Overexpression of uridine diphopspho glucuronosyltransferase 2B17 in high risk chronic lymphocytic leukemia

Michaela Gruber[1], Judith Bellemare[2], Gregor Hörmann[3], Andreas Gleiss[4], Edit Porpaczy[1], Martin Bilban[3], Trang Le[1], Sabine Zehetmayer[4], Christine Mannhalter[3], Alexander Gaiger[1], Medhat Shehata[1], Karin Fleiss[1], Cathrin Skrabs[1], Éric Lévesque[2], Katrina Vanura[1], Chantal Guillemette[2], Ulrich Jäger[1]

[1] Comprehensive Cancer Center, Department of Internal Medicine I, Division of Hematology, Medical University of Vienna, Vienna, Austria

[2] Pharmacogenomics Laboratory, CHUQ Research Center and Faculty of Pharmacy, Laval University, Québec, Canada

[3] Department of Laboratory Medicine, Medical University of Vienna, Vienna, Austria

[4] Center for Medical Statistics, Informatics and Intelligent Systems, Medical University of Vienna, Vienna, Austria

Keywords: UGT2B17, vorinostat glucuronidation, chronic lymphocytic leukemia, prognostic markers, uridine diphospho glucuronosyltransferase

Running head: *UGT2B17* overexpression in chronic lymphocytic leukemia

Word count:
Abstract: 194/200
Text: 3524/5000
Figures: 4 (paper)/7 + 4 (supplementary)
Tables: 3 (paper)/unlimited + 1 (supplementary)
References: 41/50

Corresponding author:
Ulrich Jaeger
Medical University of Vienna
Department of Internal Medicine I, Division of Hematology
Waehringer Guertel 18-20
1090 Vienna, Austria
phone: +43 1 40400 4409
e-mail: ulrich.jaeger@meduniwien.ac.at

Abstract

Uridine diphospho glucuronosyltransferase 2B17 (UGT2B17) glucuronidates androgens and xenobiotics including certain drugs. The *UGT2B17* gene is affected by a remarkable copy number variation (CNV) predisposing for solid tumors and influencing drug response.

Here, we identify a yet unobserved *UGT2B17* mRNA over-expression in poor risk chronic lymphocytic leukemia (CLL). 320 CLL patients and 449 healthy donors were analyzed in total. High *UGT2B17* was associated with established CLL prognostic factors and resulted in shorter treatment free and overall survival (HR (death) for high *UGT2B17*=2.18; 95%CI 1.18-4.01; p=0.013). The *UGT2B17* mRNA levels directly related to functional glucuronidation activity toward androgens and the anticancer drug vorinostat (R>0.9, p<0.001). *In vivo, UGT2B17* was up-regulated after treatment with fludarabine containing regimens particularly in poor responders (p=0.03).

Our data also show that the impact of mRNA levels strongly overcame the role of *UGT2B17* CNV and indicate an exclusive involvement of the 2B17 isoform within the UGT family. Gene expression profiling of a stable *UGT2B17* knock down in the cell line MEC-1 demonstrated a significant impact on key cellular processes involved in proliferation and survival. Our findings establish a relevant role of *UGT2B17* in CLL with functional consequences and potential therapeutic implications.

Introduction

Chronic lymphocytic leukemia (CLL) is characterized by a considerable heterogeneity regarding clinical presentation, need for treatment, and outcome. Many prognostic markers have been identified.[1] While most provide information about risk of progression and survival, the functional role of these markers is often unclear and therapeutic consequences are therefore lacking. Apart from the clinical Rai and Binet, staging systems and cytogenetics[2-4], molecular markers like IgV_H mutational status[5,6] and *LPL* mRNA expression have strong prognostic value.[7,8] In a pilot gene expression study with 20 CLL patients, we have identified a significant association of uridine diphospho glucuronosyltransferase 2B17 (*UGT2B17)* with these poor prognostic factors.[9]

Metabolizing enzymes of the UGT2B family conjugate various endogenous compounds, in particular steroid hormones as well as several pharmaceutical drugs.[10,11] The UGT2B genes and pseudogenes are clustered on chromosome 4q13 and display up to 95% sequence identity among each other, which is reflected in some overlap in substrate specificity but often distinct expression profile. UGT2B17 is a major androgen inactivating enzyme playing a role in local tissue-specific regulation of it's substrates.[12] Importantly, antileukemic drugs such as anthraquinones or the histone deacetylase (HDAC) inhibitor vorinostat are also subject to glucuronidation by this enzyme.[13,14] An influence of *UGT2B17* on clinical outcome after vorinostat therapy in Asian women with breast cancer has recently been reported.[15]

UGT2B17 is affected by a remarkable copy number variation (CNV) spanning a 117kb-region encompassing the entire gene.[16,17] The frequency of copy numbers shows exceptional differences between populations from Africa, Europe or Asia.[17-19] Interindividual variability in *UGT2B17* allele-frequency is accompanied by pronounced differences in gene expression characterized by over 29 times higher mRNA levels in Caucasians compared to Japanese.[20] This is by far the greatest difference in gene expression observed between these two ethnic groups.

Several reports point to a role of *UGT2B17* in cancer susceptibility. *UGT2B17* null genotype is more frequent in adenocarcinoma of the lung in women.[21] Regarding an association between the CNV and prostate cancer risk contradictory results have been obtained.[22-24] However, the presence of the *UGT2B17* deletion was recently associated with cancer recurrence and sex-steroid hormone levels suggesting a prognostic role of this gene in prostate cancer patients after initial treatment.[25] Studies on graft versus host-disease furthermore identified *UGT2B17* as a minor

histocompatibility antigen presented by different HLA-ABC molecules and being a potential trigger of antibody response in graft versus host disease.[26,27] So far, no association with leukemia has been reported.

Here we report a detailed analysis of the role of *UGT2B17* in CLL. We investigated (1) *UGT2B17* mRNA expression as a prognostic marker; (2) the impact of the *UGT2B17* copy number on CLL susceptibility and prognosis; (3) potential functional consequences of modulating *UGT2B17* expression in CLL cells as well as enzymatic activity.

Material and methods

Patients and healthy donors

Peripheral blood samples from 320 patients with CLL diagnosed between 1973 and 2011 at the Vienna General Hospital and from 449 healthy donors were analyzed. Median observation time of patients was 73.7 months (range: 0.4 - 382.3 months). The majority was untreated at time of blood collection. CLL characterization and treatment response were defined according to National Cancer Institute-Working Group guidelines.[28] Treatment free survival was defined from date of diagnosis until first day of CLL specific therapy or death. Overall survival was defined from date of CLL diagnosis until death from any cause. Patient samples were investigated for cytogenetic aberrations and IgV_H mutational status as described previously.[8] The control subjects were unrelated Caucasian men and women, who were participants in a Viennese health care program, 45 years or older, free of vascular disease and any other severe disease at the time of sample collection between 1998 and 2000. Informed consent was obtained from the study participants according to the Declaration of Helsinki and the study protocol was approved by the institutional review board (ethical approvals at the Medical University of Vienna, EK-Nos. 385/2007, 025/2009).

Cell lines

MEC-1 were obtained from Leibniz Institute DSMZ-German Collection of Microorganisms (Braunschweig, Germany) and cultured in RPMI without phenolic red (Gibco®, Life Technologies™, Grand Island, NY, USA) containing 10% of FBS Gold (PAA®, Pasching, Austria). HEK293 cells were obtained from American Type Culture collection and cultured in fetal bovine serum albumin from Sigma-Aldrich Canada

(ON, Canada). Cells were maintained in the incubator at 37°C, 5%CO_2. Additional human cell lines (n=28) were obtained only for analysis of RNA expression as listed in supplementary Figure S1.

Drugs and chemicals

Vorinostat was obtained from Merck (Darmstadt, Germany). Drugs were diluted in DMSO. Androsterone (ADT), dihydrotestosterone (DHT), androstane-3α-diol (3α-Diol) were purchased from Steraloids (Newport, RI, USA).

Quantitative real-time PCR

DNA and RNA were prepared from unsorted frozen peripheral blood mononuclear cells (PBMC). Total RNA from 240 CLL patients (143 men, 97 women; 195 Binet stage A) and from cell lines was analyzed for *UGT2B17* mRNA expression by quantitative real time PCR (qPCR) analysis using Applied Biosystems assay-on-demand and Taqman Universal Master Mix without AmpErasesUNG. Samples were run on the ABI Prism 7000 Sequence Detector (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Beta actin was used as a housekeeping gene and expression was calculated relative to the mean of 10 unrelated healthy donor samples.

UGT family mRNA expression profiling for *UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A8, UGT1A9, UGT1A10, UGT2B7, UGT2B15, UGT2B17* and *UGT2B28* was performed with cDNA from MEC-1 using SYBR® green. 20 CLL patients samples were additionally analyzed for *UGT2B7* and *UGT2B15*, isoforms showing the highest sequence homology and a largely overlapping substrate specificity with *UGT2B17*. Reactions were run on a StepOnePlus™ real-time PCR system (Applied Biosystems, Foster City, CA, USA). Primers and protocols were used as listed in supplementary Table S1.

Conventional PCR

DNA samples of 277 Austrian CLL patients (167 men, 110 women; 223 stage Binet A) and 449 healthy donors (307 men, 142 women) were compared regarding the distribution of *UGT2B17* genotype. PCR was carried out with primers for marker E (in *UGT2B17* exon 6) and marker J (flanking *UGT2B17* deletion) as well as reaction conditions according to Wilson et al[17] and run on Biometra® TGradient or MWG Biotech ® Primus 96.

UGT2B17 silencing in MEC-1 cells

For knockdown experiments five pLKO.1 clones containing shRNAs targeting human *UGT2B17* and a control vector were obtained from Sigma-Aldrich® (Vienna, Austria) and lentiviral particles were produced according to standard protocols. In brief, HEK-293T cells were cotransfected with pLKO.1 plasmids, ΔR 8.91, and pVSV-G using lipofectamine2000 (Invitrogen™, Life Technologies™, Grand Island, NY, USA) to produce recombinant VSV-G pseudotyped lentiviruses. MEC-1 cells were transduced by spin infection (800 x g, 90 minutes, 32°C) in the presence of polybrene (7 µg/ml, Sigma-Aldrich, Vienna, Austria) and selected with puromycine (PAA®, Pasching, Austria). Knockdown of *UGT2B17* was assessed by quantitative real-time PCR as described previously above. For subsequent experiments the clone inducing the best knock down was applied. After two weeks of culture, RNA for microarray analysis of three independent stably transfected MEC-1 cell lines was isolated using Trizol® (Invitrogen™) and a subsequent clean-up step with RNAeasy columns (Quiagen®, Hilden, Germany).

Enzymatic assays and detection of glucuronides by mass spectrometry

To assess enzymatic glucuronidation activity of UGT2B17 in MEC-1 and CLL patient samples, microsomes were isolated as described previously.[29] Detection of DHT-glucuronide (-G), ADT-G, 3α-Diol-17-G and 3α-Diol-3-G was performed by liquid chromatography coupled to tandem mass spectrometry as described previously.[25]

Gene expression profiling

Microarray analysis of CLL patient samples has been reported previously.[9,30] Briefly, CD19-sorted PBMNC from 10 patients treated with fludarabine and cyclophosphamide and from 10 patients treated with FC in combination with rituximab (FCR) were analyzed before and after initiation of treatment. Samples from three independent *UGT2B17* shRNA-induced knock down experiments in MEC1 were run on GeneChip Human Gene 1.0 ST Array (Affymetrix, Santa Clara, CA, USA). GeneChip datasets are available online as GEO entry GSE15490 http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSE15490 and http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=djobjcoeykgksro&acc=GSE38367, respectively.

Statistical analysis

Gene expression was given as median, quartiles and range. The influence of *UGT2B17* expression on treatment free survival (TFS) was illustrated by Kaplan-Meier plots and quantified by hazard-ratios from Cox regression models. A cut-off level for *UGT2B17* expression optimal for predicting survival was exploratively determined from ROC curve. Genotype frequencies and prognostic markers were compared between groups using chi-square tests and Odds ratios were quantified by a logistic regression model. Spearman correlation coefficients were used to quantify the association between *UGT2B17* mRNA expression and enzymatic activity. No correction for multiple testing was performed due to the exploratory character of the study. Means and confidence intervals of relative changes in expression levels were computed on the log-scale and then back-transformed. Computations were performed using SAS software Version 9.2 (SAS Institute Inc., Cary, NC, USA, 2001).

For data import handling and all further calculations of gene expression microarrays in MEC-1 we used the tools provided by the Bioconductor project in R (www.bioconductor.org, cran.r-project.org). Boxplots, histograms and a correlation plot were drawn for quality assessment. The function 'RMA' implemented in the 'Affy'-package was used for background adjustment, normalization and summarization of the data into expression sets. A filtering step was performed where genes with a small variation across groups were deleted. A heatmap of the dataset is depicted with hierarchical clustering of samples and probe sets. To compare the genes of the two groups (two-sample t-tests) the function lmFit provided by the 'limma' package was applied on the data which results in one p-value for each probe set. To adjust for multiplicity the Benjamini-Hochberg method was applied to control the false discovery rate.[31] For pathway analysis the method by Tian et al[32] was used. All computations were done by using the statistical computing environment R 2.12.0.

Results

Association of *UGT2B17* with other prognostic markers in CLL

Compared to the mean of 10 normal healthy donor PBMC samples, the relative *UGT2B17* gene expression in 240 patients ranged from 0 to 264.12 (median: 1.77). High *UGT2B17* above median (≥1.77 fold vs. healthy donors) was significantly

associated with established poor prognostic factors such as unmutated IgV_H, high *LPL* and CD38 expression, absence of deletion 13q-, and trisomy 12 (Table 1). Interestingly, we found a trend toward a negative association with deletion 17p (Odds ratio=0.61, 95%CI 0.24-1.56).

UGT2B17 mRNA expression predicts treatment free and overall survival in CLL

High *UGT2B17* expression levels above the median were significantly associated with a shorter treatment free (HR=2.25, 95% CI 1.58-3.21, p<0.001) and overall survival (HR=2.18, 95% CI 1.18-4.01, p=0.013). Treatment free survival (TFS) was significantly shorter in patients with high *UGT2B17* expression (median TFS 62.6 months vs. 154.1 months; p<0.001). Furthermore, median overall survival (OS) was doubled in patients with low *UGT2B17* compared to those with high expression (median OS 152.4 vs. 299.5 months; p=0.011) (Figure 1A).

In multivariable models the influence of *UGT2B17* on TFS was independent from sex, CD38, cytogenetics, *LPL* expression and Binet stage, but not independent from IgV_H mutational status. The effect of UGT2B17 adjusted for these variables is quantified by a hazard ratio (HR) of 1.68 (95% CI 0.89 – 3.16, p=0.110). However, analysis of different prognostic subgroups revealed a pronounced impact of *UGT2B17* among patients with mutated IgV_H surface expression (p=0.014 for the interaction; Figure 1B and supplementary Figure S2).

These data indicate that high *UGT2B17* gene expression is suitable to identify poor-risk CLL patients at early stages of the disease. The explorative search for a cut-off level of *UGT2B17* mRNA expression, which optimally discriminates with respect to TFS yielded a relative level of 4.11 fold to be recommended for further studies.

Association of *UGT2B17* genotype with mRNA expression, development and course of CLL

Practically no gene dosage effect was observed although average mRNA-expression was slightly higher in homozygous carriers of the *UGT2B17* gene (median 6.3; q1= 0.5; q3=31.3) compared to heterozygous (median 4.3; q1=0.2; q3=19.7), and remained undetectable in patients with the null genotype (Figure 1C).

There was also a trend for association of longer TFS and OS with the null genotype (median TFS 113.9 months vs. 83.9 months; HR=0.60; 95% CI 0.35-1.03; p=0.066) but the effect of genotype was weak compared to mRNA-expression (Figure 1D).

Trends of association between CNV and prognostic markers were in line with results obtained for mRNA expression, but far from statistical significance and with a much weaker impact. Of note, no patient with the null genotype (-/-) had a positive Coombs test.

Genotype frequencies in 277 CLL patients were 41.2% for carriers of two copies of the *UGT2B17* gene (+/+), 45.8% for heterozygosity (+/-), and 13.0% with no copies of *UGT2B17* or the null genotype (-/-), respectively. In 449 healthy donors we found 43.0% +/+, 46.3% +/-, and 10.7% -/- (Figure 2A). These results are consistent with other investigations on *UGT2B17* copy number variations in Caucasians.[17,21-23] The frequency of the deleted allele did not differ significantly between CLL patients and healthy subjects or between men and women. This indicates that there is no decisive influence of *UGT2B17* CNV on the development of CLL in our cohort.

UGT2B17 mRNA expression and response to therapy

Sequential microarray analysis performed in 20 patients before and 3 days after initiation of therapy with FC or RFC showed a substantial up-regulation of *UGT2B17* mRNA within 48 hours of treatment only in patients with stable or progressive disease (mean relative change=134.1%, CI 106.0%-169.7%) compared to responders to therapy (mean relative change=99.1%; CI 84.2%-116.7%) (p=0.030) (Figure 2B). However, among 72 patients treated with fludarabine-based therapy, there was no association of *pre*-treatment *UGT2B17* mRNA-level with final response or progression free survival upon these regimens.

MEC-1 as a model for functional studies

RNA from 28 different human cell lines originating from various types of solid tumors and leukemias was obtained and analyzed for *UGT2B17* (supplementary Figure S1). Among all samples tested, the CLL cell line MEC-1 showed by far the highest level of *UGT2B17* mRNA and was selected as a model for further functional studies of UGT2B17 in CLL.

UGT family profiling in MEC-1 and primary CLL cells

Expression of 11 functional UGTs (*UGT1A1, UGT1A3, UGT1A4, UGT1A6, UGT1A8, UGT1A9, UGT1A10, UGT2B7, UGT2B15, UGT2B17* and *UGT2B28*) was then investigated in MEC-1. *UGT2B17* was the only one showing a significant expression (Figure 3A). In addition, in all 19 patients samples tested the expression of

UGT2B17-homolog isoforms *UGT2B7* and *UGT2B15* was also very low (Figure 3B). This makes a significant involvement of other enzymes of the UGT family with similar substrate specificity or regulation unlikely.

Relationship between *UGT2B17* mRNA expression and enzymatic activity in CLL samples

In the 19 CLL patients samples tested, area under the curve (AUC) for vorinostat glucuronidation ranged from 0.00 to 0.73, formation of DHT-G was between 0.0 and 58.1 pg/ml/mg protein. Glucuronidation of these substrates strongly correlated to *UGT2B17* mRNA expression in all samples tested (Spearman correlation coefficient >0.9, $p<0.001$) (supplementary Figure S3).

Gene expression profiling of functional *UGT2B17* knock down in MEC-1

Using five *UGT2B17*-targeting shRNA sequences, between 30 and 90% knock down of *UGT2B17* mRNA compared to the random shRNA control were achieved in the cell line model MEC-1. The clone inducing the best knock down (90%) was NM_001077.1-1554s1c1, shRNA sequence 5'-CCG GCG TGG CAA CTA TGA TAT TTA TCT CGA GAT AAA TAT CAT AGT TGC CAC GTT TTT G-3'. Stable repression of *UGT2B17* expression was still achieved after 2 weeks of culture post-transfection. In long-term culture over several months, a knock down of approximately 70% was maintained. This shRNA-induced suppression of *UGT2B17* mRNA was directly proportional to a decrease of UGT2B17-specific enzymatic functions (Figure 3C).
Microarray gene expression profiling of three independent *UGT2B17* knock down MEC-1 cell lines revealed significant changes in expression of 864 genes compared to the negative controls (supplementary Figure S4). The genes with reduced expression upon *UGT2B17* silencing were involved in pathways related to chromatin and nucleosome organization, regulation of proliferation and apoptosis, protein and cytokine synthesis and transport as well as inflammation and immunological defense (Table 2). In contrast, a significant up-regulation of genes involved in intracellular transport and secretion including the endoplasmatic reticulum and Golgi compartments was found (Table 3). Among the top differentially expressed probe sets, we also observed several genes known to be involved in CLL susceptibility and progression, like *CASP1,*[33] CD38,[34] *GRAMD1B,*[35] *CCND2* (cyclin D2),[36] *ITGAL* (LFA-1; CD11a) [37], *GZMK*[38], *FCRL3*[39] (Figure 4). These data indicate that *UGT2B17*

knock-down leads to perturbations in gene expression which may influence CLL cell behavior.

Discussion

This is the first report of a role of a UGT metabolic enzyme in leukemia. We identified a strong association of *UGT2B17* mRNA expression with other prognostic markers in CLL. This translates into a significantly shorter treatment free and overall survival among patients with high *UGT2B17* expression levels. Furthermore, *UGT2B17* gene expression had a particular discriminating power within favorable prognostic CLL subgroups.

Our data point to a functional role of UGT2B17 in CLL. We demonstrate that the highly variable *UGT2B17* mRNA levels found in quantitative real-time PCR translate into specific enzymatic protein function. Among known substrates of the enzyme are androgens and therapeutic drugs being used for leukemia treatment like chlorambucil (unpublished data), anthraquinones[13] and the HDAC-inhibitor vorinostat.[14,40] Of note, we also observed a significant up-regulation of *UGT2B17* during treatment with standard regimens. Fludarabine and cylophosphamide are not subject to significant glucuronidation. However, constitutive over-expression or up-regulation of UGT2B17 during treatment with these drugs could lead to an increased elimination of other substrates. Such inducible and highly variable glucuronidation rates of drugs in CLL cells could be of therapeutic relevance since drug inactivation may also occur directly in leukemic cells in addition to classic metabolic tissues like liver, intestine, and kidney. Thus, the presence of UGT2B17 in leukemic cells may have an impact on antineoplastic drug metabolism in cancer cells with the potential to affect overall drug response.

Interestingly, expression profiling of various UGT genes in patients' samples and in the CLL cell line model MEC-1 revealed no significant mRNA expression or activity of any other UGT2B enzyme with a similar substrate spectrum in CLL cells. These results point to an exclusive role of the UGT2B17 isoform and suggest that its over-expression in CLL cells plays a role likely beyond glucuronidation of androgens and drugs. Indeed, UGT2B17 knock down in MEC-1 affected predominantly genes involved in intracellular trafficking, regulation of proliferation and apoptosis as well as immunological pathways, including genes known for an association with CLL. The decrease of CD38 expression in line with down-regulation of *UGT2B17* is also in agreement with the association between *UGT2B17* and CD38 observed in our CLL

patients. However, factors responsible for *UGT2B17* over-expression and its related functional consequences remain to be elucidated and are currently subject to further investigations.

We also investigated the role of *UGT2B17* copy number on CLL susceptibility and prognosis. Genotype distribution did not differ significantly between CLL patients and healthy donors among our Central European study population. This indicates that the *UGT2B17* polymorphic deletion status by itself does not have an obvious impact on CLL development. The role of *UGT2B17* expression was much more relevant for CLL prognosis than that of gene copy number alone. Therefore, the role of UGT2B17 in CLL susceptibility remains unclear. A comparative genetic study of Asian and Caucasian patients has recently been initiated to address the impact of a CNV on the different incidence of CLL in the Asian and Caucasian populations.[41]

In conclusion our findings indicate a relevant role of the *UGT2B17* pathway for CLL prognosis with potential functional and therapeutic implications

Authorship contributions

M.G. designed the research, performed experiments, collected clinical data, analyzed data and wrote the paper. J.B. performed experiments and writing on UGT2B functional experiments, G.H. did experiments and writing on lentiviral transfections, M.B. performed microarray analysis, A.Gl. and S.Z. provided expert statistical analysis, E.P. and T.L. performed experiments., C.M., A.Ga., M.S. provided samples and molecular data, K.F. and C.S. provided patient clinical information, E.L. and C.G. supervised and wrote UGT2B family functional and glucuronidation experiments, K.V. supervised *UGT2B17* PCR experiments on CLL patients and healthy donors, U.J. designed the research, supervised the study and wrote the paper. All authors critically revised the manuscript.

Acknowledgements

We thank Martin and Bernadette Hilgarth, Elisabeth Reiter, Ulrike Zeman and Particia Weiss for technical support on CLL patient sample collection and experiments, Lyne Villeneuve and Mario Harvey for technical support on UGT2B related experiments, Patrick Caron for MS-based analyses of sex-steroids and Markus Jeitler for microarray experiments. Thanks to Christine Einberger for support with clinical data collection, Peter Valent for provision of vorinostat and leukemia cell

lines as well as Thomas Grunt, Gerald Prager and Martin Filipits for provision of carcinoma cell lines. We also thank Robin Riestl for uploading our microarray data to the GEO platform.

M.G. received the EHA-ASH Translational Research Training in Hematology Award and was supported by a Marietta Blau Fellowship for a research stay in Québec, by the Hans und Blanca Moser Stiftung and by a Förderungsstipendium from the Austrian Ministry of Science. J.B. was supported by a CIHR Frederick Banting and Charles Best studentship award. E.L. holds a CIHR clinician-scientist award. C.G. holds the Canada research chair in pharmacogenomics.

This work was supported in part by Gen-AU C.h.i.l.d. II (GZ 200.136/1 - VI/1/2005 to K.V.), Initiative Krebsforschung of the Medical University of Vienna (K.V.), the Canadian Institutes of Health Research (MOP-42392 to C.G.) and the Canada Research Chair Program (C.G.).

Conflicts of interest disclosure

The authors have no conflicts of interest to declare

References

1. Gribben JG. Molecular Profiling in CLL. Hematology Am Soc Hematol Educ Program. 2008;2008:444-449.
2. Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. Blood. 1975;46:219-234.
3. Binet JL, Auquier A, Dighiero G, et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer. 1981;48:198-206.
4. Dohner H, Stilgenbauer S, Benner A, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343:1910-1916.
5. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood. 1999;94:1848-1854.
6. Damle RN, Wasil T, Fais F, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999;94:1840-1847.
7. Oppezzo P, Vasconcelos Y, Settegrana C, et al. The LPL/ADAM29 expression ratio is a novel prognosis indicator in chronic lymphocytic leukemia. Blood. 2005;106:650-657.
8. Heintel D, Kienle D, Shehata M, et al. High expression of lipoprotein lipase in poor risk B-cell chronic lymphocytic leukemia. Leukemia. 2005;19:1216-1223.
9. Bilban M, Heintel D, Scharl T, et al. Deregulated expression of fat and muscle genes in B-cell chronic lymphocytic leukemia with high lipoprotein lipase expression. Leukemia. 2006;20:1080-1088.
10. Beaulieu M, Levesque E, Hum DW, Belanger A. Isolation and characterization of a novel cDNA encoding a human UDP-glucuronosyltransferase active on C19 steroids. J Biol Chem. 1996;271:22855-22862.
11. Beaulieu M, Levesque E, Tchernof A, Beatty BG, Belanger A, Hum DW. Chromosomal localization, structure, and regulation of the UGT2B17 gene, encoding a C19 steroid metabolizing enzyme. DNA Cell Biol. 1997;16:1143-1154.
12. Chouinard S, Yueh MF, Tukey RH, et al. Inactivation by UDP-glucuronosyltransferase enzymes: the end of androgen signaling. J Steroid Biochem Mol Biol. 2008;109:247-253.
13. Turgeon D, Carrier JS, Chouinard S, Belanger A. Glucuronidation activity of the UGT2B17 enzyme toward xenobiotics. Drug Metab Dispos. 2003;31:670-676.
14. Balliet RM, Chen G, Gallagher CJ, Dellinger RW, Sun D, Lazarus P. Characterization of UGTs active against SAHA and association between SAHA glucuronidation activity phenotype with UGT genotype. Cancer Res. 2009;69:2981-2989.
15. Wong NS, Seah E, Wang LZ, et al. Impact of UDP-gluconoryltransferase 2B17 genotype on vorinostat metabolism and clinical outcomes in Asian women with breast cancer. Pharmacogenet Genomics. 2011;21:760-768.
16. McCarroll SA, Hadnott TN, Perry GH, et al. Common deletion polymorphisms in the human genome. Nat Genet. 2006;38:86-92.
17. Wilson W, 3rd, Pardo-Manuel de Villena F, Lyn-Cook BD, et al. Characterization of a common deletion polymorphism of the UGT2B17 gene linked to UGT2B15. Genomics. 2004;84:707-714.
18. Xue Y, Sun D, Daly A, et al. Adaptive Evolution of UGT2B17 Copy-Number Variation. The American Journal of Human Genetics. 2008;83:337-346.
19. Jakobsson J, Ekstrom L, Inotsume N, et al. Large differences in testosterone excretion in Korean and Swedish men are strongly associated with a UDP-

glucuronosyl transferase 2B17 polymorphism. J Clin Endocrinol Metab. 2006;91:687-693.
20. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, Cheung VG. Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet. 2007;39:226-231.
21. Gallagher CJ, Muscat JE, Hicks AN, et al. The UDP-glucuronosyltransferase 2B17 gene deletion polymorphism: sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer. Cancer Epidemiol Biomarkers Prev. 2007;16:823-828.
22. Gallagher CJ, Kadlubar FF, Muscat JE, Ambrosone CB, Lang NP, Lazarus P. The UGT2B17 gene deletion polymorphism and risk of prostate cancer. A case-control study in Caucasians. Cancer Detect Prev. 2007;31:310-315.
23. Karypidis AH, Olsson M, Andersson SO, Rane A, 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;8:147-151.
24. Olsson M, Lindstrom S, Haggkvist B, et al. The UGT2B17 gene deletion is not associated with prostate cancer risk. Prostate. 2008;68:571-575.
25. Nadeau G, Bellemare J, Audet-Walsh E, et al. Deletions of the androgen-metabolizing UGT2B genes have an effect on circulating steroid levels and biochemical recurrence after radical prostatectomy in localized prostate cancer. J Clin Endocrinol Metab. 2011;96:E1550-1557.
26. Murata M, Warren EH, Riddell SR. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. J Exp Med. 2003;197:1279-1289.
27. McCarroll SA, Bradner JE, Turpeinen H, et al. Donor-recipient mismatch for common gene deletion polymorphisms in graft-versus-host disease. Nat Genet. 2009;41:1341-1344.
28. Hallek M, Cheson BD, Catovsky D, et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood. 2008;111:5446-5456.
29. Laverdiere I, Caron P, Harvey M, Levesque E, Guillemette C. In vitro investigation of human UDP-glucuronosyltransferase isoforms responsible for tacrolimus glucuronidation: predominant contribution of UGT1A4. Drug Metab Dispos. 2011;39:1127-1130.
30. Shehata M, Demirtas D, Schnabl S, et al. Sequential gene expression profiling during treatment for identification of predictive markers and novel therapeutic targets in chronic lymphocytic leukemia. Leukemia. 2010;24:2122-2127.
31. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B-Methodological. 1995;57:289-300.
32. Tian L, Greenberg SA, Kong SW, Altschuler J, Kohane IS, Park PJ. Discovering statistically significant pathways in expression profiling studies. Proc Natl Acad Sci U S A. 2005;102:13544-13549.
33. Enjuanes A, Benavente Y, Bosch F, et al. Genetic variants in apoptosis and immunoregulation-related genes are associated with risk of chronic lymphocytic leukemia. Cancer Res. 2008;68:10178-10186.
34. Kern W, Bacher U, Haferlach C, et al. Monoclonal B-cell lymphocytosis is closely related to chronic lymphocytic leukaemia and may be better classified as early-stage CLL. Br J Haematol. 2012.

35. Slager SL, Goldin LR, Strom SS, et al. Genetic susceptibility variants for chronic lymphocytic leukemia. Cancer Epidemiol Biomarkers Prev. 2010;19:1098-1102.
36. Delmer A, Ajchenbaum-Cymbalista F, Tang R, et al. Overexpression of cyclin D2 in chronic B-cell malignancies. Blood. 1995;85:2870-2876.
37. Csanaky G, Matutes E, Vass JA, Morilla R, Catovsky D. Adhesion receptors on peripheral blood leukemic B cells. A comparative study on B cell chronic lymphocytic leukemia and related lymphoma/leukemias. Leukemia. 1997;11:408-415.
38. Aalto Y, El-Rifa W, Vilpo L, et al. Distinct gene expression profiling in chronic lymphocytic leukemia with 11q23 deletion. Leukemia. 2001;15:1721-1728.
39. Li FJ, Ding S, Pan J, et al. FCRL2 expression predicts IGHV mutation status and clinical progression in chronic lymphocytic leukemia. Blood. 2008;112:179-187.
40. Perez-Perarnau A, Coll-Mulet L, Rubio-Patino C, et al. Analysis of apoptosis regulatory genes altered by histone deacetylase inhibitors in chronic lymphocytic leukemia cells. Epigenetics. 2011;6:1228-1235.
41. Tamura K, Sawada H, Izumi Y, et al. Chronic lymphocytic leukemia (CLL) is rare, but the proportion of T-CLL is high in Japan. Eur J Haematol. 2001;67:152-157.

Table 1. Association of high *UGT2B17* (>1.77 fold HD) versus low (≤1.77 fold HD) mRNA expression with prognostic markers in CLL.

Marker	N total	CLL patients with high *UGT2B17* (N)	CLL patients with low *UGT2B17* (N)	P-value	OR/ HR	95% CI
Binet stage B or C	239	21.8% (119)	15.0% (120)	0.172	1.58	0.82 – 3.08
Positive Coombs-Test	224	8.2% (110)	1.8% (114)	**0.026**	**4.99**	1.05 – 23.64
High LPL-expression (>10.0*)	167	60.2% (93)	29.7% (74)	**<0.001**	**3.58**	1.87 – 6.84
High CD38 expression (>30%)	219	44.1% (111)	21.3% (108)	**<0.001**	**3.03**	1.67 – 5.48
Cytogenetics						
17p deletion	227	6.9% (116)	10.8% (111)	0.298	0.61	0.24 – 1.56
11q deletion	227	28.4% (116)	18.0% (111)	0.063	1.81	0.96 – 3.40
trisomy 12	227	21.6% (116)	11.7% (111)	**0.047**	**2.07**	1.00 – 4.29
13q deletion as sole abnormality	227	18.1% (116)	35.1% (111)	**0.004**	**0.41**	0.22 – 0.75
14q aberrations	227	20.9% (116)	15.3% (111)	0.279	1.46	0.74 – 2.89
normal caryotype	227	22.4% (116)	24.3% (111)	0.734	0.90	0.49 – 1.66
Unmutated IgV$_H$ (>98% sequence homology)	185	59.6% (99)	24.4% (86)	**<0.001**	**4.57**	2.42 – 8.61
Immunoglobulin V$_H$ gene usage						
1-69	183	21.8% (101)	6.1% (82)	**0.004**	**4.21**$^+$	1.50 – 11.77
3-21	183	5.0% (101)	3.7% (82)	0.532	1.59$^+$	0.37 – 6.94
3-23	183	5.9% (101)	11.0% (82)	0.414	0.64$^+$	0.21 – 1.89
UGT2B17 +/+ vs. +/-	167	50.5% (105)	41.9% (62)	0.286	1.41	0.75 – 2.66
Median treatment free survival (n events/total)	239	62.6 (81/119)	154.1 (52/120)	**<0.001**	**2.25**	1.58 – 3.27
Median overall survival (n events/total)	240	152.4 (29/120)	299.5 (19/120)	**0.013**	**2.18**	1.18 – 4.01

* fold expression in relation to healthy donor control
$^+$ vs. other

Figure 1. Impact of high *UGT2B17* mRNA expression (>1.77 fold of healthy donors) (A) and copy number variation on TFS and OS (B) in IgVH mutated subgroup on TFS (C) Lack of gene dosage effect of *UGT2B17* copy number and (D) UGT2B17 CNV and overall and treatment free survival.

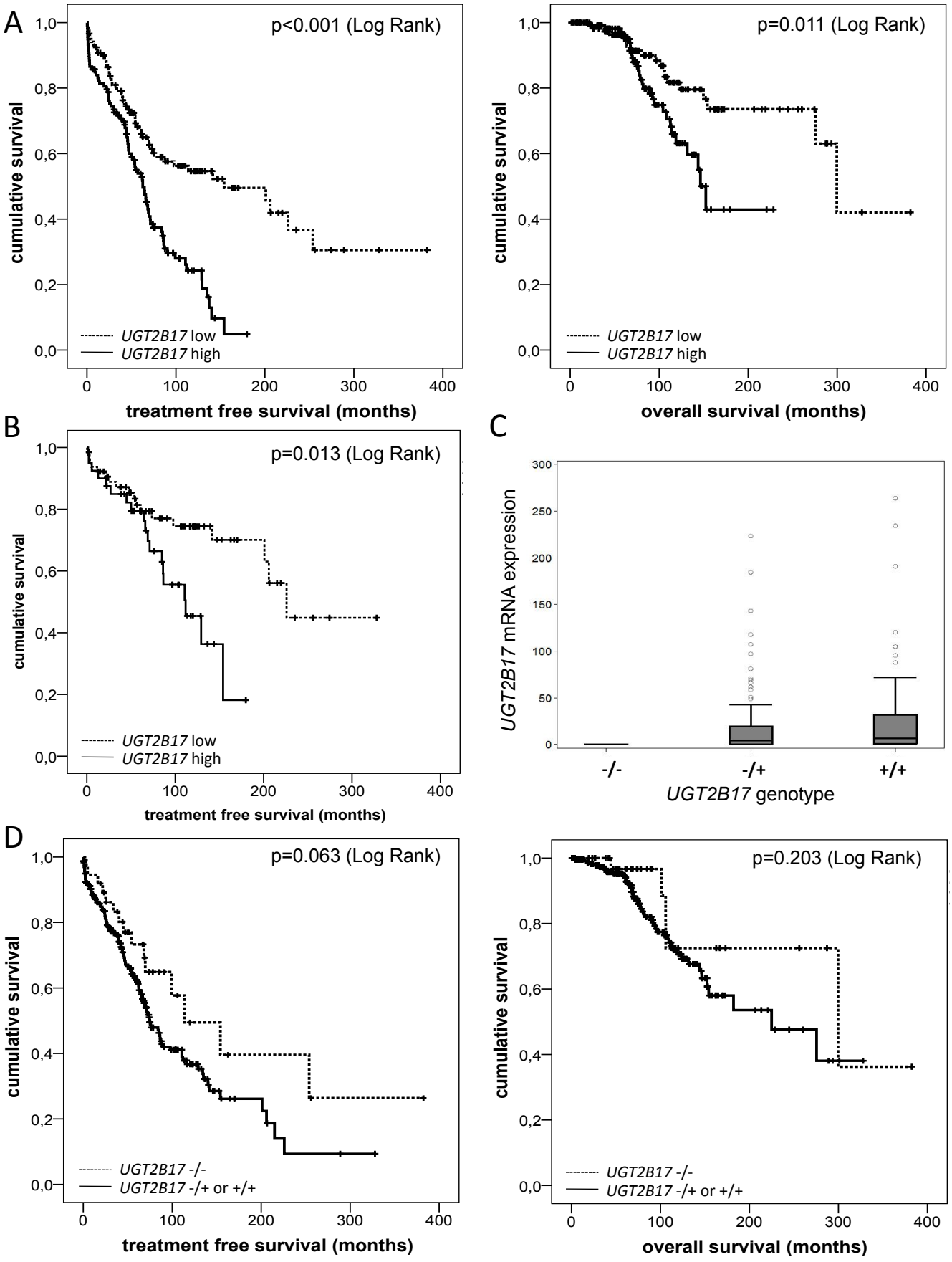
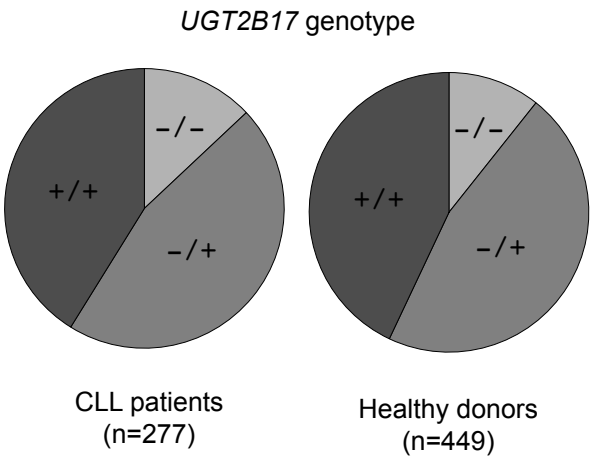


Figure 2. (A) *UGT2B17* genotype and CLL susceptibility and (B) Change in *UGT2B17* mRNA expression in CLL cells during treatment with fludarabine and cyclophosphamide.

A



B

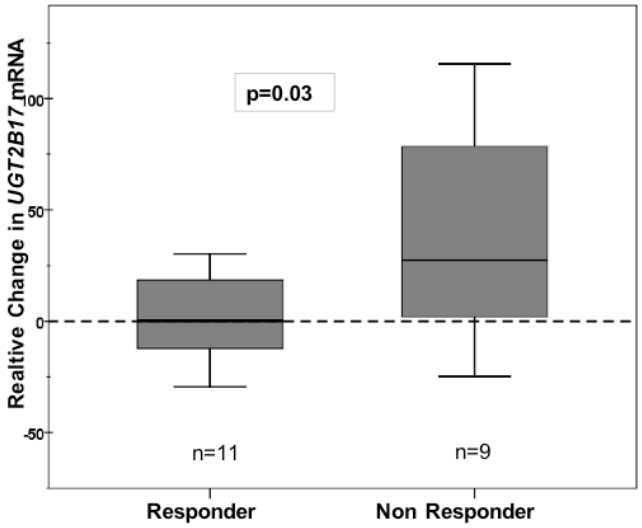


Figure 3. UGT family expression profiling in MEC-1 (A), UGT2B isoform expression in CLL patients (n=19) (B) and functional UGT2B17 shRNA knock down in MEC-1 (mean of 3 replicates)

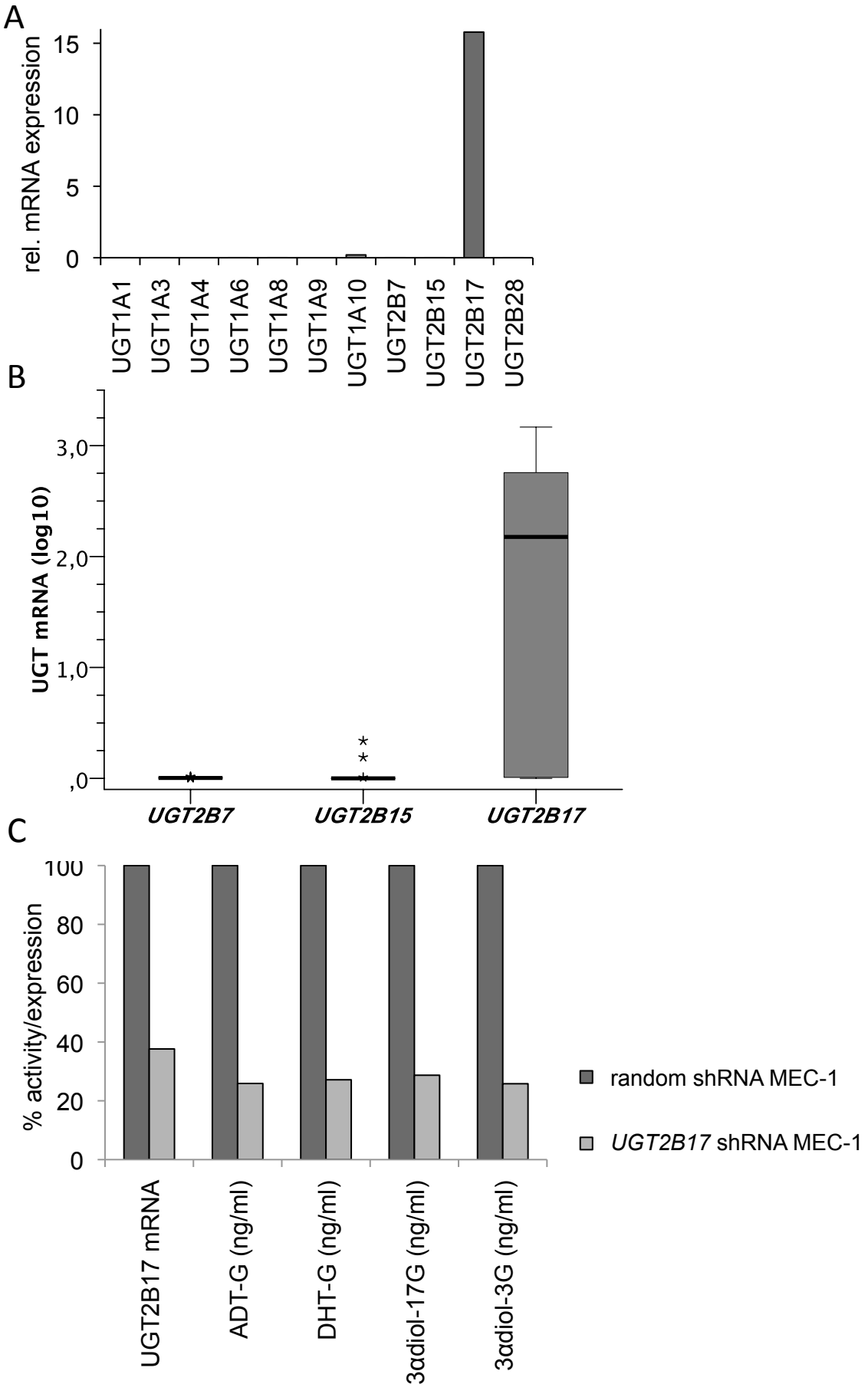
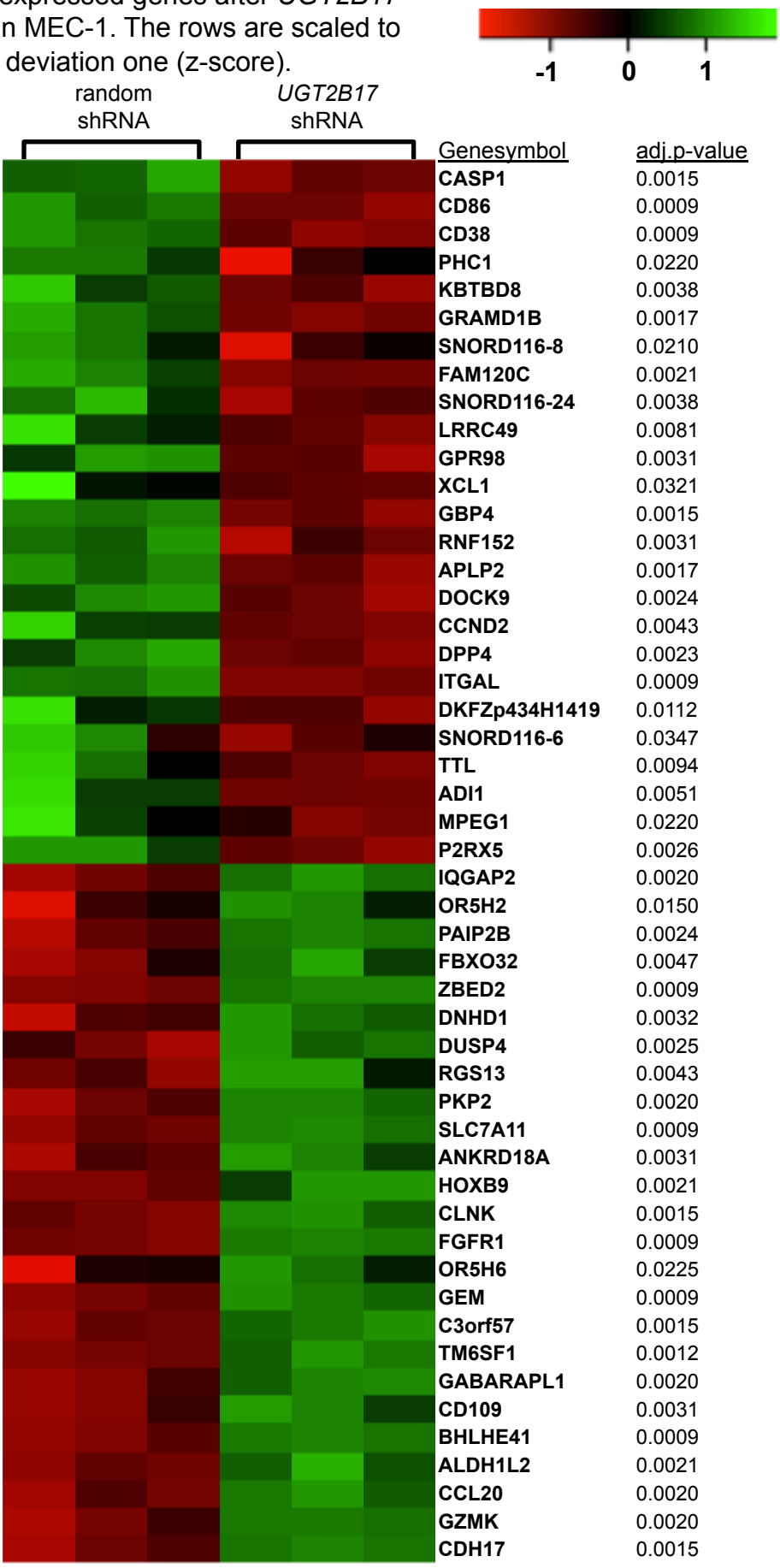


Figure 4. Top 50 differentially expressed genes after *UGT2B17* gene expression knock down in MEC-1. The rows are scaled to have mean zero and standard deviation one (z-score).



Genesymbol	adj.p-value
CASP1	0.0015
CD86	0.0009
CD38	0.0009
PHC1	0.0220
KBTBD8	0.0038
GRAMD1B	0.0017
SNORD116-8	0.0210
FAM120C	0.0021
SNORD116-24	0.0038
LRRC49	0.0081
GPR98	0.0031
XCL1	0.0321
GBP4	0.0015
RNF152	0.0031
APLP2	0.0017
DOCK9	0.0024
CCND2	0.0043
DPP4	0.0023
ITGAL	0.0009
DKFZp434H1419	0.0112
SNORD116-6	0.0347
TTL	0.0094
ADI1	0.0051
MPEG1	0.0220
P2RX5	0.0026
IQGAP2	0.0020
OR5H2	0.0150
PAIP2B	0.0024
FBXO32	0.0047
ZBED2	0.0009
DNHD1	0.0032
DUSP4	0.0025
RGS13	0.0043
PKP2	0.0020
SLC7A11	0.0009
ANKRD18A	0.0031
HOXB9	0.0021
CLNK	0.0015
FGFR1	0.0009
OR5H6	0.0225
GEM	0.0009
C3orf57	0.0015
TM6SF1	0.0012
GABARAPL1	0.0020
CD109	0.0031
BHLHE41	0.0009
ALDH1L2	0.0021
CCL20	0.0020
GZMK	0.0020
CDH17	0.0015

Table 2. Pathways with decreased gene-expression upon *UGT2B17* knock down in MEC-1

Chromatin and nucleosome organization	Set Size	NTK Stat	NTk q-value
chromatin assembly	89	6,26	0.0000
nucleosome	64	6,54	0.0000
chromosome organization and biogenesis (sensu Eukaryota)	278	6,54	0.0000
chromosome organization and biogenesis	292	6,65	0.0000
nucleosome assembly	80	6,74	0.0000
Regulation of cellular proliferation and apoptosis	Set Size	NTK Stat	NTk q-value
regulation of apoptosis	340	3,47	0.0000
Regulation of eIF4e and p70 S6 Kinase	22	4,28	0.0000
induction of apoptosis	146	4,30	0.0000
induction of programmed cell death	146	4,30	0.0000
Cell Cycle G1/S Check Point	27	4,59	0.0000
Cyclins and Cell Cycle Regulation	25	4,81	0.0000
positive regulation of programmed cell death	163	5,21	0.0000
positive regulation of apoptosis	162	5,26	0.0000
Inflammation and immunological defense	Set Size	NTK Stat	NTk q-value
lymphocyte proliferation	25	3,32	0.0000
Drug Targets for Inflammation / Immunomodulation	97	3,68	0.0000
Cells and Molecules involved in local acute inflammatory response	17	3,80	0.0000
Dendritic / Antigen Presenting Cell	169	3,82	0.0000
Leukocyte transendothelial migration	120	5,18	0.0000
Adhesion and Diapedesis of Granulocytes	15	4,47	0.0000
Protein and cytokine synthesis and transport	Set Size	NTK Stat	NTk q-value
positive regulation of protein biosynthesis	24	3,22	0.0000
positive regulation of biosynthesis	27	3,50	0.0000
positive regulation of cytokine biosynthesis	21	3,57	0.0000
positive regulation of cytokine production	25	3,69	0.0000
positive regulation of cellular biosynthesis	25	3,73	0.0000
regulation of nucleocytoplasmic transport	23	3,85	0.0000
positive regulation of protein metabolism	43	3,85	0.0000
regulation of protein transport	18	4,02	0.0000
protein complex assembly	277	6,24	0.0000
regulation of protein metabolism	211	4,37	0.0000
Other pathways	Set Size	NTK Stat	NTk q-value
integrin complex	37	4,50	0.0000
Starch and sucrose metabolism	69	4,97	0.0000
glucuronosyltransferase activity	20	5,01	0.0000
Pentose and glucuronate interconversions	16	5,16	0.0000
Porphyrin and chlorophyll metabolism	30	5,20	0.0000

Table 3. Pathways up-regulated in MEC-1 upon *UGT2B17* shRNA induced knock down

Intracellular transport and secretion involving ER and Golgi compartment	Set Size	NTK Stat	NTk q-value
ER-Golgi intermediate compartment	31	-6,71	0.0000
Golgi vesicle transport	100	-5,63	0.0000
intra-Golgi vesicle-mediated transport	16	-5,14	0.0000
retrograde vesicle-mediated transport, Golgi to ER	16	-5,14	0.0000
ER to Golgi vesicle-mediated transport	66	-5,11	0.0000
Golgi-associated vesicle	49	-4,90	0.0000
endoplasmic reticulum membrane	92	-4,65	0.0000
vesicle membrane	66	-4,63	0.0000
cytoplasmic vesicle membrane	63	-4,60	0.0000
nuclear envelope-endoplasmic reticulum network	97	-4,59	0.0000
secretory pathway	191	-4,39	0.0000
secretion	237	-4,13	0.0000
SNARE interactions in vesicular transport	35	-2,88	0.0163
Other pathways	Set Size	NTK Stat	NTk q-value
cytokine and chemokine mediated signaling pathway	23	-4,42	0.0000
protein-tyrosine kinase activity	184	-3,67	0.0000
Adherens junction	78	-2,88	0.0163

Figure S1. *UGT2B17* mRNA expression relative to healthy donor PBMC in cell lines originating from different solid tumors and leukemias

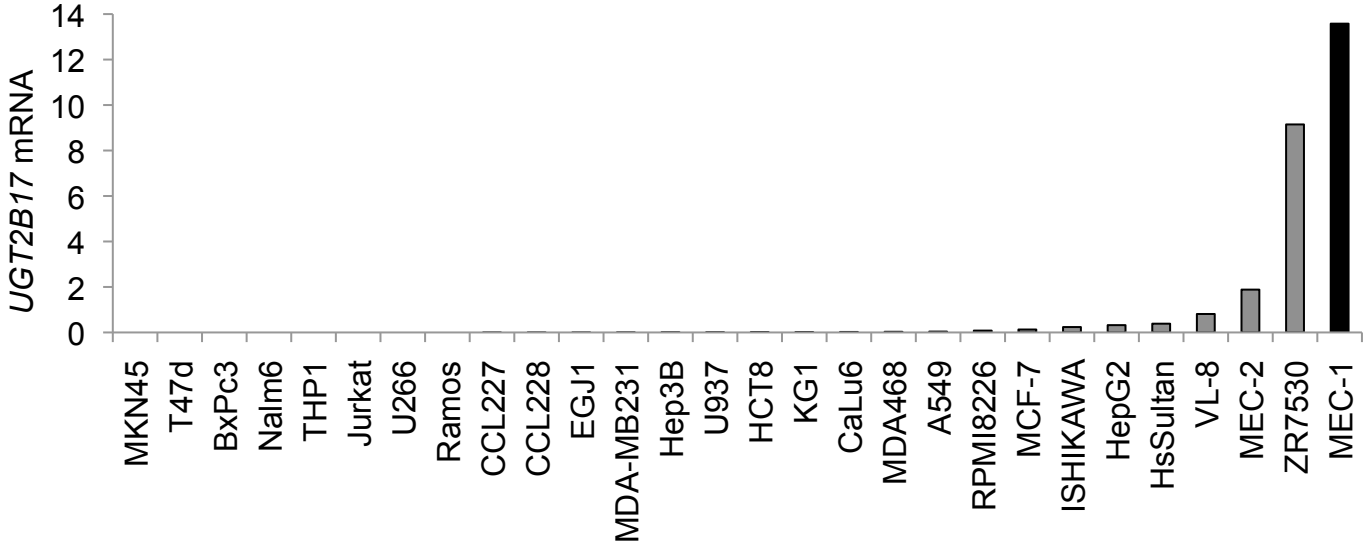


Figure S2. Impact of *UGT2B17* mRNA expression in IgV_H unmutated subgroup

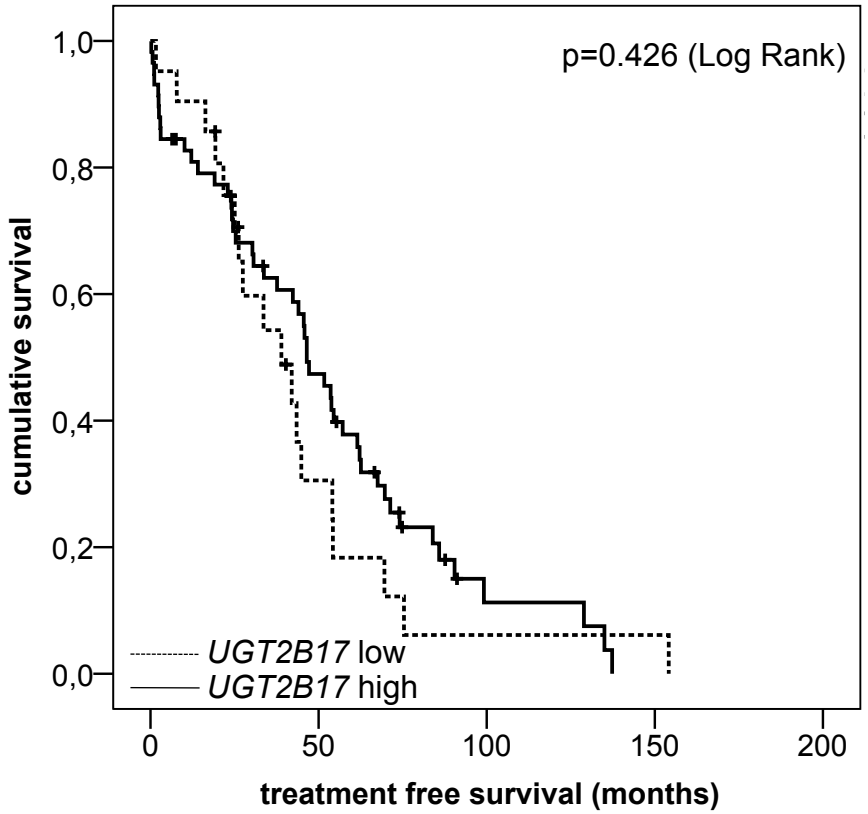
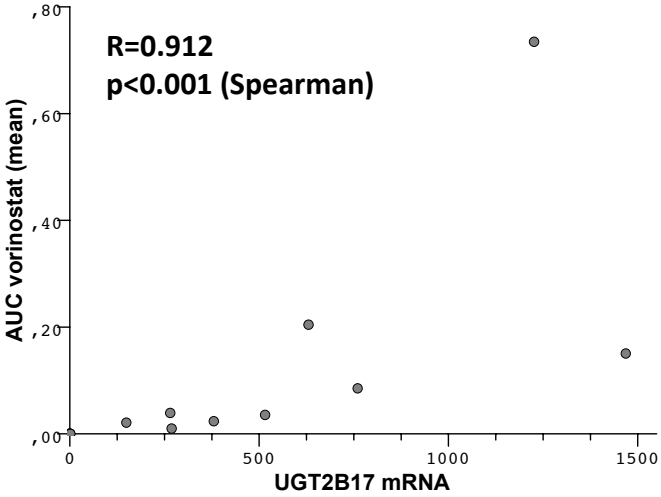


Figure S3. Correlation between UGT2B17 mRNA expression and (A) vorinostat area under the curve as well as (B) dihydrotestosterone glucuronidation in primary CLL cells (n=19)



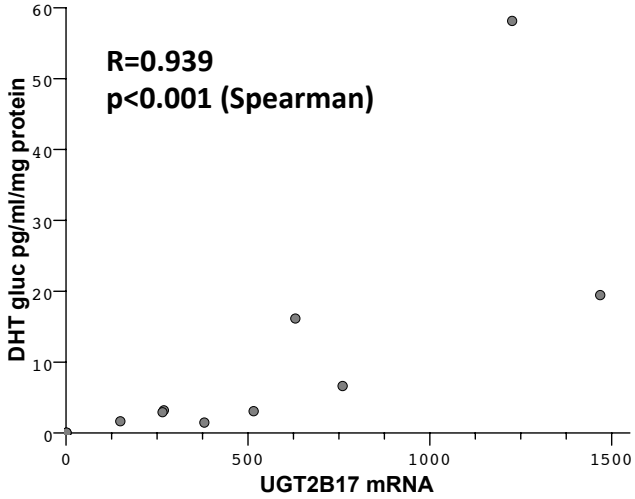


Table S1. Primers and protocols used for UGT family mRNA expression profiling by real-time PCR

mRNA	Primers	Sequence 5' → 3'	Final concentration (nM)	Reference
UGT1A1	242 - F	GAGAGAGGTGACTGTCCAGGAC	150	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT1A1	243 - R	CAAATTCCTGGGATAGTGGATTTT	150	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT1A3	3008 - F	TGAATTTGATCGCCATGTGC	600	-
UGT1A3	3010 - R	GTAGCTCCACACAAGACCTATGAT	400	-
UGT1A4	3357 - F	GTATCTTTGGCCCTTCATAGGTG	400	-
UGT1A4	3358 - R	CGCCCCACAGAGGTTAACG	200	-
UGT1A6	3354 - F	GTACTTCATCAACTGCCAGAGCC	600	-
UGT1A6	3355 - R	CAGGGAACACGGAAAACCCCTGA	400	-
UGT1A8	3011 - F	CCCATTCCCCTATGTGTTTC	600	-
UGT1A8	3012 - R	GCCCCCTGAGGATAAGTTTC	600	-
UGT1A9	389 - F	AATTCTCCAAACACCTGTTACGGAG	200	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT1A9	348 - R	ATGTAGGCTTCAAATTCCATAGGCAAC	200	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT1A10	3014 - F	TCGTACACTCTGGAAGATCAGAA	300	-
UGT1A10	3015 - R	ACTTTGTGCCTGTGCTTTCC	600	-
UGT2B7	2695 - F	AGACAATGGGGAAAGCTGACG	300	Ménard V et al. Pharmacogenet Genomics. 2011 Oct;21(10):631-41.
UGT2B7	2697 - R	TCTCCAGAGCTCTGTACAAAG	300	Ménard V et al. Pharmacogenet Genomics. 2011 Oct;21(10):631-41.
UGT2B15	223 - F	GTGTTGGGAATATTATGACTACAGTAAC	125	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT2B15	224 - R	GGGTATGTTAAATAGTTCAGCCAGT	125	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT2B17	1686 - F	TGACTTTTGGTTTCAAGC	300	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT2B17	1687 - R	TTCCATTTCCTTAGGCAA	300	Lépine J et al. J Clin Endocrinol Metab. 2010 Jun;95(6):2689-98. Epub 2010 Apr 6.
UGT2B28	3467 - F	AAGGTGCTGGTGTGGACCGGT	200	-
UGT2B28	3432 - R	GAGTTTAAGAGTGAATGCGTCATTGG	600	-
Beta-Actin	3107 - F	ACAGAGCCTCGCCTTTGCCG	400	-
Beta-Actin	3109 - R	ACCCATGCCCACCATCACGC	600	-

Figure S4. Unsupervised hierarchical clustering of 864 differentially expressed genes upon *UGT2B17* knock down in the MEC-1 cell line in 3 independent replicates for each group