétique des régulateurs transcriptionnels pourrait aussi avoir des répercussions cliniques sur la réponse à divers traitements, en modifiant la régulation des voies PK, et donc servir de marqueurs prédictifs.

Tableau 5. L'expression des UGT peut être induite par divers récepteurs nucléaires activés par des ligands endogènes et exogènes.

Récepteurs nucléaires	Ligands endogènes	Ligands exogènes	UGTs ciblées
AhR	bilirubine 12-HETE	3-Méthylcholanthrène benzopyrène, TCDD	1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A10, 2B4
AR	testostérone dihydrotestostérone	flutamide méthyltriénolone	1A1, 1A3, 2B10, 2B11, 2B15, 2B17, 2B28
CAR	3 α ,5 α -androsténol 3 α ,5 α -androstanol	artémisinine TCPOBOP	1A1, 2B7
ER α	17- β estradiol	tamoxifène	1A4, 2B15, 2B17
FXR	acides cholique, désoycholique, et chénodésoycholique	Z-guggulstérone	1A3, 2B4, 2B7, 2B10
GR	cortisone	dexaméthasone	1A1
LXR	24S-hydroxycholestérol		1A3
Nrf2	bilirubine	α -naphthoflavone olitpraz	1A1, 1A6, 1A7, 1A8, 1A10, 2B7
PPAR α	20-HETE 11,12-EET	acide clofibrrique, GW7647 gemfibrozil, bézafibrate	1A3, 1A6, 1A9, 2 B4
PPAR γ	15-désoxy-D12,14- prostaglandin-J2	ciglitazone, troglitazone GW9662	1A9
PXR	acide lithocholique	rifampicine, hyperforine	1A1
VDR	25-(OH)vitamine D3 1,25-(OH)vitamine D3		2B15, 2B17

Adapté de Hu et al., 2014, Bock 2015 et Lu et al., 2017. Ce tableau présente des exemples mais n'est pas exhaustif. AhR, récepteur aux hydrocarbures aromatiques ; AR, récepteur des androgènes ; CAR, récepteur constitutif des androstanes ; ER α , récepteur aux œstrogènes alpha ; FXR, récepteur X des farnésoïdes ; GR, récepteur des glucocorticoïdes ; LXR, récepteurs des oxystérols ; Nrf2, *nuclear factor erythroid-2-related factor 2* ; PPAR α , récepteur activé par les proliférateurs de peroxyosomes alpha ; PPAR γ , récepteur activé par les proliférateurs de peroxyosomes gamma ; PXR, *pregnane X receptor* ; VDR, récepteur de la vitamine D. EET, acide époxyeicosatriénoïque ; HETE, acide hydroxyeicosatétraénoïque ; TCDD, 2,3,7,8-tétrachlorodibenzo-p-dioxine ; TCPOBOP, 1,4-bis[2-(3,5-dichloropyridyloxy)]benzène.

3.4) Variabilité interindividuelle d'expression et d'activité de la voie de glucuronidation

Au niveau du foie, les UGT les plus exprimées en ARN, avec une forte hétérogénéité entre les sujets, sont UGT2B4 (35% du transcriptome *UGT*, différence jusqu'à 72 fois entre les individus), UGT2B10 (20%, 223 fois), UGT1A1 (11%, 9 fois), UGT2B15 (8%, 29 fois), UGT1A6 (7%, 22 fois), UGT1A4 (6%, 28 fois), UGT2B7 (5%, 506 fois) et UGT1A9 (5%, 45 fois) (Izukawa et al., 2009). Cela suggère que des mécanismes variés peuvent moduler fortement l'expression des gènes *UGT*. Une étude récente du laboratoire basée sur du séquençage d'ARN à haut débit corrobore ces résultats sauf pour UGT2B7 qui est la plus exprimée (UGT2B7 < 2B4 < 2B10 < 2B15 < 1A4 < 1A1 < 1A9 < 1A6) comme représenté dans la **Figure 6** (Tourancheau et al., 2018).

Il a aussi été montré que les niveaux entre ARN et protéine ne corrèlent pas ou peu, comme c'est le cas au foie pour UGT1A1 (coefficient de régression linéaire $r < 0,10$), UGT1A6 ($r = 0,62$), UGT1A9 ($r = 0,31$), UGT2B7 ($r = 0,28$) et UGT2B15 ($r = 0,25$) (Ohtsuki et al., 2012). Cette absence de corrélation importante entre ARN et protéine suggère que des mécanismes post-transcriptionnels différents entre les individus sont impliqués dans la régulation de l'expression des UGT. En outre, UGT2B10 est une des UGT les plus exprimées en ARN au niveau du foie mais seulement la huitième en protéine (3% des protéines UGT) (Margaillan et al., 2015b; Tourancheau et al., 2018). Plusieurs études ont quantifié l'expression hépatique des protéines UGT et ont également mis en avant une forte fluctuation inter-individuelle d'expression (**Tableau 6**).

Tableau 6. L'expression protéique des UGT au niveau hépatique est très variable entre individus.

UGT	Coefficient de variation (%)						Ratio max/min
	n = 48	n = 17	n = 9	n = 16	n = 25	n ≥ 23	n = 60
1A1	73 (2) ^a	35 (4)	43 (2)	51 (2)	18 (3)	101 (8)	23 (4)
1A3	89 (9)	45 (6)	75 (3)	100 (9)	5 (4)	99 (1)	13 (7)
1A4	47 (4)	ND	45 (5)	36 (5)	NM	43 (6)	4 (2)
1A6	52 (7)	47 (1)	43 (4)	56 (9)	NM	75 (2)	7 (8)
1A9	65 (6)	24 (5)	96 (1)	36 (7)	NM	59 (7)	4 (6)
2B4	45 (3)	NM	NM	47 (3)	NM	46 (4)	5 (3)
2B7	47 (1)	35 (2)	NM	34 (1)	7 (2)	47 (3)	6 (1)
2B10	124 (8)	NM	NM	46 (6)	NM	NM	25 (8)
2B15	47 (5)	35 (3)	NM	43 (4)	2 (1)	51 (5)	5 (5)
2B17	184 (10)	NM	NM	112 (8)	NM	NM	^b (10)
Réf.	(Margaillan et al., 2015b)	(Ohtsuki et al., 2012)	(Harbourt et al., 2012)	(Sato et al., 2014)	(Groer et al., 2014)	(Achour et al., 2014b)	(Fallon et al., 2013)

ND, UGT non détectée ou sous la limite de détection, NM, UGT non mesurée ; ^ala valeur entre parenthèse correspond au rang de l'UGT selon son expression ; ^bpour UGT2B17, il n'est pas possible de calculer de ratio entre la quantité maximale détectée chez un individu (50 pmol/mg de protéine) et la quantité minimale détectée (0 pmol/mg de protéine).

L'activité hépatique des enzymes UGT chez 39 individus corrèle très significativement avec leur expression microsomale en protéine comme c'est le cas pour UGT1A1 avec la bilirubine (coefficient de corrélation de Spearman $r_s = 0,88$) et l'estradiol (groupement 3-hydroxyle, $r_s = 0,93$), UGT1A4 avec le tacrolimus ($r_s = 0,87$), UGT1A6 avec le déféripone ($r_s = 0,91$), UGT1A9 avec le propofol ($r_s = 0,72$) et UGT2B7 avec la zidovudine ($r_s = 0,82$) et l'estradiol (groupement 17-hydroxyle, $r_s = 0,84$). Néanmoins, la variabilité interindividuelle de glucuronidation de ces substrats est importante pour UGT1A1 avec la bilirubine (CV = 61%) et l'estradiol (CV = 79%), UGT1A4 avec le tacrolimus (CV = 48%), UGT1A6 avec le déféripone (CV = 49%), UGT1A6 avec le propofol (CV = 48%) et UGT2B7 avec la zidovudine (CV = 49%) et l'estradiol (CV = 50%) (Margaillan et al., 2015b). Une autre étude sur 59 sujets a démontré une bonne corrélation entre le niveau d'expression hépatique en protéine et l'activité, mais souligne aussi une importante variabilité pour UGT1A1 et l'estradiol (CV = 72%), UGT1A3 et l'acide chénodésoxycholique (CV = 40%), UGT1A4 et la trifluopérazine (CV = 37%), UGT1A6 et

le 5-Hydroxytryptophol (CV = 30%), UGT1A9 et le propofol (CV = 36%), UGT2B7 et la zidovudine (CV = 30%) et UGT2B15 et le S-oxazépam (CV = 50%) (Achour et al., 2017).

Au rein, la variabilité en ARN chez 11 donneurs est également élevée avec des CV de 159% pour UGT1A9 et 118% pour UGT2B7 qui sont les deux UGT les plus exprimées. C'est aussi le cas en protéine pour UGT1A9 (CV = 81%) et UGT2B7 (CV = 72%) (Margaillan et al., 2015a). Pour l'intestin, les CV s'étendent de 13 à 63% pour les enzymes UGT1A (Harbourt et al., 2012). Une seconde étude sur 6 donneurs confirme cette tendance avec des CV de 54% pour UGT1A1, 57% pour UGT1A3, 57% pour UGT1A4, et 39% pour UGT1A10 alors qu'ils sont de 63% pour UGT2B7 et 33% pour UGT2B17 (Sato et al., 2014).

En résumé, la voie de glucuronidation est une voie majeure du métabolisme de phase II des médicaments en plus de contribuer à l'homéostasie de nombreux composés endogènes. L'expression en ARN et en protéine des UGT est très différente entre les individus tout comme l'activité enzymatique. L'identification et la quantification de la contribution des mécanismes génétiques, épigénétiques, post-transcriptionnels et post-traductionnels à cette variabilité interindividuelle et à l'exposition à divers médicaments permettrait le développement de potentiels biomarqueurs pharmacogénomiques.

3.5) Sources de variabilité interindividuelle de la voie de glucuronidation et réponse au traitement

Il a été montré que certains facteurs physiologiques comme l'âge ou le sexe ou des facteurs environnementaux peuvent expliquer en partie, mais pas totalement, l'hétérogénéité dans l'expression et l'activité de la voie de glucuronidation (Fujita et al., 2016; Gallagher et al., 2010; Strassburg et al., 2002). Des mécanismes post-traductionnels comme la glycosylation ou la phosphorylation affectent *in vitro* l'activité des UGT mais leur impact *in vivo* reste largement à être étudié (Barbier et al., 2000; Basu et al., 2008). De la même façon, les interactions protéiques des UGT entre elles ou avec d'autres protéines pourraient influencer leur activité (Fujiwara et al., 2016). Ces mécanismes ne sont cependant pas l'objet des travaux présentés dans cette thèse. Cette voie compte aussi de nombreux polymorphismes et autres variations génétiques, en plus de subir un important épissage alternatif. Afin

d'envisager l'utilisation de ces variations comme marqueurs prédictifs de la réponse à divers médicaments pris en charge par les UGT, l'établissement d'un lien avec l'expression et l'activité de la voie de glucuronidation ainsi que l'exposition et la réponse au traitement est essentiel.

Des processus génétiques et épigénétiques affectent la voie de glucuronidation. Les polymorphismes génétiques d'une seule paire de bases constituent une des sources principales de divergence interindividuelle de cette voie avec plusieurs centaines de variations répertoriées (**Tableau 7**). Bien que le marqueur *UGT1A1**28 ait été le variant le plus étudié en lien avec l'exposition au SN-38, d'autres variations génétiques des UGT ont été associées à l'exposition à de multiples médicaments (**Tableau 8**). Ces variations pourraient donc être associées à la réponse au traitement. C'est aussi le cas pour les UGT2B, avec par exemple le rs61750900 d'*UGT2B10* conduisant à la substitution Asp67Tyr et qui abolit presque totalement l'activité de l'enzyme envers ses substrats tels que la nicotine et la cotinine (Berg et al., 2010; Chen et al., 2007). Un autre variant, le rs2942857 altère un site d'épissage entre l'intron 2 et l'exon 3 d'*UGT2B10* et il en résulterait la formation d'un ARN messenger instable expliquant la forte diminution d'activité UGT2B10 chez les porteurs (Fowler et al., 2015; Murphy et al., 2014).

Tableau 7. Variations génétiques affectant les gènes *UGT*.

UGT	variant ^a	variants d'EA			CNV	
		n ^b	% du transcriptome ^c			
			foie	rein		intestin/côlon
1A1		5	10	ND	12	non
1A3		4	5	ND	16	non
1A4		5	9	ND	46	non
1A5		2	ND	ND	22	non
1A6	997 ^d	6	35	34	56	non
1A7		3	ND	ND	ND	non
1A8		3	ND	ND	61	non
1A9		3	25	8	ND	non
1A10		8	ND	ND	23	non
2B4	289	15	16	ND	ND	non
2B7	439	42	8	33	9	non
2B10	199	10	62	ND	ND	non
2B11	161	9	ND	ND	ND	non
2B15	143	8	6	ND	100	non
2B17	252	9	37	ND	28	oui
2B28	73	6	ND	ND	ND	oui

^aDonnées issues de www.ncbi.nlm.nih.gov/variation/view/; *GRCh38p.12 annotation release 109*, consulté le 20 février 2019. ^bDonnées issues de Tourancheau et al., 2016. ^cDonnées issues de Tourancheau et al., 2018. ^d997 variants sont répertoriés au locus *UGT1*. ND, non détecté ou sous la limite de quantification. EA, épissage alternatif ; CNV, variation du nombre de copies.

D'autre part, des variations du nombre de copies (CNV) sont observées pour les gènes *UGT2B17* et *UGT2B28* et sont parmi les CNV les plus fréquemment observées du génome malgré des différences selon les groupes ethniques (Jakobsson et al., 2008; Locke et al., 2006; McCarroll et al., 2006; Menard et al., 2009; Xue et al., 2008). Chez les Caucasiens, 44% des individus possèdent 2 copies d'*UGT2B17*, alors que 47% ne possèdent qu'une seule copie et 9% n'en possèdent aucune. Chez les Asiatiques, plus de 70% des individus ne possèdent aucune copie du gène *UGT2B17*, mais cette proportion est seulement de 4% chez les Africains (Xue et al., 2008). La délétion d'*UGT2B28* est plus rare mais 13% des Caucasiens ont au moins une délétion tandis que 2% d'entre eux ne possèdent aucune copie

du gène. Ces deux UGT sont impliquées dans le métabolisme des hormones stéroïdiennes, notamment des androgènes et des estrogènes (Beaulieu et al., 1996; Belledant et al., 2016; Levesque et al., 2001). Les CNV de ces gènes ont été associées au risque de récurrence biochimique du cancer de la prostate (Nadeau et al., 2011). Le nombre de copies d'*UGT2B17* est associé au risque de cancer de la prostate (Cai et al., 2012; Karypidis et al., 2008; Park et al., 2006).

Tableau 8. Exemples de polymorphismes des gènes *UGT* qui affectent l'exposition au médicament.

Variant (conséquence)	Médicament	Effet fonctionnel	Référence
<i>UGT1A1</i> *6 (G71R)	irinotécan	↑ exposition, ↓ glucuronidation	(Jada et al., 2007)
<i>UGT1A1</i> *28 (promoteur)	irinotécan	↑ exposition, ↓ glucuronidation	(Paoluzzi et al., 2004; Ramchandani et al., 2007)
<i>UGT1A1</i> *28 (promoteur) <i>UGT1A1</i> *60 (promoteur)	bélinostat	↑ exposition	(Goey et al., 2016)
<i>UGT1A1</i> *93 (promoteur)	irinotécan	↑ exposition	(Innocenti et al., 2009)
<i>UGT1A4</i> *3 (L48V)	olanzapine	↓ exposition	(Ghotbi et al., 2010)
<i>UGT1A9</i> rs2741049 (intron)	irinotécan, propofol, MPA	↑ glucuronidation	(Girard et al., 2006; Sandanaraj et al., 2008)
<i>UGT1A9</i> rs6714486 (intron) rs17868320 (promoteur)	MPA	↓ exposition	(Kuypers et al., 2005; Levesque et al., 2007a; van Schaik et al., 2009)
<i>UGT1A9</i> rs2741045 (promoteur), <i>UGT2B7</i> rs7438135 (intron)	MPA	↑ exposition	(Fukuda et al., 2012)
<i>UGT2B10</i> rs2942857 (site d'épissage)	RO5263397*	↑ exposition	(Fowler et al., 2015)
<i>UGT2B15</i> *2 (D85Y)	saroglitazar	↑ exposition	(Stringer et al., 2014; Stringer et al., 2013a; Stringer et al., 2013b)

*RO5263397 est une molécule en essai clinique. L'effet fonctionnel correspond à l'imapet du polymorphisme sur la glucuronidation du médicament et ses niveaux en circulation. Ce tableau présente des exemples mais n'est pas exhaustif ; MPA, acide mycophénolique.

Enfin, la méthylation de l'ADN est aussi un mécanisme de régulation de l'expression des UGT. L'expression d'*UGT1A1* est importante dans le foie où le promoteur du gène est

hypo-méthylé alors qu'elle est nulle dans le rein où une hyper-méthylation du promoteur *UGT1A1* a été démontrée (Oda et al., 2013). Cette hyper-méthylation du promoteur dans le rein, couplée à une hypo-acétylation de l'histone H3, affecte le recrutement du facteur de transcription HNF1A dont dépend l'expression d'*UGT1A1*. Cependant, au niveau hépatique, la méthylation du promoteur *UGT1A1* est hétérogène entre les individus et résulte en une expression variable de l'enzyme, notamment à cause de la méthylation du site de liaison du facteur de transcription USF qui régule aussi l'expression d'*UGT1A1* (Belanger et al., 2010; Yasar et al., 2013). L'expression d'*UGT1A1* est fluctuante dans les tumeurs de cancer du côlon et une étude *in vitro* suggère que ce phénomène soit attribuable au statut de méthylation du promoteur (Gagnon et al., 2006). Ainsi, le statut de méthylation pourrait affecter la réponse à des molécules comme le SN-38, éliminé principalement par *UGT1A1*.

Des processus transcriptionnels et post-transcriptionnels sont impliqués dans l'expression des UGT. Par exemple, les miRNA sont une source de régulation des UGT. C'est le cas des *UGT1A* dont l'expression est négativement modulée par plusieurs miRNA (Dluzen et al., 2014; Tatsumi et al., 2018; Wang et al., 2018b). Plusieurs variants localisés dans des sites de liaison de miRNA au niveau du 3'UTR, commun à toutes les *UGT1A*, ont la capacité d'altérer la régulation par ces miRNA et pourrait donc créer une variabilité de l'expression des *UGT1A* entre individus (Papageorgiou et al., 2017a). Les *UGT2B*, dont le 3'UTR est spécifique à chaque transcrit, sont aussi régulées pas les miRNA. Ainsi, *UGT2B4*, *2B7* et *2B10* sont toutes régulées par le miR-216b-5p (Dluzen et al., 2016). De nombreux miRNA régulent aussi l'expression d'*UGT2B7*, d'*UGT2B15* ou d'*UGT2B17*, et parfois même plusieurs d'entre elles (Papageorgiou et al., 2017b; Wijayakumara et al., 2015; Wijayakumara et al., 2018). Ce mécanisme de régulation des *UGT2B15* et *2B17*, impliquées dans le métabolisme des androgènes, contribue à la régulation de l'inactivation de ces stéroïdes dans le cancer de la prostate (Margaillan et al., 2016).

D'autre part, les gènes *UGT* subissent un fort épissage alternatif (**Tableau 7, Figure 6**) (Levesque et al., 2007b; Tourancheau et al., 2016; Tourancheau et al., 2018). Parmi les premiers évènements d'épissage alternatif identifiés, la substitution de l'exon canonique 5 (renommé 5a) par l'exon alternatif 5b au locus *UGT1* conduit à la formation des protéines

UGT1A_i2 possédant une séquence C-terminale différente des enzymes canoniques UGT1A_i1. Ces protéines UGT1A_i2, bien que dépourvues d'activité transférase intrinsèque, sont autant et parfois même plus exprimées que les UGT1A_i1 selon les tissus (Levesque et al., 2007b). Par exemple, les UGT1A_i2, résultant de l'utilisation de l'exon 5b alternatif, sont capables d'interagir avec les enzymes UGT1A_i1 sous formes d'homo- ou d'hétéromères et de réprimer leur activité envers des substrats endogènes comme la bilirubine ou l'estradiol et des substrats exogènes tels que le SN-38 et l'acide mycophénolique (Bellemare et al., 2010a; Bellemare et al., 2010b; Bellemare et al., 2010c; Girard et al., 2007; Rouleau et al., 2013a). De plus, les protéines UGT1A_i2 sont impliquées dans la modulation du métabolisme cellulaire, et particulièrement du stress oxydatif (Rouleau et al., 2014). Leur expression est diminuée dans les tissus de cancer du côlon comparativement aux tissus normaux. Elles semblent également responsables de changements métaboliques et phénotypiques dans des cellules de cancer du côlon, notamment dans la glycolyse et le cycle des acides tricarboxyliques (Audet-Delage et al., 2017). L'expression des ARN (*UGT1A_v2/v3*) menant aux UGT1A_i2 est variable au niveau hépatique avec des CV allant de 76 à 138% chez 18 individus pour les UGT1A exprimées, tout comme au niveau rénal avec un CV de 157% pour *UGT1A6_v2/v3* et de 108% pour *UGT1A9_v2/v3* (Tourancheau et al., 2018).

Les variations génétiques peuvent parfois entraîner des modifications de l'épissage alternatif comme c'est le cas du rs8330 et des UGT1A. Ce polymorphisme est situé dans la région 3' non traduite commune des UGT1A incluant donc 1A1, 1A6 et 1A9 qui sont les principales enzymes métabolisant l'acétaminophène (Court et al., 2001). Sa fréquence allélique varie selon le groupe ethnique de 11 à 16% pour les Asiatiques, 16 à 23% pour les Caucasiens et 39 à 50% pour les Africains. Ce variant n'affecte pas la stabilité de l'ARN messager ni l'efficacité de la traduction, mais favorise l'inclusion de l'exon 5a canonique au détriment de l'exon 5b. Ainsi, les porteurs de cette variation expriment davantage les enzymes UGT1A_i1 et moins les protéines alternatives i2 inhibitrices de l'activité, ce qui se répercute par une glucuronidation accrue de 2 à 3 fois de l'acétaminophène. Une plus faible prévalence du rs8330 est observée chez des personnes subissant une insuffisance hépatique aiguë liée à une surdose accidentelle d'acétaminophène avec un rapport de cote de 0,53 (Court et al., 2013).

Des variants d'UGT2B7 codant pour des protéines différentes en N-terminal et/ou en C-terminal ont été identifiés et sont dépourvus d'activité transférase mais peuvent moduler l'activité de l'enzyme (Menard et al., 2013a; Menard et al., 2011). Par ailleurs, l'utilisation de deux promoteurs mutuellement exclusifs aboutit à la production de transcrits contenant soit l'exon 1 canonique soit un exon alternatif (1b). Les variants contenant l'exon alternatif 1b sont principalement exprimés dans les tissus extra-hépatiques, alors que l'emploi de l'exon canonique est prédominant au foie (Menard et al., 2013b). Les variants perdant l'exon 1 ont une expression diverse avec des CV de 114% au foie et 161% au rein chez 18 donneurs sains. La présence de nombreux transcrits alternatifs avec de nouvelles séquences suggère de nouvelles fonctions potentielles. Cela semble être le cas pour un variant d'UGT2B7 qui possède un nouvel exon 2, nommé exon 2b, amenant à l'inclusion de 32 acides aminés supplémentaires entre les domaines de liaison au substrat et au cosubstrat. La protéine résultante UGT2B7_i8 est inactive mais peut interagir avec l'enzyme canonique, et augmenter son activité. Elle semble aussi pouvoir affecter la prolifération et l'adhésion cellulaire et le métabolisme cellulaire (Rouleau et al., 2016). L'expression des transcrits possédant l'exon 2b, dont celui codant pour UGT2B7_i8, a une très forte variabilité interindividuelle avec des CV de 348% au foie et de 358% au rein.

Une étude globale sur le transcriptome *UGT* a montré que l'ensemble des UGT sont concernées par l'épissage alternatif, révélant l'existence de plus de 130 variants d'épissage UGT (Tourancheau et al., 2016). Leur expression est parfois supérieure aux transcrits canoniques, selon le tissu et l'UGT concernée (**Tableau 7, Figure 6**) (Tourancheau et al., 2018). Par exemple, plus de la moitié des transcrits *UGT2B10*, gène fortement exprimé au foie, sont des transcrits alternatifs. On observe aussi une abondance variable entre des tissus normaux ou cancéreux. Dans le tissu rénal sain, la majorité des transcrits d'*UGT2B7* sont canoniques alors que dans le rein tumoral, presque uniquement des transcrits alternatifs sont détectés. Les transcrits alternatifs UGT peuvent être regroupés en trois classes principales selon qu'ils sont tronqués en C-terminal, en N-terminal ou qu'ils possèdent de nouvelles séquences codantes suggérant de nouvelles fonctions (Rouleau et al., 2016). L'expression des variants alternatifs est, comme pour les transcrits canoniques, très différente entre les individus avec des CV compris entre 46 et 348% au foie pour des variants d'UGT1A, 2B17, 2B10 et 2B15 et des CV de 108 à 358% pour des variants

d'UGT1A6, 1A9 et 2B7 (Tourancheau et al., 2018). Les variants d'épissage de gènes impliqués dans la réponse au traitement représentent donc une source de variabilité interindividuelle de réponse, mais aussi des facteurs prédictifs et des cibles thérapeutiques potentiels (Miura et al., 2012; Yi et al., 2011).

En somme, les variations génétiques héréditaires contribuent dans une certaine proportion à la variabilité de la voie de glucuronidation. Cependant, le variant *UGT1A1*28* n'explique pas toute la variabilité hépatique d'UGT1A1, comme le montre par exemple les différences d'expression au sein des groupes *UGT1A1*1/*1* (CV = 54%), *UGT1A1*1/*28* (CV = 93%) et *UGT1A1*28/*28* (CV = 81%) (Margaillan et al., 2015b). De la même façon, les différences dans l'expression protéique d'UGT2B17 peuvent être expliquées seulement à hauteur de 26% par l'âge, le sexe, 4 variants génétiques et les CNV (Bhatt et al., 2018). Cela suggère que d'autres mécanismes, comme par exemple l'épissage alternatif pourraient contribuer à l'hétérogénéité d'expression de la voie de glucuronidation. En fait, la plus grande partie des mécanismes – génétiques ou pas – expliquant cette variabilité interindividuelle des UGT demeure inconnue à ce jour.

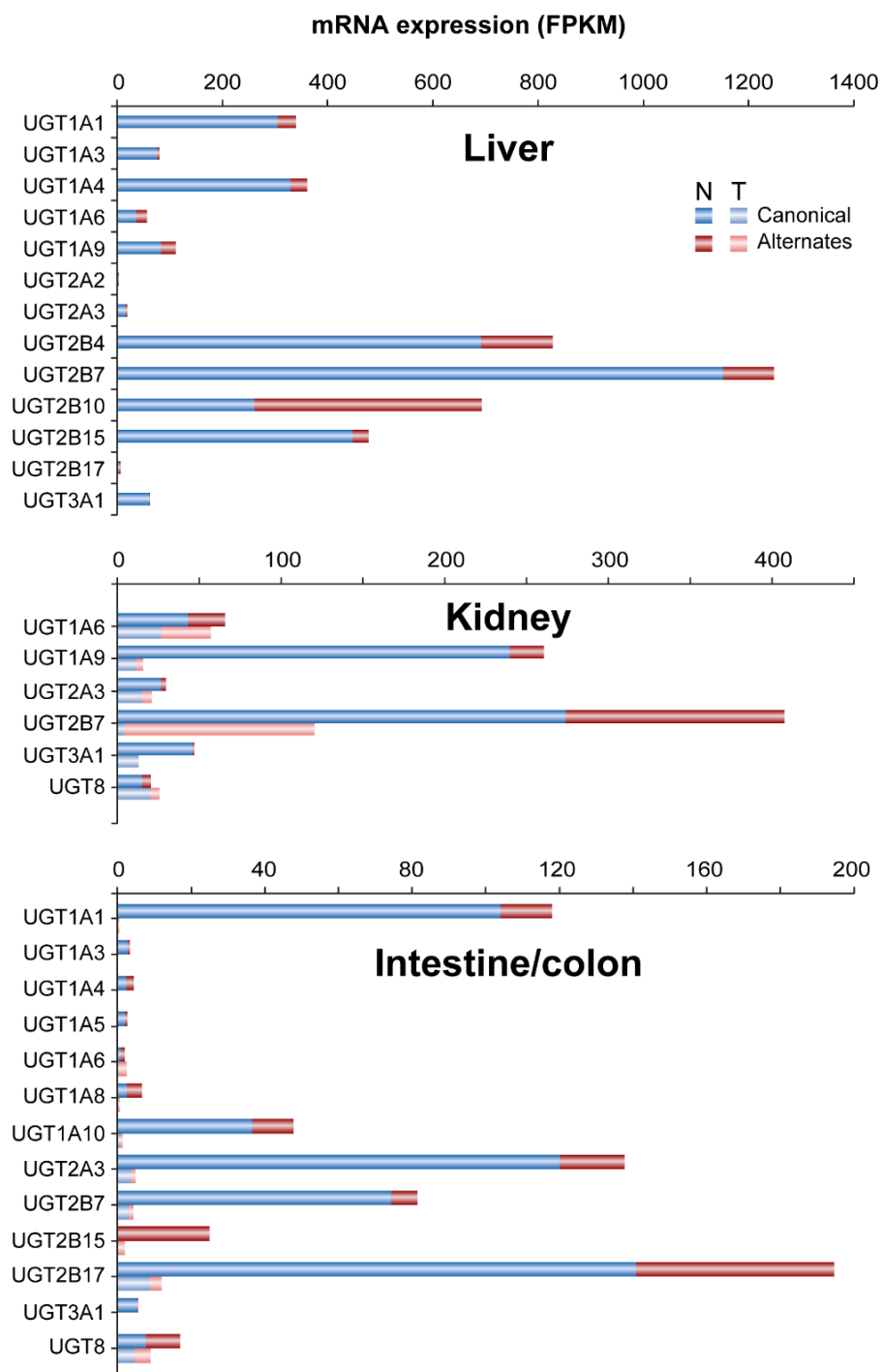


Figure 6. Quantification de l'expression des transcrits canoniques et alternatifs des UGT dans plusieurs tissus du métabolisme des médicaments sains et cancéreux.

Les données proviennent de 9 échantillons par tissus, sauf pour l'intestin/côlon normal (15 échantillons). Il n'y a pas d'échantillon tumoral de foie. FPKM, *fragments per kilobase of transcript per million mapped reads* ; N, tissu normal ; T, tissu tumoral. Figure issue de Tourancheau et al., 2018. Reproduit avec la permission du *Copyright Clearance Center* pour *Springer Nature*.

4) Hypothèses et objectifs de recherche

Objectif 1 : Découverte de nouveaux marqueurs prédictifs de la réponse à l'irinotécan et leur caractérisation moléculaire.

1.1) Découverte de marqueurs génétiques prédictifs de la réponse à l'irinotécan.

L'irinotécan a été choisi comme modèle d'étude car il est pris en charge principalement par la voie des UGT. Comme tous les traitements en oncologie, la réponse à l'irinotécan est variable entre les individus (**Tableau 3**). La génétique du patient est un des facteurs qui influencent la réponse au traitement. En effet, le marqueur *UGT1A1**28 permet d'identifier les patients susceptibles de développer des neutropénies de grades 3-4 suite à la prise d'irinotécan dans le cadre du traitement du CCRm (Cote et al., 2007; Iyer et al., 2002; Liu et al., 2008; Marcuello et al., 2004; Toffoli et al., 2006). Cependant, toutes les données disponibles ne soutiennent pas que le marqueur *UGT1A1**28 prédise la survie. En fait, il ne semble pas y avoir de marqueur validé permettant de prédire la survie sans progression ou la survie globale chez les patients atteints d'un CCRm et traités avec l'irinotécan. Quelques marqueurs du locus *UGT1* et des gènes codant pour des transporteurs ont été associés à la survie dans des études parfois de taille limitée (**Annexe A**). Surtout, peu d'études portant simultanément sur plusieurs de ces voies ont été réalisées, et les marqueurs positifs ont rarement été répliqués dans une population indépendante ayant reçu le même traitement. Dans le cadre de mon projet, une étude exhaustive des variations génétiques communes portant sur les voies PK/PD de l'irinotécan a été conduite, et les marqueurs positifs ont été répliqués dans une population indépendante.

Hypothèse : Les polymorphismes des voies PK/PD et leurs combinaisons pourraient permettre de mieux prédire la réponse au traitement.

1.2) Étude des mécanismes moléculaires en lien avec les marqueurs prédictifs.

Hypothèse : L'effet biologique n'est pas nécessairement porté par le marqueur identifié mais plutôt par un ou plusieurs variants en déséquilibre de liaison avec lui. L'identification des variants fonctionnels permettrait de mieux comprendre quels sont les mécanismes qui altèrent la réponse au médicament. Il serait aussi possible d'identifier d'autres médicaments

pris en charge par les mêmes voies et dont la réponse pourrait également être impactée par ces variants.

Objectif 2 : Étude de la fonction de nouveaux variants d'épissage.

Nos travaux suggèrent que l'épissage alternatif constitue une source importante de variabilité pour l'expression des gènes *UGT*, avec un impact significatif sur la capacité de glucuronidation du tissu (Court et al., 2013; Menard et al., 2013a; Menard et al., 2011; Rouleau et al., 2016; Tourancheau et al., 2016; Tourancheau et al., 2018). Une étude du laboratoire publiée en 2015 révèle la très grande diversité de variants d'épissage issus des 10 gènes *UGT* humains (Tourancheau et al., 2016). Ces travaux ont identifié plus de 130 nouveaux variants d'épissage dont certains possèdent de nouvelles séquences codantes internes ou en C-terminal du cadre de lecture, maintenant ainsi la séquence *UGT*. Au foie, le transcriptome d'*UGT2B10* est constitué à plus de 50% de variants alternatifs (**Tableau 7, Figure 6**). Deux de ces variants, *UGT2B10_n9* et *UGT2B10_n10* possèdent une nouvelle séquence C-terminale unique, codée par les nouveaux exons 6b et 6c. L'abondance de ces variants et la présence d'une nouvelle séquence codante justifient une étude plus approfondie sur leur rôle biologique potentiel.

Hypothèse : Les variants d'*UGT2B10* influencent la glucuronidation hépatique.

5) Justification de la méthodologie retenue

5.1) Découverte de marqueurs génétiques prédictifs de la réponse au traitement

Le premier objectif vise à identifier de nouveaux marqueurs associés à la survie des patients atteints d'un CCRm traités avec le régime FOLFIRI. En 2017 au Canada, le cancer colorectal était la 2^{ème} cause de décès par cancer chez l'homme et la 3^{ème} chez la femme ce qui représente 9400 décès, soit 12% des décès imputables au cancer² (Yu et al., 2018a). Bien que la survie relative à 5 ans soit de 64%, elle chute à seulement 13% pour les cas

² <http://www.cancer.ca/en/cancer-information/cancer-type/colorectal/colorectal-cancer>, consulté le 13 février 2019

diagnostiqués au stade métastatique de la maladie, ce qui représente plus de 20% des cas^{2,3} (Yu et al., 2018a). Il existe plusieurs combinaisons et séquences de chimiothérapie, mais les agents utilisés sont le plus souvent le 5-FU et/ou l'irinotécan et/ou l'oxaliplatine et l'apport des thérapies ciblées comme le bévacizumab a davantage accru la survie (Fujita et al., 2015; Modest et al., 2019; Saltz et al., 2000; Sanchez-Gundin et al., 2018; Sandhu et al., 2019).

L'approche par gènes candidats a été retenue plutôt que l'approche pangénome car cette dernière nécessite une cohorte de grande taille afin d'obtenir des résultats significatifs avec une puissance statistique adéquate. Une force de notre étude est de disposer d'une cohorte de découverte et d'une cohorte de validation, pour un nombre total de patients (n = 417) supérieur à la plupart des études précédentes (**Annexe A**). Cela permet de répliquer les associations initiales entre les variants et l'issue clinique, limitant ainsi le risque de faux-positifs. Les cohortes et les critères d'inclusions sont présentés en détail dans les chapitres I à III.

Nous avons sélectionné des gènes candidats selon plusieurs critères : 1) gènes de la PK de l'irinotécan, 2) gène de la PD de l'irinotécan, 3) gènes de facteurs de transcription et récepteurs nucléaires régulant les voies PK/PD, 4) gènes associés à la réponse au traitement rapportés précédemment dans des études *in vitro* ou chez le patient. Les variants fréquents (fréquence allélique mineure > 5%) de ces gènes et des régions avoisinantes (± 5 kb) ont été retenus. Afin de limiter le nombre de variations à génotyper pour chaque patient et de maximiser la couverture de criblage, une approche de variation étiquette (htSNP, *haplotype-tagging single nucleotide polymorphism*) a été choisie. Une seule variation étiquette a été génotypée pour l'ensemble des variants qui sont en déséquilibre de liaison avec elle (LD, $r^2 > 0.8$), c'est-à-dire donnant la même information génétique. Les variations étiquettes ont été génotypées principalement par spectrométrie de masse à temps de vol grâce à la technologie iPLEX de Sequenom.

Les issues cliniques principales étudiées étaient la survie globale (OS) et la survie sans progression (PFS). Les issues cliniques secondaires sont les toxicités hématologiques

³ <https://seer.cancer.gov/statfacts/html/colorect.html>, consulté le 13 février 2019

(neutropénies sévères de grades 3-4) et les toxicités gastrointestinales (diarrhées sévères de grades 3-4). Les issues cliniques sont détaillées dans le **Tableau 9**. Les associations entre variations étiquettes et survie ont été testées en utilisant des modèles proportionnels de Cox multivariés ajustés pour l'âge et le cotraitement (certains patients de la cohorte de découverte ayant reçu du bévacizumab) et des analyses de survie de Kaplan-Meier univariées. Les associations avec les toxicités sévères ont été testées avec des régressions logistiques ajustées pour les mêmes covariables. Toutes les variations étiquettes ont été testées dans la cohorte de découverte en utilisant des modèles génétiques additifs, dominants et récessifs. Seules les variations étiquettes associées à une issue clinique dans la cohorte de découverte avec une $P < 0.1$ (quel que soit le modèle génétique) ont été conservées pour la réplique dans la seconde cohorte. Le choix d'un seuil permissif à 10% ($P < 0.1$) a été fait dans le but d'inclure un maximum de marqueurs potentiels dans l'étape de réplique, où un marqueur est considéré validé s'il a le même effet que dans la cohorte de découverte au seuil de 5% ($P < 0.05$).

Tableau 9. Issues cliniques étudiées en lien avec le régime FOLFIRI.

Issue clinique	Définition
survie globale (OS)	durée entre le début du traitement FOLFIRI et le décès (de toute cause) du patient
survie sans progression (PFS)	durée entre le début du traitement FOLFIRI et la première preuve de progression de la maladie, incluant le décès, ou le dernier suivi du patient
toxicité hématologique : neutropénie (sévère : grades 3–4)	nombre de neutrophiles ($\times 10^9$) par litre : - grade 0 : $> 1,7$ (pas de neutropénie) - grade 1 : 1,5–1,7 - grade 2 : 1,0–1,5 - grade 3 : 0,5–1,0 - grade 4 : $< 0,5$
toxicité gastrointestinale : diarrhée (sévère : grades 3–4)	augmentation de la fréquence des selles par jour : - grade 0 : pas d'augmentation (pas de toxicité) - grade 1 : < 4 - grade 2 : 4–6 - grade 3 : ≥ 7 - grade 4 : pronostic vital engagé, hospitalisation nécessaire

Les définitions concernant les toxicités hématologique et gastrointestinale sont basées sur les critères du *NCI Common Terminology Criteria for Adverse Events v3.0 (2006)*.

5.2) Étude des mécanismes moléculaires en lien avec les marqueurs liés à l'issue clinique

Puisque nous avons utilisé une approche de variations étiquettes, le marqueur associé à l'issue clinique n'est pas nécessairement celui qui a l'effet biologique. La première étape consiste donc à définir, parmi l'ensemble des variants liés à la variation étiquette, lequel est le plus susceptible d'avoir un effet biologique. Nous avons considéré par ordre d'importance décroissante les variants 1) codant pour une substitution d'acide aminé, 2) localisés dans une région régulatrice (promoteur, 5'-UTR, 3'-UTR, sites d'épissage, etc.) 3) localisés dans les introns et 4) intergéniques. Des outils permettent aussi d'identifier des variants potentiellement fonctionnels comme par exemple les annotations RegulomeDB, ENCODE ou HaploReg (Boyle et al., 2012; Davis et al., 2018; ENCODE Project Consortium, 2012; Ward et al., 2016).

Après l'identification des variants potentiellement fonctionnels, nous avons voulu tester leur impact dans des modèles *in vitro* et *ex vivo*. Des données pharmacocinétiques sur les niveaux plasmatiques d'irinotécan, de SN-38 et de SN-38G sont disponibles pour un sous-groupe de 71 patients de la cohorte de validation et ont permis d'étudier l'impact du variant sur l'exposition au médicament et ses métabolites. Nous avons aussi utilisé les données de la cohorte GTEx, constituée d'individus non-atteints du cancer pour corréler variations et expression génique dans chaque tissu d'intérêt (The GTEx Consortium, 2013). D'autres jeux de données disponibles publiquement, notamment ceux de la cohorte TCGA constituée de patients atteints du cancer, ont aussi été choisis pour étudier le lien entre l'expression génique et des données cliniques (Weinstein et al., 2013). Enfin, des essais *in vitro* ont été conduits pour corroborer les associations et corrélations précédentes. En utilisant des modèles cellulaires, nous avons exprimé des protéines variantes ou modulé l'expression de gènes identifiés pour conduire des essais fonctionnels et investiguer l'impact des variations sur le phénotype des cellules cancéreuses. Nous avons donc combiné des approches *in vivo* et *in vitro* afin d'étudier l'impact moléculaire des marqueurs de la réponse au régime FOLFIRI que nous avons identifiés.

5.3) Étude de la fonction de nouveaux variants d'épissage

Nous avons choisi de nous intéresser plus particulièrement à UGT2B10 car c'est l'une des UGT les plus exprimées au foie alors que les variants d'épissage de ce gène représentent plus de la moitié de son transcriptome (**Tableau 7, Figure 6**). De plus, c'est la principale enzyme impliquée dans le métabolisme des dérivés du tabac tels que la nicotine et la cotinine ainsi que la N-glucuronidation de nombreux médicaments contenant des amines comme des antidépresseurs ou des antipsychotiques (**Tableau 4**). Ainsi, la compréhension de l'impact fonctionnel des variants d'épissage sur l'activité hépatique d'UGT2B10 pourrait permettre d'identifier différents statuts de métaboliseurs.

La première étape a été de répertorier les variants et de valider leur expression en ARN et en protéine. Nous avons choisi d'orienter notre étude vers les variants qui conduisaient à la production de protéines avec de nouvelles séquences codantes en cadre de lecture avec la séquence UGT. Ensuite, des données de séquençage de plusieurs individus ont été utilisées afin de quantifier l'abondance des différents transcrits et d'établir leur profil d'expression interindividuel. Cette quantification a été faite sur des tissus sains (cohorte GTEx) et cancéreux (cohorte TCGA) dans l'objectif d'étudier les variations selon la condition physiologique, d'autant que les UGT sont impliquées dans la détoxification de nombreux médicaments et carcinogènes.

Par la suite, nous avons établi des modèles cellulaires exprimant l'enzyme, les variants étudiés ou coexprimant enzyme et variant étant donné leur coexpression *in vivo*. Pour ce faire, nous avons utilisé les HEK293 et les HepG2. Les HEK293 ont été choisies car elles n'expriment pas d'UGT de façon endogène et représentent donc un bon modèle pour étudier une isoforme en particulier, sans interférence avec d'autres UGT. Les cellules hépatiques HepG2 sont aussi un modèle de choix car elles expriment certaines UGT, et notamment l'enzyme UGT2B10.

Ces modèles cellulaires ont permis d'étudier l'activité enzymatique et de quantifier l'influence des isoformes alternatives sur l'activité de l'enzyme canonique. Ces modèles ont aussi été utiles pour établir la localisation subcellulaire des isoformes et les possibles interactions entre variants et enzyme. Enfin, les lignées établies ont permis d'explorer les

mécanismes de régulation de l'enzyme et des variants UGT2B10, tant au niveau de l'ARN que de la protéine. L'utilisation de données de séquençage d'ARN sur des hépatocytes a permis d'étudier l'induction d'*UGT2B10* en réponse à un traitement pharmacologique.

Chapitre 1 : Combination of germline variations in irinotecan pharmacogenetic pathways associated with survival in metastatic colorectal cancer treated with FOLFIRI-based chemotherapy

1.1 Résumé

Les polymorphismes génétiques des voies de la pharmacocinétique et de l'action des médicaments ont le potentiel de modifier la survie des patients. Nous avons réalisé un criblage systématique de 11 gènes candidats associés à la pharmacologie de l'agent anticancéreux irinotécan. Le lien entre les variants génétiques et la survie globale a été établi dans deux cohortes indépendantes totalisant 417 patients CCRm, recevant un traitement de première ligne par FOLFIRI. Le génotypage par spectrométrie de masse à temps de vol a été réalisé à l'aide de la plateforme Sequenom MassARRAY. L'expression génique a été étudiée dans les modèles cellulaires et dans les tissus de côlon de patients CCRm de stade IV à partir des jeux de données TCGA (n = 87) et GSE39582 (n = 60). En combinant les deux cohortes, la variation *CESI* rs9921399T>C a été associée à une survie plus longue avec un HR de 0,40 ($P = 0,005$), tandis que les variations *ABCC1* rs17501011G>A et *UGT1* rs1113193G>A ont été associées de manière significative à une survie réduite avec des HR de 2,08 ($P = 0,003$) et 2,12 ($P = 0,004$), respectivement. Les analyses fonctionnelles ont montré que le variant intronique *ABCC1* rs17501011A est associé à une activité réduite dans une expérience de gène rapporteur dans des cellules cancéreuses de foie ($P < 0,01$) et de côlon ($P < 0,05$) par rapport à l'allèle rs17501011G. L'analyse de deux cohortes supplémentaires de patients CCRm, révèle que l'expression tumorale d'*ABCC1* est en outre associée à une survie réduite ($P \leq 0,013$). Une analyse d'haplotype a montré un effet combiné de ces polymorphismes en lien avec la survie, avec des HR de 1,98 ($P = 0,038$) à 2,97 ($P = 0,018$). Les patients porteurs de deux ou trois génotypes à risque présentaient une issue clinique défavorable avec des HR de 2,17 ($P < 0,001$) et 5,35 ($P = 0,004$), respectivement, par rapport aux porteurs d'un seul génotype à

risque. Les résultats de cette étude identifient une combinaison de déterminants génétiques potentiels pouvant contribuer à prédire la survie au CCRm. Des études additionnelles sont nécessaires pour valider ces résultats.

1.2 Abstract

Aim: Germline variants in drug-related pathways could modify survival outcome of treated metastatic colorectal cancer patients (mCRC).

Patients & Methods: We systematically analyzed the association of 285 htSNPs in 11 genes related to irinotecan pharmacology (*ABCB1*, *ABCC1*, *ABCC2*, *ABCC5*, *ABCG1*, *ABCG2*, *SLCO1B1*, *CES1*, *CES2*, *TOP1* and *UGT1*) and overall survival (OS), in two independent cohorts totalizing 417 mCRC patients receiving first-line FOLFIRI treatment. Genotyping was performed using the Sequenom MassARRAY platform. Gene expression analysis was investigated *in vitro* and in stage IV mCRC patients from the TCGA (n=87) and GSE39582 (n=60) datasets.

Results: The variation *CES1* rs9921399T>C was associated with longer OS with a hazard ratio (HR) of 0.40 ($P=0.005$) in the two cohorts whereas *ABCC1* rs17501011G>A and *UGT1* rs1113193G>A were significantly associated to shorter OS with HR of 2.08 ($P=0.003$) and 2.12 ($P=0.004$) in the combined cohort, respectively. A haplotype analysis of combined cohorts showed a combined effect of these polymorphisms on survival with HR values of 1.98 ($P=0.038$) to 2.97 ($P=0.018$). Patients carrying two or three unfavourable genotypes exhibited worse outcome with HR of 2.17 ($P<0.001$) and 5.35 ($P=0.004$), respectively, compared to carriers of one unfavourable genotype. Functional analyses showed that the intronic *ABCC1* rs17501011A variant was associated with reduced reporter-gene activity in liver ($P<0.01$) and colon ($P<0.05$) cells compared to rs17501011G. *ABCC1* tumor expression was further associated with shorter survival ($P\leq 0.013$) in two additional cohorts of mCRC.

Conclusion: Our study identified a potential combination of genetic determinants that may contribute to predict survival of mCRC. Further studies are needed to validate these findings.

SHORT COMMUNICATION

COMBINATION OF GERMLINE VARIATIONS IN IRINOTECAN PHARMACOGENETIC PATHWAYS ASSOCIATED WITH SURVIVAL IN METASTATIC COLORECTAL CANCER TREATED WITH FOLFIRI-BASED CHEMOTHERAPY

Adrien Labriet¹, Éric Lévesque², Elena De Mattia³, Erika Cecchin³, Derek Jonker⁴, Félix Couture², David Simonyan⁵, Angela Buonadonna⁶, Mario D'Andrea⁷, Lyne Villeneuve¹, Giuseppe Toffoli³ and Chantal Guillemette^{1,8*}

¹Pharmacogenomics Laboratory, Centre Hospitalier Universitaire de Québec (CHU de Québec) Research Center and Faculty of Pharmacy, Laval University, Québec, Québec, Canada; ²CHU de Québec Research Center and Faculty of Medicine, Laval University, Québec, Québec, Canada; ³Clinical and Experimental Pharmacology, “Centro di Riferimento Oncologico”- National Cancer Institute, via Franco Gallini n. 2, 33081 Aviano (PN), Italy; ⁴Division of Medical Oncology, Department of Medicine, Ottawa Hospital, University of Ottawa, Ottawa, Ontario, Canada; ⁵Clinical and Evaluative Research Platform, CHU de Québec Research Center, Québec, Canada; ⁶Medical Oncology Unit, “Centro di Riferimento Oncologico”- National Cancer Institute, via Franco Gallini n. 2, 33081, Aviano (PN), Italy; ⁷Medical Oncology Unit, “San Filippo Neri Hospital”, Via Giovanni Martinotti, 20, 00135 Rome, Italy; ⁸Canada Research Chair in Pharmacogenomics, CHU de Québec Research Center, Québec, Québec, Canada.

*Corresponding Author: Chantal Guillemette, Ph.D., CHU de Québec Research Center, R4701.5, 2705 Blvd. Laurier, Québec, Canada, G1V 4G2. E-mail: chantal.guillemette@crchudequebec.ulaval.ca

E-mail: Adrien Labriet: adrien.labriet.1@ulaval.ca; Éric Lévesque: eric.levesque@crchudequebec.ulaval.ca; Elena De Mattia: edemattia@cro.it; Erika Cecchin: ececchin@cro.it; Derek Jonker: djonker@toh.on.ca; Félix Couture: felixcou@videotron.ca; David Simonyan: David.Simonyan@crchudequebec.ulaval.ca; Angela Buonadonna: ab Buonadonna@cro.it; Mario D'Andrea: mariorosario.dandrea@asl-

rme.it; Lyne Villeneuve: Lyne.Villeneuve@crchudequebec.ulaval.ca; Giuseppe Toffoli: gtoffoli@cro.it; Chantal Guillemette: Chantal.Guillemette@crchudequebec.ulaval.ca

Keywords: irinotecan, metastatic colorectal cancer, polymorphisms, biomarkers, pharmacogenomics.

Running title : Pharmacogenetic markers of mCRC survival.

WORD COUNT:

Abstract: 256

Introduction: 317

Patients & Methods: 752

Results & Discussion: 1065

Body text: 2134

NUMBER OF:

References: 50

Tables: 1

Figures: 3

Supplementary Tables: 2

Supplementary Figures: 3

Abstract

Aim: Germline variants in drug-related pathways could modify survival outcome of treated metastatic colorectal cancer patients (mCRC).

Patients & Methods: We systematically analyzed the association of 285 htSNPs in 11 genes related to irinotecan pharmacology (*ABCB1*, *ABCC1*, *ABCC2*, *ABCC5*, *ABCG1*, *ABCG2*, *SLCO1B1*, *CES1*, *CES2*, *TOP1* and *UGT1*) and overall survival (OS), in two independent cohorts totalizing 417 mCRC patients receiving first-line FOLFIRI treatment. Genotyping was performed using the Sequenom MassARRAY platform. Gene expression analysis was investigated *in vitro* and in stage IV mCRC patients from the TCGA (n=87) and GSE39582 (n=60) datasets.

Results: The variation *CES1* rs9921399T>C was associated with longer OS with a hazard ratio (HR) of 0.40 ($P=0.005$) in the two cohorts whereas *ABCC1* rs17501011G>A and *UGT1* rs1113193G>A were significantly associated to shorter OS with HR of 2.08 ($P=0.003$) and 2.12 ($P=0.004$) in the combined cohort, respectively. A haplotype analysis of combined cohorts showed a combined effect of these polymorphisms on survival with HR values of 1.98 ($P=0.038$) to 2.97 ($P=0.018$). Patients carrying two or three unfavourable genotypes exhibited worse outcome with HR of 2.17 ($P<0.001$) and 5.35 ($P=0.004$), respectively, compared to carriers of one unfavourable genotype. Functional analyses showed that the intronic *ABCC1* rs17501011A variant was associated with reduced reporter-gene activity in liver ($P<0.01$) and colon ($P<0.05$) cells compared to rs17501011G. *ABCC1* tumor expression was further associated with shorter survival ($P\leq 0.013$) in two additional cohorts of mCRC.

Conclusion: Our study identified a potential combination of genetic determinants that may contribute to predict survival of mCRC. Further studies are needed to validate these findings.

Introduction

Prognosis of mCRC remains poor with a 5-year survival of 14% and a majority of patients cannot be cured with surgical resection [1, 2]. A significant number of new agents have been approved in this clinical setting. Systemic chemotherapy with first-line regimens containing irinotecan or oxaliplatin leads to significant improvements in survival and similar efficacy but different toxicity profiles [2-5]. Oxaliplatin is a platinum-derived DNA-damaging agent while irinotecan exerts its cytotoxic effects through inhibition of topoisomerase I in replicating cells by inducing DNA strand breaks. New agents include the monoclonal antibody bevacizumab targeting the vascular endothelial growth factor (VEGF) whereas cetuximab or panitumumab inhibit the epidermal growth factor receptor (EGFR), preventing tumor development [6, 7].

The discovery of predictive markers may assist clinicians in practice for the choice of therapy. Indeed, strategies based on germline DNA markers may help optimize treatment dosing whereas markers based on tumor biology are used for the appropriate designation of subsequent lines of palliative chemotherapy. [5, 8]. For irinotecan-based regimens, recent discoveries in genes related to the pharmacology of the drug, namely variants found in the *UGT1* locus [9-14] as well as in ATP-binding cassette (ABC) transporter genes [15-19] and in the *SLCO1B1* gene [19, 20], have been shown to impact clinical outcomes of patients by affecting drug exposure and action.

To improve our knowledge on genetic factors that can contribute to patients' survival, we systematically analyzed the association of common polymorphisms in 11 candidate genes related to irinotecan pharmacokinetics and pharmacodynamics (*ABCB1*, *ABCC1*, *ABCC2*, *ABCC5*, *ABCG1*, *ABCG2*, *SLCO1B1*, *CES1*, *CES2*, *TOP1* and *UGT1*). Our study was based on a well-characterized population of mCRC patients initiating a first-line regimen with FOLFIRI (combining folinic acid, 5-fluorouracil and irinotecan). Positive markers were replicated in an independent larger cohort, and the combined effect of positive markers on survival was assessed in the combined cohort. The biological plausibility was reinforced by functional investigations for validated markers.

Patients & methods

Patient cohorts and clinical outcomes.

A prospective cohort of 167 mCRC patients from eastern Canada was recruited from 2003 to 2012 and was set as the discovery cohort. All patients received first-line FOLFIRI and 75 patients also received the monoclonal antibody bevacizumab or another drug. Details on the treatment procedure and eligibility criteria were reported previously [21]. The replication cohort was comprised of 250 mCRC patients treated with FOLFIRI recruited in Northeast Italy between 2002 and 2005 [22]. The main characteristics of the study populations are summarized in **Table 1**. Overall survival (OS) was defined as the time (in months) between the beginning of irinotecan treatment and death from any cause. This study was conducted in accordance with the declaration of Helsinki. All participants provided a written informed consent for genetic analysis. The CHU de Quebec ethics committee and the Comitato Etico Indipendente - Centro di Riferimento Oncologico di Aviano approved the research protocols (2012-245). Additional publicly available datasets were used for expression analysis and were comprised of mCRC cases from The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>) accessed through the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>) and retrieved with TCGAbiolinks package (version 2.9.5). We also used GSE39582 obtained through the Gene Expression Omnibus (GEO) database with GEOquery package (version 2.48.0).

Genetic analysis.

Haplotype-tagging single nucleotide polymorphisms (htSNPs) were selected in *ABCB1*, *ABCC1*, *ABCC2*, *ABCC5*, *ABCG1*, *ABCG2*, *SLCO1B1*, *CES1*, *CES2*, *TOP1* and *UGT1* genes and their flanking regions (± 5 kb). They were genotyped in both cohorts to maximize the coverage of the genetic variability of the candidate genes, as described [15, 23, 24]. The list of the htSNPs is provided in **Supplementary Table 1**. For each htSNP, deviation from Hardy-Weinberg equilibrium was determined using PLINK v1.07 and those deviating from the equilibrium ($P < 0.05$) were not retained for further analysis [25]. Genotypic frequencies of htSNPs in different populations from 1000 Genomes Phase 3 were retrieved through Ensembl GRCh37 release 91. Prioritization of potential functional

markers was established using RegulomeDB and those with a score $\geq 2b$ were retained for luciferase reporter gene assay (**Supplementary Figure 1**) [26].

Luciferase reporter assays.

Fragments containing the major allele of the SNPs of interest were PCR amplified from genomic DNA using Phusion High-Fidelity DNA polymerase (New England Biolabs, Whitby, ON, Canada). Fragments were then inserted between SacI and XhoI restriction sites of pGL4.23 luciferase vector (Promega, Madison, WI). Constructs carrying minor alleles were obtained either from PCR amplification from genomic DNA or by site directed mutagenesis using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs, Whitby, ON, Canada). All primers used are listed in **Supplementary Table 2** and all constructs were sequenced using the RV primer 3. Cells were seeded in 24-well plates at a density of 1×10^5 cells/well (HepG2, HuH7 and HT-29) or 1.5×10^5 cells/wells (HCT116). After 24 h, cells were transfected with the appropriate construct (200 ng for HCT116, 400 ng for HepG2 and HuH7 and 800 ng for HT-29) and a CMV-renilla internal control construct (2 ng for HCT-116, 4 ng for HepG2 and HuH7 and 20 ng for HT-29) using lipofectamine 2000 as recommended per the manufacturer (Invitrogen, Carlsbad, CA). Medium was replaced with fresh medium 3 hours post-transfection. Cells were washed with PBS and processed for luciferase activity using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI) 24 h post-transfection. For each condition, luciferase activity was divided by the renilla internal control intensity and reported relative to the ratio obtained for the control (empty pGL4.23 vector). Three independent experiments were performed in triplicate. Differences between groups were compared using a two-tailed t-test performed in GraphPad Prism 5 software (GraphPad Software Inc., La Jolla, CA).

Statistical analysis.

Genetic associations between genotypes and OS were tested using Cox proportional hazards models adjusted for covariates (age and cotreatment for the Canadian and combined cohorts, age in the Italian cohort). Additive, dominant and recessive models were fitted independently for each htSNP. Only htSNPs associated with the outcome ($P < 0.1$) in the Canadian cohort were replicated in the Italian cohort. A htSNP was considered

validated if associated ($P < 0.05$) with the outcome in the Italian cohort using the same genetic model. A bootstrap resampling (1000 random samples) was used to validate the reliability of the Cox models in the Canadian, the Italian and the combined cohorts. Statistical analyzes were performed using SAS v9.4 and IBM SPSS Statistics version 23. Kaplan-Meier analysis was performed using GraphPad Prism 5 (GraphPad Software Inc., La Jolla, CA, USA) and survival curves were compared using the Log-rank test.

Results & discussion

We systematically analyzed 285 htSNPs of 11 candidate genes in two independent cohorts of 417 metastatic colorectal cancer (mCRC) patients receiving first-line FOLFIRI treatment. In the discovery cohort, 54 htSNPs were associated ($P < 0.1$) with OS. Of those, three were validated ($P < 0.05$) in the Italian cohort (**Figure 1A**). These 3 markers are located in intronic regions (**Supplementary Figure 1**) and their frequency varied greatly among ethnic groups (**Supplementary Figure 2**).

A first marker, *CES1* rs9921399T>C, was associated with an increased OS using a recessive model with adjusted hazard ratios (HR) of 0.53 ($P = 0.069$) in the Canadian cohort, 0.14 ($P = 0.049$) in the Italian cohort and 0.40 ($P = 0.005$) in the combined cohorts (**Figure 1A**). Kaplan-Meier analysis revealed that carriers of *CES1* rs9921399CC genotype had an improved median OS by 17 months (39 months versus 22 months, $P = 0.004$, **Figure 1B**). *CES1* predominates in the liver whereas *CES2* is also expressed in other tissues such as the small intestine [27, 28]. Despite the fact that *CES1* has an inferior affinity for conversion of irinotecan pro-drug into SN-38 compared to *CES2*, reports suggest that *CES1* would account for up to half of this reaction in the liver due to its higher expression levels [29, 30]. Our functional investigation of the intronic variant *CES1* rs12443580T>C tightly linked to the intronic rs9921399T>C did not reveal significant differences in transcriptional activity in the condition tested (**Supplementary Figure 3A**). However, a previous study supports a gene-dose effect for *CES1* on SN-38 formation and exposure in irinotecan-treated Japanese cancer patients [31].

A second marker, the rs17501011G>A located in the first intron of *ABCC1*, was associated with a reduced OS using a dominant model with HR of 1.83 ($P = 0.070$), 2.53 ($P = 0.013$) and 2.08 ($P = 0.003$) in the discovery, replication and combined cohorts, respectively (**Figure 1A**). Kaplan-Meier analysis showed that carriers of *ABCC1* rs17501011A had a reduced median OS by 6 months (17 months versus 23 months, $P = 0.007$, **Figure 1C**). We further demonstrated that the intronic *ABCC1* rs17501011A allele was associated to a significantly lower reporter gene activity compared to the G allele, in liver cells HepG2 ($P < 0.01$) and HuH7 ($P < 0.01$) and colon cancer cells HCT-116 ($P < 0.05$) and HT-29 ($P < 0.05$), supporting a functional impact (**Figure 2A**). Few genetic variations in *ABCC1* were

previously linked to increased SN-38 exposure, namely rs6498588A>T located in the 5'-flanking region, and rs35605T>C in exon 13 leading to a synonymous substitution [32, 33]. The latter was further reported to contribute to nearly 30% to plasma levels of SN-38 along with *ABCB1* rs10276036, and *UGT1A1**93 [33]. We further exposed that low tumor *ABCC1* expression was associated with shorter survival of mCRC cases from two independent datasets, the TCGA cohort (n = 87; *P* = 0.013) and GSE39582 (n = 60; *P* = 0.001) (**Figure 2B-C**). *ABCC1* encodes the MRP1 transporter, which was involved in drug resistance and treatment failure in multiple cancers [34, 35]. Here, we observed shorter survival for patients with lower *ABCC1* tumoral expression and carriers of the rs17501011A allele associated with reduced reporter gene expression. The decreased survival of mCRC cases with lower expression of *ABCC1* may result from several mechanisms. Although expressed at much lower levels in liver compared to other organs such as kidney or intestine, MRP1 is involved in the efflux of multiple endogenous compounds such as glutathione and xenobiotics including SN-38 [36-39]. First, lower *ABCC1* expression in the liver may be associated with reduced efflux to systemic circulation. MRP1 is also involved in the efflux of glutathione and reduced level of glutathione were linked to an increased sensitivity of colon cancer cells to 5-FU [40]. Thus, a reduced expression of MRP1 may increase glutathione levels and limit 5-FU efficacy. Similarly, high *ABCC1* expression in tumors was associated with longer survival in stage IV head and neck squamous cell carcinoma patients [41]. Also, locally advanced esophageal adenocarcinoma patients with high pretherapeutic *ABCC1* tumor expression were found to better respond to neoadjuvant chemotherapy combining 5-FU and cisplatin with or without paclitaxel [42]. Finally, higher expression of *ABCC1* was shown to promote tumor cell death by increasing oxidative stress through a reduction of intracellular glutathione levels [43].

A third variant, *UGT1* rs1113193G>A, was linked to shorter survival using a recessive model with HR of 2.00 (*P* = 0.040), 2.47 (*P* = 0.032) and 2.12 (*P* = 0.004) in the discovery, replication and combined cohorts (**Figure 1A**). Kaplan-Meier showed that homozygous carriers of the rs1113193A allele had a reduced OS by 10 months (13 months versus 23 months, *P* = 0.005, **Figure 1D**). The reporter gene assays with rs1113193G>A and its tightly linked rs7608713G>A variant did not reveal change of gene expression in the

condition tested (**Supplementary Figure 3B-C**). The *UGT1* locus encodes nine UGT1A functional enzymes including UGT1A1 conjugating bilirubin, also responsible for the glucuronidation of SN-38 [44-47]. Supportive of a potential functional impact of this marker or a linked variant, the *UGT1* rs1113193G>A was associated with altered bilirubin level in individuals of the Framingham Heart Study (European descent subjects) and in a study combining individuals from the Framingham Heart Study, the Rotterdam study and the Age, Gene, Environment and Susceptibility-Reykjavik study [48, 49].

Finally, we conducted a combined analysis of *CESI* rs9921399, *ABCC1* rs17501011 and *UGT1* rs1113193 markers and included a marker (*RPL28* rs4806668) recently reported by our group to be associated with survival of mCRC patients [50]. In the combined cohort, carriers of two or three unfavourable genotypes displayed shorter survival with HR values of 2.17 ($P < 0.001$) and 5.35 ($P = 0.004$), respectively, compared to carriers of one unfavourable genotype, whereas carriers of all favourable genotypes experienced prolonged survival (**Figure 3A**). A haplotype analysis confirmed an enhanced effect of these unfavourable markers on survival with HR values of 1.98 ($P = 0.038$) to 2.97 ($P = 0.018$) for carriers of unfavourable genotypes whereas HR for non-carriers was 0.31 ($P = 0.002$) as shown in **Figure 3B**.

Further studies are needed to prospectively replicate these associations. We acknowledge that no correction for multiple testing was applied due to the exploratory nature of this study. However, markers were investigated in two independent cohorts, thus reducing the possibility of false-positive associations. We conclude that the combined analysis of polymorphisms in genes related to irinotecan pharmacology may eventually help predict survival of mCRC patients.

Executive summary

-We performed a systematic analysis of 285 htSNPs in 11 genes related to irinotecan pharmacology in relation to overall survival (OS) in two independent cohorts of 417 metastatic colorectal cancer (mCRC) patients receiving first-line FOLFIRI treatment.

-*CES1* rs9921399T>C was predictive of prolonged OS whereas *ABCC1* rs17501011G>A and *UGT1* rs1113193G>A were associated with reduced OS.

-*ABCC1* rs17501011A allele caused reduced reporter-gene activity in liver ($P < 0.01$) and colon ($P < 0.05$) cells whereas lower *ABCC1* tumor expression in mCRC patients was associated with reduced OS in two independent cohorts.

-A haplotype analysis supports a combined effect of these polymorphisms on survival in mCRC patients.

-This study identifies common variants in irinotecan-related pathways that may contribute to survival of patients with colorectal cancer and requires further validation.

References

1. Noone AM, Howlander N, Krapcho M *et al.* SEER Cancer Statistics Review, 1975-2015, National Cancer Institute, Bethesda, MD, based on November 2017 SEER data submission, posted to the SEER web site, April 2018 (https://seer.cancer.gov/csr/1975_2015/).
2. Sanchez-Gundin J, Fernandez-Carballido AM, Martinez-Valdivieso L, Barreda-Hernandez D, Torres-Suarez AI. New Trends in the Therapeutic Approach to Metastatic Colorectal Cancer. *Int J Med Sci* 15(7), 659-665 (2018).
3. Drugs Approved for Colon and Rectal Cancer (Updated: August 15, 2018), accessed on April 2019 (<https://www.cancer.gov/about-cancer/treatment/drugs/colorectal>).
4. Modest DP, Pant S, Sartore-Bianchi A. Treatment sequencing in metastatic colorectal cancer. *Eur J Cancer* 109 70-83 (2019).
5. Sandhu J, Lavingia V, Fakih M. Systemic treatment for metastatic colorectal cancer in the era of precision medicine. *J Surg Oncol* 119(5), 564-582 (2019).
6. Bignucolo A, De Mattia E, Cecchin E, Roncato R, Toffoli G. Pharmacogenomics of Targeted Agents for Personalization of Colorectal Cancer Treatment. *Int J Mol Sci* 18(7), (2017).
7. Mizukami T, Izawa N, Nakajima TE, Sunakawa Y. Targeting EGFR and RAS/RAF Signaling in the Treatment of Metastatic Colorectal Cancer: From Current Treatment Strategies to Future Perspectives. *Drugs* doi:10.1007/s40265-019-01113-0 (2019).
8. Patel JN, Fong MK, Jagosky M. Colorectal Cancer Biomarkers in the Era of Personalized Medicine. *J Pers Med* 9(1), (2019).
9. Yu Q, Zhang T, Xie C *et al.* UGT1A polymorphisms associated with worse outcome in colorectal cancer patients treated with irinotecan-based chemotherapy. *Cancer Chemother Pharmacol* 82(1), 87-98 (2018).
10. Cecchin E, Innocenti F, D'andrea M *et al.* Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. *J Clin Oncol* 27(15), 2457-2465 (2009).
11. Shulman K, Cohen I, Barnett-Griness O *et al.* Clinical implications of UGT1A1*28 genotype testing in colorectal cancer patients. *Cancer* 117(14), 3156-3162 (2011).
12. Lu CY, Huang CW, Wu IC *et al.* Clinical Implication of UGT1A1 Promoter Polymorphism for Irinotecan Dose Escalation in Metastatic Colorectal Cancer Patients Treated with Bevacizumab Combined with FOLFIRI in the First-line Setting. *Transl Oncol* 8(6), 474-479 (2015).
13. Xu C, Tang X, Qu Y, Keyoumu S, Zhou N, Tang Y. UGT1A1 gene polymorphism is associated with toxicity and clinical efficacy of irinotecan-based chemotherapy in patients with advanced colorectal cancer. *Cancer Chemother Pharmacol* 78(1), 119-130 (2016).
14. Kim SY, Y SH, E KS *et al.* S-1 plus irinotecan and oxaliplatin for the first-line treatment of patients with metastatic colorectal cancer: a prospective phase II study and pharmacogenetic analysis. *Br J Cancer* 109(6), 1420-1427 (2013).

15. De Mattia E, Toffoli G, Polesel J *et al.* Pharmacogenetics of ABC and SLC transporters in metastatic colorectal cancer patients receiving first-line FOLFIRI treatment. *Pharmacogenet Genomics* 23(10), 549-557 (2013).
16. Glimelius B, Garmo H, Berglund A *et al.* Prediction of irinotecan and 5-fluorouracil toxicity and response in patients with advanced colorectal cancer. *Pharmacogenomics J* 11(1), 61-71 (2011).
17. Akiyama Y, Fujita K, Ishida H *et al.* Association of ABCC2 genotype with efficacy of first-line FOLFIRI in Japanese patients with advanced colorectal cancer. *Drug Metab Pharmacokinet* 27(3), 325-335 (2012).
18. Levi F, Karaboue A, Saffroy R *et al.* Pharmacogenetic determinants of outcomes on triplet hepatic artery infusion and intravenous cetuximab for liver metastases from colorectal cancer (European trial OPTILIV, NCT00852228). *Br J Cancer* 117(7), 965-973 (2017).
19. Teft WA, Welch S, Lenehan J *et al.* OATP1B1 and tumour OATP1B3 modulate exposure, toxicity, and survival after irinotecan-based chemotherapy. *Br J Cancer* 112(5), 857-865 (2015).
20. Huang L, Zhang T, Xie C *et al.* SLC01B1 and SLC19A1 gene variants and irinotecan-induced rapid response and survival: a prospective multicenter pharmacogenetics study of metastatic colorectal cancer. *PLoS One* 8(10), e77223 (2013).
21. Levesque E, Belanger AS, Harvey M *et al.* Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J Pharmacol Exp Ther* 345(1), 95-101 (2013).
22. Toffoli G, Cecchin E, Corona G *et al.* The role of UGT1A1*28 polymorphism in the pharmacodynamics and pharmacokinetics of irinotecan in patients with metastatic colorectal cancer. *J Clin Oncol* 24(19), 3061-3068 (2006).
23. Chen S, Laverdiere I, Tourancheau A *et al.* A novel UGT1 marker associated with better tolerance against irinotecan-induced severe neutropenia in metastatic colorectal cancer patients. *Pharmacogenomics J* 15(6), 513-520 (2015).
24. Chen S, Villeneuve L, Jonker D *et al.* ABCC5 and ABCG1 polymorphisms predict irinotecan-induced severe toxicity in metastatic colorectal cancer patients. *Pharmacogenet Genomics* 25(12), 573-583 (2015).
25. Purcell S, Neale B, Todd-Brown K *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3), 559-575 (2007).
26. Boyle AP, Hong EL, Hariharan M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 22(9), 1790-1797 (2012).
27. Ross MK, Borazjani A, Wang R, Crow JA, Xie S. Examination of the carboxylesterase phenotype in human liver. *Arch Biochem Biophys* 522(1), 44-56 (2012).
28. Merali Z, Ross S, Pare G. The pharmacogenetics of carboxylesterases: CES1 and CES2 genetic variants and their clinical effect. *Drug Metabol Drug Interact* 29(3), 143-151 (2014).
29. Hatfield MJ, Tsurkan L, Garrett M *et al.* Organ-specific carboxylesterase profiling identifies the small intestine and kidney as major contributors of activation of the anticancer prodrug CPT-11. *Biochem Pharmacol* 81(1), 24-31 (2011).

30. Sanghani SP, Quinney SK, Fredenburg TB, Davis WI, Murry DJ, Bosron WF. Hydrolysis of irinotecan and its oxidative metabolites, 7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino] carbonyloxycamptothecin and 7-ethyl-10-[4-(1-piperidino)-1-amino]-carbonyloxycamptothecin, by human carboxylesterases CES1A1, CES2, and a newly expressed carboxylesterase isoenzyme, CES3. *Drug Metab Dispos* 32(5), 505-511 (2004).
31. Sai K, Saito Y, Tatewaki N *et al.* Association of carboxylesterase 1A genotypes with irinotecan pharmacokinetics in Japanese cancer patients. *Br J Clin Pharmacol* 70(2), 222-233 (2010).
32. Li M, Seiser EL, Baldwin RM *et al.* ABC transporter polymorphisms are associated with irinotecan pharmacokinetics and neutropenia. *Pharmacogenomics J* doi:10.1038/tpj.2016.75 (2016).
33. Innocenti F, Kroetz DL, Schuetz E *et al.* Comprehensive pharmacogenetic analysis of irinotecan neutropenia and pharmacokinetics. *J Clin Oncol* 27(16), 2604-2614 (2009).
34. Lu JF, Pokharel D, Bebawy M. MRP1 and its role in anticancer drug resistance. *Drug Metab Rev* 47(4), 406-419 (2015).
35. Kunicka T, Soucek P. Importance of ABCC1 for cancer therapy and prognosis. *Drug Metab Rev* 46(3), 325-342 (2014).
36. Hilgendorf C, Ahlin G, Seithel A, Artursson P, Ungell AL, Karlsson J. Expression of thirty-six drug transporter genes in human intestine, liver, kidney, and organotypic cell lines. *Drug Metab Dispos* 35(8), 1333-1340 (2007).
37. Flens MJ, Zaman GJ, Van Der Valk P *et al.* Tissue distribution of the multidrug resistance protein. *Am J Pathol* 148(4), 1237-1247 (1996).
38. Roelofsen H, Vos TA, Schippers IJ *et al.* Increased levels of the multidrug resistance protein in lateral membranes of proliferating hepatocyte-derived cells. *Gastroenterology* 112(2), 511-521 (1997).
39. Chen ZS, Furukawa T, Sumizawa T *et al.* ATP-Dependent efflux of CPT-11 and SN-38 by the multidrug resistance protein (MRP) and its inhibition by PAK-104P. *Mol Pharmacol* 55(5), 921-928 (1999).
40. Palomares T, Carames M, Garcia-Alonso I, Alonso-Varona A. Glutathione modulation reverses the growth-promoting effect of growth factors, improving the 5-fluorouracil antitumour response in WiDr colon cancer cells. *Anticancer Res* 29(10), 3957-3965 (2009).
41. Warta R, Theile D, Mogler C *et al.* Association of drug transporter expression with mortality and progression-free survival in stage IV head and neck squamous cell carcinoma. *PLoS One* 9(9), e108908 (2014).
42. Langer R, Specht K, Becker K *et al.* Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin Cancer Res* 11(20), 7462-7469 (2005).
43. Laberge RM, Karwatsky J, Lincoln MC, Leimanis ML, Georges E. Modulation of GSH levels in ABCC1 expressing tumor cells triggers apoptosis through oxidative stress. *Biochem Pharmacol* 73(11), 1727-1737 (2007).
44. Iyer L, King CD, Whittington PF *et al.* Genetic predisposition to the metabolism of irinotecan (CPT-11). Role of uridine diphosphate glucuronosyltransferase isoform 1A1 in the glucuronidation of its active metabolite (SN-38) in human liver microsomes. *J Clin Invest* 101(4), 847-854 (1998).

45. Hanioka N, Ozawa S, Jinno H, Ando M, Saito Y, Sawada J. Human liver UDP-glucuronosyltransferase isoforms involved in the glucuronidation of 7-ethyl-10-hydroxycamptothecin. *Xenobiotica* 31(10), 687-699 (2001).
46. Gagne JF, Montminy V, Belanger P, Journault K, Gaucher G, Guillemette C. Common human UGT1A polymorphisms and the altered metabolism of irinotecan active metabolite 7-ethyl-10-hydroxycamptothecin (SN-38). *Mol Pharmacol* 62(3), 608-617 (2002).
47. Guillemette C, Levesque E, Rouleau M. Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* 96(3), 324-339 (2014).
48. Johnson AD, Kavousi M, Smith AV *et al.* Genome-wide association meta-analysis for total serum bilirubin levels. *Hum Mol Genet* 18(14), 2700-2710 (2009).
49. Lin JP, Schwaiger JP, Cupples LA *et al.* Conditional linkage and genome-wide association studies identify UGT1A1 as a major gene for anti-atherogenic serum bilirubin levels--the Framingham Heart Study. *Atherosclerosis* 206(1), 228-233 (2009).
50. Labriet A, Lévesque É, Cecchin E *et al.* Germline variability and tumor expression level of ribosomal protein gene RPL28 influence survival of metastatic colorectal cancer patients *submitted* (2019).

Conflict of interest statement

The authors declare no conflict of interest.

Funding

This work is supported by the Canadian Institutes of Health Research (CIHR, FRN-42392) and a Canada Research Chair in Pharmacogenomics (Tier I) to CG. AL is a recipient of studentship awards from the 'Fonds d'Enseignement et de Recherche (FER)' from the Faculty of Pharmacy at Laval University and from the 'Fonds de Recherche du Québec - Santé'. EL is a recipient of a CIHR clinical-scientist phase II award and Prostate Cancer Canada Rising Star Award (RS2013-55).

Acknowledgements

We would like to thank all the participants of this study. A fraction of the presented results was based upon data generated by the TCGA Research Network (<http://cancergenome.nih.gov/>).

Table 1. Characteristics of the Canadian and Italian cohorts of metastatic colorectal cancer patients.

	Canadian cohort	Italian cohort
	n (%)	n (%)
Characteristics	N = 167	N = 250
Gender		
Male	110 (66)	162 (65)
Female	57 (34)	88 (35)
Age (years)		
Mean \pm Standard Deviation	61.5 \pm 10.2	60.6 \pm 10.3
Range	29 - 86	26 - 75
FOLFIRI	167 (100)	250 (100)
Co-treatment	75 (45)	0
Bevacizumab	69 (92)	0
Other drug	6 (8)	0
Overall survival (median in months)	24	15

Figure Legends.

Figure 1. Polymorphisms associated with overall survival of metastatic colorectal cancer patients. (A) Cox proportional hazards models were adjusted for age and co-treatment in Canadian and combined cohorts, adjusted for age in Italian cohort. (B-D) Kaplan-Meier survival curves and log-rank *P*-values for *CESI* rs9921399, *ABCC1* rs17501011 and *UGT1* rs1113193. CI, Confidence Interval; Comb., combined cohorts; HR_{adj}, adjusted hazard ratio.

Figure 2. Functional impact of the *ABCC1* rs17501011 and prognostic value of *ABCC1* gene expression. (A) The *ABCC1* rs17501011A alters transcriptional activity in liver cancer cells HepG2 and HuH7 as well as in colon cancer cells HCT116 and HT-29. (B-C) Influence of *ABCC1* colorectal tumor expression on survival of mCRC patients from the TCGA (n = 87) and the GSE39582 (n = 60) datasets.

* *P* < 0.05, ** *P* < 0.01.

Figure 3. Co-occurrence of markers linked to reduced survival. (A) Carriers of several markers had a reduced overall survival compared to the reference group carrying only one marker. (B) Haplotype analysis confirmed that carriers of multiple markers had a shorter overall survival. Genotypes in yellow are associated to reduced survival, those in green to longer survival. Patients of the Canadian and Italian cohorts were combined. Only genotype combinations carried by at least 3 patients were conserved. CI, Confidence Interval; HR_{adj}, adjusted hazard ratio.

Figure 1.

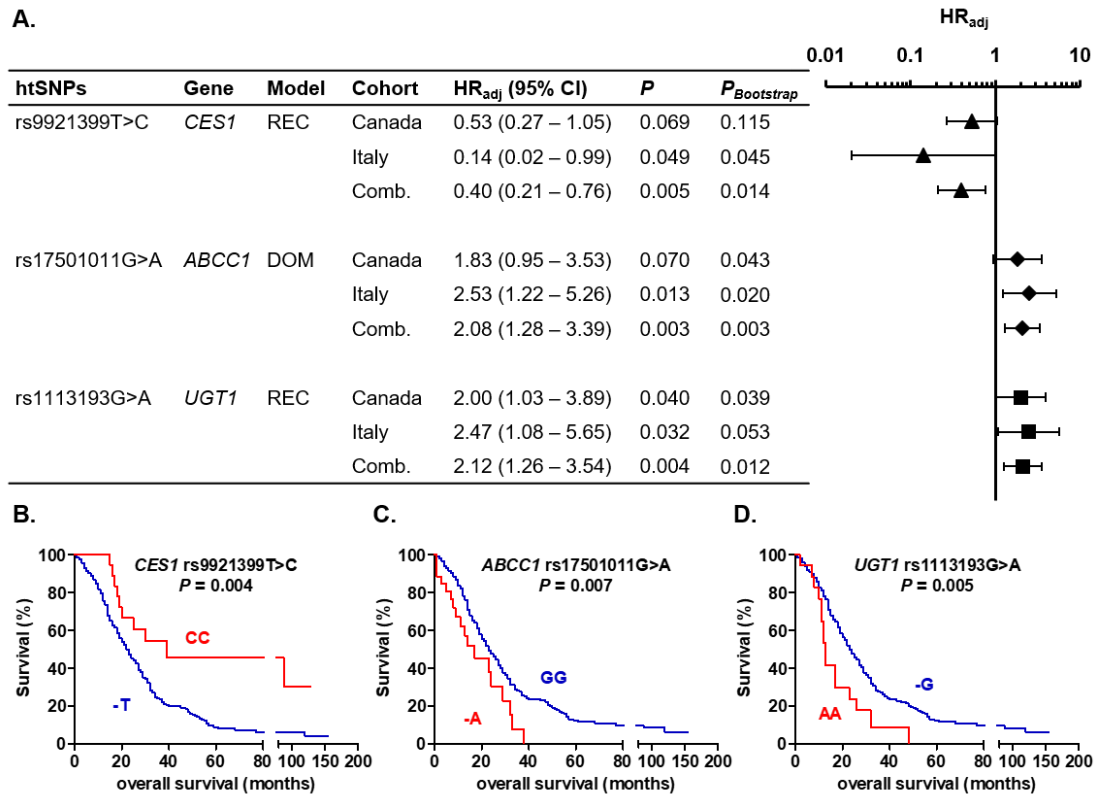
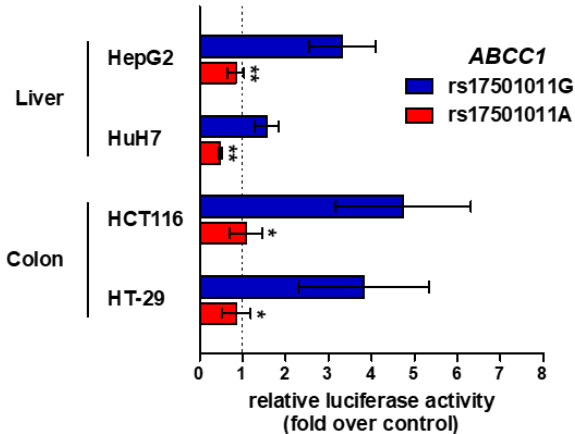
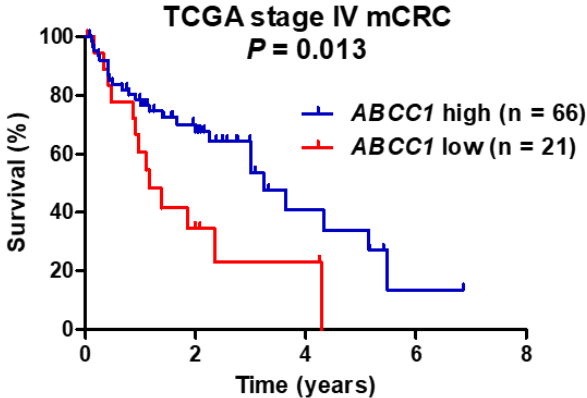


Figure 2.

A.



B.



C.

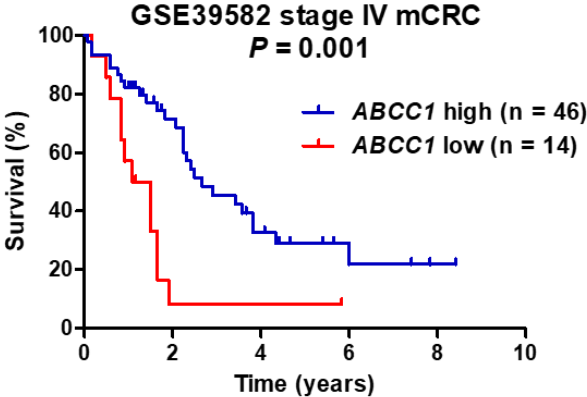
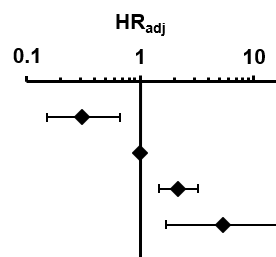


Figure 3.

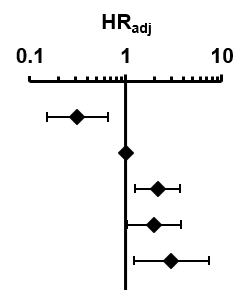
A.

Risk genotypes	n (%)	HR _{adj} (95% CI)	P	P _{Bootstrap}
0	17 (4.6)	0.31 (0.15 – 0.66)	0.002	0.003
1	313 (83.9)	1.00 (reference)	--	--
2	40 (10.7)	2.17 (1.46 – 3.21)	<0.001	0.001
3	3 (0.8)	5.35 (1.68 – 17.00)	0.004	0.001



B.

<i>CES1</i> rs9921399	<i>ABCC1</i> rs17501011	<i>UGT1</i> rs1113193	<i>RPL28</i> rs4806668	n (%)	HR _{adj} (95% CI)	P	P _{Bootstrap}
CC	GG	GG+GA	GG+GT	17 (4.6)	0.31 (0.15 – 0.66)	0.002	0.006
TC+TT	GG	GG+GA	GG+GT	311 (84.5)	1.00 (reference)	--	--
TC+TT	GA+AA	GG+GA	GG+GT	23 (6.3)	2.15 (1.26 – 3.68)	0.005	0.004
TC+TT	GG	AA	GG+GT	12 (3.2)	1.98 (1.04 – 3.77)	0.038	0.048
TC+TT	GG	GG+GA	TT	5 (1.4)	2.97 (1.21 – 7.29)	0.018	0.001



Supplementary Table 1. Haplotype-tagging single nucleotide polymorphisms (htSNPs) in the 11 candidate genes.

Genes	htSNPs
<i>ABCB1</i>	rs2235013, rs2235033, rs10276036, rs2235015, rs11983225, rs1202180, rs10259849, rs4148735, rs7802783, rs10267099, rs1045642, rs2032588, rs4148733, rs4148732, rs2214102, rs6465118, rs4728701, rs1882478, rs6949448, rs13233308, rs3789243, rs1202184, rs6946119, rs9282564, rs12720066, rs17327624, rs10264990, rs1128503, rs2032582, rs3842.
<i>ABCC1</i>	rs35588, rs212088, rs3765129, rs35605, rs9635480, rs215074, rs35629, rs12927980, rs215064, rs8059648, rs924136, rs171583, rs4148373, rs4148361, rs4148358, rs246217, rs215088, rs212083, rs4781699, rs212085, rs4148355, rs17264736, rs4781711, rs6498594, rs11075295, rs11075289, rs35626, rs246230, rs35628, rs9932506, rs2238475, rs2283512, rs3784862, rs35592, rs215100, rs246214, rs152033, rs8058040, rs12922404, rs212091, rs875740, rs8187843, rs3784867, rs4148330, rs215101, rs8053988, rs152023, rs8054670, rs212079, rs8187858, rs169985, rs17205859, rs10852377, rs215108, rs215098, rs11864374, rs4148382, rs4148380, rs3888565, rs17501011, rs2889517, rs246233, rs903880, rs3887893, rs212081, rs12927259, rs4148348, rs17501331, rs212093, rs17287570, rs4148350, rs2230671, rs2074087.
<i>ABCC2</i>	rs1885301, rs2804402, rs17216177, rs4148397, rs2804398, rs7067971, rs2273697, rs2002042, rs3740063, rs717620, rs2756109, rs8187710, rs3740066.
<i>ABCC5</i>	rs562, rs10937158, rs4148573, rs1464322, rs2292997, rs3792583, rs3749438, rs13093614, rs1554395, rs1016752, rs9838667, rs2176825.
<i>ABCG1</i>	rs425215, rs1117640, rs225378, rs7281684, rs3787970, rs915847, rs9978671, rs225440, rs914189, rs225406, rs9975333, rs6586304, rs225410, rs4148137, rs532345, rs7281093, rs3787997, rs225398, rs221948, rs1044317, rs1541290, rs225376, rs3787995, rs450808, rs3788007, rs17767083, rs4148117, rs183436, rs225391, rs182694, rs2234718, rs474142, rs13050646, rs225374.
<i>ABCG2</i>	rs2231137, rs4693924, rs2725246, rs17731538, rs2231142, rs13120400, rs2622604,

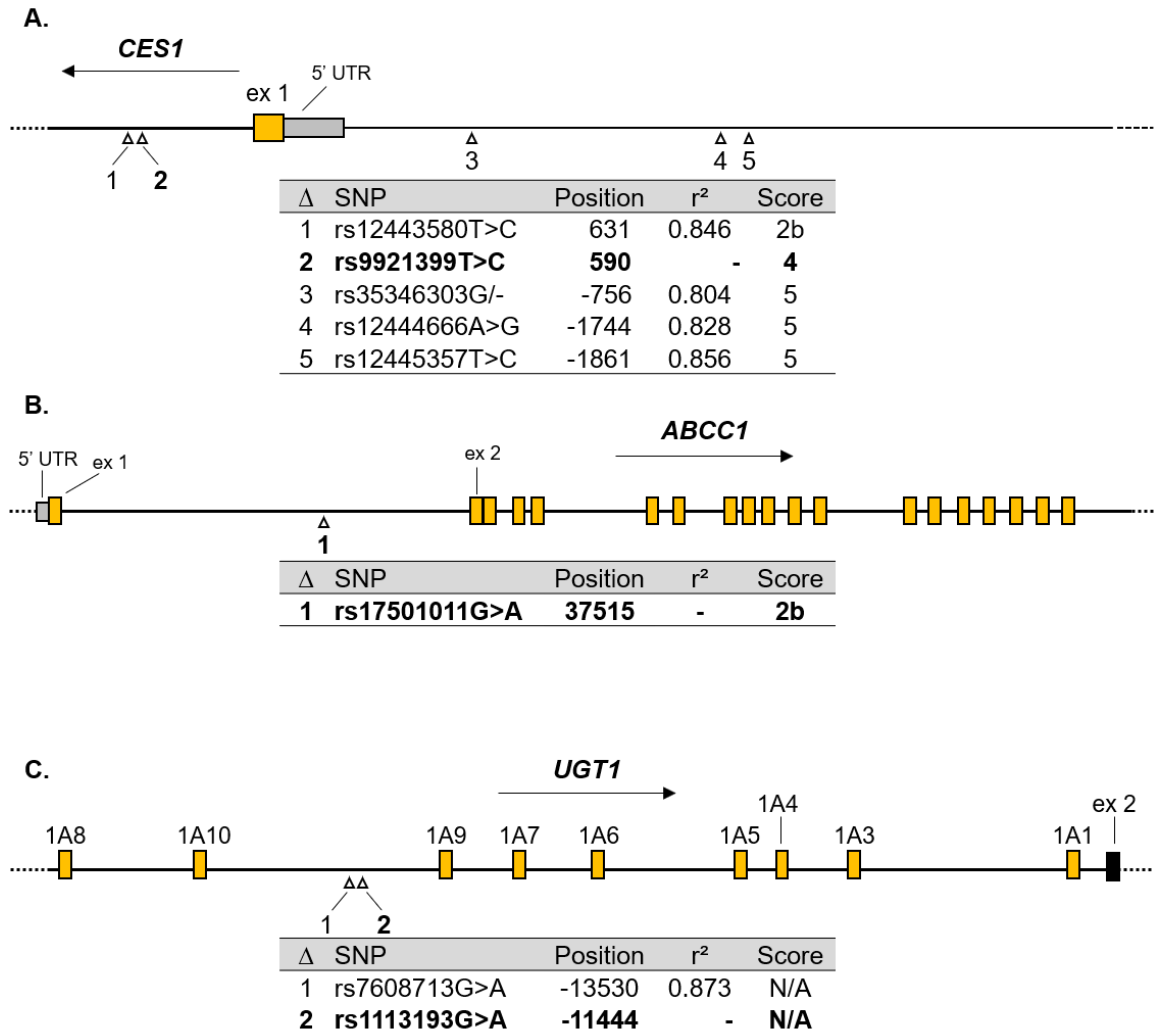
	rs2231148, rs2725256, rs2725264, rs2725270, rs2622626, rs12505410, rs17013881, rs3114018, rs2725252, rs2231135, rs2622621, rs9999111, rs2725248, rs2231156.
<i>CES1</i>	rs7498748, rs1814268, rs9921399, rs2397787, rs8192950, rs12443580, rs6499786, rs28709456.
<i>CES2</i>	rs7202523, rs4783745.
<i>SLCO1B1</i>	rs11045821, rs4363657, rs6487213, rs10841769, rs2291076, rs2010668, rs11045797, rs7139376, rs2900476, rs12830367, rs7953338, rs12427008, rs2199766, rs2417955, rs4149070, rs991262, rs16923519, rs4149057, rs4149035, rs4149033, rs11045834, rs12371604, rs1564365, rs4149022, rs4149015, rs4149056, rs2306283.
<i>TOPI</i>	rs6129750, rs6072249, rs6072269, rs17179419, rs6129758, rs6129760, rs8124620, rs4142394, rs1997833, rs6016505, rs4314052, rs2076574, rs8122094, rs11699577, rs6072262, rs79779579.
<i>UGT1</i>	rs1597942, rs1377460, rs1823803, rs2741034, rs11892031, rs7571337, rs1112310, rs1113193, rs6706988, rs17868320, rs2741044, rs3806598, rs10176426, rs2741045, rs6714486, rs3832043, rs2741048, rs2741049, rs17862857, rs2602381, rs7586110, rs17868323, rs17863778, rs11692021, rs17862859, rs6759892, rs2070959, rs1105879, rs6751673, rs17863787, rs4663326, rs7583278, rs2011404, rs28899187, rs4663965, rs11568318, rs4124874, rs10929302, rs8175347, rs28946889, rs3771342, rs2302538, rs4148328, rs10929303, rs1042640, rs8330, rs6717546, rs11563250, rs6719561.

Supplementary Table 2. Oligonucleotides used in this study.

Description of oligonucleotide	Oligonucleotide sequence 5' → 3'
Forward oligo for PCR amplification of rs12443580 + introduction of SacI restriction site for cloning	CTAGCTGGAGCTCTAGTGGCCTGACCACAGCTC
Reverse oligo for PCR amplification of rs12443580 + introduction of XhoI restriction site for cloning	CTAGCTGCTCGAGGAAACAGAGGTGGCAGTGCT
Forward oligo for PCR amplification of rs17501011 + introduction of SacI restriction site for cloning	CTAGCTGGAGCTCAGCAAGAAGTGGATCTGGAAG
Reverse oligo for PCR amplification of rs17501011 + introduction of XhoI restriction site for cloning	CTAGCTGCTCGAGACCTCAGTTTCCGCATCTGT
Forward oligo for PCR amplification of rs1113193 + introduction of SacI restriction site for cloning	CTAGCTGGAGCTCGATGGCTAGAGGCATGTCC
Reverse oligo for PCR amplification of rs1113193 + introduction of XhoI restriction site for cloning	CTAGCTGCTCGAGCCAGCTCCCAGTCTTTTGT
Forward oligo for PCR amplification of rs7608713 + introduction of SacI restriction site for cloning	CTAGCTGGAGCTCCCCATCCTCCCCTATGGTTA
Reverse oligo for PCR amplification of rs7608713 + introduction of XhoI restriction site for cloning	CTAGCTGCTCGAGTCATTTGGTATCCAGGATGG
Forward oligo for site-directed mutagenesis for rs12443580	TATTTATCCAcCCCTTTTTCGTG

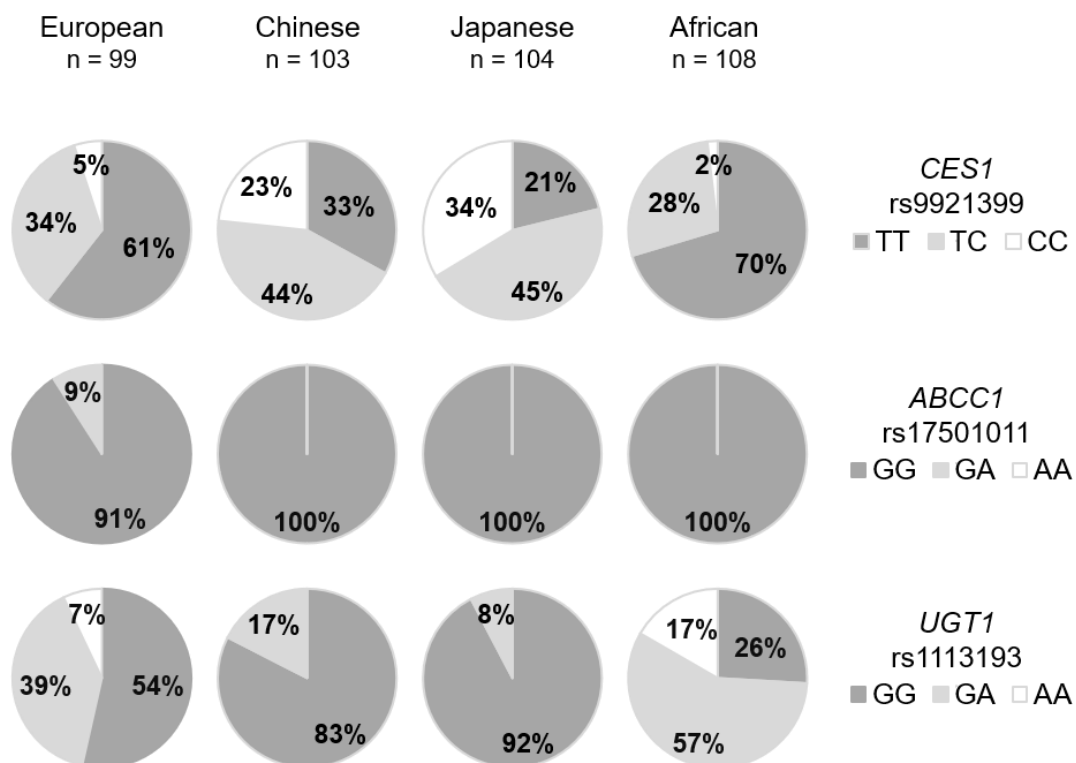
Reverse oligo for site-directed mutagenesis for rs12443580	AAGGCGCTGTTTAAAAAC
Forward oligo for site-directed mutagenesis for rs17501011	TCTGGCTCATATTTAACAAAGTGC
Reverse oligo for site-directed mutagenesis for rs17501011	GCACACAGTAGGTGCTCA
Forward oligo for site-directed mutagenesis for rs1113193	GAGACAAAGCaTTGGTTTTTCATG
Reverse oligo for site-directed mutagenesis for rs1113193	TGAAATCAGAAATACCTTGC

Supplementary Figure 1.



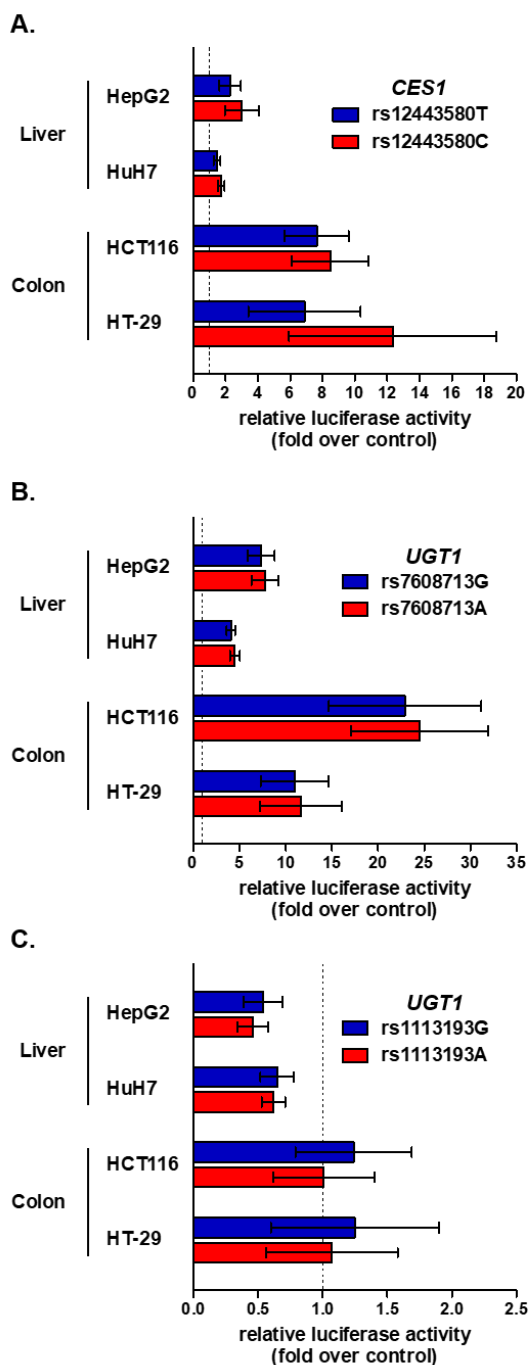
Supplementary Figure 1. (A-C) Localization relative to the translation start site of htSNPs (in bold) linked to overall survival and their associated SNPs in high linkage disequilibrium ($r^2 > 0.800$). RegulomeDB scores are shown and represent the probability for a SNP to be functional. N/A, not available. Only SNPs with score $\geq 2b$ were tested in the luciferase reporter gene assays, except for *UGT1* where both SNPs were tested.

Supplementary Figure 2.



Supplementary Figure 2. Genotypic frequencies of markers of overall survival to mCRC among several populations: European, CEU (Utah Residents with Northern and Western European Ancestry); Chinese, CHB (Han Chinese in Beijing, China); Japanese, JPT (Japanese in Tokyo, Japan) and African, YRI (Yoruba in Ibadan, Nigeria). Data were retrieved through Ensembl GRCh37 release 91. Frequency of rs4806668G>T GG, GT and TT genotypes were respectively 77%, 20% and 3% in CEU ; 72%, 27% and 1% in CHB; 76%, 22% and 2% in JPT; 9%, 33% and 58% in YRI.

Supplementary Figure 3.



Supplementary Figure 3. Luciferase reporter gene assays. No significant impact on relative luciferase activity was observed for **(A)** *CES1* rs12443580T>C, **(B)** *UGT1* rs7608713G>A and **(C)** *UGT1* rs1113193G>A.

Chapitre 2 : Germline variability and tumor expression level of ribosomal protein gene *RPL28* are associated with survival of metastatic colorectal cancer patients

2.1 Résumé

Cette étude a permis d'investiguer le potentiel de polymorphismes d'une seule paire de bases comme marqueurs prédictifs de la survie de patients atteints d'un CCRm traité avec le régime FOLFIRI (n = 417). Les gènes candidats que nous avons sélectionnés, incluant *RPL28*, ont été choisis car ils ont été associés dans des études *in vitro* ou *in vivo* à la réponse au traitement. Le rs4806668G>T du gène de la protéine ribosomique L28 (*RPL28*) est associé à une réduction de la survie sans progression et de la survie globale de 5 et 9 mois ($P = 0.002$) avec des HR de 3,36 ($P < 0.001$) et 3,07 ($P = 0.002$), respectivement. L'allèle variant rs4806668T est associé à une hausse de l'expression de *RPL28* dans des tissus normaux de côlon transverse (n = 246, $P = 0.007$). L'expression de *RPL28* est plus élevée jusqu'à 124% ($P < 0.001$) dans les tumeurs colorectales comparativement aux tissus normaux pairés provenant de trois cohortes indépendantes totalisant 78 cas. Dans deux cohortes indépendantes, les cas de CCRm ayant le plus haut niveau d'expression de *RPL28* ont une survie réduite (n = 88, $P = 0.004$ et n = 56, $P < 0.001$). Une forte expression de *RPL28* a aussi été associée à des changements dans l'expression de gènes de la voie des immunoglobulines et de la matrice extracellulaire. La répression de *RPL28*, quant à elle, réduit la prolifération des cellules de cancer du côlon de 1,4 à 5,6 fois ($P < 0.05$). L'ensemble de ces résultats suggère que la protéine ribosomique L28 a le potentiel d'influencer la survie des patients atteints d'un cancer colorectal métastatique.

2.2 Abstract

This study investigated the potential of single nucleotide polymorphisms as predictors of survival in two cohorts comprising 417 metastatic colorectal cancer (mCRC) patients treated with the FOLFIRI (folinic acid, 5-fluorouracil and irinotecan) regimen. The rs4806668G>T of the ribosomal protein gene *RPL28* was associated with shorter progression-free survival and overall survival by 5 and 9 months ($P = 0.002$), with hazard ratios of 3.36 ($P < 0.001$) and 3.07 ($P = 0.002$), respectively. The rs4806668T allele was associated with an increased *RPL28* expression in transverse normal colon tissues ($n = 246$, $P = 0.007$). *RPL28* expression was higher in colorectal tumors compared to paired normal tissues by up to 124% ($P < 0.001$) in three independent datasets. Metastatic cases with highest *RPL28* tumor expression had a reduced survival in two datasets ($n = 88$, $P = 0.004$ and $n = 56$, $P < 0.001$). High *RPL28* was further associated with changes in immunoglobulin and extracellular matrix pathways. Repression of *RPL28* reduced proliferation by 1.4-fold to 5.6-fold ($P < 0.05$) in colon cancer HCT116 and HT-29 cells. Our findings suggest that the ribosomal *RPL28* protein may influence mCRC outcome.

Germline variability and tumor expression level of ribosomal protein gene *RPL28* are associated with survival of metastatic colorectal cancer patients

Adrien Labriet¹, Éric Lévesque², Erika Cecchin³, Elena De Mattia³, Lyne Villeneuve¹, Michèle Rouleau¹, Derek Jonker⁴, Félix Couture², David Simonyan⁵, Eric P. Allain¹, Angela Buonadonna⁶, Mario D'Andrea⁷, Giuseppe Toffoli³ and Chantal Guillemette^{1,8*}

¹Pharmacogenomics Laboratory, Centre Hospitalier Universitaire de Québec (CHU de Québec) Research Center – Laval University and Faculty of Pharmacy, Laval University, Québec, Québec, Canada; ²CHU de Québec Research Center and Faculty of Medicine, Laval University, Québec, Québec, Canada; ³Clinical and Experimental Pharmacology, “Centro di Riferimento Oncologico”- National Cancer Institute, via Franco Gallini n. 2, 33081 Aviano (PN), Italy; ⁴Division of Medical Oncology, Department of Medicine, Ottawa Hospital, University of Ottawa, Ottawa, Ontario, Canada; ⁵Clinical and Evaluative Research Platform, CHU de Québec Research Center, Québec, Canada; ⁶Medical Oncology Unit, “Centro di Riferimento Oncologico”- National Cancer Institute, via Franco Gallini n. 2, 33081, Aviano (PN), Italy; ⁷Medical Oncology Unit, “San Filippo Neri Hospital”, Via Giovanni Martinotti, 20, 00135 Rome, Italy.⁸Canada Research Chair in Pharmacogenomics, CHU de Québec Research Center, Québec, Québec, Canada.

***CORRESPONDING AUTHOR:** Chantal Guillemette, Ph.D., CHU de Québec Research Center, R4701.5, 2705 Blvd. Laurier, Québec, Canada, G1V 4G2. E-mail: chantal.guillemette@crchudequebec.ulaval.ca.

Abstract: 187 words

Introduction: 298 words

Material and Methods: 1118 words

Results: 820 words

Discussion: 1224 words

Body text: 3162 words

References: 50

NUMBER OF:

Tables: 1

Figures: 5

Supplementary Tables: 6

Supplementary Figures: 2

ABSTRACT

This study investigated the potential of single nucleotide polymorphisms as predictors of survival in two cohorts comprising 417 metastatic colorectal cancer (mCRC) patients treated with the FOLFIRI (folinic acid, 5-fluorouracil and irinotecan) regimen. The rs4806668G>T of the ribosomal protein gene *RPL28* was associated with shorter progression-free survival and overall survival by 5 and 9 months ($P = 0.002$), with hazard ratios of 3.36 ($P < 0.001$) and 3.07 ($P = 0.002$), respectively. The rs4806668T allele was associated with an increased *RPL28* expression in transverse normal colon tissues ($n = 246$, $P = 0.007$). *RPL28* expression was higher in colorectal tumors compared to paired normal tissues by up to 124% ($P < 0.001$) in three independent datasets. Metastatic cases with highest *RPL28* tumor expression had a reduced survival in two datasets ($n = 88$, $P = 0.004$ and $n = 56$, $P < 0.001$). High *RPL28* was further associated with changes in immunoglobulin and extracellular matrix pathways. Repression of *RPL28* reduced proliferation by 1.4-fold to 5.6-fold ($P < 0.05$) in colon cancer HCT116 and HT-29 cells. Our findings suggest that the ribosomal *RPL28* protein may influence mCRC outcome.

INTRODUCTION

Metastatic colorectal cancer (mCRC) presents a 5-year relative survival just above 10% ¹. There are many treatment options for these patients including irinotecan-based chemotherapy. Particularly, the FOLFIRI regimen is composed of irinotecan (CPT-11) used in combination with 5-fluorouracil (5-FU) and folinic acid alone or with targeted therapies ². Both irinotecan and 5-FU are potent on actively replicating cancer cells. SN-38, the active metabolite of irinotecan, is an inhibitor of topoisomerase I (TOP1). It prevents DNA ligation by directly binding the TOP1-DNA complex, leading to replication arrest, double-strand breaks and cell death ³. 5-FU is a pyrimidine analog that exerts its effect by inhibiting the thymidylate synthetase and DNA synthesis. 5-FU can also be incorporated into RNA during synthesis and interferes with protein synthesis ⁴. The clinical response to FOLFIRI-based regimens is variable with dose-limiting toxicities occurring in a significant proportion of patients ^{5,6}. Several markers in pharmacokinetic pathways have been linked to severe toxicities. For instance, the *UGT1A1**28 polymorphism was established as a predictive marker of severe neutropenia, explained by a decreased *UGT1A1* expression. Because this gene encodes the main hepatic enzyme responsible for SN-38 inactivation and elimination, reduced expression leads to greater exposure to SN-38 and an increased risk of severe neutropenia ^{7,8}. By contrast, clinical genetic-based evidence to predict therapeutic response of mCRC patients is limited ⁹. Several genes were found to be associated with response to irinotecan and 5-FU *in vitro*, but their potential value as markers of treatment response has not been addressed in patients. These genes are related to several pathways such as detoxification of reactive oxygen species, cellular responses to stress, DNA damage recognition or DNA repair. In this study, we tested the hypothesis that genetic variability in these genes may be associated with survival of mCRC patients treated with FOLFIRI-based regimen.

RESULTS

The RPL28 rs4806668G>T variant is associated with survival of mCRC patients.

A total of 105 haplotype-tagging single nucleotide polymorphisms (htSNPs) covering the genetic variability of 17 candidate genes, which were previously associated to drug response *in vitro* or in patients (**Additional file 1: Supplementary Tables 1 and 2**) was performed in the discovery cohort comprising 167 mCRC Canadian patients. We identified 21 and 14 htSNPs associated ($P < 0.1$) with PFS and OS, respectively (**Additional file 1: Supplementary Tables 3 and 4**). In the replication cohort of 250 mCRC Italian patients, one of these htSNPs remained significantly associated ($P < 0.05$) with survival outcomes. The *RPL28* rs4806668G>T variant was associated with a reduced PFS in the Canadian cohort (hazard ratio (HR) = 3.23, $P = 0.013$), the Italian cohort (HR = 3.28, $P = 0.021$) and in combined cohorts (HR = 3.36, $P < 0.001$) using a recessive model (**Figure 1a**). This marker was also linked to shorter OS with HR of 3.09 ($P = 0.032$), 3.05 ($P = 0.030$) and 3.07 ($P = 0.002$) in the Canadian, Italian and combined cohorts, respectively, using a recessive model (**Figure 1b**). Overall, there was 9 carriers of rs4806668TT genotype. Kaplan-Meier analysis revealed that median PFS was reduced by 5 months for homozygous TT carriers compared to heterozygotes or non-carriers (9 months versus 4 months, $P = 0.002$, **Figure 1c**), and OS reduced by 9 months (23 months versus 14 months, $P = 0.002$, **Figure 1d**).

The RPL28 rs4806668G>T variant affects RPL28 gene expression that is increased in tumor tissues.

The *RPL28* rs4806668G>T is located in the 5'-untranslated (5'-UTR) region of the *RPL28* locus and is in strong linkage disequilibrium in the CEU population (European population, LD with $r^2 > 0.80$) with six other SNPs, located upstream of the 5'-UTR (**Figure 2a-b**). Four of these SNPs are predicted to affect transcription factor binding (score 2a or 2b) according to RegulomeDB (**Figure 2b**). Genotypic frequencies of rs4806668G>T was found to be highly variable among ethnic groups with 1-3% of homozygous TT in populations of Asian and European origins and reaching 58% in the African population (**Figure 2c**). To explore the potential impact of rs4806668G>T and its linked SNPs on

RPL28 gene expression, data from healthy donors of the GTEx project were used. An increased *RPL28* expression was observed for carriers of the variant allele rs4806668T ($P = 0.007$) in the transverse colon, as well as for variant alleles of SNPs in linkage disequilibrium (LD) with the rs4806668G>T (**Figure 3a**). By contrast, expression of the nearby gene *TMEM238* was not affected. Data are summarized in **Supplementary Table 5 (Additional file 1)**. In the TCGA cohort, *RPL28* expression was significantly higher by 124 % ($P < 0.001$) in colon tumors compared to paired normal tissues (**Figure 3b**). In mCRC cases from dataset GSE49355, *RPL28* expression was increased by 35 % ($P < 0.01$) in primary tumors and by 34 % ($P < 0.01$) in hepatic metastases compared to paired normal colon tissues (**Figure 3c**). Similarly, in the GSE50760 dataset, *RPL28* expression was higher by 27 % ($P < 0.05$) in primary tumors and by 30 % ($P < 0.05$) in liver metastatic tissues, compared to paired normal tissues (**Figure 3d**).

High RPL28 expression is associated with reduced survival of mCRC patients, altered gene expression and affects colon cancer cells growth and survival in vitro

Based on the dichotomization at the optimal cut-off of expression levels, mCRC patients from the TCGA cohort with high tumor expression of *RPL28* had a reduced survival compared to those with low levels ($n = 88$, $P = 0.004$; **Figure 4a**). This observation was validated in mCRC cases from the dataset GSE17538 ($n = 56$, $P < 0.001$; **Figure 4b**). Dichotomization at the median expression levels also lead to significantly reduced survival for high *RPL28* expression groups, in both TCGA ($P = 0.009$) and GSE17538 ($P = 0.009$) datasets (**Additional file 1: Supplementary Figure 1**). A differential gene expression and pathway enrichment analysis further indicated that over 773 genes were significantly altered ($FDR < 0.05$) by *RPL28* expression (high versus low, optimal cut-off separation) in cases of the TCGA cohort, including 257 down-regulated and 516 up-regulated genes (**Additional file 2: Supplementary Table 6**). From up-regulated genes, three pathways related to immunoglobulins were enriched (adjusted P -value < 0.05) in cases with high *RPL28*, whereas six pathways related to extracellular matrix (ECM) or collagen were enriched from down-regulated genes (adjusted P -value < 0.05) (**Figure 4c**). In these pathways, 14 genes comprised of mostly immunoglobulin genes were significantly upregulated ($FDR < 0.05$) in high *RPL28* cases, and a total of 25 genes, mainly collagen

genes, were significantly down regulated ($FDR < 0.05$) (**Figure 4c**). *In vitro* investigation further indicated significantly reduced proliferation by 1.4-fold to 5.6-fold ($P < 0.05$) and reduced viable cell ratios by 15 to 75% ($P < 0.05$) associated with knockdown of RPL28 in HCT116 and HT-29 colon cell lines (**Figure 5**).

DISCUSSION

We report that the germline variant rs4806668G>T in *RPL28* and tumor expression of this gene are associated with survival of mCRC patients. The rs4806668TT genotype was associated with reduced PFS and OS in two independent cohorts of mCRC. This polymorphism was also associated with an increased expression of *RPL28* in colon tissue of healthy donors, and its expression was further increased in tumor tissues compared to paired normal tissues. In two other independent cohorts of mCRC, higher *RPL28* tumor expression level was associated with poor survival.

Our findings reveal that the *RPL28* rs4806668G>T is associated with PFS and OS in mCRC cases treated with first-line FOLFIRI regimen. FOLFIRI is one of the most common chemotherapy backbones for mCRC first-line treatment, however; only few studies investigated the potential of germline polymorphisms to predict its efficacy in several cohorts. Xu et al. identified the combination of *UGT1A1**28 and *UGT1A1**6 as predictive of reduced overall survival in two cohorts of Asian mCRC cases receiving irinotecan-based regimen in second- and third-line treatment¹⁰. A meta-analysis in Caucasians concluded that *UGT1A1**28 cannot be considered as a reliable predictor of survival in CRC patients treated with irinotecan-based regimen¹¹. Most studies including a discovery and a replication set focused on genes involved in targeted therapies – mainly bevacizumab – administered along with FOLFIRI, and reported markers of survival in several genes including *IL6*, *CD39*, *CXCR4* and *MKNK1*¹²⁻¹⁶. *RPL28* is one of the 79 members of the ribosomal proteins (RPs) family. These proteins are constitutive components of the large subunit (RPL) and small subunit (RPS) of the ribosome primarily responsible for protein synthesis, which is greatly perturbed in tumor cells¹⁷. In addition to their role in ribosomal biogenesis and protein production, RPs possess ribosome-independent functions, especially in tumorigenesis, immune signaling and development¹⁸. This was notably demonstrated in breast, ovarian and pancreatic cancers as well as in osteosarcoma¹⁹⁻²². In CRC, a study showed that a higher level of the ribosomal RPS7 protein was associated with improved PFS and OS in 92 stage IV mCRC patients²³. Another study of 200 CRC patients revealed a reduced OS associated with a higher level of RPS15A protein²⁴. *RPL28* upregulation was shown to reduce OS after azacytidine treatment in patients with myelodysplasia and related

neoplasms²⁵. These studies pointed toward potential functions of ribosome-free RPs that might engage an oncogenic role or suppressing tumor cell proliferation, depending on the RPs proteins involved¹⁸. A recent study suggested that *RPL28* was part of a hub of genes correlated to microsatellite instability status in CRC, indicating a potential role for this protein in defective DNA mismatch repair system associated with CRC²⁶.

Various RPs are overexpressed in different cancer cells and are associated with the development and progression of malignant cancers. The *RPL28* htSNP rs4806668G>T and its two tightly linked SNPs (rs3810168T>C and rs3745272C>G) were associated with an increased *RPL28* expression in transverse colon of healthy individuals and were predicted to be functional based on RegulomeDB data, in support of a causal link with at least one of these SNPs. In the transverse colon, *RPL28* expression level appeared similar for GT and TT carriers and higher than in GG carriers of rs4806668. However, an allele dose effect was observed for the linked rs3810168T>C, rs73617860T>G and rs73617855T>A, with higher expression for minor homozygous carriers, in line with the recessive model of the association in our cohorts. An upregulation of *RPL28* was further noted in tumors compared to paired normal tissues suggesting an oncogenic role of this ribosomal protein in CRC. In support of this notion, repression of *RPL28* was shown to significantly impair cell proliferation of HCT116 and HT-29 colon cancer cell models, suggesting that *RPL28* higher expression may be associated to a more aggressive phenotype. Knockdown of other RPs such as *RPL9* and *RPS24* suppressed colon carcinoma cell growth *in vitro* and human CRC xenografts in nude mice^{27,28}. Similarly, overexpression of *RPS7* led to reduced proliferation of a panel of 4 CRC cell lines including HCT116 and HT-29, suggested to be mediated through repression of *HIF-1 α* and glycolysis²³.

Our analysis of transcriptomic changes associated with high *RPL28* tumor expression in mCRC patients suggested an upregulation of several genes of immunoglobulins, proposing the involvement of the complement system. Although the complement is recognized as an important actor of the immune system that contributes to the destruction of cancer cells, recent data have indicated a tumor-promoting role^{29,30}. Several studies support that malignant cells efficiently activate complement and that metastatic pathways are triggered by imbalanced complement activation and inflammation³¹. Recently,

RPL28 expression was inversely associated to viral peptide presentation by the major histocompatibility complex class I, suggesting a potential role of this ribosomal protein in the immune system processing ³². Immunosuppressive properties were proposed for RPS19, also increased in colon cancer cells and released from apoptotic tumor cells ³³. It was further suggested that RPS19 immunosuppressive role in cancer involved its interaction with the complement receptor C5aR1 (CD88), promoting tumor growth in a transgenic model of breast cancer ³⁴. Another study reported increased cell invasiveness in response to C5a-induced secretion of metalloproteinases from C5aR1-expressing cancer cells and degradation of extracellular matrix (ECM), known to contribute to neoplastic progression ^{35,36}. Indeed, we observed a down-expression of several ECM genes in high compared to low *RPL28* expressing primary tumors. A role for ECM components such as collagen, fibronectin or matrix metalloproteinases was clearly established in cancer development and metastatic spread ³⁷. A down-regulation of genes related to ECM adhesion and remodeling was also observed between hepatic CRC metastases and primary colorectal tumors ³⁸. As a lower adhesion is thought to favor metastasis and dissemination, a down-regulation of ECM related genes in the high *RPL28*/low survival group of patients might be the sign of a higher susceptibility for cancer cells to spread more efficiently and quickly from the primary tumor ³⁹. Overall, it may be envisioned that the changes in gene expression associated with *RPL28* activate cancer-related signaling pathways and exert tumor-promoting activities in mCRC patients.

A limitation of this study was that we did not apply correction for multiple testing in the polymorphism association study due to its exploratory nature. However, we studied two independent cohorts to corroborate initial findings, limiting the possibility of false-positive associations. We also included multiple datasets, including the large GTEx and TCGA cohorts, to study associations with gene expression. We acknowledge that our study design comprising only FOLFIRI-treated patients did not allow to clarify the predictive value of the positive marker and that further investigations are needed to explain the molecular mechanisms by which RPL28 affects survival and its possible interplay with ECM and immunoglobulin pathways. Given the high frequency of the rs4806668TT genotype in

Africans, it would be of interest to evaluate further the association of the rs4806668G>T with clinical outcome in this population.

In conclusion, we reported the germline variant *RPL28* rs4806668G>T, associated with increased gene expression, as a novel marker of survival in mCRC patients treated with the FOLFIRI regimen. We further established that high *RPL28* tumor expression is linked to the reduced survival of mCRC patients, possibly through an enhanced tumor growth and a remodelling of the expression of ECM and immunoglobulin pathways. This study highlights a potential role of the ribosomal protein RPL28 in colon tumorigenesis and as a possible prognostic survival marker of mCRC patients.

METHODS

Characteristics of cohorts – Two independent cohorts of mCRC patients were included in the study. They were previously described in details, including the study eligibility criteria and genomic DNA isolation procedure^{40,41}. Briefly, the discovery cohort was composed of 167 patients all treated with FOLFIRI-based regimen recruited in three centers of Eastern Canada from 2003 to 2012. These patients received FOLFIRI and 75 were also co-treated with bevacizumab or an experimental drug. The replication cohort enrolled 250 mCRC patients recruited in 13 centers of Northeastern Italy between 2002 and 2005 all treated with FOLFIRI. Clinical outcomes for this study were progression-free survival (PFS, defined as the time between the beginning of FOLFIRI treatment and the first evidence of disease progression, including death, or the last follow-up) and overall survival (OS, defined as the time between the beginning of FOLFIRI treatment and death from any cause). Characteristics of both cohorts are summarized in **Table 1**. Ethnicity was self-reported with the vast majority identified as Caucasian or White. All patients provided written informed consent and the research protocol was approved by the local ethics committees (CHU de Quebec ethics committee for the Canadian cohort and Comitato Etico Indipendente - Centro di Riferimento Oncologico di Aviano for the Italian cohort). This study was conducted in accordance with the declaration of Helsinki.

Genetic analysis – Single Nucleotide Polymorphisms (SNPs) with a minor allele frequency higher than 5% were identified from the European (CEU) population of the International HapMap project for the selected candidate genes (**Additional file 1: Supplementary Table 1**). Haplotype-tagging SNPs (htSNPs) in high linkage disequilibrium (LD, $r^2 > 0.80$) in the CEU population were selected with Haploview v4.2 (Broad Institute, Cambridge, MA, United States)⁴². A complete list of htSNPs and their associated SNPs is provided as supplementary data (**Additional file 1: Supplementary Table 2**). All htSNPs were genotyped using the Sequenom iPLEX matrix-assisted laser desorption/ionization time-of-flight mass spectrometry technology (Sequenom, San Diego, CA, United States). Genotypic frequencies and LD data were retrieved through Ensembl GRCh38 release 91 - December 2017 (<https://www.ensembl.org/index.html>) for European (CEU), Chinese

(CHB), Japanese (JPT) and African (YRI) populations. RegulomeDB v1.1 was used to investigate whether polymorphisms may have a functional impact ⁴³.

Statistical analysis – Hardy-Weinberg equilibrium was verified using PLINK v1.07 ⁴⁴ and SNPs deviating from the equilibrium ($P < 0.05$) were excluded from further analysis. Associations between htSNPs and clinical outcomes were tested using Cox proportional hazards model adjusted for age and co-treatment in the Canadian and combined cohorts and for age in the Italian cohort. Additive, dominant and recessive models were fitted independently for each htSNP. Analyses were carried out using SAS version 9.4 (SAS Institute Inc., Cary, NC, United States) and R v3.2.2 Survival package. Genetic variants were initially tested in the Canadian cohort and those with a P -value < 0.1 were genotyped in the Italian cohort. A marker was considered validated when associated to survival using the same genetic model in both cohorts ($P < 0.05$). Kaplan-Meier curves and log-rank test P -values were estimated with GraphPad Prism 5 (GraphPad Software Inc., La Jolla, CA, United States). The optimal cut-off values of gene expression for Kaplan-Meier analysis were defined as the point with the most significant log-rank test split, as determined by Cutoff Finder version 2.1 ⁴⁵.

Expression data – Normalized gene expression and P -value data from the Genotype-Tissue Expression (GTEx) project (Analysis Release V7) were obtained through the GTEx portal (<https://www.gtexportal.org/home/>) on May 29, 2019. The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>) was used to obtain data of Colon Adenocarcinoma (TCGA-COAD) and Rectum Adenocarcinoma (TCGA-READ) projects, accessed through the Genomic Data Commons Data Portal (<https://portal.gdc.cancer.gov/>) on October 2, 2018. HTseq counts (raw counts) for stage IV TCGA-COAD and TCGA-READ individuals were retrieved with TCGAbiolinks package (version 2.9.5) in R and processed for TMM normalization with edgeR package (version 3.22.3). Differential gene expression between individuals with high and low *RPL28* expression levels based on the optimal cut-off separation was then performed using edgeR package and exact test method. Ensembl gene ID were converted into gene symbols with g:Profiler ⁴⁶. Pathway enrichment analysis with significantly differentially expressed genes ($FDR < 0.05$) and $\log_2(FC)$ lower than -1.2 or higher than 1.2 was performed using Reactome 2016 library in Enrichr ^{47,48}. Expression and

clinical data publicly available for dataset GSE49355, GSE50760 and GSE17538 were obtained through GEO database with GEOquery package (version 2.48.0) ^{38,49,50}. Groups were compared using a two-tailed paired t-test or a one-way ANOVA followed by a *post-hoc* Dunnett's test in GraphPad Prism 5 software.

In vitro experiments – HCT116 and HT-29 cell lines were obtained from the American Type Culture Collection (Rockville, MD, United States) and grown as recommended. Cells were passaged less than 15 times between thawing and completion of experiments. Cells were reverse-transfected in 96-well or 24-well plates with a final concentration of 20 nM siRNA (Dharmacon, Chicago, IL, USA) using DharmaFECT2 Transfection Reagent as per the manufacturer's recommendations. Two siRNA targeting RPL28 (D-011145-02 or siRPL28-A and D-011145-03 or siRPL28-B) and a scrambled non-targeting negative control (D-001206-14 further referred as siCTR) were tested. Viable cells were counted at 24 h, 48 h, 72 h and 96 h post-transfection using trypan blue and an automated cell counter (Bio-Rad, Mississauga, ON, Canada). For each time point, fold over control was calculated by dividing the viable cell count by the viable cell count of siCTR at 24 h. Experiments were carried out three independent times in duplicate. Viable cell proliferation at 96 h was assessed by an MTS assay (CellTiter 96 AQueous One Solution Cell Proliferation Assay, Promega, Madison, WI, United States) followed by measurement of absorbance at 490 nm. Viable cell proliferation was expressed as the ratio of absorbance for each condition over absorbance for siCTR. At least three independent experiments were performed in triplicate. At 96 h, conditions were compared using a one-way ANOVA followed by a *post-hoc* Dunnett's test in GraphPad Prism 5 software. For protein expression, cells were washed with PBS 72 hours post-transfection and harvested in lysis buffer [0.05 M Tris-HCl pH 7.4, 0.15 M NaCl, 1 % (w/v) Igepal CA-630 (Sigma-Aldrich), 1 mM dithiothreitol, and Complete protease inhibitor cocktail (Roche, Laval, QC, Canada)], lysed on a rotation unit for 30 min and cleared by centrifugation for 15 min at 13000 g. Cell lysates (5 µg) were resolved on a 12 % polyacrylamide gel and transferred to a nitrocellulose membrane. RPL28 protein was detected using a rabbit polyclonal anti-RPL28 antibody (Proteintech, Group, 16649-1-AP, Rosemont, IL, United States) and GAPDH, used as a loading control, with a mouse monoclonal anti-GAPDH antibody (G8795, Sigma-Aldrich). RPL28 protein levels were determined by densitometry scanning of band intensity and normalized by the

corresponding GAPDH band intensity on immunoblots using ImageJ (<https://imagej.nih.gov/ij/>). Experiments were carried out three independent times.

Data Availability

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES

- 1 Brenner, H. & Chen, C. The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention. *Br J Cancer* **119**, 785-792, (2018).
- 2 Fujita, K., Kubota, Y., Ishida, H. & Sasaki, Y. Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J Gastroenterol* **21**, 12234-12248, (2015).
- 3 Smith, N. F., Figg, W. D. & Sparreboom, A. Pharmacogenetics of irinotecan metabolism and transport: an update. *Toxicol In Vitro* **20**, 163-175, (2006).
- 4 Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer* **3**, 330-338, (2003).
- 5 Gupta, E. *et al.* Metabolic fate of irinotecan in humans: correlation of glucuronidation with diarrhea. *Cancer Res* **54**, 3723-3725, (1994).
- 6 Hu, Z. Y., Yu, Q., Pei, Q. & Guo, C. Dose-dependent association between UGT1A1*28 genotype and irinotecan-induced neutropenia: low doses also increase risk. *Clin Cancer Res* **16**, 3832-3842, (2010).
- 7 Barbarino, J. M., Haidar, C. E., Klein, T. E. & Altman, R. B. PharmGKB summary: very important pharmacogene information for UGT1A1. *Pharmacogenet Genomics* **24**, 177-183, (2014).
- 8 Beutler, E., Gelbart, T. & Demina, A. Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? *Proc Natl Acad Sci U S A* **95**, 8170-8174, (1998).
- 9 Rodrigues, D., Longatto-Filho, A. & Martins, S. F. Predictive Biomarkers in Colorectal Cancer: From the Single Therapeutic Target to a Plethora of Options. *Biomed Res Int* **2016**, 6896024, (2016).
- 10 Xu, C. *et al.* UGT1A1 gene polymorphism is associated with toxicity and clinical efficacy of irinotecan-based chemotherapy in patients with advanced colorectal cancer. *Cancer Chemother Pharmacol* **78**, 119-130, (2016).
- 11 Liu, X., Cheng, D., Kuang, Q., Liu, G. & Xu, W. Association between UGT1A1*28 polymorphisms and clinical outcomes of irinotecan-based chemotherapies in colorectal cancer: a meta-analysis in Caucasians. *PLoS One* **8**, e58489, (2013).
- 12 Matsusaka, S. *et al.* Prognostic Impact of IL6 Genetic Variants in Patients with Metastatic Colorectal Cancer Treated with Bevacizumab-Based Chemotherapy. *Clin Cancer Res* **22**, 3218-3226, (2016).
- 13 Berger, M. D. *et al.* A Polymorphism within the Vitamin D Transporter Gene Predicts Outcome in Metastatic Colorectal Cancer Patients Treated with FOLFIRI/Bevacizumab or FOLFIRI/Cetuximab. *Clin Cancer Res* **24**, 784-793, (2018).
- 14 Tokunaga, R. *et al.* Prognostic Effect of Adenosine-related Genetic Variants in Metastatic Colorectal Cancer Treated With Bevacizumab-based Chemotherapy. *Clin Colorectal Cancer* **18**, e8-e19, (2019).
- 15 Berger, M. D. *et al.* Impact of genetic variations in the MAPK signaling pathway on outcome in metastatic colorectal cancer patients treated with first-line FOLFIRI and bevacizumab: data from FIRE-3 and TRIBE trials. *Ann Oncol* **28**, 2780-2785, (2017).
- 16 Matsusaka, S. *et al.* CXCR4 polymorphism predicts progression-free survival in metastatic colorectal cancer patients treated with first-line bevacizumab-based chemotherapy. *Pharmacogenomics J* **17**, 543-550, (2017).

- 17 Loreni, F., Mancino, M. & Biffo, S. Translation factors and ribosomal proteins control tumor onset and progression: how? *Oncogene* **33**, 2145-2156, (2014).
- 18 Zhou, X., Liao, W. J., Liao, J. M., Liao, P. & Lu, H. Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol* **7**, 92-104, (2015).
- 19 Dave, B. *et al.* Role of RPL39 in Metaplastic Breast Cancer. *J Natl Cancer Inst* **109**, (2017).
- 20 Wei, F. *et al.* Ribosomal protein L34 promotes the proliferation, invasion and metastasis of pancreatic cancer cells. *Oncotarget* **7**, 85259-85272, (2016).
- 21 Luo, S. *et al.* Highly expressed ribosomal protein L34 indicates poor prognosis in osteosarcoma and its knockdown suppresses osteosarcoma proliferation probably through translational control. *Sci Rep* **6**, 37690, (2016).
- 22 Shi, J. *et al.* Biological Function of Ribosomal Protein L10 on Cell Behavior in Human Epithelial Ovarian Cancer. *J Cancer* **9**, 745-756, (2018).
- 23 Zhang, W. *et al.* RPS7 inhibits colorectal cancer growth via decreasing HIF-1 α -mediated glycolysis. *Oncotarget* **7**, 5800-5814, (2016).
- 24 Chen, J. *et al.* Ribosomal protein S15A promotes malignant transformation and predicts poor outcome in colorectal cancer through misregulation of p53 signaling pathway. *Int J Oncol* **48**, 1628-1638, (2016).
- 25 Monika Belickova, M. *et al.* Up-regulation of ribosomal genes is associated with a poor response to azacitidine in myelodysplasia and related neoplasms. *Int J Hematol* **104**, 566-573, (2016).
- 26 Yu, C. *et al.* Identification of key genes and pathways involved in microsatellite instability in colorectal cancer. *Mol Med Rep* **19**, 2065-2076, (2019).
- 27 Baik, I. H. *et al.* Knockdown of RPL9 expression inhibits colorectal carcinoma growth via the inactivation of Id-1/NF-kappaB signaling axis. *Int J Oncol* **49**, 1953-1962, (2016).
- 28 Wang, Y. *et al.* RPS24 knockdown inhibits colorectal cancer cell migration and proliferation *in vitro*. *Gene* **571**, 286-291, (2015).
- 29 Macor, P., Capolla, S. & Tedesco, F. Complement as a Biological Tool to Control Tumor Growth. *Front Immunol* **9**, 2203, (2018).
- 30 Pio, R., Ajona, D. & Lambris, J. D. Complement inhibition in cancer therapy. *Semin Immunol* **25**, 54-64, (2013).
- 31 Reis, E. S., Mastellos, D. C., Ricklin, D., Mantovani, A. & Lambris, J. D. Complement in cancer: untangling an intricate relationship. *Nat Rev Immunol* **18**, 5-18, (2018).
- 32 Wei, J. *et al.* Ribosomal Proteins Regulate MHC Class I Peptide Generation for Immunosurveillance. *Mol Cell* **73**, 1162-1173 e1165, (2019).
- 33 Kondoh, N., Schweinfest, C. W., Henderson, K. W. & Papas, T. S. Differential expression of S19 ribosomal protein, laminin-binding protein, and human lymphocyte antigen class I messenger RNAs associated with colon carcinoma progression and differentiation. *Cancer Res* **52**, 791-796, (1992).
- 34 Markiewski, M. M. *et al.* The Ribosomal Protein S19 Suppresses Antitumor Immune Responses via the Complement C5a Receptor 1. *J Immunol* **198**, 2989-2999, (2017).
- 35 Pickup, M. W., Mouw, J. K. & Weaver, V. M. The extracellular matrix modulates the hallmarks of cancer. *EMBO Rep* **15**, 1243-1253, (2014).

- 36 Nitta, H. *et al.* Cancer cells release anaphylatoxin C5a from C5 by serine protease to enhance invasiveness. *Oncol Rep* **32**, 1715-1719, (2014).
- 37 Crotti, S. *et al.* Extracellular Matrix and Colorectal Cancer: How Surrounding Microenvironment Affects Cancer Cell Behavior? *J Cell Physiol* **232**, 967-975, (2017).
- 38 Del Rio, M. *et al.* Specific extracellular matrix remodeling signature of colon hepatic metastases. *PLoS One* **8**, e74599, (2013).
- 39 Miyaki, M. *et al.* Increased cell-substratum adhesion, and decreased gelatinase secretion and cell growth, induced by E-cadherin transfection of human colon carcinoma cells. *Oncogene* **11**, 2547-2552, (1995).
- 40 Levesque, E. *et al.* Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J Pharmacol Exp Ther* **345**, 95-101, (2013).
- 41 Toffoli, G. *et al.* The role of UGT1A1*28 polymorphism in the pharmacodynamics and pharmacokinetics of irinotecan in patients with metastatic colorectal cancer. *J Clin Oncol* **24**, 3061-3068, (2006).
- 42 Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-265, (2005).
- 43 Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-1797, (2012).
- 44 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, (2007).
- 45 Budczies, J. *et al.* Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS One* **7**, e51862, (2012).
- 46 Reimand, J. *et al.* g:Profiler-a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res* **44**, W83-89, (2016).
- 47 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, (2016).
- 48 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, (2013).
- 49 Kim, S. K. *et al.* A nineteen gene-based risk score classifier predicts prognosis of colorectal cancer patients. *Mol Oncol* **8**, 1653-1666, (2014).
- 50 Smith, J. J. *et al.* Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* **138**, 958-968, (2010).

Acknowledgements

We would like to thank all study participants and clinical staff involved in the study. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 05/29/19. The results shown here are in part based upon data generated by the TCGA Research Network (<http://cancergenome.nih.gov/>) and we thank Julien Prunier for his help with the retrieving of these data. This work was supported by the Canadian Institutes of Health Research (CIHR, FRN-42392) and the Canada Research Chair in Pharmacogenomics (Tier I) to CG. AL was supported by graduate scholarships from the “Fonds d’enseignement et de recherche” of the Faculty of pharmacy and “Fonds de Recherche du Québec - Santé”. EL received a CIHR clinical-scientist phase II award and Prostate Cancer Canada Rising Star Award (RS2013-55).

Authors' contributions

AL performed the experiments; AL, LV and DS performed statistical analysis; EL, EC, EDM, LV, DJ, FC, AB, MDA, GT and CG participated in the recruitment of patients and obtaining DNA; AL and EPA performed bioinformatics analysis; AL, MR and CG wrote the manuscript; GT and CG designed the study. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

FIGURE LEGENDS

Figure 1. *RPL28* rs4806668G>T polymorphism is associated with survival in mCRC patients treated with FOLFIRI. **a-b.** Cox proportional hazards models adjusted for age and co-treatment (Canadian cohort, n = 167) and for age (Italian cohort, n = 250) showed association of rs4806668G>T with progression-free survival (PFS) and overall survival (OS) using a recessive genetic model. **c-d.** Univariate Kaplan-Meier survival curves for PFS and OS according to rs4806668G>T genotype in combined cohorts. The percentage survival according to genotypes is shown under the graphs. HR_{adj}, adjusted hazard ratio; htSNPs, haplotype-tagging single nucleotide polymorphisms; CI, confidence interval.

Figure 2. Rs4806668 is located in the promoter region of the *RPL28* locus and is linked to several other polymorphisms (SNPs). **a.** Localization of the *RPL28* rs4806668G>T variant and its associated SNPs in strong linkage disequilibrium ($r^2 > 0.80$ in the European population). **b.** Position of the rs4806668G>T marker and its linked SNPs relative to the translation start site of *RPL28*. Scores are from RegulomeDB and represent the probability for a SNP to be functional. N/A, not available **c.** Frequencies of *RPL28* rs4806668G>T among different ethnic groups (Ensembl GRCh38 release 91).

Figure 3. Relationship between *RPL28* variants, tissue type and gene expression **a.** *RPL28* rs4806668G>T and linked SNPs are associated with an increased gene expression in transverse colon tissues of healthy individuals from the GTEx cohort. **b.** *RPL28* expression is higher in primary tumor relative to paired normal colorectal tissues (n = 50 pairs) from the TCGA cohort. **c.** *RPL28* expression is higher in primary colorectal tumors and liver metastases relative to paired normal colorectal tissues (n = 10 pairs) from the GSE49355 dataset. **d.** *RPL28* expression is higher in primary colorectal tumors and liver metastases relative to paired normal colorectal tissues (n = 18 pairs) from the GSE50760 dataset. A.U., arbitrary units; FPKM, fragments per kilobase million; N, normal tissue; PT, primary tumor tissue; M, liver metastases. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

Figure 4. *RPL28* expression level in colorectal tumor tissues is associated with survival and alters tumor transcriptome. **a.** Kaplan-Meier curves for high and low *RPL28* expression groups of stage IV mCRC individuals from the TCGA cohort (n = 88). The percentage

survival according to *RPL28* expression group is shown under the graph. **b.** Kaplan-Meier curves for high and low *RPL28* expression groups of stage IV mCRC individuals from the GSE17538 dataset (n = 56). The percentage survival according to *RPL28* expression group is shown under the graph. **c.** Significantly enriched pathways from differential gene expression in advanced stage IV mCRC cases with high *RPL28* expression compared to those with low *RPL28* expression from the TCGA cohort. Pathways in blue are enriched with down-regulated genes whereas those in red are enriched with up-regulated genes. Genes belonging to enriched pathways are shown on the right.

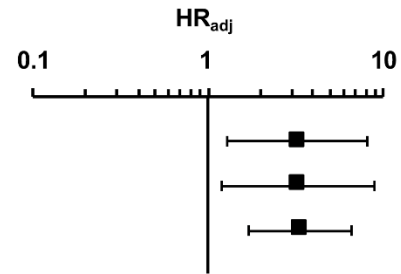
Figure 5. RPL28 expression levels influence proliferation of colorectal cancer cells. **a-b.** Efficient knock-down of RPL28 protein expression in HCT116 and HT-29 exposed to specific siRNAs (siRPL28) relative to cells transfected with scrambled non-targeting (NT) negative control (siCTR). A representative cropped Western blot is shown for each cell line and a quantification of expression levels based on densitometry of blots from three independent experiments are shown. Full-length blots are shown in **Additional file 1: Supplementary Figure 2.** **c-f.** The knock-down of RPL28 impairs cell proliferation. Viable cells were monitored up to 96 hours after siRNA transfection in (c) HCT116 and (e) HT-29 (n = 3 in duplicate). Cell proliferation was measured by MTS assays 96 hours post-transfection in (d) HCT116 and (f) HT-29 (n ≥ 3 in triplicate). Data are presented as mean ± S.E.M. * $P < 0.05$; *** $P < 0.001$.

Figure 1.

a.

Progression-free survival

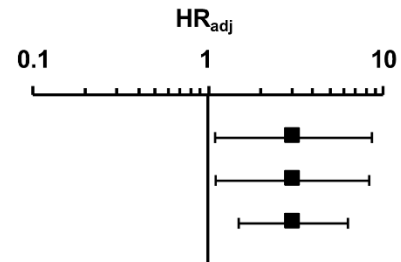
htSNP	Cohort	HR _{adj} (95% CI)	P
rs4806668G>T	Canada	3.23 (1.29 – 8.14)	0.013
	Italy	3.28 (1.20 – 8.95)	0.021
	Combined	3.36 (1.71 – 6.61)	<0.001



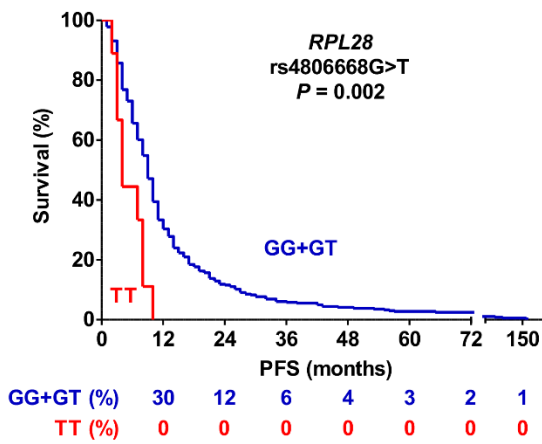
b.

Overall survival

htSNP	Cohort	HR _{adj} (95% CI)	P
rs4806668G>T	Canada	3.09 (1.10 – 8.65)	0.032
	Italy	3.05 (1.11 – 8.36)	0.030
	Combined	3.07 (1.50 – 6.30)	0.002



c.



d.

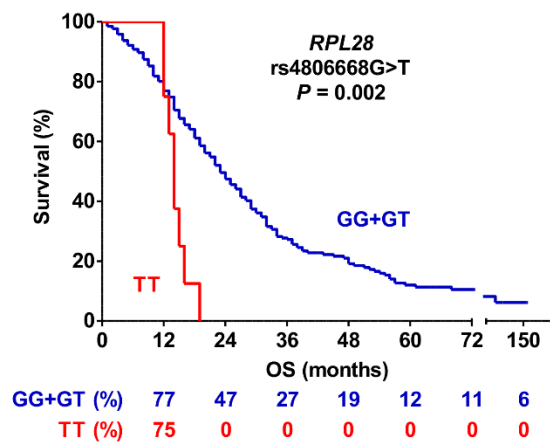


Figure 2.

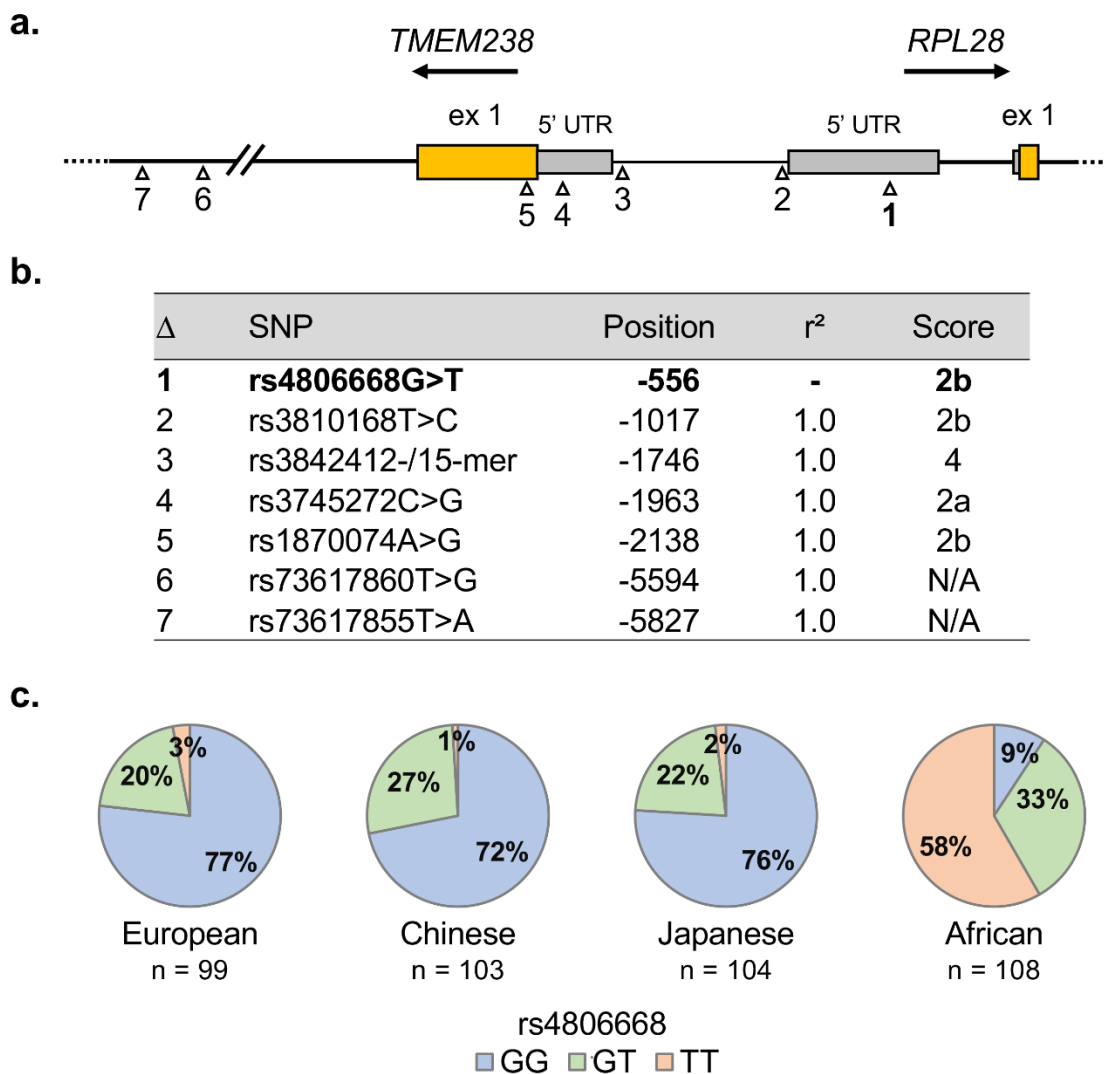


Figure 3.

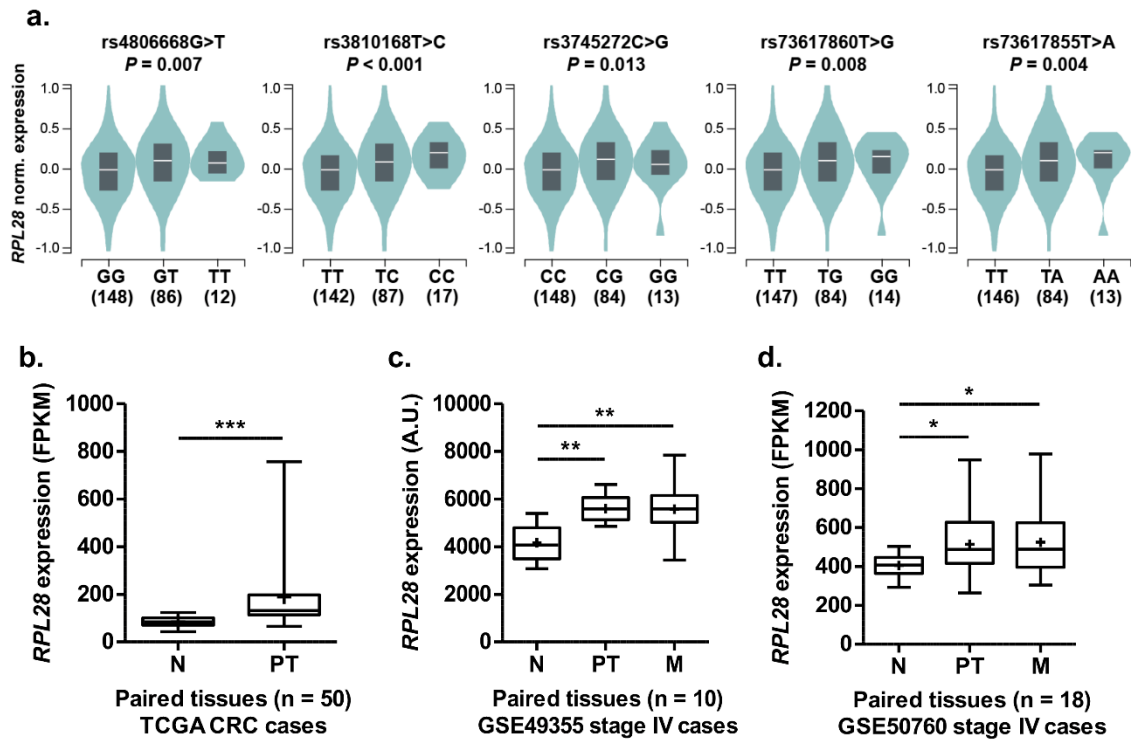


Figure 4.

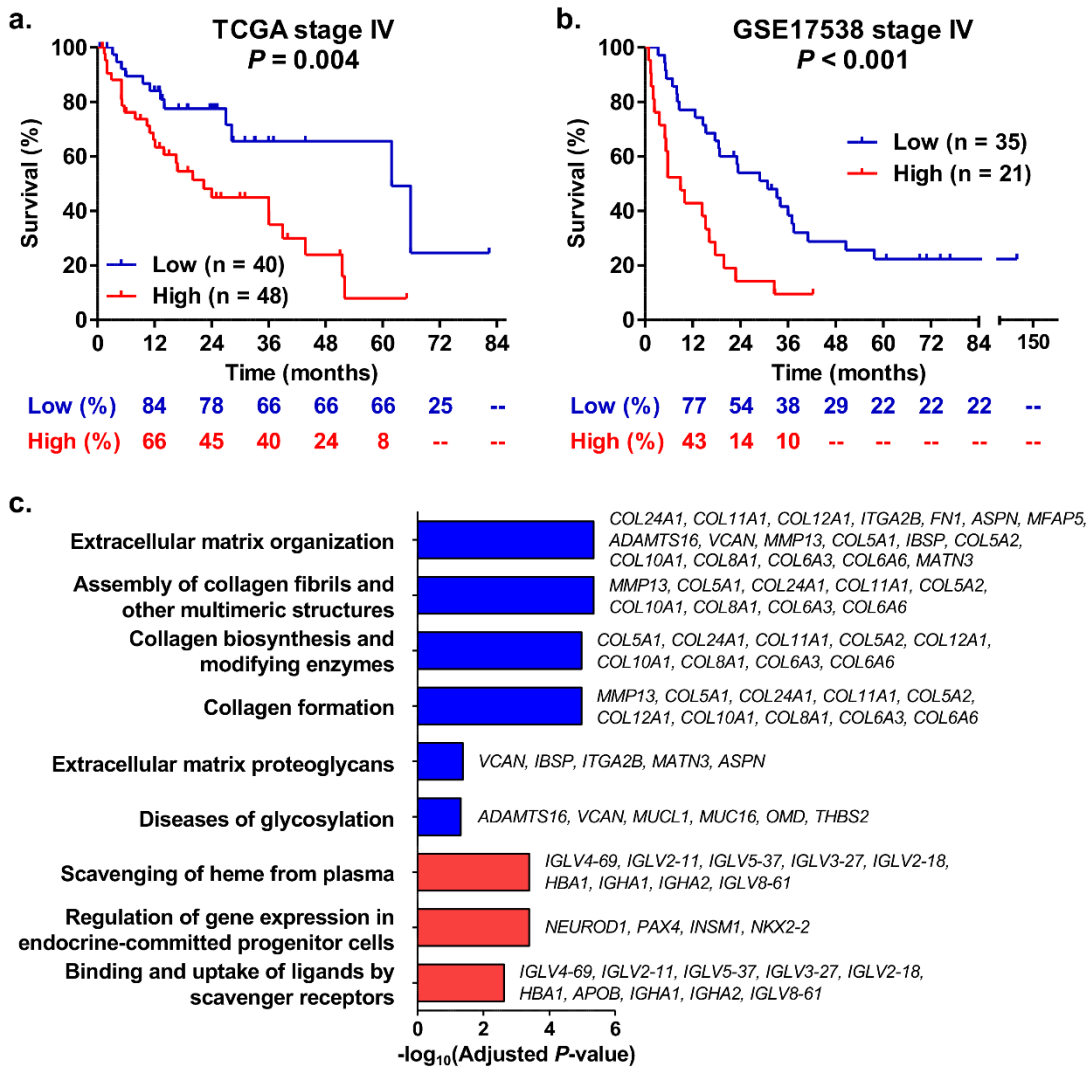


Figure 5.

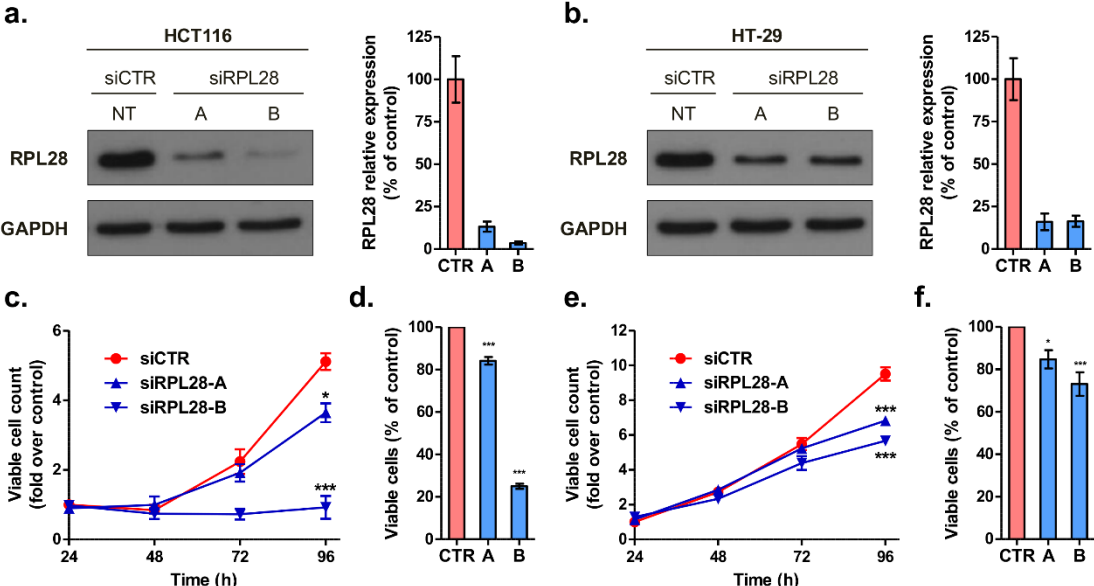


Table 1. Characteristics of the discovery and replication mCRC cohorts.

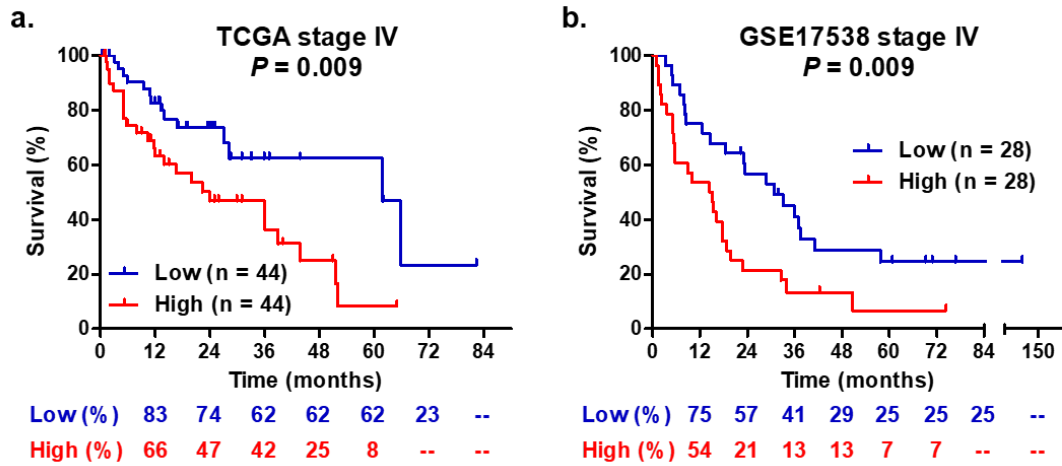
	Canadian cohort		Italian cohort	
Characteristics	N = 167	%	N = 250	%
Gender				
Male	110	66	162	65
Female	57	34	88	35
Age (years)				
Mean	61.5	--	60.6	--
Standard Deviation	10.2	--	10.3	--
Range	29 - 86	--	26 - 75	--
FOLFIRI	167	100	250	100
Co-treatment	75	44.9	0	0
Bevacizumab	69	92.0	0	0
Other drug	6	8.0	0	0
Clinical outcomes (median in months)				
Progression-free survival	11	--	7	--
Overall survival	24	--	15	--

Germline variability and tumor expression level of ribosomal protein gene *RPL28* are associated with survival of metastatic colorectal cancer patients

Adrien Labriet, Éric Lévesque, Erika Cecchin, Elena De Mattia, Lyne Villeneuve, Michèle Rouleau, Derek Jonker, Félix Couture, David Simonyan, Eric P. Allain, Angela Buonadonna, Mario D'Andrea, Giuseppe Toffoli and Chantal Guillemette

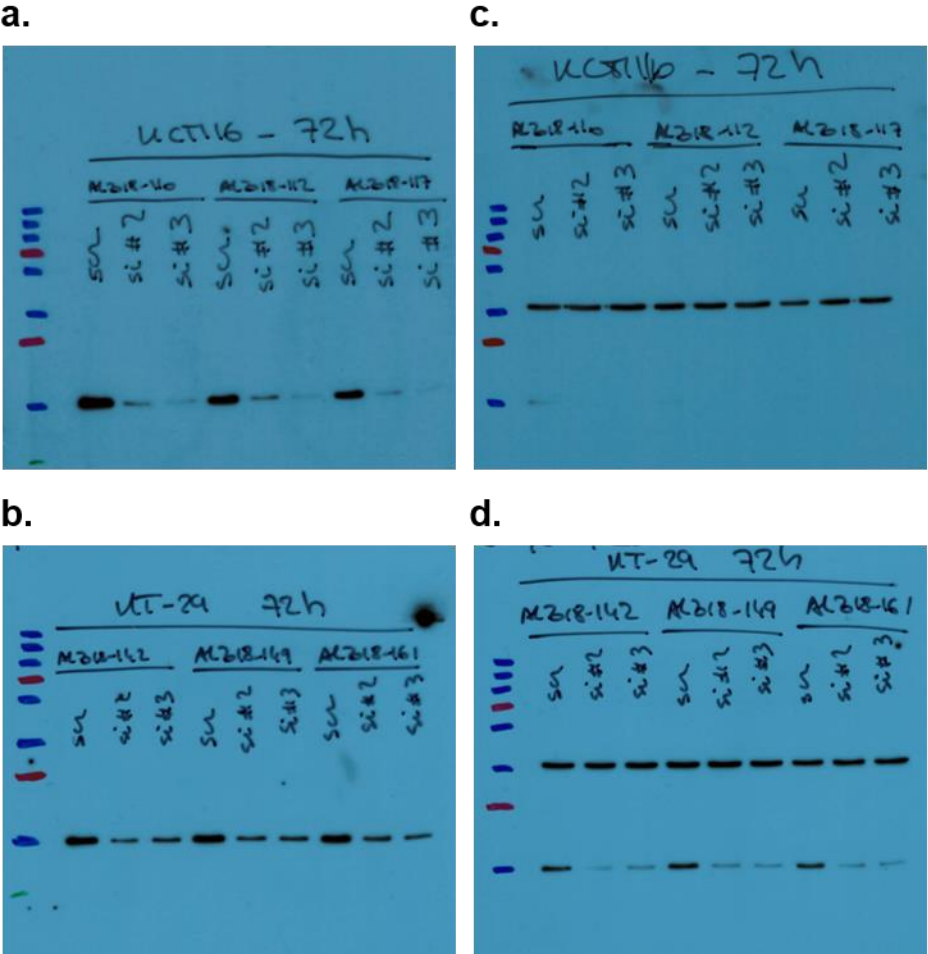
Supplementary Material

Supplementary Figure 1.



Supplementary Figure 1. *RPL28* expression level in colorectal tumor tissues is associated with survival. **A.** Kaplan-Meier curves for high and low *RPL28* expression groups (median separation) of stage IV mCRC individuals from the TCGA cohort (n = 88). The percentage survival according to *RPL28* expression group is shown under the graph. **B.** Kaplan-Meier curves for high and low *RPL28* expression groups (median separation) of stage IV mCRC individuals from the GSE17538 dataset (n = 56). The percentage survival according to *RPL28* expression group is shown under the graph. This figure is related to Figure 4a and 4b.

Supplementary Figure 2.



Supplementary Figure 2. Full-length blots related to Figure 5. Western blots for the detection of RPL28 in HCT116 (a) and HT-29 (b) are shown. GAPDH was detected on the same blots for HCT116 (c) and HT-29 (d). scr, siCTR; si#2, siRPL28-A; si#3, siRPL28-B.

Supplementary Table 1. Relation between selected genes and irinotecan and/or 5-fluorouracil response.

Gene	Gene full-name	Link with irinotecan and/or 5-FU	Ref.
<i>CDC45</i>	Cell division cycle 45	Involved in TOP1-apoptosis pathway	1
<i>MAPK9</i>	Mitogen-activated protein kinase 9	Silencing enhances SN-38 effect <i>in vitro</i> .	2
<i>PARP1</i>	Poly(ADP-ribose) polymerase 1	PARP1 inhibition increases sensitivity of colon cancer cell lines to irinotecan. Involved in the repair of TOP1-induced DNA damages.	3
<i>SART1</i>	U4/U6.U5 tri-snRNP-associated protein 1	Silencing increases sensitivity to SN-38 and 5-FU <i>in vitro</i> .	2
<i>ACTL6A</i>	Actin like 6A	Down-regulated in irinotecan-resistant cells.	4
<i>CBX3</i>	Chromobox 3	Down-regulated in irinotecan-resistant cells.	5
<i>CFL1</i>	Cofilin 1	Differentially expressed in irinotecan-resistant cells.	4,5
<i>CRYAB</i>	Crystallin alpha B	Up-regulated in irinotecan-resistant cells.	4
<i>NME1-2</i>	NME/NM23 nucleoside diphosphate kinase 1-2	Differentially expressed in irinotecan-resistant cells.	4,5
<i>PRDX2</i>	Peroxiredoxin 2	Down-regulated in irinotecan-resistant cells.	5
<i>PRDX3</i>	Peroxiredoxin 3	Up-regulated in irinotecan-resistant cells.	5
<i>PRDX4</i>	Peroxiredoxin 4	Differentially expressed in irinotecan-resistant cells.	4,5
<i>CAVIN1 (PTRF)</i>	Caveolae associated protein 1	Silencing enhances the toxicity of SN-38 and 5-FU <i>in vitro</i> .	2
<i>RPL28</i>	Ribosomal protein L28	Silencing enhances the toxicity of SN-38 and 5-FU <i>in vitro</i> .	2
<i>TDPI</i>	Tyrosyl-DNA phosphodiesterase 1	<i>TDPI</i> IVS12+79 associated with response to irinotecan-based therapy in advanced colorectal patients. Involved in the removal of TOP1-DNA adducts.	1,6
<i>GSTP1</i>	Glutathione S-transferase pi 1	<i>GSTP1</i> Ile105Val was associated with survival and toxicities in mCRC patients.	7,8
<i>TIMP1</i>	TIMP metalloproteinase inhibitor 1	TIMP1 levels were associated with survival in mCRC patients receiving combination regimen including irinotecan and 5-FU.	9,10

Supplementary Table 2. Haplotype-tagging single nucleotide polymorphisms (htSNPs) and their associated SNPs.

Genes	htSNPs	Associated SNPs	
<i>CDC45</i>	rs4141528	rs4141527 rs5748228 rs5748227 rs5748222 rs5746741	
		rs5748241 rs77362076 rs5746756 rs5748212 rs150190062	
		rs5746757 rs5748211 rs143338547 rs79114187 rs2283656	
		rs13447273 rs117554245 rs13447294 rs13447302 rs12628900	
		rs79771248 rs45596741	
		rs5993650	rs5992403 rs5993649 rs5748231 rs35880706 rs2238772
		rs59690810 rs5748226 rs2073758 rs5993651 rs5993652	
		rs5993648 rs11704146 rs739369 rs9618585 rs756657 rs756658	
		rs5746742 rs4493360 rs2073759 rs2072055 rs9606025	
		rs5993647 rs11289637 rs5748219 rs13447232 rs5993654	
rs8137899 rs2097600 rs9604966 rs1002111 rs5993657			
rs9606024 rs5746738 rs2269722 rs2073761 rs8136840			
rs36124072 rs12484607 rs2073727 rs2073737 rs2073736			
rs9606042 rs7292006 rs1057457 rs9606043 rs2073730			
rs2073731 rs5993661 rs137937633 rs2283655 rs2073735			
rs2073733 rs2006242 rs9606048 rs739370 rs9606017			
rs737858	rs5746748 rs5746749 rs5748238 rs5746744 rs5746743		
		rs1547931 rs2238771 rs5748246 rs5748218 rs5748251	
		rs5748217 rs5748216 rs5746740 rs7364291 rs5748215	
		rs2269721 rs5748254 rs1548359 rs2904558 rs756649 rs712979	
		rs9604965 rs5748259 rs2871030 rs2871031 rs5748260	
		rs9618586	rs5748243 rs7289378
rs2073736	rs2073737 rs9606042 rs9606043 rs8136840 rs2073761		
		rs5993661 rs5993657 rs2283655 rs1002111 rs2097600	
		rs2073735 rs5993654 rs2073733 rs13447232 rs11289637	
		rs9606046 rs2073759 rs5992413 rs5993665 rs9618585	
		rs11704146 rs5993652 rs5993651 rs9606048 rs739370	
		rs59690810 rs2238772 rs5993650 rs5992403 rs5993649	
		rs5748231 rs35880706 rs5748226 rs2073758 rs5993648	
		rs739369 rs756657 rs756658 rs5746742 rs4493360 rs2072055	
		rs9606025 rs5993647 rs5748219 rs8137899 rs9604966	
		rs2073732	rs5746758 rs5748252 rs5748240
<i>MAPK9</i>	rs7724543	rs11310451 rs17080136 rs33928286 rs7730064 rs7730091	
		rs13162206 rs17627593 rs6874413 rs6896513 rs6875134	
		rs11960340 rs34428073 rs35203283 rs6601104 rs9605	
		rs7710896 rs56131100 rs13178637 rs61478999	

rs6877489	rs35394934 rs17627536 rs35759987 rs55756747 rs6894217 rs1127580 rs62406991 rs3099063 rs3111525
rs4362908	rs7726664 rs4481314 rs35081878 rs11955223 rs62404860 rs1127575 rs58528869
rs11956696	rs1035413 rs6898872 rs6898670 rs33979712 rs17683933 rs35588458 rs35254089 rs35691219 rs4601008 rs4639174 rs35624369 rs36010045 rs73811323 rs73811322 rs73811321 rs56320879 rs115200228 rs77310082 rs56264608 rs74399328 rs73338764 rs79172932
rs11743530	--
rs3812067	rs7715540 rs6601106 rs6872544 rs35777469 rs56658635 rs6601105 rs13169028 rs7729984
rs12522876	rs12518159 rs57403593 rs10052124 rs12519649 rs2112593
rs7445806	rs56333381 rs4634313 rs13435 rs34715282 rs6601103
rs3812066	rs35977675 rs72811017 rs72811018 rs113747850 rs888923 rs72811021 rs72811022 rs76752307 rs77249190 rs72811024 rs34876832 rs140492570 rs17628029
rs6896347	rs11249687
rs6867398	rs62404891 rs7720684 rs55913314 rs66800744 rs140356894 rs62404890 rs730688 rs7708873 rs7722861
rs3111515	--
rs6895740	rs3111537
rs4147385	--
rs4700730	rs11249689 rs7721150 rs6879043
rs7714207	--
rs13185784	rs17627880 rs17627815 rs35067749 rs6893005
rs6868333	rs33943370 rs6886915 rs732904 rs731918 rs13178637
rs7713083	--
<i>PARP1</i> rs1805415	rs3219065 rs2117114 rs907190 rs3219082 rs2255403 rs932002 rs2027440 rs1805411 rs3219104 rs2377312 rs907191 rs907192 rs907193 rs2048425 rs34742465 rs1136410 rs2793381 rs2793658 rs2271349 rs11378196 rs34917861 rs146042055 rs1805404 rs752307 rs752309 rs1891107 rs1891108 rs2695243 rs2695242 rs892348 rs907187 rs12143101 rs56033994 rs2793379 rs1317170 rs2249731 rs76887998 rs138994074 rs2570367 rs9943081 rs4653729 rs4653728 rs1433574
rs3219110	--

rs1805410	rs2271345 rs3219095 rs3219049 rs71801596
rs8679	rs2271347 rs61835377 rs555462858 rs2793383 rs1805403 rs74973162
rs3754376	rs3835704 rs3219112 rs3219090 rs3219119 rs747657 rs1805414 rs2027439 rs2048424 rs2793382 rs2377313 rs10631977 rs59275599 rs2793380 rs2666428 rs2793657 rs1417765 rs2048426 rs144361550 rs1341336 rs11366269 rs2793378 rs1104893 rs2793377 rs35170928 rs2377013 rs11304925 rs2249844 rs1397550 rs1828446 rs1828445 rs2793654 rs2666427 rs1865222 rs1865221 rs1865220 rs2695235 rs2793655 rs2793376 rs2695236 rs575478286 rs2695237 rs9426568 rs35242305 rs2695238 rs1341333 rs1341334 rs1341335 rs2695239 rs2570368 rs2570369 rs2136875 rs2695240 rs2570370 rs1858548 rs35380305 rs1858549 rs1858550
rs12568297	--
rs6668722	rs142801988 rs146044150 rs10915984 rs141794185 rs367797484 rs12119288
rs3219123	rs78087823
rs6668851	rs6702917 rs6426551 rs7517174
rs1805401	--
rs1805405	rs2280712 rs1002153 rs1000033 rs2136876 rs1073991 rs3219043 rs4653445 rs12240196 rs2293464 rs4653731 rs3046773 rs7548007 rs7542788 rs12068649 rs3219053 rs10799349 rs1805408 rs3219058 rs6701634 rs3219031 rs3219027 rs3219073 rs3219074 rs3219023 rs3219022 rs3754370 rs2271343 rs3768347 rs3768346 rs732284 rs7522351 rs7525191 rs4653732 rs4653733 rs60698376 rs4653734 rs7531668 rs2077197 rs7527192 rs59672299 rs1109032 rs3754375 rs6665208 rs532472660 rs141615498 rs4653735 rs6679573 rs10915987 rs77173384 rs28407557 rs3219115 rs12068460 rs12025487 rs752308 rs747658 rs747659 rs6681537 rs6664761 rs2282400 rs6675427 rs6675327 rs6661762 rs1991865 rs12092726 rs6700864 rs878367
rs3219142	rs68193591

<i>SART1</i>	rs688862	rs556643 rs552130 rs550435 rs1151505 rs660118 rs12797319 rs11606402 rs1192168 rs642293 rs56239996 rs55704213 rs35239693 rs593525 rs491973 rs501353 rs11227368 rs55928757 rs1211729 rs470192 rs56154109 rs649000 rs542332 rs1308733 rs679581 rs493899 rs28578718 rs551659 rs142280625 rs4994056 rs503524 rs659982 rs661335 rs566590 rs624273 rs1786175 rs606978 rs538954 rs1151512 rs141871632 rs531612 rs12792888 rs528736 rs7105628 rs56152520 rs11227354 rs1204649 rs617791 rs3016868 rs12361552 rs674363 rs613924 rs14157 rs12794370 rs147010511 rs590318 rs10609752 rs694156 rs61893932 rs655744 rs632578
	rs7952056	rs60109456 rs11227370 rs7939253 rs735942 rs35833455 rs10791835 rs7934543 rs34844246 rs55947178 rs3759035 rs12363614 rs7128076 rs10896073 rs9735063 rs1047464 rs7113428 rs2276132 rs10791837 rs747526 rs10896074 rs147982796 rs6591195 rs150188401
	rs754532	--
	rs586921	rs543952 rs622779 rs645571 rs1204011 rs1191716 rs1151510 rs1151511 rs1151514 rs9736731 rs559161 rs677350 rs610497 rs540584 rs507672 rs506873 rs1786171 rs1786172 rs522553 rs619701 rs55644043 rs1192184 rs590531 rs549334 rs502363 rs534201 rs35911045 rs616073 rs615646 rs1786174 rs11227377 rs11227378 rs75154345 rs632640 rs7933510 rs58833868 rs684546 rs548435 rs548495 rs668210 rs140250262
<i>ACTL6A</i>	rs12631988	rs12632721 rs73048716 rs10513761 rs4855090 rs4284980 rs4855091 rs4855092 rs16830587 rs55837066 rs2292907 rs2292905 rs7643532 rs78295396 rs16830600 rs74686416 rs4475040 rs149676657 rs35045240 rs12635617 rs61086485 rs73048727 rs79000506 rs73883565 rs10513762 rs2879605 rs113041545 rs3774260 rs113646525 rs73048730 rs111353757 rs114822697 rs6772319 rs59700512 rs7613710 rs11338795 rs6783120 rs6785779 rs113425680 rs4635704 rs112009668 rs4855097 rs3732998 rs16830629 rs73883552 rs2339844 rs73883571 rs113527645 rs58728627 rs4855098 rs12631863 rs4147787 rs6788205 rs4147792
	rs9816801	rs13322072 rs10575925 rs9870205 rs7644051 rs12696481 rs7632781 rs199924842 rs6790272 rs6790039 rs201684609 rs9831673 rs4855096 rs3830546 rs6778219

	rs1132429	rs6795642 rs9818068 rs71907893 rs199924842 rs9861172 rs7610384 rs7644051 rs7635877 rs13322072 rs35792644 rs6775032 rs12492220 rs5854839 rs113724918 rs7611557 rs7630511 rs9829395 rs6797961 rs4147788 rs9882051 rs13325172 rs7616758 rs7639131 rs2077675 rs2339846 rs10937007 rs12054439 rs10937008
<i>CBX3</i>	rs10255851	rs10233641 rs10277455 rs10247665 rs36032567 rs10236009 rs1025331 rs34876113 rs33998333 rs11971002 rs11974650 rs11983671 rs33998608 rs35060854 rs6974645 rs12700716 rs71521751 rs13227109 rs11486821 rs7803845
	rs2391265	rs3757668 rs3757667 rs12672942 rs2059903 rs573090483 rs7806787 rs12672536 rs10265394 rs2074566 rs576084528 rs12113404 rs10230838 rs7798786 rs34181077 rs10951128 rs6461920 rs2115547 rs12669140 rs10275247 rs67976730 rs34954459
<i>CFL1</i>	rs635375	rs668356 rs2308219 rs656040 rs652021 rs4621 rs665306 rs667555 rs7125986 rs7947929 rs7947741 rs508548 rs1939212 rs13817 rs659857 rs545500 rs551294829 rs633877 rs538394 rs648732 rs630758 rs630755 rs630303 rs11332468 rs558114 rs60409156 rs679147
	rs11227332	--
<i>CRYAB</i>	rs2070894	rs11603779 rs14133 rs1940391 rs7950926 rs12808601 rs12785809 rs3944619 rs11214037 rs11214036 rs4441050 rs71057102 rs7127010 rs10891297 rs57414412 rs7114617 rs75228430 rs11604704 rs4936834 rs11214029 rs11214028 rs34888817 rs35581942 rs12364361 rs12364358 rs7128862 rs7125417 rs1045282 rs11600652 rs10789851 rs10789850 rs7107213 rs7124696 rs4936804 rs7112886 rs7930208 rs72994380 rs34558891 rs4936800 rs11608193
	rs12284944	rs10891299 rs10789852 rs4252591 rs57377042 rs57527690 rs10502152 rs11214043 rs34183951 rs7124407 rs7124433 rs11384304 rs5794769 rs1320664 rs34088139
	rs762550	rs1940392 rs11214033 rs7114594 rs11214031 rs610747 rs2850245 rs634983 rs626243
	rs4252588	--
<i>NME1-2</i>	rs2318782	rs11654411 rs34520637 rs2302254 rs11652807 rs11656090 rs11652793 rs368050652 rs34666604 rs12941230 rs11655050 rs35347299 rs8075231 rs8074416 rs11654212 rs11651032 rs11650103 rs2041296 rs8071647 rs6504696 rs62063055 rs35523133 rs12941552 rs2159359 rs62063023 rs3744660

rs3760467	rs78766181 rs4794206 rs1558254 rs118144056 rs74470375 rs75410762 rs74970602 rs2215290 rs56209259 rs147722568 rs2318784
rs41524846	rs545197307 rs71381337
rs16949649	rs3760468 rs7207370 rs36055638 rs34431255 rs34214448 rs11079945 rs11079944 rs12451567 rs12936304 rs12952884 rs35995934 rs11869537 rs571079699 rs28710472
rs4605213	rs8079266 rs545197307 rs2041297
rs3760469	rs11556835 rs7213328 rs28597917 rs12939771 rs35517913 rs28710472 rs9303575 rs9303574
rs2318785	--
rs9303577	rs9303578 rs34313614 rs8081088 rs59280351 rs1062386 rs3803886 rs537942439 rs11650690 rs8069453 rs7213683 rs34581166 rs10545166 rs8076679 rs12453388 rs9912882 rs11650030 rs11870136 rs35289073 rs12449426 rs17574235 rs28410310 rs12946936 rs9915099 rs9789063 rs11651110
rs12942155	rs12945077 rs59280351 rs9915554 rs12451629 rs532770255 rs7220360 rs8081088 rs8080943 rs4794210 rs34313614 rs9303578 rs9907636 rs35397252 rs9894098 rs537942439 rs11650690 rs929482 rs8069453 rs1062386 rs11079946 rs8079807 rs7213683 rs34581166 rs10545166 rs8076679 rs7222463 rs35987233 rs12453388 rs9912882 rs12453513 rs35172874 rs11650030 rs35269451 rs11870136 rs35289073 rs12449426 rs4273084 rs17574235 rs28410310 rs368640927 rs3815411 rs12946936 rs113570976 rs9915264 rs9915099 rs11656911 rs9789063 rs71300606 rs2079883 rs9910220 rs11651110 rs1860574 rs6416962 rs7207416 rs2215417 rs9905051 rs11079947 rs8076651 rs10550602
rs12937424	rs62063028 rs35292418 rs62063027 rs7208601 rs62063026 rs11868380 rs11869672 rs111481382 rs113131698
<i>PRDX2</i>	rs10427027 rs12151144 rs116813755 rs10421604 rs7507935
	rs1205171 rs12609353 rs35858668
<i>PRDX3</i>	rs1553850 rs139394880 rs61876564 rs7080535 rs71016553 rs75475521 rs74566533 rs370067252 rs7078965 rs10886414 rs11198816 rs35439312 rs1247800 rs12412287 rs1108473 rs1108471 rs1108472 rs11594600 rs34061558 rs10510054 rs4752262 rs10886417 rs10749299 rs10886421 rs10787929 rs7923896 rs61876593 rs4752263 rs1980030

	rs3740562	rs11198811 rs34060649 rs10886412 rs2271362 rs7768 rs10886413 rs1889742 rs7894367 rs1107942 rs35148933 rs10430698 rs146722239 rs7068409 rs7068525 rs7068698 rs7358070 rs10787925 rs10787926 rs35838962 rs7087504 rs10886418 rs11198819 rs7082425 rs7082346 rs10886419
	rs10886406	rs2271362
<i>PRDX4</i>	rs795493	rs795494 rs795492 rs795491 rs473769 rs518329 rs552105 rs490513 rs1548734 rs2665364 rs497797 rs795485 rs795486 rs112139616 rs570823 rs11412736 rs547679 rs496635 rs795488 rs513573 rs513572 rs477204 rs478034 rs482702 rs6629721 rs512997 rs1704732 rs565208 rs693806 rs7063461 rs528683 rs528834 rs71817516 rs557914 rs496067 rs475171 rs525443 rs35647919 rs551029 rs557637
	rs528960	rs497087 rs544971 rs35477162 rs795495 rs3838952 rs2532625 rs2109142 rs11094937 rs2665367 rs2159383 rs2665369 rs480691 rs34759783 rs476316 rs477914 rs7881384 rs5925869 rs2532626 rs518857 rs568062 rs812704 rs511120 rs140643281 rs572174 rs570466 rs569603 rs543050 rs2532630 rs2665371 rs578606 rs497967 rs552109 rs498728 rs6629723 rs2665363
	rs795489	rs477233 rs576656 rs501655 rs503572 rs73209274 rs483169 rs527749 rs113389417 rs113128237
	rs3829725	rs6629722 rs7056309 rs6629724 rs6627966 rs6629725 rs6629719 rs7052367 rs6627965 rs6629718 rs6627964 rs6629717 rs3795226 rs6629716 rs12384006
	rs564561	rs478051
	rs3795220	--
	rs525497	--
	rs561909	--
	rs499692	rs5901737 rs35250716 rs474657
<i>CAVINI</i> (<i>PTRF</i>)	rs7212299	rs7210713 rs4796582 rs6503699 rs6416922 rs77390070 rs6416923 rs112865149 rs36116679 rs7207285 rs111383744 rs7503929 rs8074832 rs12950246 rs56387576 rs9906657 rs9909871 rs8079445 rs4796649 rs4796650 rs11422354 rs7502611 rs9709260 rs59371909 rs35972838 rs7221385 rs17841509 rs7217251

rs7223784	rs560819092 rs34967087 rs12937756 rs560117712 rs8068008 rs12936165 rs35272895 rs34610643 rs34156075 rs57307201 rs537482133 rs8064439 rs8081779 rs143185919 rs35470081 rs11871801 rs11867405 rs12938897 rs12938856 rs12948909 rs35510462 rs78594996 rs72823064 rs12938957 rs72029736 rs57203370 rs7213500 rs34527783 rs56001200 rs12936869 rs35261210 rs71367913 rs6503703 rs6503704 rs12951549 rs16649 rs12943498 rs12951640 rs35111218 rs142647634 rs36005199 rs375125601 rs36075544 rs34633576 rs149422595 rs34460267 rs1032070 rs34807589
rs2883456	rs11653437 rs709631 rs2128786 rs12453307 rs963987 rs963988 rs12150088 rs34154026 rs56364076
rs1032072	rs1032071 rs58747199 rs72823071 rs151285077 rs2062213 rs12603814 rs1968866 rs8071329 rs367572409 rs371774724 rs76681937 rs2883457 rs72626001 rs191327332 rs141580568 rs140220058 rs117512511 rs72823083 rs66774412 rs115341059 rs67560218 rs116971005 rs3760387 rs140294647 rs16967794 rs72626002 rs144967292 rs7224647 rs6416924 rs138251779 rs145289234 rs201794971
rs9252	--
rs963986	rs4365333 rs72823047 rs111550283 rs8078850 rs7222712 rs1129422
<i>RPL28</i> rs4806668	rs3810168 rs3842412 rs3745272 rs1870074 rs73617860 rs73617855
rs12981911	rs10418020
rs17700376	--
rs10425596	rs11673539
rs7255648	rs7255657 rs56283327 rs78892292 rs200296013 rs142797565 rs78935016 rs74683747 rs78633024 rs75479073 rs11882786 rs3810167 rs12981704 rs141481594 rs148705147 rs76684947
<i>TDPI</i> rs9488	rs17126538 rs10645574 rs61520219 rs112332158 rs35522230 rs11851131 rs10710186 rs8003210 rs8005676 rs2401866 rs67169479 rs7150605 rs11310221
rs4143999	rs3759710 rs2181599 rs28736874 rs34348197 rs12232164 rs146704566 rs17260562 rs10498624
rs942190	rs4904629 rs7147996
rs7158356	rs73324579
rs1286927	--
rs119467003	--
rs2176075	rs746752 rs8010627 rs35972056

rs28365054	rs143694338 rs34510655 rs75021372 rs73324561 rs34080955 rs34258347 rs11849310 rs11844949 rs75530159 rs35913426 rs35782158 rs35257587 rs34705515 rs113464893 rs34458659 rs113385610 rs17126508 rs34533254 rs34254473 rs113469707 rs144733015 rs58917724 rs57055713 rs36109363 rs35819811 rs79690576 rs80153327 rs78402658 rs34614135 rs35151660 rs34086909 rs34075974 rs35211637 rs36005472 rs35320469 rs8016664 rs8020872 rs8020897 rs113492506 rs11848062 rs17126510 rs34587290 rs35666995 rs35483543 rs17126513 rs35230069 rs34612773 rs35548006 rs35766668 rs17126514 rs58176524 rs35787421 rs7150748 rs7150652 rs7145132 rs75781929 rs34567445 rs34680032 rs35178878 rs75745280 rs7153551 rs35501816 rs535016899 rs8015261 rs8015277 rs8014111 rs8014295 rs34259367 rs8008263 rs11851399 rs11851415 rs35853694 rs34601269 rs8003068 rs12881272 rs12880397 rs12879968 rs34745798 rs34277635 rs35434877 rs34507901 rs59736885 rs12101181 rs112528751 rs73324588 rs12101164 rs1399742 rs78138961 rs148839615 rs11846794 rs75911102 rs113053120 rs112036275 rs35653575 rs6575100 rs8006050 rs34657096 rs6575098 rs7151882 rs75356978 rs74710063 rs7152455 rs17126520 rs34088863 rs35664925 rs17126522 rs36082908 rs17126485 rs59577883 rs112115508 rs111546207 rs111311830 rs77461748 rs10498623 rs8003693 rs17798962 rs9919910 rs34792766 rs151254830 rs4603488 rs17126527 rs77951315 rs111232232 rs78465035 rs58017722 rs74899305 rs61073988 rs79098962 rs10130775 rs10142587 rs10142377 rs10569454 rs201118162 rs11850323 rs11849387 rs10143622 rs149550699
rs28365055	rs28422802 rs2277518 rs60362204
rs11851266	rs34290480 rs8022306 rs7154092 rs2401863 rs12884144
<i>GSTP1</i>	rs6591256 rs6591255 rs7949587 rs7949394 rs17593068 rs11311625 rs36211088 rs36211089 rs2370141 rs1079719 rs7945035 rs2370143 rs7941648 rs762803 rs7941395 rs7927657 rs7940813 rs6591254 rs6591253 rs597297 rs6591252 rs6591251 rs4930444 rs7948911 rs6591249 rs7945927 rs7108426 rs7103713 rs35534379 rs12807265 rs10896181 rs55854833 rs4930436 rs7952081 rs7948823 rs7948478 rs75649061 rs61390583 rs11605586 rs7106675 rs7106423 rs34721562 rs34802338 rs6591245 rs4320958 rs1871036 rs200602459 rs555435772 rs6591244 rs7938563 rs57808037

rs1695	rs749174 rs1871042 rs4891 rs947895 rs7952248 rs6591258 rs7104580 rs11603735 rs4360710 rs7939505 rs6591259 rs7937159 rs7950707 rs4244823 rs11227845 rs7108149 rs7108038 rs12416994 rs11603991 rs7103632 rs12790798 rs12792496 rs35247127 rs12798262 rs572499532 rs4930450 rs4024254 rs12421329 rs12793832 rs1871043
rs4147581	rs694746 rs57025443 rs11278380 rs614080 rs625978 rs656652 rs674783 rs591434 rs658768 rs645653 rs11227837 rs1262166 rs666328 rs1269981 rs1254123 rs638140 rs10444362 rs55740781 rs684928 rs12225376 rs2508459 rs35840098 rs2008922 rs57387696 rs10896180 rs10896179 rs35829185 rs611663 rs598811 rs10791906 rs1148884 rs1148883 rs12365485
rs7927381	rs593055 rs612020 rs3082142 rs1254133 rs641544 rs643318 rs643317 rs137897748 rs655848 rs656798 rs627449 rs1871041 rs1254130 rs4930443 rs594784 rs685032 rs689317 rs688878 rs676653 rs675679 rs6591247 rs4930204 rs1254125 rs7128347 rs604908 rs669956 rs11405928 rs601531 rs604232 rs667014 rs615260 rs10667536 rs657003 rs644670 rs615270 rs612162 rs598398 rs694616 rs649149 rs661624 rs687316 rs34961815 rs584338 rs595262 rs629952 rs78166531
rs1138272	rs72934504
<i>TIMPI</i> rs2070584	rs6609533 rs4898 rs6520279 rs5953060 rs5905615 rs34026683
rs6520278	rs12556415 rs5905614 rs5906434 rs6609534 rs55696512 rs111947105
rs4824621	rs4824622 rs62592125 rs150642282 rs5953061 rs5953062 rs5953063 rs5906437 rs9887335 rs5953066 rs4824424 rs4824623

Abbreviations: htSNPs, haplotype tagging SNPs; SNPs, single nucleotide polymorphisms. Associated SNPs are polymorphisms in strong linkage with the htSNP with $r^2 > 0.80$, as determined with the data of 1000 Genomes Phase 3 for the CEU population (Ensembl GRCh38 release 91 - December 2017).

Supplementary Table 3. htSNPs significantly ($P < 0.1$) associated with progression-free survival (PFS) in the Canadian cohort.

Genes	htSNPs	Base change	Model	HR_{adj} (95% CI)^a	P
<i>ACTL6A</i>	rs9816801	G>C	dominant	0.74 (0.53 - 1.05)	0.095
<i>CBX3</i>	rs2391265	T>C	dominant	1.48 (1.00 - 2.20)	0.050
<i>CFL1</i>	rs11227332	A>G	recessive	2.29 (1.05 - 5.01)	0.038
<i>GSTP1</i>	rs6591256	A>G	dominant	0.74 (0.53 - 1.04)	0.080
<i>GSTP1</i>	rs4147581	C>G	dominant	0.54 (0.37 - 0.80)	0.002
<i>MAPK9</i>	rs4362908	T>C	dominant	0.59 (0.41 - 0.86)	0.006
<i>MAPK9</i>	rs7724543	A>G	dominant	1.48 (1.02 - 2.15)	0.039
<i>MAPK9</i>	rs6896347	G>A	dominant	0.64 (0.43 - 0.96)	0.030
<i>MAPK9</i>	rs11956696	A>G	dominant	1.42 (1.01 - 2.01)	0.047
<i>MAPK9</i>	rs4147385	G>A	recessive	1.88 (0.95 - 3.72)	0.071
<i>MAPK9</i>	rs7445806	C>T	dominant	0.69 (0.49 - 0.97)	0.031
<i>NME1-2</i>	rs9303577	A>C	recessive	1.43 (0.96 - 2.12)	0.079
<i>NME1-2</i>	rs2318785	A>G	dominant	1.37 (0.97 - 1.94)	0.077
<i>PARP1</i>	rs3219110	T>C	recessive	1.39 (0.95 - 2.03)	0.087
<i>PRDX4</i>	rs795493	A>G	dominant	0.65 (0.47 - 0.92)	0.014
<i>PRDX4</i>	rs525497	C>G	dominant	1.35 (0.95 - 1.92)	0.092
<i>PRDX4</i>	rs528960	A>G	recessive	0.50 (0.24 - 1.02)	0.057
<i>PRDX4</i>	rs795489	A>G	recessive	1.44 (1.02 - 2.05)	0.039
<i>PRDX4</i>	rs561909	C>A	recessive	1.56 (0.94 - 2.59)	0.089
<i>RPL28</i>	rs4806668	G>T	recessive	3.23 (1.29 - 8.14)	0.013
<i>RPL28</i>	rs10425596	T>C	recessive	1.91 (0.93 - 3.93)	0.080

^a Adjusted for age and cotreatment. Abbreviations: CI, confidence interval; HR_{adj}, adjusted hazard ratio; htSNPs, haplotype tagging single nucleotide polymorphisms.

Supplementary Table 4. htSNPs significantly ($P < 0.1$) associated with overall survival (OS) in the Canadian cohort.

Genes	htSNPs	Base change	Model	HR_{adj} (95% CI)^a	P
<i>ACTL6A</i>	rs9816801	G>C	dominant	0.70 (0.49 - 1.01)	0.053
<i>CBX3</i>	rs2391265	T>C	dominant	1.48 (0.99 - 2.20)	0.054
<i>MAPK9</i>	rs4362908	T>C	dominant	0.58 (0.40 - 0.85)	0.005
<i>MAPK9</i>	rs7724543	A>G	dominant	1.41 (0.97 - 2.06)	0.073
<i>MAPK9</i>	rs12522876	C>A	dominant	0.69 (0.45 - 1.05)	0.082
<i>MAPK9</i>	rs6896347	G>A	dominant	0.58 (0.38 - 0.89)	0.013
<i>MAPK9</i>	rs11956696	A>G	recessive	2.03 (0.88 - 4.70)	0.099
<i>MAPK9</i>	rs4147385	G>A	recessive	1.99 (1.00 - 3.96)	0.050
<i>NME1-2</i>	rs9303577	A>C	recessive	1.62 (1.07 - 2.44)	0.023
<i>PARP1</i>	rs3219123	G>A	dominant	0.58 (0.31 - 1.07)	0.083
<i>PRDX3</i>	rs1553850	T>A	dominant	1.38 (0.96 - 1.98)	0.085
<i>PRDX4</i>	rs795493	A>G	dominant	0.71 (0.51 - 1.01)	0.053
<i>CAVIN1</i> (<i>PTRF</i>)	rs7223784	A>C	recessive	2.90 (1.33 - 6.32)	0.008
<i>RPL28</i>	rs4806668	G>T	recessive	3.09 (1.10 - 8.65)	0.032

^a Adjusted for age and cotreatment. Abbreviations: CI, confidence interval; HR_{adj}, adjusted hazard ratio; htSNPs, haplotype tagging single nucleotide polymorphisms.

Supplementary Table 5. The marker rs4806668G>T and its linked SNPs significantly altered the expression of *RPL28* in the transverse colon (GTEx dataset), but not the expression of the nearby *TMEM238* gene.

Gene	SNP	P-value
<i>RPL28</i>	rs4806668G>T	0.0067
<i>RPL28</i>	rs3810168T>C	0.00057
<i>RPL28</i>	rs3745272C>G	0.013
<i>RPL28</i>	rs73617860T>G	0.0082
<i>RPL28</i>	rs73617855T>A	0.0035
<i>TMEM238</i>	rs4806668G>T	0.92
<i>TMEM238</i>	rs3810168T>C	0.79
<i>TMEM238</i>	rs3745272C>G	0.89
<i>TMEM238</i>	rs73617860T>G	0.87
<i>TMEM238</i>	rs73617855T>A	0.90

Data were not available for rs3842412-/15-mer and rs1870074A>G. $P < 0.05$ are in bold.

Supplementary References.

- 1 Hoskins, J. M. *et al.* Irinotecan pharmacogenetics: influence of pharmacodynamic genes. *Clin Cancer Res* **14**, 1788-1796, (2008).
- 2 Allen, W. L. *et al.* A systems biology approach identifies SART1 as a novel determinant of both 5-fluorouracil and SN38 drug resistance in colorectal cancer. *Mol Cancer Ther* **11**, 119-131, (2012).
- 3 Davidson, D., Wang, Y., Aloyz, R. & Panasci, L. The PARP inhibitor ABT-888 synergizes irinotecan treatment of colon cancer cell lines. *Invest New Drugs* **31**, 461-468, (2013).
- 4 Peng, X. C. *et al.* Proteomic analysis of cell lines to identify the irinotecan resistance proteins. *J Biosci* **35**, 557-564, (2010).
- 5 Gong, F. M. *et al.* Comparative proteomic analysis of irinotecan-sensitive colorectal carcinoma cell line and its chemoresistant counterpart. *Anticancer Drugs* **22**, 500-506, (2011).
- 6 Pommier, Y., Leo, E., Zhang, H. & Marchand, C. DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem Biol* **17**, 421-433, (2010).
- 7 Kweekel, D. M. *et al.* GSTP1 Ile105Val polymorphism correlates with progression-free survival in MCRC patients treated with or without irinotecan: a study of the Dutch Colorectal Cancer Group. *Br J Cancer* **99**, 1316-1321, (2008).
- 8 Braun, M. S. *et al.* Association of molecular markers with toxicity outcomes in a randomized trial of chemotherapy for advanced colorectal cancer: the FOCUS trial. *J Clin Oncol* **27**, 5519-5528, (2009).
- 9 Sorensen, N. M. *et al.* TIMP-1 is significantly associated with objective response and survival in metastatic colorectal cancer patients receiving combination of irinotecan, 5-fluorouracil, and folinic acid. *Clin Cancer Res* **13**, 4117-4122, (2007).
- 10 Bystrom, P. *et al.* Evaluation of predictive markers for patients with advanced colorectal cancer. *Acta Oncol* **51**, 849-859, (2012).

Chapitre 3 : Improved progression-free survival in irinotecan-treated metastatic colorectal cancer patients carrying the HNF1A coding variant p.I27L

3.1 Résumé

Le facteur de transcription HNF1A (*hepatocyte nuclear factor 1-alpha*) est enrichi au foie, où il joue un rôle essentiel dans plusieurs aspects de la fonction hépatique incluant celle de détoxification. Nous avons examiné si les polymorphismes de *HNF1A* étaient associés aux issues cliniques de patients atteints du CCRm et traités en première intention avec une chimiothérapie à base d'irinotécan (FOLFIRI). Cette étude a été réalisée dans 2 cohortes indépendantes combinant 417 patients. Le marqueur intronique rs2244608A>G prédit une augmentation de la survie sans progression (PFS) dans la première cohorte avec un HR de 0,74 ($P = 0,076$) et avec un HR de 0,72 dans la deuxième cohorte ($P = 0,038$). Une forte association entre une PFS prolongée et le rs2244608A>G a aussi été observée en combinant les 2 cohortes (HR = 0,72, $P = 0,002$). Appuyant une altération de la fonction de HNF1A, les patients porteurs de l'allèle mineur G présentaient une exposition au médicament supérieure de 45% ($P = 0,032$) comparativement aux non-porteurs. Chez les Caucasiens, le rs2244608A>G est en fort déséquilibre de liaison avec le variant rs1169288c.79A>C codant pour la substitution HNF1A p.I27L. Chez les donneurs sains de la cohorte GTEx porteurs de l'allèle variant rs1169288C, nous avons observé une altération hépatique (*ABCC1*, $P = 0,009$, *ABCC2*, $P = 0,048$, *CYP3A5*, $P = 0,001$; $n = 89$) et intestinale (*TOP1*, $P = 0,004$; $n = 75$) de l'expression de plusieurs gènes impliqués dans la pharmacologie de l'irinotécan. De plus, la présence du variant rs1169288C entraîne une augmentation de l'activité du promoteur proximal *ABCC1* de 27% ($P = 0,008$) et 15% ($P = 0,041$), dans les lignées de rein humain HEK293 et de foie humain HepG2, respectivement. Nos découvertes suggèrent que le variant rs2244608A>G de *HNF1A* (ou le variant fonctionnel p.I27L qui en déséquilibre de liaison), pourrait être un marqueur pronostic potentiel.

3.2 Abstract

Hepatocyte nuclear factor 1-alpha (HNF1A) is a liver-enriched transcription factor that plays a key role in many aspects of hepatic function including detoxification processes. We examined whether *HNF1A* polymorphisms are associated with clinical outcomes in two independent cohorts combining 417 European ancestry patients with metastatic colorectal cancer (mCRC) treated with irinotecan-based chemotherapy. The intronic rs2244608A>G marker was predictive of an improved progression-free survival with a trend in the Canadian cohort and reaching significance in the Italian cohort, with hazard ratios (HR) of 0.74 and 0.72, $P = 0.076$ and 0.038 , respectively. A strong association between rs2244608A>G and improved PFS was found in the combined analysis of both cohorts (HR = 0.72; $P = 0.002$). Consistent with an altered HNF1A function, mCRC carriers of the rs2244608G minor allele displayed enhanced drug exposure by 45% ($P = 0.032$) compared to non-carriers. In Caucasians, rs2244608A>G is in strong linkage with the coding variant rs1169288c.79A>C (HNF1A p.I27L). In healthy donors, we observed an altered hepatic (*ABCC1*, $P = 0.009$, *ABCC2*, $P = 0.048$ and *CYP3A5*, $P = 0.001$; $n = 89$) and intestinal (*TOPI1*, $P = 0.004$; $n = 75$) gene expression associated with the rs1169288C allele. In addition, the rs1169288C polymorphism could significantly increase the *ABCC1* promoter activity by 27% ($P = 0.008$) and 15% ($P = 0.041$) in the human kidney HEK293 and the human liver HepG2 cell lines, respectively. Our findings suggest that the *HNF1A* rs2244608, or the tightly linked functional coding variant p.I27L, might be a potential prognostic marker with irinotecan-based regimens.

Improved progression-free survival in irinotecan-treated metastatic colorectal cancer patients carrying the HNF1A coding variant p.I27L

Adrien Labriet^{1,2§}, Elena De Mattia^{3§}, Erika Cecchin³, Éric Lévesque^{4,5}, Derek Jonker⁶, Félix Couture^{4,5}, Angela Buonadonna⁷, Mario D'Andrea⁸, Lyne Villeneuve^{1,2}, Giuseppe Toffoli³ and Chantal Guillemette^{1,2*}

§Authors contributed equally

¹Pharmacogenomics Laboratory, Centre Hospitalier Universitaire (CHU) de Québec Research Center, Québec, QC, Canada;

²Faculty of Pharmacy, Laval University, Québec, QC, Canada;

³Clinical and Experimental Pharmacology, IRCCS National Cancer Institute 'Centro di Riferimento Oncologico', Aviano, Italy;

⁴Centre Hospitalier Universitaire (CHU) de Québec Research Center, Québec, QC, Canada;

⁵Faculty of Medicine, Laval University, Québec, QC, Canada;

⁶Division of Medical Oncology, Department of Medicine, Ottawa Hospital, University of Ottawa, Ottawa, ON, Canada;

⁷Medical Oncology Unit, IRCCS National Cancer Institute 'Centro di Riferimento Oncologico', Aviano, Italy;

⁸Medical Oncology Unit, San Filippo Neri Hospital, Rome, Italy;

*Corresponding Author: Chantal Guillemette, Ph.D., CHU de Québec Research Center, R4701.5, 2705 Blvd. Laurier, Québec, Canada, G1V 4G2. E-mail: chantal.guillemette@crchudequebec.ulaval.ca

KEYWORDS: metastatic colorectal cancer, hepatocyte nuclear factor 1-alpha, irinotecan, polymorphism, progression-free survival

Running Title: HNF1A p.I27L predicts progression-free survival in metastatic colorectal cancer patients

Abstract: 252

Introduction: 505

Materials and Methods: 1237

Results: 934

Discussion: 1301

References: 44

NUMBER OF:

Tables: 2

Figures: 7

Supplementary Tables: 4

Supplementary Figures: 1

ABSTRACT

Hepatocyte nuclear factor 1-alpha (HNF1A) is a liver-enriched transcription factor that plays a key role in many aspects of hepatic function including detoxification processes. We examined whether *HNF1A* polymorphisms are associated with clinical outcomes in two independent cohorts combining 417 European ancestry patients with metastatic colorectal cancer (mCRC) treated with irinotecan-based chemotherapy. The intronic rs2244608A>G marker was predictive of an improved progression-free survival with a trend in the Canadian cohort and reaching significance in the Italian cohort, with hazard ratios (HR) of 0.74 and 0.72, $P = 0.076$ and 0.038, respectively. A strong association between rs2244608A>G and improved PFS was found in the combined analysis of both cohorts (HR = 0.72; $P = 0.002$). Consistent with an altered HNF1A function, mCRC carriers of the rs2244608G minor allele displayed enhanced drug exposure by 45% ($P = 0.032$) compared to non-carriers. In Caucasians, rs2244608A>G is in strong linkage with the coding variant rs1169288c.79A>C (HNF1A p.I27L). In healthy donors, we observed an altered hepatic (*ABCC1*, $P = 0.009$, *ABCC2*, $P = 0.048$ and *CYP3A5*, $P = 0.001$; $n = 89$) and intestinal (*TOPI1*, $P = 0.004$; $n = 75$) gene expression associated with the rs1169288C allele. In addition, the rs1169288C polymorphism could significantly increase the *ABCC1* promoter activity by 27% ($P = 0.008$) and 15% ($P = 0.041$) in the human kidney HEK293 and the human liver HepG2 cell lines, respectively. Our findings suggest that the *HNF1A* rs2244608, or the tightly linked functional coding variant p.I27L, might be a potential prognostic marker with irinotecan-based regimens.

INTRODUCTION

The camptothecin derivative irinotecan (CPT-11) is a widely used chemotherapeutic agent for the treatment of solid tumors, particularly for metastatic carcinoma of the colon or rectum. A variety of first-line regimens such as irinotecan, 5-fluorouracil (5-FU) and leucovorin (FOLFIRI) used for metastatic colorectal cancer (mCRC) in combination with targeted therapy improves outcomes (Fujita et al., 2015). Its active metabolite SN-38 formed by the action of carboxylesterases inhibits the replication enzyme topoisomerase I (TOP1), inducing DNA damage and cell death in actively replicating cells leading to S-phase-specific cell killing. The clinical benefits from irinotecan-based regimens are variable among patients whereas sequencing and method of administration influence toxicity. Predictive and prognostic biomarkers are critical in identifying those who most benefit from the chemotherapy. A number of genetic variants in genes involved in irinotecan pharmacology have been described and may help predict benefit of irinotecan-based chemotherapy. Most notably, the polymorphism *UGT1A1**28, leading to a decreased *UGT1A1* enzyme expression and a reduced glucuronidation of the active metabolite SN-38, is a well-established marker of irinotecan-induced severe neutropenia and help guide dosing for treatment of mCRC (Beutler et al., 1998; Barbarino et al., 2014). In addition to help in decreasing the incidence of adverse drug reactions, pharmacogenetic markers can also be useful in predicting irinotecan responsiveness. Irinotecan pharmacology is complex and involves activation into SN-38 by carboxylesterases (CES), inactivation of SN-38 by UDP-glucuronosyltransferases (UGT)-dependent glucuronidation and cytochrome P450 (CYP)-dependent oxidation and hepatic transport of irinotecan and its metabolites by multiple transporters, all potentially contributing to the bioavailability and pharmacological activity of SN-38 (Smith et al., 2006).

Whereas genetic variants in genes regulating irinotecan pharmacology have been investigated, the role of nuclear receptors and other transcriptional factors has not been investigated, despite their potentially broader influence on drug metabolism (Cecchin et al., 2016; De Mattia et al., 2016). Hepatocyte nuclear factor 1-alpha (HNF1A) is a master hepatic transcriptional regulator that influences the expression of numerous genes critical to liver functions. Consistent with an important role in hepatic functions, *HNF1A* locus and

genetic variations in *HNF1A* are associated with maturity-onset diabetes of the young (MODY3) (Yamagata et al., 1996), type 2 diabetes (Holmkvist et al., 2006; Voight et al., 2010), and coronary artery disease (Liu et al., 2014; Zhou et al., 2017). *HNF1A* also regulates a number of key genes involved in drug metabolism and disposition including *UGT* and ATP-binding cassette sub-family C member 2 (*ABCC2*) expression (Odom et al., 2004; Odom et al., 2007; Aleksunes et al., 2009; Qadri et al., 2009; Hu et al., 2014). Here, we hypothesized that polymorphisms in *HNF1A* may influence irinotecan pathways and predict clinical outcomes in mCRC patients. We tested the predictive significance of haplotype-tagging single nucleotide polymorphisms (htSNPs) in two independent cohorts of mCRC patients of European descent from Canada and Italy treated with the FOLFIRI regimen, based on irinotecan (CPT-11) in combination with folinic acid (leucovorin) and 5-fluorouracil (5-FU). We also study the link with drug exposure in mCRC patients and tissular expression of genes potentially targeted by *HNF1A* and relevant to irinotecan pharmacology.

MATERIALS AND METHODS

Patient characteristics. A first cohort of 167 mCRC Canadian patients and a second cohort comprising 250 mCRC Italian patients were studied. Eligibility criteria included no prior irinotecan-based chemotherapy, histologically confirmed mCRC, a life expectancy of at least 3 months and a good performance status (Eastern Cooperative Oncology Group/World Health Organization Performance Status ≤ 2). White patients from the Canadian cohort were recruited from 2003 to 2012 in three medical centers of eastern Canada and those from the Italian cohort were enrolled from 2002 to 2005 in thirteen medical centers of Northeast Italy. All the patients from the Canadian cohort and the majority ($> 90\%$) of the Italian patients were treated with the modified FOLFIRI regimen (irinotecan 180 mg/m² intravenously for 2 hours on day 1 plus 5-fluorouracil (5-FU) 400 mg/m² bolus followed by continuous infusion for 46 hours of 5-FU 2400 mg/m² plus leucovorin 200 mg/m²) every two weeks. The remaining Italian patients received a FOLFIRI regimen (irinotecan 180 mg/m² intravenously for 2 hours on day 1 plus 5-FU 400 mg/m² bolus followed by 5-FU 600 mg/m² continuous infusion during 22 hours on days 1 and 2 + leucovorin 200 mg/m² on days 1 and 2 every 2 weeks). Sixty-nine patients from the Canadian cohort received bevacizumab as a co-treatment and six others received another drug or a placebo. Genomic DNA isolation from blood samples was as described (Toffoli et al., 2006; Levesque et al., 2013). For a subset of Italian mCRC patients with available *HNF1A* genotype (n = 49), pharmacokinetics data including total plasma concentration of irinotecan and its metabolites, SN-38 and SN-38G assessed on a serial blood samples collected after drug administration and using high-performance liquid chromatography, were available (Toffoli et al., 2006). Biliary index was also calculated with the formula: [CPT-11 area under the curve (AUC) X (SN-38 AUC/SN-38G AUC)] and glucuronidation ratio with the formula: (SN-38G AUC/SN-38 AUC). Overall survival (OS), progression-free survival (PFS) and response rate data in relation to patient characteristics and toxicity in both cohorts are shown in **Table 1**. This study was carried out in accordance with the regulatory framework for chemical and biological hazards of the CHU de Québec and Centro di Riferimento Oncologico di Aviano. All subjects provided written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Comitato Etico

Indipendente- Centro di Riferimento Oncologico di Aviano and the CHU de Quebec ethics committees.

Genetic analysis. Single-nucleotide polymorphisms in the *HNF1A* gene and 5 kb flanking regions were identified using the CEU population of the International HapMap Project information (<http://www.hapmap.org>). Haplotype-tagging single nucleotide polymorphisms (htSNPs) were selected using Haploview v4.2 to tag for SNPs in high linkage disequilibrium ($r^2 > 0.8$) (Broad Institute, Cambridge, MA, USA) (Barrett et al., 2005). Allele frequencies and linkage disequilibrium (LD) data from the 1000 Genomes Project Phase 3 were obtained through the 1000 Genomes Browser – Ensembl (<http://grch37.ensembl.org/index.html>), assembly GRCh37.p13, accessed on March 22, 2017. The list of htSNPs ($n = 13$) and their associated SNPs is provided in **Supplementary Table 1**. Genotyping of htSNPs was performed by Sequenom iPLEX matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Sequenom, San Diego, CA, USA) and using the Illumina BeadXpress platform (Illumina Inc., San Diego, CA, USA). Experimental design of PCR primers, extension primers and genotyping conditions were defined with the SpectroDESIGNER software (Sequenom).

Statistical analysis. Deviation from Hardy-Weinberg equilibrium was calculated for each htSNP using PLINK v1.07 (Purcell et al., 2007) and those deviating from the equilibrium ($P < 0.05$) were not retained for further analysis. Overall survival (OS) was defined as the period of time between initiation of irinotecan treatment and death. Progression-free survival (PFS) was defined as the period of time between initiation of irinotecan treatment and the first evidence of disease progression, death or last-follow-up. For response rate, patients with complete or partial response were compared to patients with stable disease or progression. Genetic associations with OS and PFS were tested using Cox proportional hazards model (SAS version 9.2). Associations with response rate and severe toxicities (neutropenia and diarrhea) were assessed by logistic regression. Hazard ratios (HRs) and odds ratios (ORs) were adjusted for covariates including age, and co-treatment for the Canadian cohort and adjusted for gender, age, cancer primary site, stage at diagnosis, radical surgery and adjuvant chemotherapy in the Italian cohort, as performed in previous reports (Toffoli et al., 2006; Cecchin et al., 2009; Levesque et al., 2013; Chen et al., 2015a;

Chen et al., 2015b). Additive, dominant and recessive models were fitted independently. Each marker was tested independently in the two cohorts and then the results were compared. A marker was considered validated when the same effect using the same genetic model was observed in both populations, a trend observed in the Canadian cohort ($P < 0.1$) replicated in the Italian cohort with a P value < 0.05 . Analysis of pharmacokinetics parameters was performed using GraphPad Prism 5 (GraphPad Software Inc., La Jolla, CA, USA) and differences according to genotypes were tested using a two-tailed Mann-Whitney test. Gene expression data in tissues and associated genotypes were obtained from the GTEx Project portal (<http://www.gtexportal.org/home/> on March 16, 2017). Analyses were carried out with GraphPad Prism 5 and differences among groups were tested using a two-tailed Mann-Whitney test. The datasets were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000424.v6.p1; project ID 13346. Gene expression values were quantile normalized, as detailed on the GTEx portal. Median expression of genes relevant to irinotecan metabolism in RPKM (reads per kilo base per million mapped reads) from 119 normal liver samples was collected from the GTEx portal. Spearman's correlation was used to assess correlation between gene expression.

In vitro functional assays. Sequences of primers used in this study are provided in **Supplementary Table 4**. The *HNF1A* cDNA was PCR amplified from a construct (FR_HNF1A #31104) kindly provided by Gerhart Ryffel (Addgene, Cambridge, MA, USA) (Senkel et al., 2005) that was subcloned into pcDNA3 EcoRI-EcoRV (Invitrogen, Carlsbad, CA, USA). This construct (pHNF1A_L27_N487) carried the minor alleles *rs1169288C* and *rs2264196A*, encoding respectively a leucine at position 27 and an asparagine at position 487. Mutagenesis was performed to obtain the major alleles (pHNF1A_L27_S487 or pHNF1A_I27_S487 carrying *rs2464196G* encoding a serine at position 487) and either the minor or the major allele of *rs1169288A>C* using the Q5 Site-directed mutagenesis kit (New England Biolabs, Whitby, ON, Canada). A portion of the *ABCC1* promoter (-2.1 kb to -0.3 kb) was PCR amplified from genomic DNA and subcloned into pGL3-Basic vector (Promega Corp., Madison, WI, USA) between SacI and XhoI restriction sites upstream of the firefly luciferase gene. Sanger sequencing was used to verify the sequence of all constructs. HEK293 and HepG2 cells were obtained from ATCC

(Rockville, MD, USA). Cell lines were grown as recommended and plated into 24-wells plates at a density of 1.0×10^5 cells/well prior to transfection 24h later. The pcDNA3_HNF1A constructs (200 ng) were co-transfected with the promoter/pGL3-Basic firefly luciferase reporter vector (200 ng) and pRL-CMV renilla luciferase vector as an internal control (20 ng, Promega Corp., Madison, WI, USA) using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) and OptiMEM-reduced serum media (Life Technologies, Burlington, ON, Canada) as per the manufacturer's instructions. Cells were lysed 48h after transfection and assessed for luciferase activity using Dual-luciferase reporter assay kit (Promega Corp., Madison, WI, USA). Firefly luciferase activity is expressed as the ratio of firefly/renilla activities relatively to the corresponding control. Differences among conditions were tested by a Student's two-tailed t-test.

RESULTS

Two cohorts comprising 167 mCRC Canadian patients and 250 mCRC Italian patients were studied, for a total of 417 mCRC patients all treated with the FOLFIRI regimen (leucovorin, 5-fluorouracil and irinotecan) (**Table 1**). Thirteen *HNFI*A htSNPs were assessed for their association with overall survival (OS), progression-free survival (PFS), response rate, severe gastrointestinal toxicities and severe neutropenia (listed in **Supplementary Table 1**). None of the tested *HNFI*A markers were associated with severe diarrhea in any of the two cohorts examined (data not shown). A few htSNPs were significantly associated with severe neutropenia in the Canadian cohort but not in the Italian cohort (**Supplementary Table 2**) whereas some htSNPs associated with OS and response rate were significant only in the Italian cohort (**Supplementary Table 3**).

For *HNFI*A intronic variant rs2244608A>G, a trend for an improved PFS in the Canadian population with a hazard ratio (HR) of 0.74 (95% CI = 0.53 – 1.03, $P = 0.076$) was observed. The same effect using the same genetic model was significant in the Italian cohort with a HR of 0.72 (95% CI = 0.53 – 0.98, $P = 0.038$) (**Table 2**). A strong association between rs2244608A>G and improved PFS was also found when both cohorts were combined and analyzed with the same genetic model, adjusted for age and co-treatment, with a HR of 0.72 (95% CI = 0.59 – 0.89, $P = 0.002$). Consistent with a potential significant influence of this genetic variation, in the analysis of its association with the pharmacokinetics parameters of irinotecan in a subset of patients of the prospective Italian cohort, carriers of the *rs2244608G* allele ($n = 24$) presented an enhanced blood exposure to the active metabolite SN-38 (increase of 45% of SN-38 AUC, $P = 0.032$) compared to carriers of homozygous *rs2244608AA* genotype ($n = 25$) whereas no significant changes were noted for CPT-11 AUC and SN-38G AUC (**Fig.1A-C**). Carriers of the *rs2244608G* allele also had a 41% increased biliary index ($P = 0.021$) and a 24% decreased glucuronidation ratio ($P = 0.035$) compared to homozygote carriers of *rs2244608A* allele (**Fig.1D-E**).

The minor allele frequency of rs2244608A>G is 0.32 in the CEU population, 0.34 in the Canadian cohort and 0.33 in the Italian cohort. The htSNP rs2244608A>G is located in the first intron of *HNFI*A and is in strong LD with 24 other SNPs in individuals of European

descent ($r^2 > 0.80$ from 1000 Genomes Phase 3 CEU population) (**Fig.2**), namely with the non-synonymous coding variant rs1169288c.79A>C located in the first exon of *HNF1A* with r^2 of 0.98. The minor allele *rs1169288C* introduces the substitution of the isoleucine 27 to a leucine (p.I27L) in the dimerization domain of HNF1A (**Fig.2**) and therefore could functionally contribute to changes observed in circulating SN-38 levels for carriers of the minor *rs2244608G* allele in almost complete linkage with the minor *rs1169288C* (p.27L) allele. However, the degree of linkage between the positive marker rs2244608A>G and the coding variant rs1169288A>C varies significantly across ethnic groups. The strongest linkage was observed in individuals of European descent ($r^2 = 0.98$), followed by Asians (Japanese $r^2 = 0.77$ and Chinese $r^2 = 0.69$) and Africans ($r^2 = 0.23$) (**Fig.3**). All other linked polymorphisms are intronic or located in the untranscribed region upstream of the *HNF1A* gene, and may influence the expression of *HNF1A*.

The analysis of liver tissues from 119 healthy donors indicated that *HNF1A* as well as *ABCC1* and *ABCC2* are expressed (data not shown). In support of a potential functional significance of the coding *HNF1A* variant rs1169288 and based on the analysis of liver samples from 89 healthy donors with available genotype data, we observed an enhanced hepatic expression of *ABCC1* ($P = 0.009$) and a lower expression of *ABCC2* ($P = 0.048$) and *CYP3A5* ($P = 0.001$) in carriers of the minor *rs1169288C* allele ($n = 38$ individuals) compared to those carrying the rs1169288AA genotype ($n = 51$ individuals) (**Fig.4**). In addition, compared to homozygous carriers of *rs1169288A* allele ($n = 42$ individuals), expression of the irinotecan drug target gene *TOP1* was significantly higher in the small intestine ($P = 0.004$) in carriers of the minor allele *rs1169288C* ($n = 33$ individuals) (**Fig.4**). We also noted that the hepatic expression of *HNF1A* is inversely correlated to *ABCC1* (Spearman $\rho = -0.571$, $P < 0.0001$) and a trend for a positive correlation with *ABCC2* (Spearman $\rho = 0.193$, $P = 0.058$) was observed (**Fig.5A-B**). In line, a negative correlation between *ABCC1* and *ABCC2* expressions was observed (Spearman $\rho = -0.310$, $P = 0.002$, **Fig.5C**). When accounting for the rs1169288 genotype, the negative correlation between expression of *ABCC1* and *HNF1A* remained significant in both genotype groups. The negative correlation between *ABCC1* and *ABCC2* was also preserved but was significant only in homozygous carriers of the reference *rs1169288A* allele ($n = 51$), likely due to the smaller sample size for *rs1169288C* carriers ($n = 38$) (**Fig.5D**). Relationship to other

pathways relevant to irinotecan is provided in **Fig.S1**, but none reached significance. A trend was observed for carriers of *rs1169288C* allele and a decreased *HNF1A* expression ($P = 0.082$).

Based on the prominent association between *HNF1A* and *ABCC1* observed in human liver samples, we then tested whether the Ile27Leu substitution affects the ability of transcription factor *HNF1A* to induce transcription of a reporter gene construct containing the *ABCC1* promoter. *In vitro* luciferase assays revealed a significant impact of the Ile27Leu substitution when tested in the human HEK293 kidney cells and the liver HepG2 cell line with a significant increased in reporter gene activity associated with *HNF1A* p.L27 by 27% ($P = 0.008$) and 15% ($P = 0.041$), respectively, compared to cells expressing the *HNF1A* p.I27 protein (**Fig.6**).

DISCUSSION

We report that the intronic htSNP rs2244608A>G located in the *HNF1A* gene may be predictive of an improved PFS in two independent cohorts of 417 mCRC white patients treated with irinotecan-based regimen. In support of a better response to chemotherapy, an enhanced systemic exposure to the active metabolite of irinotecan, SN-38, was observed in a subset of mCRC patients carrying the minor *rs2244608G* allele associated with prolonged PFS. The rs2244608A>G was not associated with severe toxicities in the Canadian and the Italian cohorts, suggesting that the increase in SN-38 exposure for carriers of *rs2244608G* allele may not induce severe side effects. In individuals of European descent, rs2244608A>G marker is almost in complete linkage ($r^2 = 0.98$) with the non-synonymous coding variant rs1169288A>C (p.I27L) located in exon 1, which could explain the observed phenotype. This is reinforced by an altered hepatic expression of metabolic and transport genes involved in irinotecan pharmacology, in favour of superior blood levels of the anticancer drug coupled with enhanced expression of irinotecan target gene *TOP1* in the intestine, in carriers of the *rs1169288C* allele strongly linked to the *rs2244608G* marker. In functional *in vitro* studies, the *HNF1A* *rs1169288C* p.L27 polymorphism could significantly increase the *ABCC1* promoter activity in human HepG2 hepatocellular carcinoma cell line and HEK293 kidney cells. This is the first study to demonstrate the relationship between variability in the transcription factor *HNF1A* gene and clinical outcome in the context of irinotecan-based regimen.

HNF1A plays an important part in the pathogenesis of various diseases particularly diabetes and cardiovascular disease. Genetic variations in the *HNF1A* gene, and especially the coding variant rs1169288A>C (p.I27L), have been previously linked to type 2 diabetes (T2D) (Holmkvist et al., 2006; Gaulton et al., 2015), gestational diabetes mellitus (Shaat et al., 2006), cholelithiasis (Richter et al., 2012), coronary artery diseases (CAD) (Wang et al., 2015; Zhou et al., 2017), serum lipid levels (Babaya et al., 2003; Zhou et al., 2017) and levels of the biomarker of inflammation C-reactive protein (CRP) (Reiner et al., 2008; Reiner et al., 2009), suggesting a functional impact of the codon 27 amino acid substitution. However, this does not preclude the potential impact of the intronic rs2244608 marker (or other SNPs in high LD), which can influence gene expression, stability or splicing of the

mRNA. The Ile27Leu substitution is located in the N-terminal dimerization domain of HNF1A that mediates its homodimerization and heterodimerization with HNF1B and the dimerization cofactor DcoH (Narayana et al., 2001). This conservative substitution may be anticipated to have little impact on the structure of HNF1A; however, isoleucine 27 is highly conserved across species, supporting a functional significance (Narayana et al., 2001; Setoh et al., 2015).

Of key genes involved in irinotecan pharmacology, the hepatic expression of *CYP3A5* and the transporters *ABCC1* and *ABCC2* was significantly influenced by the rs1169288A>C. The expression of these genes may be regulated by HNF1A, either directly or indirectly. In line, the expression of *ABCC2* is regulated by HNF1A in liver cells (Odom et al., 2004; Qadri et al., 2009). In addition, results of our *in vitro* investigations allowed to further substantiate evidence for a functional impact of variant HNF1A p.I27L (rs1169288A>C) on the expression of *ABCC1* in human liver and kidney cells, with an increased transcriptional induction by HNF1A p.L27 using the *ABCC1* promoter. Consistent with the elevated SN-38 AUC, higher *ABCC1* (MRP1) expression and lower *ABCC2* (MRP2) and *CYP3A5* expressions was observed for *rs224608G* carriers, strongly linked to the *rs1169288C* allele. MRP1 is on the basolateral membrane of hepatocytes (Emi et al., 2013) and mediates efflux of SN-38 into blood (Chen et al., 1999) whereas MRP2 is located on the apical membrane of hepatocytes (Emi et al., 2012) and is involved in the efflux of several endogenous compounds and drugs, including irinotecan and its metabolites, to the bile (Chu et al., 1997; Sugiyama et al., 1998). The superior exposure to SN-38 in carriers of *rs224608G* may thus be explained through an enhanced redirection of SN-38 from hepatocytes to blood circulation and a lower excretion of irinotecan metabolites in the bile. In line with this hypothesis, Wang *et al.* demonstrated that hepatic inhibition of MRP2 expression with hesperidin in rats lead to an increased SN-38 exposure while biliary excretion of SN-38 was reduced (Wang et al., 2016). Furthermore, the rs1169288 polymorphism has been linked to altered irinotecan pharmacokinetics in a study of 85 patients with solid tumors (Rosner et al., 2008). The strong negative correlation between hepatic expression of *ABCC1* and *HNF1A* and the nearly significant positive correlation between *ABCC2* and *HNF1A* observed in the GTEx cohort sustain a functional interaction, yet an inverse regulatory influence of HNF1A on these two transporters. We could thus speculate that a decreased

expression or function of *HNF1A* would lead to increase *ABCC1* expression and potentially decreased *ABCC2* expression. This might contribute to an enhancement of SN-38 systemic exposure and a reduced elimination of the drug, leading to a prolonged PFS, as represented in **Figure 7**. Consistent with this notion, Holmkvist *et al.* demonstrated a decreased transcriptional induction of *SLC2A2* promoter (Solute Carrier Family 2 Member 2 or GLUT2) by HNF1A p.L27 in HeLa cells (Holmkvist *et al.*, 2006). Although there are inconsistencies in the literature, previous reports also support a link between genetic variations in the *ABCC1* and *ABCC2* genes, altered irinotecan pharmacokinetics and clinical outcomes in colon and lung cancer patients (Rosner *et al.*, 2008; Innocenti *et al.*, 2009; Akiyama *et al.*, 2012; Li *et al.*, 2016).

The impact of variations in HNF1A expression and/or function is likely to be multifactorial and may impact multiple pathways relevant to irinotecan pharmacokinetics and pharmacodynamics (Smith *et al.*, 2006). In the liver of healthy individuals, we further found a positive correlation between *HNF1A* and *CYP3A5* expressions, in line with a previous observation by Noll *et al.* (Noll *et al.*, 2016), and an association with reduced expression of *CYP3A5* for the *rs1169288C* allele. As *CYP3A5* oxidizes irinotecan in inactive metabolites, reduction of its expression may increase the proportion of irinotecan available for conversion into SN-38 by carboxylesterases, resulting in higher SN-38 exposure. Lastly, carriers of *rs1169288C* also presented higher expression of the irinotecan drug target gene *TOP1* in the small intestine, which could explain, at least in part, a better response to TOP1 inhibitor treatment and improved PFS (**Fig.7**). Additional functional studies will be required to test these hypotheses.

The strengths of our study include the examination of two independent cohorts of mCRC comprising 417 white patients and the analysis of pharmacokinetics and transcriptomic data to further investigate the potential mechanisms underlying the association. Given its exploratory nature, we did not apply correction for multiple testing. Besides, the study design involved an independent set of patients used to replicate findings observed in the training set, which reduces the likelihood of false-positive results while maximizing the detection of true-positive associations. Whereas this may result in potential false-positive associations, the high stringency of multiple testing may have led to false-negative

associations due to the number of patients included. The linkage disequilibrium between rs2244608A>G and rs1169288A>C was almost complete in individuals of European descent but is lower in African and Asian populations implying that our findings may be most particularly significant in individuals of European descent. While recognizing the limitations of our study, further investigations are warranted to replicate our findings and substantiate the functional impact of the HNF1A p.I27L coding variation on pathways relevant to drug metabolism, transport and action.

In conclusion, we report the identification of rs2244608A>G, strongly linked to the functional coding variant *HNF1A* rs1169288A>C (p.I27L), as a predictive marker of improved PFS in mCRC patients treated with irinotecan-based chemotherapy. Prolonged PFS may likely be explained by a greater exposure to SN-38, possibly due to an altered transcriptional regulation by HNF1A of genes involved in irinotecan pharmacology including *CYP3A5*, *ABCC1* and *ABCC2* transporters and *TOPI*.

ABBREVIATIONS

AUC, area under the curve

CEU population, Utah residents with Northern and Western European ancestry population

FOLFIRI, folinic acid (leucovorin) + 5-fluorouracil (5-FU) + irinotecan

HR, hazard ratio

htSNP, haplotype-tagging single nucleotide polymorphism

LD, linkage disequilibrium

mCRC, metastatic colorectal cancer

OR, odds ratio

OS, overall survival

PFS, progression-free survival

ACKNOWLEDGEMENTS

The authors would like to thank all study participants as well as Dr Michèle Rouleau for critical discussion, Frédéric Fournier from the CHU de Québec for retrieving data from The Genotype-Tissue Expression (GTEx) Project and David Simonyan for support with statistics. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from: the GTEx Portal and dbGaP accession number phs000424.v6.p1 on 03/16/2017.

Conflicts of interest

The authors declare no conflict of interest.

Financial support

This work was supported by the Canadian Institutes of Health Research (CIHR) (C.G.; CIHR MOP-42392) and the Canada Research Chair Program (C.G.) and the ‘Associazione Italiana per la Ricerca sul Cancro’ (AIRC; Special Program Molecular Clinical Oncology, 5x1000, [No. 12214]) (G.T.). A.L. was a recipient of a studentship award from ‘Fonds d’Enseignement et de Recherche (FER)’ from Faculty of Pharmacy at Laval University. E.L. is a recipient of a CIHR clinical-scientist phase II award and Prostate Cancer Canada Rising Star Award (RS2013-55). C.G. holds a Canada Research Chair in Pharmacogenomics (Tier I).

Author contributions

Conceptualization: CG, GT; *Methodology:* All authors; *Investigation:* AL, ED, EC, EL, DJ, FC, AB, MD, LV, CG; *Formal Analysis:* AL, ED, EC, EL, LV, GT, CG; *Writing – Review & Editing:* All authors; *Supervision:* CG, GT; *Funding Acquisition:* CG, GT.

REFERENCES

- Akiyama, Y., Fujita, K., Ishida, H., Sunakawa, Y., Yamashita, K., Kawara, K., et al. (2012). Association of ABCC2 genotype with efficacy of first-line FOLFIRI in Japanese patients with advanced colorectal cancer. *Drug Metab Pharmacokinet* 27(3), 325-335. doi: 10.2133/dmpk.DMPK-11-RG-128.
- Aleksunes, L.M., Yeager, R.L., and Klaassen, C.D. (2009). Application of multivariate statistical procedures to identify transcription factors that correlate with MRP2, 3, and 4 mRNA in adult human livers. *Xenobiotica* 39(7), 514-522. doi: 10.1080/00498250902952514.
- Babaya, N., Ikegami, H., Fujisawa, T., Nojima, K., Itoi-Babaya, M., Inoue, K., et al. (2003). Association of I27L polymorphism of hepatocyte nuclear factor-1 alpha gene with high-density lipoprotein cholesterol level. *J Clin Endocrinol Metab* 88(6), 2548-2551. doi: 10.1210/jc.2002-021891.
- Barbarino, J.M., Haidar, C.E., Klein, T.E., and Altman, R.B. (2014). PharmGKB summary: very important pharmacogene information for UGT1A1. *Pharmacogenet Genomics* 24(3), 177-183. doi: 10.1097/fpc.0000000000000024.
- Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21(2), 263-265. doi: 10.1093/bioinformatics/bth457.
- Beutler, E., Gelbart, T., and Demina, A. (1998). Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? *Proc Natl Acad Sci U S A* 95(14), 8170-8174.
- Cecchin, E., De Mattia, E., and Toffoli, G. (2016). Nuclear receptors and drug metabolism for the personalization of cancer therapy. *Expert Opin Drug Metab Toxicol* 12(3), 291-306. doi: 10.1517/17425255.2016.1141196.
- Cecchin, E., Innocenti, F., D'Andrea, M., Corona, G., De Mattia, E., Bion, P., et al. (2009). Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. *J Clin Oncol* 27(15), 2457-2465. doi: 10.1200/jco.2008.19.0314.
- Chen, S., Laverdiere, I., Tourancheau, A., Jonker, D., Couture, F., Cecchin, E., et al. (2015a). A novel UGT1 marker associated with better tolerance against irinotecan-induced severe neutropenia in metastatic colorectal cancer patients. *Pharmacogenomics J* 15(6), 513-520. doi: 10.1038/tpj.2015.12.
- Chen, S., Villeneuve, L., Jonker, D., Couture, F., Laverdiere, I., Cecchin, E., et al. (2015b). ABCC5 and ABCG1 polymorphisms predict irinotecan-induced severe toxicity in metastatic colorectal cancer patients. *Pharmacogenet Genomics* 25(12), 573-583. doi: 10.1097/fpc.0000000000000168.
- Chen, Z.S., Furukawa, T., Sumizawa, T., Ono, K., Ueda, K., Seto, K., et al. (1999). ATP-Dependent efflux of CPT-11 and SN-38 by the multidrug resistance protein (MRP) and its inhibition by PAK-104P. *Mol Pharmacol* 55(5), 921-928.
- Chu, X.Y., Kato, Y., Niinuma, K., Sudo, K.I., Hakusui, H., and Sugiyama, Y. (1997). Multispecific organic anion transporter is responsible for the biliary excretion of the camptothecin derivative irinotecan and its metabolites in rats. *J Pharmacol Exp Ther* 281(1), 304-314.

- De Mattia, E., Cecchin, E., Roncato, R., and Toffoli, G. (2016). Pregnane X receptor, constitutive androstane receptor and hepatocyte nuclear factors as emerging players in cancer precision medicine. *Pharmacogenomics* 17(14), 1547-1571. doi: 10.2217/pgs-2016-0095.
- Emi, Y., Harada, Y., and Sakaguchi, M. (2013). Involvement of a di-leucine motif in targeting of ABCC1 to the basolateral plasma membrane of polarized epithelial cells. *Biochem Biophys Res Commun* 441(1), 89-95. doi: 10.1016/j.bbrc.2013.10.013.
- Emi, Y., Yasuda, Y., and Sakaguchi, M. (2012). A cis-acting five-amino-acid motif controls targeting of ABCC2 to the apical plasma membrane domain. *J Cell Sci* 125(Pt 13), 3133-3143. doi: 10.1242/jcs.099549.
- Fujita, K., Kubota, Y., Ishida, H., and Sasaki, Y. (2015). Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J Gastroenterol* 21(43), 12234-12248. doi: 10.3748/wjg.v21.i43.12234.
- Gaulton, K.J., Ferreira, T., Lee, Y., Raimondo, A., Magi, R., Reschen, M.E., et al. (2015). Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat Genet* 47(12), 1415-1425. doi: 10.1038/ng.3437.
- Holmkvist, J., Cervin, C., Lyssenko, V., Winckler, W., Anevski, D., Cilio, C., et al. (2006). Common variants in HNF-1 alpha and risk of type 2 diabetes. *Diabetologia* 49(12), 2882-2891. doi: 10.1007/s00125-006-0450-x.
- Hu, D.G., Meech, R., McKinnon, R.A., and Mackenzie, P.I. (2014). Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab Rev* 46(4), 421-458. doi: 10.3109/03602532.2014.973037.
- Innocenti, F., Kroetz, D.L., Schuetz, E., Dolan, M.E., Ramirez, J., Relling, M., et al. (2009). Comprehensive pharmacogenetic analysis of irinotecan neutropenia and pharmacokinetics. *J Clin Oncol* 27(16), 2604-2614. doi: 10.1200/jco.2008.20.6300.
- Levesque, E., Belanger, A.S., Harvey, M., Couture, F., Jonker, D., Innocenti, F., et al. (2013). Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J Pharmacol Exp Ther* 345(1), 95-101. doi: 10.1124/jpet.112.202242.
- Li, M., Seiser, E.L., Baldwin, R.M., Ramirez, J., Ratain, M.J., Innocenti, F., et al. (2016). ABC transporter polymorphisms are associated with irinotecan pharmacokinetics and neutropenia. *Pharmacogenomics J*. doi: 10.1038/tpj.2016.75.
- Liu, R., Liu, H., Gu, H., Teng, X., Nie, Y., Zhou, Z., et al. (2014). A Polymorphism in Hepatocyte Nuclear Factor 1 Alpha, rs7310409, Is Associated with Left Main Coronary Artery Disease. *Biochem Res Int* 2014, 924105. doi: 10.1155/2014/924105.
- Narayana, N., Hua, Q., and Weiss, M.A. (2001). The dimerization domain of HNF-1alpha: structure and plasticity of an intertwined four-helix bundle with application to diabetes mellitus. *J Mol Biol* 310(3), 635-658. doi: 10.1006/jmbi.2001.4780.
- Noll, E.M., Eisen, C., Stenzinger, A., Espinet, E., Muckenhuber, A., Klein, C., et al. (2016). CYP3A5 mediates basal and acquired therapy resistance in different subtypes of pancreatic ductal adenocarcinoma. *Nat Med* 22(3), 278-287. doi: 10.1038/nm.4038.
- Odom, D.T., Dowell, R.D., Jacobsen, E.S., Gordon, W., Danford, T.W., MacIsaac, K.D., et al. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* 39(6), 730-732. doi: 10.1038/ng2047.

- Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., et al. (2004). Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303(5662), 1378-1381. doi: 10.1126/science.1089769.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., et al. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81(3), 559-575. doi: 10.1086/519795.
- Qadri, I., Hu, L.J., Iwahashi, M., Al-Zuabi, S., Quattrochi, L.C., and Simon, F.R. (2009). Interaction of hepatocyte nuclear factors in transcriptional regulation of tissue specific hormonal expression of human multidrug resistance-associated protein 2 (abcc2). *Toxicol Appl Pharmacol* 234(3), 281-292. doi: 10.1016/j.taap.2008.10.005.
- Reiner, A.P., Barber, M.J., Guan, Y., Ridker, P.M., Lange, L.A., Chasman, D.I., et al. (2008). Polymorphisms of the HNF1A gene encoding hepatocyte nuclear factor-1 alpha are associated with C-reactive protein. *Am J Hum Genet* 82(5), 1193-1201. doi: 10.1016/j.ajhg.2008.03.017.
- Reiner, A.P., Gross, M.D., Carlson, C.S., Bielinski, S.J., Lange, L.A., Fornage, M., et al. (2009). Common coding variants of the HNF1A gene are associated with multiple cardiovascular risk phenotypes in community-based samples of younger and older European-American adults: the Coronary Artery Risk Development in Young Adults Study and The Cardiovascular Health Study. *Circ Cardiovasc Genet* 2(3), 244-254. doi: 10.1161/circgenetics.108.839506.
- Richter, D., Harsch, S., Strohmeyer, A., Hirobe-Jahn, S., Schimmel, S., Renner, O., et al. (2012). MALDI-TOF mass spectrometry screening of cholelithiasis risk markers in the gene of HNF1alpha. *J Proteomics* 75(12), 3386-3399. doi: 10.1016/j.jprot.2012.04.036.
- Rosner, G.L., Panetta, J.C., Innocenti, F., and Ratain, M.J. (2008). Pharmacogenetic pathway analysis of irinotecan. *Clin Pharmacol Ther* 84(3), 393-402. doi: 10.1038/clpt.2008.63.
- Senkel, S., Lucas, B., Klein-Hitpass, L., and Ryffel, G.U. (2005). Identification of target genes of the transcription factor HNF1beta and HNF1alpha in a human embryonic kidney cell line. *Biochim Biophys Acta* 1731(3), 179-190. doi: 10.1016/j.bbaexp.2005.10.003.
- Setoh, K., Terao, C., Muro, S., Kawaguchi, T., Tabara, Y., Takahashi, M., et al. (2015). Three missense variants of metabolic syndrome-related genes are associated with alpha-1 antitrypsin levels. *Nat Commun* 6, 7754. doi: 10.1038/ncomms8754.
- Shaat, N., Karlsson, E., Lernmark, A., Ivarsson, S., Lynch, K., Parikh, H., et al. (2006). Common variants in MODY genes increase the risk of gestational diabetes mellitus. *Diabetologia* 49(7), 1545-1551. doi: 10.1007/s00125-006-0258-8.
- Smith, N.F., Figg, W.D., and Sparreboom, A. (2006). Pharmacogenetics of irinotecan metabolism and transport: an update. *Toxicol In Vitro* 20(2), 163-175. doi: 10.1016/j.tiv.2005.06.045.
- Sugiyama, Y., Kato, Y., and Chu, X. (1998). Multiplicity of biliary excretion mechanisms for the camptothecin derivative irinotecan (CPT-11), its metabolite SN-38, and its glucuronide: role of canalicular multispecific organic anion transporter and P-glycoprotein. *Cancer Chemother Pharmacol* 42 Suppl, S44-49. doi: 10.1007/s002800051078.
- Toffoli, G., Cecchin, E., Corona, G., Russo, A., Buonadonna, A., D'Andrea, M., et al. (2006). The role of UGT1A1*28 polymorphism in the pharmacodynamics and

- pharmacokinetics of irinotecan in patients with metastatic colorectal cancer. *J Clin Oncol* 24(19), 3061-3068. doi: 10.1200/jco.2005.05.5400.
- Voight, B.F., Scott, L.J., Steinthorsdottir, V., Morris, A.P., Dina, C., Welch, R.P., et al. (2010). Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 42(7), 579-589. doi: 10.1038/ng.609.
- Wang, X., Rao, Z., Qin, H., Zhang, G., Ma, Y., Jin, Y., et al. (2016). Effect of hesperidin on the pharmacokinetics of CPT-11 and its active metabolite SN-38 by regulating hepatic Mrp2 in rats. *Biopharm Drug Dispos* 37(7), 421-432. doi: 10.1002/bdd.2024.
- Wang, X.B., Han, Y.D., Cui, N.H., Gao, J.J., Yang, J., Huang, Z.L., et al. (2015). Associations of lipid levels susceptibility loci with coronary artery disease in Chinese population. *Lipids Health Dis* 14, 80. doi: 10.1186/s12944-015-0079-1.
- Yamagata, K., Oda, N., Kaisaki, P.J., Menzel, S., Furuta, H., Vaxillaire, M., et al. (1996). Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature* 384(6608), 455-458. doi: 10.1038/384455a0.
- Zhou, Y.J., Yin, R.X., Hong, S.C., Yang, Q., Cao, X.L., and Chen, W.X. (2017). Association of the HNF1A polymorphisms and serum lipid traits, the risk of coronary artery disease and ischemic stroke. *J Gene Med* 19(1-2). doi: 10.1002/jgm.2941.

FIGURE LEGENDS

Figure 1. Association of *HNF1A* rs2244608A>G with pharmacokinetics parameters.

Differences according to genotypes of 49 mCRC patients from the Italian cohort were tested using a two-tailed Mann-Whitney test. The bars indicate mean values for each group. Biliary index was calculated with the formula: [CPT-11 AUC X (SN-38 AUC/SN-38G AUC)] and glucuronidation ratio with the formula: (SN-38G AUC/SN-38 AUC). AUC, area under the curve.

Figure 2. The *HNF1A* marker rs2244608A>G associated with improved PFS is tightly linked ($r^2 = 0.98$) to a number of other SNPs in the *HNF1A* gene including the non-synonymous coding variant rs1169288A>C (p.I27L).

A. Genetic variants in strong linkage disequilibrium (LD) ($r^2 > 0.80$) with the intronic marker rs2244608 significantly associated with outcome in both population of mCRC patients treated with irinotecan are represented on the *HNF1A* gene locus. The positive intronic marker rs2244608 and the non-synonymous rs1169288 variant, represented in bold, are tightly linked ($r^2 = 0.98$) in individuals of European descent. **B.** Representation of the HNF1A protein and coordinates in amino acids of its main protein domains. **C.** Position (relative to the translation start site) and LD values are shown for polymorphisms presented in A. LD values are from the 1000 Genomes Phase 3 CEU population (European descent).

Figure 3. Frequency of the non-synonymous rs1169288A>C variation in diverse ethnic groups and its linkage with the intronic marker rs2244608A>G associated with outcome.

Linkage disequilibrium (r^2) is given for each population. Data were obtained from the 1000 Genomes Phase 3 project. n = number of genotyped alleles. The rs2244608A>G marker is associated with outcome in mCRC patients of European descent.

Figure 4. The *HNF1A* rs1169288C variant allele (p.27L) is associated with altered expression of *CYP3A5*, *ABC* transporter genes and irinotecan drug target *TOP1*.

Gene expression and genotypes were from individuals of the GTEx cohort. *P*-values were calculated by a two-tailed Mann-Whitney test. **A-C.** Normal livers (n = 89 individuals): rs1169288AA, n = 51; rs1169288AA/AC, n = 38; **D.** Normal small intestines (n = 75 individuals): rs1169288AA, n = 42; rs1169288AC/CC, n = 33.

Figure 5. Correlation of the hepatic expression of *HNF1A*, *ABCC1* and *ABCC2* and the influence of the rs1169288 genotype.

A-C. Correlations are shown for 97 individuals of the GTEx project. **D.** Spearman correlation ρ coefficients and *P*-values for genotyped individuals carrying rs1169288AA (homozygous for p.27I; n = 51) and rs1169288AC/CC (heterozygous and homozygous p.27L; n = 38) genotypes.

Figure 6. Functional impact of the *HNF1A* rs1169288 coding variant. Transcriptional induction of the *ABCC1* promoter by HNF1A p.Ile27 and p.Leu27 in liver (**A**) and kidney (**B**) cells. Cells were co-transfected with a reporter vector containing the firefly luciferase gene downstream of the *ABCC1* promoter and an *HNF1A* construct (p.Ile27 or p.Leu27). The bars represent the mean \pm S.D. of three independent experiments performed in triplicate. Statistical analysis used Student's two-tailed t-test.

Figure 7. Proposed model for the underlying mechanism of an improved progression-free survival associated with the *HNF1A* rs1169288C allele.

Table 1. Characteristics of the study cohorts of mCRC Caucasian patients treated with irinotecan-based chemotherapy (FOLFIRI regimens).

Characteristics	Canadian cohort		Italian cohort	
	N = 167	%	N = 250	%
Gender				
Male	110	66	162	65
Female	57	34	88	35
Age (years)				
Mean	61.5	--	60.6	--
Standard Deviation	10.2	--	10.3	--
Range	29 - 86	--	26 - 75	--
FOLFIRI				
Co-treatment	75	44.9	0	0
Bevacizumab	69	92.0	0	0
Other drug	6	8.0	0	0
Stage at diagnosis				
I	4	2.4	5	2.0
II	12	7.2	20	8.0
III	45	26.9	65	26.0
IV	100	59.9	160	64.0
Unknown	6	3.6	0	0
Outcome				
PFS (median in months)	11	--	7	--
OS (median in months)	24	--	15	--
Tumor response rate				
CR + PR	78	50.0	103	43.3
SD + PD	78	50.0	135	56.7
Severe toxicities (grades 3-4)				
Neutropenia	28	16.8	35	14.0
Diarrhea	24	14.4	21	8.4

Stage at diagnosis evaluated by TNM system. CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; PFS, progression-free survival; OS, overall survival.

Table 2. Association between *HNF1A* polymorphisms and progression-free survival (PFS).

htSNPs ^a	Model	Canadian cohort	Italian cohort
		HR ^b (95% CI); p-value	HR ^c (95% CI); p-value
rs2244608A>G	Dominant	0.74 (0.53 – 1.03); 0.076	0.72 (0.53 – 0.98); 0.038
rs1169286T>C	Dominant	0.66 (0.46 - 0.95); 0.024	Non-significant
rs2393791T>C	Dominant	0.69 (0.50 – 0.97); 0.032	Non-significant
rs12427353G>C	Dominant	1.36 (0.97 – 1.90); 0.074	Non-significant
rs2071190 T>A	Additive	Non-significant	1.35 (1.04 – 1.74); 0.022
rs1169302T>G	Recessive	Non-significant	1.54 (1.03 – 2.31); 0.038
rs2259820C>T	Dominant	0.69 (0.50 - 0.96); 0.026	Non-significant
rs1169307C>T	Recessive	1.57 (1.02 - 2.40); 0.040	Non-significant
rs735396T>C	Dominant	Not in HWE	0.72 (0.54 – 0.97); 0.033

^aAdditional genotyped htSNPs were not significantly associated with PFS (rs1882149, rs3999413, rs1169303, rs1169293) or were already tagged ($r^2 \geq 0.98$) by another htSNP presented here. ^bAdjusted for age and co-treatment. ^cAdjusted for gender, age, cancer site, stage at diagnosis, radical surgery and adjuvant chemotherapy. Base changes are given for the + strand. Abbreviations: 95% CI, 95% confidence interval; HR, hazard ratio; htSNP, haplotype-tagging single nucleotide polymorphism; HWE, Hardy-Weinberg equilibrium.

Figure 1.

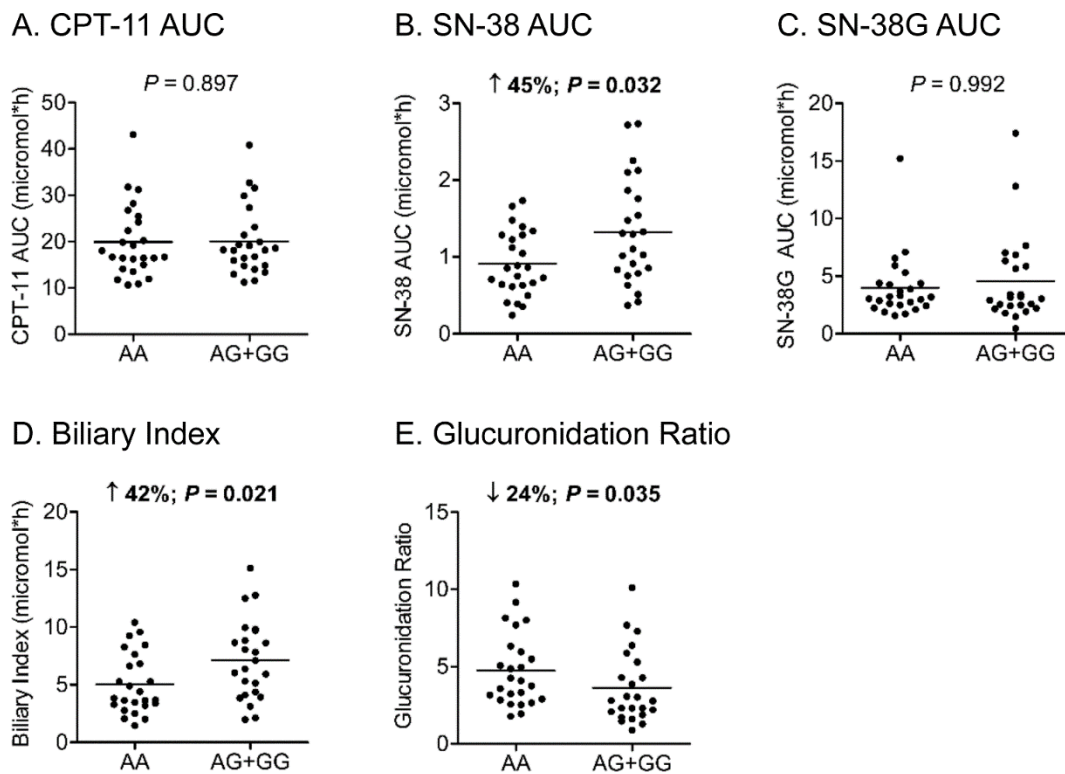
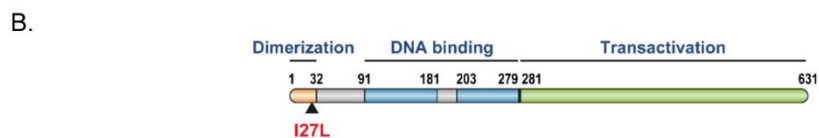
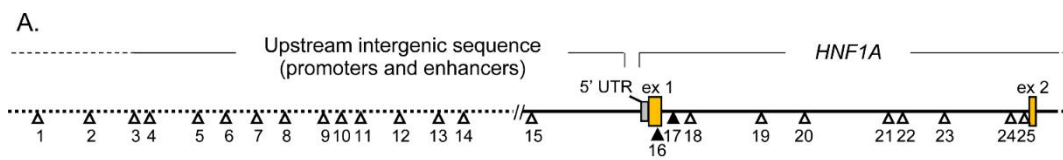


Figure 2.



C.

Δ	SNP	Base change	Position	r^2	Location
1	rs11065358	C>T	-32076	0.85	intergenic
2	rs11065359	A>G	-30449	0.84	intergenic
3	rs1696359	C>T	-28012	0.82	intergenic
4	rs2650000	C>A	-27609	0.84	intergenic
5	rs2178463	G>A	-26850	0.84	intergenic
6	rs2701194	G>A	-26493	0.84	intergenic
7	rs2701175	A>C	-24900	0.84	intergenic
8	rs1732391	T>C	-24230	0.84	intergenic
9	rs6489786	G>A	-18696	0.84	intergenic
10	rs7954039	C>A	-17917	0.84	intergenic
11	rs7954331	T>G	-17914	0.84	intergenic
12	rs71076693	-T	-14923	0.84	intergenic
13	rs2251468	A>C	-11445	0.84	intergenic
14	rs2243616	T>G	-10201	0.84	intergenic
15	rs142632970	A>G	-3226	0.84	intergenic
16	rs1169288	A>C	79	0.98	exon 1
17	rs2244608	A>G	417	--	intron 1
18	rs3830659	-ACTC	965	0.89	intron 1
19	rs1169284	T>C	3355	0.95	intron 1
20	rs1183910	G>A	4236	0.95	intron 1
21	rs11065384	C>T	6714	0.95	intron 1
22	rs11065385	G>A	6815	0.95	intron 1
23	rs2243458	C>T	7919	0.95	intron 1
24	rs1169292	C>T	9907	0.84	intron 1
25	rs1169294	G>A	10023	0.86	intron 1

Figure 3.

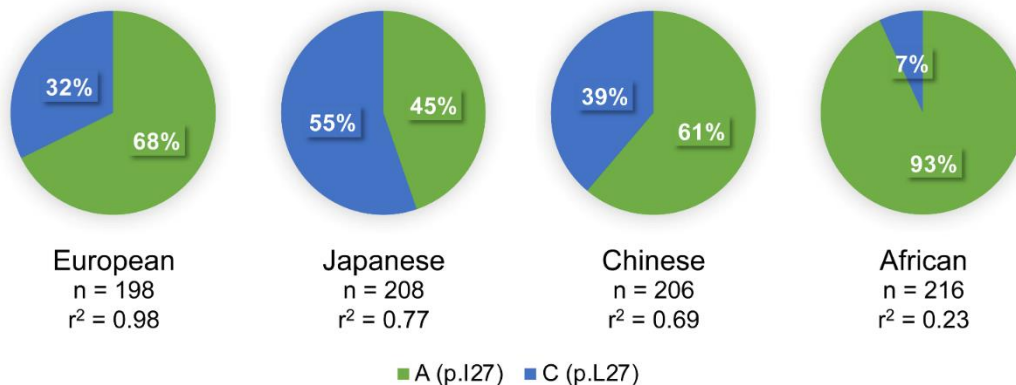


Figure 4.

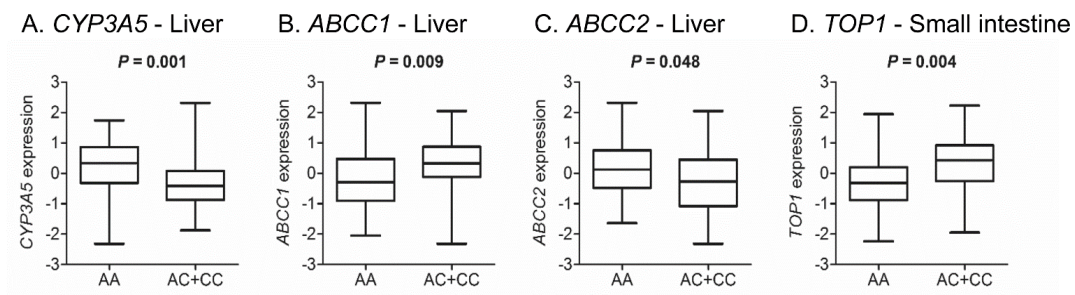


Figure 5.

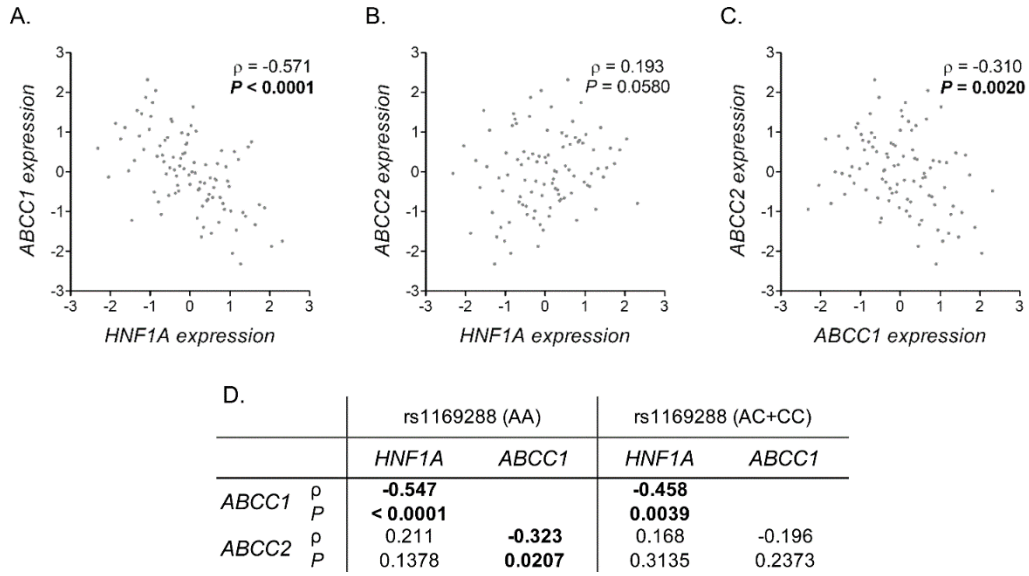


Figure 6.

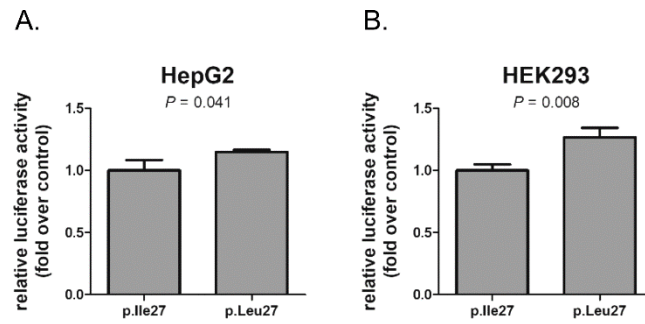
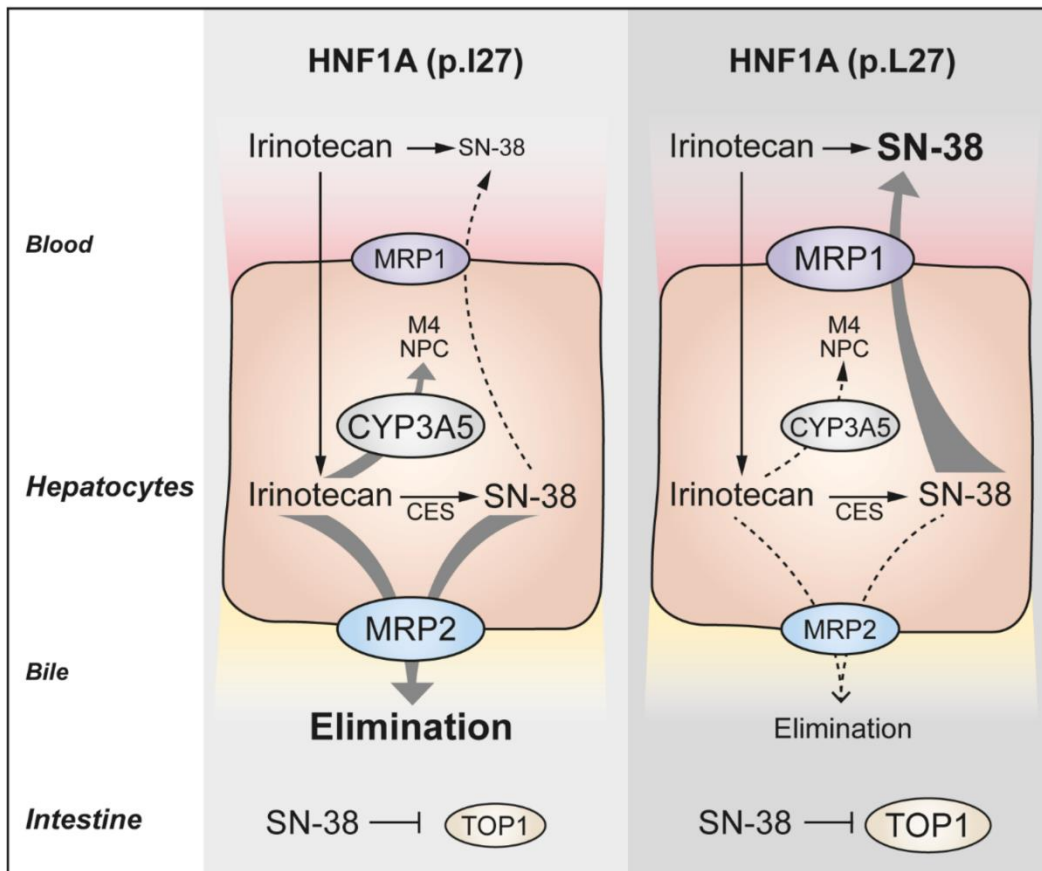


Figure 7.



Improved progression-free survival in irinotecan-treated metastatic colorectal cancer patients carrying the HNF1A coding variant p.I27L

Adrien Labriet^{1,2§}, Elena De Mattia^{3§}, Erika Cecchin³, Éric Lévesque^{4,5}, Derek Jonker⁶, Félix Couture^{4,5}, Angela Buonadonna⁷, Mario D'Andrea⁸, Lyne Villeneuve^{1,2}, Giuseppe Toffoli³ and Chantal Guillemette^{1,2*}

*Corresponding Author: Chantal Guillemette, Ph.D., CHU de Québec Research Center, R4701.5, 2705 Blvd. Laurier, Québec, Canada, G1V 4G2. E-mail: chantal.guillemette@crchudequebec.ulaval.ca

Supplementary Tables

Supplementary Table 1. *HNF1A* haplotype-tagging single nucleotide polymorphisms (htSNPs) in moderate to weak linkage disequilibrium ($r^2 < 0.8$), and their corresponding associated SNPs ($r^2 > 0.8$) as determined with the data of 1000 Genomes Phase 3 for the CEU population (European descent).

htSNPs	Associated SNPs				
rs2244608	rs1169288	rs2243458	rs142632970	rs7954039	rs2178463
	rs1169284	rs3830659	rs2243616	rs6489786	rs2650000
	rs1183910	rs1169294	rs2251468	rs1732391	rs11065359
	rs11065384	rs11065358	rs71076693	rs2701175	rs1696359
	rs11065385	rs1169292	rs7954331	rs2701194	
rs1169286	-				
rs2393791	rs9738226	rs7979478	rs7139079		
	rs2393776	rs7979473	rs7135337		
	rs7970695	rs7310409	rs7953249		
	rs2393775	rs1169289	rs11065365		
	rs546442868	rs2464190			
rs12427353	rs34565150	rs67075643	rs71453526		
	rs148285095	rs12425790	rs112508908		
	rs5801412	rs7131696	rs7965349		
	rs55834942	rs7957197			
	rs55879344	rs61953352			
rs2071190	-				
rs1169302	rs1169299				
rs2259820	rs2464196	rs2257764	rs2264778		
	rs1169301	rs2264750	rs2258287		
	rs1169300	rs1182933	rs1169281		
	rs1169314	rs3213545	rs1169279		
	rs1169315	rs1169291			
rs1169307	rs2257813				
	rs2264779				
rs735396	rs1169309	rs2259816			
	rs1169306	rs2464195	rs2257962		
	rs1169310	rs2259852	rs2254971		
	rs1169311	rs34521980	rs4278584		
	rs1169312	rs2264782			
rs1882149	rs41279096	rs2393777	rs73228105	rs7303401	
	rs56383076	rs201044504	rs7961178	rs7132164	
	rs2393774	rs73228108	rs2251556	rs73226290	
	rs55691424	rs55956549	rs73226296	rs73226268	
	rs73214155	rs2254779	rs7135458		
rs3999413	rs1169304	rs1169280	rs12811390		

	rs3999412	rs2259697	
	rs10849828	rs12819210	
	rs11065393	rs535217845	
	rs2258227	rs11065403	
rs1169303	rs2258043		
	rs2701180		
	rs2245407	rs1971317	rs12313339
	rs113154140	rs73226270	rs113277649
rs1169293	rs959400	rs1696360	rs2708086
	rs73226274	rs570954040	
	rs9919745	rs1617471	

Abbreviations: htSNP, haplotype-tagging SNP; SNP, single nucleotide polymorphism. Associated SNPs are polymorphisms in strong linkage with the htSNP with $r^2 > 0.8$, as determined with the data of 1000Genomes Phase 3 for the CEU population.

Supplementary Table 2. Association between *HNF1A* htSNPs and severe neutropenia in the Canadian cohort. None of the htSNPs were significant in the Italian cohort using the same genetic model.

htSNPs	Model	OR^a (95% CI); P-value
rs1169286T>C	Dominant	0.45 (0.19 - 1.07); 0.070
rs2393791T>C	Dominant	0.45 (0.20 - 1.04); 0.061
rs12427353G>C	Recessive	8.58 (1.30 - 56.57); 0.025
rs2071190T>A	Recessive	2.70 (0.85 - 8.51); 0.091
rs1169302T>G	Dominant	0.48 (0.21 - 1.11); 0.085
rs2259820C>T	Dominant	0.32 (0.13 - 0.78); 0.012
rs1169303A>C	Dominant	0.35 (0.15 - 0.84); 0.018
rs1169307C>T	Recessive	3.38 (1.31 - 8.73); 0.012

^aAdjusted for age and co-treatment. Base changes are given for the + strand. Abbreviations: 95% CI, 95% confidence interval; OR, odds ratio; htSNP, haplotype-tagging single nucleotide polymorphism.

Supplementary Table 3. Association between *HNF1A* htSNPs and overall survival and response rate in the Italian cohort. None of the htSNPs were significant in the Canadian cohort for these endpoints using the same genetic model.

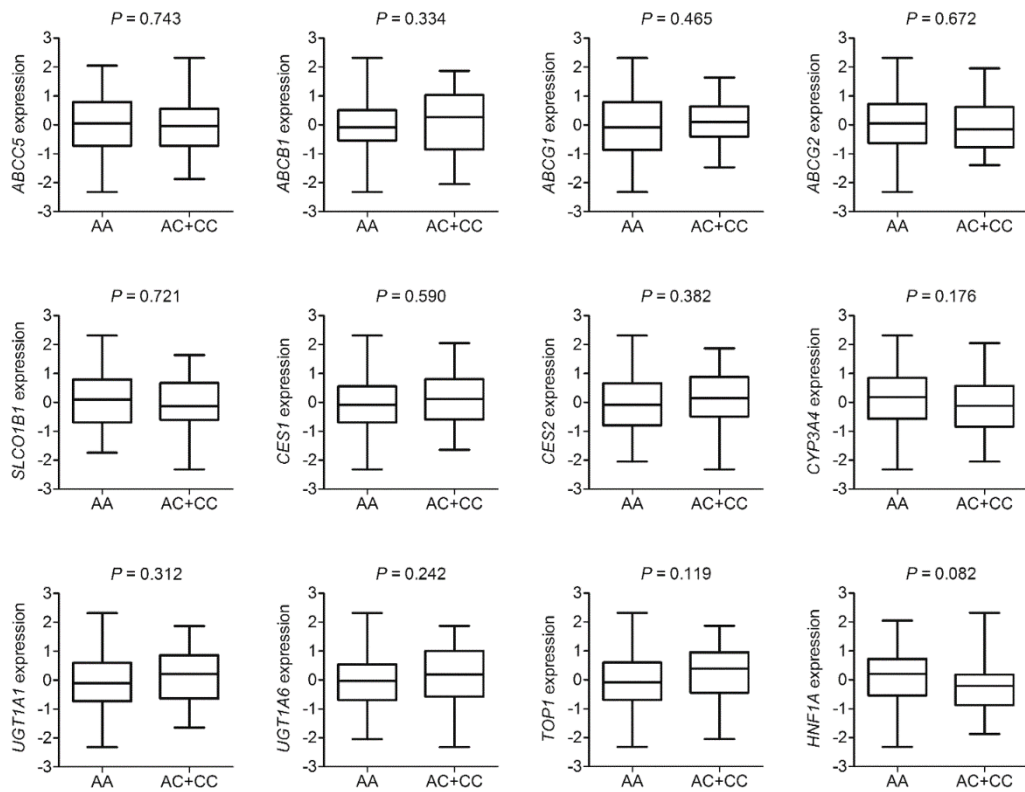
Overall survival	Model	HR^a (95% CI); P-value
rs2393791T>C	Dominant	0.61 (0.42 – 0.88); 0.008
rs1169302T>G	Recessive	1.78 (1.13 – 2.79); 0.013
rs1169307C>T	Recessive	1.69 (1.07 – 2.67); 0.026
Response rate	Model	OR^a (95% CI); P-value
rs12427353G>C	Dominant	0.20 (0.05 – 0.78); 0.020

^aAdjusted for gender, age, cancer site, stage at diagnosis, radical surgery and adjuvant chemotherapy. Base changes are given for the + strand. Abbreviations: 95% CI, 95% confidence interval; HR, hazard ratio; htSNP, haplotype-tagging single nucleotide polymorphism; OR, odds ratio.

Supplementary Table 4. Sequences of primers used for *in vitro* studies.

Primers	Sequence (5' → 3')
<i>HNF1A</i> -EcoRI-forward (PCR and cloning of <i>HNF1A</i> ORF)	CTAGCTGGAATTCCCATGGAGCAAAGCTCATT
<i>HNF1A</i> -EcoRV-reverse (PCR and cloning of <i>HNF1A</i> ORF)	CTAGCTGGATATCCCCAAGCAGGCAGTACAGG
<i>rs1169288A</i> -forward (site directed mutagenesis)	AGAGGCACTGaTCCAGGCACT
<i>rs1169288A</i> -reverse (site directed mutagenesis)	TTGCTCAGCCCTGACTCC
<i>rs2464196G</i> -forward (site directed mutagenesis)	GTGACCCAGAgCCCCTTCATG
<i>rs2464196G</i> -reverse (site directed mutagenesis)	ATGGCTCTGCACAGGTGG
prom <i>ABCC1</i> -Sac-forward (PCR and cloning of <i>ABCC1</i> promoter)	CTAGCTGGAGCTCGCACTGAGATTACATGCGTGA
prom <i>ABCC1</i> -Xho1-reverse (PCR and cloning of <i>ABCC1</i> promoter)	CTAGCTGCTCGAGCCCAGATCCTCCAAGGCTTA

Supplementary Figure 1.



Supplementary Figure 1. Association between rs1169288A>C genotypes and hepatic expression of genes relevant to irinotecan pharmacology.

P -values were calculated by a two-tailed Mann-Whitney test. Normal livers: rs1169288AA, $n = 51$; rs1169288AA/AC, $n = 38$.

Chapitre 4 : Post-transcriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing

4.1 Résumé

L'enzyme de détoxification UDP-glucuronosyltransférase UGT2B10 est spécialisée dans la glucuronidation des amines de nombreux médicaments et xénobiotiques. Ses substrats préférentiels sont les amines tertiaires aliphatiques et amines hétérocycliques, comme les composés carcinogéniques contenus dans le tabac, de nombreux antidépresseurs et des antipsychotiques. Nous avons émis l'hypothèse que l'épissage alternatif constitue un mécanisme de régulation de la voie d'UGT2B10. Son transcriptome a été étudié dans des tissus normaux et tumoraux humains et le plus haut niveau d'expression a été retrouvé au niveau du foie, dans lequel 10 transcrits alternatifs ont été détectés et représentent 50% du transcriptome *UGT2B10*, tant au niveau du tissu normal (n = 50) que dans les tumeurs de carcinome hépatocellulaire (n = 44). Un sous-ensemble abondant de transcrits correspond à des messagers possédant une nouvelle séquence exonique. Deux de ces transcrits présentent une nouvelle séquence C-terminale de 10 et 65 acides aminés, respectivement. Leur expression hépatique est très variable et corrèle avec les niveaux du transcrit canonique et elle est 3,5 fois plus élevée dans les tumeurs. La détection de peptides spécifiques par spectrométrie de masse soutient leur traduction en protéine. À l'aide de modèles cellulaires, nous avons démontré une colocalisation avec l'enzyme et une modulation de la conjugaison de l'amitriptyline et de la lévométoprolol (40% – 70%, $P < 0.01$) d'une façon spécifique au contexte cellulaire. Contrairement à l'enzyme UGT2B10, les protéines alternatives sont sensibles à la dégradation par le protéasome et sont stabilisées par son inhibition. De plus, un remodelage de l'épissage d'*UGT2B10* induit par un traitement pharmacologique a été observé dans le modèle cellulaire hépatique HepaRG, en faveur de l'expression des variants et au détriment de celle du transcrit canonique. Nos découvertes soutiennent une contribution majeure de l'épissage alternatif sur la régulation de l'expression hépatique d'UGT2B10, avec un impact sur l'activité enzymatique.

4.2 Abstract

The detoxification enzyme UDP-glucuronosyltransferase UGT2B10 is specialized in the N-linked glucuronidation of many drugs and xenobiotics. Preferred substrates possess tertiary aliphatic amines and heterocyclic amines such as tobacco carcinogens and several antidepressants and anti-psychotics. We hypothesized that alternative splicing (AS) constitutes a mean to regulate steady state levels of UGT2B10 and enzyme activity. We established the transcriptome of *UGT2B10* in normal and tumoral tissues of multiple individuals. Highest expression was in the liver, where ten AS transcripts represented 50% of the *UGT2B10* transcriptome in 50 normal livers and 44 hepatocellular carcinomas. One abundant class of transcripts involves a novel exonic sequence and leads to two alternative (alt.) variants with novel in-frame C-termini of 10 or 65 amino acids. Their hepatic expression was highly variable among individuals, correlated with canonical transcript levels, and was 3.5 fold higher in tumors. Evidence for their translation in liver tissues was acquired by mass spectrometry. In cell models, they co-localized with the enzyme and influenced the conjugation of amitriptyline and levomedetomidine by repressing or activating the enzyme (40-70%; $P < 0.01$), in a cell context-specific manner. A high turnover rate for the alt. proteins, regulated by the proteasome, was observed in contrast to the more stable UGT2B10 enzyme. Moreover, a drug-induced remodelling of *UGT2B10* splicing was demonstrated in the HepaRG hepatic cell model, which favored alt. variants expression over the canonical transcript. Our findings support a significant contribution of AS in the regulation of UGT2B10 expression in the liver with an impact on enzyme activity.

Post-transcriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing

Adrien Labriet, Eric P. Allain, Michèle Rouleau, Yannick Audet-Delage, Lyne Villeneuve and Chantal Guillemette

Pharmacogenomics Laboratory, Centre Hospitalier Universitaire (CHU) de Québec Research Center – Université Laval, Québec, QC, Canada and Faculty of Pharmacy, Université Laval, Québec, Canada

Canada Research Chair in Pharmacogenomics (CG)

Short title: Regulation of UGT2B10 by alternative splicing

Corresponding Author:

Chantal Guillemette, Ph.D.

Pharmacogenomics Laboratory

CHU de Québec Research Center – Université Laval

2705 Boul. Laurier, R4701.5

Québec, Canada, G1V 4G2

Tel. (418) 654-2296

E-mail: Chantal.Guillemette@crchudequebec.ulaval.ca

Abstract: 248 words

Introduction: 722

Material and Methods: 1811

Results: 1566

Discussion: 1686

Body text: 5785

References: 60

NUMBER OF:

Tables: 0

Figures: 7

Supplemental Tables: 1

Supplemental Figures: 6

Non-standard abbreviations:

AS: Alternative splicing

ER: Endoplasmic reticulum

MS-MRM: Mass spectrometry – multiple reaction monitoring

RT-PCR: Reverse transcription - PCR

RNA-seq: RNA sequencing

TPM: transcripts per million

UDP-GlcA: UDP-glucuronic acid

UGT: UDP-glucuronosyltransferases

Abstract

The detoxification enzyme UDP-glucuronosyltransferase UGT2B10 is specialized in the N-linked glucuronidation of many drugs and xenobiotics. Preferred substrates possess tertiary aliphatic amines and heterocyclic amines such as tobacco carcinogens and several antidepressants and anti-psychotics. We hypothesized that alternative splicing (AS) constitutes a mean to regulate steady state levels of UGT2B10 and enzyme activity. We established the transcriptome of *UGT2B10* in normal and tumoral tissues of multiple individuals. Highest expression was in the liver, where ten AS transcripts represented 50% of the *UGT2B10* transcriptome in 50 normal livers and 44 hepatocellular carcinomas. One abundant class of transcripts involves a novel exonic sequence and leads to two alternative (alt.) variants with novel in-frame C-termini of 10 or 65 amino acids. Their hepatic expression was highly variable among individuals, correlated with canonical transcript levels, and was 3.5 fold higher in tumors. Evidence for their translation in liver tissues was acquired by mass spectrometry. In cell models, they co-localized with the enzyme and influenced the conjugation of amitriptyline and levomedetomidine by repressing or activating the enzyme (40-70%; $P < 0.01$), in a cell context-specific manner. A high turnover rate for the alt. proteins, regulated by the proteasome, was observed in contrast to the more stable UGT2B10 enzyme. Moreover, a drug-induced remodelling of *UGT2B10* splicing was demonstrated in the HepaRG hepatic cell model, which favored alt. variants expression over the canonical transcript. Our findings support a significant contribution of AS in the regulation of UGT2B10 expression in the liver with an impact on enzyme activity.

Introduction

N-linked glucuronidation is an important inactivation route for amine-containing drugs and xenobiotics (Kaivosaaari et al., 2011; Kato et al., 2013). Two of the 19 UDP-glucuronosyltransferases (UGTs), UGT2B10 and UGT1A4, are the main drivers of N-linked glucuronidation (Chen et al., 2008a; Kerdpin et al., 2009; Kato et al., 2013). Long described as an orphan UGT, the discovery that UGT2B10 is crucial for the detoxification of tobacco carcinogens has raised much attention and rationalized the importance to characterize this unique UGT (Chen et al., 2007; Kaivosaaari et al., 2007; Chen et al., 2008b; Berg et al., 2010; Murphy et al., 2014). UGT2B10 is one of the main liver UGT enzymes, based on mRNA (quantitative PCR and deep RNA-sequencing data) and protein levels (quantitative mass spectrometry-based proteomics data) (Court et al., 2012; Fallon et al., 2013; Margaillan et al., 2015; Tourancheau et al., 2017). Expression has also been reported in the breast, testis, gallbladder, tongue and tonsils although at much lower levels than hepatic expression (Haakensen et al., 2010; Jones and Lazarus, 2014).

UGT2B10 displays a preference for tertiary aliphatic amines and heterocyclic amines. These structures are found in several clinically used drugs such as antihistamines, anti-psychotics and anti-depressants including several of the tricyclic class such as imipramine and amitriptyline (Kaivosaaari et al., 2008; Kaivosaaari et al., 2011; Kato et al., 2013; Kazmi et al., 2015; Pattanawongsa et al., 2016) whereas endogenous substrates have yet to be identified. Although UGT1A4 substrate preference significantly overlaps with that of UGT2B10, the latter presents a greater affinity and clearance for many tertiary cyclic amines at therapeutic concentration (Kato et al., 2013). One structural determinant of the specificity towards amine substrates may be the residues Pro40 of UGT1A4 and Leu34 of UGT2B10 located in their substrate binding domain, a position that is otherwise a strictly conserved histidine residue (His40, coordinates of UGT1A1) in all human UGTs (Kerdpin et al., 2009).

Genetic studies in smokers have established a direct link between single nucleotide polymorphisms (SNPs) at the *UGT2B10* gene locus (4q13.2), UGT2B10 activity and metabolism of nicotine and cotinine (Chen et al., 2007; Chen et al., 2012; Murphy et al., 2014; Patel et al., 2015; Ware et al., 2016; Murphy, 2017). Two relatively frequent

polymorphisms are associated with significant decreased nicotine and cotinine glucuronidation in humans. One variant creates the missense Asp67Tyr coding variant (rs61750900) that abolishes UGT2B10 glucuronidation activity (Chen et al., 2007), while the other variant (rs116294140/rs2942857) alters the splice acceptor site between intron 2 and exon 3, and is thought to create an unstable mRNA (Murphy et al., 2014; Fowler et al., 2015). Consistently, individuals homozygous for either of these SNPs have very low to undetected nicotine and cotinine glucuronides in urine, further indicating that UGT2B10 is the main UGT responsible for glucuronidation of nicotine and cotinine (Chen et al., 2007; Murphy et al., 2014). It also supports the main role of this detoxification pathway in tobacco carcinogenesis. These genetic variations display an important ethnic bias. Whereas the Asp67Tyr variation is most frequent in Caucasians (nearly 10%) (Chen et al., 2007; Murphy et al., 2014), the splice variant rs2942857 prevails in approximately 35% of African Americans and in up to 50% of subjects of African origins (NCBI; Murphy et al., 2014). Of interest, the splice variant was also reported to affect the metabolism of the antipsychotic preclinical drug RO5263397 whose main clearance route is N-glucuronidation (Fowler et al., 2015). It is thus likely that these SNPs also affect the glucuronidation of prescribed drugs conjugated by UGT2B10 such as amitriptyline (Kaivosari et al., 2008; Kato et al., 2013).

Recent studies by our group revealed that alternative splicing (AS) largely expands the UGT transcriptome (Tourancheau et al., 2016; Tourancheau et al., 2017). In turn, alternative (alt.) isoforms modulate the activity of UGT enzymes, suggesting that AS programs may contribute to interindividual variability in xenobiotics metabolism and cancer susceptibility (Bellemare et al., 2010b; Menard et al., 2013; Rouleau et al., 2014; Rouleau et al., 2016). Given the clinical importance of UGT2B10, by virtue of its detoxification functions towards tobacco by-products, anti-depressors and anti-psychotics, we hypothesized that AS represents a mean to regulate steady-state levels of UGT2B10 mRNA and enzyme activity. The goal of this study was to investigate in more details the expression of UGT2B10 and evaluate the influence of alternative isoforms on the metabolic functions of UGT2B10.

Materials and Methods

Analysis of UGT2B10 mRNA expression

To carefully analyze *UGT2B10* expression and AS patterns, raw RNA sequencing (RNA-seq) data were downloaded from public databases and realigned to the recently established UGT transcriptome (Tourancheau et al., 2016; GTEx (<http://www.gtexportal.org/home/>) and TCGA (<https://gdc.cancer.gov/>). RNA-seq datasets were obtained from dbGaP at <http://www.ncbi.nlm.nih.gov/gap> through dbGaP accession number phs000424.v6.p1; project ID 13346. The GTEx normal liver data (downloaded on June 7, 2017) were from 50 healthy individuals [29 males; 18 females; 3 unknown; median age: 55 years old (range 21-69 years old), mostly Caucasians]. The TCGA cancer liver data (downloaded on August 22, 2017), all hepatocellular carcinoma, were from 44 individuals [29 males, 15 females; median age: 62 years old (range 18-82 years old), 16 Caucasians, 22 Asians, and 5 African-Americans]. GTEx RNA-seq data of bladder (n=6), breast (n=51), colon (n=49), lung (n=48), kidney (n=35) prostate (n=50), and skin (n=53) were similarly obtained. HepaRG RNA-seq data (GSE 71446, accessed March 9, 2017; (Li et al., 2015)) were downloaded from the NCBI Gene expression omnibus database. For each individual RNA-seq data, quality of FASTQ files was assessed with FastQC before quality trimming with Trimmomatic v0.36 (Bolger et al., 2014). UGT transcript quantification was done with trimmed reads from each sample and a custom UGT transcript annotation (Tourancheau et al., 2016) using Kallisto v0.43 (Bray et al., 2016). Data (counts) was then upper-quantile normalized and further adjusted using housekeeping genes as described (de Jonge et al., 2007) using the EDASeq and RUVSeq packages for R version 3.2.2. Differential expression of UGT isoforms was assessed using the edgeR package for R. Normalized counts were then converted to counts per million (CPM) or transcripts per million (TPM) using transcript length. Reverse Transcription-PCR (RT-PCR) analysis of *UGT2B10* transcripts was as described (Tourancheau et al., 2016), using primer sequences provided in Table S1. The Basic Local Alignment Search Tool (BLAST) of the National Center for Biotechnology Information (NCBI, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) served to search sequence similarity between the novel UGT2B10 sequences and other genes in humans and other species. The Protein BLAST (blastp) suite was used to search "Non-

redundant protein sequences" and "Reference proteins" whereas Translated nucleotide (tblastn) suite was used to search the "Nucleotide collection" and "Reference RNA sequences" with the unique amino acid sequences of alt. proteins.

Expression vectors and human cell models (HEK293 and HepG2)

To study alt. transcripts and proteins, expression vectors were produced from the *UGT2B10_v1* pcDNA6 construct (Beaulieu et al., 1998) using the Q5 Site-Directed Mutagenesis kit (New England Biolabs Ltd., Whitby, ON, Canada). Sequences of purified primers used for mutagenesis are provided in Supplemental Table 1. *UGT2B10_v1* coding sequence was also cloned in the pcDNA3.1 vector to produce co-expression cell models (below). All constructs were verified by Sanger sequencing. For the expression of alt. proteins tagged with V5-his (for immunoprecipitations studies), the stop codon of each coding sequence cloned in the pcDNA6 vector was removed using the Q5 Site-Directed Mutagenesis kit to permit in-frame V5-his expression.

HEK293 and HepG2 cell lines were obtained from the American Type Culture Collection (Manassas, VA, USA) and grown as described (Levesque et al., 1997; Menard et al., 2013). HEK293 cells (UGT negative; 2×10^6 cells in a 10 cm-plate) were transfected with $1 \mu\text{g}$ of each construct using Effectene (Qiagen, Toronto, ON, Canada), and HepG2 cells (UGT positive; 1×10^7 cells) were transfected with $20 \mu\text{g}$ of each construct by electroporation with the Neon Transfection System (Invitrogen, ThermoFisher Scientific, Ottawa, ON, Canada) as per manufacturers' instructions. HEK293 cells stably expressing the *UGT2B10* cDNAs were established by supplementing cell culture media with blasticidin (Wisent, St-Bruno, QC, Canada; $10 \mu\text{g/ml}$). Clones were selected based on *UGT2B10* expression detected by immunoblotting with antibodies specified below. HEK293 cells co-expressing *UGT2B10_v1* (encoding the canonical enzyme) and alt. *UGT2B10* (with novel sequences in C-termini) were established by subsequent transfection of HEK293 clones expressing alt. sequences with the *UGT2B10_v1*-pcDNA3.1 construct and selection with G418 (Invitrogen, 1 mg/ml). HepG2 cells, which express *UGT2B10_v1* endogenously, were transfected with the constructs expressing alt. variants and clone selection was with blasticidin. Control HEK293 and HepG2 cells were produced by transfection with the parental vector pcDNA6 and selection as above.

Analysis of protein expression

Antibodies. The rabbit polyclonal anti-UGT2B10 #1845 produced in-house against GST-UGT2B11 (aa 60-140) by Dr Alain Bélanger's group (Chouinard et al., 2006), was used for the immunodetection of UGT2B10 in HepG2 (1:5000) and HEK293 (1:10 000) cell models. Our analysis revealed that this antibody detects UGT2B10, UGT2B11 and less efficiently UGT2B28 (Supplemental Figure 1). The monoclonal anti-UGT2B10 antibody (Abcam ab57685) was used for immunoprecipitation in human liver. Cell compartment-specific antibodies used for immunofluorescence (IF) were anti-58K Golgi protein (1:100; ab27043, Abcam Inc., Toronto, ON, Canada), anti-protein disulfide isomerase (PDI: 1:100; ab2792, Abcam) for the ER, and anti-lamin (1:200; sc-376248, Santa Cruz Biotechnology, Dallas, TX, USA). DNA was stained with DRAQ5 (1:2000; ThermoFisher Scientific). Anti-calnexin was from Enzo Life Sciences (Farmingdale, NY, USA; ADI-SPA-860; 1:5000).

Mass spectrometry – multiple reaction monitoring (MS-MRM). Detection of peptides unique to alt. UGT2B10 was as described (Rouleau et al., 2016) with minor modifications. Briefly, human liver S9 fraction (8 mg protein; Xenotech LLC, Lenexa, KS, USA) was lysed for 45 min on ice in a total volume of 4 ml Lysis Buffer containing 0.05 M Tris-HCl pH 7.4, 0.15 M NaCl, 1% (w/v) Igepal CA-630 (Sigma-Aldrich), 1 mM dithiothreitol, and Complete protease inhibitor cocktail (Roche, Laval, QC, Canada). Lysates were centrifuged for 15 min at 13,000 g, and UGT2B10 was immunoprecipitated with 10 µg of the monoclonal anti-UGT2B10 (Abcam ab57685) for 1 h at 4°C on an orbital shaker. Protein complexes were captured by an overnight incubation at 4°C with Protein G-coated magnetic beads (200 µl Dynabeads, ThermoFisher Scientific). Beads were washed in lysis buffer and with 50 mM ammonium bicarbonate and stored at –20°C until analysis. Tryptic digests of UGT2B10 were prepared and analyzed by MS-coupled multiple reaction monitoring on a 6500 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer (Sciex, Concord, ON, Canada) as described (Rouleau et al., 2016). Briefly, MS analyses were conducted with an ionspray voltage of 2500 V in positive ion mode. Peptides were desalted on a 200 µm × 6 mm chip trap ChromXP C18 column, 3 µm (Eksigent, Sciex), at 2 µl/min solvent A (0.1% formic acid). Peptides were then eluted at a flow rate of 1 µl/min

with a 30-min linear gradient from 5 to 40% solvent B (acetonitrile with 0.1% formic acid) and a 10-min linear gradient from 40 to 95% solvent B. MRM-MS analyses were performed using the four most intense transitions for each of the target peptides for the light and heavy forms. The UGT2B10 signature peptides were detected in tryptic digests of the immunoprecipitated UGT2B10 samples, and peptide identity was confirmed by co-injection of isotopically labeled [$^{13}\text{C}_6,^{15}\text{N}_2$]Lys and [$^{13}\text{C}_6,^{15}\text{N}_4$]Arg synthetic peptides (Pierce Protein Biotechnology, ThermoFisher Scientific).

Glucuronidation assays – For enzymatic assays in intact cells (*in situ* assays), two cell models were employed (HEK293 and HepG2 cells). Cells were seeded in 24-well plates (HEK293: 8×10^4 cells/well, HepG2: 2.25×10^5 cells/well). Assays were initiated 72 h after seeding by replacing the culture medium with fresh medium (1 ml/well) containing a UGT2B10 substrate (amitriptyline, 7.5 μM and 150 μM ; levomedetomidine, 7.5 μM and 75 μM). All substrates were obtained from Sigma-Aldrich. Cells were incubated for 4 h, media were then collected and stored at -20°C until glucuronide (G) quantification by high-performance liquid chromatography tandem mass spectrometry. Assays were replicated in at least two independent experiments in triplicates. Separation of amitriptyline and levomedetomidine was performed onto an ACE Phenyl column 3 μM packing material, 100×4.6 mm (Canadian Life Science, ON, Canada). Isocratic condition with 70% methanol/30% water/3 mM ammonium formate with a flow rate of 0.9 ml/min was used to elute amitriptyline-G. A linear gradient was used to elute levomedetomidine-G, with 5% methanol/95% water/1 mM ammonium formate as initial conditions followed by a 10 min-linear gradient to 90% methanol/10% water/ 1 mM ammonium formate. The glucuronides were quantified by tandem mass spectrometry (API 6500; Biosystems-Sciex, Concord, ON, Canada). The following mass ion transitions (m/z) were used: 377.1 \rightarrow 201.1 for levomedetomidine-G and 454.2 \rightarrow 191.1 for amitriptyline-G. Glucuronidation activity (area/hour/mg protein/UGT level) was normalized for the expression of the UGT2B10 enzyme in each cell model, determined by densitometry scanning of band intensity on immunoblots with antibody #1845. Cycloheximide glucuronidation assays were conducted using microsomes isolated from human liver (Xenotech LLC), HepG2-pcDNA6 described above, commercial supersomes expressing UGT1A and UGT2B isoenzymes (Corning,

Woburn, MA, USA), and microsomes of UGT2B11-expressing HEK293 prepared in-house as described (Lepine et al., 2004). Glucuronidation assays were conducted at 37°C using 20 µg membrane proteins and a final concentration of 200 µM cycloheximide (Sigma-Aldrich). Cycloheximide-G quantification was conducted as specified above for the UGT2B10 substrates, using a linear gradient with 10% methanol/90% water/1 mM ammonium formate as the initial conditions followed by a linear gradient to 85% methanol/15% water/ 1 mM ammonium formate in 5 min.

Immunofluorescence – Subcellular distribution of UGT2B10 enzyme and alternative isoforms was carried out in HEK293 cells stably expressing either protein and detected with the anti-UGT2B10 antibody #1845 (1:1000) as described (Rouleau et al., 2016).

Immunoprecipitation - Stable HEK293 *UGT2B10_v1* cells (2×10^6 cells) were seeded in 10 cm-plates and transiently transfected with alt. UGT2B10 V5-tagged constructs (3 µg) using Effectene as per manufacturer's instructions (Qiagen). 36 h post-transfection, cells were processed for cross-linking and immunoprecipitation with the polyclonal anti-V5 (1:600; NB600-380; Novus Biologicals) as described (Rouleau et al., 2016).

Protein and mRNA stability in cell models - HEK293 cell models were seeded in 6-well plates (4×10^5 cells/well) and grown for 48 h. HepG2 cell models were seeded in 6-well plates (1.5×10^6 cells/well) and grown for 24 h. Cells were rinsed and incubated for 16h with media containing 1 µM MG132 (Calbiochem, EMD Millipore, Etobicoke, ON, Canada) or vehicle (DMSO). Cells were harvested and protein extracts were prepared in Lysis Buffer. Cells were lysed for 30 min on a rotation unit, homogenized by pipetting up and down through fine needles (18G and 20G, 10 times each) on ice and cleared by centrifugation for 15 min at 13,000 g prior to analysis by immunoblotting using anti-UGT2B10 #1845. Duplicate cell samples were harvested for RNA extraction and RT-PCR analysis of UGT2B10 variants as described (Rouleau et al., 2016). Assays were done twice. UGT2B10 proteins half-lives were determined by treatment with cycloheximide (20 µg/ml) for 0-16 h as described (Turgeon et al., 2003). Cells were washed in PBS, collected and lysed by scraping in Lysis Buffer, prior to analysis by immunoblotting.

Assessment of the glycosylation status of UGT2B10 proteins expressed in HEK293 cells - Microsomes (20 µg proteins) prepared from HEK293 stably expressing each UGT2B10 isoform were treated with Endo H and O-glycosidase obtained from New England Biolabs Inc. (Ipswich, MA, USA) as described (Girard-Bock et al., 2016). Assays were performed in two independent experiments.

Results

UGT2B10 gene expression predominates in human liver and is regulated by alternative splicing

For a quantitative assessment of UGT2B10 expression, we performed a realignment of public GTEX and TCGA RNA-seq data from human tissue samples to the fully annotated UGT variant sequence database. This database was created based on RNA-seq experiments previously conducted with several human tissues (Tourancheau et al., 2016). Data indicated that the UGT2B10 transcriptome is comprised of one canonical and 10 alt. transcripts arising from the single *UGT2B10* gene. The UGT2B10 alt. transcripts are created by partial intronization of exon 1-2, exon skipping of exon 4-6 and inclusion of novel terminal exons 6b or 6c. UGT2B10 transcripts and encoded isoforms are depicted in **Figure 1**. Data revealed *UGT2B10* as one of the highest expressed UGTs in normal liver samples whereas variants arising from alternative splicing represented nearly 50% of the *UGT2B10* hepatic transcriptome (**Figure 2A**). Some variant classes were remarkably abundant, at levels comparable to those of the canonical *UGT2B10_v1* transcript encoding the UGT2B10 enzyme, namely those with N-terminal truncations and with novel C-terminal sequences (**Figure 2A**). Total *UGT2B10* expression was 2.15-fold ($P=0.013$) higher in hepatocellular carcinoma (HCC; $n=44$) relative to normal livers ($n=50$), and was highly variable among individuals (coefficients of variation of 93 to 117%) (**Figure 2B**). In contrast to the high hepatic levels, expression of *UGT2B10* in other normal tissues surveyed was much lower, with values below 1 TPM in the bladder, breast, colon, kidney, lung, prostate, and skin (data not shown).

Alternative splicing creates UGT2B10 variants with novel in-frame C-terminal sequences detected in human liver at the protein level

Splicing events generating *UGT2B10* transcripts ($n9$ and $n10$) harbouring novel C-terminal sequences were of particular interest for this study. These alt. variants are produced by intronization of parts of the canonical exon 6 and inclusion of a novel exon 6c (**Figure 1**). These AS events appeared specific to human and were not noted in other species based on BLAST searches in the non-redundant nucleotide collection and RNA Reference sequence

databases. Transcripts were abundant in human livers and were 3.5 fold higher in hepatic tumors relative to normal tissues. They represented on average 15% of total *UGT2B10* expression in normal liver tissues and 25% in HCC tissues (**Figure 2A-B**). The considerable interindividual variability for these alt. transcripts was higher than for the total and canonical *UGT2B10* hepatic expression. We also noted a significant positive correlation between the expression of the canonical and alt. transcripts ($\rho=0.854-0.915$, $P\leq 0.001$) (**Figure 2C**).

The putative UGT2B10 proteins encoded by these alt. transcripts are referred to as UGT2B10 isoforms 4 and 5, named i4 and i5 respectively. They are predicted to retain both the substrate and the co-substrate (UDP-GlcA) binding domains coupled to a novel alt. C-sequence (**Figure 1**). Isoform 4 lacks the 43 C-terminal amino acids including the transmembrane domain and the positively charged C-terminal tail of the UGT2B10 enzyme that are replaced by 10 novel amino acids encoded by exon 6c (**Figure 3A**). As for isoform 5, because a smaller portion of exon 6 is intronized, the encoded protein is predicted to retain 18 of the 24 transmembrane domain residues and to be extended by 65 novel amino acids, of which half are encoded by a frame-shift in exon 6, and half by exon 6c (**Figure 1, 3A**).

Validation of endogenous protein expression was possible for UGT2B10_i5 through the identification of a unique peptide sequence by targeted mass spectrometry-multiple reaction monitoring (MRM) of liver UGT2B10 immunoprecipitated from multiple donors (**Figure 3B,C**). This result is in line with the detection of the corresponding transcripts by PCR in human livers (**Supplemental Figure 2A**). The endogenous expression of UGT2B10_i4 could not be addressed by this approach given the short and hydrophobic nature of the C-terminal unique sequence.

Alternative isoforms with novel C-terminal sequences co-localized with the UGT2B10 enzyme and modified its activity *in vitro*

The alt. UGT2B10 isoforms were stably expressed in the embryonic kidney cell line HEK293, devoid of endogenous UGT expression, alone or with the canonical UGT2B10 enzyme. As well, their expression was examined in the liver cell model HepG2 that

expresses endogenously the UGT2B10 enzyme and conjugates substrates of the enzyme such as amitriptyline and levomedetomidine. We initially confirmed protein expression using immunoblot and immunofluorescence experiments using an anti-UGT2B10 antibody (#1845) that targets amino acids encoded by exon 1 and therefore recognizes the three UGT2B10 isoforms (**Figure 4A**). Indeed, we detected UGT2B10_i4 and UGT2B10_i5, near their predicted molecular weights (MW) of 57 and 66 kDa in cell models (**Figure 4A**).

The subcellular distribution of each protein was examined in the HEK293 models where the canonical and alt. proteins were detected and largely restricted to an endoplasmic reticulum (ER) localization. This was confirmed by the co-localization with the ER marker protein disulfide isomerase (PDI) (**Figure 4B**). Each isoform also displayed minor perinuclear and Golgi localization (**Supplemental Figure 3**). The glycosylation status of each protein was studied by subjecting microsomes from HEK293 cell models to Endo H glycosidase, which cleaves N-linked sugars on asparagine acquired by ER-resident enzymes, and to O-glycosidase, which removes serine and threonine O-linked complex sugars acquired in the Golgi. Each UGT2B10 protein was sensitive to Endo H treatment, revealed by the shift to a higher mobility protein band upon treatment (**Figure 4C**). In contrast, their mobility was not affected by a treatment with O-glycosidase.

Enzymatic assays in intact cells were subsequently conducted in the hepatic cell model HepG2 that expresses the endogenous UGT2B10 enzyme. Compared to the reference cell line (stably transfected with pcDNA6), expression of alt. UGT2B10_i4 isoform enhanced glucuronidation of the UGT2B10 enzyme with two drug substrates, amitriptyline and levomedetomidine, by 1.5 to 2 fold (**Figure 5A,B**). In turns, the presence of alt. UGT2B10_i5 significantly impaired the glucuronidation of amitriptyline and tended to reduce that of levomedetomidine as well, by 20-25%. Glucuronidation assays conducted with the UGT negative cells HEK293 stably expressing either alt. UGT2B10 isoforms revealed no transferase activity for amitriptyline and levomedetomidine. When co-expressed with the UGT2B10 enzyme, we observed a significant inhibition by 23 to 65% of glucuronidation activity by HEK293 cells in the presence of the alt. i4 or i5 proteins (**Figure 5A,B**). Since we observed a co-localization of alt. isoforms and the UGT2B10 enzyme in the ER, we addressed their potential interaction as a possible regulatory

mechanism. Immunoprecipitations were conducted with an anti-V5 epitope antibody using cell models stably expressing the UGT2B10 enzyme and transiently expressing either alt. isoform tagged with the V5 epitope. The UGT2B10 enzyme was immunoprecipitated with each alt. isoform, indicating their ability to form complexes (**Figure 5C**).

Alternative isoforms have shorter half-lives than the UGT2B10 enzyme and are targeted for degradation by the proteasome

Protein stability was evaluated in both cell models. The UGT2B10 enzyme displayed a half-life over 16 hours in both HEK293 (exogenous expression) and HepG2 (endogenous expression) cell models whereas the alt. isoforms displayed superior turnover rates. The alt. UGT2B10_i5 was the least stable, with short half-lives of 1.9 and 0.7 hours in HEK293 and HepG2, respectively (**Figure 6A**). The turnover rate of isoform i4 differed between the two cell models, and was 11.5 hours in HEK293 but much shorter in HepG2 (1.5 hours). We noted in HepG2 cells a significant recovery of alt. proteins expression, even rising above those of non-treated cells by 16h after initiation of the cycloheximide treatment. This was not observed for the UGT2B10 enzyme, nor for any of the UGT2B10 proteins in HEK293 cells. This observation implied a possible inactivation in HepG2 cells, which was confirmed by the detection of two glucuronides of cycloheximide (G1 and G2) with HepG2 microsomes (**Supplemental Figure 5**). This was further validated using microsomal fractions of pooled human livers and UGT supersomes. In these experiments, UGT1A9 and UGT2B7 most efficiently glucuronidated cycloheximide (G1 and G2), whereas UGT1A3 and UGT1A4 (G1) as well as UGT1A1 and UGT2B4 (G1 and G2) also conjugated some cycloheximide (**Supplemental Figure 5**), with some of them detected in HepG2 cells and not in the HEK293 cell model. As a consequence, this may lead to an inaccurate assessment of UGT proteins half-lives using this approach in the HepG2 model.

Accordingly, the difference in protein stability between the UGT2B10 enzyme and alt. proteins was further addressed by proteasomal inhibition. Whereas the enzyme levels were nearly unperturbed by inhibition of the ubiquitin-proteasome system, alt. UGT2B10 and more particularly isoform i5, were stabilized, indicating that they were degraded via the ubiquitin-proteasome system. Proteasomal inhibition for 16 h increased the ratio of alt. isoform / UGT2B10 enzyme in HepG2 liver cell models, whereas stabilization was more

modest in HEK293 (**Figure 6B**). Increased levels of alt. isoforms were not derived from enhanced transcription, verified at the mRNA level (**Supplemental Figure 2B**).

Differential induction of UGT2B10 alternative transcripts in liver cells by phenobarbital and a CAR agonist

HepaRG cells constitute a good surrogate system to study hepatic functions and response to drug treatments. An analysis of public HepaRG RNA-seq data (Li et al., 2015) with the exhaustive UGT transcriptome revealed an expression of canonical and alt. *UGT2B10* in comparable proportions to the human liver (not shown). In HepaRG cells treated with the constitutive androstane receptor (CAR) agonist CITCO, a significant and preferential induction of the alt. UGT2B10 transcripts (1.4 fold, $P=0.003$) occurred, whereas the enzyme-coding *v1* transcript was not significantly altered (**Figure 7**). Phenobarbital did not significantly induce expression of *UGT2B10* in wild type HepaRG cells. In contrast, in HepaRG cells with a CAR knock-out, alt. transcripts were significantly upregulated by 1.8 fold ($P\leq 0.001$) by phenobarbital, and 1.4 fold ($P=0.011$) by the CAR agonist, whereas *v1* was not significantly perturbed (**Figure 7**).

Discussion

The recent expansion of the pharmacogene transcriptome by AS has shed light on a novel mechanism regulating drug metabolism and clearance (Bellemare et al., 2010a; Bellemare et al., 2010b; Guillemette et al., 2010; Guillemette et al., 2014; Rouleau et al., 2014; Chhibber et al., 2016; Rouleau et al., 2016; Tourancheau et al., 2016; Tourancheau et al., 2017). Our study of the *UGT2B10* transcriptome, encoding a key detoxification enzyme specialized in N-glucuronidation of multiple harmful xenobiotics (Kaivosaari et al., 2007; Kaivosaari et al., 2011; Kato et al., 2013), demonstrated that AS accounts for a large proportion of *UGT2B10* gene expression and especially in the liver. This observation held true in liver tumors, where *UGT2B10* expression was enhanced by two-fold in hepatocellular carcinoma. Indeed, our analysis of next-generation sequencing data revealed that *UGT2B10* expression prevails in the liver whereas in all other tissues surveyed, including the lung, its expression was low to undetected. This is consistent with the expression determined at the RNA level in several human tissues, including those of the aerodigestive tract (Ohno and Nakajin, 2011; Court et al., 2012; Jones and Lazarus, 2014). It supports that a main detoxification site of UGT2B10-dependent N-glucuronidation is the liver, where UGT2B10 is one of the most abundant UGT enzyme based on proteomics data (Fallon et al., 2013; Sato et al., 2014; Margaiilan et al., 2015). Besides, AS also provides an explanation for the multiple observations reporting a lack of correlation between mRNA and protein expression, such as in hepatocellular carcinoma where the RNA expression remained equivalent between tumor tissues and adjacent normal tissues whereas glucuronidation activity was drastically decreased (Lu et al., 2015).

With a focus on one abundant class of hepatic alt. *UGT2B10* variants containing a novel 3' terminal exon that were confirmed at the protein level in human liver samples and in heterologous expression models, we expose their regulated expression and influence on UGT2B10 enzyme activity *in vitro*. In fact, the transcriptional regulation of *UGT2B10* has been poorly studied. A response element for the bile acid sensing farnesoyl X receptor (FXR) was recently uncovered in the *UGT2B10* promoter region and participated in the induction of *UGT2B10* by the FXR agonists GW4064 and chenodeoxycholic acid (Lu et al., 2017a). In the cell model HepaRG, a surrogate to human primary hepatocytes in drug

metabolism studies (Antherieu et al., 2012), RNA-seq data revealed drug-induced regulation of the *UGT2B10* transcriptome by both the constitutive androstane receptor (CAR) and pregnane X receptor (PXR). The superior induction of alt. *UGT2B10* by CAR and PXR agonists observed herein, especially in HepaRG cells devoid of CAR expression, further raises the possibility of a PXR-dependent remodelling of splicing events at the *UGT2B10* locus that may significantly influence UGT2B10 detoxification activity. Whether other receptors such as FXR, previously reported to regulate *UGT2B10* transcription (Lu et al., 2017a), also influence splicing remain to be addressed. When expressed in human cells, alt. UGT2B10 acted as regulators of the glucuronidation activity of the UGT2B10 enzyme, possibly conveyed by heterologous complexes formed between the enzyme and alternative proteins; a regulatory mechanism documented for other human UGTs (Bushey and Lazarus, 2012; Menard et al., 2013; Rouleau et al., 2014; Rouleau et al., 2016). Our findings further suggest a cell-specific influence given that in HepG2 cells, an increased N-glucuronidation of the UGT2B10 substrates amitriptyline and levomedetomidine by the endogenous enzyme was observed, in contrast with a repression of enzyme activity in HEK293 cells. The endogenous expression of additional UGTs other than UGT2B10 in HepG2 cells and different protein turnover rates could be among factors influencing their functions. Likewise, the impact of AS on UGT2B10 activity has been documented previously by findings of a common polymorphism (rs116294140) that disrupts a splice site in exon 3 and introduces a premature stop codon possibly triggering non-sense mRNA decay (Fowler et al., 2015). This polymorphism, particularly more frequent among African Americans, significantly reduced N-glucuronidation of drugs such as RO5263397 as well as nicotine and cotinine (Murphy et al., 2014; Fowler et al., 2015). In fact, occurrence of this splice site variant as well as the coding variant Asp67Tyr (rs61750900) were estimated to collectively explain over 24% of interindividual variability in cotinine glucuronidation (Patel et al., 2015). Our results support that alternative splicing at the *UGT2B10* locus may be a major factor contributing to this variability in the constitutive expression of the gene, with a potential impact on responses to substrates of the UGT2B10 pathway.

The alternate UGT2B10 proteins with novel in-frame C-terminal sequences are predicted to include the entire putative catalytic domain (Radomska-Pandya et al., 2010), although

their enzyme activity could not be confirmed in standard *in vitro* assay conditions when expressed in HEK293 human cells. This could be due to an inadequate topology of the alt. proteins essential for enzyme function. *In silico* analysis of the novel exonic sequence with NCBI BLAST tools did not reveal a match with other nucleotide or amino acid sequences of any organism. The splicing event at the *UGT2B10* locus appears specific to humans. This is consistent with the low to undetectable activity towards preferred UGT2B10 substrates such as N-heterocyclic amines and aliphatic tertiary amines in most primates and other mammals, suggesting that *UGT2B10* expression occurs preferentially in humans (Kaivosari et al., 2007; Kaivosari et al., 2008; Zhou et al., 2010; Lu et al., 2017b). The novel appended amino acid sequence encoded in the alternative frame by exon 6 is however related to that of putative alternative variants of several other human proteins including CLCN3 (chloride exchange transporter 3), ALG9 (alpha-1, 2-mannosyltransferase), and CHFR (E3-ubiquitin ligase) supporting that the new sequence may encode a conserved domain (**Supplemental Figure 6**).

As for subcellular localization, our immunofluorescence data indicated that the alt. isoforms, partially or completely lacking the transmembrane domain and devoid of the positively charged lysine tail (**Figure 3**) are nonetheless ER-resident proteins when expressed in HEK293 human cells. Although studied in a limited set of UGTs, the high sequence similarity of ER retention elements among UGTs supports a common ER retention mechanism. ER residency of UGT enzymes is mediated by at least four protein regions: the N-terminal signal peptide, a short hydrophobic patch in the N-terminal substrate binding domain, the C-terminal transmembrane region and the cytoplasmic dilysine motif containing the ER retention signal "KXKXX" (K, lysines, X, any amino acids) (Jackson et al., 1993; Ciotti et al., 1998; Ouzzine et al., 1999; Barre et al., 2005). In line with our observations, these structural features may be partially redundant, as none appears strictly essential to ER retention. The Endo H sensitivity and lack of sensitivity towards O-glycosidase for each of the UGT2B10 enzyme and alt. UGT2B10 also support ER residency. Consistent with their co-localization, the potential of alt. UGT2B10 to interact with the UGT2B10 enzyme suggests that this may be a mechanism underlying their regulatory function.

The preferential stabilization of alt. isoforms by proteasomal inhibition and their shorter half-lives relative to the UGT2B10 enzyme suggest that distinct pathways may govern the turnover of alt. isoforms and the enzyme. This also suggests that a small variation in RNA levels has the potential to affect AS protein expression level. The significant abundance of alt. transcripts encoding these proteins may indicate that hepatic cells are poised to adapt levels of regulatory isoforms in response to various endogenous or exogenous stimuli. This is supported by the preferential mRNA expression of alternates in HepaRG cells treated with nuclear receptors agonists. While the structural determinants of UGT protein stability are scarcely known, mechanisms modulating UGT turnover may significantly contribute to the regulation of their detoxifying functions, a well-documented aspect for some drug metabolizing CYPs (Zhukov and Ingelman-Sundberg, 1999; Kim et al., 2016 and references therein). When expressed in the HEK293 model, half-lives of UGTs were more than 12-16 hours for UGT2B4, UGT2B7, UGT2B10 and UGT2B15 whereas UGT1A1 and UGT2B17 were more labile with half-lives less than 3 hours (Turgeon et al., 2001; Rouleau et al., 2016); this study). However, given that cycloheximide is glucuronidated by several UGT enzymes as demonstrated herein, alt. UGT protein half-lives established by translational inhibition with this compound in cells expressing UGTs likely constitute an inaccurate estimate that will be influenced by the enzymes expressed.

A limitation of our study is the lack of detection of alt. UGT2B10 by immunoblotting in pooled liver microsomes from 50 individuals using the polyclonal anti-UGT2B10 antibody #1845. Their expression varied widely among individuals (CV of 155%), with some having no or barely detectable hepatic alt. *UGT2B10* variants. Levels of alt. proteins in our human liver pool may be too low for detection by immunoblotting. In contrast, the detection of alt. UGT2B10 by MS-MRM was performed following an immunoprecipitation step that enriched UGT2B10, thus improving the sensitivity of detection. The high turnover rate of the alt. proteins, as observed in the two cell models, may also influence our ability to detect them in human livers by immunoblotting. Additional experiments are required to ascertain the expression of alt. UGT2B10 proteins in individual human liver samples and their expression ratio relative to the UGT2B10 enzyme.

In conclusion, our study reveals that AS creates a diversified *UGT2B10* transcriptome and represents half of the UGT2B10 expression in the human liver, with a wide interindividual variability. Alternate UGT2B10 proteins may significantly influence the UGT2B10-dependent detoxification of amine-containing drugs such as anti-psychotic and tobacco metabolites and are expected to modulate endogenous substrates of the UGT2B10 enzyme, which are currently unknown. Our study further highlights a long-term stability of the UGT2B10 enzyme that contrasts with the lability of alt. proteins, the latter being regulated by proteasomal degradation. Most interestingly, we further expose a preferential induction by PXR and CAR inducers of alt. *UGT2B10* with novel in-frame C-terminal sequences in hepatic cells, implying a fine regulation of the AS process by xenosensing transcription factors. Our study highlights an important regulatory role of AS in UGT2B10 expression and detoxification functions that may explain part of the significant variability in N-glucuronidation, largely mediated by the UGT2B10 pathway in the liver. We thus believe that interindividual differences in the clinical response to UGT2B10 substrates are likely to be understood through the AS process affecting both the constitutive and inducible expression of UGT2B10.

Acknowledgements

Authors wish to thank Dr H. Wang and B. Mackowiak (U. of Maryland) for their collaboration with the HepaRG data and Patrick Caron, Véronique Turcotte, Anne-Marie Duperré, Camille Girard-Bock and Andréa Fournier for their technical assistance. We also thank the Proteomics Platform of the CHU de Québec Research Center for their services.

Authorship contribution

Participated in research design: Labriet, Allain, Rouleau, Audet-Delage, Guillemette

Conducted experiments: Labriet, Allain, Audet-Delage, Villeneuve

Performed data analysis: Labriet, Allain, Rouleau, Audet-Delage, Villeneuve, Guillemette

Wrote or contributed to the writing of the manuscript: Labriet, Rouleau, Guillemette

Financial disclosure

All authors declare they have no competing interest

References

- Antherieu S, Chesne C, Li R, Guguen-Guillouzo C, and Guillouzo A (2012) Optimization of the HepaRG cell model for drug metabolism and toxicity studies. *Toxicol In Vitro* **26**:1278-1285.
- Barre L, Magdalou J, Netter P, Fournel-Gigleux S, and Ouzzine M (2005) The stop transfer sequence of the human UDP-glucuronosyltransferase 1A determines localization to the endoplasmic reticulum by both static retention and retrieval mechanisms. *FEBS J* **272**:1063-1071.
- Beaulieu M, Levesque E, Hum DW, and Belanger A (1998) Isolation and characterization of a human orphan UDP-glucuronosyltransferase, UGT2B11. *Biochem Biophys Res Commun* **248**:44-50.
- Bellemare J, Rouleau M, Girard H, Harvey M, and Guillemette C (2010a) 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.
- Bellemare J, Rouleau M, Harvey M, and Guillemette C (2010b) Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *J Biol Chem* **285**:3600-3607.
- Berg JZ, Mason J, Boettcher AJ, Hatsukami DK, and Murphy SE (2010) Nicotine metabolism in African Americans and European Americans: variation in glucuronidation by ethnicity and UGT2B10 haplotype. *J Pharmacol Exp Ther* **332**:202-209.
- Bolger AM, Lohse M, and Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114-2120.
- Bray NL, Pimentel H, Melsted P, and Pachter L (2016) Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* **34**:525-527.
- 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.
- Chen G, Blevins-Primeau AS, Dellinger RW, Muscat JE, and Lazarus P (2007) Glucuronidation of nicotine and cotinine by UGT2B10: loss of function by the UGT2B10 Codon 67 (Asp>Tyr) polymorphism. *Cancer Res* **67**:9024-9029.
- Chen G, Dellinger RW, Gallagher CJ, Sun D, and Lazarus P (2008a) Identification of a prevalent functional missense polymorphism in the UGT2B10 gene and its association with UGT2B10 inactivation against tobacco-specific nitrosamines. *Pharmacogenet Genomics* **18**:181-191.
- Chen G, Dellinger RW, Sun D, Spratt TE, and Lazarus P (2008b) Glucuronidation of tobacco-specific nitrosamines by UGT2B10. *Drug Metab Dispos* **36**:824-830.
- Chen G, Giambone NE, and Lazarus P (2012) Glucuronidation of trans-3'-hydroxycotinine by UGT2B17 and UGT2B10. *Pharmacogenet Genomics* **22**:183-190.
- Chhibber A, French CE, Yee SW, Gamazon ER, Theusch E, Qin X, Webb A, Papp AC, Wang A, Simmons CQ, Konkashbaev A, Chaudhry AS, Mitchel K, Stryke D, Ferrin TE, Weiss ST, Kroetz DL, Sadee W, Nickerson DA, Krauss RM, George AL, Schuetz EG, Medina MW, Cox NJ, Scherer SE, Giacomini KM, and Brenner SE

- (2016) Transcriptomic variation of pharmacogenes in multiple human tissues and lymphoblastoid cell lines. *Pharmacogenomics J. In press.*
- Chouinard S, Pelletier G, Belanger A, and Barbier O (2006) Isoform-specific regulation of uridine diphosphate-glucuronosyltransferase 2B enzymes in the human prostate: differential consequences for androgen and bioactive lipid inactivation. *Endocrinology* **147**:5431-5442.
- Ciotti M, Cho JW, George J, and Owens IS (1998) Required buried alpha-helical structure in the bilirubin UDP-glucuronosyltransferase, UGT1A1, contains a nonreplaceable phenylalanine. *Biochemistry* **37**:11018-11025.
- Court MH, Zhang X, Ding X, Yee KK, Hesse LM, and Finel M (2012) Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica* **42**:266-277.
- de Jonge HJ, Fehrmann RS, de Bont ES, Hofstra RM, Gerbens F, Kamps WA, de Vries EG, van der Zee AG, te Meerman GJ, and ter Elst A (2007) Evidence based selection of housekeeping genes. *PLoS One* **2**:e898.
- Fallon JK, Neubert H, Hyland R, Goosen TC, and Smith PC (2013) Targeted quantitative proteomics for the analysis of 14 UGT1As and -2Bs in human liver using NanoUPLC-MS/MS with selected reaction monitoring. *J Proteome Res* **12**:4402-4413.
- Fowler S, Kletzl H, Finel M, Manevski N, Schmid P, Tuerck D, Norcross RD, Hoener MC, Spleiss O, and Iglesias VA (2015) A UGT2B10 splicing polymorphism common in african populations may greatly increase drug exposure. *J Pharmacol Exp Ther* **352**:358-367.
- Girard-Bock C, Benoit-Biancamano MO, Villeneuve L, Desjardins S, and Guillemette C (2016) A Rare UGT2B7 Variant Creates a Novel N-Glycosylation Site at Codon 121 with Impaired Enzyme Activity. *Drug Metab Dispos* **44**:1867-1871.
- Guillemette C, Levesque E, Harvey M, Bellemare J, and Menard V (2010) UGT genomic diversity: beyond gene duplication. *Drug Metab Rev* **42**:24-44.
- Guillemette C, Levesque E, and Rouleau M (2014) Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther* **96**:324-339.
- Haakensen VD, Biong M, Lingjaerde OC, Holmen MM, Frantzen JO, Chen Y, Navjord D, Romundstad L, Luders T, Bukholm IK, Solvang HK, Kristensen VN, Ursin G, Borresen-Dale AL, and Helland A (2010) Expression levels of uridine 5'-diphospho-glucuronosyltransferase genes in breast tissue from healthy women are associated with mammographic density. *Breast Cancer Res* **12**:R65.
- Jackson MR, Nilsson T, and Peterson PA (1993) Retrieval of transmembrane proteins to the endoplasmic reticulum. *J Cell Biol* **121**:317-333.
- Jones NR and Lazarus P (2014) UGT2B gene expression analysis in multiple tobacco carcinogen-targeted tissues. *Drug Metab Dispos* **42**:529-536.
- Kaivosaari S, Finel M, and Koskinen M (2011) N-glucuronidation of drugs and other xenobiotics by human and animal UDP-glucuronosyltransferases. *Xenobiotica* **41**:652-669.
- Kaivosaari S, Toivonen P, Aitio O, Sipila J, Koskinen M, Salonen JS, and Finel M (2008) Regio- and stereospecific N-glucuronidation of medetomidine: the differences between UDP glucuronosyltransferase (UGT) 1A4 and UGT2B10 account for the complex kinetics of human liver microsomes. *Drug Metab Dispos* **36**:1529-1537.

- Kaivosaaari S, Toivonen P, Hesse LM, Koskinen M, Court MH, and Finel M (2007) Nicotine glucuronidation and the human UDP-glucuronosyltransferase UGT2B10. *Mol Pharmacol* **72**:761-768.
- Kato Y, Izukawa T, Oda S, Fukami T, Finel M, Yokoi T, and Nakajima M (2013) Human UDP-glucuronosyltransferase (UGT) 2B10 in drug N-glucuronidation: substrate screening and comparison with UGT1A3 and UGT1A4. *Drug Metab Dispos* **41**:1389-1397.
- Kazmi F, Barbara JE, Yerino P, and Parkinson A (2015) A long-standing mystery solved: the formation of 3-hydroxydesloratadine is catalyzed by CYP2C8 but prior glucuronidation of desloratadine by UDP-glucuronosyltransferase 2B10 is an obligatory requirement. *Drug Metab Dispos* **43**:523-533.
- Kerdpin O, Mackenzie PI, Bowalgaha K, Finel M, and Miners JO (2009) Influence of N-terminal domain histidine and proline residues on the substrate selectivities of human UDP-glucuronosyltransferase 1A1, 1A6, 1A9, 2B7, and 2B10. *Drug Metab Dispos* **37**:1948-1955.
- Kim SM, Wang Y, Nabavi N, Liu Y, and Correia MA (2016) Hepatic cytochromes P450: structural degons and barcodes, posttranslational modifications and cellular adapters in the ERAD-endgame. *Drug Metab Rev* **48**:405-433.
- Lepine J, Bernard O, Plante M, Tetu B, Pelletier G, Labrie F, Belanger A, and Guillemette C (2004) Specificity and regioselectivity of the conjugation of estradiol, estrone, and their catecholestrogen and methoxyestrogen metabolites by human uridine diphospho-glucuronosyltransferases expressed in endometrium. *J Clin Endocrinol Metab* **89**:5222-5232.
- Levesque E, Beaulieu M, Green MD, Tephly TR, Belanger A, and Hum DW (1997) Isolation and characterization of UGT2B15(Y85): a UDP-glucuronosyltransferase encoded by a polymorphic gene. *Pharmacogenetics* **7**:317-325.
- Li D, Mackowiak B, Brayman TG, Mitchell M, Zhang L, Huang SM, and Wang H (2015) Genome-wide analysis of human constitutive androstane receptor (CAR) transcriptome in wild-type and CAR-knockout HepaRG cells. *Biochem Pharmacol* **98**:190-202.
- Lu D, Wang S, Xie Q, Guo L, and Wu B (2017a) Transcriptional Regulation of Human UDP-Glucuronosyltransferase 2B10 by Farnesoid X Receptor in Human Hepatoma HepG2 Cells. *Mol Pharm* **14**:2899-2907.
- Lu D, Xie Q, and Wu B (2017b) N-glucuronidation catalyzed by UGT1A4 and UGT2B10 in human liver microsomes: Assay optimization and substrate identification. *J Pharm Biomed Anal* **145**:692-703.
- Lu L, Zhou J, Shi J, Peng XJ, Qi XX, Wang Y, Li FY, Zhou FY, Liu L, and Liu ZQ (2015) Drug-Metabolizing Activity, Protein and Gene Expression of UDP-Glucuronosyltransferases Are Significantly Altered in Hepatocellular Carcinoma Patients. *PLoS One* **10**:e0127524.
- Margaillan G, Rouleau M, Klein K, Fallon JK, Caron P, Villeneuve L, Smith PC, Zanger UM, and Guillemette C (2015) Multiplexed Targeted Quantitative Proteomics Predicts Hepatic Glucuronidation Potential. *Drug Metab Dispos* **43**:1331-1335.
- Menard V, Collin P, Margaillan G, and Guillemette C (2013) Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos* **41**:2197-2205.

- Murphy SE (2017) Nicotine Metabolism and Smoking: Ethnic Differences in the Role of P450 2A6. *Chem Res Toxicol* **30**:410-419.
- Murphy SE, Park SS, Thompson EF, Wilkens LR, Patel Y, Stram DO, and Le Marchand L (2014) Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis* **35**:2526-2533.
- NCBI dbSNP build 151 accessed Dec. 8, 2017
https://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi
- Ohno S and Nakajin S (2011) Quantitative analysis of UGT2B28 mRNA expression by real-time RT-PCR and application to human tissue distribution study. *Drug Metab Lett* **5**:202-208.
- Ouzzine M, Magdalou J, Burchell B, and Fournel-Gigleux S (1999) An internal signal sequence mediates the targeting and retention of the human UDP-glucuronosyltransferase 1A6 to the endoplasmic reticulum. *J Biol Chem* **274**:31401-31409.
- Patel YM, Stram DO, Wilkens LR, Park SS, Henderson BE, Le Marchand L, Haiman CA, and Murphy SE (2015) The contribution of common genetic variation to nicotine and cotinine glucuronidation in multiple ethnic/racial populations. *Cancer Epidemiol Biomarkers Prev* **24**:119-127.
- Pattanaowongsa A, Nair PC, Rowland A, and Miners JO (2016) Human UDP-Glucuronosyltransferase (UGT) 2B10: Validation of Cotinine as a Selective Probe Substrate, Inhibition by UGT Enzyme-Selective Inhibitors and Antidepressant and Antipsychotic Drugs, and Structural Determinants of Enzyme Inhibition. *Drug Metab Dispos* **44**:378-388.
- Radomska-Pandya A, Bratton SM, Redinbo MR, and Miley MJ (2010) The crystal structure of human UDP-glucuronosyltransferase 2B7 C-terminal end is the first mammalian UGT target to be revealed: the significance for human UGTs from both the 1A and 2B families. *Drug Metab Rev* **42**:133-144.
- Rath A, Glibowicka M, Nadeau VG, Chen G, and Deber CM (2009) Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proc Natl Acad Sci U S A* **106**:1760-1765.
- Rouleau M, Roberge J, Bellemare J, and Guillemette C (2014) Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Mol Pharmacol* **85**:29-36.
- Rouleau M, Tourancheau A, Girard-Bock C, Villeneuve L, Vaucher J, Duperre AM, Audet-Delage Y, Gilbert I, Popa I, Droit A, and Guillemette C (2016) Divergent Expression and Metabolic Functions of Human Glucuronosyltransferases through Alternative Splicing. *Cell Rep* **17**:114-124.
- Sato Y, Nagata M, Tetsuka K, Tamura K, Miyashita A, Kawamura A, and Usui T (2014) Optimized methods for targeted peptide-based quantification of human uridine 5'-diphosphate-glucuronosyltransferases in biological specimens using liquid chromatography-tandem mass spectrometry. *Drug Metab Dispos* **42**:885-889.
- Tourancheau A, Margaillan G, Rouleau M, Gilbert I, Villeneuve L, Levesque E, Droit A, and Guillemette C (2016) Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J* **16**:60-70.

- Tourancheau A, Rouleau M, Guauque-Olarte S, Villeneuve L, Gilbert I, Droit A, and Guillemette C (2017) Quantitative profiling of the UGT transcriptome in human drug-metabolizing tissues. *Pharmacogenomics J. In press.*
- Turgeon D, Carrier JS, Levesque E, Hum DW, and Belanger A (2001) Relative enzymatic activity, protein stability, and tissue distribution of human steroid-metabolizing UGT2B subfamily members. *Endocrinology* **142**:778-787.
- Turgeon D, Chouinard S, Belanger P, Picard S, Labbe JF, Borgeat P, and Belanger A (2003) Glucuronidation of arachidonic and linoleic acid metabolites by human UDP-glucuronosyltransferases. *J Lipid Res* **44**:1182-1191.
- Ware JJ, Chen X, Vink J, Loukola A, Minica C, Pool R, Milaneschi Y, Mangino M, Menni C, Chen J, Peterson RE, Auro K, Lytikainen LP, Wedenoja J, Stiby AI, Hemani G, Willemsen G, Hottenga JJ, Korhonen T, Heliövaara M, Perola M, Rose RJ, Paternoster L, Timpson N, Wassenaar CA, Zhu AZ, Davey Smith G, Raitakari OT, Lehtimäki T, Kahonen M, Koskinen S, Spector T, Penninx BW, Salomaa V, Boomsma DI, Tyndale RF, Kaprio J, and Munafò MR (2016) Genome-Wide Meta-Analysis of Cotinine Levels in Cigarette Smokers Identifies Locus at 4q13.2. *Sci Rep* **6**:20092.
- Zhou D, Guo J, Linnenbach AJ, Booth-Genthe CL, and Grimm SW (2010) Role of human UGT2B10 in N-glucuronidation of tricyclic antidepressants, amitriptyline, imipramine, clomipramine, and trimipramine. *Drug Metab Dispos* **38**:863-870.
- Zhukov A and Ingelman-Sundberg M (1999) Relationship between cytochrome P450 catalytic cycling and stability: fast degradation of ethanol-inducible cytochrome P450 2E1 (CYP2E1) in hepatoma cells is abolished by inactivation of its electron donor NADPH-cytochrome P450 reductase. *Biochem J* **340** (Pt 2):453-458.

Footnotes

This work was supported by the Canadian Institutes of Health Research [FRN-42392]; and the Canada Research Chair in Pharmacogenomics (Tier I); AL and EPA were supported by graduate scholarships from the "Fonds d'enseignement et de recherche" (FER) of the Faculty of pharmacy, Laval University (graduate scholarship); YAD was supported by a graduate scholarship from the "Fonds de Recherche en Santé du Québec". The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by National Cancer Institute (NCI), National Human Genome Research Institute (NHGRI), National Heart, Lung, and Blood Institute, National Institute on Drug Abuse, National Institute of Mental Health, and National Institute of Neurological Disorders and Stroke and The Cancer Genome Atlas (TCGA) is managed by the NCI and NHGRI (<http://cancergenome.nih.gov>).

Figure legends

Figure 1. Schematic overview of *UGT2B10* mRNA transcripts and encoded proteins.

Top. The *UGT2B10* gene is comprised of 6 canonical exons and 2 alternative exons 6b and 6c. *Bottom.* The canonical *UGT2B10_v1* transcript encodes the UGT2B10_i1 enzyme. Alternative splicing produces 3 classes of variants, with N-terminal or C-terminal truncations, and novel C-terminal sequence. The N-terminally truncated *UGT2B10_n3* and *_n4* variants are reported in the NCBI RefSeq *UGT2B10* gene entry (<https://www.ncbi.nlm.nih.gov/gene/7365>) to encode, respectively, isoforms i3 and i2 in RefSeq. Consequently, the isoform encoded by the transcripts *n9* and *n10* were named i4 and i5 respectively. Exons are represented by colored boxes and skipped exons by dashed boxes. Thinner parts of exons 1 and 6 represent untranslated regions. UGT2B10 is the UGT2B with the largest 3'untranslated region and only the first 279 of 1424 nucleotides of exon 6 are coding in the canonical mRNA. Note that the predicted molecular weight of mature proteins is smaller by 2.5 kDa due to the cleavage of the signal peptide.

Figure 2. Alternative splicing diversifies the hepatic *UGT2B10* transcriptome. A.

Relative levels of canonical (*v1*) and alt. transcript classes in the normal human liver and in hepatocellular carcinoma. The Δ C-term+new sequence class is predominantly comprised of *n9* and *n10* transcripts whereas transcript *n8* represented less than 2% in normal and tumor tissues. B. Interindividual variability in *UGT2B10* expression for all transcripts (Total), canonical (*v1*) and alt. transcripts (*n9/n10*). Boxes represent 25-75 percentiles, whiskers 10-90 percentiles. Median is indicated by the horizontal line and mean by a "+". CV, coefficient of variation; FC, fold change; N: normal tissues; T: tumor tissues; TPM: transcripts per million. C. Correlation between *UGT2B10_v1* and *n9/n10* expression in normal livers and hepatocellular carcinoma. All expression data were derived by a realignment of RNA-Seq data from GTEx (n = 50) and TCGA (n = 44) to the fully annotated UGT variant sequence (Tourancheau et al., 2016).

Figure 3. Alternate UGT2B10 protein is expressed in the human liver. A.

C-terminal amino acid sequences of the UGT2B10 enzyme (i1) and alt. isoforms i4 and i5. Sequences unique to each alt. proteins are italicized. TMD: transmembrane domain; K⁺: ER retention signal. B. Experimental approach for the detection of UGT2B10_i5 by

immunoprecipitation and multiple reaction monitoring (MRM). **C.** The common UGT2B10 peptide NSWNFK (left) and the alt. specific peptide LLGSSNNPPILASQR (right) were detected in tryptic digests of UGT2B10 immunoprecipitated from human liver samples (upper chromatograms). The chromatograms of control peptides (lower chromatograms) labelled with stable isotopes mixed with the immunopurified UGT2B10 confirmed the identity of i1 and i5 peptides. Representative chromatograms are shown (n = 2).

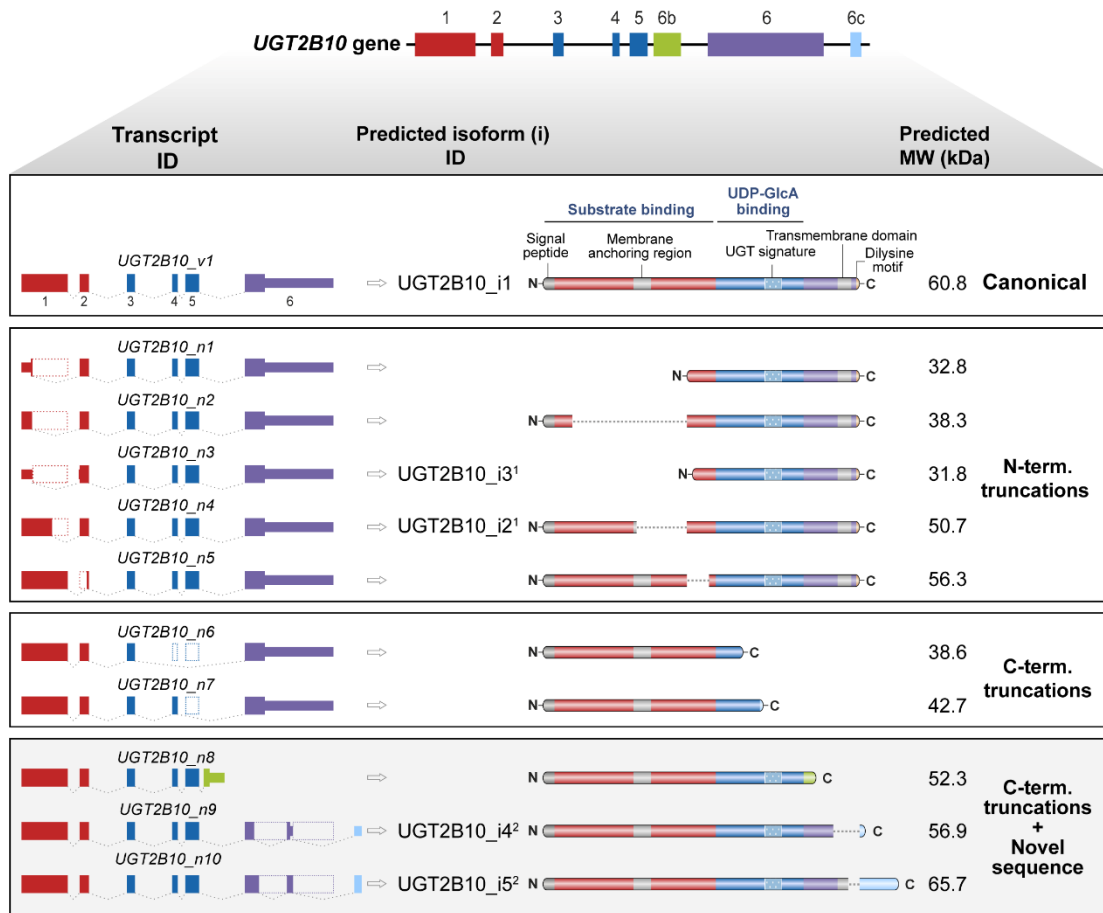
Figure 4. Interaction of UGT2B10 isoforms in the ER. **A.** Stable expression of UGT2B10 alternative isoforms in HEK293 and HepG2 cell models. UGT2B10 was detected in microsomal fractions of each cell model (20 µg proteins) and of human liver (L) using the anti-UGT2B10 antibody. The expression ratio of the UGT2B10 enzyme relative to alt. isoforms is estimated to more than 4 fold in each model. Note that the apparent molecular weight of the enzyme was lower than predicted, likely explained by the influence of the hydrophobic amino acids in the transmembrane domains (TMD) that perturb the interactions with SDS and the shape of the denatured protein (Rath et al., 2009). **B.** UGT2B10 enzyme and alt. isoforms (i4 and i5) co-localized with the ER marker protein disulfide isomerase (PDI). Merged images of UGT2B10 proteins labelled with anti-UGT2B10 (green), PDI (red), and nuclei (blue) are shown. Separate images and co-localization with other subcellular markers are shown in **Supplemental Figure 3**. Bar = 10 µm. **C.** Glycosylation status of the UGT2B10 enzyme and alt. proteins. UGT2B10 proteins in microsomes from HEK293 cell models subjected to endoglycosidase H (Endo H) or O-glycosidase were detected by immunoblotting.

Figure 5. The alternative UGT2B10 isoforms modulate the glucuronidation activity of the UGT2B10 enzyme. **A-B.** *In situ* cell glucuronidation assays of amitriptyline and levomedetomidine in HEK293 and HepG2 cell models that co-express the alt. isoforms i4 and i5. Activity is presented relative to that in cells expressing only the UGT2B10 enzyme (control). Assays were conducted twice in triplicates. *t*, $P \leq 0.1$; *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. **C.** Co-immunoprecipitation of the UGT2B10 enzyme with alt. isoforms i4 and i5 expressed in HEK293. Alternative isoforms were C-terminally tagged with the epitope V5 and immunoprecipitated (IP) with the anti-V5 antibody.

Figure 6. Stability of UGT2B10 proteins. **A.** The half-life of UGT2B10 proteins was determined by translational inhibition with cycloheximide (CHX) in HEK293 and HepG2 cell models. Protein levels were measured by densitometry scanning of immunoblots from two independent assays and are expressed as % of untreated cells. Half-lives were averaged from the two independent assays. A representative experiment is shown. An independent replicate assay is presented in **Supplemental Figure 4.** **B.** UGT2B10 alternative isoforms are stabilized by proteasomal inhibition. HEK293 and HepG2 cell models were exposed to vehicle only (DMSO; lane "-") or MG132 (1 μ M; lanes "+") for 16 h. UGT2B10 was detected by immunoblotting total cell lysates (40 μ g) with anti-UGT2B10.

Figure 7. Differential induction of alt. UGT2B10 by phenobarbital and a CAR agonist in HepaRG cells. Expression data of *UGT2B10* canonical and alt. transcripts *n9* and *n10* (encoding i4 and i5) in HepaRG cells wild type (WT) or with a knock-out of the constitutive androstane receptor (CAR KO) were obtained from the public RNA-seq data GSE 71446. Cells were treated either with vehicle (DMSO), phenobarbital (PB, 1 mM) or the CAR agonist CITCO (1 μ M) for 24 h as described (Li et al., 2015). **, $P \leq 0.01$; ***, $P \leq 0.001$.

Figure 1.



¹Nomenclature from RefSeq *Homo sapiens* annotation release 108 (<https://www.ncbi.nlm.nih.gov/gene/7365>)

²Nomenclature from Tourancheau et al. (2017)

Focus of this study

Figure 2.

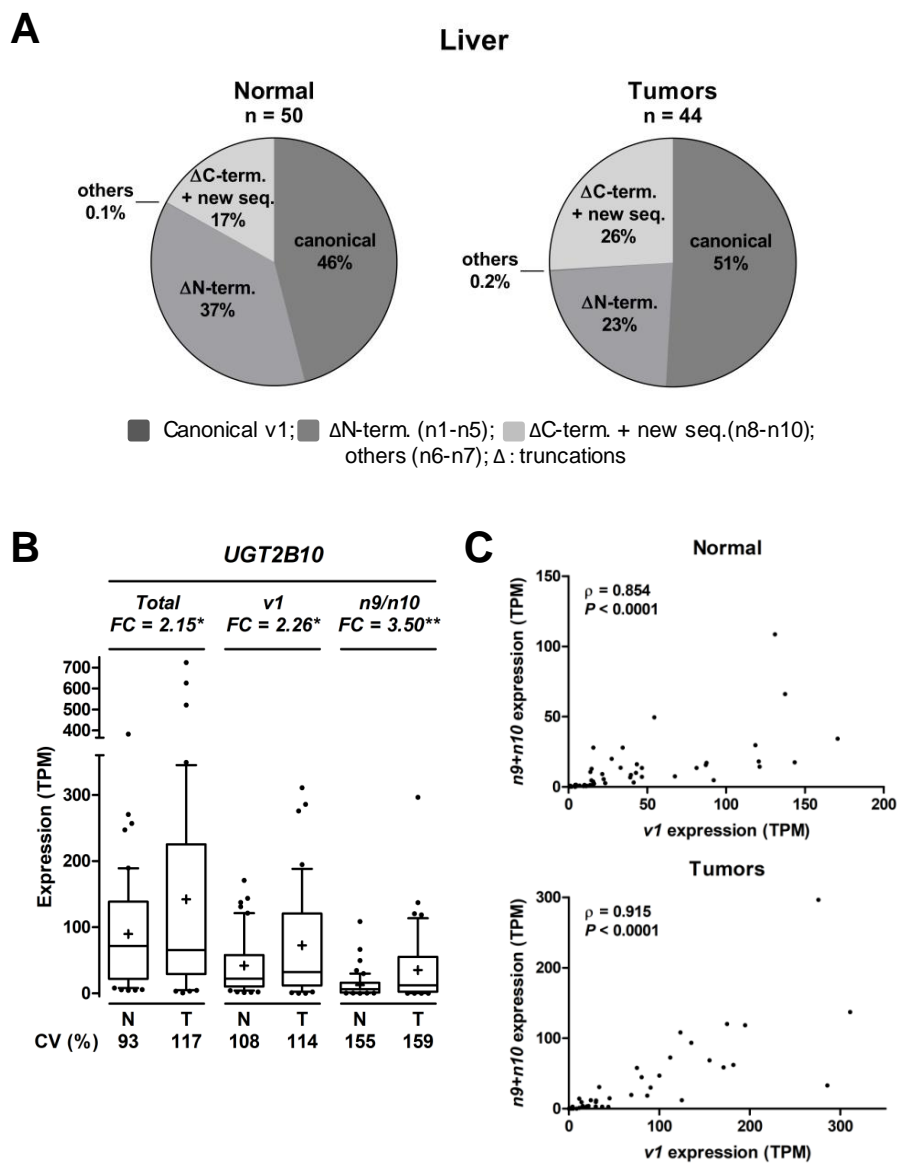


Figure 3.

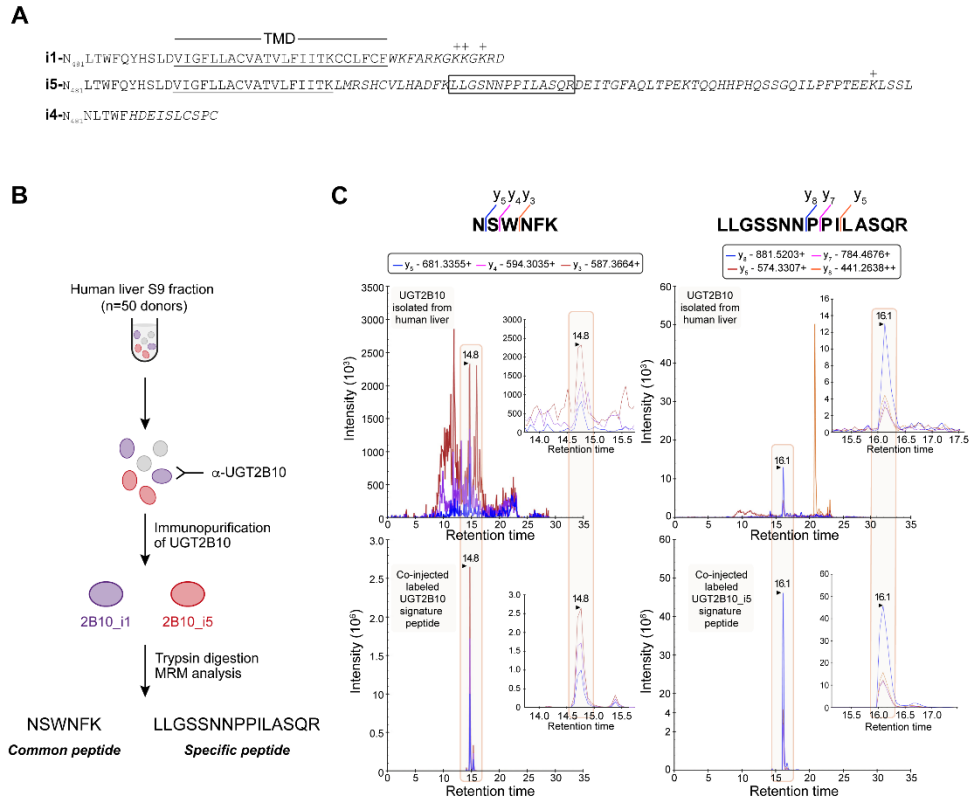


Figure 4.

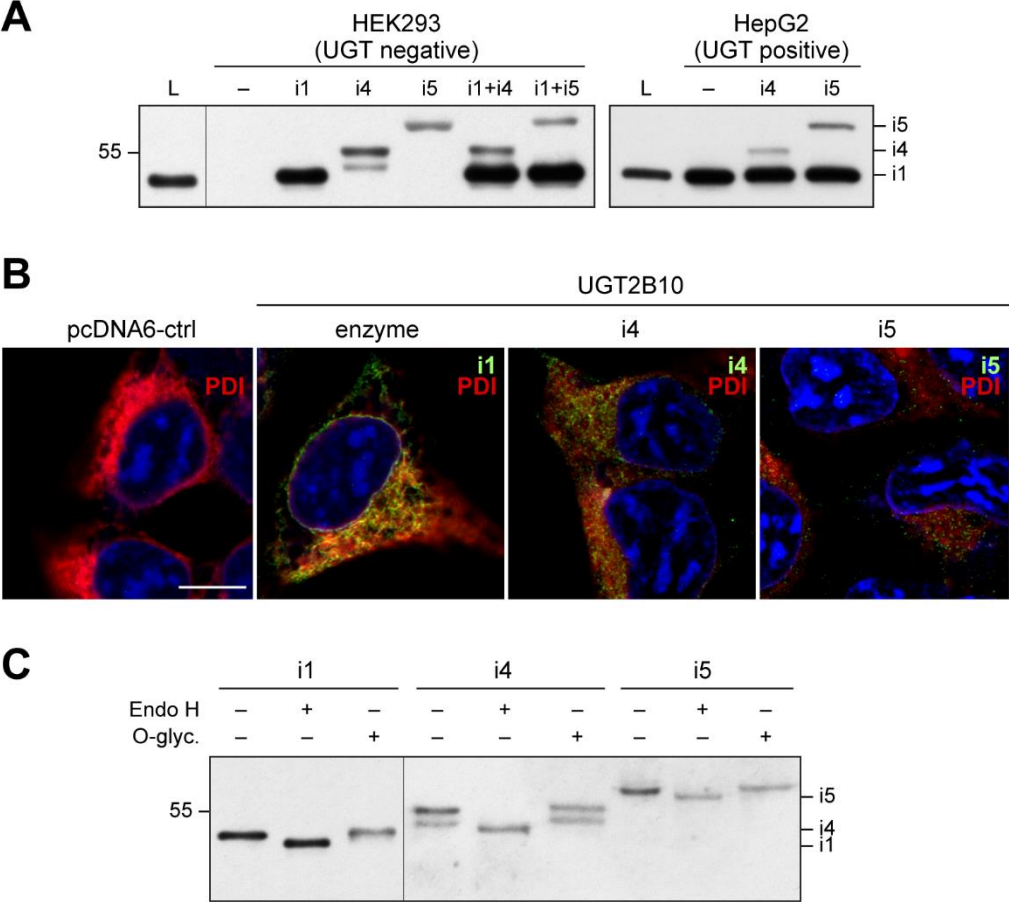


Figure 5.

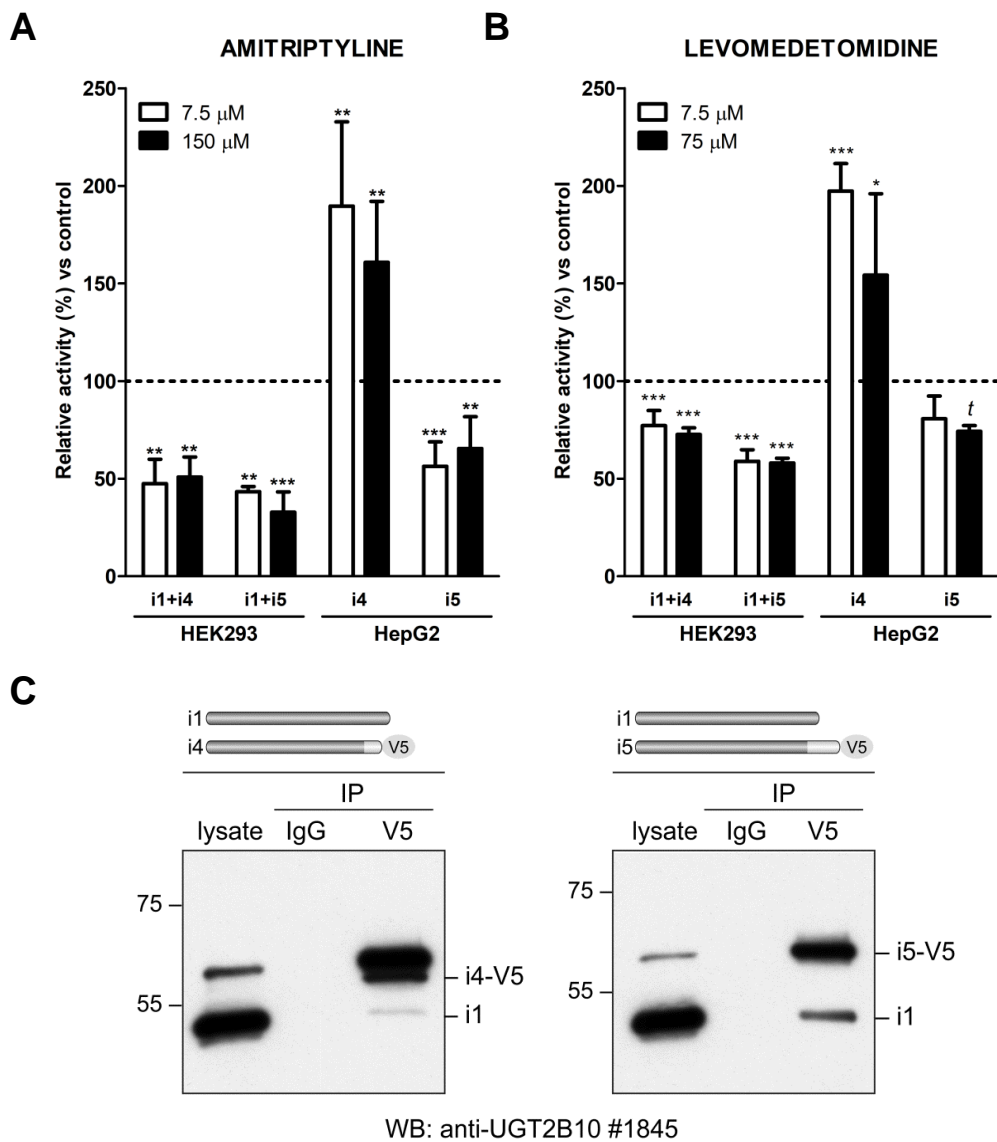


Figure 6.

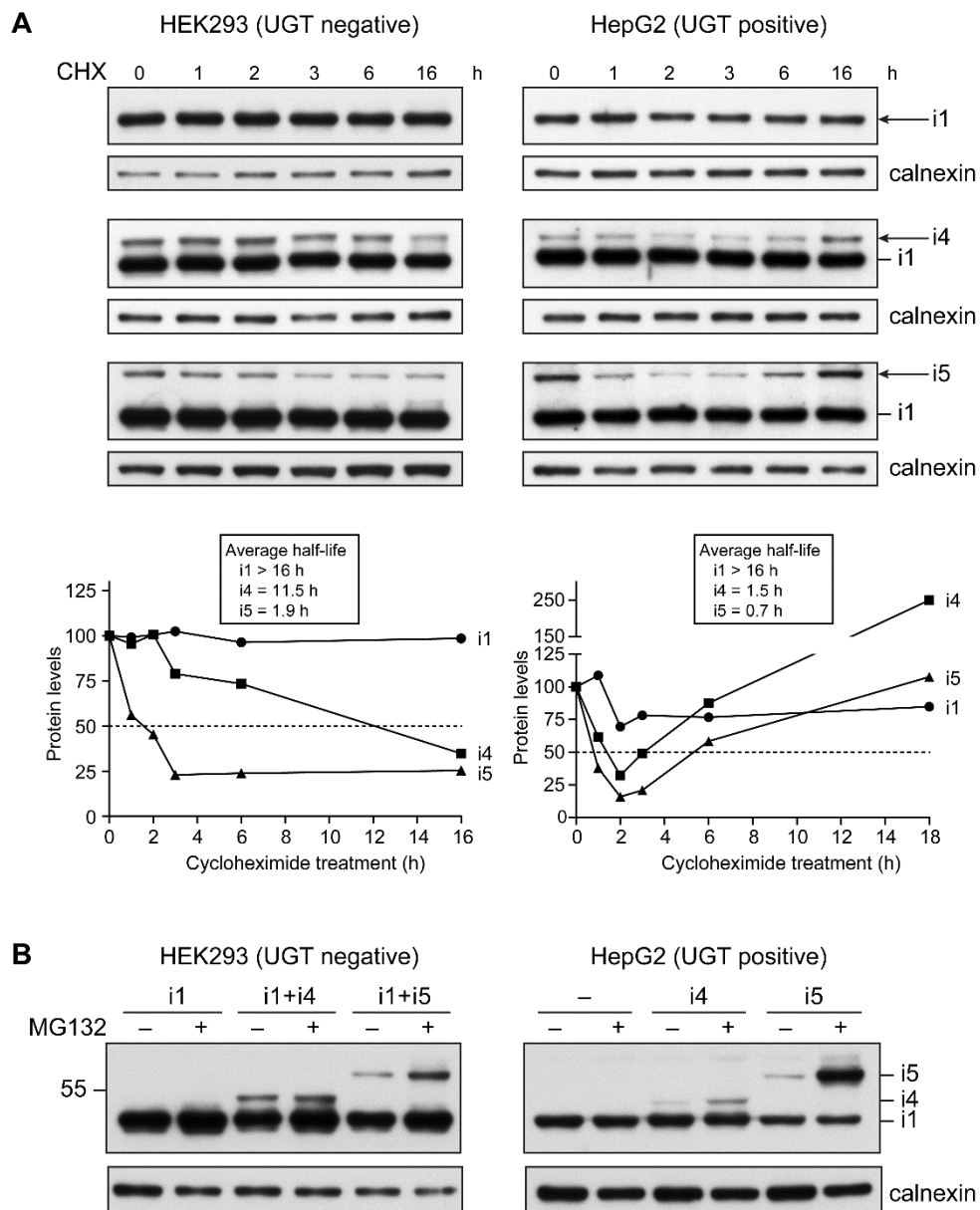
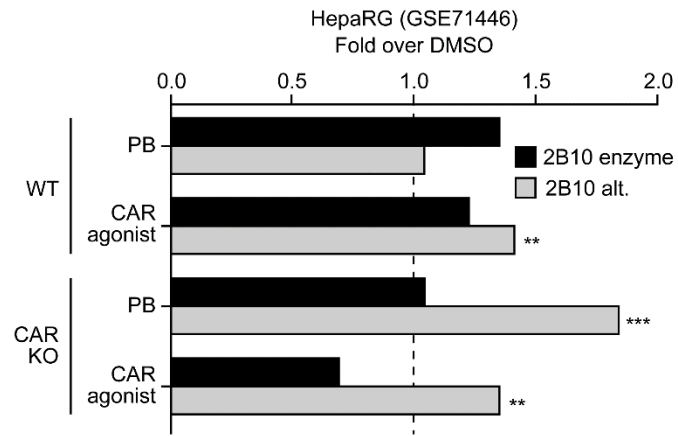


Figure 7.



Posttranscriptional regulation of UGT2B10 hepatic expression and activity by alternative splicing

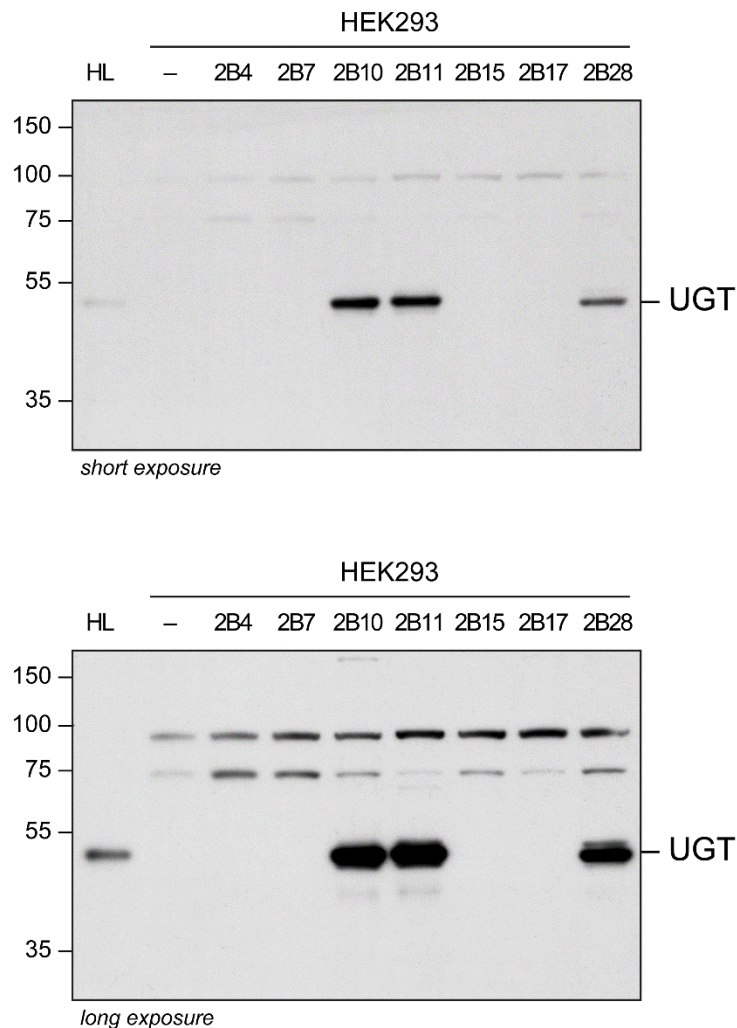
Adrien Labriet, Eric P. Allain, Michèle Rouleau, Yannick Audet-Delage, Lyne Villeneuve and Chantal Guillemette

Drug Metabolism and Disposition

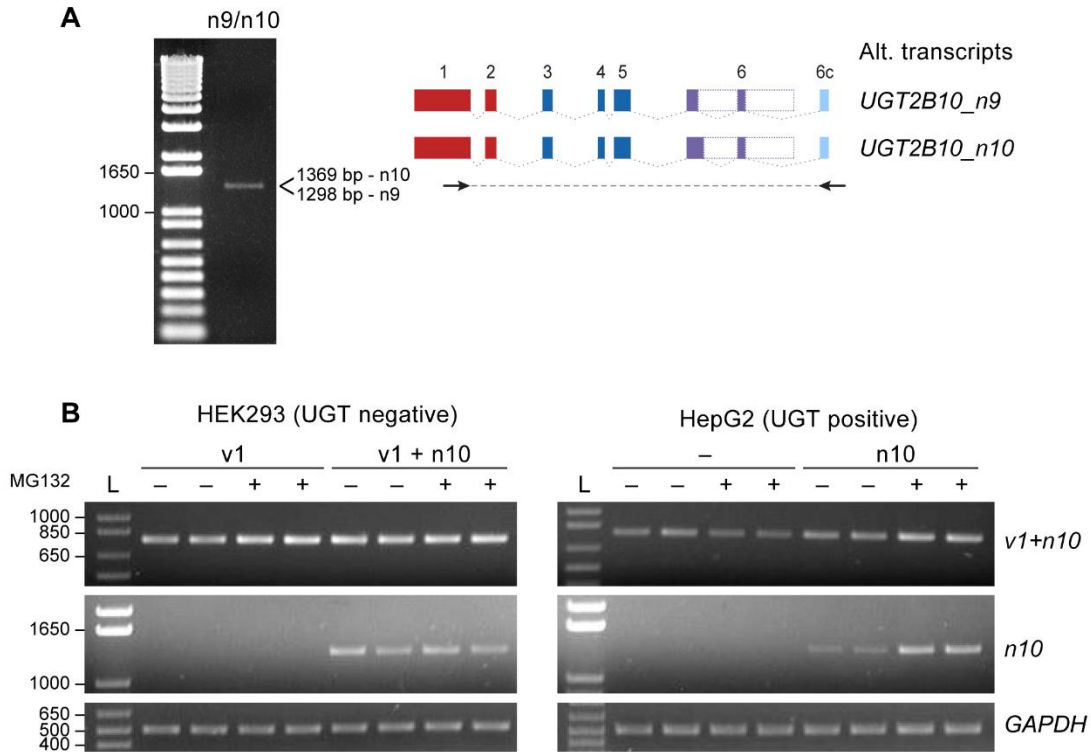
Supplemental Table and Figures

Supplemental Table 1. Primers used in this study.

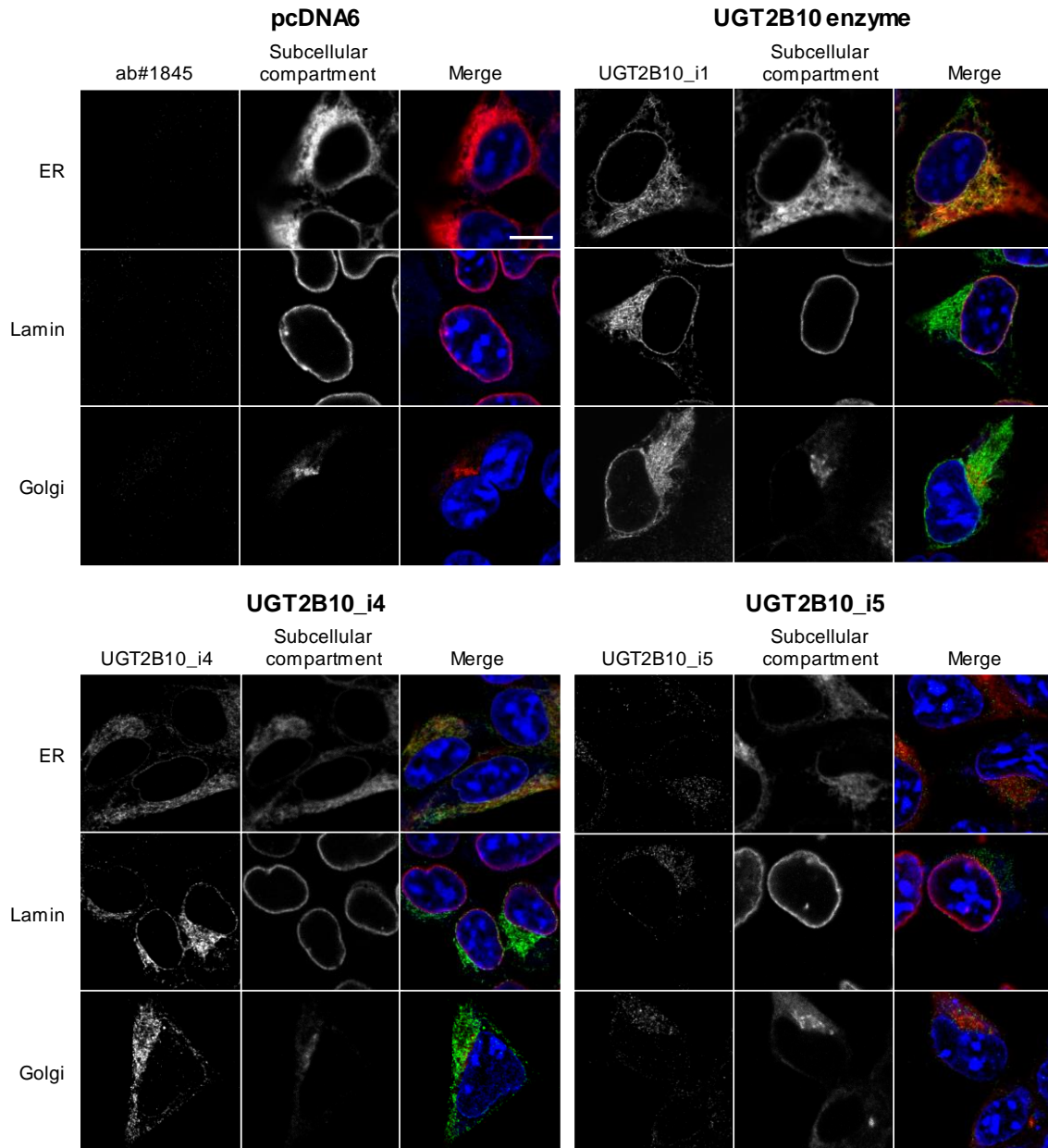
Target	Primer	Sequence (5' to 3')
<i>Cloning</i>		
UGT2B10_n9	4238F	TGTTCTCCATGCTGACTCGAGTCTAGAGGGCCC
	4239R	CAATGAGATCTCATCATGGAACCAGGTGAGGTTG
UGT2B10_n10	4240F	AACAATCCTCCCATTTTAGCATCCCAAAGGGATG AGATTACAGGTTTTGCTCGAGTCTAGAGGGCCC TGAGCCCAGGAGTTTGAAATCAGCATGGAGAACA
	4241R	CAATGAGATCTCATCAACTTTGTGATGATAAATA GCACGG
	4242F	TGGACAAATTCTTCCATTTCTACAGAGGAAAAA CTTCTTCATTATAACTCGAGTCTAGAGGGCCC GAGCTCTGATGAGGATGGTGTGTTGTGTTTTCTC
	4243R	TGGAGTCAGTTGTGCAAAACCTGTAATCTCATCC CTTTG
<i>Validation</i>		
v1, n9, n10		
Exon 1	1894F	AGATTTGACATCGTTTTTGCAGATGCTTA
Exon 4	1895R	CCTTCATGTGAGCAATATTATCAGGTTGATCAAA
n9, n10		
Exon 1	4203F	GTTAAGAGATTGTCAGAAATTCA
Exon 6c	4211R	TCTCTGGAGTCAGTTGTGC
GAPDH	1904F	TGGGTGTGAACCATGAG
	1905R	CCCAGCGTCAAAGGTGG



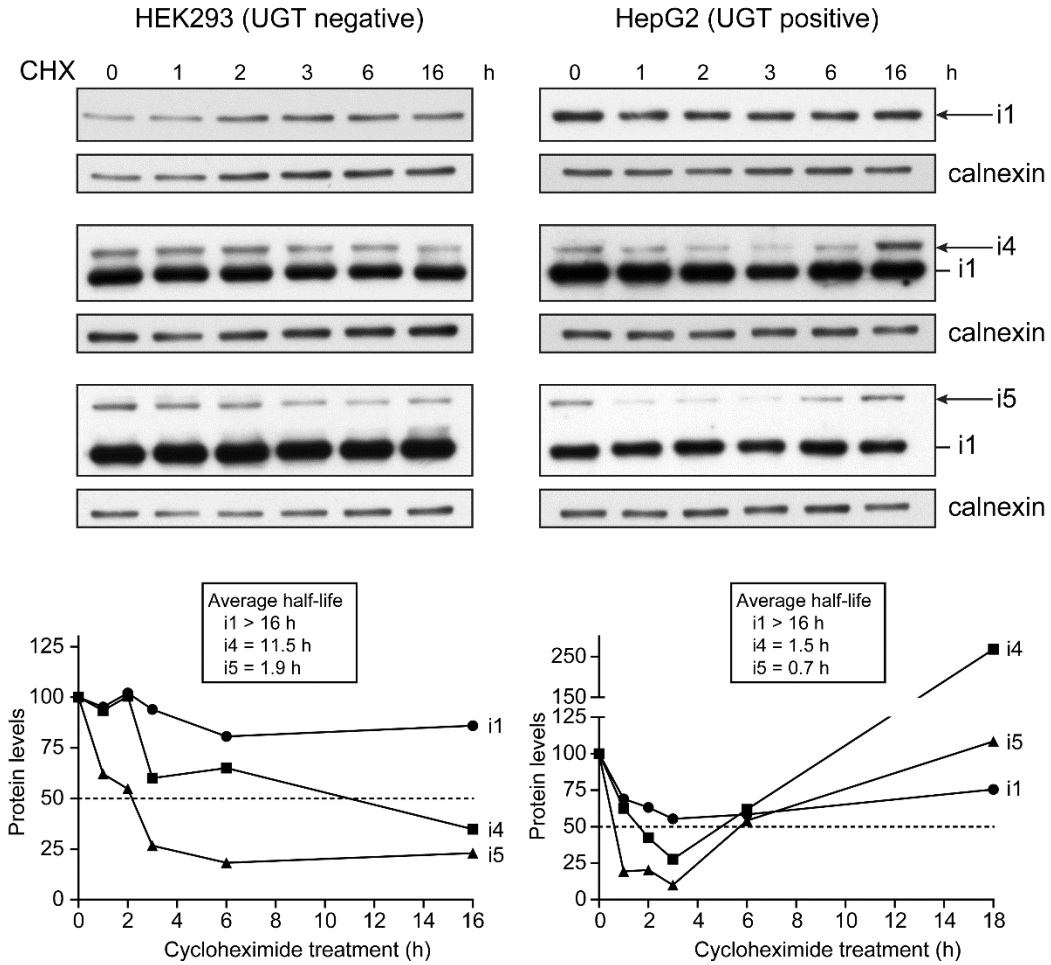
Supplemental Figure 1. Specificity of anti-UGT2B10 #1845 for human UGT enzymes. The microsomal fraction prepared from HEK293 cells (UGT negative) stably expressing each of the seven UGT2B enzymes or transfected with the empty vector pcDNA6 (-) was probed with the anti-UGT2B10 to demonstrate specificity. HL, human liver microsomes. Short and long exposures of the same immunoblot is provided to ascertain specificity.



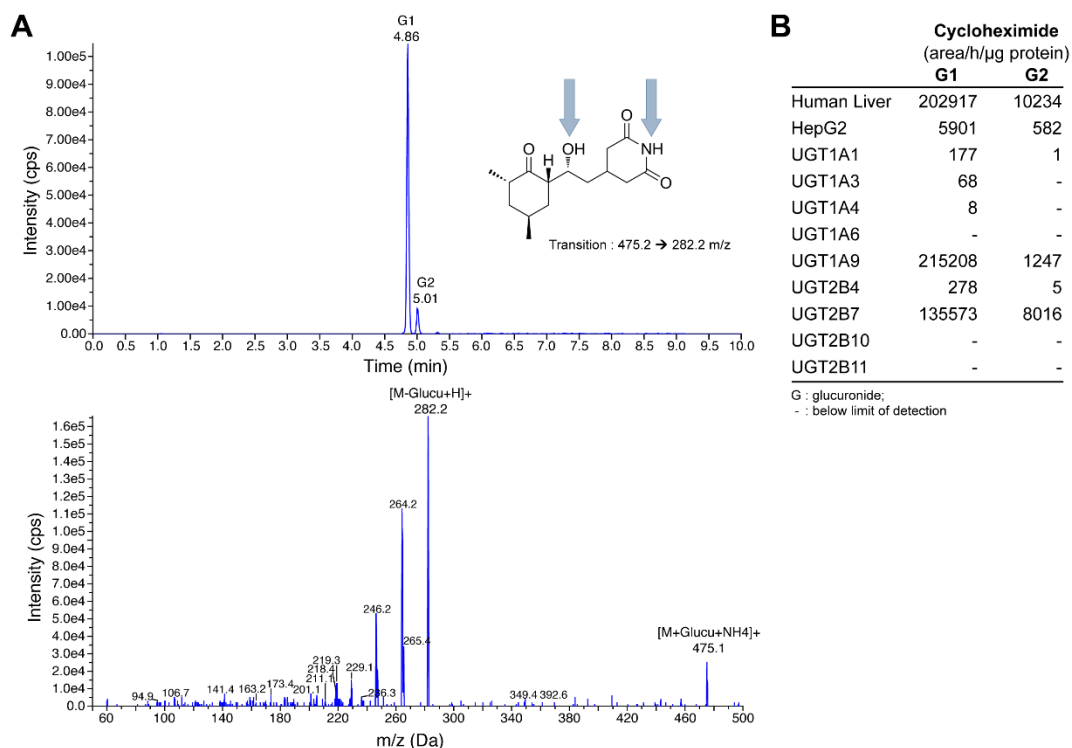
Supplemental Figure 2. RT-PCR amplification of *UGT2B10* transcripts. **A.** Validation of *UGT2B10_n9* and *UGT2B10_n10* expression in human liver. Primers (arrows; 4203F and 4211R) were located in exon 1 and in exon 6c common to both variants. PCR products corresponding to *n9* and *n10* could not be distinguished because they differed by only 71 base pairs (bp). The identity of PCR products as *n9* and *n10* was verified by Sanger sequencing. **B.** RT-PCR detection of transcripts in cell models exposed to MG132. PCR strategies specified on the right either amplified both *v1* and *n10* or only *n10*. GAPDH served as a positive and loading control. L: molecular weight ladder.



Supplemental Figure 3. Subcellular localization of UGT2B10 and alternative isoforms expressed in HEK293. UGT2B10 proteins were detected with #1845, and cellular compartments with specific markers as described in Materials and Methods. In merged images, nuclei are stained with DRAQ5. Bar represents 10 μ m. Control HEK293 cells were transfected with the empty vector pcDNA6.



Supplemental Figure 4. Stability of UGT2B10 proteins - Replicate assay. The half-life of UGT2B10 proteins were determined by translational inhibition with cycloheximide (CHX) in HEK293 and HepG2 cell models. Protein levels were measured by densitometry scanning of immunoblots from two independent assays. Half-lives were averaged from the two biological replicates. Related to Fig.6A.



Supplemental Figure 5. Analysis of cycloheximide glucuronidation by high-performance liquid chromatography tandem mass spectrometry. **A.** Human liver microsomes produced two cycloheximide glucuronides (G1 and G2). Representative chromatograms of cycloheximide glucuronides (top) and their fragmentation pattern (bottom). Putative glucuronidation sites are indicated by blue arrows in the chemical structure of cycloheximide. **B.** Glucuronidation activity of microsomes from human liver, HepG2 pcDNA6, HEK-UGT2B11 and of commercial UGT supersomes. Two independent assays were conducted in triplicate.

CLUSTAL O(1.2.4) multiple sequence alignment

```
UGT2B10_i5          510 LMRSHCVLHADFKLLGSSNPPILASQRDEITGFAQLTPEKTQQ
SP|P51790-5|CLCN3_i4      6 LGESHYVVQAGLQLLGSSNPPALASQVAEITGTHYTMNGGSI
SP|Q96EP1-3|CHFR_i3     135 EMVPCCVAQAGLKLLGSSDPPTLASQSIVITGSGGGGISPKGS
NP_001339349|ALG9_ii    401 ASRSQFVAQAGLKLLGSSNPPTSASQNIGITGLSHRAWPLTSP
XP_011508885|MITD1_iX5  150 QTGFHHVAQAGLKLLGSSSPPTLASQSAGITGMSHCTQPHYLY
          *  :*.:::*****..**  ***  ***
```

Supplemental Figure 6. Sequence similarity of the UGT2B10_i5 unique sequence with other alt. human proteins. A BLASTp search with the unique C-terminal UGT2B10_i5 amino acid sequence revealed a conserved region encoded by the frame shift in exon 6 (boxed). asterisk: conserved amino acid; colon: conservative substitution. Alignment was produced with the Clustal Omega tool available at www.Uniprot.org.

Discussion

La glucuronidation constitue une voie majeure de la phase II du métabolisme de nombreux médicaments qui sont actuellement utilisés en clinique ou en cours de développement. Elle est impliquée dans l'inactivation et l'élimination de ces molécules, et a le potentiel d'influencer l'exposition et possiblement la réponse de l'organisme à celles-ci. Une importante variabilité d'expression et d'activité de cette voie est observée entre les individus. De nombreuses sources de variabilité – génétique, épigénétique, post-transcriptionnelle et post-traductionnelle – affectent la voie des UGT. Certains variants génétiques, comme par exemple le polymorphisme *UGT1A1*28* qui réduit l'expression et l'activité d'UGT1A1, permettent d'expliquer en partie la réponse au traitement des patients. Néanmoins, la majeure partie des mécanismes à l'origine de la variabilité interindividuelle de l'activité de glucuronidation demeure inconnue. Dans le cadre de ma thèse, je me suis intéressé à l'influence 1) des polymorphismes génétiques et 2) de l'épissage alternatif sur l'activité de la voie de glucuronidation. Le premier objectif consistait à évaluer si des variations génétiques de gènes candidats pouvaient affecter la réponse au traitement. Le deuxième objectif cherchait à mieux comprendre la possible influence de l'épissage alternatif sur l'activité de glucuronidation.

Les polymorphismes génétiques de la voie de glucuronidation et des voies associées comme nouveaux marqueurs prédictifs de la réponse au traitement FOLFIRI.

Un avantage important des polymorphismes génétiques est la facilité de leur détection par séquençage de l'ADN, à un coût désormais faible. L'ADN peut être extrait à partir d'une simple prise de sang, d'une manière peu invasive. Les polymorphismes représentent donc des marqueurs de choix de la réponse à un traitement, c'est-à-dire son efficacité ou son innocuité. Dans le cadre du traitement du CCRm, le régime FOLFIRI (irinotécan + 5-FU + acide folinique) constitue toujours aujourd'hui une base de chimiothérapie, à laquelle vient s'ajouter une thérapie ciblée (Fujita et al., 2015; Modest et al., 2019; Sanchez-Gundin et al., 2018; Sandhu et al., 2019). L'irinotécan est susceptible d'engendrer des neutropénies sévères chez les patients porteurs homozygotes du polymorphisme *UGT1A1*28*, qui présentent une activité de glucuronidation moindre envers le SN-38 et sont donc plus

exposés au métabolite actif (Beutler et al., 1998; Iyer et al., 2002). Il existe différents outils approuvés pour la détection en clinique du génotype *UGT1A1**28⁴ afin d'ajuster la dose initiale d'irinotécan à administrer au patient (Dean, 2012b). Malgré son association reconnue avec le risque de neutropénies sévères de grades 3-4, le variant *UGT1A1**28 n'est pas associé à la survie des patients traités avec l'irinotécan de manière reproductible (Liu et al., 2013). Le 5-FU peut également causer des toxicités sévères chez les porteurs de variants du gène *DPYD*, codant pour l'enzyme dihydropyrimidine déshydrogénase (DPD) responsable de l'inactivation du médicament (Lee et al., 2014; Meulendijks et al., 2015; Toffoli et al., 2015). La détection du polymorphisme *DPYD**2A est désormais obligatoire au Québec⁵, et il est aussi recommandé de chercher la présence d'autres variants additionnels altérant l'activité DPD avant toute administration de fluoropyrimidines. Des recommandations de dose ont été émises en fonction des différents profils génétiques (Amstutz et al., 2018; Dean, 2012a). Les gènes des voies de la pharmacocinétique des médicaments peuvent aider à prédire le risque de toxicité, car ils modulent l'exposition à la molécule active.

Puisqu'ils régulent l'exposition au traitement, les variants des gènes de la PK de l'irinotécan ont été étudiés pour leur association avec l'efficacité du traitement. Néanmoins, la majorité de ces travaux ne se concentraient que sur un nombre restreint de gènes ou de variants étudiés seulement dans une cohorte (**Annexe A, Sections A-C**). Cette observation soutient la nécessité de conduire un criblage détaillé de ces gènes avec une étape de découverte et une étape de validation des marqueurs identifiés. En fait cette étape de validation dans une cohorte indépendante n'a été menée que pour les études les plus récentes portant sur des marqueurs de la réponse aux thérapies ciblées et publiées entre 2015 et aujourd'hui (**Annexe A – Section D**). De plus, nous avons inclus dans notre étude les variations de gènes additionnels potentiellement impliqués dans l'effet du médicament. Ils ont été sélectionnés sur la base d'études *in vivo* ou *in vitro* de la littérature supportant un effet possible sur l'exposition ou sur la réponse au traitement. Enfin, les récepteurs nucléaires et facteurs de transcription régulent une portion importante des voies PK/PD, et

⁴ <https://www.ncbi.nlm.nih.gov/gtr/>, consulté le 11 mai 2019.

⁵ https://www.inesss.qc.ca/fileadmin/doc/INESSS/Rapports/Oncologie/INESSS_DPYD_Traitements.pdf, consulté le 31 mai 2019.

pourraient donc être associés aussi bien à l'efficacité du traitement qu'au risque de toxicités induites par celui-ci. Nous avons aussi cherché à combiner les marqueurs des différents gènes pour tenter d'améliorer la prédiction du risque, ce qui a été fait en lien avec les toxicités mais plus rarement en lien avec la survie, du fait du peu de gènes inclus dans les études. Les associations entre les variants communs des différents gènes que nous avons inclus et les issues cliniques ont été testées dans une cohorte de découverte puis validées dans une cohorte indépendante. Nos deux cohortes d'étude sont indépendantes mais semblables puisqu'elles sont composées de patients atteints du CCRm traité en première intention par le régime FOLFIRI. La découverte de marqueurs prédictifs de l'efficacité du FOLFIRI est un défi important pour le traitement du CCRm car la réponse à ce traitement est variable entre les individus, mais reste encore un régime standard de chimiothérapie dans ce contexte clinique.

Une première étude portant sur les voies PK a révélé que le polymorphisme du gène *UGT1* rs1113193G>A est associé à une survie réduite des patients CCRm traités avec le régime FOLFIRI (**Chapitre 1, Figure 1**). L'outil de prédiction HaploReg⁶ localise le variant rs7608713G>A, génétiquement lié au rs1113193G>A, dans une région « *enhancer* » de l'ADN des cellules du tissu hépatique. De plus, les deux variants sont associés dans divers tissus du projet GTEx⁷ à une hausse de l'expression du gène *USP40*, localisé en amont du locus *UGT1* mais dans l'orientation antisens. Ces observations supportent un impact fonctionnel potentiel du rs1113193G>A et/ou du rs7608713G>A sur l'expression des gènes proximaux, incluant *UGT1*. Ces deux polymorphismes sont situés entre les exons 1 d'*UGT1A10* et d'*UGT1A9*, en amont d'*UGT1A1*. L'allèle rs1113193A a été liée à une diminution des niveaux sanguins de bilirubine totale, suggérant qu'elle puisse augmenter l'expression d'*UGT1A1* (Johnson et al., 2009). En effet, *UGT1A1* est la seule isoforme qui catalyse l'inactivation de la bilirubine et est la principale responsable de l'inactivation hépatique du SN-38 (Bock, 2015; Guillemette et al., 2014). Il serait aussi possible qu'un des polymorphismes augmente l'expression d'*UGT1A10*, qui est fortement exprimée dans le tractus gastrointestinal mais pas au foie, et qui est capable d'inactiver le SN-38 (Gregory

⁶ <https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>, consulté le 13 mai 2019.

⁷ <https://gtexportal.org/home/>, consulté le 13 mai 2019.

et al., 2004; Oda et al., 2014; Oguri et al., 2004; Troberg et al., 2017). L'association entre le rs1113193G>A et la survie des patients CCRm pourrait s'expliquer par une augmentation de la capacité d'inactivation du SN-38 et une exposition réduite au médicament des patients porteurs du marqueur *UGTI* rs1113193A.

Nous avons également mis en évidence le marqueur *CESI* rs9921399T>C associé à la survie prolongée des patients (**Chapitre 1, Figure 1**). La carboxylestérase 1 est impliquée dans la conversion de l'irinotécan en SN-38. Bien qu'ayant une affinité et une vélocité pour le SN-38 inférieure à la carboxylestérase 2, son expression au foie est jusqu'à 50 fois plus élevée et elle contribue donc de façon importante à l'hydrolyse de l'irinotécan (Hatfield et al., 2011; Ross et al., 2012). Une expression élevée de la carboxylestérase 1 pourrait résulter en une survie prolongée du patient en permettant une exposition plus importante au SN-38. D'ailleurs, un effet gène-dose sur la formation de SN-38 a été observé chez des patients japonais traités avec l'irinotécan (Sai et al., 2010). Cela appuie une contribution de la carboxylestérase 1 dans la réponse à l'irinotécan.

Le variant rs17501011G>A du gène *ABCC1* a été associé à une survie réduite des patients CCRm (**Chapitre 1, Figure 1**). Ce gène code pour le transporteur MRP1 qui est exprimé au foie, mais faiblement comparativement à l'intestin ou au rein (Flens et al., 1996; Hilgendorf et al., 2007; Roelofsen et al., 1997). Ce transporteur est notamment responsable de l'efflux du SN-38 vers la circulation sanguine (Chen et al., 1999). Nous avons montré que le variant rs17501011A est associé à une diminution de l'expression d'un gène rapporteur dans deux modèles de foie (HepG2 et HuH-7) et deux modèles de côlon (HCT116 et HT-29). Bien que les mécanismes moléculaires exacts restent à être étudiés, cela suggère qu'au foie, la diminution de l'expression d'*ABCC1* pourrait limiter l'efflux du SN-38 et donc favoriser sa conversion en métabolite inactif par *UGT1A1*. En utilisant deux cohortes indépendantes pour lesquelles les données d'expression génique sont accessibles publiquement, nous avons montré que les patients dont l'expression d'*ABCC1* dans les tumeurs primaires de côlon est plus faible ont une survie réduite.

Une étude sur les carcinomes épidermoïdes de la tête et du cou a montré que les patients avec une expression tumorale plus faible d'*ABCC1* avait une survie réduite (Warta et al., 2014). Il a aussi été montré qu'une surexpression tumorale pré-thérapeutique d'*ABCC1*

était associée à une meilleure réponse après chimiothérapie néoadjuvante au 5-FU/cisplatine chez des patients atteints d'un cancer de l'œsophage (Langer et al., 2005). En fait, MRP1 est aussi responsable de l'efflux de glutathion, ce qui a été associé à la fois à une plus grande sensibilité pour des cellules de cancer du côlon au 5-FU mais aussi à l'induction de l'apoptose dans des cellules cancéreuses via une augmentation du stress oxydatif (Laberge et al., 2007; Palomares et al., 2009). En somme, MRP1 pourrait donc être impliqué dans la sensibilité des cellules cancéreuses au 5-FU et dans le transport du SN-38, expliquant son association avec la survie des patients CCRm traités avec le régime FOLFIRI.

Une seconde étude portant sur les voies affectant l'action du traitement a identifié le variant rs4806668G>T du gène codant pour la protéine ribosomique RPL28 comme un marqueur de survie réduite chez les patients CCRm (**Chapitre 2, Figure 1**). Ce gène a été inclus dans notre liste de candidats car une étude *in vitro* a démontré que la répression de RPL28 entraîne une sensibilité accrue au SN-38 et au 5-FU, tous deux compris dans le régime FOLFIRI (Allen et al., 2012). Les 79 protéines ribosomiques constituent la grande (RPL) et la petite (RPS) sous-unité du ribosome et sont directement liées à la capacité de synthèse protéique des cellules⁸. Une altération de cette synthèse et du métabolisme cellulaire est fréquemment observée dans les cellules cancéreuses car leur prolifération et le développement tumoral diffèrent des tissus normaux (Lindqvist et al., 2018; White-Gilbertson et al., 2009). Cela suggère qu'une modification dans la voie de la synthèse protéique puisse modifier le développement tumoral et la réponse au traitement.

D'autre part, les protéines ribosomiques sont de plus en plus étudiées pour leurs fonctions extra-ribosomiques, sous forme libre, c'est-à-dire indépendante du ribosome. Par exemple, parmi les 79 protéines ribosomiques, plusieurs d'entre elles sont capables d'activer la voie p53 et jouent donc un rôle dans plusieurs cancers, parfois en réponse à des agents pharmacologiques tel que le cisplatine (Llanos et al., 2010; Zhou et al., 2015). On peut donc penser que la présence du variant génétique rs4806668G>T de *RPL28*, couplée au traitement de chimiothérapie puisse influencer la survie des patients. D'ailleurs, nous avons

⁸ <http://ribosome.med.miyazaki-u.ac.jp/>, consulté le 12 mai 2019.

montré que l'allèle rs4806668T est associée à une hausse de l'expression du gène dans des tissus sains de côlon. De plus, la répression de RPL28 par interférence à l'ARN dans les modèles cellulaires de cancer du côlon HCT116 et HT-29 diminuait leur capacité de prolifération (**Chapitre 2, Figures 3 et 5**). Une étude a aussi démontré que la répression de RPL28 dans les HCT116 augmentait leur sensibilité au 5-FU et SN-38 (Allen et al., 2012). Ces observations soutiennent qu'un niveau plus élevé de RPL28 pourrait être associé à une survie plus faible. Effectivement, une plus forte expression tumorale de *RPL28* est associée à une survie réduite dans deux cohortes de patients CCRm (**Chapitre 2, Figure 4**). L'ensemble de ces résultats supportent un rôle de RPL28 dans la réponse au traitement et la progression de la maladie.

En appui, une répression des voies de la matrice extracellulaire et une surexpression de gènes d'immunoglobulines sont observées dans les tumeurs primaires de patients CCRm exprimant davantage *RPL28*. Cela suggère une modulation importante de plusieurs voies cellulaires, qui pourrait expliquer la différence de survie entre les patients. Un remodelage de la matrice extracellulaire ainsi que l'activité du complément et de l'inflammation sont des mécanismes connus dans la progression du cancer (Crotti et al., 2017; Del Rio et al., 2013; Janakiram et al., 2014; Reis et al., 2018). L'étape suivante serait de comprendre comment la hausse de l'expression de *RPL28* affecte la survie et son lien avec les autres voies altérées. Une étude approfondie de l'impact de ces voies dans la maladie pourrait permettre d'identifier des cibles pour améliorer la réponse des patients.

Au-delà des voies PK/PD, notre troisième étude portait sur le facteur de transcription HNF1A. Nous avons montré que le variant rs2244608A>G du facteur de transcription HNF1A est associé à la survie des patients CCRm traités avec le régime FOLFIRI (**Chapitre 3, Tableau 2**). Agissant comme régulateurs de l'expression des voies PK/PD de nombreux médicaments, les facteurs de transcription et récepteurs nucléaires apparaissent comme des marqueurs potentiels de la prédiction de la réponse au médicament (Cecchin et al., 2016; De Mattia et al., 2016; Mbatchi et al., 2018). Le métabolisme de l'irinotécan est principalement hépatique, et HNF1A est un régulateur clé de l'expression hépatique de nombreux gènes (Odom et al., 2004). D'ailleurs, l'expression hépatique de gènes de la pharmacologie de l'irinotécan (*CYP3A5*, *ABCC1* et *ABCC2*) était altérée chez des individus

sains porteurs du variant *HNF1A* rs2244608A>G (**Chapitre 3, Figure 4**). Nous avons démontré que les patients CCRm porteurs du marqueur de *HNF1A* étaient davantage exposés au SN-38 (**Chapitre 3, Figure 1**). Cela suggère que les variants de régulateurs transcriptionnels pourraient modifier la réponse au médicament en affectant l'expression de multiples gènes des voies PK/PD de ce médicament, et potentiellement permettre de prédire la survie et/ou les toxicités.

Deux études réalisées en collaboration avec le groupe de recherche du Dr Toffoli sur les deux cohortes CCRm viennent appuyer cette idée. Dans la première, nous avons observé que le variant rs11574077A>G du récepteur nucléaire *VDR* était associé à une diminution de l'efficacité de l'élimination du SN-38 et une augmentation des toxicités gastrointestinales induites par ce traitement (De Mattia et al., 2018). Comme *HNF1A* au niveau hépatique, le récepteur *VDR* est effectivement impliqué dans la régulation de plusieurs gènes de la pharmacologie de l'irinotécan incluant notamment des transporteurs ABC et SLC, surtout au niveau de l'intestin (Prakash et al., 2015). Une étude précédente avait associé un variant de *PXR* au risque de toxicités hématologiques et à une dégradation accrue du SN-38 chez des patients atteints d'un CCRm traité à base d'irinotécan (Mbatchi et al., 2016). Plus récemment, une seconde étude a démontré que le variant *NR1I2* rs1054190C>T est associé à une survie réduite alors que le variant *VDR* rs7299460C>T prédit une survie plus longue (De Mattia et al., 2019). Ces résultats supportent le fait que les récepteurs nucléaires puissent être impliqués à la fois dans l'efficacité du traitement et les effets indésirables qu'il peut induire. Il pourrait aussi être pertinent d'investiguer le potentiel prédictif des variants des régulateurs transcriptionnels que nous avons découverts sur la réponse à d'autres molécules prises en charge par les mêmes voies PK que l'irinotécan.

Nos différentes études soutiennent également le fait que la combinaison de plusieurs marqueurs permettrait une meilleure prédiction de la réponse au médicament. Dans le cas du régime FOLFIRI, nous avons montré que la combinaison de plusieurs marqueurs des voies PK/PD associés individuellement à la réponse (*CESI* rs9921399T>C, *UGT1* rs1113193G>A, *ABCC1* rs17501011G>A et *RPL28* rs4806668G>T) présentait un effet combiné sur la survie globale (**Chapitre 1, Figure 3**). Ce type d'analyse permet de mieux

définir des profils de risque qu'avec l'utilisation d'un seul marqueur mais nécessite de disposer de cohortes de tailles suffisantes. Dans le cas du régime FOLFIRI, cette approche a déjà fait ses preuves puisque la combinaison de marqueurs du locus *UGT1* et de gènes de transporteurs *ABC* permet d'améliorer la prédiction du risque de neutropénies sévères comparativement à l'utilisation du seul marqueur *UGT1A1*28* (Chen et al., 2015a; Chen et al., 2015b; Levesque et al., 2013). L'intérêt des multi-marqueurs a aussi été démontré dans le cadre du traitement avec la warfarine, pour laquelle 40% de la variation de la dose initiale requise peut être déterminée grâce à la variabilité des deux gènes *CYP2C9* et *VKORC1* (Dean, 2012c; Wadelius et al., 2005). La FDA a inclus dans la monographie du médicament des recommandations de dosage selon les génotypes de *VKORC1* et *CYP2C9* et les consortiums DPWG et CPIC ont également émis des lignes directrices (Dean, 2012c; Johnson et al., 2017). Il existe plusieurs tests génétiques qui permettent de déterminer simultanément les génotypes *VKORC1* et *CYP2C9* et ainsi ajuster la dose de warfarine selon le profil génétique et clinique du patient. Il serait donc pertinent d'inclure dans un même test génétique les différents variants liés à l'efficacité du traitement à base d'irinotécan, et ceux liés aux toxicités décrits ci-dessus, incluant le polymorphisme *UGT1A1*28*.

Nos investigations ont permis d'associer l'expression des gènes *ABCC1* et *RPL28* dans les tumeurs primaires à la survie des patients CCRm dans deux cohortes indépendantes (**Chapitre 1, Figure 2 et Chapitre 2, Figure 4**). Cependant, contrairement aux polymorphismes génétiques, l'analyse de l'expression génique tumorale est plus contraignante car elle requiert l'extraction d'ARN à partir d'une biopsie de la tumeur. De plus, ce prélèvement localisé ne permet pas de tenir compte ni de l'hétérogénéité tumorale ni de son évolution au cours du temps après la biopsie (Gerlinger et al., 2012). Pour les cancers gastrointestinaux et particulièrement le cancer colorectal, il a été montré que les biopsies liquides représentaient une alternative prometteuse dans la recherche de biomarqueurs (Pantel et al., 2017). Cette technique permet d'isoler des CTC (cellules tumorales circulantes) à partir d'une prise de sang. Il est ainsi possible de mesurer l'expression des gènes d'intérêt dans ce type de cellules tumorales ou d'identifier des mutations somatiques (Alix-Panabieres et al., 2016; Lianidou, 2016). Ainsi, la mesure de

l'expression de gènes associés à la survie dans ces cellules pourrait permettre d'améliorer la prédiction de la réponse des patients.

Les marqueurs génétiques que nous avons identifiés ont des fréquences différentes dans les populations à l'étude et entre les groupes ethniques. Le marqueur *HNF1A* rs2244608G a une fréquence d'un peu moins de 35% dans nos cohortes d'origine caucasienne mais seulement de 11% dans les populations d'origine africaine. Les porteurs homozygotes de l'allèle rs4806668T et les porteurs de l'allèle rs17501011A représentent respectivement 3% et 9% des individus de nos cohortes à l'étude. Ces fréquences sont similaires à celles observées dans les populations d'origine caucasienne. Il serait pertinent d'investiguer la valeur prédictive de ces marqueurs dans des populations où leur fréquence génotypique est supérieure aux populations d'origine caucasienne. Par exemple la fréquence du génotype *RPL28* rs4806668TT est de 58% dans les populations d'origine africaine, celle d'*UGT1* rs1113193AA y est de 17% alors que 23 à 34% des personnes d'origine asiatique sont porteuses homozygotes de l'allèle rs9921399C de *CESI*. Selon leur fréquence dans les divers groupes ethniques, les marqueurs associés à la survie pourraient permettre d'identifier des sous-groupes de patients quantitativement plus ou moins importants et donc des populations qui sont susceptibles de mieux répondre au FOLFIRI.

Notre étude a identifié plusieurs marqueurs génétiques associés à la survie des patients traités avec le FOLFIRI démontrant aussi que l'approche par gènes candidats est valide dans l'optique de découvrir de nouveaux marqueurs prédictifs de la réponse au traitement. D'ailleurs, bien que les mutations somatiques (par exemple *KRAS* et *BRAF*) aient été les premiers marqueurs prédictifs de l'efficacité des thérapies ciblées nouvellement utilisées dans le traitement du CCRm (bévacicumab, cétuximab, panitumumab), de récentes études démontrent la validité de variations génétiques héréditaires à prédire l'efficacité de ces thérapies (Berger et al., 2018; Matsusaka et al., 2017; Matsusaka et al., 2016; Tokunaga et al., 2019a; Tokunaga et al., 2019b). Ces études sont basées sur un nombre restreint de gènes candidats qui ont été sélectionnés *a priori* car ils sont impliqués dans les voies de la prise en charge par la cellule de ces molécules et de la réponse thérapeutique. Ces dernières études incluent de façon similaire à nos études une cohorte de découverte et une de réplication, confirmant la pertinence de valider les marqueurs identifiés.

Néanmoins, les études pangénomiques pourraient permettre d'identifier des variants additionnels prédictifs de la réponse dans des gènes qui ne sont pas *a priori* liés à la PK/PD du traitement. Une des conditions essentielles serait de disposer d'une cohorte de taille suffisante pour avoir une puissance statistique satisfaisante. En fait, à ce jour, aucune étude pangénomique cherchant à lier variants génétiques et efficacité ou innocuité du régime FOLFIRI n'a été publiée. Malgré tout, quatre études pangénomiques visant à identifier un lien entre variants génétiques et réponse à l'irinotécan ont été effectuées sur des cohortes de tailles modestes. Pour deux d'entre-elles, les auteurs ont mis en place des stratégies multi-étapes, statistiques ou expérimentales, pour découvrir des marqueurs prédictifs de l'efficacité ou du risque de toxicités et tenter d'outrepasser les limitations statistiques inhérentes à une étude à l'échelle du génome entier (Kim et al., 2010; Takahashi et al., 2014). Deux autres études, conduites par le même groupe, ont été menées en lien avec les toxicités induites par l'irinotécan chez 101 patients atteints d'un cancer du poumon non à petites cellules avancé ou la réponse au traitement chez 139 patients atteints d'un cancer du poumon à petites cellules (Han et al., 2014; Han et al., 2013). Cependant, les auteurs soulignent le fait que la puissance statistique de ces études est faible et représente une limitation majeure, nécessitant donc des validations futures. En fait, ces résultats supportent davantage la pertinence de notre choix de l'approche par gènes candidats.

Les points forts de nos travaux sont 1) un criblage exhaustif des variations communes (> 5%) des voies PK/PD de la réponse au traitement ; 2) la validation des découvertes dans une seconde cohorte indépendante ; 3) des investigations *in vitro* ou à partir de jeux de données publiques pour explorer les mécanismes sous-jacents. Cependant, certaines limites sont aussi à considérer car l'approche par gènes candidats a certainement exclu des gènes associés à la survie et à la toxicité. De même, la réplication des résultats de façon prospective est un prérequis indispensable afin de valider les marqueurs pour permettre leur intégration possible en clinique. Enfin, le régime FOLFIRI n'est plus administré seul de nos jours mais il constitue toujours une base de chimiothérapie standard couplée à des anticorps dont les cibles sont différentes du FOLFIRI. Il est donc toujours pertinent de découvrir des marqueurs de la réponse au FOLFIRI. Dans une future étude de validation, il serait intéressant d'intégrer aussi les marqueurs prédictifs de la réponse aux thérapies ciblées dont les mécanismes de prise en charge et d'action sont différents de ceux de la chimiothérapie.

Nos travaux ont montré pour la première fois que la combinaison de marqueurs des voies PK et PD, situés dans les gènes *UGT1*, *CES1*, *ABCC1* et *RPL28* pouvait prédire l'efficacité du traitement par le régime FOLFIRI de patients atteints du CCRm. Nous avons aussi démontré que certains variants de gènes codant pour des régulateurs transcriptionnels, tels que *HNF1A*, *VDR* ou *PXR* sont de potentiels biomarqueurs de l'efficacité du traitement alors que d'autres sont associés aux toxicités du traitement.

L'épissage alternatif est un mécanisme qui pourrait altérer le métabolisme et la réponse au traitement.

Les gènes de la voie de glucuronidation sont tous soumis à l'épissage alternatif (Tourancheau et al., 2016). Parmi les 130 variants d'épissage récemment découverts, certains produisent des protéines tronquées en N-terminal ou C-terminal ou pouvant posséder de nouvelles séquences codantes. Par exemple, les *UGT1A_i2* sont tronquées en C-terminal et possèdent une séquence spécifique de 10 acides aminés (Girard et al., 2007; Levesque et al., 2007b). La protéine *UGT2B7_i8* possède une nouvelle séquence interne entre les domaines de liaison au substrat et au cosubstrat (Rouleau et al., 2016). Les travaux présentés dans cette thèse se sont concentrés sur les protéines alternatives *UGT2B10_i4* et *i5* qui sont tronquées en C-terminal et possèdent de nouvelles séquences. Le transcriptome hépatique d'*UGT2B10* semble particulièrement intéressant à étudier car il est constitué à hauteur d'environ 50% de variants alternatifs dans le tissu hépatique humain normal et cancéreux. Les transcrits *UGT2B10_n9* et *UGT2B10_n10*, possédant de nouvelles séquences codantes, découlent d'évènements d'exclusion partielle de l'exon 6 canonique et de l'inclusion du nouvel exon 6c. Le variant *n9* code pour la protéine *UGT2B10_i4* et le variant *n10* pour *UGT2B10_i5*. De plus, la séquence d'acides aminés unique à *UGT2B10_i5* se rapproche de celles retrouvées dans diverses protéines alternatives prédites, incluant des variants du transporteur *CLCN3*, de la mannosyltransférase *ALG9* et de l'ubiquitine ligase *E3* (**Chapitre 4, Figure supplémentaire 6**). Ces transcrits *UGT2B10* représentent 15% du transcriptome hépatique normal et 25% du transcriptome hépatique cancéreux. Ces chiffres soulignent l'abondance des transcrits alternatifs générés par le gène *UGT2B10* (**Chapitre 4, Figures 1 et 2**). C'est aussi le cas de nombreux gènes impliqués dans les voies pharmacologiques, avec 965 variants rapportés à ce jour qui proviennent des

57 gènes *CYP*, dont pas moins de 281 pour le seul locus *CYP2* (Annalora et al., 2017). Cette abondance, quantitative et qualitative, suggère un rôle biologique fonctionnel pour au moins une partie de ces variants d'épissage.

Nous avons confirmé l'expression protéique de l'enzyme UGT2B10 et de la protéine UGT2B10_i5 au niveau du foie, soutenant une fonction biologique possible de cette protéine (**Chapitre 4, Figure 3**). De la même façon, plus de la moitié des transcrits alternatifs *CYP* sont prédits comme menant à l'obtention de protéines (Annalora et al., 2017). Les protéines UGT2B10_i4 et i5 ont par la suite été exprimées dans deux modèles cellulaires, exprimant de façon endogène (HepG2) ou pas (HEK293) l'enzyme UGT2B10. Des essais enzymatiques utilisant deux substrats d'UGT2B10, l'amitriptyline et la lévométopromidol, à deux concentrations différentes n'ont pas permis de démontrer une activité transférase des protéines alternatives. En revanche, dans les systèmes de coexpression avec l'enzyme canonique, l'expression des variants d'UGT2B10 altère fortement la capacité de glucuronidation (**Chapitre 4, Figure 5A et 5B**). Ces phénomènes sont semblables à ceux observés pour d'autres variants UGT, notamment pour les UGT1A_i2 (tronquées en C-terminal) et pour UGT2B7_i8 (nouvelle séquence interne), qui n'ont pas d'activité intrinsèque connue à ce jour mais qui agissent comme modulateurs de la glucuronidation (Bellemare et al., 2010b; Girard et al., 2007; Levesque et al., 2007b; Rouleau et al., 2016). Nos résultats ainsi que les précédentes études soutiennent que les protéines UGT résultant de l'épissage alternatif jouent un rôle de régulateur de la capacité de détoxification dans les cellules.

Nous avons aussi voulu comprendre par quels mécanismes les protéines alternatives modifient l'activité enzymatique. Nous avons montré que l'enzyme, tout comme les isoformes i4 et i5, colocalisent au niveau du réticulum endoplasmique (**Chapitre 4, Figure 3B**). Des expériences de co-immunoprécipitation supportent une interaction entre l'enzyme et les variants i4 et i5, suggérant que la modulation de l'activité transférase puisse impliquer une interaction entre ces isoformes. Ces mécanismes sont similaires à ceux observés pour les enzymes UGT1A et les variants tronqués i2 qui sont capables d'interaction (Bellemare et al., 2010a; Bellemare et al., 2010b; Rouleau et al., 2016). De plus, l'intensité de la modulation de la glucuronidation par les UGT1A semble directement

dépendante du niveau d'expression des isoformes alternatives tronquées i2 (Rouleau et al., 2013b). Une étude menée sur les UGT1A1, 1A7 et 1A8 a démontré que l'expression des protéines alternatives i2 altère la vitesse de formation des glucuronides, c'est-à-dire la vélocité de l'enzyme, mais pas l'affinité pour les substrats testés ni le cosubstrat UDPGlcA (Bellemare et al., 2010b). Cependant, dans le cas d'UGT2B10, le mécanisme moléculaire par lequel les isoformes alternatives affectent la glucuronidation reste à être élucidé.

Les mécanismes de régulation des différentes isoformes pourraient nous aiguiller sur les similitudes et divergences potentielles de leurs fonctions. *In vitro*, l'enzyme UGT2B10 est plus stable que les protéines alternatives mais les différences observées entre les deux modèles cellulaires laissent penser que la stabilité des isoformes est variable selon le contexte cellulaire (**Chapitre 4, Figure 6A**). L'inhibition pharmacologique du protéasome n'affecte pas les niveaux d'enzyme mais augmente considérablement ceux des protéines alternatives, dont la dégradation par ce mécanisme cellulaire semble ainsi leur être propre (**Chapitre 4, Figure 6B**). Enfin, dans le modèle hépatique HepaRG, nous avons observé une induction significative des variants alternatifs n9 et n10 en réponse aux traitements pharmacologiques avec le phénobarbital et un agoniste du récepteur CAR, alors que le transcrite canonique n'était pas significativement induit par ces traitements (**Chapitre 4, Figure 7**). En fait, la stabilité réduite des protéines alternatives et leur induction préférentielle lors d'un traitement pharmacologique soutiennent l'idée qu'elles puissent agir comme de fins et rapides régulateurs de la capacité de glucuronidation dans les tissus. Il serait aussi pertinent d'étudier leur régulation dans un contexte d'inhibition thérapeutique du protéasome, qui semble être impliqué dans la régulation des niveaux de protéines alternatives UGT2B10_i4 et i5. En effet, le bortézomib et l'ixazomib sont deux inhibiteurs du protéasome utilisés dans le cadre du traitement contre le myélome multiple (Moreau et al., 2016; Richardson et al., 2005). S'ils stabilisent effectivement les isoformes variantes chez les patients traités avec ces médicaments, cela pourrait aussi avoir des conséquences sur le métabolisme des dérivés du tabac ou de médicaments pris en charge par UGT2B10, tels que les antihistaminiques, les antidépresseurs ou les antipsychotiques (Guillemette et al., 2014; Kato et al., 2013).

L'épissage alternatif est un mécanisme dynamique, dont l'expression est variable en réponse à de multiples stimuli (Pleiss et al., 2007; Younis et al., 2010). Il est donc pertinent de se questionner sur les changements métaboliques que pourraient induire les variants UGT, à la fois sur la glucuronidation mais aussi de façon plus étendue. Une possibilité serait d'avoir recours à l'établissement d'une signature métabolomique caractéristique des variants. Par exemple, il serait possible d'extrapoler la capacité de glucuronidation d'une UGT à partir d'une prise de sang, en fonction de métabolites dont les niveaux circulants auront préalablement été démontrés comme étant affectés par le niveau d'activité de cette enzyme. Au-delà de l'impact du niveau d'expression des protéines alternatives sur la capacité de glucuronidation, cette approche pourrait aussi permettre de tenir compte de multiples facteurs altérant l'activité. Dans le cas d'UGT2B10, cette approche permettrait par exemple de tenir compte de l'impact des polymorphismes rs61750900 et rs2942857 connus pour réduire drastiquement l'activité d'UGT2B10 envers ses substrats (Chen et al., 2007; Fowler et al., 2015; Murphy et al., 2014). En fait, l'approche métabolomique nous permettrait d'évaluer la capacité d'un individu à métaboliser des substrats d'UGT2B10 en tenant compte de tous les facteurs qui peuvent affecter l'activité de l'enzyme, incluant l'épissage alternatif.

L'utilisation de la métabolomique pour prédire la réponse à des molécules prises en charge par les UGT apparaît donc comme une future application d'intérêt. En fait, cela n'a jamais été fait pour aucune UGT mais la pharmaco-métabolomique, bien que naissante, est une composante prometteuse de la médecine de précision. Elle vise à mesurer simultanément l'impact de variations dans le génome, le transcriptome et le protéome, qui se reflète sur le métabolome et à l'associer à la réponse à un traitement (Beger et al., 2016). Une étude *in vitro* portant sur la sensibilité de cellules de mésothéliome pleural malin aux antifolates, qui inhibent les enzymes de la synthèse des purines et des pyrimidines, a montré une hausse de la choline et une baisse de la glycine dans les cellules sensibles au traitement par le pémétréxed (Sato et al., 2018). Chez l'homme, une étude portant sur la réponse à la chimio-radiothérapie néoadjuvante (5-FU + cisplatine + radiothérapie) chez 26 patients traités pour un carcinome épidermoïde de l'œsophage démontre l'applicabilité de la méthode (Fujigaki et al., 2018). Parmi 322 métabolites mesurés dans le sang des patients, 19 étaient diminués dans le groupe des patients ayant une réponse pathologique complète et pourraient

éventuellement représenter une signature prédictive de la réponse au traitement. Dans le cas du variant d'épissage AR-V7, impliqué dans la résistance au traitement du cancer de la prostate, il a été montré que son expression dans la lignée cellulaire de cancer de la prostate LNCaP induisait de forts changements du métabolome (Shafi et al., 2015). Plusieurs métabolites sont altérés spécifiquement en réponse à l'expression du AR-V7 comme c'est le cas pour le citrate qui est fortement diminué. Cette observation se transpose *in vivo* puisque les auteurs ont observé des niveaux similaires de citrate dans des tissus prostatiques bénins et de tumeurs primaires, qui sont fortement diminués dans les tumeurs métastatiques résistantes à la castration. La métabolomique, conduite parallèlement à des analyses génomiques et transcriptomiques, pourrait donc permettre d'identifier des voies impliquées dans la réponse au traitement, la progression de la maladie ou qui seraient de nouvelles cibles thérapeutiques.

Au-delà de la glucuronidation, les travaux du laboratoire ont montré que le niveau d'expression de variants d'épissage alternatif affecte largement le métabolome cellulaire *in vitro*. La lignée cellulaire HT115, un modèle de cancer du côlon, exprime de façon endogène les transcrits canoniques codant pour les UGT1A_i1 et les variants alternatifs tronqués codant pour les UGT1A_i2 selon une proportion de 1 pour 0.8 (Bellemare et al., 2010c; Rouleau et al., 2014). La perturbation de cet équilibre par répression stable de l'expression des i2 conduit à une altération du niveau de 58 métabolites (Audet-Delage et al., 2017). Notamment, une baisse de plusieurs acides aminés et leurs dérivés a été observée, en parallèle d'une hausse du niveau de plusieurs sucres tels que le glucose, fructose ou xylose. Cela permet donc de définir une signature témoignant de l'expression *in vitro* des isoformes alternatives i2. D'une façon similaire, l'expression de la protéine alternative UGT2B7_i8 dans les lignées HEK293 et HepG2 entraîne une modification du métabolome de ces cellules (Rouleau et al., 2016). Dans les deux lignées, une augmentation générale des acides aminés et notamment de la glutamine a été observée. Une réduction des purines et pyrimidines a été mesurée spécifiquement dans les HEK293 exprimant la protéine alternative, alors que plusieurs molécules de la glycolyse et du cycle de Krebs sont diminuées dans les HepG2. Ainsi, il serait intéressant de définir une signature métabolomique reflétant l'expression des variants alternatifs UGT2B10 dans nos modèles cellulaires. En fonction des métabolites altérés, cela pourrait nous éclairer sur d'éventuels

nouveaux substrats ou cosubstrats des différentes isoformes, ou les voies métaboliques sur lesquelles les protéines alternatives pourraient avoir un impact.

Puisque les variants alternatifs UGT semblent avoir la capacité de modifier le métabolome cellulaire, leurs rôles dans la perturbation du métabolisme cellulaire et sur le risque d'apparition d'une condition pathologique semblent des avenues d'étude pertinentes. En effet les perturbations rapportées en lien avec les variants tronqués UGT1A_i2 et la protéine UGT2B7_i8 possédant une nouvelle séquence interne affectent des métabolites de voies majeures du métabolisme cellulaire que sont la glycolyse, le cycle de Krebs et la synthèse d'acides aminés (Audet-Delage et al., 2017; Rouleau et al., 2016). D'ailleurs, ces mêmes études ont montré des changements phénotypiques associés à la surexpression des UGT1A_i2 dans les HT115 et d'UGT2B7_i8 dans les HEK293, notamment en lien avec l'adhésion et la prolifération cellulaire. Or, ces deux phénotypes sont altérés dans le cancer, renforçant l'idée que l'épissage alternatif des UGT puisse jouer un rôle dans la transformation néoplasique et l'évolution de la maladie.

Les UGT ont été associées au risque et au pronostic de plusieurs cancers, de par leur rôle dans le métabolisme de divers carcinogènes et molécules endogènes telles que les hormones (Guillemette et al., 2014). L'épissage alternatif joue un rôle dans l'apparition et le développement du cancer et dans la réponse au traitement (Climente-Gonzalez et al., 2017; Oltean et al., 2014; Siegfried et al., 2018). L'enzyme UGT2B10 est responsable de l'inactivation et l'élimination de composés et carcinogènes du tabac impliqués dans le cancer du poumon, tels que la nicotine, la cotinine et les nitrosamines (Chen et al., 2007; Chen et al., 2008; Murphy et al., 2018). Les protéines alternatives, qui affectent l'activité de l'enzyme, pourraient donc aussi modifier la prédisposition et le développement du cancer.

Plus de 130 variants d'épissage ont été identifiés par le laboratoire. Si nous voulons étudier plus spécifiquement l'impact d'un seul variant d'épissage ou comprendre les mécanismes régulant sa transcription, sa traduction et sa dégradation, il est nécessaire d'établir une méthode de priorisation des variants les plus susceptibles d'avoir un rôle fonctionnel. Les études sur les protéines alternatives d'UGT2B10 et d'UGT2B7 possédant de nouvelles séquences ont souligné l'impact fonctionnel de ces variants. Néanmoins, les variants tronqués comme les UGT1A_i2 ne sont pas à négliger car leur impact fonctionnel semble

aussi important. D'ailleurs d'autres variants tronqués possèdent une fonction biologique, comme par exemple le récepteur aux androgènes AR-V7 qui perd le domaine de liaison aux hormones et est donc constitutivement actif, avec les conséquences discutées précédemment (Antonarakis et al., 2014; Hu et al., 2009). De façon semblable, le variant *CYP2C8_v3* tronqué en N-terminal code pour une protéine qui est majoritairement localisée à la mitochondrie alors que l'enzyme canonique est plutôt située au réticulum endoplasmique (Bajpai et al., 2014). La protéine alternative est impliquée dans le métabolisme de l'acide arachidonique et semble impacter les niveaux d'espèces réactives de l'oxygène et la respiration mitochondriale. Ainsi, il apparaît judicieux de choisir les variants prioritaires selon leur capacité à produire une protéine.

Les variants alternatifs pourraient être testés selon leur potentiel à produire une protéine alternative, qui sera ensuite exprimée de façon stable dans des modèles cellulaires et investiguée quant à son impact sur la glucuronidation et le métabolisme cellulaire. L'étude de la stabilité et de la régulation des transcrits pourrait être effectuée par des approches de séquençage à haut débit (Chan et al., 2018; Lugowski et al., 2018). La capacité des différents transcrits à produire une protéine pourrait être évaluée grâce à une adaptation de la méthode VAMP-seq (*Variant Abundance by Massively Parallel Sequencing*) récemment décrite (Matreyek et al., 2018). Brièvement, la technique initiale consiste à inclure dans un plasmide toutes les combinaisons possibles d'ARN codant pour une protéine, en substituant chacun des acides aminés possibles, en amont d'une protéine fluorescente. Après transfection, les cellules sont triées par niveau de fluorescence (par exemple faible, moyen, fort) puis séquencées pour déterminer quels variants sont exprimés. On peut donc en déduire l'abondance de chaque protéine alternative à partir d'un pool de cellules transfectées avec toutes les combinaisons de variants possibles. Une fois ce criblage initial effectué, les isoformes protéiques effectivement produites pourraient alors être étudiées dans des modèles cellulaires transfectés ou exprimant de façon endogène les transcrits alternatifs d'intérêt comme c'est le cas pour les HT-115 avec les isoformes tronquées d'UGT1A ou les HepaRG qui expriment les variants d'UGT2B10 que nous avons étudiés.

Pour terminer, il convient d'adresser les forces et limites de mes travaux. Tout d'abord nous avons validé l'expression des transcrits alternatifs dans les tissus hépatiques humains en

utilisant deux techniques distinctes, soit par séquençage à haut débit et par PCR. Nous avons ensuite validé l'expression de la protéine UGT2B10_i5 possédant une nouvelle séquence C-terminale par spectrométrie de masse. Ensuite, nous avons combiné des analyses sur de multiples tissus hépatiques humains à des expérimentations dans deux modèles cellulaires différents pour étudier les variants alternatifs. Cependant, la détection par immunobuvardage dans des tissus humains des protéines alternatives demeure à être démontrée. Cette absence de signal pourrait être due à un niveau d'expression trop faible pour être détecté par l'anticorps polyclonal que nous avons utilisé. En outre, la validation de l'expression d'UGT2B10_i4 dans les tissus humains n'a pas pu être effectuée car le seul peptide spécifique à cette isoforme est fortement hydrophobe, empêchant sa détection par spectrométrie de masse. Nos travaux ont démontré pour la première fois la contribution majeure de l'épissage alternatif sur la diversité du transcriptome d'UGT2B10. Ils suggèrent aussi un rôle biologique important des isoformes alternatives, notamment leur influence sur la capacité de détoxification des cellules.

Conclusion

Les résultats présentés dans cette thèse soutiennent nos hypothèses initiales. Premièrement, les polymorphismes génétiques de la voie de glucuronidation et d'autres gènes de la pharmacologie du traitement ont le potentiel d'affecter la réponse clinique des patients. Ils pourraient ainsi représenter des marqueurs prédictifs de l'efficacité du traitement. Deuxièmement, l'épissage alternatif influence la capacité de glucuronidation et pourrait contribuer à une proportion significative de la variabilité de la voie des UGT et du métabolisme cellulaire. Ces deux sources de variabilité contribuent à l'hétérogénéité interindividuelle de l'expression et de l'activité de la voie de glucuronidation. Ils constituent donc de potentiels marqueurs du métabolisme et de la réponse à diverses molécules, notamment des médicaments pris en charge par cette voie.

Dans une optique de médecine personnalisée, ce type de marqueurs pourrait permettre de mieux prédire la réponse à un traitement. Les polymorphismes génétiques semblent les marqueurs les plus aisés à employer car l'information qu'ils fournissent reste constante tout au long de la vie de l'individu. L'expression des variants d'épissage alternatif peut varier au cours du temps en réponse à un traitement pharmacologique ou une condition pathologique, ce qui rend plus difficile son utilisation clinique, mais pas impossible comme le montre l'exemple du AR-V7. Ce mécanisme pourrait contribuer de façon significative à la réponse à divers médicaments en modulant la capacité de glucuronidation. Une meilleure compréhension des facteurs impliqués dans l'expression des isoformes alternatives est donc nécessaire. L'établissement d'une signature prédictive reflétant l'ensemble de la variabilité affectant la réponse à un traitement, en incluant l'impact des polymorphismes génétiques et de l'épissage alternatif, constituerait la suite de ces travaux.

Enfin, au-delà de la prise en charge de nombreux médicaments, les UGT sont aussi responsables du métabolisme de multiples molécules endogènes incluant les hormones mais aussi exogènes tels que des carcinogènes et polluants environnementaux. Les mécanismes de variabilité de la voie de glucuronidation pourraient donc avoir des répercussions plus étendues sur le métabolisme cellulaire ou le risque de développer certaines maladies, comme le cancer. L'étude de l'impact de la variabilité de la voie de

glucuronidation et de ses conséquences sur la perturbation d'autres voies cellulaires pourrait aider à identifier de nouvelles cibles thérapeutiques dans diverses maladies, ou déterminer quelle alternative thérapeutique pourrait être envisagée chez les personnes présentant un profil prédictif de réponse défavorable à la thérapie.

Annexe A : Germline polymorphisms associated to survival of mCRC patients treated with irinotecan-containing regimen

Marker	Impact	n =	Study population	Candidate genes (number of SNP)	Regimen	Reference
SECTION A: One gene – one cohort						
<i>GSTP1</i> rs1695AG+GG	↑ PFS	141	mainly Caucasian	<i>GSTP1</i> (1)	CAPIRI	(Kweekel et al., 2008)
<i>EZH2</i> rs3757441CC	↓ PFS & OS	110	Not specified	<i>EZH2</i> (4)	FOLFIRI + BV	(Crea et al., 2012)
<i>MLH1</i> rs1800734A	↑ PFS	125	Chinese	<i>MLH1</i> (1)	FOLFIRI	(Wang et al., 2018a)
<i>UGT1A1</i> *28 ; Haplotype <i>UGT1</i> rs3755321T, rs3821242C, rs4124874G, rs3755319C	↓ PFS & OS	164	Chinese	<i>UGT1</i> (15)	containing irinotecan	(Yu et al., 2018b)
<i>UGT1A1</i> *28	↑ PFS	250	Recruited in Italy	<i>UGT1</i> (5)	FOLFIRI	(Cecchin et al., 2009)
<i>UGT1A1</i> *28	↓ OS	329	Israeli	<i>UGT1</i> (1)	containing irinotecan	(Shulman et al., 2011)
<i>VEGF</i> rs1570360GG	↓ PFS	40	mainly Caucasian	<i>VEGF</i> (8)	FOLFIRI + BV	(Formica et al., 2011)
<i>VEGF</i> rs833061TT	↓ PFS	111	Recruited in Italy	<i>VEGF</i> (4)	FOLFIRI + BV	(Loupakis et al., 2011)
<i>VEGF</i> rs1570360GG	↓ OS	173	Not specified	<i>VEGF</i> (5)	FOLFIRI + BV & XELIRI + BV	(Koutras et al., 2012)
<i>VEGFA</i> -2549 del/del + rs1570360GG + rs833061TT	↑ PFS	98	Recruited in France	<i>VEGFA</i> (5)	FOLFIRI + CT	(Rollin et al., 2015)
<i>UGT1A1</i> *28	↓ PFS	70	Chinese	<i>UGT1</i> (1)	FOLFIRI + BV	(Lu et al., 2015)
<i>VEGFA</i> rs83306T	↑ PFS	125	Chinese	<i>VEGFA</i> (2)	FOLFIRI + BV	(Cui et al., 2017)

<i>CD133</i> rs3130CC	↓ OS	58	Greek study	<i>CD133</i> (2)	containing irinotecan + BV	(Aravantinos et al., 2015)
SECTION B: One gene - two cohorts						
<i>UGT1A1</i> *28 + <i>UGT1A1</i> *6	↓ OS ↓ OS	114 69	Uygur Han	<i>UGT1</i> (2)	containing irinotecan	(Xu et al., 2016)
SECTION C: Multi-genes - one cohort						
<i>ABCB1</i> rs2032582GG	↓ OS	250	Recruited in Italy	<i>ABCB1</i> (9), <i>ABCC1</i> (10), <i>ABCC2</i> (10), <i>ABCG2</i> (7), <i>SLCO1B1</i> (3)	FOLFIRI	(De Mattia et al., 2013)
Haplotype <i>ABCB1</i> rs1128503T + rs2032582T + rs1045642T	↓ OS	140	mainly Caucasian	<i>ABCB1</i> (3), <i>MTHFR</i> (2), <i>TYMS</i> (1), <i>UGT1</i> (2)	containing irinotecan	(Glimelius et al., 2011)
<i>ABCB1</i> rs1045642TT <i>ABCB1</i> rs2032582TT <i>VKORC1</i> rs9923231TT	↑ PFS ↑ PFS ↑ OS	52	European	34 genes(207)	OPTILIV (<i>KRAS</i> WT pts)	(Levi et al., 2017)
<i>ABCC2</i> rs717620C	↑ PFS	61	Japanese	12 genes(29)	FOLFIRI	(Akiyama et al., 2012)
<i>ABCC2</i> rs717620TT <i>SLCO1B1</i> rs2306283GG <i>UGT1A1</i> *28	↓ PFS ↑ PFS ↓ PFS	103	mainly Caucasian	<i>ABCB1</i> (1), <i>ABCG2</i> (2), <i>ABCC2</i> (2), <i>ABCC5</i> (1), <i>SLCO1B1</i> (3), <i>UGT1</i> (1), <i>CES1</i> (2), <i>CYP3A4</i> (1) <i>CYP3A5</i> (1)	FOLFIRI +/- BV	(Teft et al., 2015)
<i>CYP3A5</i> rs776746GG	↑ PFS	82	Chinese	33 <i>CYP</i> genes(79)	FOLFIRI	(Dong et al., 2015)
<i>DPYD</i> rs1801159G	↑ PFS	661	Chinese	<i>UGT1A</i> (7), <i>DPYD</i> (3)	containing irinotecan	(Liu et al., 2017)
<i>GSTP1</i> rs1695AA	↓ PFS	60	Not specified	10 genes(13)	containing irinotecan	(Zarate et al., 2010)
<i>SHMT1</i> rs1979277T	↑ PFS & OS	85	Recruited in Hungary	<i>MTHFR</i> (1), <i>SHMT1</i> (1), <i>TYMS</i> (2)	FOLFIRI + BV	(Budai et al., 2012)

<i>SLCO1B1</i> rs2306283A	↑ PFS	137	Chinese	<i>SLCO1B1</i> (2), <i>SLC19A1</i> (1)	containing irinotecan (mainly FOLFIRI)	(Huang et al., 2013)
<i>XRCC1</i> rs25487AA	↓ OS	43	Turkish	<i>XPD</i> (1), <i>XRCC1</i> (1)	containing irinotecan	(Artac et al., 2010)
<i>pri-miR26a-1</i> rs7372209TT <i>pri-miR-100</i> rs1834306TT	↓ PFS ↓ PFS	61	Caucasian	14 genes(18)	containing irinotecan	(Boni et al., 2011)
<i>VEGFR2</i> rs12505758C	↓ PFS	424	Recruited in Italy	<i>VGFA</i> (3), <i>VEGFR1</i> (2), <i>VEGFR2</i> (3), <i>EPAS1</i> (1)	FOLFIRI + BV	(Loupakis et al., 2013)
<i>VEGFA</i> rs25648C	↑ OS	45	Recruited in France	<i>UGT1</i> (1), <i>TYMS</i> (1), <i>VEGFA</i> (3)	FOLFIRI + BV	(Becouarn et al., 2014)
<i>CCL5</i> rs2280789G <i>CCR5</i> rs1799988T	↓ OS ↓ PFS	244	Recruited in Germany and Austria	<i>CCL5</i> (1), <i>CCR5</i> (2)	FOLFIRI + CT (<i>KRAS</i> WT)	(Suenaga et al., 2019)
<i>UGT1A1</i> *28 <i>UGT1A1</i> *60	↓ OS	42	Asians	<i>UGT1</i> (11), <i>CYP2A6</i> (3)	containing irinotecan	(Kim et al., 2013)
<i>TYMS</i> 3'UTR 6+/6+ <i>XRCC3</i> rs861539CC and their combination	↓ PFS	146	Recruited in Italy	<i>TYMS</i> (2), <i>MTHFR</i> (2), <i>ERCC1</i> (1), <i>XPD</i> (2), <i>XRCC1</i> (1), <i>XRCC3</i> (1), <i>UGT1</i> (1)	FOLFIRI	(Ruzzo et al., 2008)
<i>RASSF1</i> rs2236947CC	↑ OS	150	Recruited in Germany and Austria	<i>RASSF1</i> (2), <i>LATS2</i> (1), <i>TAZ</i> (1)	FOLFIRI + CT (<i>RAS</i> WT left-sided mCRC)	(Sebio et al., 2018)
<i>EGFR</i> rs11568315 S/S <i>EGF</i> rs4444903GG	↑ OS	110	Recruited in Italy	<i>EGF</i> (1), <i>EGFR</i> (3), <i>CCND1</i> (1), <i>FCGR2A</i> (1), <i>FCGR3A</i> (1)	irinotecan + CT	(Graziano et al., 2008)
<i>FCGR2A</i> rs1801274G + <i>FCGR3A</i> rs396991T	↓ PFS	51	Recruited in France	<i>FCGR2A</i> (1), <i>FCGR3A</i> (1)	containing irinotecan + CT	(Bibeau et al., 2009)

<i>TYMS</i> + <i>FCGR3A</i> favorable genotypes	↑ OS	52	Recruited in France	<i>EGF</i> (1), <i>EGFR</i> (3), <i>UGT1</i> (1), <i>TYMS</i> (3), <i>MTHFR</i> (2), <i>FCGR2A</i> (1), <i>FCGR3A</i> (1)	CETUFTIRI	(Etienne-Grimaldi et al., 2012)
<i>EGFR</i> rs712830CC <i>CCND1</i> rs603965G	↑ PFS ↓ PFS	34	Mainly recruited in France	<i>EGF</i> (1), <i>EGFR</i> (4), <i>CCND1</i> (1), <i>FCGR2A</i> (1), <i>FCGR3A</i> (1)	mainly containing irinotecan + CT (<i>KRAS</i> WT)	(Dahan et al., 2011)
<i>FCGR3A</i> rs396991GG <i>CCND1</i> rs603965G	↓ OS ↓ OS	58	Mainly recruited in France	<i>EGF</i> (1), <i>EGFR</i> (4), <i>CCND1</i> (1), <i>FCGR2A</i> (1), <i>FCGR3A</i> (1)	mainly containing irinotecan + CT	(Dahan et al., 2011)
<i>AREG</i> rs11942466AA <i>AREG</i> rs9996584T <i>EGFR</i> rs712829T	↓ PFS & OS ↑ OS ↑ PFS & OS	74	Recruited in Spain	<i>EGF</i> (8), <i>EGFR</i> (3), <i>AREG/B</i> (9), <i>EREG</i> (2)	containing irinotecan + CT or PN (<i>BRAF</i> & <i>KRAS</i> WT)	(Sebio et al., 2014)
<i>PDGFR-β</i> rs2302273T	↑ PFS	424	Recruited in Italy	<i>PDGFR-β</i> (2), <i>RGS5</i> (2), <i>CSPG4/NG2</i> (2), <i>RALBP1</i> (2)	FOLFIRI + BV	(Volz et al., 2015)
<i>UBE2M</i> rs895374A	↑ PFS	153	Recruited in Germany and Austria	<i>c-CBL</i> (4), <i>EPS15</i> (3), <i>NAE1</i> (3), <i>SH3KBP1</i> (4), <i>SGIP1</i> (3), <i>UBE2M</i> (2), <i>UBE2L3</i> (1)	FOLFIRI + CT (<i>BRAF</i> & <i>KRAS</i> WT)	(Stintzing et al., 2015)
SECTION D: Multi-genes - two or more cohorts						
<i>TLR1</i> rs5743618TT	↓ PFS & OS Not replicated	228 297	Italian	<i>TLR1</i> (2), <i>TLR2</i> (2), <i>TLR6</i> (2), <i>TAK1</i> (3)	FOLFIRI + BV	(Okazaki et al., 2016)
<i>TBK1</i> rs7486100T	↓ OS ↓ PFS	95 93	Recruited in Italy, Germany and Austria	<i>CCL2</i> (1), <i>CCR2</i> (1), <i>HRG</i> (3), <i>PIGF</i> (1), <i>NFKB1</i> (1), <i>TBK1</i> (2), <i>CCL18</i> (1), <i>IRF3</i> (2)	FOLFIRI + BV (<i>KRAS</i> WT)	(Sunakawa et al., 2015)
<i>IL6</i> rs2069837G	↓ PFS ↓ PFS	223 270	Recruited in Italy, Germany and Austria	<i>IL-6</i> (2), <i>STAT3</i> (2)	FOLFIRI + BV	(Matusaka et al., 2016)

<i>CXCR4</i> rs2228014CC	↑ PFS ↑ PFS	397 229	Recruited in Italy	<i>CXCR4</i> (1), <i>CXCL12</i> (2)	FOLFIRI + BV FOLFOXIRI + BV	(Matsusaka et al., 2017)
<i>CD39</i> rs11188513C	↓ OS ↓ OS	107 215	Recruited in Italy, Germany and Austria	<i>ENTPD1</i> (2), <i>NT5E</i> (1), <i>ADORA2A</i> (1), <i>ADORA2B</i> (1), <i>HIF1A</i> (2)	FOLFIRI + BV	(Tokunaga et al., 2019b)
<i>MKNK1</i> rs8602AA	↓ PFS ↓ PFS	236 94	Recruited in Italy, Germany and Austria	<i>MKNK1</i> (1), <i>EIF4E</i> (1), <i>EIF4G1</i> (1), <i>4EBP1</i> (1), <i>HNRNPA1</i> (1), <i>SPRY2</i> (1)	FOLFIRI + BV (<i>KRAS</i> WT)	(Berger et al., 2017)
<i>GC</i> rs4588AA	↓ OS ↓ OS	287 227	Recruited in Italy, Germany and Austria	<i>GC</i> (1), <i>CYP24A1</i> (1), <i>CYP27B1</i> (1), <i>VDR</i> (1), <i>DKK1</i> (1), <i>CST5</i> (1)	FOLFIRI + BV	(Berger et al., 2018)
<i>PIN1</i> rs2233678C	↑ PFS Not replicated	228 276	Recruited in Italy	<i>PIN1</i> (2), <i>TP53</i> (1), <i>AATF</i> (1), <i>CSNK2A1</i> (1), <i>CSNK2B</i> (1), <i>TOP1</i> (1), <i>RBBP8</i> (1)	FOLFIRI + BV	(Suenaga et al., 2018)
Several SNP related to AMPK	PFS & OS	723 (total)	Recruited in North America and Europe	13 genes(19)	containing irinotecan + BV or CT	(Tokunaga et al., 2019a)

OPTILIV: hepatic artery infusion of irinotecan, oxaliplatin and 5-fluorouracil with intravenous cetuximab; CAPIRI or XELIRI: capecitabine (Xeloda) + irinotecan; FOLFIRI: folinic acid + 5-fluorouracil + irinotecan; FOLFOXIRI : folinic acid + 5-fluorouracil + oxaliplatin + irinotecan; CETUFTIRI: cetuximab + tegafur/uracil + folinic acid + irinotecan. BV: bevacizumab; CT: cetuximab, PN: panitumumab. Pts: patients; WT: wild-type.

Bibliographie

- Achour, B., Barber, J., and Rostami-Hodjegan, A., Expression of hepatic drug-metabolizing cytochrome p450 enzymes and their intercorrelations: a meta-analysis. *Drug Metab Dispos*, 2014a. 42(8): p. 1349-56.
- Achour, B., Dantonio, A., Niosi, M., Novak, J.J., Fallon, J.K., Barber, J., Smith, P.C., Rostami-Hodjegan, A., and Goosen, T.C., Quantitative Characterization of Major Hepatic UDP-Glucuronosyltransferase Enzymes in Human Liver Microsomes: Comparison of Two Proteomic Methods and Correlation with Catalytic Activity. *Drug Metab Dispos*, 2017. 45(10): p. 1102-1112.
- Achour, B., Russell, M.R., Barber, J., and Rostami-Hodjegan, A., Simultaneous quantification of the abundance of several cytochrome P450 and uridine 5'-diphospho-glucuronosyltransferase enzymes in human liver microsomes using multiplexed targeted proteomics. *Drug Metab Dispos*, 2014b. 42(4): p. 500-10.
- Akiyama, Y., Fujita, K., Ishida, H., Sunakawa, Y., Yamashita, K., Kawara, K., Miwa, K., Saji, S., and Sasaki, Y., Association of ABCC2 genotype with efficacy of first-line FOLFIRI in Japanese patients with advanced colorectal cancer. *Drug Metab Pharmacokinet*, 2012. 27(3): p. 325-35.
- Alix-Panabieres, C. and Pantel, K., Clinical Applications of Circulating Tumor Cells and Circulating Tumor DNA as Liquid Biopsy. *Cancer Discov*, 2016. 6(5): p. 479-91.
- Allen, W.L., Stevenson, L., Coyle, V.M., Jithesh, P.V., Proutski, I., Carson, G., Gordon, M.A., Lenz, H.J., Van Schaeybroeck, S., Longley, D.B., and Johnston, P.G., A systems biology approach identifies SART1 as a novel determinant of both 5-fluorouracil and SN38 drug resistance in colorectal cancer. *Mol Cancer Ther*, 2012. 11(1): p. 119-31.
- Amstutz, U., Henricks, L.M., Offer, S.M., Barbarino, J., Schellens, J.H.M., Swen, J.J., Klein, T.E., McLeod, H.L., Caudle, K.E., Diasio, R.B., and Schwab, M., Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Dihydropyrimidine Dehydrogenase Genotype and Fluoropyrimidine Dosing: 2017 Update. *Clin Pharmacol Ther*, 2018. 103(2): p. 210-216.
- Annalora, A.J., Marcus, C.B., and Iversen, P.L., Alternative Splicing in the Cytochrome P450 Superfamily Expands Protein Diversity to Augment Gene Function and Redirect Human Drug Metabolism. *Drug Metab Dispos*, 2017. 45(4): p. 375-389.
- Antonarakis, E.S., Lu, C., Lubber, B., Wang, H., Chen, Y., Nakazawa, M., Nadal, R., Paller, C.J., Denmeade, S.R., Carducci, M.A., Eisenberger, M.A., and Luo, J., Androgen Receptor Splice Variant 7 and Efficacy of Taxane Chemotherapy in Patients With Metastatic Castration-Resistant Prostate Cancer. *JAMA Oncol*, 2015. 1(5): p. 582-91.
- Antonarakis, E.S., Lu, C., Wang, H., Lubber, B., Nakazawa, M., Roeser, J.C., Chen, Y., Mohammad, T.A., Chen, Y., Fedor, H.L., Lotan, T.L., Zheng, Q., De Marzo, A.M., Isaacs, J.T., Isaacs, W.B., Nadal, R., Paller, C.J., Denmeade, S.R., Carducci, M.A., Eisenberger, M.A., and Luo, J., AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med*, 2014. 371(11): p. 1028-38.
- Aoyama, T., Yamano, S., Waxman, D.J., Lapenson, D.P., Meyer, U.A., Fischer, V., Tyndale, R., Inaba, T., Kalow, W., Gelboin, H.V., and et al., Cytochrome P-450 hPCN3, a novel cytochrome P-450 IIIA gene product that is differentially expressed

- in adult human liver. cDNA and deduced amino acid sequence and distinct specificities of cDNA-expressed hPCN1 and hPCN3 for the metabolism of steroid hormones and cyclosporine. *J Biol Chem*, 1989. 264(18): p. 10388-95.
- Aparicio, T., Lavau-Denes, S., Phelip, J.M., Maillard, E., Jouve, J.L., Gargot, D., Gasmi, M., Locher, C., Adhoute, X., Michel, P., Khemissa, F., Lecomte, T., Provencal, J., Breysacher, G., Legoux, J.L., Lepere, C., Charneau, J., Cretin, J., Chone, L., Azzedine, A., Bouche, O., Sobhani, I., Bedenne, L., and Mitry, E., Randomized phase III trial in elderly patients comparing LV5FU2 with or without irinotecan for first-line treatment of metastatic colorectal cancer (FFCD 2001-02). *Ann Oncol*, 2016. 27(1): p. 121-7.
- Aranda, E., Valladares, M., Martinez-Villacampa, M., Benavides, M., Gomez, A., Massutti, B., Marcuello, E., Constenla, M., Camara, J.C., Carrato, A., Duenas, R., Reboredo, M., Navarro, M., and Diaz-Rubio, E., Randomized study of weekly irinotecan plus high-dose 5-fluorouracil (FUIRI) versus biweekly irinotecan plus 5-fluorouracil/leucovorin (FOLFIRI) as first-line chemotherapy for patients with metastatic colorectal cancer: a Spanish Cooperative Group for the Treatment of Digestive Tumors Study. *Ann Oncol*, 2009. 20(2): p. 251-7.
- Aravantinos, G., Isaakidou, A., Karantanos, T., Sioziou, A., Theodoropoulos, G.E., Pektasides, D., and Gazouli, M., Association of CD133 polymorphisms and response to bevacizumab in patients with metastatic colorectal cancer. *Cancer Biomark*, 2015. 15(6): p. 843-50.
- Artac, M., Bozcuk, H., Pehlivan, S., Akcan, S., Pehlivan, M., Sever, T., Ozdogan, M., and Savas, B., The value of XPD and XRCC1 genotype polymorphisms to predict clinical outcome in metastatic colorectal carcinoma patients with irinotecan-based regimens. *J Cancer Res Clin Oncol*, 2010. 136(6): p. 803-9.
- Audet-Delage, Y., Rouleau, M., Rouleau, M., Roberge, J., Miard, S., Picard, F., Tetu, B., and Guillemette, C., Cross-Talk between Alternatively Spliced UGT1A Isoforms and Colon Cancer Cell Metabolism. *Mol Pharmacol*, 2017. 91(3): p. 167-177.
- Bajpai, P., Srinivasan, S., Ghosh, J., Nagy, L.D., Wei, S., Guengerich, F.P., and Avadhani, N.G., Targeting of splice variants of human cytochrome P450 2C8 (CYP2C8) to mitochondria and their role in arachidonic acid metabolism and respiratory dysfunction. *J Biol Chem*, 2014. 289(43): p. 29614-30.
- Bank, P.C.D., Caudle, K.E., Swen, J.J., Gammal, R.S., Whirl-Carrillo, M., Klein, T.E., Relling, M.V., and Guchelaar, H.J., Comparison of the Guidelines of the Clinical Pharmacogenetics Implementation Consortium and the Dutch Pharmacogenetics Working Group. *Clin Pharmacol Ther*, 2018. 103(4): p. 599-618.
- Barbier, O., Girard, C., Breton, R., Belanger, A., and Hum, D.W., N-glycosylation and residue 96 are involved in the functional properties of UDP-glucuronosyltransferase enzymes. *Biochemistry*, 2000. 39(38): p. 11540-52.
- Barua, A.B. and Sidell, N., Retinoyl beta-glucuronide: a biologically active interesting retinoid. *J Nutr*, 2004. 134(1): p. 286S-289S.
- Basu, N.K., Kole, L., Basu, M., Chakraborty, K., Mitra, P.S., and Owens, I.S., The major chemical-detoxifying system of UDP-glucuronosyltransferases requires regulated phosphorylation supported by protein kinase C. *J Biol Chem*, 2008. 283(34): p. 23048-61.
- Battaglia, E., Nowell, S., Drake, R.R., Mizeracka, M., Berg, C.L., Magdalou, J., Fournel-Gigleux, S., Gollan, J.L., Lester, R., and Radominska, A., Two kinetically-distinct

- components of UDP-glucuronic acid transport in rat liver endoplasmic reticulum. *Biochim Biophys Acta*, 1996. 1283(2): p. 223-31.
- Beaulieu, M., Levesque, E., Hum, D.W., and Belanger, A., Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids. *J Biol Chem*, 1996. 271(37): p. 22855-62.
- Becouarn, Y., Cany, L., Pulido, M., Beyssac, R., Texereau, P., Le Morvan, V., Bechade, D., Brunet, R., Aitouferoukh, S., Lalet, C., Mathoulin-Pelissier, S., Fonck, M., and Robert, J., FOLFIRI(R) and bevacizumab in first-line treatment for colorectal cancer patients: safety, efficacy and genetic polymorphisms. *BMC Res Notes*, 2014. 7: p. 260.
- Beer, T.M., Armstrong, A.J., Rathkopf, D.E., Lorient, Y., Sternberg, C.N., Higano, C.S., Iversen, P., Bhattacharya, S., Carles, J., Chowdhury, S., Davis, I.D., de Bono, J.S., Evans, C.P., Fizazi, K., Joshua, A.M., Kim, C.S., Kimura, G., Mainwaring, P., Mansbach, H., Miller, K., Noonberg, S.B., Perabo, F., Phung, D., Saad, F., Scher, H.I., Taplin, M.E., Venner, P.M., and Tombal, B., Enzalutamide in metastatic prostate cancer before chemotherapy. *N Engl J Med*, 2014. 371(5): p. 424-33.
- Beger, R.D., Dunn, W., Schmidt, M.A., Gross, S.S., Kirwan, J.A., Cascante, M., Brennan, L., Wishart, D.S., Oresic, M., Hankemeier, T., Broadhurst, D.I., Lane, A.N., Suhre, K., Kastenmuller, G., Sumner, S.J., Thiele, I., Fiehn, O., and Kaddurah-Daouk, R., Metabolomics enables precision medicine: "A White Paper, Community Perspective". *Metabolomics*, 2016. 12(10): p. 149.
- Belanger, A., Hum, D.W., Beaulieu, M., Levesque, E., Guillemette, C., Tchernof, A., Belanger, G., Turgeon, D., and Dubois, S., Characterization and regulation of UDP-glucuronosyltransferases in steroid target tissues. *J Steroid Biochem Mol Biol*, 1998. 65(1-6): p. 301-10.
- Belanger, A.S., Tojcic, J., Harvey, M., and Guillemette, C., Regulation of UGT1A1 and HNF1 transcription factor gene expression by DNA methylation in colon cancer cells. *BMC Mol Biol*, 2010. 11: p. 9.
- Belledant, A., Hovington, H., Garcia, L., Caron, P., Brisson, H., Villeneuve, L., Simonyan, D., Tetu, B., Fradet, Y., Lacombe, L., Guillemette, C., and Levesque, E., The UGT2B28 Sex-steroid Inactivation Pathway Is a Regulator of Steroidogenesis and Modifies the Risk of Prostate Cancer Progression. *Eur Urol*, 2016. 69(4): p. 601-609.
- Bellemare, J., Rouleau, M., Girard, H., Harvey, M., and Guillemette, C., Alternatively spliced products of the UGT1A gene interact with the enzymatically active proteins to inhibit glucuronosyltransferase activity in vitro. *Drug Metab Dispos*, 2010a. 38(10): p. 1785-9.
- Bellemare, J., Rouleau, M., Harvey, M., and Guillemette, C., Modulation of the human glucuronosyltransferase UGT1A pathway by splice isoform polypeptides is mediated through protein-protein interactions. *J Biol Chem*, 2010b. 285(6): p. 3600-7.
- Bellemare, J., Rouleau, M., Harvey, M., Tetu, B., and Guillemette, C., Alternative-splicing forms of the major phase II conjugating UGT1A gene negatively regulate glucuronidation in human carcinoma cell lines. *Pharmacogenomics J*, 2010c. 10(5): p. 431-41.

- Beretta, G.L., Perego, P., and Zunino, F., Targeting topoisomerase I: molecular mechanisms and cellular determinants of response to topoisomerase I inhibitors. *Expert Opin Ther Targets*, 2008. 12(10): p. 1243-56.
- Berg, J.Z., von Weymarn, L.B., Thompson, E.A., Wickham, K.M., Weisensel, N.A., Hatsukami, D.K., and Murphy, S.E., UGT2B10 genotype influences nicotine glucuronidation, oxidation, and consumption. *Cancer Epidemiol Biomarkers Prev*, 2010. 19(6): p. 1423-31.
- Berger, M.D., Stintzing, S., Heinemann, V., Cao, S., Yang, D., Sunakawa, Y., Matsusaka, S., Ning, Y., Okazaki, S., Miyamoto, Y., Suenaga, M., Schirripa, M., Hanna, D.L., Soni, S., Puccini, A., Zhang, W., Cremolini, C., Falcone, A., Loupakis, F., and Lenz, H.J., A Polymorphism within the Vitamin D Transporter Gene Predicts Outcome in Metastatic Colorectal Cancer Patients Treated with FOLFIRI/Bevacizumab or FOLFIRI/Cetuximab. *Clin Cancer Res*, 2018. 24(4): p. 784-793.
- Berger, M.D., Stintzing, S., Heinemann, V., Yang, D., Cao, S., Sunakawa, Y., Ning, Y., Matsusaka, S., Okazaki, S., Miyamoto, Y., Suenaga, M., Schirripa, M., Soni, S., Zhang, W., Falcone, A., Loupakis, F., and Lenz, H.J., Impact of genetic variations in the MAPK signaling pathway on outcome in metastatic colorectal cancer patients treated with first-line FOLFIRI and bevacizumab: data from FIRE-3 and TRIBE trials. *Ann Oncol*, 2017. 28(11): p. 2780-2785.
- Berman, E., Jhanwar, S., Hedvat, C., Arcila, M.E., Wahab, O.A., Levine, R., Maloy, M., Ma, W., and Albitar, M., Resistance to imatinib in patients with chronic myelogenous leukemia and the splice variant BCR-ABL1(35INS). *Leuk Res*, 2016. 49: p. 108-12.
- Bernard, P., Goudonnet, H., Artur, Y., Desvergne, B., and Wahli, W., Activation of the mouse TATA-less and human TATA-containing UDP-glucuronosyltransferase 1A1 promoters by hepatocyte nuclear factor 1. *Mol Pharmacol*, 1999. 56(3): p. 526-36.
- Bertilsson, G., Heidrich, J., Svensson, K., Asman, M., Jendeberg, L., Sydow-Backman, M., Ohlsson, R., Postlind, H., Blomquist, P., and Berkenstam, A., Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction. *Proc Natl Acad Sci U S A*, 1998. 95(21): p. 12208-13.
- Bertilsson, L., Dahl, M.L., Dalen, P., and Al-Shurbaji, A., Molecular genetics of CYP2D6: clinical relevance with focus on psychotropic drugs. *Br J Clin Pharmacol*, 2002. 53(2): p. 111-22.
- Beutler, E., Gelbart, T., and Demina, A., Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? *Proc Natl Acad Sci U S A*, 1998. 95(14): p. 8170-4.
- Bhatt, D.K., Basit, A., Zhang, H., Gaedigk, A., Lee, S.B., Claw, K.G., Mehrotra, A., Chaudhry, A.S., Pearce, R.E., Gaedigk, R., Broeckel, U., Thornton, T.A., Nickerson, D.A., Schuetz, E.G., Amory, J.K., Leeder, J.S., and Prasad, B., Hepatic Abundance and Activity of Androgen- and Drug-Metabolizing Enzyme UGT2B17 Are Associated with Genotype, Age, and Sex. *Drug Metab Dispos*, 2018. 46(6): p. 888-896.
- Bibeau, F., Lopez-Crapez, E., Di Fiore, F., Thezenas, S., Ychou, M., Blanchard, F., Lamy, A., Penault-Llorca, F., Frebourg, T., Michel, P., Sabourin, J.C., and Boissiere-Michot, F., Impact of Fc{gamma}RIIa-Fc{gamma}RIIIa polymorphisms and

- KRAS mutations on the clinical outcome of patients with metastatic colorectal cancer treated with cetuximab plus irinotecan. *J Clin Oncol*, 2009. 27(7): p. 1122-9.
- Blumberg, B., Sabbagh, W., Jr., Juguilon, H., Bolado, J., Jr., van Meter, C.M., Ong, E.S., and Evans, R.M., SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Genes Dev*, 1998. 12(20): p. 3195-205.
- Bock, K.W., Roles of human UDP-glucuronosyltransferases in clearance and homeostasis of endogenous substrates, and functional implications. *Biochem Pharmacol*, 2015. 96(2): p. 77-82.
- Boni, V., Zarate, R., Villa, J.C., Bandres, E., Gomez, M.A., Maiello, E., Garcia-Foncillas, J., and Aranda, E., Role of primary miRNA polymorphic variants in metastatic colon cancer patients treated with 5-fluorouracil and irinotecan. *Pharmacogenomics J*, 2011. 11(6): p. 429-36.
- Bosio, A., Binczek, E., and Stoffel, W., Molecular cloning and characterization of the mouse CGT gene encoding UDP-galactose ceramide-galactosyltransferase (cerebroside synthetase). *Genomics*, 1996. 35(1): p. 223-6.
- Bosma, P.J., Chowdhury, J.R., Bakker, C., Gantla, S., de Boer, A., Oostra, B.A., Lindhout, D., Tytgat, G.N., Jansen, P.L., Oude Elferink, R.P., and et al., The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N Engl J Med*, 1995. 333(18): p. 1171-5.
- Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., Cherry, J.M., and Snyder, M., Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*, 2012. 22(9): p. 1790-7.
- Budai, B., Komlosi, V., Adleff, V., Pap, E., Reti, A., Nagy, T., Kralovanszky, J., Lang, I., and Hitre, E., Impact of SHMT1 polymorphism on the clinical outcome of patients with metastatic colorectal cancer treated with first-line FOLFIRI+bevacizumab. *Pharmacogenet Genomics*, 2012. 22(1): p. 69-72.
- Cai, L., Huang, W., and Chou, K.C., Prostate cancer with variants in CYP17 and UGT2B17 genes: a meta-analysis. *Protein Pept Lett*, 2012. 19(1): p. 62-9.
- Caudle, K.E., Klein, T.E., Hoffman, J.M., Muller, D.J., Whirl-Carrillo, M., Gong, L., McDonagh, E.M., Sangkuhl, K., Thorn, C.F., Schwab, M., Agundez, J.A., Freimuth, R.R., Huser, V., Lee, M.T., Iwuchukwu, O.F., Crews, K.R., Scott, S.A., Wadelius, M., Swen, J.J., Tyndale, R.F., Stein, C.M., Roden, D., Relling, M.V., Williams, M.S., and Johnson, S.G., Incorporation of pharmacogenomics into routine clinical practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) guideline development process. *Curr Drug Metab*, 2014. 15(2): p. 209-17.
- Cecchin, E., De Mattia, E., and Toffoli, G., Nuclear receptors and drug metabolism for the personalization of cancer therapy. *Expert Opin Drug Metab Toxicol*, 2016. 12(3): p. 291-306.
- Cecchin, E., Innocenti, F., D'Andrea, M., Corona, G., De Mattia, E., Biason, P., Buonadonna, A., and Toffoli, G., Predictive role of the UGT1A1, UGT1A7, and UGT1A9 genetic variants and their haplotypes on the outcome of metastatic colorectal cancer patients treated with fluorouracil, leucovorin, and irinotecan. *J Clin Oncol*, 2009. 27(15): p. 2457-65.
- Chai, X., Zeng, S., and Xie, W., Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond. *Expert Opin Drug Metab Toxicol*, 2013. 9(3): p. 253-66.

- Chan, L.Y., Mugler, C.F., Heinrich, S., Vallotton, P., and Weis, K., Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *Elife*, 2018. 7.
- Chen, G., Blevins-Primeau, A.S., Dellinger, R.W., Muscat, J.E., and Lazarus, P., Glucuronidation of nicotine and cotinine by UGT2B10: loss of function by the UGT2B10 Codon 67 (Asp>Tyr) polymorphism. *Cancer Res*, 2007. 67(19): p. 9024-9.
- Chen, G., Dellinger, R.W., Gallagher, C.J., Sun, D., and Lazarus, P., Identification of a prevalent functional missense polymorphism in the UGT2B10 gene and its association with UGT2B10 inactivation against tobacco-specific nitrosamines. *Pharmacogenet Genomics*, 2008. 18(3): p. 181-91.
- Chen, S., Laverdiere, I., Tourancheau, A., Jonker, D., Couture, F., Cecchin, E., Villeneuve, L., Harvey, M., Court, M.H., Innocenti, F., Toffoli, G., Levesque, E., and Guillemette, C., A novel UGT1 marker associated with better tolerance against irinotecan-induced severe neutropenia in metastatic colorectal cancer patients. *Pharmacogenomics J*, 2015a. 15(6): p. 513-20.
- Chen, S., Villeneuve, L., Jonker, D., Couture, F., Laverdiere, I., Cecchin, E., Innocenti, F., Toffoli, G., Levesque, E., and Guillemette, C., ABCC5 and ABCG1 polymorphisms predict irinotecan-induced severe toxicity in metastatic colorectal cancer patients. *Pharmacogenet Genomics*, 2015b. 25(12): p. 573-83.
- Chen, Y., Tang, Y., Guo, C., Wang, J., Boral, D., and Nie, D., Nuclear receptors in the multidrug resistance through the regulation of drug-metabolizing enzymes and drug transporters. *Biochem Pharmacol*, 2012. 83(8): p. 1112-26.
- Chen, Z., Shi, T., Zhang, L., Zhu, P., Deng, M., Huang, C., Hu, T., Jiang, L., and Li, J., Mammalian drug efflux transporters of the ATP binding cassette (ABC) family in multidrug resistance: A review of the past decade. *Cancer Lett*, 2016. 370(1): p. 153-64.
- Chen, Z.S., Furukawa, T., Sumizawa, T., Ono, K., Ueda, K., Seto, K., and Akiyama, S.I., ATP-Dependent efflux of CPT-11 and SN-38 by the multidrug resistance protein (MRP) and its inhibition by PAK-104P. *Mol Pharmacol*, 1999. 55(5): p. 921-8.
- Ciotti, M., Basu, N., Brangi, M., and Owens, I.S., Glucuronidation of 7-ethyl-10-hydroxycamptothecin (SN-38) by the human UDP-glucuronosyltransferases encoded at the UGT1 locus. *Biochem Biophys Res Commun*, 1999. 260(1): p. 199-202.
- Ciotti, M., Cho, J.W., George, J., and Owens, I.S., Required buried alpha-helical structure in the bilirubin UDP-glucuronosyltransferase, UGT1A1, contains a nonreplaceable phenylalanine. *Biochemistry*, 1998. 37(31): p. 11018-25.
- Climente-Gonzalez, H., Porta-Pardo, E., Godzik, A., and Eyraes, E., The Functional Impact of Alternative Splicing in Cancer. *Cell Rep*, 2017. 20(9): p. 2215-2226.
- Cocucci, E., Kim, J.Y., Bai, Y., and Pabla, N., Role of Passive Diffusion, Transporters, and Membrane Trafficking-Mediated Processes in Cellular Drug Transport. *Clin Pharmacol Ther*, 2017. 101(1): p. 121-129.
- Cote, J.F., Kirzin, S., Kramar, A., Mosnier, J.F., Diebold, M.D., Soubeyran, I., Thirouard, A.S., Selves, J., Laurent-Puig, P., and Ychou, M., UGT1A1 polymorphism can predict hematologic toxicity in patients treated with irinotecan. *Clin Cancer Res*, 2007. 13(11): p. 3269-75.

- Court, M.H., Duan, S.X., von Moltke, L.L., Greenblatt, D.J., Patten, C.J., Miners, J.O., and Mackenzie, P.I., Interindividual variability in acetaminophen glucuronidation by human liver microsomes: identification of relevant acetaminophen UDP-glucuronosyltransferase isoforms. *J Pharmacol Exp Ther*, 2001. 299(3): p. 998-1006.
- Court, M.H., Freytsis, M., Wang, X., Peter, I., Guillemette, C., Hazarika, S., Duan, S.X., Greenblatt, D.J., and Lee, W.M., The UDP-glucuronosyltransferase (UGT) 1A polymorphism c.2042C>G (rs8330) is associated with increased human liver acetaminophen glucuronidation, increased UGT1A exon 5a/5b splice variant mRNA ratio, and decreased risk of unintentional acetaminophen-induced acute liver failure. *J Pharmacol Exp Ther*, 2013. 345(2): p. 297-307.
- Court, M.H., Hazarika, S., Krishnaswamy, S., Finel, M., and Williams, J.A., Novel polymorphic human UDP-glucuronosyltransferase 2A3: cloning, functional characterization of enzyme variants, comparative tissue expression, and gene induction. *Mol Pharmacol*, 2008. 74(3): p. 744-54.
- Court, M.H., Zhang, X., Ding, X., Yee, K.K., Hesse, L.M., and Finel, M., Quantitative distribution of mRNAs encoding the 19 human UDP-glucuronosyltransferase enzymes in 26 adult and 3 fetal tissues. *Xenobiotica*, 2012. 42(3): p. 266-77.
- Crea, F., Fornaro, L., Paolicchi, E., Masi, G., Frumento, P., Loupakis, F., Salvatore, L., Cremolini, C., Schirripa, M., Graziano, F., Ronzoni, M., Ricci, V., Farrar, W.L., Falcone, A., and Danesi, R., An EZH2 polymorphism is associated with clinical outcome in metastatic colorectal cancer patients. *Ann Oncol*, 2012. 23(5): p. 1207-13.
- Cremolini, C., Loupakis, F., Antoniotti, C., Lupi, C., Sensi, E., Lonardi, S., Mezi, S., Tomasello, G., Ronzoni, M., Zaniboni, A., Tonini, G., Carlomagno, C., Allegrini, G., Chiara, S., D'Amico, M., Granetto, C., Cazzaniga, M., Boni, L., Fontanini, G., and Falcone, A., FOLFOXIRI plus bevacizumab versus FOLFIRI plus bevacizumab as first-line treatment of patients with metastatic colorectal cancer: updated overall survival and molecular subgroup analyses of the open-label, phase 3 TRIBE study. *Lancet Oncol*, 2015. 16(13): p. 1306-15.
- Crotti, S., Piccoli, M., Rizzolio, F., Giordano, A., Nitti, D., and Agostini, M., Extracellular Matrix and Colorectal Cancer: How Surrounding Microenvironment Affects Cancer Cell Behavior? *J Cell Physiol*, 2017. 232(5): p. 967-975.
- Cui, W., Li, F., Yuan, Q., Chen, G., Chen, C., and Yu, B., Role of VEGFA gene polymorphisms in colorectal cancer patients who treated with bevacizumab. *Oncotarget*, 2017. 8(62): p. 105472-105478.
- Dahan, L., Norguet, E., Etienne-Grimaldi, M.C., Formento, J.L., Gasmi, M., Nanni, I., Gaudart, J., Garcia, S., Ouafik, L., Seitz, J.F., and Milano, G., Pharmacogenetic profiling and cetuximab outcome in patients with advanced colorectal cancer. *BMC Cancer*, 2011. 11: p. 496.
- Daly, A.K., Significance of the minor cytochrome P450 3A isoforms. *Clin Pharmacokinet*, 2006. 45(1): p. 13-31.
- Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., Onate, K.C., Graham, K., Miyasato, S.R., Dreszer, T.R., Strattan, J.S., Jolanki, O., Tanaka, F.Y., and Cherry, J.M., The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res*, 2018. 46(D1): p. D794-D801.

- de Jong, F.A., Scott-Horton, T.J., Kroetz, D.L., McLeod, H.L., Friberg, L.E., Mathijssen, R.H., Verweij, J., Marsh, S., and Sparreboom, A., Irinotecan-induced diarrhea: functional significance of the polymorphic ABCC2 transporter protein. *Clin Pharmacol Ther*, 2007. 81(1): p. 42-9.
- De Mattia, E., Cecchin, E., Montico, M., Labriet, A., Guillemette, C., Dreussi, E., Roncato, R., Bignucolo, A., Buonadonna, A., D'Andrea, M., Coppola, L., Lonardi, S., Levesque, E., Jonker, D., Couture, F., and Toffoli, G., Association of STAT-3 rs1053004 and VDR rs11574077 With FOLFIRI-Related Gastrointestinal Toxicity in Metastatic Colorectal Cancer Patients. *Front Pharmacol*, 2018. 9: p. 367.
- De Mattia, E., Cecchin, E., Roncato, R., and Toffoli, G., Pregnane X receptor, constitutive androstane receptor and hepatocyte nuclear factors as emerging players in cancer precision medicine. *Pharmacogenomics*, 2016. 17(14): p. 1547-71.
- De Mattia, E., Polesel, J., Roncato, R., Labriet, A., Bignucolo, A., Dreussi, E., Romanato, L., Guardascione, M., Buonadonna, A., D'Andrea, M., Lévesque, E., Jonker, D., Couture, F., Guillemette, C., Cecchin, E., and Toffoli, G., Germline polymorphisms in the nuclear receptors PXR and VDR as novel prognostic markers in metastatic colorectal cancer patients treated with FOLFIRI. Submitted, 2019.
- De Mattia, E., Toffoli, G., Polesel, J., D'Andrea, M., Corona, G., Zagonel, V., Buonadonna, A., Dreussi, E., and Cecchin, E., Pharmacogenetics of ABC and SLC transporters in metastatic colorectal cancer patients receiving first-line FOLFIRI treatment. *Pharmacogenet Genomics*, 2013. 23(10): p. 549-57.
- Dean, L., Fluorouracil Therapy and DPYD Genotype, in *Medical Genetics Summaries*, V. Pratt, et al., Editors. 2012a: Bethesda MD.
- Dean, L., Irinotecan Therapy and UGT1A1 Genotype, in *Medical Genetics Summaries*, V. Pratt, et al., Editors. 2012b: Bethesda MD.
- Dean, L., Warfarin Therapy and VKORC1 and CYP Genotype, in *Medical Genetics Summaries*, V. Pratt, et al., Editors. 2012c: Bethesda MD.
- Deenen, M.J., Cats, A., Beijnen, J.H., and Schellens, J.H., Part 4: pharmacogenetic variability in anticancer pharmacodynamic drug effects. *Oncologist*, 2011. 16(7): p. 1006-20.
- Del Rio, M., Mollevi, C., Vezzio-Vie, N., Bibeau, F., Ychou, M., and Martineau, P., Specific extracellular matrix remodeling signature of colon hepatic metastases. *PLoS One*, 2013. 8(9): p. e74599.
- Deo, A.K., Prasad, B., Balogh, L., Lai, Y., and Unadkat, J.D., Interindividual variability in hepatic expression of the multidrug resistance-associated protein 2 (MRP2/ABCC2): quantification by liquid chromatography/tandem mass spectrometry. *Drug Metab Dispos*, 2012. 40(5): p. 852-5.
- Dluzen, D.F., Sun, D., Salzberg, A.C., Jones, N., Bushey, R.T., Robertson, G.P., and Lazarus, P., Regulation of UDP-glucuronosyltransferase 1A1 expression and activity by microRNA 491-3p. *J Pharmacol Exp Ther*, 2014. 348(3): p. 465-77.
- Dluzen, D.F., Sutliff, A.K., Chen, G., Watson, C.J., Ishmael, F.T., and Lazarus, P., Regulation of UGT2B Expression and Activity by miR-216b-5p in Liver Cancer Cell Lines. *J Pharmacol Exp Ther*, 2016. 359(1): p. 182-93.
- Dodds, H.M., Haaz, M.C., Riou, J.F., Robert, J., and Rivory, L.P., Identification of a new metabolite of CPT-11 (irinotecan): pharmacological properties and activation to SN-38. *J Pharmacol Exp Ther*, 1998. 286(1): p. 578-83.

- Dong, N., Meng, F., Wu, Y., Wang, M., Cui, Y., and Zhang, S., Genetic polymorphisms in cytochrome P450 and clinical outcomes of FOLFIRI chemotherapy in patients with metastatic colorectal cancer. *Tumour Biol*, 2015. 36(10): p. 7691-8.
- Dotzlaw, H., Leygue, E., Watson, P., and Murphy, L.C., The human orphan receptor PXR messenger RNA is expressed in both normal and neoplastic breast tissue. *Clin Cancer Res*, 1999. 5(8): p. 2103-7.
- ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature*, 2012. 489(7414): p. 57-74.
- Etienne-Grimaldi, M.C., Bennouna, J., Formento, J.L., Douillard, J.Y., Francoual, M., Hennebelle, I., Chatelut, E., Francois, E., Faroux, R., El Hannani, C., Jacob, J.H., and Milano, G., Multifactorial pharmacogenetic analysis in colorectal cancer patients receiving 5-fluorouracil-based therapy together with cetuximab-irinotecan. *Br J Clin Pharmacol*, 2012. 73(5): p. 776-85.
- Etienne-Grimaldi, M.C., Boyer, J.C., Thomas, F., Quaranta, S., Picard, N., Lorient, M.A., Narjoz, C., Poncet, D., Gagnieu, M.C., Ged, C., Broly, F., Le Morvan, V., Bouquie, R., Gaub, M.P., Philibert, L., Ghiringhelli, F., and Le Guellec, C., UGT1A1 genotype and irinotecan therapy: general review and implementation in routine practice. *Fundam Clin Pharmacol*, 2015. 29(3): p. 219-37.
- Evans, W.E. and Relling, M.V., Pharmacogenomics: translating functional genomics into rational therapeutics. *Science*, 1999. 286(5439): p. 487-91.
- Falcone, A., Ricci, S., Brunetti, I., Pfanner, E., Allegrini, G., Barbara, C., Crino, L., Benedetti, G., Evangelista, W., Fanchini, L., Cortesi, E., Picone, V., Vitello, S., Chiara, S., Granetto, C., Porcile, G., Fioretto, L., Orlandini, C., Andreuccetti, M., and Masi, G., Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J Clin Oncol*, 2007. 25(13): p. 1670-6.
- Fallon, J.K., Neubert, H., Hyland, R., Goosen, T.C., and Smith, P.C., Targeted quantitative proteomics for the analysis of 14 UGT1As and -2Bs in human liver using NanoUPLC-MS/MS with selected reaction monitoring. *J Proteome Res*, 2013. 12(10): p. 4402-13.
- Fallon, J.K., Smith, P.C., Xia, C.Q., and Kim, M.S., Quantification of Four Efflux Drug Transporters in Liver and Kidney Across Species Using Targeted Quantitative Proteomics by Isotope Dilution NanoLC-MS/MS. *Pharm Res*, 2016. 33(9): p. 2280-8.
- Filipski, K.K., Mechanic, L.E., Long, R., and Freedman, A.N., Pharmacogenomics in oncology care. *Front Genet*, 2014. 5: p. 73.
- Flens, M.J., Zaman, G.J., van der Valk, P., Izquierdo, M.A., Schroeijers, A.B., Scheffer, G.L., van der Groep, P., de Haas, M., Meijer, C.J., and Scheper, R.J., Tissue distribution of the multidrug resistance protein. *Am J Pathol*, 1996. 148(4): p. 1237-47.
- Formelli, F., Barua, A.B., and Olson, J.A., Bioactivities of N-(4-hydroxyphenyl) retinamide and retinoyl beta-glucuronide. *FASEB J*, 1996. 10(9): p. 1014-24.
- Formica, V., Palmirotta, R., Del Monte, G., Savonarola, A., Ludovici, G., De Marchis, M.L., Grenga, I., Schirru, M., Guadagni, F., and Roselli, M., Predictive value of VEGF gene polymorphisms for metastatic colorectal cancer patients receiving first-

- line treatment including fluorouracil, irinotecan, and bevacizumab. *Int J Colorectal Dis*, 2011. 26(2): p. 143-51.
- Fowler, S., Kletzl, H., Finel, M., Manevski, N., Schmid, P., Tuerck, D., Norcross, R.D., Hoener, M.C., Spleiss, O., and Iglesias, V.A., A UGT2B10 splicing polymorphism common in african populations may greatly increase drug exposure. *J Pharmacol Exp Ther*, 2015. 352(2): p. 358-67.
- Frances, B., Gout, R., Monsarrat, B., Cros, J., and Zajac, J.M., Further evidence that morphine-6 beta-glucuronide is a more potent opioid agonist than morphine. *J Pharmacol Exp Ther*, 1992. 262(1): p. 25-31.
- Francois, E., Berdah, J.F., Chamorey, E., Lesbats, G., Teissier, E., Codoul, J.F., Badetti, J.L., Hebert, C., and Mari, V., Use of the folinic acid/5-fluorouracil/irinotecan (FOLFIRI 1) regimen in elderly patients as a first-line treatment for metastatic colorectal cancer: a Phase II study. *Cancer Chemother Pharmacol*, 2008. 62(6): p. 931-6.
- Fuchs, C.S., Marshall, J., Mitchell, E., Wierzbicki, R., Ganju, V., Jeffery, M., Schulz, J., Richards, D., Soufi-Mahjoubi, R., Wang, B., and Barrueco, J., Randomized, controlled trial of irinotecan plus infusional, bolus, or oral fluoropyrimidines in first-line treatment of metastatic colorectal cancer: results from the BICC-C Study. *J Clin Oncol*, 2007. 25(30): p. 4779-86.
- Fujigaki, S., Nishiumi, S., Kobayashi, T., Suzuki, M., Iemoto, T., Kojima, T., Ito, Y., Daiko, H., Kato, K., Shouji, H., Honda, K., Azuma, T., and Yoshida, M., Identification of serum biomarkers of chemoradiosensitivity in esophageal cancer via the targeted metabolomics approach. *Biomark Med*, 2018. 12(8): p. 827-840.
- Fujita, K., Kubota, Y., Ishida, H., and Sasaki, Y., Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J Gastroenterol*, 2015. 21(43): p. 12234-48.
- Fujita, K., Masuo, Y., Okumura, H., Watanabe, Y., Suzuki, H., Sunakawa, Y., Shimada, K., Kawara, K., Akiyama, Y., Kitamura, M., Kunishima, M., Sasaki, Y., and Kato, Y., Increased Plasma Concentrations of Unbound SN-38, the Active Metabolite of Irinotecan, in Cancer Patients with Severe Renal Failure. *Pharm Res*, 2016. 33(2): p. 269-82.
- Fujita, K., Sugiura, T., Okumura, H., Umeda, S., Nakamichi, N., Watanabe, Y., Suzuki, H., Sunakawa, Y., Shimada, K., Kawara, K., Sasaki, Y., and Kato, Y., Direct inhibition and down-regulation by uremic plasma components of hepatic uptake transporter for SN-38, an active metabolite of irinotecan, in humans. *Pharm Res*, 2014. 31(1): p. 204-15.
- Fujiwara, R., Yokoi, T., and Nakajima, M., Structure and Protein-Protein Interactions of Human UDP-Glucuronosyltransferases. *Front Pharmacol*, 2016. 7: p. 388.
- Fukuda, T., Goebel, J., Cox, S., Maseck, D., Zhang, K., Sherbotie, J.R., Ellis, E.N., James, L.P., Ward, R.M., and Vinks, A.A., UGT1A9, UGT2B7, and MRP2 genotypes can predict mycophenolic acid pharmacokinetic variability in pediatric kidney transplant recipients. *Ther Drug Monit*, 2012. 34(6): p. 671-9.
- Gaedigk, A., Simon, S.D., Pearce, R.E., Bradford, L.D., Kennedy, M.J., and Leeder, J.S., The CYP2D6 activity score: translating genotype information into a qualitative measure of phenotype. *Clin Pharmacol Ther*, 2008. 83(2): p. 234-42.
- Gagne, J.F., Montminy, V., Belanger, P., Journault, K., Gaucher, G., and Guillemette, C., Common human UGT1A polymorphisms and the altered metabolism of irinotecan

- active metabolite 7-ethyl-10-hydroxycamptothecin (SN-38). *Mol Pharmacol*, 2002. 62(3): p. 608-17.
- Gagnon, J.F., Bernard, O., Villeneuve, L., Tetu, B., and Guillemette, C., Irinotecan inactivation is modulated by epigenetic silencing of UGT1A1 in colon cancer. *Clin Cancer Res*, 2006. 12(6): p. 1850-8.
- Gallagher, C.J., Balliet, R.M., Sun, D., Chen, G., and Lazarus, P., Sex differences in UDP-glucuronosyltransferase 2B17 expression and activity. *Drug Metab Dispos*, 2010. 38(12): p. 2204-9.
- Gardner-Stephen, D., Heydel, J.M., Goyal, A., Lu, Y., Xie, W., Lindblom, T., Mackenzie, P., and Radominska-Pandya, A., Human PXR variants and their differential effects on the regulation of human UDP-glucuronosyltransferase gene expression. *Drug Metab Dispos*, 2004. 32(3): p. 340-7.
- Gerlinger, M., Rowan, A.J., Horswell, S., Math, M., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N.Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C.R., Nohadani, M., Eklund, A.C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, P.A., and Swanton, C., Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*, 2012. 366(10): p. 883-892.
- Gerrits, C.J., de Jonge, M.J., Schellens, J.H., Stoter, G., and Verweij, J., Topoisomerase I inhibitors: the relevance of prolonged exposure for present clinical development. *Br J Cancer*, 1997. 76(7): p. 952-62.
- Ghotbi, R., Mannheimer, B., Aklillu, E., Suda, A., Bertilsson, L., Eliasson, E., and Osby, U., Carriers of the UGT1A4 142T>G gene variant are predisposed to reduced olanzapine exposure--an impact similar to male gender or smoking in schizophrenic patients. *Eur J Clin Pharmacol*, 2010. 66(5): p. 465-74.
- Girard, H., Levesque, E., Bellemare, J., Journault, K., Caillier, B., and Guillemette, C., Genetic diversity at the UGT1 locus is amplified by a novel 3' alternative splicing mechanism leading to nine additional UGT1A proteins that act as regulators of glucuronidation activity. *Pharmacogenet Genomics*, 2007. 17(12): p. 1077-89.
- Girard, H., Villeneuve, L., Court, M.H., Fortier, L.C., Caron, P., Hao, Q., von Moltke, L.L., Greenblatt, D.J., and Guillemette, C., The novel UGT1A9 intronic I399 polymorphism appears as a predictor of 7-ethyl-10-hydroxycamptothecin glucuronidation levels in the liver. *Drug Metab Dispos*, 2006. 34(7): p. 1220-8.
- Glimelius, B., Garmo, H., Berglund, A., Fredriksson, L.A., Berglund, M., Kohnke, H., Bystrom, P., Sorbye, H., and Wadelius, M., Prediction of irinotecan and 5-fluorouracil toxicity and response in patients with advanced colorectal cancer. *Pharmacogenomics J*, 2011. 11(1): p. 61-71.
- Goey, A.K., Sissung, T.M., Peer, C.J., Trepel, J.B., Lee, M.J., Tomita, Y., Ehrlich, S., Bryla, C., Balasubramaniam, S., Piekarz, R., Steinberg, S.M., Bates, S.E., and Figg, W.D., Effects of UGT1A1 genotype on the pharmacokinetics, pharmacodynamics, and toxicities of belinostat administered by 48-hour continuous infusion in patients with cancer. *J Clin Pharmacol*, 2016. 56(4): p. 461-73.
- Gong, Q.H., Cho, J.W., Huang, T., Potter, C., Gholami, N., Basu, N.K., Kubota, S., Carvalho, S., Pennington, M.W., Owens, I.S., and Popescu, N.C., Thirteen UDPglucuronosyltransferase genes are encoded at the human UGT1 gene complex locus. *Pharmacogenetics*, 2001. 11(4): p. 357-68.

- Goto, S., Taniguchi, M., Muraoka, M., Toyoda, H., Sado, Y., Kawakita, M., and Hayashi, S., UDP-sugar transporter implicated in glycosylation and processing of Notch. *Nat Cell Biol*, 2001. 3(9): p. 816-22.
- Graziano, F., Ruzzo, A., Loupakis, F., Canestrari, E., Santini, D., Catalano, V., Bissoni, R., Torresi, U., Floriani, I., Schiavon, G., Andreoni, F., Maltese, P., Rulli, E., Humar, B., Falcone, A., Giustini, L., Tonini, G., Fontana, A., Masi, G., and Magnani, M., Pharmacogenetic profiling for cetuximab plus irinotecan therapy in patients with refractory advanced colorectal cancer. *J Clin Oncol*, 2008. 26(9): p. 1427-34.
- Gregory, P.A., Lewinsky, R.H., Gardner-Stephen, D.A., and Mackenzie, P.I., Regulation of UDP glucuronosyltransferases in the gastrointestinal tract. *Toxicol Appl Pharmacol*, 2004. 199(3): p. 354-63.
- Grinfeld, J., Gerrard, G., Alikian, M., Alonso-Dominguez, J., Ale, S., Valganon, M., Nteliopoulos, G., White, D., Marin, D., Hedgley, C., O'Brien, S., Clark, R., Goldman, J.M., Milojkovic, D., Apperley, J.F., and Foroni, L., A common novel splice variant of SLC22A1 (OCT1) is associated with impaired responses to imatinib in patients with chronic myeloid leukaemia. *Br J Haematol*, 2013. 163(5): p. 631-9.
- Groer, C., Busch, D., Patrzyk, M., Beyer, K., Busemann, A., Heidecke, C.D., Drozdik, M., Siegmund, W., and Oswald, S., Absolute protein quantification of clinically relevant cytochrome P450 enzymes and UDP-glucuronosyltransferases by mass spectrometry-based targeted proteomics. *J Pharm Biomed Anal*, 2014. 100: p. 393-401.
- Guillemette, C., Levesque, E., Harvey, M., Bellemare, J., and Menard, V., UGT genomic diversity: beyond gene duplication. *Drug Metab Rev*, 2010. 42(1): p. 24-44.
- Guillemette, C., Levesque, E., and Rouleau, M., Pharmacogenomics of human uridine diphospho-glucuronosyltransferases and clinical implications. *Clin Pharmacol Ther*, 2014. 96(3): p. 324-39.
- Guillemette, C., Millikan, R.C., Newman, B., and Housman, D.E., Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 and association with breast cancer among African Americans. *Cancer Res*, 2000. 60(4): p. 950-6.
- Haaz, M.C., Rivory, L., Riche, C., Vernillet, L., and Robert, J., Metabolism of irinotecan (CPT-11) by human hepatic microsomes: participation of cytochrome P-450 3A and drug interactions. *Cancer Res*, 1998. 58(3): p. 468-72.
- Han, J.Y., Lee, Y.S., Shin, E.S., Hwang, J.A., Nam, S., Hong, S.H., Ghang, H.Y., Kim, J.Y., Yoon, S.J., and Lee, J.S., A genome-wide association study of survival in small-cell lung cancer patients treated with irinotecan plus cisplatin chemotherapy. *Pharmacogenomics J*, 2014. 14(1): p. 20-7.
- Han, J.Y., Shin, E.S., Lee, Y.S., Ghang, H.Y., Kim, S.Y., Hwang, J.A., Kim, J.Y., and Lee, J.S., A genome-wide association study for irinotecan-related severe toxicities in patients with advanced non-small-cell lung cancer. *Pharmacogenomics J*, 2013. 13(5): p. 417-22.
- Hanioka, N., Ozawa, S., Jinno, H., Ando, M., Saito, Y., and Sawada, J., Human liver UDP-glucuronosyltransferase isoforms involved in the glucuronidation of 7-ethyl-10-hydroxycamptothecin. *Xenobiotica*, 2001. 31(10): p. 687-99.
- Harbourt, D.E., Fallon, J.K., Ito, S., Baba, T., Ritter, J.K., Glish, G.L., and Smith, P.C., Quantification of human uridine-diphosphate glucuronosyl transferase 1A isoforms

- in liver, intestine, and kidney using nanobore liquid chromatography-tandem mass spectrometry. *Anal Chem*, 2012. 84(1): p. 98-105.
- Hatfield, M.J., Tsurkan, L., Garrett, M., Shaver, T.M., Hyatt, J.L., Edwards, C.C., Hicks, L.D., and Potter, P.M., Organ-specific carboxylesterase profiling identifies the small intestine and kidney as major contributors of activation of the anticancer prodrug CPT-11. *Biochem Pharmacol*, 2011. 81(1): p. 24-31.
- Heinemann, V., von Weikersthal, L.F., Decker, T., Kiani, A., Vehling-Kaiser, U., Al-Batran, S.E., Heintges, T., Lerchenmuller, C., Kahl, C., Seipelt, G., Kullmann, F., Stauch, M., Scheithauer, W., Hielscher, J., Scholz, M., Muller, S., Link, H., Niederle, N., Rost, A., Hoffkes, H.G., Moehler, M., Lindig, R.U., Modest, D.P., Rossius, L., Kirchner, T., Jung, A., and Stintzing, S., FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. *Lancet Oncol*, 2014. 15(10): p. 1065-75.
- Hilgendorf, C., Ahlin, G., Seithel, A., Artursson, P., Ungell, A.L., and Karlsson, J., Expression of thirty-six drug transporter genes in human intestine, liver, kidney, and organotypic cell lines. *Drug Metab Dispos*, 2007. 35(8): p. 1333-40.
- Horie, N., Aiba, H., Oguro, K., Hojo, H., and Takeishi, K., Functional analysis and DNA polymorphism of the tandemly repeated sequences in the 5'-terminal regulatory region of the human gene for thymidylate synthase. *Cell Struct Funct*, 1995. 20(3): p. 191-7.
- Hoskins, J.M., Goldberg, R.M., Qu, P., Ibrahim, J.G., and McLeod, H.L., UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters. *J Natl Cancer Inst*, 2007. 99(17): p. 1290-5.
- Hu, D.G., Meech, R., McKinnon, R.A., and Mackenzie, P.I., Transcriptional regulation of human UDP-glucuronosyltransferase genes. *Drug Metab Rev*, 2014. 46(4): p. 421-58.
- Hu, R., Dunn, T.A., Wei, S., Isharwal, S., Veltri, R.W., Humphreys, E., Han, M., Partin, A.W., Vessella, R.L., Isaacs, W.B., Bova, G.S., and Luo, J., Ligand-independent androgen receptor variants derived from splicing of cryptic exons signify hormone-refractory prostate cancer. *Cancer Res*, 2009. 69(1): p. 16-22.
- Huang, L., Zhang, T., Xie, C., Liao, X., Yu, Q., Feng, J., Ma, H., Dai, J., Li, M., Chen, J., Zang, A., Wang, Q., Ge, S., Qin, K., Cai, J., and Yuan, X., SLCO1B1 and SLC19A1 gene variants and irinotecan-induced rapid response and survival: a prospective multicenter pharmacogenetics study of metastatic colorectal cancer. *PLoS One*, 2013. 8(10): p. e77223.
- Huang, W., Zhang, J., Chua, S.S., Qatanani, M., Han, Y., Granata, R., and Moore, D.D., Induction of bilirubin clearance by the constitutive androstane receptor (CAR). *Proc Natl Acad Sci U S A*, 2003. 100(7): p. 4156-61.
- Humerickhouse, R., Lohrbach, K., Li, L., Bosron, W.F., and Dolan, M.E., Characterization of CPT-11 hydrolysis by human liver carboxylesterase isoforms hCE-1 and hCE-2. *Cancer Res*, 2000. 60(5): p. 1189-92.
- Ichikawa, S., Sakiyama, H., Suzuki, G., Hidari, K.I., and Hirabayashi, Y., Expression cloning of a cDNA for human ceramide glucosyltransferase that catalyzes the first glycosylation step of glycosphingolipid synthesis. *Proc Natl Acad Sci U S A*, 1996. 93(10): p. 4638-43.

- Innocenti, F., Kroetz, D.L., Schuetz, E., Dolan, M.E., Ramirez, J., Relling, M., Chen, P., Das, S., Rosner, G.L., and Ratain, M.J., Comprehensive pharmacogenetic analysis of irinotecan neutropenia and pharmacokinetics. *J Clin Oncol*, 2009. 27(16): p. 2604-14.
- Iyer, L., Das, S., Janisch, L., Wen, M., Ramirez, J., Karrison, T., Fleming, G.F., Vokes, E.E., Schilsky, R.L., and Ratain, M.J., UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J*, 2002. 2(1): p. 43-7.
- Izukawa, T., Nakajima, M., Fujiwara, R., Yamanaka, H., Fukami, T., Takamiya, M., Aoki, Y., Ikushiro, S., Sakaki, T., and Yokoi, T., Quantitative analysis of UDP-glucuronosyltransferase (UGT) 1A and UGT2B expression levels in human livers. *Drug Metab Dispos*, 2009. 37(8): p. 1759-68.
- Jackson, N.A., Barrueco, J., Soufi-Mahjoubi, R., Marshall, J., Mitchell, E., Zhang, X., and Meyerhardt, J., Comparing safety and efficacy of first-line irinotecan/fluoropyrimidine combinations in elderly versus nonelderly patients with metastatic colorectal cancer: findings from the bolus, infusional, or capecitabine with camptostar-celecoxib study. *Cancer*, 2009. 115(12): p. 2617-29.
- Jada, S.R., Lim, R., Wong, C.I., Shu, X., Lee, S.C., Zhou, Q., Goh, B.C., and Chowbay, B., Role of UGT1A1*6, UGT1A1*28 and ABCG2 c.421C>A polymorphisms in irinotecan-induced neutropenia in Asian cancer patients. *Cancer Sci*, 2007. 98(9): p. 1461-7.
- Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., Bras, J.M., Schymick, J.C., Hernandez, D.G., Traynor, B.J., Simon-Sanchez, J., Matarin, M., Britton, A., van de Leemput, J., Rafferty, I., Bucan, M., Cann, H.M., Hardy, J.A., Rosenberg, N.A., and Singleton, A.B., Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*, 2008. 451(7181): p. 998-1003.
- Janakiram, N.B. and Rao, C.V., The role of inflammation in colon cancer. *Adv Exp Med Biol*, 2014. 816: p. 25-52.
- Johnson, A.D., Kavousi, M., Smith, A.V., Chen, M.H., Dehghan, A., Aspelund, T., Lin, J.P., van Duijn, C.M., Harris, T.B., Cupples, L.A., Uitterlinden, A.G., Launer, L., Hofman, A., Rivadeneira, F., Stricker, B., Yang, Q., O'Donnell, C.J., Gudnason, V., and Witteman, J.C., Genome-wide association meta-analysis for total serum bilirubin levels. *Hum Mol Genet*, 2009. 18(14): p. 2700-10.
- Johnson, J.A., Caudle, K.E., Gong, L., Whirl-Carrillo, M., Stein, C.M., Scott, S.A., Lee, M.T., Gage, B.F., Kimmel, S.E., Perera, M.A., Anderson, J.L., Pirmohamed, M., Klein, T.E., Limdi, N.A., Cavallari, L.H., and Wadelius, M., Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline for Pharmacogenetics-Guided Warfarin Dosing: 2017 Update. *Clin Pharmacol Ther*, 2017. 102(3): p. 397-404.
- Kaneda, S., Takeishi, K., Ayusawa, D., Shimizu, K., Seno, T., and Altman, S., Role in translation of a triple tandemly repeated sequence in the 5'-untranslated region of human thymidylate synthase mRNA. *Nucleic Acids Res*, 1987. 15(3): p. 1259-70.
- Kapoor, R., Tan-Koi, W.C., and Teo, Y.Y., Role of pharmacogenetics in public health and clinical health care: a SWOT analysis. *Eur J Hum Genet*, 2016. 24(12): p. 1651-1657.
- Karypidis, A.H., Olsson, M., Andersson, S.O., Rane, A., and Ekstrom, L., Deletion polymorphism of the UGT2B17 gene is associated with increased risk for prostate

- cancer and correlated to gene expression in the prostate. *Pharmacogenomics J*, 2008. 8(2): p. 147-51.
- Kato, Y., Izukawa, T., Oda, S., Fukami, T., Finel, M., Yokoi, T., and Nakajima, M., Human UDP-glucuronosyltransferase (UGT) 2B10 in drug N-glucuronidation: substrate screening and comparison with UGT1A3 and UGT1A4. *Drug Metab Dispos*, 2013. 41(7): p. 1389-97.
- Kawakami, K. and Watanabe, G., Identification and functional analysis of single nucleotide polymorphism in the tandem repeat sequence of thymidylate synthase gene. *Cancer Res*, 2003. 63(18): p. 6004-7.
- Kawato, Y., Aonuma, M., Hirota, Y., Kuga, H., and Sato, K., Intracellular roles of SN-38, a metabolite of the camptothecin derivative CPT-11, in the antitumor effect of CPT-11. *Cancer Res*, 1991. 51(16): p. 4187-91.
- Kerdpin, O., Mackenzie, P.I., Bowalgaha, K., Finel, M., and Miners, J.O., Influence of N-terminal domain histidine and proline residues on the substrate selectivities of human UDP-glucuronosyltransferase 1A1, 1A6, 1A9, 2B7, and 2B10. *Drug Metab Dispos*, 2009. 37(9): p. 1948-55.
- Kim, J.C., Kim, S.Y., Cho, D.H., Roh, S.A., Choi, E.Y., Jo, Y.K., Jung, S.H., Na, Y.S., Kim, T.W., and Kim, Y.S., Genome-wide identification of chemosensitive single nucleotide polymorphism markers in colorectal cancers. *Cancer Sci*, 2010. 101(4): p. 1007-13.
- Kim, S.Y., Y, S.H., E, K.S., Kong, S.Y., Shin, A., Baek, J.Y., and Jung, K.H., S-1 plus irinotecan and oxaliplatin for the first-line treatment of patients with metastatic colorectal cancer: a prospective phase II study and pharmacogenetic analysis. *Br J Cancer*, 2013. 109(6): p. 1420-7.
- Kobayashi, T., Sleeman, J.E., Coughtrie, M.W., and Burchell, B., Molecular and functional characterization of microsomal UDP-glucuronic acid uptake by members of the nucleotide sugar transporter (NST) family. *Biochem J*, 2006. 400(2): p. 281-9.
- Koch, I., Weil, R., Wolbold, R., Brockmoller, J., Hustert, E., Burk, O., Nuessler, A., Neuhaus, P., Eichelbaum, M., Zanger, U., and Wojnowski, L., Interindividual variability and tissue-specificity in the expression of cytochrome P450 3A mRNA. *Drug Metab Dispos*, 2002. 30(10): p. 1108-14.
- Koutras, A.K., Antonacopoulou, A.G., Eleftheraki, A.G., Dimitrakopoulos, F.I., Koumariou, A., Varthalitis, I., Fostira, F., Sgouros, J., Briasoulis, E., Bournakis, E., Bafaloukos, D., Bompolaki, I., Galani, E., Kalogeras, K.T., Pectasides, D., Fountzilas, G., and Kalofonos, H.P., Vascular endothelial growth factor polymorphisms and clinical outcome in colorectal cancer patients treated with irinotecan-based chemotherapy and bevacizumab. *Pharmacogenomics J*, 2012. 12(6): p. 468-75.
- Kuehl, P., Zhang, J., Lin, Y., Lamba, J., Assem, M., Schuetz, J., Watkins, P.B., Daly, A., Wrighton, S.A., Hall, S.D., Maurel, P., Relling, M., Brimer, C., Yasuda, K., Venkataramanan, R., Strom, S., Thummel, K., Boguski, M.S., and Schuetz, E., Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat Genet*, 2001. 27(4): p. 383-91.
- Kuypers, D.R., Naesens, M., Vermeire, S., and Vanrenterghem, Y., The impact of uridine diphosphate-glucuronosyltransferase 1A9 (UGT1A9) gene promoter region single-nucleotide polymorphisms T-275A and C-2152T on early mycophenolic acid dose-

- interval exposure in de novo renal allograft recipients. *Clin Pharmacol Ther*, 2005. 78(4): p. 351-61.
- Kweekel, D.M., Koopman, M., Antonini, N.F., Van der Straaten, T., Nortier, J.W., Gelderblom, H., Punt, C.J., and Guchelaar, H.J., GSTP1 Ile105Val polymorphism correlates with progression-free survival in MCRC patients treated with or without irinotecan: a study of the Dutch Colorectal Cancer Group. *Br J Cancer*, 2008. 99(8): p. 1316-21.
- Laakkonen, L. and Finel, M., A molecular model of the human UDP-glucuronosyltransferase 1A1, its membrane orientation, and the interactions between different parts of the enzyme. *Mol Pharmacol*, 2010. 77(6): p. 931-9.
- Laberge, R.M., Karwatsky, J., Lincoln, M.C., Leimanis, M.L., and Georges, E., Modulation of GSH levels in ABCC1 expressing tumor cells triggers apoptosis through oxidative stress. *Biochem Pharmacol*, 2007. 73(11): p. 1727-37.
- Langer, R., Specht, K., Becker, K., Ewald, P., Bekesch, M., Sarbia, M., Busch, R., Feith, M., Stein, H.J., Siewert, J.R., and Hofler, H., Association of pretherapeutic expression of chemotherapy-related genes with response to neoadjuvant chemotherapy in Barrett carcinoma. *Clin Cancer Res*, 2005. 11(20): p. 7462-9.
- Lee, A.M., Shi, Q., Pavey, E., Alberts, S.R., Sargent, D.J., Sinicrope, F.A., Berenberg, J.L., Goldberg, R.M., and Diasio, R.B., DPYD variants as predictors of 5-fluorouracil toxicity in adjuvant colon cancer treatment (NCCTG N0147). *J Natl Cancer Inst*, 2014. 106(12).
- Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T., and Kliever, S.A., The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *J Clin Invest*, 1998. 102(5): p. 1016-23.
- Levesque, E., Belanger, A.S., Harvey, M., Couture, F., Jonker, D., Innocenti, F., Cecchin, E., Toffoli, G., and Guillemette, C., Refining the UGT1A haplotype associated with irinotecan-induced hematological toxicity in metastatic colorectal cancer patients treated with 5-fluorouracil/irinotecan-based regimens. *J Pharmacol Exp Ther*, 2013. 345(1): p. 95-101.
- Levesque, E., Delage, R., Benoit-Biancamano, M.O., Caron, P., Bernard, O., Couture, F., and Guillemette, C., The impact of UGT1A8, UGT1A9, and UGT2B7 genetic polymorphisms on the pharmacokinetic profile of mycophenolic acid after a single oral dose in healthy volunteers. *Clin Pharmacol Ther*, 2007a. 81(3): p. 392-400.
- Levesque, E., Girard, H., Journault, K., Lepine, J., and Guillemette, C., Regulation of the UGT1A1 bilirubin-conjugating pathway: role of a new splicing event at the UGT1A locus. *Hepatology*, 2007b. 45(1): p. 128-38.
- Levesque, E., Turgeon, D., Carrier, J.S., Montminy, V., Beaulieu, M., and Belanger, A., Isolation and characterization of the UGT2B28 cDNA encoding a novel human steroid conjugating UDP-glucuronosyltransferase. *Biochemistry*, 2001. 40(13): p. 3869-81.
- Levi, F., Karaboue, A., Saffroy, R., Desterke, C., Boige, V., Smith, D., Hebbar, M., Innominato, P., Taieb, J., Carvalho, C., Guimbaud, R., Focan, C., Bouchahda, M., Adam, R., Ducreux, M., Milano, G., and Lemoine, A., Pharmacogenetic determinants of outcomes on triplet hepatic artery infusion and intravenous cetuximab for liver metastases from colorectal cancer (European trial OPTILIV, NCT00852228). *Br J Cancer*, 2017. 117(7): p. 965-973.

- Li, H., Wang, Z., Tang, K., Zhou, H., Liu, H., Yan, L., Guan, W., Chen, K., Xu, H., and Ye, Z., Prognostic Value of Androgen Receptor Splice Variant 7 in the Treatment of Castration-resistant Prostate Cancer with Next generation Androgen Receptor Signal Inhibition: A Systematic Review and Meta-analysis. *Eur Urol Focus*, 2018. 4(4): p. 529-539.
- Li, N., Zhang, Y., Hua, F., and Lai, Y., Absolute difference of hepatobiliary transporter multidrug resistance-associated protein (MRP2/Mrp2) in liver tissues and isolated hepatocytes from rat, dog, monkey, and human. *Drug Metab Dispos*, 2009. 37(1): p. 66-73.
- Li, Q. and Shu, Y., Role of solute carriers in response to anticancer drugs. *Mol Cell Ther*, 2014. 2: p. 15.
- Li, T. and Chiang, J.Y., Mechanism of rifampicin and pregnane X receptor inhibition of human cholesterol 7 alpha-hydroxylase gene transcription. *Am J Physiol Gastrointest Liver Physiol*, 2005. 288(1): p. G74-84.
- Lianidou, E.S., Gene expression profiling and DNA methylation analyses of CTCs. *Mol Oncol*, 2016. 10(3): p. 431-42.
- Lim, Y.P., Kuo, S.C., Lai, M.L., and Huang, J.D., Inhibition of CYP3A4 expression by ketoconazole is mediated by the disruption of pregnane X receptor, steroid receptor coactivator-1, and hepatocyte nuclear factor 4alpha interaction. *Pharmacogenet Genomics*, 2009. 19(1): p. 11-24.
- Lindqvist, L.M., Tandoc, K., Topisirovic, I., and Furic, L., Cross-talk between protein synthesis, energy metabolism and autophagy in cancer. *Curr Opin Genet Dev*, 2018. 48: p. 104-111.
- Liu, C.Y., Chen, P.M., Chiou, T.J., Liu, J.H., Lin, J.K., Lin, T.C., Chen, W.S., Jiang, J.K., Wang, H.S., and Wang, W.S., UGT1A1*28 polymorphism predicts irinotecan-induced severe toxicities without affecting treatment outcome and survival in patients with metastatic colorectal carcinoma. *Cancer*, 2008. 112(9): p. 1932-40.
- Liu, D., Li, J., Gao, J., Li, Y., Yang, R., and Shen, L., Examination of multiple UGT1A and DPYD polymorphisms has limited ability to predict the toxicity and efficacy of metastatic colorectal cancer treated with irinotecan-based chemotherapy: a retrospective analysis. *BMC Cancer*, 2017. 17(1): p. 437.
- Liu, X., Cheng, D., Kuang, Q., Liu, G., and Xu, W., Association between UGT1A1*28 polymorphisms and clinical outcomes of irinotecan-based chemotherapies in colorectal cancer: a meta-analysis in Caucasians. *PLoS One*, 2013. 8(3): p. e58489.
- Llanos, S. and Serrano, M., Depletion of ribosomal protein L37 occurs in response to DNA damage and activates p53 through the L11/MDM2 pathway. *Cell Cycle*, 2010. 9(19): p. 4005-12.
- Locke, D.P., Sharp, A.J., McCarroll, S.A., McGrath, S.D., Newman, T.L., Cheng, Z., Schwartz, S., Albertson, D.G., Pinkel, D., Altshuler, D.M., and Eichler, E.E., Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. *Am J Hum Genet*, 2006. 79(2): p. 275-90.
- Lokiec, F., du Sorbier, B.M., and Sanderink, G.J., Irinotecan (CPT-11) metabolites in human bile and urine. *Clin Cancer Res*, 1996. 2(12): p. 1943-9.
- Loupakis, F., Cremolini, C., Yang, D., Salvatore, L., Zhang, W., Wakatsuki, T., Bohanes, P., Schirripa, M., Benhaim, L., Lonardi, S., Antoniotti, C., Aprile, G., Graziano, F., Ruzzo, A., Lucchesi, S., Ronzoni, M., De Vita, F., Tonini, G., Falcone, A., and Lenz, H.J., Prospective validation of candidate SNPs of VEGF/VEGFR pathway in

- metastatic colorectal cancer patients treated with first-line FOLFIRI plus bevacizumab. *PLoS One*, 2013. 8(7): p. e66774.
- Loupakis, F., Ruzzo, A., Salvatore, L., Cremolini, C., Masi, G., Frumento, P., Schirripa, M., Catalano, V., Galluccio, N., Canestrari, E., Vincenzi, B., Santini, D., Bencardino, K., Ricci, V., Manzoni, M., Danova, M., Tonini, G., Magnani, M., Falcone, A., and Graziano, F., Retrospective exploratory analysis of VEGF polymorphisms in the prediction of benefit from first-line FOLFIRI plus bevacizumab in metastatic colorectal cancer. *BMC Cancer*, 2011. 11: p. 247.
- Lu, C.Y., Huang, C.W., Wu, I.C., Tsai, H.L., Ma, C.J., Yeh, Y.S., Chang, S.F., Huang, M.L., and Wang, J.Y., Clinical Implication of UGT1A1 Promoter Polymorphism for Irinotecan Dose Escalation in Metastatic Colorectal Cancer Patients Treated with Bevacizumab Combined with FOLFIRI in the First-line Setting. *Transl Oncol*, 2015. 8(6): p. 474-9.
- Lugowski, A., Nicholson, B., and Rissland, O.S., Determining mRNA half-lives on a transcriptome-wide scale. *Methods*, 2018. 137: p. 90-98.
- Luo, F.R., Paranjpe, P.V., Guo, A., Rubin, E., and Sinko, P., Intestinal transport of irinotecan in Caco-2 cells and MDCK II cells overexpressing efflux transporters Pgp, cMOAT, and MRP1. *Drug Metab Dispos*, 2002a. 30(7): p. 763-70.
- Luo, G., Cunningham, M., Kim, S., Burn, T., Lin, J., Sinz, M., Hamilton, G., Rizzo, C., Jolley, S., Gilbert, D., Downey, A., Mudra, D., Graham, R., Carroll, K., Xie, J., Madan, A., Parkinson, A., Christ, D., Selling, B., LeCluyse, E., and Gan, L.S., CYP3A4 induction by drugs: correlation between a pregnane X receptor reporter gene assay and CYP3A4 expression in human hepatocytes. *Drug Metab Dispos*, 2002b. 30(7): p. 795-804.
- Mackenzie, P.I., Bock, K.W., Burchell, B., Guillemette, C., Ikushiro, S., Iyanagi, T., Miners, J.O., Owens, I.S., and Nebert, D.W., Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily. *Pharmacogenet Genomics*, 2005. 15(10): p. 677-85.
- MacKenzie, P.I., Rogers, A., Elliot, D.J., Chau, N., Hulin, J.A., Miners, J.O., and Meech, R., The novel UDP glycosyltransferase 3A2: cloning, catalytic properties, and tissue distribution. *Mol Pharmacol*, 2011. 79(3): p. 472-8.
- Mackenzie, P.I., Rogers, A., Treloar, J., Jorgensen, B.R., Miners, J.O., and Meech, R., Identification of UDP glycosyltransferase 3A1 as a UDP N-acetylglucosaminyltransferase. *J Biol Chem*, 2008. 283(52): p. 36205-10.
- Maiello, E., Gebbia, V., Giuliani, F., Paoletti, G., Gebbia, N., Borsellino, N., Carteni, G., Pezzella, G., Manzione, L., Romito, S., Lopez, M., and Colucci, G., FOLFIRI regimen in advanced colorectal cancer: the experience of the Gruppo Oncologico dell'Italia Meridionale (GOIM). *Ann Oncol*, 2005. 16 Suppl 4: p. iv56-60.
- Mandola, M.V., Stoehlmacher, J., Muller-Weeks, S., Cesarone, G., Yu, M.C., Lenz, H.J., and Ladner, R.D., A novel single nucleotide polymorphism within the 5' tandem repeat polymorphism of the thymidylate synthase gene abolishes USF-1 binding and alters transcriptional activity. *Cancer Res*, 2003. 63(11): p. 2898-904.
- Marcuello, E., Altes, A., Menoyo, A., Del Rio, E., Gomez-Pardo, M., and Baiget, M., UGT1A1 gene variations and irinotecan treatment in patients with metastatic colorectal cancer. *Br J Cancer*, 2004. 91(4): p. 678-82.
- Marcuello, E., Paez, D., Pare, L., Salazar, J., Sebino, A., del Rio, E., and Baiget, M., A genotype-directed phase I-IV dose-finding study of irinotecan in combination with

- fluorouracil/leucovorin as first-line treatment in advanced colorectal cancer. *Br J Cancer*, 2011. 105(1): p. 53-7.
- Margaillan, G., Levesque, E., and Guillemette, C., Epigenetic regulation of steroid inactivating UDP-glucuronosyltransferases by microRNAs in prostate cancer. *J Steroid Biochem Mol Biol*, 2016. 155(Pt A): p. 85-93.
- Margaillan, G., Rouleau, M., Fallon, J.K., Caron, P., Villeneuve, L., Turcotte, V., Smith, P.C., Joy, M.S., and Guillemette, C., Quantitative profiling of human renal UDP-glucuronosyltransferases and glucuronidation activity: a comparison of normal and tumoral kidney tissues. *Drug Metab Dispos*, 2015a. 43(4): p. 611-9.
- Margaillan, G., Rouleau, M., Klein, K., Fallon, J.K., Caron, P., Villeneuve, L., Smith, P.C., Zanger, U.M., and Guillemette, C., Multiplexed Targeted Quantitative Proteomics Predicts Hepatic Glucuronidation Potential. *Drug Metab Dispos*, 2015b. 43(9): p. 1331-5.
- Markowski, M.C., Silberstein, J.L., Eshleman, J.R., Eisenberger, M.A., Luo, J., and Antonarakis, E.S., Clinical Utility of CLIA-Grade AR-V7 Testing in Patients With Metastatic Castration-Resistant Prostate Cancer. *JCO Precis Oncol*, 2017. 2017.
- Mathijssen, R.H., van Alphen, R.J., Verweij, J., Loos, W.J., Nooter, K., Stoter, G., and Sparreboom, A., Clinical pharmacokinetics and metabolism of irinotecan (CPT-11). *Clin Cancer Res*, 2001. 7(8): p. 2182-94.
- Matreyek, K.A., Starita, L.M., Stephany, J.J., Martin, B., Chiasson, M.A., Gray, V.E., Kircher, M., Khechaduri, A., Dines, J.N., Hause, R.J., Bhatia, S., Evans, W.E., Relling, M.V., Yang, W., Shendure, J., and Fowler, D.M., Multiplex assessment of protein variant abundance by massively parallel sequencing. *Nat Genet*, 2018. 50(6): p. 874-882.
- Matsusaka, S., Cao, S., Hanna, D.L., Sunakawa, Y., Ueno, M., Mizunuma, N., Zhang, W., Yang, D., Ning, Y., Stintzing, S., Sebio, A., Stremitzer, S., Yamauchi, S., Parekh, A., Okazaki, S., Berger, M.D., El-Khoueiry, R., Mendez, A., Ichikawa, W., Loupakis, F., and Lenz, H.J., CXCR4 polymorphism predicts progression-free survival in metastatic colorectal cancer patients treated with first-line bevacizumab-based chemotherapy. *Pharmacogenomics J*, 2017. 17(6): p. 543-550.
- Matsusaka, S., Hanna, D.L., Cao, S., Zhang, W., Yang, D., Ning, Y., Sunakawa, Y., Okazaki, S., Berger, M.D., Miyamoto, Y., Parekh, A., Stintzing, S., Loupakis, F., and Lenz, H.J., Prognostic Impact of IL6 Genetic Variants in Patients with Metastatic Colorectal Cancer Treated with Bevacizumab-Based Chemotherapy. *Clin Cancer Res*, 2016. 22(13): p. 3218-26.
- Mbatchi, L.C., Brouillet, J.P., and Evrard, A., Genetic variations of the xenoreceptors NR1I2 and NR1I3 and their effect on drug disposition and response variability. *Pharmacogenomics*, 2018. 19(1): p. 61-77.
- Mbatchi, L.C., Robert, J., Ychou, M., Boyer, J.C., Del Rio, M., Gassiot, M., Thomas, F., Tubiana, N., and Evrard, A., Effect of Single Nucleotide Polymorphisms in the Xenobiotic-sensing Receptors NR1I2 and NR1I3 on the Pharmacokinetics and Toxicity of Irinotecan in Colorectal Cancer Patients. *Clin Pharmacokinet*, 2016. 55(9): p. 1145-57.
- McCarroll, S.A., Hadnott, T.N., Perry, G.H., Sabeti, P.C., Zody, M.C., Barrett, J.C., Dallaire, S., Gabriel, S.B., Lee, C., Daly, M.J., and Altshuler, D.M., Common deletion polymorphisms in the human genome. *Nat Genet*, 2006. 38(1): p. 86-92.

- McDonnell, A.M. and Dang, C.H., Basic review of the cytochrome p450 system. *J Adv Pract Oncol*, 2013. 4(4): p. 263-8.
- McLean, K.J., Luciakova, D., Belcher, J., Tee, K.L., and Munro, A.W., Biological diversity of cytochrome P450 redox partner systems. *Adv Exp Med Biol*, 2015. 851: p. 299-317.
- Meech, R. and Mackenzie, P.I., Determinants of UDP glucuronosyltransferase membrane association and residency in the endoplasmic reticulum. *Arch Biochem Biophys*, 1998. 356(1): p. 77-85.
- Menard, V., Collin, P., Margaillan, G., and Guillemette, C., Modulation of the UGT2B7 enzyme activity by C-terminally truncated proteins derived from alternative splicing. *Drug Metab Dispos*, 2013a. 41(12): p. 2197-205.
- Menard, V., Eap, O., Harvey, M., Guillemette, C., and Levesque, E., Copy-number variations (CNVs) of the human sex steroid metabolizing genes UGT2B17 and UGT2B28 and their associations with a UGT2B15 functional polymorphism. *Hum Mutat*, 2009. 30(9): p. 1310-9.
- Menard, V., Eap, O., Roberge, J., Harvey, M., Levesque, E., and Guillemette, C., Transcriptional diversity at the UGT2B7 locus is dictated by extensive pre-mRNA splicing mechanisms that give rise to multiple mRNA splice variants. *Pharmacogenet Genomics*, 2011. 21(10): p. 631-41.
- Menard, V., Levesque, E., Chen, S., Eap, O., Joy, M.S., Ekstrom, L., Rane, A., and Guillemette, C., Expression of UGT2B7 is driven by two mutually exclusive promoters and alternative splicing in human tissues: changes from prenatal life to adulthood and in kidney cancer. *Pharmacogenet Genomics*, 2013b. 23(12): p. 684-96.
- Meulendijks, D., Henricks, L.M., Sonke, G.S., Deenen, M.J., Froehlich, T.K., Amstutz, U., Largiader, C.R., Jennings, B.A., Marinaki, A.M., Sanderson, J.D., Kleibl, Z., Kleiblova, P., Schwab, M., Zanger, U.M., Palles, C., Tomlinson, I., Gross, E., van Kuilenburg, A.B., Punt, C.J., Koopman, M., Beijnen, J.H., Cats, A., and Schellens, J.H., Clinical relevance of DPYD variants c.1679T>G, c.1236G>A/HapB3, and c.1601G>A as predictors of severe fluoropyrimidine-associated toxicity: a systematic review and meta-analysis of individual patient data. *Lancet Oncol*, 2015. 16(16): p. 1639-50.
- Miley, M.J., Zielinska, A.K., Keenan, J.E., Bratton, S.M., Radomska-Pandya, A., and Redinbo, M.R., Crystal structure of the cofactor-binding domain of the human phase II drug-metabolism enzyme UDP-glucuronosyltransferase 2B7. *J Mol Biol*, 2007. 369(2): p. 498-511.
- Miller, D.A. and DeLuca, H.F., Biosynthesis of retinoyl-beta-glucuronide, a biologically active metabolite of all-trans-retinoic acid. *Arch Biochem Biophys*, 1986. 244(1): p. 179-86.
- Miura, K., Fujibuchi, W., and Unno, M., Splice isoforms as therapeutic targets for colorectal cancer. *Carcinogenesis*, 2012. 33(12): p. 2311-9.
- Modest, D.P., Pant, S., and Sartore-Bianchi, A., Treatment sequencing in metastatic colorectal cancer. *Eur J Cancer*, 2019. 109: p. 70-83.
- Monaghan, G., Ryan, M., Seddon, R., Hume, R., and Burchell, B., Genetic variation in bilirubin UPD-glucuronosyltransferase gene promoter and Gilbert's syndrome. *Lancet*, 1996. 347(9001): p. 578-81.

- Montgomery, R.B., Mostaghel, E.A., Vessella, R., Hess, D.L., Kalhorn, T.F., Higano, C.S., True, L.D., and Nelson, P.S., Maintenance of intratumoral androgens in metastatic prostate cancer: a mechanism for castration-resistant tumor growth. *Cancer Res*, 2008. 68(11): p. 4447-54.
- Moreau, P., Masszi, T., Grzasko, N., Bahlis, N.J., Hansson, M., Pour, L., Sandhu, I., Ganly, P., Baker, B.W., Jackson, S.R., Stoppa, A.M., Simpson, D.R., Gimsing, P., Palumbo, A., Garderet, L., Cavo, M., Kumar, S., Touzeau, C., Buadi, F.K., Laubach, J.P., Berg, D.T., Lin, J., Di Bacco, A., Hui, A.M., van de Velde, H., and Richardson, P.G., Oral Ixazomib, Lenalidomide, and Dexamethasone for Multiple Myeloma. *N Engl J Med*, 2016. 374(17): p. 1621-34.
- Moscovitz, J.E., Kalgutkar, A.S., Nulick, K., Johnson, N., Lin, Z., Goosen, T.C., and Weng, Y., Establishing Transcriptional Signatures to Differentiate PXR-, CAR-, and AhR-Mediated Regulation of Drug Metabolism and Transport Genes in Cryopreserved Human Hepatocytes. *J Pharmacol Exp Ther*, 2018. 365(2): p. 262-271.
- Murphy, S.E., Park, S.S., Thompson, E.F., Wilkens, L.R., Patel, Y., Stram, D.O., and Le Marchand, L., Nicotine N-glucuronidation relative to N-oxidation and C-oxidation and UGT2B10 genotype in five ethnic/racial groups. *Carcinogenesis*, 2014. 35(11): p. 2526-33.
- Murphy, S.E., von Weymarn, L.B., Parenteau, M., Stepanov, I., Tiirikainen, M., LeMarchand, L., and Park, S.L., Influence of UGT2B10 Genotype on Urinary Excretion of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol- N-glucuronide by African American Smokers. *Chem Res Toxicol*, 2018. 31(3): p. 168-175.
- Nadeau, G., Bellemare, J., Audet-Walsh, E., Flageole, C., Huang, S.P., Bao, B.Y., Douville, P., Caron, P., Fradet, Y., Lacombe, L., Guillemette, C., and Levesque, E., Deletions of the androgen-metabolizing UGT2B genes have an effect on circulating steroid levels and biochemical recurrence after radical prostatectomy in localized prostate cancer. *J Clin Endocrinol Metab*, 2011. 96(9): p. E1550-7.
- Nakamura, A., Nakajima, M., Yamanaka, H., Fujiwara, R., and Yokoi, T., Expression of UGT1A and UGT2B mRNA in human normal tissues and various cell lines. *Drug Metab Dispos*, 2008. 36(8): p. 1461-4.
- Nakatomi, K., Yoshikawa, M., Oka, M., Ikegami, Y., Hayasaka, S., Sano, K., Shiozawa, K., Kawabata, S., Soda, H., Ishikawa, T., Tanabe, S., and Kohno, S., Transport of 7-ethyl-10-hydroxycamptothecin (SN-38) by breast cancer resistance protein ABCG2 in human lung cancer cells. *Biochem Biophys Res Commun*, 2001. 288(4): p. 827-32.
- Nelson, D.R., Zeldin, D.C., Hoffman, S.M., Maltais, L.J., Wain, H.M., and Nebert, D.W., Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. *Pharmacogenetics*, 2004. 14(1): p. 1-18.
- Nie, Y.L., Meng, X.G., Liu, J.Y., Yan, L., Wang, P., Bi, H.Z., Kan, Q.C., and Zhang, L.R., Histone Modifications Regulate the Developmental Expression of Human Hepatic UDP-Glucuronosyltransferase 1A1. *Drug Metab Dispos*, 2017. 45(12): p. 1372-1378.
- Nigam, S.K., What do drug transporters really do? *Nat Rev Drug Discov*, 2015. 14(1): p. 29-44.

- Nozawa, T., Minami, H., Sugiura, S., Tsuji, A., and Tamai, I., Role of organic anion transporter OATP1B1 (OATP-C) in hepatic uptake of irinotecan and its active metabolite, 7-ethyl-10-hydroxycamptothecin: in vitro evidence and effect of single nucleotide polymorphisms. *Drug Metab Dispos*, 2005. 33(3): p. 434-9.
- Oda, S., Fukami, T., Yokoi, T., and Nakajima, M., Epigenetic regulation is a crucial factor in the repression of UGT1A1 expression in the human kidney. *Drug Metab Dispos*, 2013. 41(10): p. 1738-43.
- Oda, S., Fukami, T., Yokoi, T., and Nakajima, M., Epigenetic regulation of the tissue-specific expression of human UDP-glucuronosyltransferase (UGT) 1A10. *Biochem Pharmacol*, 2014. 87(4): p. 660-7.
- Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., Volkert, T.L., Schreiber, J., Rolfe, P.A., Gifford, D.K., Fraenkel, E., Bell, G.I., and Young, R.A., Control of pancreas and liver gene expression by HNF transcription factors. *Science*, 2004. 303(5662): p. 1378-81.
- Oguri, T., Takahashi, T., Miyazaki, M., Isobe, T., Kohno, N., Mackenzie, P.I., and Fujiwara, Y., UGT1A10 is responsible for SN-38 glucuronidation and its expression in human lung cancers. *Anticancer Res*, 2004. 24(5A): p. 2893-6.
- Ohno, S. and Nakajin, S., Determination of mRNA expression of human UDP-glucuronosyltransferases and application for localization in various human tissues by real-time reverse transcriptase-polymerase chain reaction. *Drug Metab Dispos*, 2009. 37(1): p. 32-40.
- Ohno, S. and Nakajin, S., Quantitative analysis of UGT2B28 mRNA expression by real-time RT-PCR and application to human tissue distribution study. *Drug Metab Lett*, 2011. 5(3): p. 202-8.
- Ohtsuki, S., Schaefer, O., Kawakami, H., Inoue, T., Liehner, S., Saito, A., Ishiguro, N., Kishimoto, W., Ludwig-Schwellinger, E., Ebner, T., and Terasaki, T., Simultaneous absolute protein quantification of transporters, cytochromes P450, and UDP-glucuronosyltransferases as a novel approach for the characterization of individual human liver: comparison with mRNA levels and activities. *Drug Metab Dispos*, 2012. 40(1): p. 83-92.
- Okazaki, S., Loupakis, F., Stintzing, S., Cao, S., Zhang, W., Yang, D., Ning, Y., Sunakawa, Y., Stremtzer, S., Matsusaka, S., Berger, M.D., Parekh, A., West, J.D., Miyamoto, Y., Suenaga, M., Schirripa, M., Cremolini, C., Falcone, A., Heinemann, V., DePaolo, R.W., and Lenz, H.J., Clinical Significance of TLR1 I602S Polymorphism for Patients with Metastatic Colorectal Cancer Treated with FOLFIRI plus Bevacizumab. *Mol Cancer Ther*, 2016. 15(7): p. 1740-5.
- Okuyama, Y., Hazama, S., Nozawa, H., Kobayashi, M., Takahashi, K., Fujikawa, K., Kato, T., Nagata, N., Kimura, H., Oba, K., Sakamoto, J., and Mishima, H., Prospective phase II study of FOLFIRI for mCRC in Japan, including the analysis of UGT1A1 28/6 polymorphisms. *Jpn J Clin Oncol*, 2011. 41(4): p. 477-82.
- Oltean, S. and Bates, D.O., Hallmarks of alternative splicing in cancer. *Oncogene*, 2014. 33(46): p. 5311-8.
- Osborne, M.J., Coutinho de Oliveira, L., Volpon, L., and Borden, K.L.B., Backbone assignment of the apo-form of the human C-terminal domain of UDP-glucuronosyltransferase 1A (UGT1A). *Biomol NMR Assign*, 2018.
- Osborne, M.J., Coutinho de Oliveira, L., Volpon, L., Zahreddine, H.A., and Borden, K.L.B., Overcoming Drug Resistance through the Development of Selective

- Inhibitors of UDP-Glucuronosyltransferase Enzymes. *J Mol Biol*, 2019. 431(2): p. 258-272.
- Paez, D., Tobena, M., Fernandez-Plana, J., Sebio, A., Virgili, A.C., Cirera, L., Barnadas, A., Riera, P., Sullivan, I., and Salazar, J., Pharmacogenetic clinical randomised phase II trial to evaluate the efficacy and safety of FOLFIRI with high-dose irinotecan (HD-FOLFIRI) in metastatic colorectal cancer patients according to their UGT1A 1 genotype. *Br J Cancer*, 2019. 120(2): p. 190-195.
- Palomares, T., Carames, M., Garcia-Alonso, I., and Alonso-Varona, A., Glutathione modulation reverses the growth-promoting effect of growth factors, improving the 5-fluorouracil antitumour response in WiDr colon cancer cells. *Anticancer Res*, 2009. 29(10): p. 3957-65.
- Pantel, K. and Alix-Panabieres, C., Liquid biopsy in 2016: Circulating tumour cells and cell-free DNA in gastrointestinal cancer. *Nat Rev Gastroenterol Hepatol*, 2017. 14(2): p. 73-74.
- Paoluzzi, L., Singh, A.S., Price, D.K., Danesi, R., Mathijssen, R.H., Verweij, J., Figg, W.D., and Sparreboom, A., Influence of genetic variants in UGT1A1 and UGT1A9 on the in vivo glucuronidation of SN-38. *J Clin Pharmacol*, 2004. 44(8): p. 854-60.
- Papageorgiou, I. and Court, M.H., Identification and validation of microRNAs directly regulating the UDP-glucuronosyltransferase 1A subfamily enzymes by a functional genomics approach. *Biochem Pharmacol*, 2017a. 137: p. 93-106.
- Papageorgiou, I. and Court, M.H., Identification and validation of the microRNA response elements in the 3'-untranslated region of the UDP glucuronosyltransferase (UGT) 2B7 and 2B15 genes by a functional genomics approach. *Biochem Pharmacol*, 2017b. 146: p. 199-213.
- Park, J., Chen, L., Ratnashinge, L., Sellers, T.A., Tanner, J.P., Lee, J.H., Dossett, N., Lang, N., Kadlubar, F.F., Ambrosone, C.B., Zachariah, B., Heysek, R.V., Patterson, S., and Pow-Sang, J., Deletion polymorphism of UDP-glucuronosyltransferase 2B17 and risk of prostate cancer in African American and Caucasian men. *Cancer Epidemiol Biomarkers Prev*, 2006. 15(8): p. 1473-8.
- Park, J.W., Reed, J.R., Brignac-Huber, L.M., and Backes, W.L., Cytochrome P450 system proteins reside in different regions of the endoplasmic reticulum. *Biochem J*, 2014. 464(2): p. 241-9.
- Paschalis, A., Sharp, A., Welti, J.C., Neeb, A., Raj, G.V., Luo, J., Plymate, S.R., and de Bono, J.S., Alternative splicing in prostate cancer. *Nat Rev Clin Oncol*, 2018. 15(11): p. 663-675.
- Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C., Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol Cell*, 2007. 27(6): p. 928-37.
- Prakash, C., Zuniga, B., Song, C.S., Jiang, S., Cropper, J., Park, S., and Chatterjee, B., Nuclear Receptors in Drug Metabolism, Drug Response and Drug Interactions. *Nucl Receptor Res*, 2015. 2.
- Radominska-Pandya, A., Bratton, S.M., Redinbo, M.R., and Miley, M.J., 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*, 2010. 42(1): p. 133-44.

- Radomska-Pandya, A., Czernik, P.J., Little, J.M., Battaglia, E., and Mackenzie, P.I., Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab Rev*, 1999. 31(4): p. 817-99.
- Ramchandani, R.P., Wang, Y., Booth, B.P., Ibrahim, A., Johnson, J.R., Rahman, A., Mehta, M., Innocenti, F., Ratain, M.J., and Gobburu, J.V., The role of SN-38 exposure, UGT1A1*28 polymorphism, and baseline bilirubin level in predicting severe irinotecan toxicity. *J Clin Pharmacol*, 2007. 47(1): p. 78-86.
- Raynal, C., Pascussi, J.M., Leguelinel, G., Breuker, C., Kantar, J., Lallemand, B., Poujol, S., Bonnans, C., Joubert, D., Hollande, F., Lumbroso, S., Brouillet, J.P., and Evrard, A., Pregnane X Receptor (PXR) expression in colorectal cancer cells restricts irinotecan chemosensitivity through enhanced SN-38 glucuronidation. *Mol Cancer*, 2010. 9: p. 46.
- Reis, E.S., Mastellos, D.C., Ricklin, D., Mantovani, A., and Lambris, J.D., Complement in cancer: untangling an intricate relationship. *Nat Rev Immunol*, 2018. 18(1): p. 5-18.
- Relling, M.V. and Evans, W.E., Pharmacogenomics in the clinic. *Nature*, 2015. 526(7573): p. 343-50.
- Richardson, P.G., Sonneveld, P., Schuster, M.W., Irwin, D., Stadtmauer, E.A., Facon, T., Harousseau, J.L., Ben-Yehuda, D., Lonial, S., Goldschmidt, H., Reece, D., San-Miguel, J.F., Blade, J., Boccadoro, M., Cavenagh, J., Dalton, W.S., Boral, A.L., Esseltine, D.L., Porter, J.B., Schenkein, D., and Anderson, K.C., Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N Engl J Med*, 2005. 352(24): p. 2487-98.
- Ritter, J.K., Chen, F., Sheen, Y.Y., Tran, H.M., Kimura, S., Yeatman, M.T., and Owens, I.S., A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J Biol Chem*, 1992. 267(5): p. 3257-61.
- Rivory, L.P., Riou, J.F., Haaz, M.C., Sable, S., Vuilhorgne, M., Commercon, A., Pond, S.M., and Robert, J., Identification and properties of a major plasma metabolite of irinotecan (CPT-11) isolated from the plasma of patients. *Cancer Res*, 1996. 56(16): p. 3689-94.
- Roelofsen, H., Vos, T.A., Schippers, I.J., Kuipers, F., Koning, H., Moshage, H., Jansen, P.L., and Muller, M., Increased levels of the multidrug resistance protein in lateral membranes of proliferating hepatocyte-derived cells. *Gastroenterology*, 1997. 112(2): p. 511-21.
- Rollin, J., Payance, A., Gouilleux-Gruart, V., Boisdron-Celle, M., Azzopardi, N., Morel, A., Gruel, Y., Paintaud, G., Gamelin, E., Watier, H., and Lecomte, T., Significant effect of VEGFA polymorphisms on the clinical outcome of metastatic colorectal cancer patients treated with FOLFIRI-cetuximab. *Pharmacogenomics*, 2015. 16(18): p. 2035-43.
- Ross, M.K., Borazjani, A., Wang, R., Crow, J.A., and Xie, S., Examination of the carboxylesterase phenotype in human liver. *Arch Biochem Biophys*, 2012. 522(1): p. 44-56.
- Rouleau, M., Collin, P., Bellemare, J., Harvey, M., and Guillemette, C., Protein-protein interactions between the bilirubin-conjugating UDP-glucuronosyltransferase UGT1A1 and its shorter isoform 2 regulatory partner derived from alternative splicing. *Biochem J*, 2013a. 450(1): p. 107-14.

- Rouleau, M., Roberge, J., Bellemare, J., and Guillemette, C., Dual roles for splice variants of the glucuronidation pathway as regulators of cellular metabolism. *Mol Pharmacol*, 2014. 85(1): p. 29-36.
- Rouleau, M., Roberge, J., Falardeau, S.A., Villeneuve, L., and Guillemette, C., The relative protein abundance of UGT1A alternative splice variants as a key determinant of glucuronidation activity in vitro. *Drug Metab Dispos*, 2013b. 41(4): p. 694-7.
- Rouleau, M., Tourancheau, A., Girard-Bock, C., Villeneuve, L., Vaucher, J., Duperre, A.M., Audet-Delage, Y., Gilbert, I., Popa, I., Droit, A., and Guillemette, C., Divergent Expression and Metabolic Functions of Human Glucuronosyltransferases through Alternative Splicing. *Cell Rep*, 2016. 17(1): p. 114-124.
- Rowland, A., Mackenzie, P.I., and Miners, J.O., Transporter-mediated uptake of UDP-glucuronic acid by human liver microsomes: assay conditions, kinetics, and inhibition. *Drug Metab Dispos*, 2015. 43(1): p. 147-53.
- Rowland, A., Miners, J.O., and Mackenzie, P.I., The UDP-glucuronosyltransferases: their role in drug metabolism and detoxification. *Int J Biochem Cell Biol*, 2013. 45(6): p. 1121-32.
- Ruzzo, A., Graziano, F., Loupakis, F., Santini, D., Catalano, V., Bissoni, R., Ficarelli, R., Fontana, A., Andreoni, F., Falcone, A., Canestrari, E., Tonini, G., Mari, D., Lippe, P., Pizzagalli, F., Schiavon, G., Alessandrini, P., Giustini, L., Maltese, P., Testa, E., Menichetti, E.T., and Magnani, M., Pharmacogenetic profiling in patients with advanced colorectal cancer treated with first-line FOLFIRI chemotherapy. *Pharmacogenomics J*, 2008. 8(4): p. 278-88.
- Ryan, C.J., Smith, M.R., de Bono, J.S., Molina, A., Logothetis, C.J., de Souza, P., Fizazi, K., Mainwaring, P., Piulats, J.M., Ng, S., Carles, J., Mulders, P.F., Basch, E., Small, E.J., Saad, F., Schrijvers, D., Van Poppel, H., Mukherjee, S.D., Suttman, H., Gerritsen, W.R., Flaig, T.W., George, D.J., Yu, E.Y., Efsthathiou, E., Pantuck, A., Winquist, E., Higano, C.S., Taplin, M.E., Park, Y., Kheoh, T., Griffin, T., Scher, H.I., and Rathkopf, D.E., Abiraterone in metastatic prostate cancer without previous chemotherapy. *N Engl J Med*, 2013. 368(2): p. 138-48.
- Sai, K., Kaniwa, N., Ozawa, S., and Sawada, J.I., A new metabolite of irinotecan in which formation is mediated by human hepatic cytochrome P-450 3A4. *Drug Metab Dispos*, 2001. 29(11): p. 1505-13.
- Sai, K., Saito, Y., Tatewaki, N., Hosokawa, M., Kaniwa, N., Nishimaki-Mogami, T., Naito, M., Sawada, J., Shirao, K., Hamaguchi, T., Yamamoto, N., Kunitoh, H., Tamura, T., Yamada, Y., Ohe, Y., Yoshida, T., Minami, H., Ohtsu, A., Matsumura, Y., Saijo, N., and Okuda, H., Association of carboxylesterase 1A genotypes with irinotecan pharmacokinetics in Japanese cancer patients. *Br J Clin Pharmacol*, 2010. 70(2): p. 222-33.
- Saltz, L.B., Cox, J.V., Blanke, C., Rosen, L.S., Fehrenbacher, L., Moore, M.J., Maroun, J.A., Ackland, S.P., Locker, P.K., Pirotta, N., Elfring, G.L., and Miller, L.L., Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. *N Engl J Med*, 2000. 343(13): p. 905-14.
- Sanchez-Gundin, J., Fernandez-Carballido, A.M., Martinez-Valdivieso, L., Barreda-Hernandez, D., and Torres-Suarez, A.I., New Trends in the Therapeutic Approach to Metastatic Colorectal Cancer. *Int J Med Sci*, 2018. 15(7): p. 659-665.
- Sandanaraj, E., Jada, S.R., Shu, X., Lim, R., Lee, S.C., Zhou, Q., Zhou, S., Goh, B.C., and Chowbay, B., Influence of UGT1A9 intronic I399C>T polymorphism on SN-38

- glucuronidation in Asian cancer patients. *Pharmacogenomics J*, 2008. 8(3): p. 174-85.
- Sandhu, J., Lavingia, V., and Fakih, M., Systemic treatment for metastatic colorectal cancer in the era of precision medicine. *J Surg Oncol*, 2019. 119(5): p. 564-582.
- Sanghani, S.P., Quinney, S.K., Fredenburg, T.B., Davis, W.I., Murry, D.J., and Bosron, W.F., Hydrolysis of irinotecan and its oxidative metabolites, 7-ethyl-10-[4-N-(5-aminopentanoic acid)-1-piperidino] carbonyloxycamptothecin and 7-ethyl-10-[4-(1-piperidino)-1-amino]-carbonyloxycamptothecin, by human carboxylesterases CES1A1, CES2, and a newly expressed carboxylesterase isoenzyme, CES3. *Drug Metab Dispos*, 2004. 32(5): p. 505-11.
- Santoro, A., Comandone, A., Rimassa, L., Granetti, C., Lorusso, V., Oliva, C., Ronzoni, M., Siena, S., Zuradelli, M., Mari, E., Pressiani, T., and Carnaghi, C., A phase II randomized multicenter trial of gefitinib plus FOLFIRI and FOLFIRI alone in patients with metastatic colorectal cancer. *Ann Oncol*, 2008. 19(11): p. 1888-93.
- Santos, A., Zanetta, S., Cresteil, T., Deroussent, A., Pein, F., Raymond, E., Vernillet, L., Risse, M.L., Boige, V., Gouyette, A., and Vassal, G., Metabolism of irinotecan (CPT-11) by CYP3A4 and CYP3A5 in humans. *Clin Cancer Res*, 2000. 6(5): p. 2012-20.
- Sato, Y., Matsuda, S., Maruyama, A., Nakayama, J., Miyashita, T., Udagawa, H., Umemura, S., Yanagihara, K., Ochiai, A., Tomita, M., Soga, T., Tsuchihara, K., and Makinoshima, H., Metabolic Characterization of Antifolate Responsiveness and Non-responsiveness in Malignant Pleural Mesothelioma Cells. *Front Pharmacol*, 2018. 9: p. 1129.
- Sato, Y., Nagata, M., Tetsuka, K., Tamura, K., Miyashita, A., Kawamura, A., and Usui, T., Optimized methods for targeted peptide-based quantification of human uridine 5'-diphosphate-glucuronosyltransferases in biological specimens using liquid chromatography-tandem mass spectrometry. *Drug Metab Dispos*, 2014. 42(5): p. 885-9.
- Satoh, T., Hosokawa, M., Atsumi, R., Suzuki, W., Hokusui, H., and Nagai, E., Metabolic activation of CPT-11, 7-ethyl-10-[4-(1-piperidino)-1-piperidino]carbonyloxycamptothecin, a novel antitumor agent, by carboxylesterase. *Biol Pharm Bull*, 1994. 17(5): p. 662-4.
- Scher, H.I., Lu, D., Schreiber, N.A., Louw, J., Graf, R.P., Vargas, H.A., Johnson, A., Jendrisak, A., Bambury, R., Danila, D., McLaughlin, B., Wahl, J., Greene, S.B., Heller, G., Marrinucci, D., Fleisher, M., and Dittamore, R., Association of AR-V7 on Circulating Tumor Cells as a Treatment-Specific Biomarker With Outcomes and Survival in Castration-Resistant Prostate Cancer. *JAMA Oncol*, 2016. 2(11): p. 1441-1449.
- Sebio, A., Paez, D., Salazar, J., Berenguer-Llargo, A., Pare-Brunet, L., Lasa, A., Del Rio, E., Tobena, M., Martin-Richard, M., Baiget, M., and Barnadas, A., Intergenic polymorphisms in the amphiregulin gene region as biomarkers in metastatic colorectal cancer patients treated with anti-EGFR plus irinotecan. *Pharmacogenomics J*, 2014. 14(3): p. 256-62.
- Sebio, A., Stintzing, S., Heinemann, V., Sunakawa, Y., Zhang, W., Ichikawa, W., Tsuji, A., Takahashi, T., Parek, A., Yang, D., Cao, S., Ning, Y., Stremtizer, S., Matsusaka, S., Okazaki, S., Barzi, A., Berger, M.D., and Lenz, H.J., A genetic variant in *Rassf1a* predicts outcome in mCRC patients treated with cetuximab plus chemotherapy:

- results from FIRE-3 and JACCRO 05 and 06 trials. *Pharmacogenomics J*, 2018. 18(1): p. 43-48.
- Shafi, A.A., Putluri, V., Arnold, J.M., Tsouko, E., Maity, S., Roberts, J.M., Coarfa, C., Frigo, D.E., Putluri, N., Sreekumar, A., and Weigel, N.L., Differential regulation of metabolic pathways by androgen receptor (AR) and its constitutively active splice variant, AR-V7, in prostate cancer cells. *Oncotarget*, 2015. 6(31): p. 31997-2012.
- Shulman, K., Cohen, I., Barnett-Griness, O., Kuten, A., Gruber, S.B., Lejbkowitz, F., and Rennert, G., Clinical implications of UGT1A1*28 genotype testing in colorectal cancer patients. *Cancer*, 2011. 117(14): p. 3156-62.
- Siegfried, Z. and Karni, R., The role of alternative splicing in cancer drug resistance. *Curr Opin Genet Dev*, 2018. 48: p. 16-21.
- Silvestris, N., Simone, G., Partipilo, G., Scarpi, E., Lorusso, V., Brunetti, A.E., Maiello, E., Paradiso, A., and Mangia, A., CES2, ABCG2, TS and Topo-I primary and synchronous metastasis expression and clinical outcome in metastatic colorectal cancer patients treated with first-line FOLFIRI regimen. *Int J Mol Sci*, 2014. 15(9): p. 15767-77.
- Skof, E., Rebersek, M., Hlebanja, Z., and Ocvirk, J., Capecitabine plus Irinotecan (XELIRI regimen) compared to 5-FU/LV plus Irinotecan (FOLFIRI regimen) as neoadjuvant treatment for patients with unresectable liver-only metastases of metastatic colorectal cancer: a randomised prospective phase II trial. *BMC Cancer*, 2009. 9: p. 120.
- Smith, N.F., Figg, W.D., and Sparreboom, A., Pharmacogenetics of irinotecan metabolism and transport: an update. *Toxicol In Vitro*, 2006. 20(2): p. 163-75.
- Sneitz, N., Court, M.H., Zhang, X., Laajanen, K., Yee, K.K., Dalton, P., Ding, X., and Finel, M., Human UDP-glucuronosyltransferase UGT2A2: cDNA construction, expression, and functional characterization in comparison with UGT2A1 and UGT2A3. *Pharmacogenet Genomics*, 2009. 19(12): p. 923-34.
- Souglakos, J., Androulakis, N., Syrigos, K., Polyzos, A., Ziras, N., Athanasiadis, A., Kakolyris, S., Tsousis, S., Kouroussis, C., Vamvakas, L., Kalykaki, A., Samonis, G., Mavroudis, D., and Georgoulis, V., FOLFOXIRI (folinic acid, 5-fluorouracil, oxaliplatin and irinotecan) vs FOLFIRI (folinic acid, 5-fluorouracil and irinotecan) as first-line treatment in metastatic colorectal cancer (MCC): a multicentre randomised phase III trial from the Hellenic Oncology Research Group (HORG). *Br J Cancer*, 2006. 94(6): p. 798-805.
- Souglakos, J., Pallis, A., Kakolyris, S., Mavroudis, D., Androulakis, N., Kouroussis, C., Agelaki, S., Xenidis, N., Milaki, G., and Georgoulis, V., Combination of irinotecan (CPT-11) plus 5-fluorouracil and leucovorin (FOLFIRI regimen) as first line treatment for elderly patients with metastatic colorectal cancer: a phase II trial. *Oncology*, 2005. 69(5): p. 384-90.
- Sprowl, J.A. and Sparreboom, A., Uptake carriers and oncology drug safety. *Drug Metab Dispos*, 2014. 42(4): p. 611-22.
- Stingl, J.C., Bartels, H., Viviani, R., Lehmann, M.L., and Brockmoller, J., Relevance of UDP-glucuronosyltransferase polymorphisms for drug dosing: A quantitative systematic review. *Pharmacol Ther*, 2014. 141(1): p. 92-116.
- Stintzing, S., Zhang, W., Heinemann, V., Neureiter, D., Kemmerling, R., Kirchner, T., Jung, A., Folwaczny, M., Yang, D., Ning, Y., Sebio, A., Stremitzer, S., Sunakawa, Y., Matsusaka, S., Yamauchi, S., Loupakis, F., Cremolini, C., Falcone, A., and

- Lenz, H.J., Polymorphisms in Genes Involved in EGFR Turnover Are Predictive for Cetuximab Efficacy in Colorectal Cancer. *Mol Cancer Ther*, 2015. 14(10): p. 2374-81.
- Strassburg, C.P., Oldhafer, K., Manns, M.P., and Tukey, R.H., Differential expression of the UGT1A locus in human liver, biliary, and gastric tissue: identification of UGT1A7 and UGT1A10 transcripts in extrahepatic tissue. *Mol Pharmacol*, 1997. 52(2): p. 212-20.
- Strassburg, C.P., Strassburg, A., Kneip, S., Barut, A., Tukey, R.H., Rodeck, B., and Manns, M.P., Developmental aspects of human hepatic drug glucuronidation in young children and adults. *Gut*, 2002. 50(2): p. 259-65.
- Stringer, F., DeJongh, J., Scott, G., and Danhof, M., A model-based approach to analyze the influence of UGT2B15 polymorphism driven pharmacokinetic differences on the pharmacodynamic response of the PPAR agonist sipoglitazar. *J Clin Pharmacol*, 2014. 54(4): p. 453-61.
- Stringer, F., Ploeger, B.A., DeJongh, J., Scott, G., Urquhart, R., Karim, A., and Danhof, M., Evaluation of the impact of UGT polymorphism on the pharmacokinetics and pharmacodynamics of the novel PPAR agonist sipoglitazar. *J Clin Pharmacol*, 2013a. 53(3): p. 256-63.
- Stringer, F., Scott, G., Valbuena, M., Kinley, J., Nishihara, M., and Urquhart, R., The effect of genetic polymorphisms in UGT2B15 on the pharmacokinetic profile of sipoglitazar, a novel anti-diabetic agent. *Eur J Clin Pharmacol*, 2013b. 69(3): p. 423-30.
- Suenaga, M., Schirripa, M., Cao, S., Zhang, W., Yang, D., Cremolini, C., Lonardi, S., Bergamo, F., Ning, Y., Yamamoto, N., Okazaki, S., Berger, M.D., Miyamoto, Y., Gopez, R., Jr., Barzi, A., Yamaguchi, T., Stintzing, S., Heinemann, V., Loupakis, F., Falcone, A., and Lenz, H.J., Potential role of PIN1 genotypes in predicting benefit from oxaliplatin-based and irinotecan-based treatment in patients with metastatic colorectal cancer. *Pharmacogenomics J*, 2018. 18(5): p. 623-632.
- Suenaga, M., Stintzing, S., Cao, S., Zhang, W., Yang, D., Ning, Y., Okazaki, S., Berger, M.D., Miyamoto, Y., Schirripa, M., Soni, S., Barzi, A., Heinemann, V., and Lenz, H.J., Role of CCL5 and CCR5 gene polymorphisms in epidermal growth factor receptor signalling blockade in metastatic colorectal cancer: analysis of the FIRE-3 trial. *Eur J Cancer*, 2019. 107: p. 100-114.
- Sugatani, J., Uchida, T., Kurosawa, M., Yamaguchi, M., Yamazaki, Y., Ikari, A., and Miwa, M., Regulation of pregnane X receptor (PXR) function and UGT1A1 gene expression by posttranslational modification of PXR protein. *Drug Metab Dispos*, 2012. 40(10): p. 2031-40.
- Sunakawa, Y., Stintzing, S., Cao, S., Heinemann, V., Cremolini, C., Falcone, A., Yang, D., Zhang, W., Ning, Y., Stremtzer, S., Matsusaka, S., Yamauchi, S., Parekh, A., Okazaki, S., Berger, M.D., Graver, S., Mendez, A., Scherer, S.J., Loupakis, F., and Lenz, H.J., Variations in genes regulating tumor-associated macrophages (TAMs) to predict outcomes of bevacizumab-based treatment in patients with metastatic colorectal cancer: results from TRIBE and FIRE3 trials. *Ann Oncol*, 2015. 26(12): p. 2450-6.
- Tagen, M., Zhuang, Y., Zhang, F., Harstead, K.E., Shen, J., Schaiquevich, P., Fraga, C.H., Panetta, J.C., Waters, C.M., and Stewart, C.F., P-glycoprotein, but not multidrug

- resistance protein 4, plays a role in the systemic clearance of irinotecan and SN-38 in mice. *Drug Metab Lett*, 2010. 4(4): p. 195-201.
- Takahashi, H., Sai, K., Saito, Y., Kaniwa, N., Matsumura, Y., Hamaguchi, T., Shimada, Y., Ohtsu, A., Yoshino, T., Doi, T., Okuda, H., Ichinohe, R., Takahashi, A., Doi, A., Odaka, Y., Okuyama, M., Saijo, N., Sawada, J., Sakamoto, H., and Yoshida, T., Application of a combination of a knowledge-based algorithm and 2-stage screening to hypothesis-free genomic data on irinotecan-treated patients for identification of a candidate single nucleotide polymorphism related to an adverse effect. *PLoS One*, 2014. 9(8): p. e105160.
- Tatsumi, N., Tokumitsu, S., Nakano, M., Fukami, T., and Nakajima, M., miR-141-3p commonly regulates human UGT1A isoforms via different mechanisms. *Drug Metab Pharmacokinet*, 2018. 33(4): p. 203-210.
- Teasdale, R.D. and Jackson, M.R., Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. *Annu Rev Cell Dev Biol*, 1996. 12: p. 27-54.
- Teft, W.A., Welch, S., Lenehan, J., Parfitt, J., Choi, Y.H., Winquist, E., and Kim, R.B., OATP1B1 and tumour OATP1B3 modulate exposure, toxicity, and survival after irinotecan-based chemotherapy. *Br J Cancer*, 2015. 112(5): p. 857-65.
- Teufel, A., Steinmann, S., Siebler, J., Zanke, C., Hohl, H., Adami, B., Schroeder, M., Klein, O., Hohler, T., Galle, P.R., Heike, M., and Moehler, M., Irinotecan plus folinic acid/continuous 5-fluorouracil as simplified bimonthly FOLFIRI regimen for first-line therapy of metastatic colorectal cancer. *BMC Cancer*, 2004. 4: p. 38.
- The GTEx Consortium, The Genotype-Tissue Expression (GTEx) project. *Nat Genet*, 2013. 45(6): p. 580-5.
- Toffoli, G., Cecchin, E., Corona, G., Russo, A., Buonadonna, A., D'Andrea, M., Pasetto, L.M., Pessa, S., Errante, D., De Pangher, V., Giusto, M., Medici, M., Gaion, F., Sandri, P., Galligioni, E., Bonura, S., Boccalon, M., Bion, P., and Frustaci, S., The role of UGT1A1*28 polymorphism in the pharmacodynamics and pharmacokinetics of irinotecan in patients with metastatic colorectal cancer. *J Clin Oncol*, 2006. 24(19): p. 3061-8.
- Toffoli, G., Cecchin, E., Gasparini, G., D'Andrea, M., Azzarello, G., Basso, U., Mini, E., Pessa, S., De Mattia, E., Lo Re, G., Buonadonna, A., Nobili, S., De Paoli, P., and Innocenti, F., Genotype-driven phase I study of irinotecan administered in combination with fluorouracil/leucovorin in patients with metastatic colorectal cancer. *J Clin Oncol*, 2010. 28(5): p. 866-71.
- Toffoli, G., Giodini, L., Buonadonna, A., Berretta, M., De Paoli, A., Scalone, S., Miolo, G., Mini, E., Nobili, S., Lonardi, S., Pella, N., Lo Re, G., Montico, M., Roncato, R., Dreussi, E., Gagno, S., and Cecchin, E., Clinical validity of a DPYD-based pharmacogenetic test to predict severe toxicity to fluoropyrimidines. *Int J Cancer*, 2015. 137(12): p. 2971-80.
- Toffoli, G., Sharma, M.R., Marangon, E., Posocco, B., Gray, E., Mai, Q., Buonadonna, A., Polite, B.N., Miolo, G., Tabaro, G., and Innocenti, F., Genotype-Guided Dosing Study of FOLFIRI plus Bevacizumab in Patients with Metastatic Colorectal Cancer. *Clin Cancer Res*, 2017. 23(4): p. 918-924.
- Togawa, H., Shinkai, S., and Mizutani, T., Induction of human UGT1A1 by bilirubin through AhR dependent pathway. *Drug Metab Lett*, 2008. 2(4): p. 231-7.

- Tokunaga, R., Cao, S., Naseem, M., Battaglin, F., Lo, J.H., Arai, H., Loupakis, F., Stintzing, S., Puccini, A., Berger, M.D., Soni, S., Zhang, W., Mancao, C., Salhia, B., Mumenthaler, S.M., Weisenberger, D.J., Liang, G., Cremolini, C., Heinemann, V., Falcone, A., Millstein, J., and Lenz, H.J., AMPK variant, a candidate of novel predictor for chemotherapy in metastatic colorectal cancer: A meta-analysis using TRIBE, MAVERICC and FIRE3. *Int J Cancer*, 2019a.
- Tokunaga, R., Cao, S., Naseem, M., Lo, J.H., Battaglin, F., Puccini, A., Berger, M.D., Soni, S., Millstein, J., Zhang, W., Stintzing, S., Loupakis, F., Cremolini, C., Heinemann, V., Falcone, A., and Lenz, H.J., Prognostic Effect of Adenosine-related Genetic Variants in Metastatic Colorectal Cancer Treated With Bevacizumab-based Chemotherapy. *Clin Colorectal Cancer*, 2019b. 18(1): p. e8-e19.
- Tourancheau, A., Margaillan, G., Rouleau, M., Gilbert, I., Villeneuve, L., Levesque, E., Droit, A., and Guillemette, C., Unravelling the transcriptomic landscape of the major phase II UDP-glucuronosyltransferase drug metabolizing pathway using targeted RNA sequencing. *Pharmacogenomics J*, 2016. 16(1): p. 60-70.
- Tourancheau, A., Rouleau, M., Guauque-Olarte, S., Villeneuve, L., Gilbert, I., Droit, A., and Guillemette, C., Quantitative profiling of the UGT transcriptome in human drug-metabolizing tissues. *Pharmacogenomics J*, 2018. 18(2): p. 251-261.
- Tracy, T.S., Chaudhry, A.S., Prasad, B., Thummel, K.E., Schuetz, E.G., Zhong, X.B., Tien, Y.C., Jeong, H., Pan, X., Shireman, L.M., Tay-Sontheimer, J., and Lin, Y.S., Interindividual Variability in Cytochrome P450-Mediated Drug Metabolism. *Drug Metab Dispos*, 2016. 44(3): p. 343-51.
- Troberg, J., Jarvinen, E., Ge, G.B., Yang, L., and Finel, M., UGT1A10 Is a High Activity and Important Extrahepatic Enzyme: Why Has Its Role in Intestinal Glucuronidation Been Frequently Underestimated? *Mol Pharm*, 2017. 14(9): p. 2875-2883.
- Tucker, T.G., Milne, A.M., Fournel-Gigleux, S., Fenner, K.S., and Coughtrie, M.W., Absolute immunoquantification of the expression of ABC transporters P-glycoprotein, breast cancer resistance protein and multidrug resistance-associated protein 2 in human liver and duodenum. *Biochem Pharmacol*, 2012. 83(2): p. 279-85.
- Turgeon, D., Carrier, J.S., Levesque, E., Beatty, B.G., Belanger, A., and Hum, D.W., Isolation and characterization of the human UGT2B15 gene, localized within a cluster of UGT2B genes and pseudogenes on chromosome 4. *J Mol Biol*, 2000. 295(3): p. 489-504.
- Turgeon, D., Chouinard, S., Belanger, P., Picard, S., Labbe, J.F., Borgeat, P., and Belanger, A., Glucuronidation of arachidonic and linoleic acid metabolites by human UDP-glucuronosyltransferases. *J Lipid Res*, 2003. 44(6): p. 1182-91.
- Van Cutsem, E., Kohne, C.H., Lang, I., Folprecht, G., Nowacki, M.P., Cascinu, S., Shchepotin, I., Maurel, J., Cunningham, D., Tejpar, S., Schlichting, M., Zubel, A., Celik, I., Rougier, P., and Ciardiello, F., Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall survival according to tumor KRAS and BRAF mutation status. *J Clin Oncol*, 2011. 29(15): p. 2011-9.
- van der Wouden, C.H., Cambon-Thomsen, A., Cecchin, E., Cheung, K.C., Davila-Fajardo, C.L., Deneer, V.H., Dolzan, V., Ingelman-Sundberg, M., Jonsson, S., Karlsson, M.O., Kriek, M., Mitropoulou, C., Patrinos, G.P., Pirmohamed, M., Samwald, M.,

- Schaeffeler, E., Schwab, M., Steinberger, D., Stingl, J., Sunder-Plassmann, G., Toffoli, G., Turner, R.M., van Rhenen, M.H., Swen, J.J., and Guchelaar, H.J., Implementing Pharmacogenomics in Europe: Design and Implementation Strategy of the Ubiquitous Pharmacogenomics Consortium. *Clin Pharmacol Ther*, 2017. 101(3): p. 341-358.
- van der Wouden, C.H., Swen, J.J., Samwald, M., Mitropoulou, C., Schwab, M., and Guchelaar, H.J., A brighter future for the implementation of pharmacogenomic testing. *Eur J Hum Genet*, 2016. 24(12): p. 1658-1660.
- van Dorp, E.L., Morariu, A., and Dahan, A., Morphine-6-glucuronide: potency and safety compared with morphine. *Expert Opin Pharmacother*, 2008. 9(11): p. 1955-61.
- van Schaik, R.H., van Agteren, M., de Fijter, J.W., Hartmann, A., Schmidt, J., Budde, K., Kuypers, D., Le Meur, Y., van der Werf, M., Mamelok, R., and van Gelder, T., UGT1A9 -275T>A/-2152C>T polymorphisms correlate with low MPA exposure and acute rejection in MMF/tacrolimus-treated kidney transplant patients. *Clin Pharmacol Ther*, 2009. 86(3): p. 319-27.
- Vegran, F., Boidot, R., Oudin, C., Riedinger, J.M., Bonnetain, F., and Lizard-Nacol, S., Overexpression of caspase-3s splice variant in locally advanced breast carcinoma is associated with poor response to neoadjuvant chemotherapy. *Clin Cancer Res*, 2006. 12(19): p. 5794-800.
- Volz, N.B., Stintzing, S., Zhang, W., Yang, D., Ning, Y., Wakatsuki, T., El-Khoueiry, R.E., Li, J.E., Kardosh, A., Loupakis, F., Cremolini, C., Falcone, A., Scherer, S.J., and Lenz, H.J., Genes involved in pericyte-driven tumor maturation predict treatment benefit of first-line FOLFIRI plus bevacizumab in patients with metastatic colorectal cancer. *Pharmacogenomics J*, 2015. 15(1): p. 69-76.
- Wadelius, M., Chen, L.Y., Downes, K., Ghori, J., Hunt, S., Eriksson, N., Wallerman, O., Melhus, H., Wadelius, C., Bentley, D., and Deloukas, P., Common VKORC1 and GGCX polymorphisms associated with warfarin dose. *Pharmacogenomics J*, 2005. 5(4): p. 262-70.
- Wallace, B.D. and Redinbo, M.R., Xenobiotic-sensing nuclear receptors involved in drug metabolism: a structural perspective. *Drug Metab Rev*, 2013. 45(1): p. 79-100.
- Wang, D., Zhang, X., Zhang, Y., Wu, Y., Guan, X., Zhu, W., Wang, M., Qi, C., and Shen, B., Association of MLH1 single nucleotide polymorphisms with clinical outcomes of first-line irinotecan-based chemotherapy in colorectal cancer. *Onco Targets Ther*, 2018a. 11: p. 8083-8088.
- Wang, J., Huang, M., Hu, H., Yu, L., and Zeng, S., Pregnane X receptor-mediated transcriptional activation of UDP-glucuronosyltransferase 1A1 by natural constituents from foods and herbs. *Food Chem*, 2014. 164: p. 74-80.
- Wang, J.C., Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol*, 2002. 3(6): p. 430-40.
- Wang, P., Nie, Y.L., Wang, S.J., Yang, L.L., Yang, W.H., Li, J.F., Li, X.T., and Zhang, L.R., Regulation of UGT1A expression by miR-298 in human livers from the Han Chinese population and in human cell lines. *Epigenomics*, 2018b. 10(1): p. 43-57.
- Wang, X., Rao, Z., Qin, H., Zhang, G., Ma, Y., Jin, Y., Han, M., Shi, A., Wang, Y., and Wu, X., Effect of hesperidin on the pharmacokinetics of CPT-11 and its active metabolite SN-38 by regulating hepatic Mrp2 in rats. *Biopharm Drug Dispos*, 2016. 37(7): p. 421-432.

- Ward, L.D. and Kellis, M., HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res*, 2016. 44(D1): p. D877-81.
- Warta, R., Theile, D., Mogler, C., Herpel, E., Grabe, N., Lahrmann, B., Plinkert, P.K., Herold-Mende, C., Weiss, J., and Dyckhoff, G., Association of drug transporter expression with mortality and progression-free survival in stage IV head and neck squamous cell carcinoma. *PLoS One*, 2014. 9(9): p. e108908.
- Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M., The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet*, 2013. 45(10): p. 1113-20.
- Werck-Reichhart, D. and Feyereisen, R., Cytochromes P450: a success story. *Genome Biol*, 2000. 1(6): p. REVIEWS3003.
- Westlind-Johnsson, A., Malmebo, S., Johansson, A., Otter, C., Andersson, T.B., Johansson, I., Edwards, R.J., Boobis, A.R., and Ingelman-Sundberg, M., Comparative analysis of CYP3A expression in human liver suggests only a minor role for CYP3A5 in drug metabolism. *Drug Metab Dispos*, 2003. 31(6): p. 755-61.
- White-Gilbertson, S., Kurtz, D.T., and Voelkel-Johnson, C., The role of protein synthesis in cell cycling and cancer. *Mol Oncol*, 2009. 3(5-6): p. 402-8.
- Wijayakumara, D.D., Hu, D.G., Meech, R., McKinnon, R.A., and Mackenzie, P.I., Regulation of Human UGT2B15 and UGT2B17 by miR-376c in Prostate Cancer Cell Lines. *J Pharmacol Exp Ther*, 2015. 354(3): p. 417-25.
- Wijayakumara, D.D., Mackenzie, P.I., McKinnon, R.A., Hu, D.G., and Meech, R., Regulation of UDP-Glucuronosyltransferase 2B15 by miR-331-5p in Prostate Cancer Cells Involves Canonical and Noncanonical Target Sites. *J Pharmacol Exp Ther*, 2018. 365(1): p. 48-59.
- Wrighton, S.A., Brian, W.R., Sari, M.A., Iwasaki, M., Guengerich, F.P., Raucy, J.L., Molowa, D.T., and Vandenbranden, M., Studies on the expression and metabolic capabilities of human liver cytochrome P450III_{A5} (HLp3). *Mol Pharmacol*, 1990. 38(2): p. 207-13.
- Wu, B., Morrow, J.K., Singh, R., Zhang, S., and Hu, M., Three-dimensional quantitative structure-activity relationship studies on UGT1A9-mediated 3-O-glucuronidation of natural flavonols using a pharmacophore-based comparative molecular field analysis model. *J Pharmacol Exp Ther*, 2011. 336(2): p. 403-13.
- Xiao, L., Zhu, L., Li, W., Li, C., Cao, Y., Ge, G., and Sun, X., New Insights into SN-38 Glucuronidation: Evidence for the Important Role of UDP Glucuronosyltransferase 1A9. *Basic Clin Pharmacol Toxicol*, 2018. 122(4): p. 424-428.
- Xu, C., Tang, X., Qu, Y., Keyoumu, S., Zhou, N., and Tang, Y., UGT1A1 gene polymorphism is associated with toxicity and clinical efficacy of irinotecan-based chemotherapy in patients with advanced colorectal cancer. *Cancer Chemother Pharmacol*, 2016. 78(1): p. 119-30.
- Xue, Y., Sun, D., Daly, A., Yang, F., Zhou, X., Zhao, M., Huang, N., Zerjal, T., Lee, C., Carter, N.P., Hurles, M.E., and Tyler-Smith, C., Adaptive evolution of UGT2B17 copy-number variation. *Am J Hum Genet*, 2008. 83(3): p. 337-46.
- Yasar, U., Greenblatt, D.J., Guillemette, C., and Court, M.H., Evidence for regulation of UDP-glucuronosyltransferase (UGT) 1A1 protein expression and activity via DNA methylation in healthy human livers. *J Pharm Pharmacol*, 2013. 65(6): p. 874-83.

- Ychou, M., Hohenberger, W., Thezenas, S., Navarro, M., Maurel, J., Bokemeyer, C., Shacham-Shmueli, E., Rivera, F., Kwok-Keung Choi, C., and Santoro, A., A randomized phase III study comparing adjuvant 5-fluorouracil/folinic acid with FOLFIRI in patients following complete resection of liver metastases from colorectal cancer. *Ann Oncol*, 2009. 20(12): p. 1964-70.
- Yeh, Y.S., Tsai, H.L., Huang, C.W., Wang, J.H., Lin, Y.W., Tang, H.C., Sung, Y.C., Wu, C.C., Lu, C.Y., and Wang, J.Y., Prospective analysis of UGT1A1 promoter polymorphism for irinotecan dose escalation in metastatic colorectal cancer patients treated with bevacizumab plus FOLFIRI as the first-line setting: study protocol for a randomized controlled trial. *Trials*, 2016. 17: p. 46.
- Yi, Q. and Tang, L., Alternative spliced variants as biomarkers of colorectal cancer. *Curr Drug Metab*, 2011. 12(10): p. 966-74.
- Younis, I., Berg, M., Kaida, D., Dittmar, K., Wang, C., and Dreyfuss, G., Rapid-response splicing reporter screens identify differential regulators of constitutive and alternative splicing. *Mol Cell Biol*, 2010. 30(7): p. 1718-28.
- Yu, I.S. and Cheung, W.Y., Metastatic Colorectal Cancer in the Era of Personalized Medicine: A More Tailored Approach to Systemic Therapy. *Can J Gastroenterol Hepatol*, 2018a. 2018: p. 9450754.
- Yu, Q., Zhang, T., Xie, C., Qiu, H., Liu, B., Huang, L., Peng, P., Feng, J., Chen, J., Zang, A., and Yuan, X., UGT1A polymorphisms associated with worse outcome in colorectal cancer patients treated with irinotecan-based chemotherapy. *Cancer Chemother Pharmacol*, 2018b. 82(1): p. 87-98.
- Zanger, U.M., Fischer, J., Raimundo, S., Stuvén, T., Evert, B.O., Schwab, M., and Eichelbaum, M., Comprehensive analysis of the genetic factors determining expression and function of hepatic CYP2D6. *Pharmacogenetics*, 2001. 11(7): p. 573-85.
- Zanger, U.M. and Schwab, M., Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther*, 2013. 138(1): p. 103-41.
- Zarate, R., Rodriguez, J., Bandres, E., Patino-Garcia, A., Ponz-Sarvisé, M., Viudez, A., Ramirez, N., Bitarte, N., Chopitea, A., and Gacia-Foncillas, J., Oxaliplatin, irinotecan and capecitabine as first-line therapy in metastatic colorectal cancer (mCRC): a dose-finding study and pharmacogenomic analysis. *Br J Cancer*, 2010. 102(6): p. 987-94.
- Zduniak, K., Agrawal, A., Agrawal, S., Hossain, M.M., Ziolkowski, P., and Weber, G.F., Osteopontin splice variants are differential predictors of breast cancer treatment responses. *BMC Cancer*, 2016. 16: p. 441.
- Zhang, L., Gong, C., Lau, S.L., Yang, N., Wong, O.G., Cheung, A.N., Tsang, J.W., Chan, K.Y., and Khoo, U.S., SpliceArray profiling of breast cancer reveals a novel variant of NCOR2/SMRT that is associated with tamoxifen resistance and control of ER α transcriptional activity. *Cancer Res*, 2013. 73(1): p. 246-55.
- Zhou, X., Liao, W.J., Liao, J.M., Liao, P., and Lu, H., Ribosomal proteins: functions beyond the ribosome. *J Mol Cell Biol*, 2015. 7(2): p. 92-104.