UCSF
UC San Francisco Electronic Theses and Dissertations

**Title**
# Novel Computational Methods to Discover Genes Linked to Drug Response

**Permalink**
https://escholarship.org/uc/item/59g8108v

**Author**
Goswami, Srijib

**Publication Date**
2015

Peer reviewed|Thesis/dissertation

eScholarship.org
Powered by the California Digital Library
University of California

Novel Computational Methods to Discover Genes Linked to Drug Response

by

Srijib Goswami

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Pharmaceutical Sciences and Pharmacogenomics

in the

## Acknowledgements

My successful completion of a Ph.D. would not be possible without the unwavering support from my family, friends, colleagues and mentors. All of my accomplishments are a direct result of their faith and confidence in my abilities, especially during difficult times when success seemed faint and distant. My journey as a Ph.D. student was filled with invaluable interactions, all of which were integral in this achievement.

First, Dr. Kathleen M. Giacomini has been an outstanding mentor throughout my graduate career, and I cannot thank her enough for her constant support and guidance. Dr. Giacomini was critical in shaping my abilities as a scientist, influencing how I analyze problems and teaching me how to put my research findings into clinical perspective, a skill that will serve me during any phase of my career. Most importantly, Dr. Giacomini gave me the confidence to be bold and take calculated risks. As we all know, confidence is a critical determinant of success. I would like to thank Dr. Giacomini for helping me build a high level of self-confidence, which enabled me to trust in my skills and abilities.

I would like to thank my incredible mentor, Dr. Radojka Savic. Over the past 5 years, I have learned a tremendous amount from Dr. Savic, especially in quantitative pharmacology. My countless discussions with Dr. Savic have influenced the way I think and analyze data, and have certainly made me a better scientist. Dr. Savic inspired me to pursue my next career endeavor, entering the

realm of entrepreneurship and starting a company. I am very grateful for her optimism, confidence, and invaluable mentorship.

I would also like to thank my oral qualifying exam committee, consisting of Drs. Leslie Benet, Nadav Ahituv, Laura Bull, and Pui Kwok, for taking the time to review and provide critical feedback on my dissertation proposal. In addition to Dr. Giacomini, I would also like to thank Dr. Savic, and Dr. Benet for their efforts on my dissertation committee. Both took the time to meet with me on several occasions, in order to ensure that my science was of the highest quality. I could not complete this thesis without their helpful discussions and advice.

All members of the Giacomini and Savic laboratories deserve special thanks for their support and making my time at UCSF very memorable. I would like to especially thank Dr. Sook Wah Yee and Dr. Ron Keizer for their insightful scientific discussions and technical support. Daily interactions with lab members including Lawrence Lin, Chris Wen, Arik Zur, Ethan Geier, Xiaomin Liang, Adrian Stecula, Sneha Gupta, Jennifer Hibma and Imke Bartelink created a fun environment, motivating me to conduct high quality research.

The successful completion of a Ph.D. is not possible without the support from loved ones. The graduate journey is arduous, and can be mentally taxing to close friends and family. For this reason, I want to thank my brother, Abhishek Goswami, for being my number one role model ever since I was a little kid. His care, wisdom and support as an amazing older brother enabled me to successfully pursue and compete this doctoral degree. I am also incredible grateful to Ankita Iyer, my best friend, soul mate, and life partner. My day-to-day

interactions with Ankita steered me out of the low points of graduate school, successfully converting all of my stress-induced frowns into smiles. Needless to say, achieving this goal would have been impossible without her love, confidence and support.

Finally, I would like to thank my incredible father and mother, Premendra Goswami and Sampa Goswami. Their unconditional love was the stimulus behind any of my successes. My parents moved from India when I was two years old, and worked incredibly hard so that my brother and I can live a happy and prosperous life. They have succeeded beyond measure, all while creating a loving environment that nurtures academic interest. My mother is an incredibly intelligent, strong and caring woman. Her empathy, devoted love and tasty home cooked meals have maintained my mental sanity during late nights filled critical thinking and scientific analysis. My father has been the foundation to my support system. He has constantly preached for me to work hard and aim high in life, all while maintaining a healthy balance. He has encouraged me to smile, and enjoy all of the positive aspects that life has to offer. My father's calmness, optimistic attitude, and spiritual nature are key qualities that I hope to one day emulate. Ever since I was little, my father's dream was for me to successfully earn a Ph.D., so I am delighted and honored to receive this doctoral degree not just for me, but also for him.

**Novel Computational Methods to Discover Genes Linked to Drug Response**

**Srijib Goswami**

## Abstract

Metformin is used first line for treatment of type 2 diabetes (T2D) and is one of the most frequently prescribed drugs worldwide. As the global incidence of T2D rapidly increases, the low cost of metformin makes this treatment option particularly attractive in developing nations. Understanding metformin's efficacy in different patient populations with diverse genetic backgrounds will be critical in managing this deleterious metabolic disorder. The major goal of this dissertation research was to use novel, quantitative approaches to elucidate genetic and non-genetic components that predict metformin disposition and glycemic response. As a first goal, the role of transcription factor variants on metformin pharmacokinetics and pharmacodynamics was investigated. From this analysis, five variants in SP1 were significantly associated with changes in treatment HbA1c ($p < 0.01$) and metformin secretory clearance ($p < 0.05$). Genetic variants in transcription factors PPAR-alpha and HNF4-alpha were significantly associated with HbA1c change only, but were not significantly associated with pharmacokinetics. A plausible biological mechanism by which genetic variants affected the pharmacological variation of metformin was determined using gene expression levels linked to genetic variants (eQTLs). The focus was on transporter expression. From this study, we discovered that genomic regions

proximal to metformin transporters were linked to expression levels of SLC47A1, SLC22A3, and SLC22A2, with a potential transcription factor-binding hypothesis for SP1. We also found variants in transcription factor HNF4-alpha were the most influential trans-eQTLs, accounting for expression level variation in both SLC47A1 and SLC22A1. Finally, we developed a mathematical model to quantify disease progression on metformin therapy using HbA1c data with the goal of explaining long-term HbA1c variability through the investigation of genetic, demographic, and clinical factors. From this analysis, we found two SNPs in CSMD1 (rs2617102, rs2954625) and one SNP in *SLC22A2* (rs316009) as significantly influencing the long-term variance in HbA1c. Overall, this dissertation research enhances our current knowledge of the pharmacogenetic landscape by expanding the set of pharmacologically relevant genes and providing a pharmacokinetic and biological basis for some of these genes. Future research will continue to focus on replication and uncovering the mechanism driving the pharmacological genes highlighted here.

## Table of Contents

Abstract ..... vi

Chapter 1: Introduction and an Overview of Metformin Pharmacokinetics, Pharmacodynamics and Pharmacogenomics ..... 1

Chapter 2: Genetic Variants in Transcription Factors Associate with the Pharmacokinetics and Pharmacodynamics of Metformin ..... 28

Chapter 3: Discovery of Expression Quantitative Trait Loci (eQTLs) of Metformin Transporters ..... 65

Chapter 4: A Longitudinal HbA1c Model Elucidates Genes Linked to Disease Progression on Metformin ..... 83

Chapter 5: Summary and Conclusions ..... 118

## List of Tables

Table 1.1 Summary of genes and variants involved in metformin pharmacogenomics ........ 15

Table 2.1 List of transporter genetic variants investigated in the population pharmacokinetic model of metformin ........ 38

Table 2.2 Baseline characteristics of patients with type 2 diabetes and healthy volunteers dosed with metformin ........ 40

Table 2.3 Prioritized genetic variants in *SP1* and *AP2* associated with metformin pharmacodynamics and pharmacokinetics ........ 42

Table 2.4 Population pharmacokinetic model derived estimates and bootstrap results for pharmacokinetic parameters of metformin ........ 46

Table 3.1 Summary of statistically significant cis-eQTLs. ........ 72

Table 3.2. *HNF4-alpha* SNPs (P < 0.05) associated with metformin transporters. ........ 74

Table 4.1 Baseline characteristics of patients with type 2 diabetes ........ 88

Table 4.2 Population pharmacodynamic model derived estimates and bootstrap results for model parameters ........ 96

Table 4.3 Summary of top genetic variants included in final mathematical model of metformin ........ 101

## List of Figures

Figure 1.1 Pharmacokinetics pathway of metformin. .......................................... 4

Figure 1.2 Pharmacodynamics pathway of metformin. ....................................... 8

Figure 2.1. Transcription factors regulating metformin PK/PD. ......................... 33

Figure 2.2 Final pharmacokinetic model and visual predictive check. ................ 43

Figure 2.3 Summary of covariate effects on population pharmacokinetic parameters. ........................................................................................... 44

Figure 2.4. Influence of top *SP1* variants on metformin pharmacodynamics and pharmacokinetics using multivariate regression. ............................................... 49

Figure 2.5. Manhattan plots of top SNPs for *HNF4-alpha* and *PPAR-alpha*....... 51

Figure 3.1. Proposed mechanism of SNP influence on SLC expression. ........... 69

Figure 3.2. Summary of gene expression variation of metformin transporters. .. 71

Figure 3.3. Top cis-eQTLs associated with expression levels of *SLC47A1* and *SLC22A3*. .............................................................................................. 73

Figure 3.4. Association between top *HNF4-alpha* eQTLs with *SLC47A1* expression level. ........................................................................................ 75

Figure 3.5. Proposed mechanism of an *HNF4-alpha* SNP effect on metformin secretory clearance. ................................................................................ 76

Figure 4.1 Workflow of longitudinal modeling, genetic analysis and the potential clinical impact on individualizing metformin therapy. .......................................... 86

Figure 4.2 Model mechanics and interplay of disease progression, HbA1c synthesis rate and %HbA1c level over time. ...................................................... 90

Figure 4.3 Longitudinal HbA1c levels over time and model based visual predictive check. ........................................................................................ 98

Figure 4.4 Top genetic and demographic covariates on long term HbA1c levels. ....................................................................................................... 102

Figure 4.5 Effect of SNP combinations in *CSMD1, SLC22A2, and WWOX* on the dynamics of HbA1c levels. ....................................................................... 103

## Chapter 1

## Introduction and an Overview of Metformin Pharmacokinetics, Pharmacodynamics and Pharmacogenomics*

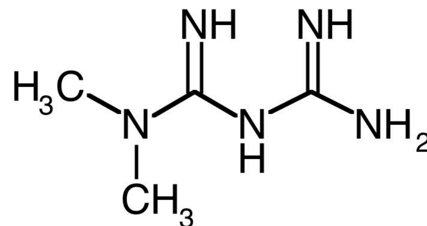
### A review of metformin pharmacokinetics (PK), pharmacodynamics (PD) and pharmacogenetics (PgX)

For this introduction, I appended sections to a review that I had published in the *Journal of Pharmacogenetics and Genomics* that describes metformin PK, PD and PGx up until 2012. After that I provide an overview of the goals of my dissertation research, highlighting the aims that will be addressed in subsequent chapters.

.

### Introduction to metformin

Metformin is a first line therapy for type 2 diabetes (T2D, formerly "non-insulin-dependent diabetes mellitus"), and is one of the most commonly prescribed drugs worldwide. As a biguanide agent, metformin lowers both basal and postprandial plasma glucose levels[1,2]. It can be used as a monotherapy or in combination with other anti-diabetic agents including sulfonylureas, alpha-glucosidase inhibitors, insulin, thiazolidinediones, DPP-4




* This chapter is a modified version of the material published in *Pharmacogenetics and Genomics* in Li Gong, Kathleen M. Giacomini, Russ B. Altman, Teri Klein.

inhibitors as well as GLP-1 agonists. Metformin works by inhibiting hepatic glucose production, reducing intestinal glucose absorption and improving glucose uptake and utilization. Besides lowering blood glucose level, metformin may have additional health benefits, including weight reduction, lowering plasma lipid levels, and prevention of some vascular complications[3]. As obesity rates in the United States rise, the use of metformin is also increasing. Metformin is also used for other indications such as polycystic ovary syndrome[1]. Metformin has also been studied in several cancers. In one meta-analysis, metformin was associated with a decreased risk of cancer incidence compared with other treatments among diabetic patients[4]. Also for Alzheimer's disease (AD), one article concluded that there was a slightly higher risk of AD in long-term users of metformin[5].

Overall, metformin is well tolerated by the majority of patients. However, the glycemic response to metformin is quite variable. Some patients respond extremely well while others show no benefit[6]. The following sections will provide an overview of metformin PK, PD and PGx.

**A review of metformin pharmacokinetics**

Metformin is not metabolized[1,7] and is excreted unchanged in the urine with a half-life of approximately 5 hours[7]. The population mean for renal clearance is 510 +/- 120 mL/min. Active tubular secretion in the kidney is the principle route of metformin elimination. The drug is widely distributed into body tissues including intestine, liver and kidney via various organic cation

transporters[7]. There is large inter-individual variability in metformin pharmacokinetics as measured by differences in trough steady-state metformin plasma concentrations, which range from 54-4133 ng/ml[8].

The absorption of metformin is incomplete with some literature suggesting a dose-dependent absolute bioavailability that decreases with increasing doses. This effect is probably attributable to a relatively slow absorption rate together with an effective absorption window that is confined to the small intestine, where a saturable transport mechanism may occur[1,9]. At normal doses, the bioavailability (F) ranges between 50 and 60%. The intestinal absorption of metformin may be mediated by plasma membrane monoamine transporter (PMAT, encoded by gene *SLC29A4*), which is expressed on the luminal side of enterocytes[10] (Figure 1.1).

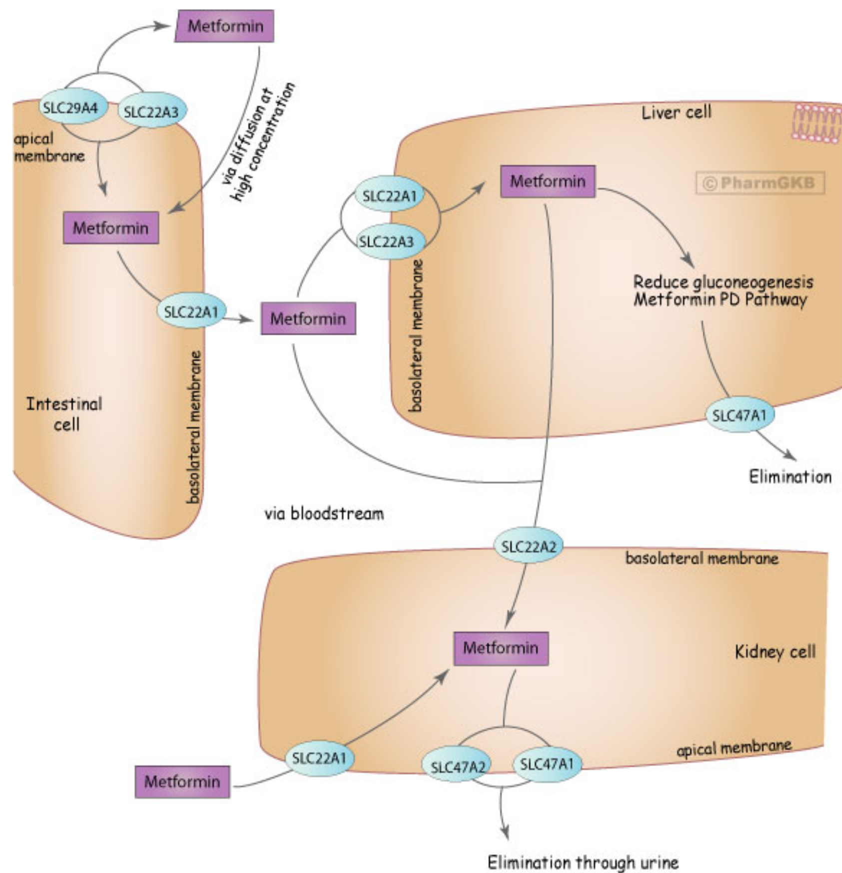



**Figure 1.1 Pharmacokinetics pathway of metformin.**
Stylized cells depicting genes involved in the transport and clearance of metformin. A fully interactive version is available online at http://www.pharmgkb.org/pathway/PA165948259.

However, there are currently no *in vivo* data on the role of PMAT in the disposition and pharmacological effect of metformin. OCT3 (gene *SLC22A3*) is also expressed in the brush border of the enterocytes and may contribute to metformin absorption[7,11]. In fact, a recent study from our laboratory in Oct3-/- mice demonstrated that the transporter plays an important role in metformin bioavailability[12]. Additionally, OCT1 (gene *SLC22A1*), which is expressed on the basolateral membrane of enterocytes, may facilitate transfer of metformin into the

interstitial fluids surrounding portal capillaries[11]. The role of OCT1 and OCT3 in intestinal transport of metformin remains to be defined.

The liver plays an important role in the pharmacological action of metformin. The hepatic uptake of metformin is mediated primarily by OCT1 (*SLC22A1*), and to a lesser extent by OCT3 (*SLC22A3*). Both transporters are expressed on the basolateral membrane of hepatocytes[7,13–15]. In *Oct1*-deficient mice, the hepatic metformin concentration in the liver is significantly lower than control mice, suggesting that OCT1 is essential for hepatic uptake of metformin[16]. Metformin is also a good substrate for human multidrug and toxin extrusion 1 (MATE1, encoded by gene *SLC47A1*) and MATE2-K (gene *SLC47A2*)[13,17,18,19]. MATE1 *(SLC47A1)* is highly expressed in the liver, kidney and skeletal muscle[20] and may contribute to the excretion of metformin from both liver and kidney. However, MATE1's role in hepatic secretion has been questioned, as biliary excretion of metformin seems to be insignificant in humans[7]. Data from a *Mate1* knockout mouse study suggest that, at least in rodents, biliary excretion of metformin occurs and is mediated in part by MATE1[21].

The uptake of metformin from circulation into renal epithelial cells is primarily facilitated by OCT2 (gene *SLC22A2*)[13], which is expressed predominantly at the basolateral membrane in the proximal tubule. Renal excretion of metformin from tubule cell to lumen is mediated via MATE1 (*SLC47A1*), MATE2, and MATE2-K (*SLC47A2*)[17,18,22,23]. MATE1 and MATE2 are expressed on the apical membrane of renal proximal tubule cells and studies in

healthy individuals suggest that they contribute to the renal excretion of metformin[24].

OCT1 also appears to be expressed on the apical and subapical side of both the proximal and distal tubules in the kidney, and may play an important role in metformin reabsorption in kidney tubules[25]. Plasma membrane monoamine transporter (PMAT, gene *SLC29A4*) is expressed on the apical membrane of renal epithelial cells, and may play a role in renal reabsorption of metformin[26]. However, there are no in vivo data yet supporting this role. Additionally, P-gp (gene *ABCB1*) and BCRP (gene *ABCG2*) are involved in the efflux of metformin across placental apical membranes[27].

Since metformin is not metabolized in the liver, drug-drug interactions via the inhibition of metformin transporters (OCTs and MATEs) are clinically relevant. Genetic polymorphisms in these transporter genes are also likely to have a direct impact on metformin pharmacokinetics and variability in drug responses (see *Pharmacogenomics* section). Recent drug-drug interaction studies suggest that proton pump inhibitors inhibit metformin uptake *in vitro* by inhibiting OCT1, OCT2 and OCT3. Oral anti-diabetic drugs repaglinide and rosiglitazone also inhibited OCT1 mediated metformin transport *in vitro*[28]. The H2 blocker, cimetidine is associated with reduced renal tubular secretion and increased systemic exposure to metformin when both drugs are co-administered[29]. Inhibition of MATEs, but not OCT2[30], is the likely mechanism underlying the drug-drug interactions with cimetidine in renal elimination[23]. A recent study suggests the potential for a transporter mediated drug-drug interaction between metformin and

specific tyrosine kinase inhibitors (e.g. imatinib, nilotinib, gefitinib, and erlotinib), which may have clinical implications on the disposition, efficacy, and toxicity of metformin[31]. Several of the compounds were shown to inhibit the transporters *in vitro* at clinically relevant concentrations.

**A review of metformin pharmacodynamics**

Metformin lowers both basal and postprandial plasma glucose. It works mainly by suppressing excessive hepatic glucose production, through a reduction in gluconeogenesis[32]. Other potential effects of metformin include: an increase in glucose uptake, increase in insulin signaling, decrease in fatty acid and triglyceride synthesis, and an increase in fatty acid beta-oxidation. Metformin may also increase glucose utilization in peripheral tissues, and possibly reduce food intake and intestinal glucose absorption. As metformin does not stimulate endogenous insulin secretion, it does not cause hypoglycemia or hyperinsulinemia, which are common side effects associated with other anti-diabetic drugs.

The molecular mechanisms underlying metformin action appear to be complex and remain a topic of much debate. Metformin's pharmacologic effects reflect its effects on cellular energy content. Studies in the early 2000's suggested that metformin inhibits Complex I in the mitochondria of the liver, reducing ATP levels and increasing the AMP/ATP ratio[33]; however, recent studies suggest that metformin reduces glycolytic energy by affecting hepatic and intestinal thiamine disposition[34]. Further, a study published in *Nature*

demonstrates that the drug inhibits glycerophosphate dehydrogenase, disrupting the electron transport chain and the proton gradient needed for ATP generation[35].

Irrespective of how the drug reduces cellular energy, there is general agreement that metformin administration results in phosphorylation and activation of AMP-activated protein kinase (AMPK) in the liver, which in turn may lead to diverse pharmacologic effects, including inhibition of glucose and lipid synthesis[2,36]. Although the specific route of AMPK phosphorylation is not yet clear, molecular components LKB1/STK11 and ATM have been demonstrated to play a role in the phosphorylation of AMPK in the presence of metformin[36] (Figure 1.2).

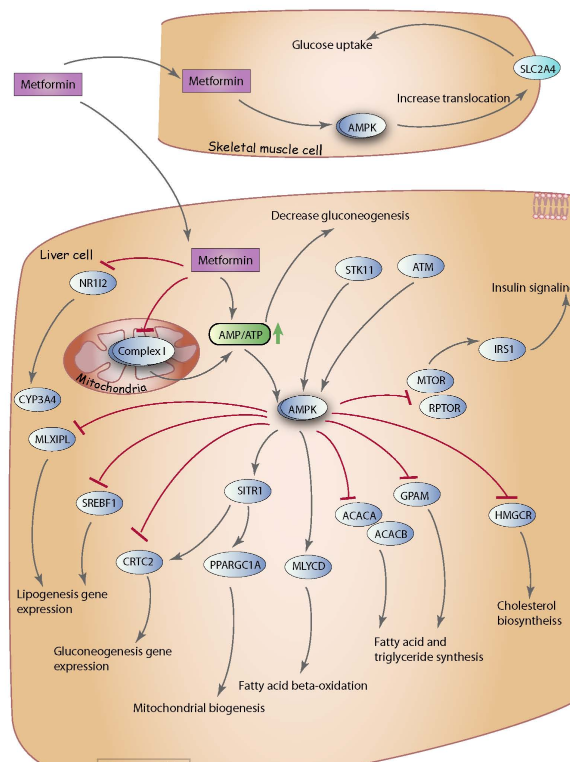



**Figure 1.2 Pharmacodynamics pathway of metformin.**
Stylized cell depicting the mechanism of action of metformin. A fully interactive version is available online at http://www.pharmgkb.org/pathway/PA165948566.

However, ATM, LKB1 and AMPK are not the direct targets of metformin[37]. A recent study using liver-specific AMPK-knockout mice demonstrated that inhibition of hepatic glucose production by metformin is preserved, suggesting that metformin may inhibit hepatic gluconeogenesis in an LKB1- and AMPK-independent manner[38]. In separate study in *Oct-1* knockout mice, metformin both activated AMPK and reduced gluconeogenesis[16]. Another group has also concluded that metformin inhibits hepatic gluconeogenesis through AMPK dependent regulation of SHP[39]. Furthermore, a reduction in gluconeogenesis may occur both ways, in an AMPK dependent and independent manner.

Although the direct target is not fully clear, metformin specifically inhibits ATP production suggesting that this inhibition may activate AMPK by increasing the cellular AMP:ATP ratio[33,37,40,41]. AMPK is a major cellular regulator of lipid and glucose metabolism. The activated AMPK phosphorylates and inactivates HMG-CoA reductase (encoded by gene *HMGCR*), mTOR (target of rapamycin); ACC-2 (encoded by gene *ACACB*); ACC (encoded by gene *ACACA*), glycerol-3-phosphate acyltransferase (encoded by gene *GPAM*), and carbohydrate response element binding protein[36,42]. Activation of AMPK by metformin also suppresses the expression of SREBP-1 (encoded by gene *SREBF1*), a key lipogenic transcription factor[43]. Phosphorylated AMPK also activates SiRT1 and increases Pgc-1a (encoded by gene *PPARGC1A*) expression in the nucleus, leading to the downstream activation of mitochondrial biogenesis. Metformin disrupts the co-activation of PXR with SRC1, resulting in down regulation of CYP3A4 gene expression[44]. Finally, activated AMPK results in an increase in

glucose uptake in skeletal muscle via increasing the GLUT4 (encoded by gene *SLC2A4*) translocation activity[16]. The overall pharmacological effect of AMPK activation in the liver includes the stimulation of fatty acid (FA) oxidation with inhibition of cholesterol and triglyceride synthesis. Peripheral effects include stimulation of FA oxidation and glucose uptake in skeletal muscle as well as a systemic increase in insulin sensitivity[40]. However, the role of metformin in insulin-mediated glucose uptake has been debated[45].

Given the increased risk of cancer in type 2 diabetes patients, metformin has also been evaluated for its tumor suppression ability and its potential to protect from cancer[46]. Population studies have shown that metformin is associated with a significant reduction of neoplasia in multiple cancer types (cancer of the breast and prostate, in particular)[47]. Metformin may also inhibit the growth of cancer cells. The mechanisms underlying this protective effect are not well understood and may involve activation of multiple pathways as well as changes in glucose utilization and energy production[2,46]. The cell cycle arrest in metformin treated breast cancer cells seems to involve activation of AMPK, down regulation of cyclin D1, and requires p27Kip1 or p21Cip1[48,49]. Metformin was reported to suppress HER2 (ERBB2) oncoprotein overexpression via inhibition of the mTOR effector p70S6K1/ RPS6KB1 in human breast carcinoma cells[50].

**A review of metformin pharmacogenomics**

The role of genetic factors in predicting response variation to metformin has been the subject of many investigations. Multiple studies reported

associations between genomic variations of metformin transporters and its pharmacokinetics and response, and a few have explored the role of pharmacodynamic genes/variants in drug efficacy. However, the clinical relevance of these variants remains to be established in large-scale studies. Currently, no validated genetic predictor is used in the clinic.

Over the past few years, progress has been made in understanding the effect of common genetic polymorphisms in transporter genes on modulation of metformin pharmacokinetics. Much work has been done with the organic cation transporter family (*SLC22A* family) (reviewed by Nies et al.[51]). OCT1 (gene *SLC22A1*) is essential for the hepatic uptake of metformin[7]. In one study with 20 healthy volunteers, several genetic variants of *OCT1*: R61C (rs12208357), G401S (rs34130495), 420del (rs142448543 or rs34305973 or rs35191146), and G465R (rs34059508) had a significant effect on the pharmacokinetics of metformin after oral administration. Individuals carrying any of the reduced function OCT1 alleles demonstrated a higher area under the concentration–time curve (AUC), higher maximal plasma concentration ($C_{max}$) and a lower volume of distribution (V) compared to individuals carrying wild type alleles[16]. A subsequent study in 103 healthy male Caucasians demonstrated the impact of these low activity alleles on pharmacokinetics, with increased renal clearances ($CL_r$) and decreased hepatic uptake. However, unlike the earlier study, the reduced function allele did not lead to differences in metformin exposure (AUC)[25]. A recent study by Christensen et al. demonstrated that reduced functional alleles of *OCT1* were associated with decreased trough steady-state concentration of

metformin and a reduction in the absolute decrease in Hb1Ac during the initiation as well as maintenance period[8]. Overall, replication of OCT1 low activity alleles on governing metformin disposition highlights the importance of these genetic variants on pharmacokinetic impact and may be taken into consideration for metformin therapy.

Studies in healthy volunteers have tested the effect of genetic variants in OCT2 *(*gene *SLC22A2)* on metformin pharmacokinetics. Genetic variants of OCT2 (c.596C>T, c602C>T, and c.808G>T (rs316019)) were associated with differences in pharmacokinetics, compared to the reference genotype with an increase in AUC and $C_{max}$ and a decrease in renal clearance[52]. A follow up study in 15 healthy Chinese participants observed that rs316019 (808G>T) is associated with a reduced $CL_r$ but not overall drug exposure[30]. Interestingly, in a separate healthy volunteer study consisting of Caucasian and African-Americans, individuals heterozygous for the variant 808G/T had higher metformin renal clearances than the reference group[53]. Similar to OCT1, the impact of OCT2 genetic variants is replicated, strongly suggesting the importance of these alleles on determining metformin exposure. There has been limited research conducted on genetic variants of MATE1 or MATE2K for explaining differences in metformin pharmacokinetics. In a more recent study, MATE2-K variants (rs578427 and rs34834489) showed that common promoter haplotypes of MATE2-K were associated with increased renal and secretory clearance[5].

As far as DDIs are concerned, the importance of MATE1 and MATE2K are clear. In one healthy volunteer study, administration of a MATE inhibitor,

pyrimethamine, caused significant increases in metformin $C_{max}$ and AUC. *In vivo* studies also demonstrated the importance of the rodent Mate1 in modulating the pharmacokinetics of metformin through gene knockout[22].

In addition to pharmacokinetics, a number of studies have been conducted investigating the role of genetic variants on metformin pharmacodynamics and response. Despite having an effect on renal clearance, well-established genetic polymorphisms of OCT1 and OCT2 that alter metformin disposition do not sufficiently explain the broad variation in clinical efficacy[54,55]. A pharmacogenetic study in healthy volunteers demonstrated significant clinical effects of reduced function OCT1 variants (R61C, G401S, 420del and G465R), causing an impaired response to a glucose tolerance test[16]. A prospective study conducted in patients with polycystic ovary syndrome concluded that genetic variations in OCT1 may be associated with heterogeneity in metabolic response to metformin[56]. Interestingly, the minor allele of an intronic variant of MATE1/*SLC47A1*, rs2289669 G>A, was significantly associated with a greater reduction in hemoglobin A1c (HbA1c), in a cohort of 116 metformin users, despite the lack of association between the polymorphism and metformin $CL_r$ or other pharmacokinetic parameters[57]. In a meta-analysis by Jablonski et al., the minor allele of rs8065082, a SNP in LD with MATE1 intronic SNP rs2289669, was associated with reduced diabetes incidence in patients taking metformin[58]. A recent study by Choi et al., also observed the association between this *MATE1* intronic variant with a change in HbA1c level, almost at the level of statistical significance[59]. This evidence, along with the relatively large sample sizes,

provides strong support for the functional impact of this MATE1 intronic variant. However, the missing link between pharmacokinetics and a reduction in HbA1c requires further research regarding this SNP and it's mechanistic role.

Recently, a study by Choi et al. showed diabetic patients who were homozygous for g.-130G>A (rs12943590) in MATE2-K had a significantly poorer response to metformin treatment, assessed by the relative change in glycated hemoglobin[59]. In addition to the aforementioned transporters, the effect of variations in *OCT3* has also been investigated. An *in vitro* study showed that OCT3 (gene *SLC22A3*) may also play a role in the therapeutic action of metformin[14]. OCT inhibitor, such as OCT3-specific short hairpin RNA, significantly reduced the activating effect of metformin on AMPK in skeletal muscle cells. Also, genetic variants of OCT3 (T400I (rs8187725), V423F and T44M (rs8187715)) significantly impacted metformin uptake and kinetics. In addition to transporters, a SNP in serine racemase (*SRR*), rs391300, demonstrated an association with serum fasting plasma glucose (FPG), postprandial plasma glucose (PPG), and cholesterol (CHO) in 402 Chinese patients and 171 healthy controls taking metformin[60]. This discovery, although yet to be replicated, provides promise for the discovery of other genetic variants that may affect clinical outcome. The exploration of gene-gene interaction may be a promising new area of research. In a study by Becker et al.[61], an interaction between two polymorphisms, rs622342 in OCT1 and rs2289669 in MATE1 was reported, suggesting that interactions between genes in the metformin pathway may impact metformin response. However, this study is small and the importance

of epistatic mechanism remains to be replicated. More recently in a study by Stocker et al. findings suggested that promoter variants of MATE1 (rs2252281) and MATE2 (rs12943590) were important determinants of metformin disposition and response in healthy volunteers and diabetic patients. Interestingly, the renal and secretory clearances of metformin were higher (22% and 26%) respectively in carriers of variant MATE2, who were also MATE1 reference. These pharmacokinetic results were consistent with metformin response, with variant carriers of MATE1 and MATE2 having an enhanced and reduced response respectively[62]. Table 1.1 summarizes some of the important and more recent PGx findings for metformin.

**Table 1.1** Summary of genes and variants involved in metformin pharmacogenomics

| Gene | Variant | Associated phenotype | PMID | PK/PD |
|---|---|---|---|---|
| *SLC22A1* (OCT1) | Reduced function alleles: R61C (rs12208357), G401S (rs34130495), 420del (rs142448543 or rs34305973 or rs35191146), and G465R (rs34059508) | High AUC, higher maximal plasma concentration (Cmax), and lower oral volume of distribution (V/F) | 17609683 | PK |
| | | Impaired response to a glucose tolerance test | 17609683 | PD |
| | | Increased renal clearances (CLr) and decreased hepatic uptake, no exposure changes (AUC) | 19536068 | PK |
| | | Reduced lipid response (total cholesterol and triglycerides) and insulin responses to metformin in PCOS women | 20660041 | PD |
| | rs622342 | Decreased CL renal reference and lower and CL sec reference | 23873119 | PK |
| | rs72552763 deletion, rs34130495 and other reduced function alleles | Reduced trough metformin steady-state concentration, and association with the initial absolute decrease in HbA1c | 21989078 | PK, PD |
| *SLC22A2* (OCT2) | 596C>T, 602C>T, and 808G>T (rs316019) | Increase in AUC and Cmax and a decrease in renal clearance | 18401339 | PK |
| | rs316019 (808G>T) | Reduced CLr but no effect on overall drug exposure | 18551044 | PK |
| | rs316019 (808G>T) | Higher metformin renal clearances | 19483665 | PK |

| *SLC22A3* (OCT3) | T400I (rs8187725), V423F and T44M (rs8187715) | Reduced metformin uptake | 20859243 | PK, PD |
|---|---|---|---|---|
| *SLC47A1* (MATE1) | rs2289669 | Reduction in hemoglobin A1c (HbA1c) | 19228809 | PD |
| | | No observed association with metformin CLr or other pharmacokinetic parameters | 19228809 | PK |
| | rs2252281 | Enhanced metformin response in healthy volunteers (post-metformin glucose tolerance) and in patients (relative change in HbA1c levels) | 23267855 | PD, PK |
| | rs8065082 | Reduced diabetes incidence | 20682687 | PD |
| *SLC47A2* (MATE2-K) | rs12943590 (-130G>A) | Poorer response to metformin treatment, assessed by the relative change in glycated hemoglobin | 21956618 | PD |
| | | Renal and secretory clearance of metformin were higher in carriers of variant who were also reference MATE 1 (rs2252281); altered metformin glucose tolerance | 23267855 | PD, PK |
| | rs34834489 | Increased renal clearance and secretion clearance of metformin when administered in healthy individuals | 23652408 | PK |
| | rs578427 | Increased renal clearance and secretion clearance of metformin when administered in healthy individuals | 23652408 | PK |
| *SRR* | rs391300 | Associated with levels of FPG, PPG, and CHO | 21933224 | PD |
| *ATM* | rs11212617 | Associated with metformin treatment success (Hba1c < 7%) | 21186350 | PD |
| *LKB/STK11* | rs8111699 | C allele associated with a significantly decreased chance of ovulation in PCOS women | 18681789; 18000088 | PD |
| *CAPN* | rs3792269 | Decreased response to metformin in people with Diabetes Mellitus | 25327507 | PD |
| *PPARG* | rs1801282 | Increased likelihood of short and long-term weight loss when treated with metformin | 22179955 | PD |
| *SP1* | rs784888 | Decreased post-HbA1c levels when treated with metformin in people with diabetes | 24853734 | PD |
| | | Decreased metformin secretory clearance when exposed to metformin | 24853734 | PK |
| *PPARA* | rs149711321 | Decreased post-HbA1c levels when treated with metformin in patients with diabetes | 24853734 | PD |

The first genome wide association study on metformin response by GoDARTs and UKPDS and WTCCC2 investigated 1024 Scottish individuals with T2D, and was replicated in two cohorts including 1,783 Scottish individuals and

1,113 individuals from a UK prospective study[63]. The study discovered that common variants near the *ATM* (Ataxia Telangiectasia Mutated) locus were associated with glycemic response to metformin. The genes near this locus include: *CUL5, NPAT, C11org65, EXPH5, ACAT1, and KDELC2*. The minor allele (C) of the most strongly associated SNP, rs11212617, had a population frequency of 44% and was associated with treatment success (achieving HbA1c < 7%). In the meta-analysis, SNP rs11212617 was significantly associated with treatment success with an odds ratio of 1.35. Despite the strong association, the SNP only accounts for 2.5% of the observed variability in glycemic response. *ATM* was selected as a causative gene due to its role in insulin resistance, increased risk of diabetes and its role in AMPK activation. Furthermore, *in vitro* functional studies performed by this group demonstrated that inhibition of ATM by a chemical inhibitor (KU-55933) attenuated the metformin-induced phosphorylation and activation of AMPK. However, recent data suggest that this ATM inhibitor also inhibits OCT1 and may have acted through inhibition of metformin uptake rather than inhibition of ATM[64]. Furthermore, a study by Florez et al. reported that the association between rs11212617 with metformin response was not confirmed in the diabetes prevention program (DPP) cohort[65]. Overall, this recent finding strongly suggests that the effect of ATM on activating AMPK and altering pharmacological outcomes is far from conclusive.

In addition to the treatment of diabetes, metformin is used in the treatment of insulin resistance in individuals with polycystic ovary syndrome (PCOS). A small study demonstrated that a polymorphism in *LKB/STK11 (*rs8111699) is

associated with ovulatory response to treatment with metformin alone in a prospective randomized trial; with the C allele associated with a significantly decreased chance of ovulation in PCOS women treated with metformin[66,67].

**Research focus**

Although many new drugs have been developed for T2D, metformin is still widely accepted as first line therapy due to its low incidence of micro- and macro-vascular events and its beneficial effects on plasma lipids and body weight. There is no validated genetic predictor of metformin response or pharmacokinetics, and the data suggest that epistatic mechanisms (gene-gene interactions) may be important. Investigation of genetic variants in specific patient populations (e.g. stratification by ethnicity), considerations of response dynamics, as well as the investigation of gene-gene and gene-environment interactions may elucidate important determinants governing the pleiotropic nature of response. Overall, enhancing our understanding of the factors underlying response variability will have a significant impact on the diabetic community, resulting in the identification of disease subtypes and biomarkers that may lead to downstream dosage adjustments or administration of other, more personalized anti-diabetic drugs.

**Research hypothesis**: Our research aim is to use novel, quantitative approaches in order to elucidate crucial genetic and non-genetic components in pharmacological pathways that predict metformin disposition and glycemic response. At a high level, the goal of this research is to combine pharmacometric

approaches with genetic analysis techniques to describe and quantify the pleiotropic nature of metformin pharmacokinetics and pharmacodynamics. A special focus of this research will be on interrogating genetic variants in transporter and transcription factor genes. Secondly, the biological mechanism of genetic variants will be explored by analyzing the link of prioritized variants with gene expression levels of metformin transporters. Finally, longitudinal changes in HbA1c levels in patients with heterogeneous genetic and demographic makeup will be quantified using semi-mechanistic modeling approaches. The downstream vision of this research is to translate research findings into clinical practice, enabling clinicians to provide personalized treatment advice to TII Diabetes patients. My dissertation research had three aims:

**Aim 1)** To investigate the roles of transcription factor variants on metformin pharmacokinetics and pharmacodynamics. A pharmacokinetic model was developed to investigate the role of prioritized genetic variants on the kinetics of metformin.

**Aim 2)** To discover expression quantitative trait loci (eQTLs) that are linked to the gene expression levels of metformin transporters. A total of sixty kidney tissue samples were used to identify *cis* and *trans* genomic regions linked to gene expression levels of primary metformin transporters.

**Aim 3)** To develop a longitudinal HbA1c model that characterizes the underlying disease progression and the time course of glycosylated hemoglobin (HbA1c) in relation to a patient's exposure to metformin therapy (PK).

Chapter 2 addresses Aim 1, where we investigated the role of gene expression modulators (transcription factors) of metformin transporter genes on the pharmacokinetics and pharmacodynamics of metformin. We explored a pharmacokinetic mechanism for these SNP associations by developing a population pharmacokinetic model of metformin in healthy volunteers and Type 2 diabetic patients. Chapter 3 aims to profile SNPs that are linked to the gene expression levels of metformin transporter genes (e.g. *SLC22A1*, *SLC22A2* etc.) by performing a computational analysis on genetic and gene expression levels in kidney tissues. Finally, in chapter 3, we develop a longitudinal model to characterize and quantify disease progression on metformin therapy, with the goal of explaining temporal HbA1c variability through the investigation of genetic, demographic and clinical factors.

## REFERENCES


1. Scarpello, J. H. & Howlett, H. C. Metformin therapy and clinical uses. *Diab. Vasc. Dis. Res.* **5,** 157–67 (2008).

2. Viollet, B., Guigas, B., Sanz Garcia, N., Leclerc, J., Foretz, M. & Andreelli, F. Cellular and molecular mechanisms of metformin: an overview. *Clin. Sci. (Lond).* **122,** 253–70 (2012).

3. DeFronzo, R. A. & Goodman, A. M. Efficacy of metformin in patients with non-insulin-dependent diabetes mellitus. The Multicenter Metformin Study Group. *N. Engl. J. Med.* **333,** 541–49 (1995).

4. Decensi, A., Puntoni, M., Goodwin, P., Cazzaniga, M., Gennari, A., Bonanni, B. & Gandini, S. Metformin and cancer risk in diabetic patients: a systematic review and meta-analysis. *Cancer Prev. Res. (Phila).* **3,** 1451–61 (2010).

5. Chung, J.-Y., Cho, S. K., Kim, T. H., Kim, K. H., Jang, G. H., Kim, C. O., Park, E.-M., Cho, J.-Y., Jang, I.-J. & Choi, J. H. Functional characterization

of MATE2-K genetic variants and their effects on metformin pharmacokinetics. *Pharmacogenet. Genomics* **23,** 365–73 (2013).

6. Reitman, M. L. & Schadt, E. E. Pharmacogenetics of metformin response: a step in the path toward personalized medicine. *J. Clin. Invest.* **117,** 1226–29 (2007).

7. Graham, G. G., Punt, J., Arora, M., Day, R. O., Doogue, M. P., Duong, J. K., Furlong, T. J., Greenfield, J. R., Greenup, L. C., Kirkpatrick, C. M., Ray, J. E., Timmins, P. & Williams, K. M. Clinical pharmacokinetics of metformin. *Clin. Pharmacokinet.* **50,** 81–98 (2011).

8. Christensen, M. M. H., Brasch-Andersen, C., Green, H., Nielsen, F., Damkier, P., Beck-Nielsen, H. & Brosen, K. The pharmacogenetics of metformin and its impact on plasma metformin steady-state levels and glycosylated hemoglobin A1c. *Pharmacogenet. Genomics* **21,** 837–50 (2011).

9. Meinicke, T. Bioequivalence of Glucophage® (Metformin) tablets from Europe and the United States tested in healthy volunteers. *J. Bioequiv. Availab.* **6,** 61–66 (2014).

10. Zhou, M., Xia, L. & Wang, J. Metformin transport by a newly cloned proton-stimulated organic cation transporter (plasma membrane monoamine transporter) expressed in human intestine. *Drug Metab. Dispos.* **35,** 1956–62 (2007).

11. Müller, J., Lips, K. S., Metzner, L., Neubert, R. H. H., Koepsell, H. & Brandsch, M. Drug specificity and intestinal membrane localization of human organic cation transporters (OCT). *Biochem. Pharmacol.* **70,** 1851–60 (2005).

12. Chen, E. The role of organic cation transporters in the pharmacokinetics and pharmacodynamics of metformin. (University of California, 2015). at <http://search.proquest.com/docview/1655359770?accountid=14525>

13. Takane, H., Shikata, E., Otsubo, K., Higuchi, S. & Ieiri, I. Polymorphism in human organic cation transporters and metformin action. *Pharmacogenomics* **9,** 415–22 (2008).

14. Chen, L., Pawlikowski, B., Schlessinger, A., More, S. S., Stryke, D., Johns, S. J., Portman, M. A., Chen, E., Ferrin, T. E., Sali, A. & Giacomini, K. M. Role of organic cation transporter 3 (SLC22A3) and its missense variants in the pharmacologic action of metformin. *Pharmacogenet. Genomics* **20,** 687–99 (2010).

15. Nies, A. T., Koepsell, H., Winter, S., Burk, O., Klein, K., Kerb, R., Zanger, U. M., Keppler, D., Schwab, M. & Schaeffeler, E. Expression of organic cation transporters OCT1 (SLC22A1) and OCT3 (SLC22A3) is affected by genetic factors and cholestasis in human liver. *Hepatology* **50,** 1227–40 (2009).

16. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 ( OCT1 ) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

17. Tsuda, M., Terada, T., Ueba, M., Sato, T., Masuda, S., Katsura, T. & Inui, K. Involvement of human multidrug and toxin extrusion 1 in the drug interaction between cimetidine and metformin in renal epithelial cells. *J. Pharmacol. Exp. Ther.* **329,** 185–91 (2009).

18. Sato, T., Masuda, S., Yonezawa, A., Tanihara, Y., Katsura, T. & Inui, K.-I. Transcellular transport of organic cations in double-transfected MDCK cells expressing human organic cation transporters hOCT1/hMATE1 and hOCT2/hMATE1. *Biochem. Pharmacol.* **76,** 894–903 (2008).

19. Tanihara, Y., Masuda, S., Sato, T., Katsura, T., Ogawa, O. & Inui, K.-I. Substrate specificity of MATE1 and MATE2-K, human multidrug and toxin extrusions/H(+)-organic cation antiporters. *Biochem. Pharmacol.* **74,** 359–71 (2007).

20. Otsuka, M., Matsumoto, T., Morimoto, R., Arioka, S., Omote, H. & Moriyama, Y. A human transporter protein that mediates the final excretion step for toxic organic cations. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 17923–48 (2005).

21. Ito, S., Kusuhara, H., Kuroiwa, Y., Wu, C., Moriyama, Y., Inoue, K., Kondo, T., Yuasa, H., Nakayama, H., Horita, S. & Sugiyama, Y. Potent and specific inhibition of mMate1-mediated efflux of type I organic cations in the liver and kidney by pyrimethamine. *J. Pharmacol. Exp. Ther.* **333,** 341–50 (2010).

22. Tsuda, M., Terada, T., Mizuno, T., Katsura, T., Shimakura, J. & Inui, K. Targeted disruption of the multidrug and toxin extrusion 1 (mate1) gene in mice reduces renal secretion of metformin. *Mol. Pharmacol.* **75,** 1280–86 (2009).

23. Ito, S., Kusuhara, H., Yokochi, M., Toyoshima, J., Inoue, K., Yuasa, H. & Sugiyama, Y. Competitive inhibition of the luminal efflux by multidrug and toxin extrusions, but not basolateral uptake by organic cation transporter 2, is the likely mechanism underlying the pharmacokinetic drug-drug

interactions caused by cimetidine in the kidney. *J. Pharmacol. Exp. Ther.* **340,** 393–403 (2012).

24. Kusuhara, H., Ito, S., Kumagai, Y., Jiang, M., Shiroshita, T., Moriyama, Y., Inoue, K., Yuasa, H. & Sugiyama, Y. Effects of a MATE protein inhibitor, pyrimethamine, on the renal elimination of metformin at oral microdose and at therapeutic dose in healthy subjects. *Clin. Pharmacol. Ther.* **89,** 837–44 (2011).

25. Tzvetkov, M. V, Vormfelde, S. V, Balen, D., Meineke, I., Schmidt, T., Sehrt, D., Sabolić, I., Koepsell, H. & Brockmöller, J. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. *Clin. Pharmacol. Ther.* **86,** 299–306 (2009).

26. Xia, L., Engel, K., Zhou, M. & Wang, J. Membrane localization and pH-dependent transport of a newly cloned organic cation transporter (PMAT) in kidney cells. *Am. J. Physiol. Renal Physiol.* **292,** F682–90 (2007).

27. Hemauer, S. J., Patrikeeva, S. L., Nanovskaya, T. N., Hankins, G. D. V & Ahmed, M. S. Role of human placental apical membrane transporters in the efflux of glyburide, rosiglitazone, and metformin. *Am. J. Obstet. Gynecol.* **202,** 1–7 (2010).

28. Bachmakov, I., Glaeser, H., Fromm, M. F. & König, J. Interaction of oral antidiabetic drugs with hepatic uptake transporters: focus on organic anion transporting polypeptides and organic cation transporter 1. *Diabetes* **57,** 1463–69 (2008).

29. Somogyi, A., Stockley, C., Keal, J., Rolan, P. & Bochner, F. Reduction of metformin renal tubular secretion by cimetidine in man. *Br. J. Clin. Pharmacol.* **23,** 545–51 (1987).

30. Wang, Z.-J., Yin, O. Q. P., Tomlinson, B. & Chow, M. S. S. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. *Pharmacogenet. Genomics* **18,** 637–45 (2008).

31. Minematsu, T. & Giacomini, K. M. Interactions of tyrosine kinase inhibitors with organic cation transporters and multidrug and toxic compound extrusion proteins. *Mol. Cancer Ther.* **10,** 531–39 (2011).

32. Hundal, R. S., Krssak, M., Dufour, S., Laurent, D., Lebon, V., Chandramouli, V., Inzucchi, S. E., Schumann, W. C., Petersen, K. F., Landau, B. R. & Shulman, G. I. Mechanism by which metformin reduces glucose production in type 2 diabetes. *Diabetes* **49,** 2063–69 (2000).

33. Owen, M. R., Doran, E. & Halestrap, A. P. Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. *Biochem. J.* **348 Pt 3,** 607–14 (2000).

34. Chen, L., Shu, Y., Liang, X., Chen, E. C., Yee, S. W., Zur, A. A., Li, S., Xu, L., Keshari, K. R., Lin, M. J., Chien, H.-C., Zhang, Y., Morrissey, K. M., Liu, J., Ostrem, J., Younger, N. S., Kurhanewicz, J., Shokat, K. M., Ashrafi, K. & Giacomini, K. M. OCT1 is a high-capacity thiamine transporter that regulates hepatic steatosis and is a target of metformin. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 9983–88 (2014).

35. Madiraju, A. K., Erion, D. M., Rahimi, Y., Zhang, X.-M., Braddock, D. T., Albright, R. A., Prigaro, B. J., Wood, J. L., Bhanot, S., MacDonald, M. J., Jurczak, M. J., Camporez, J.-P., Lee, H.-Y., Cline, G. W., Samuel, V. T., Kibbey, R. G. & Shulman, G. I. Metformin suppresses gluconeogenesis by inhibiting mitochondrial glycerophosphate dehydrogenase. *Nature* **510,** 542–46 (2014).

36. Foretz, M. & Viollet, B. Regulation of hepatic metabolism by AMPK. *J. Hepatol.* **54,** 827–29 (2011).

37. Hardie, D. G. Neither LKB1 nor AMPK are the direct targets of metformin. *Gastroenterology* **131,** 974–75 (2006).

38. Foretz, M., Hébrard, S., Leclerc, J., Zarrinpashneh, E., Soty, M., Mithieux, G., Sakamoto, K., Andreelli, F. & Viollet, B. Metformin inhibits hepatic gluconeogenesis in mice independently of the LKB1/AMPK pathway via a decrease in hepatic energy state. *J. Clin. Invest.* **120,** 2355–69 (2010).

39. Kim, Y. D., Park, K.-G., Lee, Y.-S., Park, Y.-Y., Kim, D.-K., Nedumaran, B., Jang, W. G., Cho, W.-J., Ha, J., Lee, I.-K., Lee, C.-H. & Choi, H.-S. Metformin inhibits hepatic gluconeogenesis through AMP-activated protein kinase-dependent regulation of the orphan nuclear receptor SHP. *Diabetes* **57,** 306–14 (2007).

40. Shackelford, D. B. & Shaw, R. J. The LKB1–AMPK pathway: metabolism and growth control in tumour suppression. *Nat. Rev. Cancer* **9,** 563–75 (2009).

41. El-Mir, M. Y., Nogueira, V., Fontaine, E., Avéret, N., Rigoulet, M. & Leverve, X. Dimethylbiguanide inhibits cell respiration via an indirect effect targeted on the respiratory chain complex I. *J. Biol. Chem.* **275,** 223–28 (2000).

42. Kawaguchi, T., Osatomi, K., Yamashita, H., Kabashima, T. & Uyeda, K. Mechanism for fatty acid 'sparing' effect on glucose-induced transcription:

regulation of carbohydrate-responsive element-binding protein by AMP-activated protein kinase. *J. Biol. Chem.* **277,** 3829–35 (2002).

43. Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., Musi, N., Hirshman, M. F., Goodyear, L. J. & Moller, D. E. Role of AMP-activated protein kinase in mechanism of metformin action. *J. Clin. Invest.* **108,** 1167–74 (2001).

44. Krausova, L., Stejskalova, L., Wang, H., Vrzal, R., Dvorak, Z., Mani, S. & Pavek, P. Metformin suppresses pregnane X receptor (PXR)-regulated transactivation of CYP3A4 gene. *Biochem. Pharmacol.* **82,** 1771–80 (2011).

45. Natali, A. & Ferrannini, E. Effects of metformin and thiazolidinediones on suppression of hepatic glucose production and stimulation of glucose uptake in type 2 diabetes: a systematic review. *Diabetologia* **49,** 434–41 (2006).

46. Papanas, N., Maltezos, E. & Mikhailidis, D. P. Metformin and cancer: licence to heal? *Expert Opin. Investig. Drugs* **19,** 913–17 (2010).

47. Giovannucci, E., Harlan, D. M., Archer, M. C., Bergenstal, R. M., Gapstur, S. M., Habel, L. A., Pollak, M., Regensteiner, J. G. & Yee, D. Diabetes and cancer: a consensus report. *Diabetes Care* **33,** 1674–85 (2010).

48. Zhuang, Y. & Miskimins, W. K. Cell cycle arrest in metformin treated breast cancer cells involves activation of AMPK, downregulation of cyclin D1, and requires p27Kip1 or p21Cip1. *J. Mol. Signal.* **3,** 14–18 (2008).

49. Ben Sahra, I., Laurent, K., Loubat, A., Giorgetti-Peraldi, S., Colosetti, P., Auberger, P., Tanti, J. F., Le Marchand-Brustel, Y. & Bost, F. The antidiabetic drug metformin exerts an antitumoral effect in vitro and in vivo through a decrease of cyclin D1 level. *Oncogene* **27,** 3576–86 (2008).

50. Vazquez-Martin, A., Oliveras-Ferraros, C. & Menendez, J. A. The antidiabetic drug metformin suppresses HER2 (erbB-2) oncoprotein overexpression via inhibition of the mTOR effector p70S6K1 in human breast carcinoma cells. *Cell Cycle* **8,** 88–96 (2009).

51. Keppler, D. Multidrug Resistance Proteins (MRPs,ABCCs): Importance for pathopysiology and Drug Therapy. *Handb. Exp. Pharmacol.* **201,** 299–323 (2011).

52. Song, I. S., Shin, H. J., Shim, E. J., Jung, I. S., Kim, W. Y., Shon, J. H. & Shin, J. G. Genetic variants of the organic cation transporter 2 influence the disposition of metformin. *Clin. Pharmacol. Ther.* **84,** 559–62 (2008).

53. Chen, Y., Li, S., Brown, C., Cheatham, S., Castro, R. A., Leabman, M. K., Urban, T. J., Chen, L., Yee, S. W., Choi, J. H., Huang, Y., Brett, C. M., Burchard, E. G. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 2 on the renal elimination of metformin. *Pharmacogenet. Genomics* **19,** 497–504 (2009).

54. Zhou, K., Donnelly, L. A., Kimber, C. H., Donnan, P. T., Doney, A. S. F., Leese, G., Hattersley, A. T., McCarthy, M. I., Morris, A. D., Palmer, C. N. A. & Pearson, E. R. Reduced-function SLC22A1 polymorphisms encoding organic cation transporter 1 and glycemic response to metformin: A GoDARTS study. *Diabetes* **58,** 1434–39 (2009).

55. Shikata, E., Yamamoto, R., Takane, H., Shigemasa, C., Ikeda, T., Otsubo, K. & Ieiri, I. Human organic cation transporter (OCT1 and OCT2) gene polymorphisms and therapeutic effects of metformin. *J. Hum. Genet.* **52,** 117–22 (2007).

56. Gambineri, A., Tomassoni, F., Gasparini, D. I., Di Rocco, A., Mantovani, V., Pagotto, U., Altieri, P., Sanna, S., Fulghesu, A. M. & Pasquali, R. Organic cation transporter 1 polymorphisms predict the metabolic response to metformin in women with the polycystic ovary syndrome. *J. Clin. Endocrinol. Metab.* **95,** E204–08 (2010).

57. Becker, M. L., Visser, L. E., van Schaik, R. H. N., Hofman, A., Uitterlinden, A. & Stricker, B. H. C. Genetic variation in the multidrug and toxin extrusion 1 transporter protein influences the glucose-lowering effect of metformin in patients with diabetes: a preliminary study. *Diabetes* **58,** 745–49 (2009).

58. Jablonski, K. A., McAteer, J. B., De Bakker, P. I. W., Franks, P. W., Pollin, T. I., Hanson, R. L., Saxena, R., Fowler, S., Shuldiner, A. R., Knowler, W. C., Altshuler, D. & Florez, J. C. Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* **59,** 2672–81 (2010).

59. Choi, J. H., Yee, S. W., Ramirez, H., Morrissey, K. M., Jang, G. H., Joski, P. J., Mefford, J., Hesselson, S. E., Schlessinger, A., Jenkins, G., Castro, R., Johns, S. J., Stryke, D., Sali, A., Ferrin, T. E., Witte, J. S., Kwok, P.-Y., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L. & Giacomini, K. M. A common 5'-UTR variant in MATE2-K is associated with poor response to metformin. *Clin. Pharmacol. Ther.* **90,** 674–84 (2011).

60. Dong, M., Gong, Z.-C., Dai, X.-P., Lei, G.-H., Lu, H.-B., Fan, L., Qu, J., Zhou, H.-H. & Liu, Z.-Q. Serine racemase rs391300 G/A polymorphism influences the therapeutic efficacy of metformin in Chinese patients with diabetes mellitus type 2. *Clin. Exp. Pharmacol. Physiol.* **38,** 824–29 (2011).

61. Becker, M. L., Visser, L. E., van Schaik, R. H. N., Hofman, A., Uitterlinden, A. G. & Stricker, B. H. C. Interaction between polymorphisms in the OCT1 and MATE1 transporter and metformin response. *Pharmacogenet. Genomics* **20,** 38–44 (2010).

62. Stocker, S. L., Morrissey, K. M., Yee, S. W., Castro, R., Xu, L., Dahlin, A., Ramirez, H., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L., Brett, C. M. & Giacomini, K. M. The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **93,** 186–94 (2013).

63. Zhou, K., Bellenguez, C., Spencer, C. C., Bennett, A. J., Coleman, R. L., Tavendale, R., Hawley, S., Donnelly, L., Schofield, C., Groves, C. J., Burch, L., Carr, F., Strange, A., Freeman, C., Blackwell, J. M., Bramon, E., Brown, M., Casas, J. P., Corvin, A., Craddock, N., Deloukas, P., Dronov, S., Duncanson, A., Edkins, S., Gray, E., Hunt, S., Jankowski, J., Langford, C., Markus, H. S., Mathew, C. G., Plomin, R., Rautanen, A., Sawcer, S. J., Samani, N. J., Trembath, R., Viswanathan, A. C., Wood, N. W., Harries, L. W., Hattersley, A. T., Doney, A. S. F., Colhoun, H., Morris, A. D., Sutherland, C., Hardie, D. G., Peltonen, L., McCarthy, M. I., Holman, R. R., Palmer, C. N., Donnelly, P. & Pearson, E. R. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat. Genet.* **43,** 117–20 (2011).

64. Yee, S. W., Chen, L. & Giacomini, K. M. The role of ATM in response to metformin treatment and activation of AMPK. *Nat. Genet.* **44,** 359–60 (2012).

65. Florez, J. C., Barrett-Connor, E., Jablonski, K., Knowler, W. C., Taylor, A., Shuldiner, A. R., Mather, K., Pollin, T. I., Horton, E. & White, N. H. The C allele of ATM rs11212617 does not associate with metformin response in the diabetes prevention program. *Diabetes Care* **35,** 1864–67 (2012).

66. Goldenberg, N. & Glueck, C. J. Is pharmacogenomics our future? Metformin, ovulation and polymorphism of the STK11 gene in polycystic ovary syndrome. *Pharmacogenomics* **9,** 1163–65 (2008).

67. Legro, R. S., Barnhart, H. X., Schlaff, W. D., Carr, B. R., Diamond, M. P., Carson, S. A., Steinkampf, M. P., Coutifaris, C., McGovern, P. G., Cataldo, N. A., Gosman, G. G., Nestler, J. E., Giudice, L. C., Ewens, K. G., Spielman, R. S., Leppert, P. C. & Myers, E. R. Ovulatory response to treatment of polycystic ovary syndrome is associated with a polymorphism in the STK11 gene. *J. Clin. Endocrinol. Metab.* **93,** 792–800 (2008).

# Chapter 2

# Genetic Variants in Transcription Factors Associate with the Pharmacokinetics and Pharmacodynamics of Metformin†

## INTRODUCTION

Metformin is a first-line therapy for type 2 diabetes, and is one of the most commonly prescribed drugs worldwide[1–9]. Despite 50 years of clinical use, its mechanism of action remains controversial. It has been well-established that metformin activates AMP-activated protein kinase, which may contribute to many of the pharmacological outcomes of metformin, including the inhibition of gluconeogenesis, reduction of glucose absorption, and enhancement of glucose uptake and utilization[2,6,10].

There is considerable variability in the glycemic response and pharmacokinetic characteristics of metformin. In terms of pharmacokinetics, metformin is not metabolized, and is excreted unchanged in the urine, with a half-life of roughly 5 hours[2,5,6,9]. The pharmacokinetic variability of metformin is unusually high for a renally cleared drug. In particular, mean plasma concentrations of metformin fluctuate between 0.4 and 1.3 mg/L at a dose of 1000 mg twice daily[1,2,5,6,11–16]. Metformin relies on facilitated transport for uptake into various tissues as well as for renal elimination. Specifically, transporters that mediate metformin elimination and tissue distribution include organic

---

† This chapter was published in *Clinical Pharmacology and Therapeutics* in June 2014. The following authors contributed to this work:
Srijib Goswami, Sook Wah Yee, Sophie Stocker, Jonathan D. Mosley, Michiaki Kubo, Rich Castro, Joel A. Mefford, Christopher Wen, Xiaomin Liang, John Witte, Claire Brett, Shiro Maeda, Monique M. Hedderson, Melissa D. Simpson, Robert L. Davis, Dan M. Roden, Kathleen M. Giacomini, Radojka M. Savic

cation transporters (OCTs) and multidrug and toxin extrusion proteins (MATEs), and may contribute to the wide variation in metformin pharmacokinetics. Pharmacokinetic variability contributes to variation in response to metformin as various research groups have observed dose-response relationships with fasting plasma glucose and HbA1c levels[17–19]. Metformin response variability is substantial, as >30% of patients receiving metformin are classified as poor responders[1,5,9].

To date, many pharmacogenetic studies have focused on the relationship between genetic variants in transporters and metformin pharmacokinetic parameters, and there has been one genome wide association study for metformin response[1,5,11–16,20–23]. For example, OCT1 is a major determinant of metformin uptake into hepatocytes, and genetic polymorphisms of OCT1 have been associated with reduced response and changes in metformin pharmacokinetics in healthy subjects and diabetic patients[1,7,24]. Recently, promoter variants of MATE1 and MATE2K, transporters that determine the efflux of metformin into the urine, were also shown to associate with metformin disposition and response in healthy subjects and diabetic patients[5,14,25]. Understanding genetic predictors of variability of both its response and disposition is important in the rational use of metformin for the treatment of patients with type 2 diabetes.

Though genetic studies have demonstrated associations with SNPs in transporters with metformin pharmacokinetics and pharmacodynamics, each individual SNP only accounts for a small fraction of the variation in HbA1c among

type 2 diabetic patients. This is not surprising given that metformin disposition is governed by multiple transporters rather than a single transporter. With this in mind, I proposed to study genetic variants in transcription factors that may regulate the expression levels of multiple metformin transporters and thus have larger effects on metformin disposition and response than variants in a single transporter. A subset of transcription factors have been shown to modulate the expression levels of OCTs (*SLC22*) and MATEs (*SLC47*), which are involved in determining metformin pharmacokinetics[26]. For example, transfection of *HNF4-alpha* has been shown to increase transcript levels of OCT1 in hepatocytes[3,7,8]. *SP1* has been implicated in modulating mRNA levels of MATE1[27–29]. *AP2* has been shown to have a repressive effect on MATE1 gene expression[3,28,30]. Other transcription factors have also been linked to modulating the expression levels of OCTs and MATEs involved in metformin disposition[7,24,27,29,31,32]. To date, the impact of transcription factor polymorphisms on metformin pharmacokinetics and response phenotypes has not been studied. Our hypothesis is that, compared to genetic variants in transporter genes, genetic variants in transcription factors will have a stronger impact on overall metformin plasma and tissue levels. This is because transcription factors modulate expression levels of a system of transporters leading to stronger effect sizes on pharmacological outcomes.

In this study, I first investigated the effect of genetic variants in a subset of genes on metformin response, specifically HbA1c levels in type 2 diabetic patients. Genes included were relevant metformin transcription factors, cited in the literature and demonstrated to play a modulatory role on key

metformin transporters. Subsequently, for the most significant transcription factor variants associated with HbA1c change, I further investigated their relationship with metformin pharmacokinetics using two approaches: 1) In a subset of healthy subjects with abundant pharmacokinetic measurements and available urine data, I used multiple linear regression to investigate the effect of the top transcription factor variants on measured metformin secretory clearance, which is a major route of metformin elimination. 2) Using data combined from type 2 diabetic patients and healthy subjects, I then developed a population pharmacokinetic model to investigate the effect of prioritized transcription factor variants identified to be significantly associated with secretory clearance, on various metformin pharmacokinetic parameters, attempting to explain the variability of parameter estimates in the context of well-studied transporter variants and ethnic variation[1,5,13].

This study suggests the importance of transcription factors and genetic variants in transcription factors on the pharmacological outcomes of metformin, namely HbA1c levels and metformin pharmacokinetic parameters. Variants in *SP1* exhibited the strongest association with both metformin pharmacokinetics and pharmacodynamics.

## METHODS

### Healthy human subjects

Data from four healthy volunteer studies from the University of California, San Francisco were pooled for this study, as previously described[5,34]. Healthy

male and female subjects were recruited directly from the Study of Pharmacogenetics in Ethnically Diverse Populations (IRB 10-03167) and participants were enrolled only after informed consent was provided. Volunteers of European American, African American, or Asian American ancestry in the study were at least 18 years of age and not taking any medications other than vitamins and/or oral contraceptives. Studies 6112, 6113, and 865 followed similar protocols. Healthy subjects were dosed with 1,000 mg of metformin, followed by an 850 mg dose of metformin on the second day of the study. Participants from study 767 were given a single 850 mg dose of metformin. The screening visit included a comprehensive medical history, physical examination, and laboratory studies. Volunteers with values two standard deviations from normal or a positive pregnancy test were excluded. During the short duration of the study in healthy volunteers, metformin levels in the liver may not have reached steady state.

**Patients with type 2 diabetes**

Diabetic patients of European American, African American, or Asian American ancestry were recruited into a multicenter retrospective study as described previously[5,14]. All patients were metformin naive, had HbA1c levels measured before and after initiation of metformin therapy (between 3 and 18 months), and had a medication possession ratio of >80%. The IRBs of Marshfield Clinic Research Foundation, Kaiser Permanente South East, Kaiser Georgia, and Vanderbilt approved this study and informed consent was obtained. In diabetic patients, metformin was administered for at least three months and

steady-state would have been achieved.

### Selection of transcription factor genes and variants

The candidate gene study consisted of genes and SNPs from transcription factors known to modulate levels of metformin related transporters (Figure 2.1).

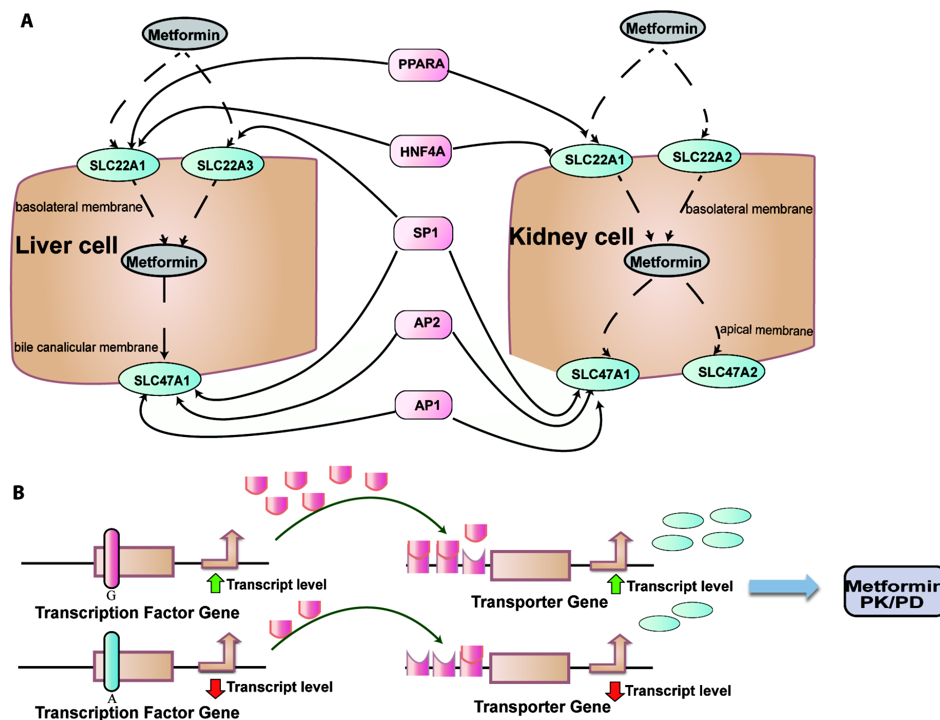



**Figure 2.1. Transcription factors regulating metformin PK/PD.**
A: A cell diagram that depicts a putative network of transcription factors working in concert to modulate gene expression levels of metformin transporters. B: A high-level gene diagram that highlights a mechanism by which a single nucleotide polymorphism change in a transcription factor gene may modulate the pharmacological outcome of metformin.

PharmGKB was the resource used to determine the list of transporters used in this study[2,33]. Transcription factors were selected based on evidence

from previous publications linking the transcription factor to one of the metformin related transporters[3,30,7,28,34]. Special importance was placed on transcription factors that had a regulatory link to MATE1 and OCT1, two transporters that play a crucial role in metformin pharmacokinetics and response. For example, our laboratory previously defined *AP1* and *AP2* as regulators of MATE1[3]. *SP1* was also selected as a potential regulator of MATE1. Functional involvement of *SP1* was confirmed by *SP1* overexpression, a mutational analysis of *SP1* binding sites, and an electrophoretic mobility shift assay. A separate study by our laboratory also suggested a role of *SP1* on OCT3[30]. *HNF4-alpha* and *PPAR-alpha* were also selected in the final list due to strong evidence linking the transcription factors to OCT1[7,27]. Because the goal of the study was to generate hypotheses about transcription factors that are involved in metformin response and pharmacokinetics, we took a less stringent approach to multiple comparisons testing. That is, we corrected for the number of genes tested (i.e. 5 genes) rather than the number of SNPs. Accordingly, the p-value was set at $p < 0.01$ for treatment HbA1c levels. A total of 5 transcription factors were selected, including *SP1, AP2, AP1, HNF4-alpha* and *PPAR-alpha*. After filtering out SNPs with low minor allele frequencies (MAF < 5%) in the type 2 diabetes cohort, we selected SNPs in transcription factors within 50,000 base pairs upstream and downstream of each transcription factor gene. Imputation was performed using data from the 1000 Genomes project.

**Phenotype selection for multivariate regression analysis**

For the pharmacodynamic analysis, a minimum treatment HbA1c value between 3 and 18 months post initiation of metformin, referred to as the treatment HbA1c, was selected as the phenotype of interest. The lower limit of this range takes into consideration metformin's delayed effect on HbA1c levels. Top associated SNPs with treatment HbA1c were selected and subsequently tested on metformin secretory clearance using multivariate regression. Pharmacokinetic parameters from a healthy subject study were determined previously using non-compartmental analysis (NCA)[5]. Secretory clearance was the primary parameter of interest since this parameter is assumed to be most sensitive to changes in transporter function and expression level and it was calculated using the following formula:

$$CL_{SR} = CL_R - CL_{Cr}$$

where $CL_{SR}$ is secretory clearance, $CL_R$ is renal clearance, and $CL_{Cr}$ is creatinine clearance.

**Genotyping, quality control and imputation**

Isolated DNA samples from healthy human volunteers and patients with T2D were genotyped on Illumina OmniExpress1.0 genotyping array at the RIKEN institute in Japan. Genotype data quality control was done using the standard protocol that has been previously described in other GWA studies. Imputation on relevant genes was performed using the IMPUTE2 software. SNPs associated with metformin pharmacokinetics and pharmacodynamics were filtered by excluding variants with low imputed scores (imputed score < 0.5).

### Linear regression analysis

Linear regression was performed using PLINK (v1.07), assuming an additive genetic model[35]. Imputation was performed using IMPUTE2 software (version 2)[36]. Variants with < 5% minor allele frequency (MAF) were excluded from the regression analysis. The statistical base model included clinically relevant covariates. For the regression analysis with treatment HbA1c, the statistical base model was adjusted for principal components, baseline HbA1c levels, average serum creatinine level, time to baseline, time to treatment HbA1c collection, other T2D drugs patients were administered, clinical site, age and gender. Transcription factor variants were then tested on the statistical base model. Pharmacokinetic data were used to provide mechanistic support for treatment HbA1c associated genetic variants. Top variants associated with treatment HbA1c with an adjusted P value <0.01 were filtered. These SNPs were then tested against metformin secretory clearance. In healthy subjects, the statistical model corrected for principal components and age.

### Population pharmacokinetic modeling of metformin and final model selection

Data from 5 studies (patient study, and healthy volunteer studies 6112, 6113, 865, and 767), which includes healthy volunteers and type 2 diabetic patients, were analyzed using non-linear mixed effect modeling (NONMEM 7) with first order conditional estimation method with interaction (FOCE-I). Model selection was informed by using the objective function value (OFV, -2log

likelihood) and visual inspection of diagnostic plots. Different base models with varying transit compartments to capture the absorption of metformin were evaluated. The final structural model was parameterized in terms of apparent clearance (CL/F), apparent central and peripheral volumes of distribution (Vc/F and Vp/F), apparent inter-compartmental clearance (Q/F), mean transit time (MTT), and a first order absorption ($k_a$). Inter-individual variability was estimated for CL/F, Vc/F and Q/F. The structural model was first built using healthy volunteer data due to the availability of abundant plasma samples. Once the structural model was established, patient data were then added and parameters were re-estimated to provide the base model for covariate inspection. The Stepwise Covariate Model (SCM) tool in Perl Speaks NONMEM (PsN) was used to develop the final model with statistically significant covariates on metformin pharmacokinetic parameters. Along with the top transcription factor variants, a comprehensive list of the transporter variants that were tested in SCM is shown in Table 2.1.

**Table 2.1** List of transporter genetic variants investigated in the population pharmacokinetic model of metformin

| Variant | Gene | Relevance/Association in literature | # Samples in Study |
|---|---|---|---|
| rs35191146 | *SLC22A1* | ↑AUC ↑ Cmax ↑CLr ↓ Vd ↓OGTT[1,15] | N=21<br>N=103 |
| rs34130495 | *SLC22A1* | ↑AUC ↑ Cmax ↑CLr ↓ Vd ↓OGTT[1,15] | N=21<br>N=103 |
| rs34059508 | *SLC22A1* | ↑AUC ↑ Cmax ↑CLr ↓ Vd ↓OGTT[1,15] | N=21<br>N=103 |
| rs12208357 | *SLC22A1* | ↑AUC ↑ Cmax ↑CLr ↓ Vd ↓OGTT[1,15] | N=21<br>N=103 |
| rs1867351 | *SLC22A1* | ↑Renal Clearance (CLr)[1] | N=103 |
| rs622342 | *SLC22A1* | Response: ↑ HbA1c[20] | N=99 |
| rs2289669 | *SLC47A1* | Response: ↓HbA1c[12,37] | N=116 |
| rs8065082 | *SLC47A1* | ↓Diabetes incidence[12,37] | N=116 |
| rs316019 | *SLC22A2* | ↑AUC ↑Cmax ↑CLr [13,21,22] | N=26<br>N=15<br>N=23 |
| rs12943590 | *SLC47A2* | Response: ↓HbA1c ↑ CLr[14] | N=57 |
| rs2252281 | *SLC47A1* | Altered glucose tolerance[5] | N=57 |
| rs555754 | *SLC22A3* | ↑ Luciferase activity for Minor allele<br>↑mRNA of OCT3[30] | - |
| rs683369 | *SLC22A1* | ↓Response to imatinib mesylate [23] | - |

AUC = Exposure of metformin for the particular study cited. Cmax = Maximum plasma concentration of metformin. CLr = Metformin renal clearance. Vd = Metformin volume of distribution. OGTT = Oral glucose tolerance test.

Finally, a bootstrap was performed with 1000 samples to obtain 95% confidence intervals of all pharmacokinetic parameters used to characterize the final model.

### Gene expression correlations

Gene expression data were collected from control kidney samples from The Cancer Genome Atlas database. From this publically available online portal,

a total of 65 kidney tissue samples were available for analysis, which was retrieved on June of 2012. Gene expression correlations were calculated using Pearson correlation analysis.

### Transcription factor binding analysis

The likelihood of *SP1*, *PPAR-alpha, HNF4-alpha,* and *AP2* binding to DNA sequences 50 kilobases upstream and downstream of the 6 transporter genes was determined using available online tool FIMO[38]. Transcription factor *AP1* was not available in dataset. Probability matrices were generated for each transcription factor using other online sources including JASPAR and NUBISCAN[38,39]. Once the probability matrix was established, FIMO was then used to scan the likelihood of a given transcription factor binding to DNA regions proximal to the transporter genes.

## RESULTS

### Characteristics of type 2 diabetic patients and healthy subjects

Baseline characteristics of patients and healthy subjects are summarized in Table 2.2. Clinical data included longitudinal HbA1c measurements from 440 Type 2 diabetic patients. A total of 2382 metformin plasma samples in healthy subjects (102) and patients (133) were used to develop a population pharmacokinetic model. Of the 102 healthy subjects, 57 subjects also had available urine samples, which allowed for the collection of creatinine levels and the subsequent calculation of metformin secretory clearance.

**Table 2.2** Baseline characteristics of patients with type 2 diabetes and healthy volunteers dosed with metformin

| | Type 2 Diabetic Patients | | | Healthy Volunteers | |
|---|---|---|---|---|---|
| Characteristic | Marshfield Clinic | Kaiser South East | Vanderbilt | Study No. 6112/6113 | Study No. 865/767 |
| N | 149 | 133 | 162 | 57 | 45 |
| Available PK data, N | 0 | 133 | 0 | 57 | 45 |
| Male, N | 65 | 44 | 78 | 21 | 23 |
| Female, N | 84 | 89 | 84 | 36 | 22 |
| European American (%) | 100 | 18 | NA | 32 | 84 |
| African American (%) | - | 76 | NA | 58 | 11 |
| Asian American (%) | - | 4 | NA | 10 | 5 |
| Other (%) | - | 2 | NA | - | - |
| Quantitative traits | | | | | |
| Age (years) | 57 (23-90) | 58 (33-79) | 59 (33-81) | 25 (18-45) | 31 (18-44) |
| Average body weight (kg) | 98 (51-212) | 98 (58-182) | 93 (34-184) | 73 (49-136) | 70 (44-112) |
| Baseline HbA1c (%) | 7.5 (5.5-12.8) | 7.7 (5.2-11.9) | 7.6 (5.8-14.9) | - | - |
| Average metformin daily dose (mg) | 990 (500-2000) | 1000 (250-2500) | 1000 (250-2000) | NA | NA |
| Metformin dose (mg) (Healthy Volunteers) | NA | NA | NA | 1850 | 1850/850 |

Quantitative data shown reflect the median (range). From the Kaiser South East cohort, 4 patients did not have reported baseline HbA1c levels.
Study 6112/6113/865: Healthy volunteers in these studies were administered 1000 mg of metformin, followed by 850 mg of metformin after a 12 hour interval. Healthy subjects in studies 6112/6113 also had urine data available to calculate metformin secretory clearance. Please see references 5,8,9.
Study 767: Healthy volunteers were given a single dose of metformin (850 mg). Please see reference 9 in paper.
Vanderbilt ethnicity is not reported (NA), and principal components were used for the genetic analysis across all studies.

All healthy volunteers and 133 patients with sparse pharmacokinetic information were used to build the population pharmacokinetic model.

**Top transcription factor variants from multivariate regression approach**

A total of five transcription factors were selected (*AP1, AP2, SP1, HNF4-alpha* and *PPAR-alpha*) for multivariate linear regression with treatment HbA1c levels. Among the tested genetic variants that met our criteria (see Methods section), 40 SNPs were associated with HbA1c change 3 months after metformin initiation, adjusted for baseline levels (Treatment HbA1c). Among the 40 SNPs, a multivariate linear regression on metformin secretory clearance was performed as one method for investigating a pharmacokinetic mechanism. A total of six SNPs in two genes were significantly associated with metformin secretory clearance and treatment HbA1c levels using a multivariate regression model. Of these six genetic variants, five were located in the *SP1* region and one was located in the *AP2* region (Table 2.3)

**Table 2.3** Prioritized genetic variants in *SP1* and *AP2* associated with metformin pharmacodynamics and pharmacokinetics

| SNP | Gene | FRQ | BETA | P | Phenotype |
|---|---|---|---|---|---|
| chr13:74559166:D | *AP2* | 0.06 | 0.45 | 0.003 | A1C |
| chr13:74559166:D | *AP2* | 0.06 | 185.93 | 0.004 | CLSR |
| rs784892 | *SP1* | 0.91 | -0.32 | 0.008 | A1C |
| rs784892 | *SP1* | 0.82 | -76.94 | 0.01 | CLSR |
| rs2694855 | *SP1* | 0.92 | -0.40 | 0.008 | A1C |
| rs2694855 | *SP1* | 0.86 | -98.63 | 0.01 | CLSR |
| rs2683511 | *SP1* | 0.90 | -0.32 | 0.01 | A1C |
| rs2683511 | *SP1* | 0.83 | -86.75 | 0.01 | CLSR |
| rs10747673 | *SP1* | 0.12 | 0.38 | 0.01 | A1C |
| rs10747673 | *SP1* | 0.18 | 105.98 | 0.02 | CLSR |
| rs784888 | *SP1* | 0.92 | -0.36 | 0.01 | A1C |
| rs784888 | *SP1* | 0.86 | -106.23 | 0.04 | CLSR |

FRQ = Frequency of the associated allele in the dataset. CLSR= Metformin secretory clearance. A1C = treatment HbA1c. BETA = Slope of association

**Population Pharmacokinetic Model**

In addition to metformin secretory clearance, a population based modeling approach was used to investigate the effect of prioritized transcription factor variants on the systemic plasma levels of metformin in both patients and healthy subjects. The six-transcription factor SNPs (associated with metformin PK and PD) identified from the multiple linear regression analysis along with 13 genetic variants in metformin transporters and demographic variables were investigated.

A 2-compartment model with a delayed absorption (1-transit compartment) best described the data. A schematic of the final model with included covariates is shown in Figure 2.2.

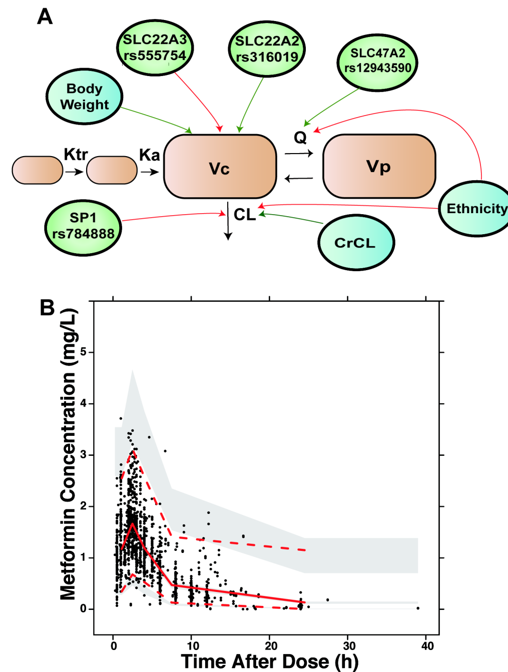



**Figure 2.2 Final pharmacokinetic model and visual predictive check.**
A: A 2-compartment model with delayed absorption best characterized the data. All parameters reflect ratios over the bioavailability of metformin (F). The model structure includes the final model covariates determined by a stepwise covariate analysis (SCM). CrCL= Creatinine clearance. Vc = Apparent central volume of distribution. Vp =Apparent peripheral volume of distribution. CL= Apparent clearance of metformin. Ktr = transit rate constant. Ka = absorption rate constant. B: Visual predicted check of the final population pharmacokinetic model. The shaded regions indicate the 95$^{th}$ and 5$^{th}$ percentiles (ends) and the range of median simulated profiles (center) of simulated predictions from the visual predictive check. Overlaid back points are combined healthy volunteer (n=102) and type 2 diabetic patient (n=133) data observations. The solid line indicates the median of the observed data, with red lines indicating the 95$^{th}$ and 5$^{th}$ percentile of observations.

Pharmacokinetic profiles were overall quite similar between healthy subjects and patients, as has been previously observed[40]. The final model consisted of statistically significant covariates on central compartment volume ($V_c/F$), apparent clearance (CL/F), and metformin peripheral flow (Q/F). Covariates that affected $V_c/F$ were body weight, OCT3 variant rs555754 and OCT2 variant rs316019.

Based on different combinations of these covariates, $V_c/F$ may decrease by as much as 54% and increase by as much as 260% (Figure 2.3).

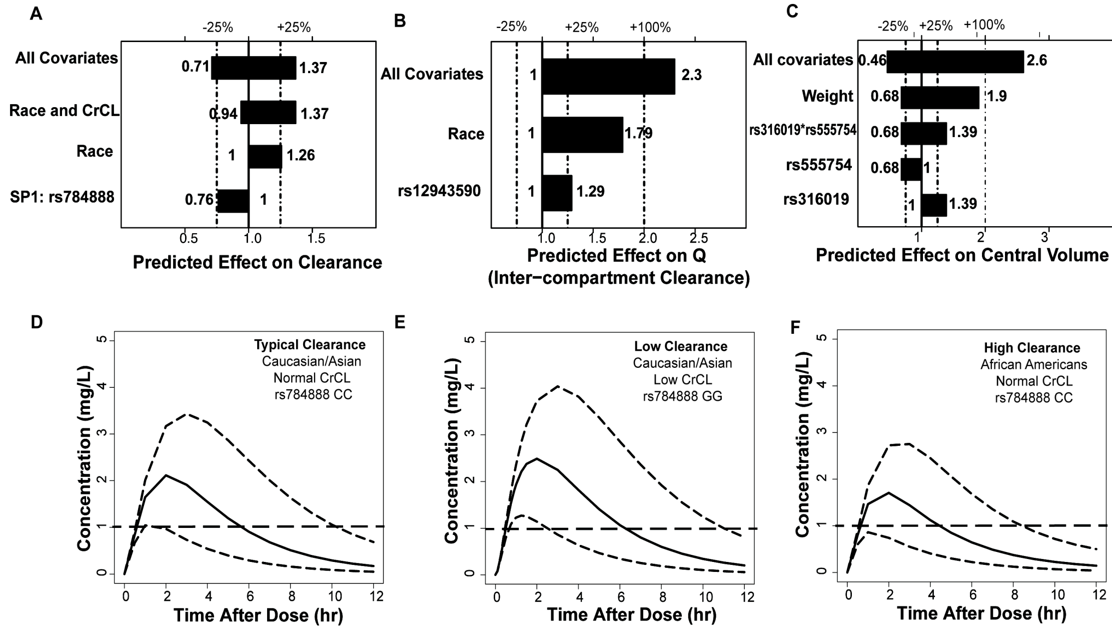



**Figure 2.3 Summary of covariate effects on population pharmacokinetic parameters.**

CrCL=creatinine clearance. A: Predicted effects of covariates on apparent clearance (CL/F) with the reference ethnicity identified as European Americans. B: Predicted effect of covariates on inter-compartment clearance (Q/F). C: Predicted effects of covariates on central compartment volume ($V_c/F$). D, E, and F: Simulations of pharmacokinetic profiles following a single 850 mg dose of metformin with variability estimates on clearance. Simulations were based on the predicted metformin clearance estimates. Predicted clearance estimates are based on the covariates described in figure 2.3A. Ethnicity, creatinine clearance (CrCL), and *SP1* variant status are shown in each figure. Dashed line at 1 mg/L indicates the lower target of metformin concentration based on therapeutic range. D: Black solid line=typical value of clearance for a Caucasian with normal creatinine clearance (80-130 mL/min), and homozygous (CC) SNP rs784888. Dashed black lines indicate the 97.5% and 2.5% of the Inter-individual variability (ETA) distribution for CL/F for this patient. E: Black solid line=typical value of clearance for a Caucasian with low creatinine clearance (<70 mL/min), and homozygous (GG) for SNP rs784888. Dashed lines indicate the 97.5% and 2.5% of the Inter-individual variability (ETA) distribution for CL/F for this patient. F: Black solid line=typical value of clearance for an African American with normal creatinine clearance (80-130 mL/min), and homozygous (CC) for SNP rs784888. Dashed lines indicate the 97.5% and 2.5% of the Inter-individual variability (ETA) distribution for CL/F for this patient.

For peripheral flow (Q/F), ethnicity and the MATE2 variant rs12943590 were significant. Based on different combinations of ethnicity and the rs12943590 variant, Q/F may increase by as much as 230%. In Figure 2.3, the reference ethnic population was European American to more clearly show the range of covariate effects on Q/F and CL/F. Finally for CL/F, creatinine clearance (CrCL), ethnicity, and the *SP1* variant rs784888 were statistically significant. Different combination of these covariates may decrease CL/F by as much as 29% and increase it by as much as 37%. Final model parameter estimates are summarized in table 2.4.

**Table 2.4** Population pharmacokinetic model derived estimates and bootstrap results for pharmacokinetic parameters of metformin

| Final Model Parameter | Median (%RSE)[1] | Median (90% CI)[2] |
|---|---|---|
| Total clearance CL/F (L/h) | 78.4 (6) | 77.5 (70-87) |
| Central volume of distribution Vc/F (L) | 76.8 (15) | 76.7 (53-95) |
| Peripheral flow Q/F (L/h) | 18.1 (9) | 18.8 (16-32) |
| Peripheral volume of distribution Vp/F (L) | 413 (43) | 419 (286-3293) |
| Mean transit time (h) | 0.207 (24) | 0.203 (0.14-0.37) |
| Absorption rate $K_a$ (1/h) | 0.312 (5) | 0.31 (0.28-0.34) |
| **Between-subject variability (% variance)** | | |
| Between-subject variability (CL) | 50 (10) | 50 (40-60) |
| Between-subject variability (Vc) | 45 (13) | 45 (34-64) |
| Between-subject variability (Q) | 41 (23) | 40 (27-49) |
| **Covariance of parameters** | | |
| Correlation CL-V | 0.05 (5) | 0.10 (-0.33-0.29) |
| **Residual error model** | | |
| **Studies 6112/6113** | | |
| Proportional error (%) | 14 (12) | 0.14 (0.11-0.16) |
| Additive error | 0.02 (32) | 0.02 (0.01-0.02) |
| **Study 865** | | |
| Proportional error (%) | 12 (6) | 12 (11-13) |
| Additive error | 0.01 | - |
| **Study 787** | | |
| Proportional error (%) | 21 (13) | 20 (16-25) |
| Additive error | 0.01 (20) | 0.007 (0.003-0.012) |
| **Patient data** | | |
| Proportional error (%) | 20 (13) | 0.2 (0.16-0.24) |
| Additive error | 0.01 | - |

[1]Typical value of pharmacokinetic parameter in final model. RSE= Relative standard error (%), also known as the precision of the population pharmacokinetic parameter estimate.
[2]Confidence interval for the population pharmacokinetic parameter following bootstrap results. Reference ethnicity of model output is African Americans, due to high proportion of African Americans in the cohort.

The final model explained 5% of the variability in CL/F, 23% of the variability in $V_c$/F, and 13% in Q/F. The *SP1* variant rs784888 and ethnicity reduced the variance of inter-individual variability (IIV) in CL/F by 1.6% and 2.3% respectively. The covariates body weight, OCT3 variant rs555765, and OCT2 variant rs316019 reduced the variance of IIV in $V_c$/F by 41.3%, 7.8%, and 7.4% respectively. Finally, for Q/F, covariates MATE2 variant rs12943590 and ethnicity reduced the variance of IIV in Q/F by 20% and 41% respectively. Based on model estimates, an extreme case of high metformin clearance (low exposure) is brought on by an African American (AA) patient with normal renal function and homozygous reference for the *SP1* variant rs784888. The performance of the final model, as determined by basic goodness of fit plots, parameter precision, bootstrapping, and visual predictive checks was adequate (Table 2.4 and Figure 2.2). The final equations for CL/F, $V_c$/F, and Q/F were (Eqs 1,2,3):

$$(1)\ \frac{CL}{F}_{TV} = 78.4\times(1+\ \theta_{CRCL}(CRCL\ -112))\times(1+\ \theta_{Ethnicity,CL})\times(1+\ \theta_{rs784888}\times(SP1-0))$$

$$(2)\ \frac{Vc}{F}_{TV} = 76.8\times(1+\ \theta_{rs555754}(OCT3_{rs555754}-0))\times\left(1+\ \theta_{WT}\times(WT-75)\right)\times(1+\theta_{rs316019})$$

$$(3)\ \frac{Q}{F}_{TV} = 18.1\times(1+\ \theta_{rs12943590})\times(1+\ \theta_{Ethnicity,Q})$$

where 78.4 L/h is the typical metformin apparent clearance for an African American with a creatinine clearance of 112 mL/min, and homozygous reference (CC genotype) for the *SP1* variant, rs784888. The imputed *SP1* variant can take on a value between 0 and 2, which quantifies the presence of a minor allele. Similarly, 76.8 L is the typical central compartment volume for an African American weighing 75 kg (WT), and is homozygous reference for both rs555754

(*SLC22A3*) and rs316019 (*SLC22A2*). $\Theta_{CRCL,}$ $\Theta_{Ethnicity,CL,}$ $\Theta_{rs784888,}$ $\Theta_{rs555754,}$ $\Theta_{WT,}$ $\Theta_{rs316019,}$ $\Theta_{rs12943590,}$ and $\Theta_{Ethnicity,Q,}$ are the corresponding effect sizes for creatinine clearance, ethnicity on CL/F, *SP1* variant rs784888, OCT3 variant rs555754, body weight, OCT2 variant rs316019, MATE2 variant rs12943590 and ethnicity on Q/F, respectively. The number of subjects in the dataset with an imputed value of greater than or equal to 1 (for the minor allele) for variants rs784888, rs555754, rs316019 and rs12943590 were 54, 131, 106, and 60, respectively. African Americans were used as the reference ethnic population in the model due to a large representation of African Americans in our cohort. In Figure 2.3, the reference ethnic population was changed to European American (EA) to more clearly show the range of covariate effects on each pharmacokinetic parameter.

**Genetic variants in *SP1* were critical determinants of variation in pharmacokinetics and pharmacodynamics of metformin**

Our top finding was in the transcription factor, *SP1*, a gene previously noted to potentially modulate transcript levels of OCT3 and MATE1[28,30]. Our regression results showed multiple independent SNPs in the *SP1* locus associated with pharmacodynamic and pharmacokinetic phenotypes of metformin. Five of the six total variants linked to metformin pharmacokinetics and pharmacodynamics were in the *SP1* gene region. Most of the SNPs were in noncoding regions including intronic, upstream and downstream regions of SP1. The most strongly associated SNP, rs784892, is in the intronic region of the

downstream gene, *AMHR2* (Anti-Mullerian Hormone Receptor, Type II). This SNP strongly associated with both metformin pharmacodynamics (HbA1c) and pharmacokinetics (metformin secretory clearance), with beta coefficients of -0.32 HbA1c per G allele (P = 0.008) and -76.9 mL/min per G allele (P = 0.02) respectively (Figure 2.4).

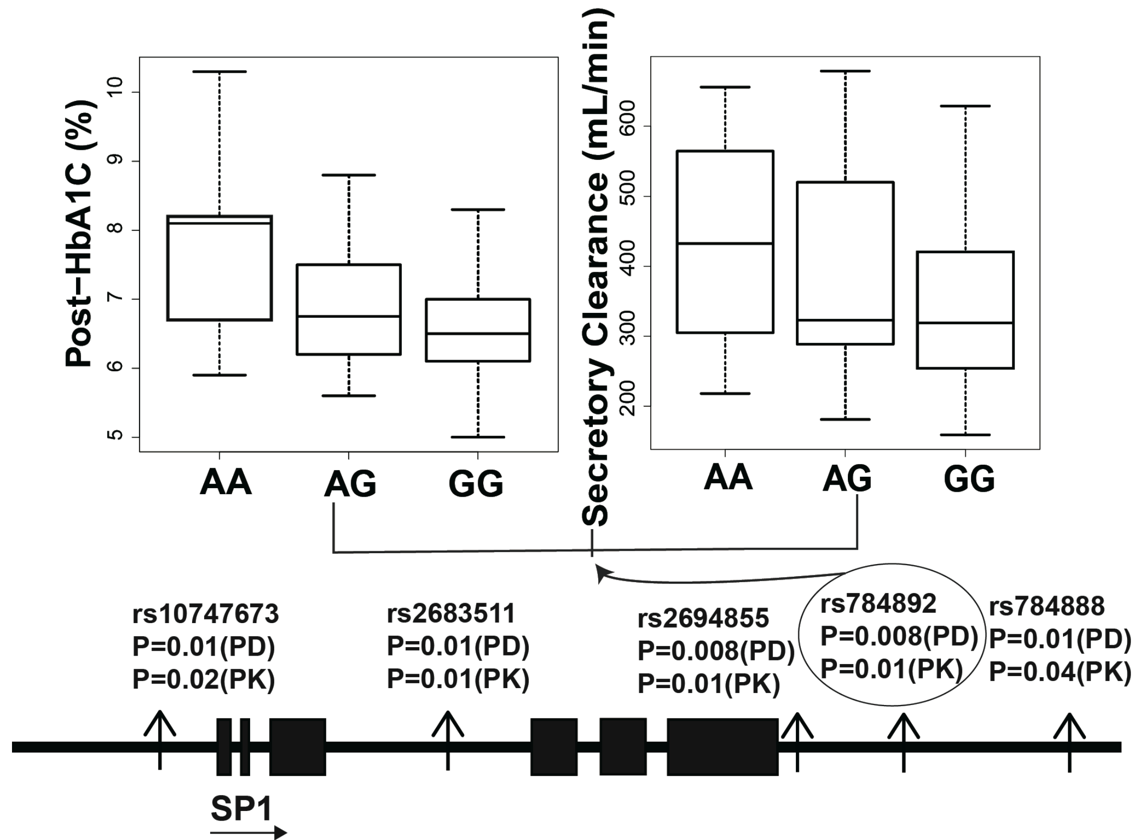



**Figure 2.4. Influence of top *SP1* variants on metformin pharmacodynamics and pharmacokinetics using multivariate regression.**
PD = Pharmacodynamics, variant association with treatment HbA1c level, as defined in the text. PK= Pharmacokinetics, variant association with measured metformin secretory clearance (mL/min). G=Associated allele of rs784892. Gene diagram summarizes the chromosome location of top associated SNPs in relation to exonic regions. N=440 patients for pharmacodynamics analysis and N=57 healthy subjects for the pharmacokinetic analysis.

To further investigate the significance of *SP1* variants on secretory clearance, a separate analysis was performed on creatinine clearance. From this analysis, no statistical significance or trend was observed between *SP1* variants and creatinine clearance. The rs784892 variant has a combined minor allele frequency of approximately 11% across all ethnic groups, with African Americans (~35%) having a higher frequency than European Americans (<1%). However, the effect of race was accounted for using principal components in our multivariate analysis, and a separate analysis in African Americans was performed to ensure that the variant has a significant effect on both Treatment HbA1c levels and secretory clearance in African Americans. Statistical significance was observed for both phenotypes in the African American cohort.

Results from the final population pharmacokinetic model of metformin determined that rs784888, a SNP less than 50 kilobases downstream of *SP1*, was an important predictor of metformin apparent clearance. The rs784888 G allele led to a 12% reduction in metformin apparent clearance. This variant has a frequency of approximately 9% (AA=42%, EA=<1%). Similar to rs784892, a separate analysis confirmed this finding in African Americans.

**_PPAR-alpha_, and _HNF4-alpha_ were major genes with polymorphisms associated with variation in pharmacodynamics, independent of pharmacokinetics**

A total of 17 variants in *PPAR-alpha* and 6 variants in *HNF4-alpha* associated with metformin pharmacodynamics (P<0.01) (Figure 2.5).

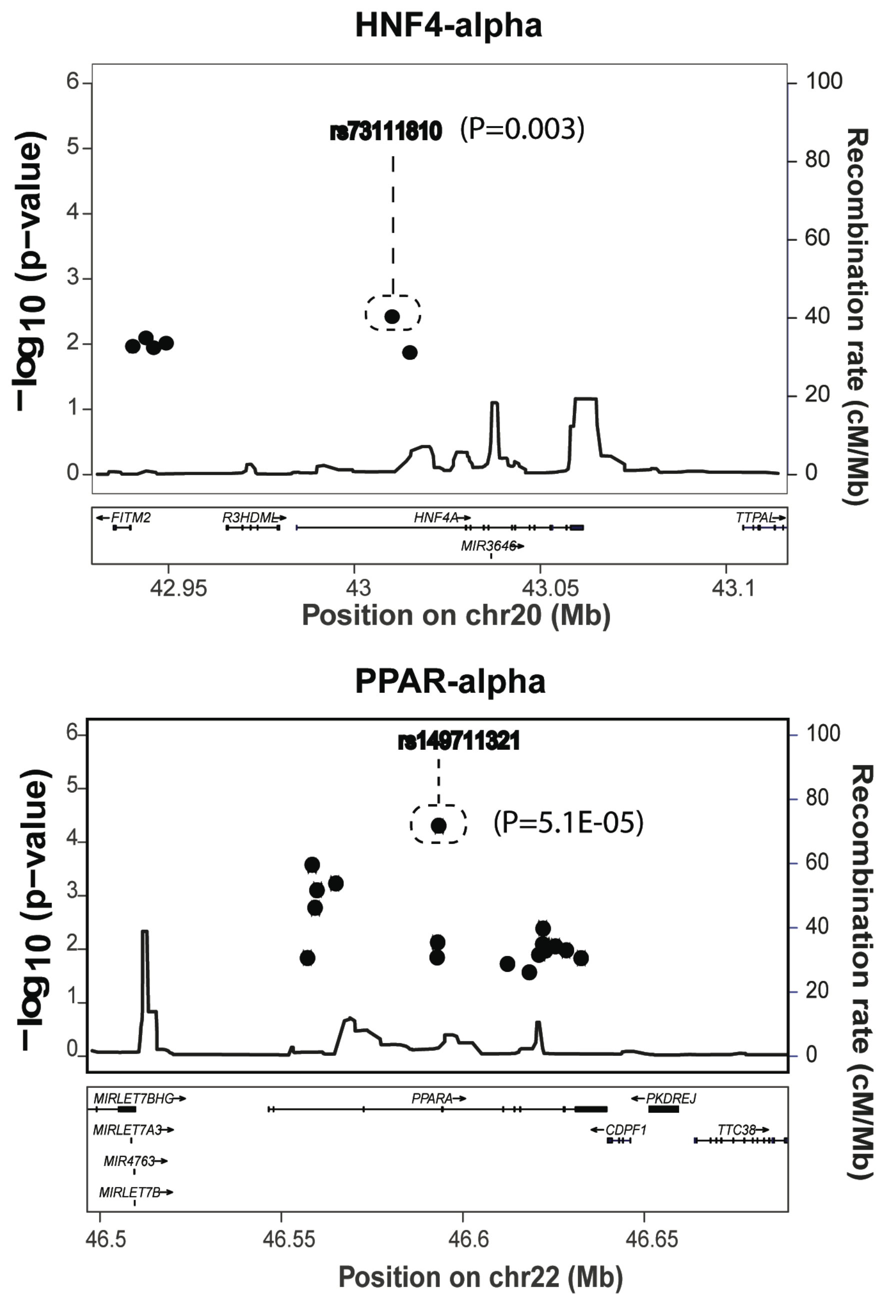



**Figure 2.5. Manhattan plots of top SNPs for *HNF4-alpha* and *PPAR-alpha*.** A zoomed-in view of genetic polymorphisms in HNF4-alpha and PPAR-alpha associated with treatment HbA1c levels. Circles represent the location and the -$\log_{10}$ P value of the association. Recombination rates are also overlaid on the figure, with each peak representing relatively high recombination rates for that region.

The most significant variant from our pharmacodynamic analysis was intronic SNP rs149711321 in *PPAR-alpha* (P=1E-05). This variant has a minor allele frequency of 6% reported for both European American and African American ethnicities.  Out of the 23 total variants in *PPAR-alpha* and *HNF4-alpha*, none were significantly associated with metformin pharmacokinetics in healthy subjects.

**Deletion in *AP2*, a repressor, was associated with changes in metformin pharmacokinetics and pharmacodynamics**

A total of 11 variants in *AP2* were linked to changes in HbA1c levels in type 2 diabetic patients receiving metformin. Of these, 1 imputed deletion variant in the intronic region of *AP2*, associated with an increase in metformin secretory clearance (P=0.004) and an increase in treatment HbA1c levels (P=0.003). The minor allele frequency of this variant was 6%, with a slope of 0.45 percentage change in HbA1c per minor allele. This variant was not associated with the pharmacokinetics of metformin in our population pharmacokinetic model. However, it is important to note that the model did not specifically include metformin secretory clearance as a pharmacokinetic parameter due to the absence of urine data.

**Significant impact of other covariates in the population pharmacokinetic model on metformin pharmacokinetic parameters**

Transporter variants previously associated with metformin pharmacokinetics from non-compartmental approaches were also significantly associated with metformin kinetics (CL/F, $V_c$/F Q/F) in the population pharmacokinetic model. In particular, a MATE2/*SLC47A2* variant (rs12943590) was significantly associated with peripheral flow (Figure 2.2), potentially increasing the uptake of metformin into peripheral tissues by up to 30%. A gene/gene interaction effect of an OCT3 variant (rs555754) and an OCT2 variant (rs316019) on metformin central volume was also observed. This interaction reflected a range of effect sizes on $V_c$/F. Depending on the combination of minor alleles from rs555754 and rs316019, metformin $V_c$/F may decrease by as much as 32% and increase by as much as 39% (Figure 2.3). Although significant in the final model, the downstream clinical significance of these variants will require replication by other studies before the findings can be clinically translated.

The effect of ethnicity was investigated in 123 African Americans, 69 European Americans, and 22 Asian American type 2 diabetic patients using our population pharmacokinetic model. Ethnicity was found to be a significant predictor of metformin flow, specifically, the pharmacokinetic parameters CL/F and Q/F. For both parameters, African Americans had significantly higher values compared to European Americans (Figure 2.3). Compared to African Americans, European Americans had approximately a 26% lower metformin clearance and Asian Americans had a 22% lower clearance. For Q/F, European Americans were predicted to have a 46% lower peripheral flow compared to African Americans. These effects were independent of creatinine clearance and body

weight. Simulations in Figure 2.3 compare the effect of typical and extreme values of metformin clearance following an 850 mg dose (based on different covariate combinations). The impact is shown on metformin plasma levels, with a line drawn at 1 mg/L to indicate a lower concentration target for metformin.

## DISCUSSION

Previous pharmacogenetic studies of metformin pharmacokinetics have focused on a few non-synonymous variants in transporter genes[1,5,13,14]. Pharmacogenetic investigation of variants in gene expression modulators of key transporters involved in metformin pharmacokinetics is a novel approach in understanding the variability in response to metformin. This study tested the effect of genetic variants in key transcription factor genes on metformin pharmacodynamics with a focus on glycemic response to metformin in type 2 diabetic patients. Subsequently, the top hits associated with metformin response were examined for a pharmacokinetic mechanism in both type 2 diabetic patients and healthy subjects using two approaches. Four important findings emerged from our combined PK/PD analysis: 1) SNPs in *SP1* associated with metformin pharmacodynamics in type 2 diabetic patients and this association had a pharmacokinetic basis. 2) SNPs in *AP2*, in particular a deletion variant, associated with the pharmacokinetics and pharmacodynamics of metformin. 3) SNPs in *PPAR-alpha* and *HNF4-alpha* associated with the pharmacodynamics of metformin, but did not have a pharmacokinetic mechanism. 4) Finally, African

Americans were observed to have greater apparent clearances compared with European Americans and Asian Americans.

To date, the role of *SP1* on metformin pharmacology has not been investigated despite previous studies indicating that *SP1* modulates the gene expression of MATE1 and OCT3, two important metformin transporters.[28,30] Our findings strongly suggest an important role of *SP1* in governing metformin disposition and response. Using standard regression, we observed that 5 genetic variants in *SP1* associated with metformin PK and PD. Multiple independent effects were observed even after accounting for linkage disequilibrium. A population-based approach demonstrated that one such SNP located downstream of *SP1*, rs784888, significantly affected metformin apparent clearance, therefore impacting systemic plasma levels of metformin. Comparing the effect size of this *SP1* variant on CL/F with the effect sizes of transporter variants observed in previous studies revealed that the effect size for the *SP1* variant (12%) was greater than those previously reported for variants in transporters and CL/F[1,5,11]. Interestingly, rs784888 is in strong linkage disequilibrium with rs147778161 ($r^2$>0.8), an intronic SNP of *SP1*, which was initially removed due to not passing our initial pharmacodynamic cutoff (P<0.01). The observed associations for *SP1* are biologically plausible, considering the expression of MATE1 on the apical membrane in the proximal tubule of the kidney and its role in metformin renal secretion. The role of OCT3 in the renal elimination of metformin is still not fully elucidated, therefore, it is not known

whether SNPs in *SP1* may modulate metformin pharmacokinetics by regulating the expression of OCT3 in addition to MATE1 and other metformin transporters. Furthermore, we performed a separate transcription factor binding analysis using a transcription factor binding tool (FIMO), and found that in addition to regulating the expression of MATE1 and OCT3, *SP1* may also modulate levels of OCT2 and MATE2-K, transporters that are known to play a very important role in metformin elimination from the kidney. Though speculative, we propose that genetic variants in *SP1* may affect the binding affinity or the expression level of the transcription factor, which could then have a combined effect on MATE1, OCT2, OCT3, and MATE2-K levels, globally affecting the pharmacokinetic and pharmacodynamic outcomes of metformin. Furthermore, the regression results suggest a clinically significant impact of *SP1* variants on metformin pharmacokinetics and pharmacodynamics. For example, typical patients homozygous for the minor allele in rs784892, achieved on average treatment HbA1c levels 1.1% lower than patients homozygous for the reference allele. This finding, if replicated, would have enormous clinical significance, given that metformin reduces HbA1c levels by 1.12% on average (i.e. from 8.0% A1c to 6.9% A1c) within the first year of therapy[27,41,42]. Furthermore, in healthy subjects, there was a 98 mL/min reduction in metformin secretory clearance on average in homozygous carriers of the variant allele, compared to homozygous carriers of the reference allele. This pharmacokinetic mechanism supports our pharmacodynamic finding, where a lower metformin secretory clearance is expected to increase metformin exposure and hence reduce HbA1c levels to a

greater extent than the reference allele. In our population-based approach, rs784892 trended in the same direction as observed in our regression analysis for metformin secretory clearance, but was only borderline significant from stepwise covariate modeling building, and was not included in the final model based on our strict criteria. However, *SP1* variant rs784888 was included. This variant had a significant impact on the apparent clearance of metformin, and was predicted to lower metformin clearance by up to 24% in homozygous carriers of the minor allele. A lower clearance of metformin was predicted to increase metformin exposure, potentially leading to a more favorable response to metformin.

Interestingly, one deletion variant in *AP2* associated with an increase in metformin secretory clearance and a reduction in the glycemic response to metformin. This finding was similar to those observed for polymorphisms in *SP1*, and consistent with previous studies in our laboratory showing that *AP2* is an important repressor of *SLC47A1* (MATE1) transcription[3,5]. This finding is biologically plausible, with high expression of *AP2* and MATE1 found in the kidney.[43] Furthermore, we performed additional gene expression analysis using data from The Cancer Genome Atlas (TCGA) to investigate correlations between high priority transcription factors and transporter genes. Gene expression data, along with our supplemental motif binding analysis suggested a regulatory role of *AP2* on *SLC47A1* and *SLC47A2*. In addition to *SP1*, *AP2* may also play an important role in modulating metformin pharmacokinetics and

pharmacodynamics via its role as a global regulator of multiple transporter genes.

*PPAR-alpha* and *HNF4-alpha* are multifunctional transcription factors primarily expressed in the liver[7,27]. *PPAR-alpha* is a known regulator of lipid metabolism in the liver. In a study by the Diabetes Prevention Program, one SNP in *PPAR-alpha* was associated with diabetes incidence and another variant showed significant interaction with metformin intervention[37]. Mutations in *HNF4-alpha* have also been linked to Type 2 diabetes[44]. Moreover, there is literature evidence to suggest that both *PPAR-alpha* and *HNF4-alpha* are important regulators of *SLC22A1* (OCT1), a key transporter mediating the uptake of metformin into the liver, the primary site of action[5,27]. Our regression results suggest that *PPAR-alpha* and *HNF4-alpha* are important modulators of metformin pharmacodynamics, perhaps independent of pharmacokinetics. A total of 23 genetic variants in *PPAR-alpha* and *HNF4-alpha* associated with treatment HbA1c levels, none of which were explained by a pharmacokinetic mechanism. Furthermore, gene expression levels for both transcription factors were strongly correlated with OCT1 expression, as well as other metformin transporters in the liver. *PPAR-alpha* and *HNF4-alpha* may be involved in regulating OCT1 and other genes in the liver that play a role in metformin disposition and glycemic response.

In addition to genetics, ethnicity was a significant predictor of metformin pharmacokinetics, affecting inter-compartmental clearance (Q/F) and apparent clearance (CL/F). African Americans had significantly higher mean CL/F and Q/F

estimates compared to European Americans and Asian Americans. The effects of creatinine clearance and body weight did not confound this effect observed in our model. Furthermore, based on our simulations, for African Americans to achieve similar metformin exposure to European Americans, a 26% increase in dose should be considered. This is taking into account similar creatinine clearances and *SP1* genotype status for both ethnicities. This means that where a European American individual would start with an 850 mg dose of metformin, an African American would require at least a 1000 mg dose of the drug to attain similar exposure levels. The impact of ethnicity on metformin response still needs to be investigated. Also, further studies and different approaches are required to validate this observation for metformin pharmacokinetics.

Overall, this study demonstrates that genetic variants in key transcription factor genes, along with transporters and ethnicity, are important determinants of metformin pharmacokinetics and pharmacodynamics. Transcription factors may regulate gene expression levels either through enhancer or repressor activity. In some cases, e.g., *SP1* and *AP2*, the association of genetic variants with pharmacodynamics may be mediated through pharmacokinetic mechanisms, which ultimately control systemic blood levels of the drug. For example, *SP1* may regulate the expression of a system of transporters in the kidney involved in metformin elimination. In other cases, *PPAR-alpha* for example, the observed mechanisms for the effects of SNPs are unclear. *PPAR-alpha* SNPs may be modulating metformin pharmacodynamics independent of the effects on systemic levels of metformin. Clearly, future studies are needed to further clarify the

biological roles of *SP1, AP2, HNF4-alpha* and *PPAR-alpha* in the disposition and action of metformin.

## REFERENCES


1. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 ( OCT1 ) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

2. Gong, L., Goswami, S., Giacomini, K. M., Altman, R. B. & Klein, T. Metformin pathways: pharmacokinetics and pharmacodynamics. *Pharmacogenet Genomics* **22,** 820–27 (2013).

3. Ha Choi, J., Wah Yee, S., Kim, M. J., Nguyen, L., Ho Lee, J., Kang, J.-O., Hesselson, S., Castro, R. A., Stryke, D., Johns, S. J., Kwok, P.-Y., Ferrin, T. E., Goo Lee, M., Black, B. L., Ahituv, N. & Giacomini, K. M. Identification and characterization of novel polymorphisms in the basal promoter of the human transporter, MATE1. *Pharmacogenet. Genomics* **19,** 770–80 (2009).

4. Alexander, G. C., Sehgal, N. L., Moloney, R. M. & Stafford, R. S. National trends in treatment of type 2 diabetes mellitus, 1994-2007. *Arch. Intern. Med.* **168,** 2088–94 (2008).

5. Stocker, S. L., Morrissey, K. M., Yee, S. W., Castro, R., Xu, L., Dahlin, A., Ramirez, H., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L., Brett, C. M. & Giacomini, K. M. The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **93,** 186–94 (2013).

6. Graham, G. G., Punt, J., Arora, M., Day, R. O., Doogue, M. P., Duong, J. K., Furlong, T. J., Greenfield, J. R., Greenup, L. C., Kirkpatrick, C. M., Ray, J. E., Timmins, P. & Williams, K. M. Clinical pharmacokinetics of metformin. *Clin. Pharmacokinet.* **50,** 81–98 (2011).

7. Saborowski, M., Kullak-Ublick, G. A. & Eloranta, J. J. The human organic cation transporter-1 gene is transactivated by hepatocyte nuclear factor-4alpha. *J. Pharmacol. Exp. Ther.* **317,** 778–85 (2006).

8. Kamiyama, Y., Matsubara, T., Yoshinari, K., Nagata, K., Kamimura, H. & Yamazoe, Y. Role of human hepatocyte nuclear factor 4alpha in the expression of drug-metabolizing enzymes and transporters in human

hepatocytes assessed by use of small interfering RNA. *Drug Metab. Pharmacokinet.* **22,** 287–98 (2007).

9. Cook, M. N., Girman, C. J., Stein, P. P. & Alexander, C. M. Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with type 2 diabetes in UK primary care. *Diabet. Med.* **24,** 350–58 (2007).

10. Owen, M. R., Doran, E. & Halestrap, A. P. Evidence that metformin exerts its anti-diabetic effects through inhibition of complex 1 of the mitochondrial respiratory chain. *Biochem. J.* **348 Pt 3,** 607–14 (2000).

11. Zhou, K., Bellenguez, C., Spencer, C. C., Bennett, A. J., Coleman, R. L., Tavendale, R., Hawley, S., Donnelly, L., Schofield, C., Groves, C. J., Burch, L., Carr, F., Strange, A., Freeman, C., Blackwell, J. M., Bramon, E., Brown, M., Casas, J. P., Corvin, A., Craddock, N., Deloukas, P., Dronov, S., Duncanson, A., Edkins, S., Gray, E., Hunt, S., Jankowski, J., Langford, C., Markus, H. S., Mathew, C. G., Plomin, R., Rautanen, A., Sawcer, S. J., Samani, N. J., Trembath, R., Viswanathan, A. C., Wood, N. W., Harries, L. W., Hattersley, A. T., Doney, A. S. F., Colhoun, H., Morris, A. D., Sutherland, C., Hardie, D. G., Peltonen, L., McCarthy, M. I., Holman, R. R., Palmer, C. N., Donnelly, P. & Pearson, E. R. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat. Genet.* **43,** 117–20 (2011).

12. Becker, M. L., Visser, L. E., van Schaik, R. H. N., Hofman, A., Uitterlinden, A. & Stricker, B. H. C. Genetic variation in the multidrug and toxin extrusion 1 transporter protein influences the glucose-lowering effect of metformin in patients with diabetes: a preliminary study. *Diabetes* **58,** 745–49 (2009).

13. Chen, Y., Li, S., Brown, C., Cheatham, S., Castro, R. A., Leabman, M. K., Urban, T. J., Chen, L., Yee, S. W., Choi, J. H., Huang, Y., Brett, C. M., Burchard, E. G. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 2, OCT2, on the renal elimination of metformin. *Pharmacogenet. Genomics* **5,** 1337–54 (2011).

14. Choi, J. H., Yee, S. W., Ramirez, H., Morrissey, K. M., Jang, G. H., Joski, P. J., Mefford, J., Hesselson, S. E., Schlessinger, A., Jenkins, G., Castro, R., Johns, S. J., Stryke, D., Sali, A., Ferrin, T. E., Witte, J. S., Kwok, P.-Y., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L. & Giacomini, K. M. A common 5'-UTR variant in MATE2-K is associated with poor response to metformin. *Clin. Pharmacol. Ther.* **90,** 674–84 (2011).

15. Tzvetkov, M. V, Vormfelde, S. V, Balen, D., Meineke, I., Schmidt, T., Sehrt, D., Sabolić, I., Koepsell, H. & Brockmöller, J. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3

on the renal clearance of metformin. *Clin. Pharmacol. Ther.* **86,** 299–306 (2009).

16. Takane, H., Shikata, E., Otsubo, K., Higuchi, S. & Ieiri, I. Polymorphism in human organic cation transporters and metformin action. *Pharmacogenomics* **9,** 415–22 (2008).

17. Scarpello, J. H. & Howlett, H. C. Metformin therapy and clinical uses. *Diab. Vasc. Dis. Res.* **5,** 157–67 (2008).

18. Fujioka, K., Brazg, R. L., Raz, I., Bruce, S., Joyal, S., Swanink, R. & Pans, M. Efficacy, dose-response relationship and safety of once-daily extended-release metformin (Glucophage® XR) in type 2 diabetic patients with inadequate glycaemic control despite prior treatment with diet and exercise. *Diabetes, Obes. Metab.* **7,** 28–39 (2005).

19. Garber, A. J., Duncan, T. G., Goodman, A. M., Mills, D. J. & Rohlf, J. L. Efficacy of metformin in type II diabetes: results of a double-blind, placebo-controlled, dose-response trial. *Am. J. Med.* **103,** 491–97 (1997).

20. Becker, M. L., Visser, L. E., van Schaik, R. H., Hofman, A., Uitterlinden, A. & Stricker, B. H. Genetic variation in the organic cation transporter 1 is associated with metformin response in patients with diabetes mellitus. *Pharmacogenomics J* **9,** 242–47 (2009).

21. Song, I. S., Shin, H. J., Shim, E. J., Jung, I. S., Kim, W. Y., Shon, J. H. & Shin, J. G. Genetic variants of the organic cation transporter 2 influence the disposition of metformin. *Clin. Pharmacol. Ther.* **84,** 559–62 (2008).

22. Wang, Z.-J., Yin, O. Q. P., Tomlinson, B. & Chow, M. S. S. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. *Pharmacogenet. Genomics* **18,** 637–45 (2008).

23. Kim, D. H. D., Sriharsha, L., Xu, W., Kamel-Reid, S., Liu, X., Siminovitch, K., Messner, H. & Lipton, J. H. Clinical relevance of a pharmacogenetic approach using multiple candidate genes to predict response and resistance to imatinib therapy in chronic myeloid leukemia. *Clin. cancer Res.* **15,** 4750–58 (2009).

24. Kamiyama, Y., Matsubara, T., Yoshinari, K., Nagata, K., Kamimura, H. & Yamazoe, Y. Role of human hepatocyte nuclear factor 4alpha in the expression of drug-metabolizing enzymes and transporters in human hepatocytes assessed by use of small interfering RNA. *Drug Metab. Pharmacokinet.* **22,** 287–98 (2007).

25. Chung, J.-Y., Cho, S. K., Kim, T. H., Kim, K. H., Jang, G. H., Kim, C. O., Park, E.-M., Cho, J.-Y., Jang, I.-J. & Choi, J. H. Functional characterization of MATE2-K genetic variants and their effects on metformin pharmacokinetics. *Pharmacogenet. Genomics* **23,** 365–73 (2013).

26. Shu, Y., Bello, C. L., Mangravite, L. M., Feng, B. & Giacomini, K. M. Functional characteristics and steroid hormone-mediated regulation of an organic cation transporter in Madin-Darby canine kidney cells. *J. Pharmacol. Exp. Ther.* **299,** 392–98 (2001).

27. Nie, W., Sweetser, S., Rinella, M. & Green, R. M. Transcriptional regulation of murine Slc22a1 (Oct1) by peroxisome proliferator agonist receptor-alpha and -gamma. *Am. J. Physiol. Gastrointest. Liver Physiol.* **288,** G207–12 (2005).

28. Kajiwara, M., Terada, T., Asaka, J., Ogasawara, K., Katsura, T., Ogawa, O., Fukatsu, A., Doi, T. & Inui, K. Critical roles of Sp1 in gene expression of human and rat H+/organic cation antiporter MATE1. *Am. J. Physiol. Renal Physiol.* **293,** F1564–70 (2007).

29. Cho, S. K., Yoon, J. S., Lee, M. G., Lee, D. H., Lim, L., Park, K., Park, M. S. & Chung, J.-Y. Rifampin enhances the glucose-lowering effect of metformin and increases OCT1 mRNA levels in healthy participants. *Clin. Pharmacol. Ther.* **89,** 416–21 (2011).

30. Chen, L., Hong, C., Chen, E. C., Yee, S. W., Xu, L., Almof, E. U., Wen, C., Fujii, K., Johns, S. J., Stryke, D., Ferrin, T. E., Simko, J., Chen, X., Costello, J. F. & Giacomini, K. M. Genetic and epigenetic regulation of the organic cation transporter 3, SLC22A3. *Pharmacogenomics J.* **13,** 110–20 (2013).

31. Asaka, J., Terada, T., Ogasawara, K., Katsura, T. & Inui, K. Characterization of the basal promoter element of human organic cation transporter 2 gene. *J Pharmacol Exp Ther* **321,** 684–89 (2007).

32. O'Brien, V. P., Bokelmann, K., Ramírez, J., Jobst, K., Ratain, M. J., Brockmöller, J. & Tzvetkov, M. V. Hepatocyte nuclear factor 1 regulates the expression of the organic cation transporter 1 via binding to an evolutionary conserved region in intron 1 of the OCT1 gene. *J. Pharmacol. Exp. Ther.* **347,** 181–92 (2013).

33. Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., Altman, R. B. & Klein, T. E. Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **92,** 414–27 (2012).

34. Chen, L., Pawlikowski, B., Schlessinger, A., More, S. S., Stryke, D., Johns, S. J., Portman, M. A., Chen, E., Ferrin, T. E., Sali, A. & Giacomini, K. M.

Role of organic cation transporter 3 (SLC22A3) and its missense variants in the pharmacologic action of metformin. *Pharmacogenet. Genomics* **20,** 687–99 (2010).

35. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. W., Daly, M. J. & Sham, P. C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–75 (2007).

36. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3 (Bethesda).* **1,** 457–70 (2011).

37. Jablonski, K. A., McAteer, J. B., De Bakker, P. I. W., Franks, P. W., Pollin, T. I., Hanson, R. L., Saxena, R., Fowler, S., Shuldiner, A. R., Knowler, W. C., Altshuler, D. & Florez, J. C. Common variants in 40 genes assessed for diabetes incidence and response to metformin and lifestyle intervention in the diabetes prevention program. *Diabetes* **59,** 2672–81 (2010).

38. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27,** 1017–28 (2011).

39. Podvinec, M., Kaufmann, M. R., Handschin, C. & Meyer, U. NUBIScan, an in silico approach for prediction of nuclear receptor response elements. *Mol. Endocrinol.* **16,** 1269–79 (2002).

40. Sambol, N. C., Chiang, J., O’Conner, M., Liu, C. Y., Lin, E. T., Goodman, M., Benet, L. Z. & Karam, J. H. Pharmacokinetics and pharmacodynamics of metformin in healthy subjects and patients with noninsulin-dependent diabetes mellitus. *J. Clin. Pharmacol.* **36,** 1012–21 (1996).

41. Little, R. R. & Rohlfing, C. L. Analytical goals for HbA1c: are HbA1c results good enough for optimal use? *J. Diabetes* **3,** 3–6 (2011).

42. Hirst, J., Farmer, A. J., Ali, R., Roberts, N. W. & Stevens, R. J. Quantifying the effect of metformin treatment and dose on glycemic control. *Diabetes Care* **35,** 446–54 (2012).

43. Suda, S., Rai, T., Sohara, E., Sasaki, S. & Uchida, S. Postnatal expression of KLF12 in the inner medullary collecting ducts of kidney and its trans-activation of UT-A1 urea transporter promoter. *Biochem. Biophys. Res. Commun.* **344,** 246–52 (2006).

44. Mohlke, K. L. & Boehnke, M. The role of HNF4A variants in the risk of type 2 diabetes. *Curr. Diab. Rep.* **5,** 149–56 (2005).

## Chapter 3

## Discovery of Expression Quantitative Trait Loci (eQTLs) of Metformin Transporters

### INTRODUCTION

To date, pharmacogenetic research on metformin response has primarily focused on investigating the role of non-synonymous variants in kidney and liver specific transporters[1–3]. Despite promising findings, researchers are still unable to explain a significant portion of the response variation observed in type 2 diabetes (T2D) patients on metformin. Variation in the overall protein expression levels of solute carrier (SLC) transporters governing metformin disposition may be crucial in determining variation in metformin response[3,4]. Gene expression is one of the major determinants of how much protein is expressed; hence understanding the mechanistic nature of how genetic variants influence gene expression variation of crucial transporters may aid in analyzing the underlying physiology that governs metformin response[5].

Gene expression varies significantly between individuals; it is believed that expression quantitative trait loci (eQTLs), which are genome loci that are associated with gene expression, may account for a significant portion of this inter-individual variability in expression[6]. Identification of regulatory elements in metformin transporter genes is therefore essential for understanding the factors that control the expression level of these genes. eQTLs will highlight genomic regions (e.g. transcription binding sites, enhancer sites, etc.) that modulate the expression of SLC transporters involved in metformin disposition and elimination

and provide a biological basis for the pharmacological variation observed. With the use of various pattern-matching algorithms, this analysis may also highlight novel transcription factors that regulate gene expression.

Gene expression regulation in individuals has been extensively studied in human lymphoblastoid cell lines and a number of primary tissues[7–13]. While eQTLs have been identified in several human tissue types, there are a limited number of studies to date that have identified eQTLs in the kidney. The discovery of eQTLs can indicate genomic regions that regulate gene expression, identify transcription factors involved in this mechanism, and increase overall understanding of the underlying cellular biology in the kidney in relation to the pharmacological variation of metformin. The goal of the current study was to discover eQTLs of the transporters that modulate metformin disposition: OCT1, OCT2, OCT3, MATE1 and MATE2-K. The clinical basis of identified eQTLs were then investigated by genotyping healthy volunteers and linking the variant to pharmacokinetic phenotypes. From this study, genomic regions proximal to metformin transporters were linked to expression levels of *SLC47A1*, *SLC22A3*, and *SLC22A2*, with a potential transcription factor-binding hypothesis (e.g. *SP1*). Additionally, variants in transcription factor *HNF4-alpha* were the most influential trans-eQTLs, correlating with expression level variation in both *SLC47A1* and *SLC22A1*. Our findings provide a mechanistic basis of the identified variants that affect metformin pharmacology.

## METHODS

### Kidney tissue samples

Kidney tissues were purchased from commercial vendors and consist of renal cortex samples from Caucasian males and females excised during surgery. A BioTrove qPCR instrument was used to determine gene expression levels for several transporters, enzymes, and transcription factors. An Affymetrix Axiom array designed to finely map confirmed disease associated regions and ADME (*Absorption, Disposition, Metabolism, and Excretion*) genes was used to generate genotyping data. The transporter genes of focus that were extracted for analysis include *SLC22A1*, *SLC22A2*, *SLC22A3*, *SLC47A1* and *SLC47A2*.

### Post processing of expression and genotype data

Raw expression data were evaluated by principal component analysis to identify and remove extreme outliers in gene expression patterns. Expression data from the aforementioned metformin transporters were extracted and normalized by calculating the geometric mean of housekeeping genes. To ensure high quality genotype information, the data were filtered based on genotype call rate, minor allele frequency, relatedness of patients, and gender confirmation.

### Computational analysis

A linear regression analysis was performed for the transporters of interest in order to interrogate SNPs in both *cis* and *trans* regions. A *cis* region was

defined here by containing all SNPs within a 50 kilobase (kb) region proximal to the transporter gene of interest. A *trans* region in this context contains SNPs in a 20 kb region surrounding a transcription factor gene of interest. Transcription factors were included when highly expressed in the kidney and meeting one or more of the following criteria: 1) Previously implicated in literature as a genetic regulator of SLC transporters, 2) Having a strong correlation with gene expression levels of an SLC transporter, or 3) Computationally identified in TRANSFAC to bind to an evolutionarily conserved, non-exonic region within 50 kb of a metformin transporter gene[14]. Evolutionary conserved regions (ECRs) were determined using an ECR browser[15]. Association was performed using linear regression in PLINK (a whole genome association analysis toolset) assuming an additive genetic model (Figure 3.1)[16]. Genotypes were encoded as 0, 1, or 2, depending on the number of minor alleles present. The additive model examined an association between genotype count and expression level. For example, a linear trend in gene expression was expected with the addition of each minor allele (i.e., 0, 1, or 2).

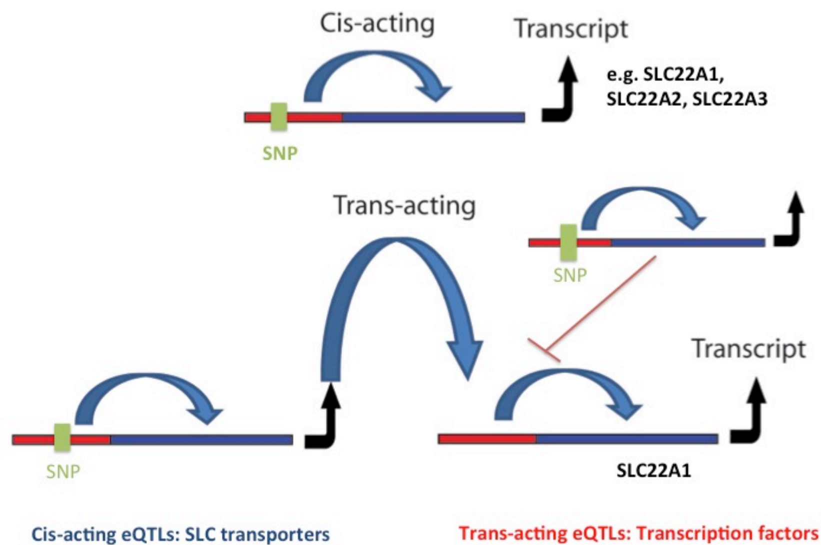



**Figure 3.1. Proposed mechanism of SNP influence on SLC expression.** *Cis*-acting and *trans*-acting mechanisms either inhibit or induce gene expression levels of the transporter.

To adjust for the number of SNPs tested, other statistical measures were used to further investigate statistical significance, including the empirical P-value and False Discovery Rate (FDR).

**Linkage disequilibrium (LD) analysis**

Significant SNPs were collected and expanded for further analysis. Statistical significance of a SNP from the association analysis does not necessarily indicate that the SNP is causative. Hence, we investigated all SNPs in LD to top associated SNPs ($R^2 \geq 0.8$) using the HapMap data conversion tool with 1000 genomes. Upon obtaining the expanded list of *cis* region SNPs (*cis*-eQTLs), I functionally annotated each SNP (e.g., whether the SNP is in a coding region, near a TSS, TES, etc.). Computational exploration of the biological

mechanism was performed to investigate the affinity of transcription factor binding to DNA regions consisting of the interrogated polymorphism. A number of online tools such as TRANSFAC, JASPAR, TFSEARCH, and TOMTOM were used for this investigation[14,17–19].

**Clinical basis of eQTLs**

In addition to examining the mechanistic nature of the statistical association, the clinical relevance of the genetic polymorphism was also investigated in order to link the impact of the identified eQTL to a pharmacokinetic or pharmacodynamic phenotype and therefore also link the genetic mechanism to the clinical outcome. Previously described healthy volunteer studies at UCSF provided information on secretion clearance (CLsr), renal clearance, exposure (AUC), and fasting plasma glucose (FPG)[1,20]. After finalizing the top eQTLs, the SNPs were subsequently genotyped in healthy volunteer samples with the goal of revealing an association between the SNP and a pharmacological phenotype of metformin.

## RESULTS

**Summary of gene expression levels across metformin transporters**

A total of 60 kidney tissue samples were used for this analysis. Gene expression values were normalized to housekeeping genes (delta Ct normalized). Overall, there was marked variation observed in expression levels across all metformin transporters. The mean delta Ct values for *SLC22A1*,

*SLC22A2*, *SLC22A3*, *SLC47A1*, and *SLC47A2* were 9.8, 4.47, 7.46, 6.6, and 9.13, respectively (higher delta Ct value corresponds to lower expression). Since the mRNA transcript level is inversely proportional to the delta Ct value, the negative value of delta Ct normalized is summarized in Figure 3.2, so that higher expression levels may correspond to higher numbers. *SLC22A2* had the largest variation with a %CV of 32%, followed by *SLC47A1* (26%) and *SLC22A3* (20.7%).

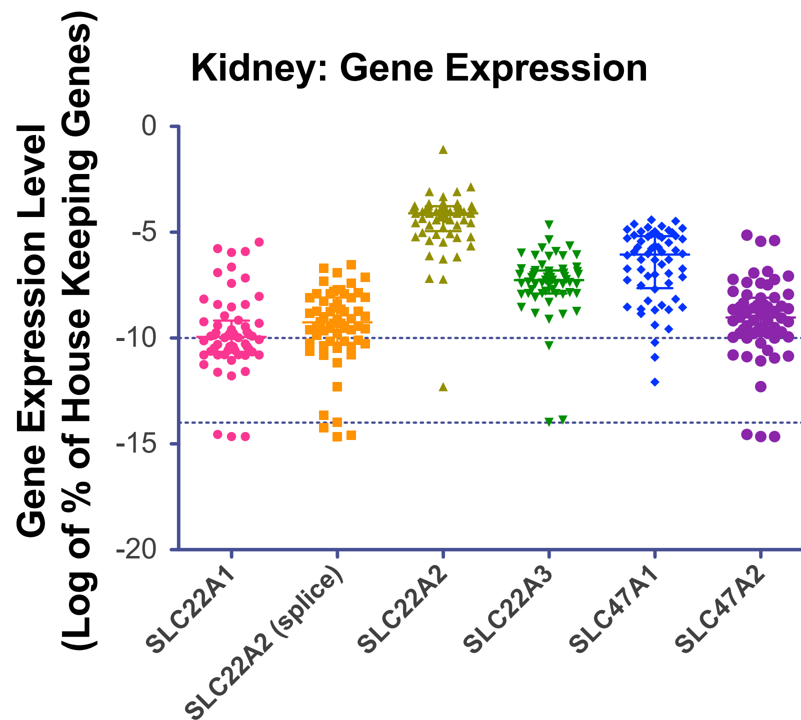



**Figure 3.2. Summary of gene expression variation of metformin transporters.**
Logarithm of % of housekeeping genes is the inverse of delta Ct normalized. In real time PCR, the Ct (cycle threshold) is defined as the number of cycles required for the fluorescent signal to cross the threshold. The Ct levels are inversely proportional to the amount of the targeted mRNA in the sample.

### *Cis*-eQTL Analysis

A regression analysis of genetic variants less than 50 kb from each transporter gene was performed to identify *cis*-eQTLs of metformin transporters. A statistical filter of $p < 0.05$ was used to prioritize variants, with the top variants listed in Table 3.1.

**Table 3.1** Summary of statistically significant cis-eQTLs.

| Transporter | SNP | p value | Kruskal-Wallis statistic | Identified TF | TF Expressed in Kidney? |
|---|---|---|---|---|---|
| *SLC47A1* | rs2250486 | 0.01 | yes | *Sox-5* | Yes |
| *SLC47A1* | rs2247436 | 0.02 | no | *SP1/TATA/ FOXA1/A2* | Yes |
| *SLC47A1* | rs6587200 | 0.03 | no | *GATA-1* | Yes |
| *SLC22A2* | rs604258 | 0.02 | no | NA | NA |
| *SLC22A2* | rs576075 | 0.04 | no | NA | NA |
| *SLC22A3* | rs600987 | 0.05 | yes | *ADR1* | Yes |

TF= Transcription factor. TFSEARCH is an online tool that was used to identify transcription factors that exhibit differential DNA binding once the SNP is introduced in the cis-eQTL containing DNA sequence. NA = No identified TF from TFSEARCH. The column 'TF Expressed in the Kidney' summarizes whether the transcription factor (TF) identified has known expression in the kidney.

The most significant variant was rs2250486 ($p = 0.01$), an intronic variant in the *SLC47A1* gene (Figure 3.3). This variant, along with *SLC22A3* variants (rs60098, $p < 0.05$) remained significant even when exploring an independent statistical measure (i.e. Kruskal-Wallis test). Both variants were also significant after performing a multiple testing procedure using permutation correction, in which the expression level was permuted 10,000 times.

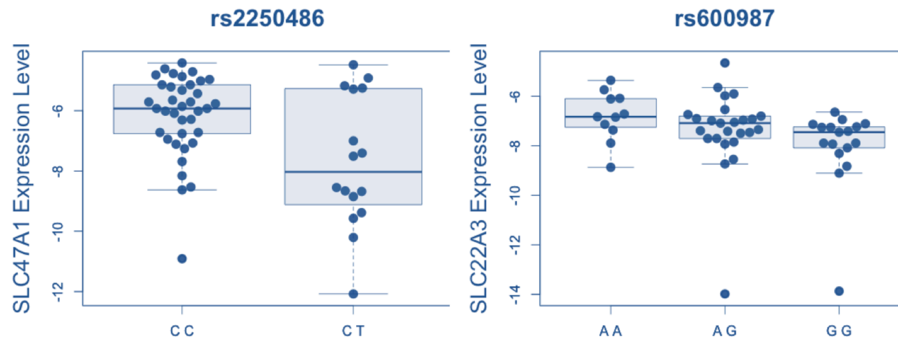



**Figure 3.3. Top cis-eQTLs associated with expression levels of *SLC47A1* and *SLC22A3*.**
Expression level units reflect the negative value of delta-Ct normalized. For SNP rs2250486, homozygous TT alleles were not present.

The initial variant list was subsequently expanded to include SNPs in strong linkage disequilibrium ($r^2 \geq 0.8$) to the statistically associated variants. Transcription factor binding algorithms (i.e., TFSEARCH and JASPAR) were applied to the DNA region (±10 base pairs) containing the variant of interest in order to identify the transcription factors that have a lower probability of binding to the genomic region once the variant has been introduced. This particular analysis highlighted *Sox-5, SP1, FOXA1, TATA, GATA-1*, and *ADR-1* as transcription factors that demonstrate (*in silico*) differential binding once the variant is introduced into the system.

**Summary of top trans-eQTLs**

An eQTL analysis of genetic variants less than 20 kilobases from a subset of transcription factor genes was performed to identify significantly associated *trans* eQTLs of metformin transporters. The major result from this analysis was in transcription factor gene *HNF4-alpha*, where SNPs in this transcription factor affected the expression levels of *SLC47A1* as well as *SLC22A1* (Table 3.2).

**Table 3.2.** *HNF4-alpha* SNPs (P < 0.05) associated with metformin transporters.

| **Main SNP Significant (P < 0.05)** | **LD SNPs to Main SNP** | **Transporter Association** |
|---|---|---|
| rs2093248 (p = 0.03) | rs911358<br>rs2093247 | OCT1 |
| rs3212198 (p = 0.02) | rs3212199 | MATE1 |
| rs6093978 (p = 0.04) | rs6073432<br>rs6073433 | MATE1 |
| rs6073432 (p = 0.04) | rs34956692<br>rs6093978<br>rs6073433 | MATE1 |
| rs6031606 (p = 0.04) | rs6103738<br>rs6017344 ($r^2$ = 0.68) | OCT1 |

Transporter association indicates the transporter mRNA levels that are significantly associated with the main trans SNP. LD=Linkage disequilibrium.

The list was subsequently expanded to include variants that had a strong LD to the primary SNP of interest. The most significant trans eQTLs (rs3212198 and rs6093978) were in the intronic regions of *HNF4-alpha* (p < 0.01) (Figure 3.4).

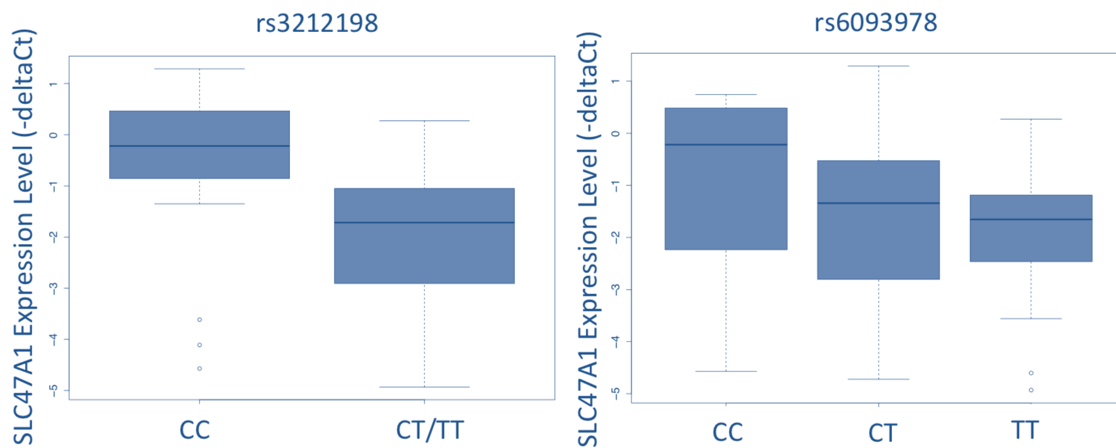



**Figure 3.4. Association between top *HNF4-alpha* eQTLs with *SLC47A1* expression level.**

### *HNF4-alpha* was linked to metformin secretory clearance

A clinical analysis was performed on prioritized *cis* and *trans* eQTLs in healthy volunteers in order to investigate the role of these variants and transcription factors on relevant pharmacokinetic phenotypes. Of the 16 variants investigated, only one variant (rs2093248) had a significant association with the pharmacokinetics of metformin and a significant association with metformin secretory clearance ($P < 0.05$). A proposed pharmacological mechanism describing how this SNP may impact metformin pharmacokinetics is shown in Figure 3.5.

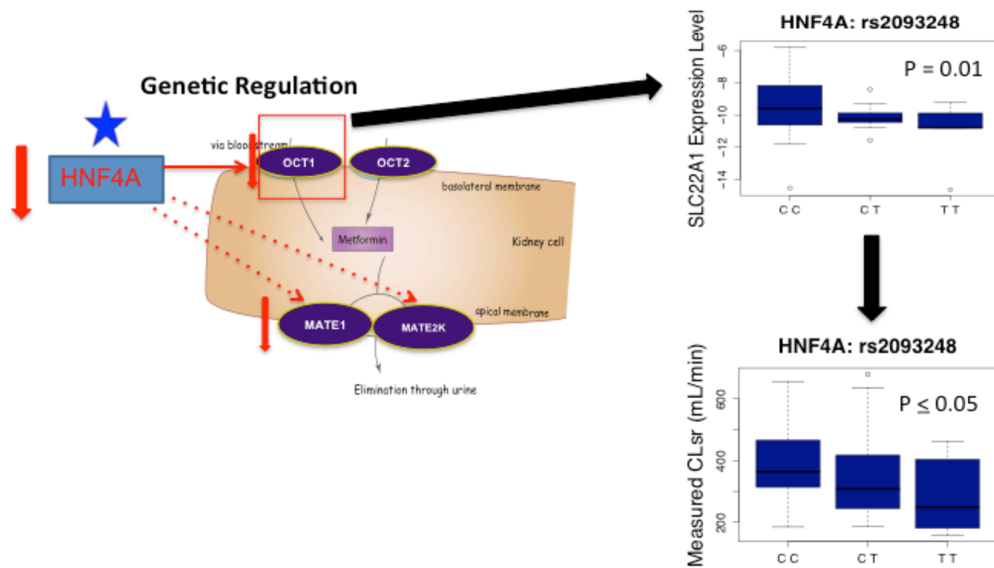



**Figure 3.5. Proposed mechanism of an *HNF4-alpha* SNP effect on metformin secretory clearance.**
CLsr = secretory clearance of metformin. *HNF4-alpha* regulates a system of transporters expressed in the proximal tubule of the kidney (i.e. OCT1, MATE1, MATE2K), ultimately having a downstream impact on metformin pharmacokinetics (i.e. metformin secretory clearance).

## DISCUSSION

This study explores the link between genetic variants and the gene expression levels of transporters involved in metformin disposition and elimination. It was hypothesized that this mechanistic relationship will ultimately influence the pharmacological variation of metformin. Two findings emerged from this analysis: 1) Genomic regions proximal to metformin transporters were linked to expression levels of *SLC47A1*, *SLC22A3*, and *SLC22A2*, with a potential transcription factor binding hypothesis and 2) Variants in transcription factor

HNF4-alpha were the most influential trans-eQTLs, correlating with expression level variation in both *SLC47A1* and *SLC22A1*.

From the *cis*-eQTL analysis, a total of 6 variants were significantly associated with the expression levels of *SLC22A2*, *SLC22A3*, and *SLC47A1*. Further investigation of these proximal genomic regions indicated transcription factors that may be novel regulatory factors of these genes. From this analysis, three *cis* regions less than 50 kb from *SLC47A1* and hypothesized to play a mechanistic role were discovered. These regions also demonstrated differential binding of different transcription factors (e.g., *Sox-5, GATA-1, SP1,* and *TATA*). However, the pharmacological role of these transcription factors remains unclear. For example, TATA binding proteins are known to bind to a region of DNA typically found 30 base pairs upstream of a gene called the TATA box and may modulate expression of these genes. The variant rs2247436 is located in the intron of the *SLC47A1* gene. This intronic region may be a site of heavy regulatory activity, as multiple transcription factors such as *SP1*, *FOXA1*, and *FOXA2* may to bind to this particular DNA region[21]. Of these transcription factors, *SP1* is of particular interest as it has been previously shown to modulate expression levels of MATE1[22]. Additionally, we have shown in a previous study that the *SP1* gene may modulate metformin pharmacodynamics and pharmacokinetics[2]. It is important to note that the same *SP1* variants associated with metformin PK/PD are not linked to expression levels of metformin transporters in this study. Nevertheless, this is the first study to demonstrate a

mechanistic link between *SP1* and gene expression levels of *SLC47A1* through a cis-eQTL based hypothesis.

Trans-eQTL analysis highlighted the potential importance of *HNF4-alpha* in regulating the expression levels of *SLC22A1* and *SLC47A1*. A total of 2 variants were significantly associated with the expression level of SLC22A1 and 3 variants were associated with the expression level of *SLC47A1*. *HNF4-alpha* is a multifunctional transcription factor primarily expressed in the liver[23]. Mutations in *HNF4-alpha* have also been linked to T2D[24]. Literature evidence also suggests that *HNF4-alpha* is an important regulator of *SLC22A1* (OCT1), and a key transporter in mediating the uptake of metformin into the liver, which is the primary site of action[25]. Similar to *SP1*, our lab has previously shown a link between *HNF4-alpha* and metformin pharmacokinetics and pharmacodynamics[2].

In this study, I provide eQTL-based support to further demonstrate the importance of this transcription factor on a cellular level. To investigate the clinical significance of these eQTLs in healthy volunteers, this particular variant was genotyped in healthy volunteers and a regression analysis was performed with metformin secretory clearance. Results indicated that the T allele of the eQTL rs2093248 was associated with a lower metformin secretory clearance in an additive manner. This directionality of the minor allele follows both mechanistic and pharmacological expectations as the T allele was linked to lower expression levels of *SLC22A1*. Although the expression of *SLC22A1* in the proximal tubule of the kidney has not been confirmed, this analysis did show reasonable levels of this transporter in the kidney. Furthermore, the *trans*-eQTL

based analysis also provided a link between *HNF4-alpha* and *SLC47A1*, a transporter that has been well profiled in the kidney. A proposed mechanism by which a SNP in *HNF4-alpha* may influence metformin secretory clearance suggests that it does so through the modulation of gene expression levels of *SLC22A1* and *SLC47A1* in renal proximal tubule cells in the kidney.

In summary, the discovery of kidney specific eQTLs of metformin transporters provided insight into important genomic regions and transcription factors that regulate gene expression levels. These key findings may ultimately affect the pharmacological outcome of patients on metformin therapy. Experimental validation and replication of such genomic regions and transcription factors are a critical next step that will detail our understanding of the underlying biology that governs metformin response variation.

**REFERENCES**


1. Stocker, S. L., Morrissey, K. M., Yee, S. W., Castro, R., Xu, L., Dahlin, A., Ramirez, H., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L., Brett, C. M. & Giacomini, K. M. The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **93,** 186–94 (2013).

2. Goswami, S., Yee, S. W., Stocker, S., Mosley, J. D., Kubo, M., Castro, R., Mefford, J., Wen, C., Liang, X., Witte, J., Brett, C., Maeda, S., Simpson, M. D., Hedderson, M. M., Davis, R. L., Roden, D. M., Giacomini, K. M. & Savic, R. M. Genetic variants in transcription factors are associated with the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **96,** 370–90 (2014).

3. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 ( OCT1 ) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

4. Tzvetkov, M. V, Vormfelde, S. V, Balen, D., Meineke, I., Schmidt, T., Sehrt, D., Sabolić, I., Koepsell, H. & Brockmöller, J. The effects of genetic polymorphisms in the organic cation transporters OCT1, OCT2, and OCT3 on the renal clearance of metformin. *Clin. Pharmacol. Ther.* **86,** 299–306 (2009).

5. Guo, Y., Xiao, P., Lei, S., Deng, F., Xiao, G. G., Liu, Y., Chen, X., Li, L., Wu, S., Chen, Y., Jiang, H., Tan, L., Xie, J., Zhu, X., Liang, S. & Deng, H. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochim. Biophys. Sin. (Shanghai).* **40,** 426–36 (2008).

6. Cheung, V. G., Conlin, L. K., Weber, T. M., Arcaro, M., Jen, K.-Y., Morley, M. & Spielman, R. S. Natural variation in human gene expression assessed in lymphoblastoid cells. *Nat. Genet.* **33,** 422–35 (2003).

7. Göring, H. H. H., Curran, J. E., Johnson, M. P., Dyer, T. D., Charlesworth, J., Cole, S., Jowett, J. B. M., Abraham, L. J., Rainwater, D. L., Comuzzie, A. G., Mahaney, M. C., Almasy, L., MacCluer, J. W., Kissebah, A. H., Collier, G. R., Moses, E. K. & Blangero, J. Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. *Nat. Genet.* **39,** 1208–16 (2007).

8. Stranger, B. E., Nica, C., Forrest, M. S., Dimas, A., Bird, C. P., Beazley, C., Ingle, C. E., Dunning, M., Flicek, P., Koller, D., Montgomery, S., Tavare, S., Deloukas, P. & Dermitzakis, E. T. Population genomics of human gene expression. *Nat.Genet.* **39,** 1217–24 (2007).

9. Dimas, A. S., Deutsch, S., Stranger, B. E., Montgomery, S. B., Borel, C., Attar-Cohen, H., Ingle, C., Beazley, C., Gutierrez Arcelus, M., Sekowska, M., Gagnebin, M., Nisbett, J., Deloukas, P., Dermitzakis, E. T. & Antonarakis, S. E. Common regulatory variation impacts gene expression in a cell type-dependent manner. *Science* **325,** 1246–50 (2009).

10. Stranger, B. E., Forrest, M. S., Clark, A. G., Minichiello, M. J., Deutsch, S., Lyle, R., Hunt, S., Kahl, B., Antonarakis, S. E., Tavaré, S., Deloukas, P. & Dermitzakis, E. T. Genome-wide associations of gene expression variation in humans. *PLoS Genet.* **1,** 695–704 (2005).

11. Gibbs, J. R., van der Brug, M. P., Hernandez, D. G., Traynor, B. J., Nalls, M. A., Lai, S. L., Arepalli, S., Dillman, A., Rafferty, I. P., Troncoso, J., Johnson, R., Zielke, H. R., Ferrucci, L., Longo, D. L., Cookson, M. R. & Singleton, A. B. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet.* **6,** 29–31 (2010).

12. Myers, A. J., Gibbs, J. R., Webster, J. A., Rohrer, K., Zhao, A., Marlowe, L., Kaleem, M., Leung, D., Bryden, L., Nath, P., Zismann, V. L., Joshipura, K., Huentelman, M. J., Hu-Lince, D., Coon, K. D., Craig, D. W., Pearson, J. V, Holmans, P., Heward, C. B., Reiman, E. M., Stephan, D. & Hardy, J. A survey of genetic human cortical gene expression. *Nat. Genet.* **39,** 1494–99 (2007).

13. Schadt, E. E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P. Y., Kasarskis, A., Zhang, B., Wang, S., Suver, C., Zhu, J., Millstein, J., Sieberts, S., Lamb, J., GuhaThakurta, D., Derry, J., Storey, J. D., Avila-Campillo, I., Kruger, M. J., Johnson, J. M., Rohl, C. A., van Nas, A., Mehrabian, M., Drake, T. A., Lusis, A. J., Smith, R. C., Guengerich, F. P., Strom, S. C., Schuetz, E., Rushmore, T. H. & Ulrich, R. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6,** 107–13 (2008).

14. Chekmenev, D. S., Haid, C. & Kel, A. E. P-Match: transcription factor binding site search by combining patterns and weight matrices. *Nucleic Acids Res.* **33,** W432–37 (2005).

15. Ovcharenko, I., Nobrega, M. A., Loots, G. G. & Stubbs, L. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. *Nucleic Acids Res.* **32,** 280–86 (2004).

16. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. W., Daly, M. J. & Sham, P. C. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–75 (2007).

17. Podvinec, M., Kaufmann, M. R., Handschin, C. & Meyer, U. NUBIScan, an in silico approach for prediction of nuclear receptor response elements. *Mol. Endocrinol.* **16,** 1269–79 (2002).

18. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27,** 1017–28 (2011).

19. Sandelin, A., Alkema, W., Engström, P., Wasserman, W. W. & Lenhard, B. JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32,** D91–94 (2004).

20. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 (OCT1) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

21. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. & Haussler, A. D. The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

22. Kajiwara, M., Terada, T., Asaka, J., Ogasawara, K., Katsura, T., Ogawa, O., Fukatsu, A., Doi, T. & Inui, K. Critical roles of Sp1 in gene expression of human and rat H+/organic cation antiporter MATE1. *Am. J. Physiol. Renal Physiol.* **293,** F1564–70 (2007).

23. Saborowski, M., Kullak-Ublick, G. A. & Eloranta, J. J. The human organic cation transporter-1 gene is transactivated by hepatocyte nuclear factor-4alpha. *J. Pharmacol. Exp. Ther.* **317,** 778–85 (2006).

24. Mohlke, K. L. & Boehnke, M. The role of HNF4A variants in the risk of type 2 diabetes. *Curr. Diab. Rep.* **5,** 149–56 (2005).

25. Lu, H., Gonzalez, F. J. & Klaassen, C. Alterations in hepatic mRNA expression of phase II enzymes and xenobiotic transporters after targeted disruption of hepatocyte nuclear factor 4 alpha. *Toxicol. Sci.* **118,** 380–90 (2010).

# Chapter 4

# A Longitudinal HbA1c Model Elucidates Genes Linked to Disease Progression on Metformin

## INTRODUCTION

Metformin is the first line of therapy for treatment of type 2 diabetes (T2D) and is one of the most frequently prescribed drugs worldwide[1–3]. As the global incidence of T2D rapidly increases, the low cost of metformin makes this treatment option particularly attractive in developing nations. Understanding metformin's efficacy in different patient populations with diverse genetic backgrounds will be critical in managing this deleterious metabolic disorder. Metformin lowers both basal and postprandial glucose in patients with T2D and works by inhibiting hepatic glucose production, reducing intestinal glucose absorption, and improving glucose uptake and utilization[2,4]. Glycosylated hemoglobin (HbA1c) is formed through a non-enzymatic and irreversible reaction between hemoglobin and glucose; HbA1c is the primary surrogate biomarker for long-term glycemic control and drug response, reflecting the average glucose levels circulating in the blood over previous months[5]. This biomarker has been shown to be more reliable than fasting plasma glucose in terms of assessing long-term efficacy (measured by mortality, cardiovascular related outcomes, etc.) due to the high diurnal variation associated with fasting plasma glucose. Several studies have shown that HbA1c is strongly linked to adverse T2D-related cardiovascular outcomes and mortality[6–8].

Metformin response is highly variable; greater than 30% of patients taking metformin are considered non-responders and require other medication such as sulfonylureas and insulin instead of metformin. Baseline HbA1c levels vary significantly in the T2D population (5.5-15%)[9,10]. Most studies have been primarily focused on uncovering the effect of genetic variants in pharmacokinetic (PK) genes on static pharmacological phenotypes of metformin and fail to address the pleiotropic nature of metformin response[2,10–14]. One of the largest studies to date, which consisted of a genome-wide association study on metformin response in individuals from the United Kingdom, identified variants near the Ataxia Telangiectasia Mutated locus associated with the ability to achieve HbA1c below 7% in the first 18 months of metformin treatment[15]. Although significant, the dichotomous phenotype did not account for longitudinal fluctuations in the patient population. Although many studies have demonstrated associations between single nucleotide polymorphisms (SNPs) in biologically relevant genes with metformin PK and pharmacodynamics (PD), each variant only accounts for a small fraction of the variation in HbA1c levels.

Additionally, there have been no studies on the effect of genetic and demographic variables on long-term changes of HbA1c in patients on metformin. These covariates may influence the drug's efficacy or the patient's underlying disease progression and once accounted for, may make it easier to detect responders and non-responders to metformin[16]. The traditional approach considers a glycemic HbA1c change from baseline to evaluate the effectiveness of the drug. This approach, however, effectively collapses the time dimension in

the data by disregarding the actual trajectory of biomarker and the disease status over time. As a result, it not only ignores crucial information on disease progression, but also lumps together the short-term effects of a treatment with the long-term effects on the disease.

Longitudinal disease progression analysis allows for a quantitative assessment of drug treatment effect on the time-course of the disease/biomarker. Computational methods use mathematical models to describe or predict changes in the disease status as a function of time[16]. This is distinctly and clinically advantageous in that these methods allow researchers to understand the role of genes as well as any relevant demographic predictors on specific response curve characteristics (such as disease progression and the long-term dynamics of therapeutic effects). Non-linear Mixed Effect Analysis (NLME) is a very powerful statistical approach used for this longitudinal analysis. NLME effectively enhances the signal-to-noise ratio and enables the utility of all data points, irrespective of study design[17,18,19].

To date, current models that capture the time-course of HbA1c in relation to metformin therapy have been limited by small sample sizes and sparse measurements[16,20,21]. Furthermore, a comprehensive genetic analysis linking genetic variants to long-term HbA1c fluctuations has not yet been performed and consequently, there is no current knowledge regarding the influence of genetics on long-term HbA1c dynamics.

The overall aim of this research is to explain the variance in long-term response, linking genes, demographics, and clinical factors to the upward

trajectory of HbA1c levels using a rich, long-term HbA1c dataset from patients on metformin (Figure 4.1). Ultimately, this research also aims to use this predictive model as a tool to better identify non-responders to metformin therapy early on during the course of treatment.

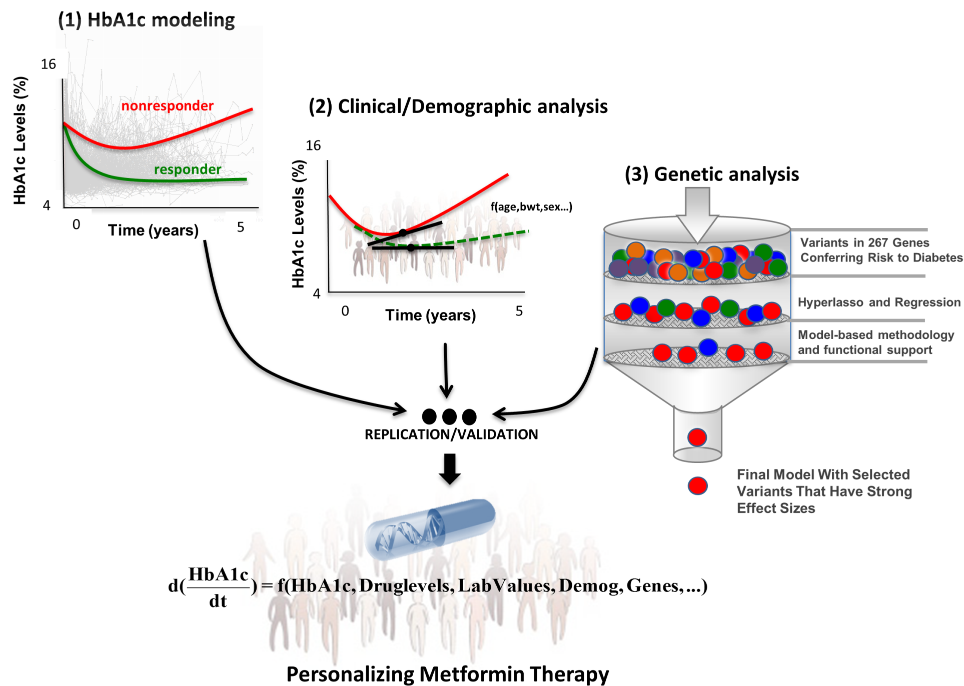



**Figure 4.1 Workflow of longitudinal modeling, genetic analysis and the potential clinical impact on individualizing metformin therapy.**

## METHODS

### Patients with type 2 diabetes

Diabetic patients of European American, African American, and Asian American ancestry were recruited into a multicenter retrospective study as described previously[13,22]. All patients were metformin-naive, had HbA1c levels

measured before and after initiation of metformin therapy (between 3 and 18 months), and had a medication possession ratio greater than 80%. The institutional review boards (IRBs) of Marshfield Clinic Research Foundation, Kaiser Permanente Northern California, Kaiser Permanente South East, Georgia, and Vanderbilt approved this study and informed consent was obtained. In diabetic patients, metformin was administered for at least three months to achieve steady state. Patients were in the study for an average of 2.8 years (median = 1.43 years) with approximately 7.4 (median = 5) HbA1c samples obtained. The median metformin dose across the patient population was 1000 mg (Table 4.1). Patients were genotyped using an Illumina OmniExpress1.0 genotyping array except for the samples from Kaiser Permanente Northern California, which were genotyped using Illumina OmniExpressExome at the RIKEN institute in Japan. Genotype data quality control was done using the standard protocol that has been previously described in other genome-wide association (GWA) studies.

‘

**Table 4.1** Baseline characteristics of patients with type 2 diabetes

| **Clinical Site** | **N** |
|---|---|
| Total patients | 1056 |
| Kaiser South East Georgia | 154 |
| Marshfield Clinic | 150 |
| Vanderbilt | 251 |
| Kaiser Northern California | 501 |
| **Categorical Variable** | **N** |
| Males | 415 |
| Females | 641 |
| European Americans[1] | 376 |
| African Americans | 665 |
| Asian Americans and others | 15 |
| **Quantitative Variable** | **Median (range)** |
| Age (years) | 55 (23-90) |
| Body weight (kg) | 96 (34-212) |
| Average serum creatinine (mg/dL) | 0.91 (0.5-2.0) |
| Baseline HbA1c (%) | 7.6 (5.6-17.9) |
| Metformin daily dose (mg) | 1000 (200-2500) |
| # HbA1c samples/patient | 5 (1-45) |
| Years on study | 1.43 (0.28-13.5) |

[1]Ethnicities shown here are all self-reported.

**Development of mathematical model**

Patient data were analyzed using non-linear mixed effect modeling (NONMEM 7) with first order conditional estimation method with interaction (FOCE-I). Several semi-mechanistic approaches were explored to best describe the longitudinal HbA1c versus time profiles. Available PK information was taken into account in the model structure by investigating either surrogate exposure

(average serum creatinine level) or imputed exposure (based on individual metformin clearance values). Model selection was determined by using the objective function value (OFV, -2 times the log of the likelihood) and visual inspection of diagnostic plots. The selected longitudinal HbA1c profiles were described by the following equation:

$$(1)\ \frac{d(HbA1c)}{dt} = K_{IN}(t) * (1 - Metf_{EFFECT}) - K_{OUT} * HbA1c$$

where $Metf_{EFFECT}$, and $K_{out}$ represent metformin's effect from baseline and the loss rate of HbA1c, respectively. $K_{IN}(t)$ represents the synthesis rate of HbA1c: a non-linear, time sensitive parameter dependent on the baseline synthesis rate and the extent of patient-specific disease progression. Patients with higher disease progression estimates will result in a faster synthesis rate of HbA1c. The effect of disease progression estimates on $K_{IN}(t)$ and longitudinal HbA1c levels can be visualized in Figure 4.2.

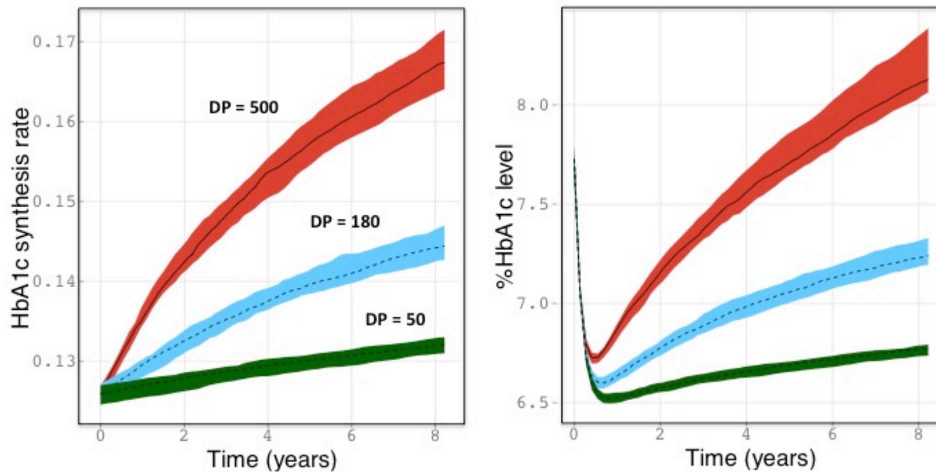



**Figure 4.2 Model mechanics and interplay of disease progression, HbA1c synthesis rate and %HbA1c level over time.**
The mathematical relationship between the disease progression parameter and HbA1c levels are shown here. DP = Model based disease progression parameter. HbA1c synthesis rate = %HbA1c level/day. The red, blue and green highlight a disease progression parameter estimate of 500,180, and 50, respectively. The intuitive representation of these estimates is shown on the right plot. Simulated median HbA1c levels are plotted as a function of time and varying disease progression estimates. Hypothetic values of DP are visualized here on the time course of HbA1c.

With this model structure, disease progression effectively captures the increased trajectory of HbA1c, which may result from the progression in diabetes or a reduction in the reversible effect of metformin. Between-subject variability (BSV) was estimated for baseline HbA1c, the magnitude of metformin's effect, and the disease progression parameter. An additive and proportional estimate was used to characterize the residual error model.

## Demographic analysis

Using the mathematical model described above, an agnostic stepwise forward selection ($P < 0.05$) and backward elimination ($P < 0.01$) were applied to identify statistically significant demographic and clinical covariates on model parameter estimates, which helped guide the selection of the demographic-corrected final model. The subsequent demographic-corrected mathematical model served as a basis to investigate the effect of genetic variants on the variance of long-term response.

## Selection of candidate genes

A comprehensive list of candidate genes was selected using the GWAS Integrator tool on the HuGE Navigator[23]. The browser selected genes that were previously identified from a GWA study conferring risk to one or more T2D traits. Additionally, genes in metformin's PK/PD pathway were also included, as highlighted in PharmGKB[24]. A final list consisting of 267 T2D-linked genes was selected for the genetic analysis. Genetic polymorphisms in a 50-kilobase region surrounding each gene were investigated on disease progression and other model parameters.

## Genetic analyses of model parameters

A penalized regression-based approach called hyperlasso was implemented to statistically prioritize the variants associated with phenotypes outputted from the mathematical model (e.g. disease progression, metformin

effect, and baseline). This methodology was originally proposed by Hoggart et al., and is a generalization of Lasso[25,26]. The penalty parameter used for this analysis reflects the number of SNPs tested as well as the total sample size. Using hyperlasso is computationally inexpensive and is therefore suitable for large-scale genomic studies. Hyperlasso also considers linkage disequilibrium patterns across the tested variants and ultimately produces a list of statistically associated variants with limited correlation. The hyperlasso analysis was performed with and without the inclusion of self reported ethnicity as a covariate in the model. An additional regression analysis, assuming an additive genetic model, was performed on the variants identified by hyperlasso. Variants with less than a 5% minor allele frequency (MAF) and missing genotype call rates greater than 10% were excluded from this analysis.

**Model based genetic analysis of identified variants**

The top SNPs from hyperlasso were subsequently investigated in the developed mathematical model described above. Model based analyses are advantageous because they account for correlations across various model parameters as well as potential SNP/SNP interactions. Two key steps were taken to select the final mathematical model; (1) removal of non-significant SNPs, which resulted from a univariate analysis of each variant in the demographic-adjusted mathematical model, and (2) removal of variants from the full genetic model that had very low, clinically irrelevant effect sizes. Goodness of fit plots and changes in the objective function value were evaluated to examine the

appropriateness of the final model with incorporated genetics. Simulations using this model were performed to predict the longitudinal impact of each SNP at various time points during therapy, as well as the impact of combinations of SNPs on long-term HbA1c levels.

### Functional annotation of top variants

Upon obtaining the final variants identified from the concluding mathematical model (resulting from the hyperlasso and model-based methodology), we then determined whether the variants have potential regulatory functions by searching the following databases: Regulome database and NCBI eQTL. The steps were as follows: first, identify the SNPs in linkage disequilibrium (LD, with $r^2 \geq 0.8$) to the 9 variants using HaploReg v2 and reference populations for European (EUR) and African American (AFR)[27]; next, we determined whether the top variants and SNPs in LD are located in a regulatory region using RegulomeDB (this database is used as a resource to guide interpretation of regulatory variants). In addition, we also used the NCBI eQTL database to determine whether the top SNPs are linked to gene expression levels of pharmacologically relevant genes.

## RESULTS

### Summary of data

Baseline characteristics of patients with T2D are summarized in Table 4.1. A total of 7822 HbA1c measurements from 1056 patients were used to develop a

mathematical model of longitudinal HbA1c levels. The average length of time that each patient was under study was 2.78 years (median of 1.43 years, range of 0.28-13.5 years). Mean HbA1c samples provided per patient available for analysis was 7.5 (median of 5, range of 1-45). Of the 1056 patients, HbA1c measurements were available for 202 patients 5 years following metformin initiation, 123 patients 7 years after metformin initiation, and 28 patients 10 years after.

**Mathematical model development**

A turnover HbA1c model with a reversible metformin effect on the synthesis rate of HbA1c best characterized the data. The upward trajectory (disease progression) of HbA1c over time was modeled by implementing a separate compartment that represented the HbA1c synthesis rate: $K_{IN}(t)$. In the model structure, the $K_{IN}$ kinetics were influenced by the disease progression parameter. The disease progression parameter generates a nonlinear increase of $K_{IN}$ over time, especially with high estimates of disease progression. A time dependent increase in the HbA1c synthesis rate captured the upward HbA1c patterns observed in the data. Between-subject variability (BSV) was estimated for baseline HbA1c, the magnitude of metformin's effect (an individual's maximum HbA1c relative change from baseline), and disease progression. The inclusion of a full covariance block between all BSV parameters resulted in a significant improvement in the likelihood ratio. The selection of the model was based on the objective function value and visual predictive checks of the

longitudinal HbA1c data. Through simulations, the model predicts that the onset of disease progression for patients on metformin is approximately 321 days; at this point, levels increased at a rate of 0.1 HbA1c level [0.07%-0.13%] per year through the first three years. For patients not on metformin, the model predicts that HbA1c levels would increase at a steady state rate of approximately 0.16 HbA1c level [0.08%-0.22%] per year. Mathematical model parameters along with clinically derived parameters are summarized in Table 4.2.

**Table 4.2** Population pharmacodynamic model derived estimates and bootstrap results for model parameters

| Final Model Parameter | Median (%RSE)[1] | Median (90% CI)[2] |
|---|---|---|
| Baseline HbA1c Level (%) | 7.74 (1) | 7.73 (7.6-7.8) |
| Half life of effect (days) | 40.9 (6) | 41.2 (36.8-45.7) |
| Metformin effect magnitude *EFF* | 13.1% (5) | 13.0 (12.1-14.4) |
| Disease progression estimate *DISPR*[3] (All patients) | 82.2 (67) | 75.3 (32.6-249) |
| Boxcox transformation parameter on Baseline | 2.38 (9) | 2.41 (1.99-2.78) |
| Boxcox transformation parameter on *DISPR* | -0.246 (15) | -0.26 (-0.31- -0.20) |
| **Between-subject variability (% variance)** | | |
| Between-subject variability (Baseline) | 16.9 (3) | 16.6 (15.9-17.8) |
| Between-subject variability (Metformin Effect Magnitude $METF_{EFF}$) | 76.4 (4) | 75.9 (71.7-81.6) |
| Between-subject variability (Disease Progression *DISPR*) | 324 (17) | 390 (164-418) |
| **Covariance of parameters (%)** | | |
| Correlation Baseline- $METF_{EFF}$ | 0.114 (1) | 0.11 (0.101-0.136) |
| Correlation Baseline-DISPR | 0.033 (3.6) | 0.03 (-0.07-0.14) |
| Correlation DISPR- $METF_{EFF}$ | 0.204 (21) | 0.31 (-0.42-0.95) |
| **Residual error model** | | |
| Proportional error (%) | 0.098 (3) | 0.098 (0.092-0.101) |
| Additive error | 0.1 (FIXED) | 0.1 (NA) |
| **Derived clinical parameters** | **Simulated median (90% CI)** | |
| Estimated onset of disease progression[4] | 321 (309-332) days | |
| Estimated yearly rate of HbA1c increase on Metformin[4] | 0.1 %HbA1c (0.07-0.13) | |
| Estimated yearly rate of HbA1c increase not on Metformin[4] | 0.16 %HbA1c (0.08-0.22) | |

[1]Typical value of parameter in final model. RSE= Relative standard error (%), also known as the precision of the parameter estimate.
[2]Confidence interval for the population pharmacodynamic parameter following bootstrap results.
Covariance of parameters are shown in untransformed format.

[3]DISPR is the disease progression model parameter that affects the synthesis rate of HbA1c and longitudinal HbA1c levels through the following equations. (1) $DADT(A1) = K_{ON}*(1+DISPR) - K_{LOSS}*A(1)$ and (2) $DADT(A2) = A(1)*(1-METF_{EFF}) - KOUT*A(2)$. Where A(1) represents the synthesis rate of HbA1c ($K_{SYN}$), and A(2) represents HbA1c levels.
[4]Yearly rate of HbA1c increase was based on simulated median yearly increase over the first three years following the onset of disease progression (i.e. 321 days). The median and 90%CI of the onset and yearly rate of HbA1c increase were calculated across simulations. For example, each simulation provided a median, which was then summarized across 1000 simulations.

**Final demographic/clinical covariate model**

A stepwise forward ($p < 0.05$) and backward ($p < 0.01$) unbiased selection process was performed to account for statistically significant predictors of baseline HbA1c, magnitude of metformin's effect, and disease progression. As determined by model diagnostics, the demographic-corrected mathematical model adequately described the data (Figure 4.3).

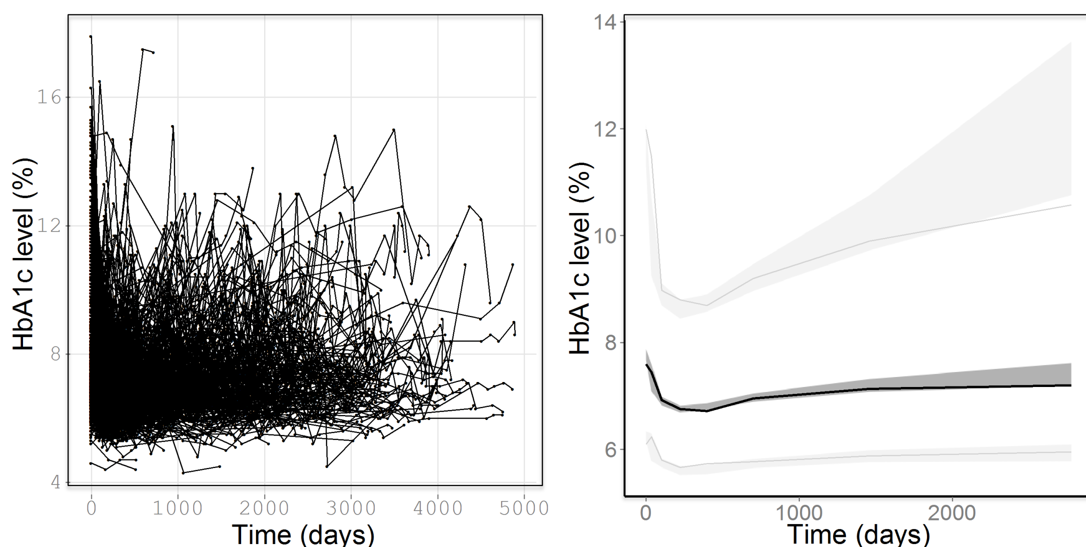



**Figure 4.3 Longitudinal HbA1c levels over time and model based visual predictive check.**
The plot to the left shows raw HbA1c observations over time. On the right plot, a visual predictive check is shown, where the solid black line highlights the median observed profiles. The shaded regions indicate the $95^{th}$ and $5^{th}$ percentiles (ends) and the range of median simulated profiles (center) of simulated predictions from the visual predictive check.

Intuitively, average serum creatinine (a key surrogate for metformin drug exposure) was a significant predictor on $Metf_{EFFECT}$, the model parameter that captured the magnitude of metformin's effect from baseline, with an estimated effect size of 3% change per 0.1 mg/dL change in serum creatinine level. Performing simulations results in a 0.77-0.96 change in %HbA1c level from baseline (at 2 years) for patients with average serum creatinine levels ranging from 0.6-1.3 mg/dL. This response characteristic is anticipated pharmacologically as the average exposure of metformin is expected to increase by approximately (18%-20%) between this range of serum creatinine values (0.6-1.3 mg/dL) for males and females of age 50.

Additionally, body weight and clinical site were significant covariates on the $Metf_{EFFECT}$ model parameter. Body weight was estimated to result in a 6% decrease in metformin's effect per 10-kilogram increase in body weight. This results in a 0.99% and 0.80% change in HbA1c level from baseline (at 2 years) for patients with body weights of 66 kg and 140 kg (5$^{th}$ and 95$^{th}$ percentile, respectively). For clinical site, Vanderbilt and Kaiser Georgia had a 16% and 30% lower estimate on the metformin effect parameter when compared to Kaiser Northern California, respectively.

Age was also a significant covariate on the disease progression parameter, with a negative correlation observed between age and disease progression. On average, the relative change in HbA1c at 2 years for patients between the ages of 49 and 64 years was predicted to fall between 0.76% and 0.84%, respectively.

**Genetic analysis: hyperlasso methodology on model parameters**

A total of 267 genes were selected and approximately 12000 variants within a 50-kilobase region around each gene were extracted for analysis. Of the variants investigated, a total of 21 SNPs were linked to the disease progression parameter. Of the 21 variants, 5 were eliminated due to minor allele frequencies that were less than 5%. Of the remaining 16 variants, 11 were intronic [*CSMD1*(4), *ADCY5*(1), *PRKAG*(1), *SLC22A2*(1), *EMILIN2*(1), *SULF1*(1), *FTO*(1), *WWOX*(1)], 1 was missense [*SREBF1*], and 4 were located within 50

kilobases upstream or downstream of each gene [*VPS13C*(1), *KCNK16*(1), *PPARG*(1), *FOXN3*(1)].

**Genetic analysis: model-based approach for variant selection**

Of the prioritized 16 variants from hyperlasso, a model-based methodology was implemented to further filter SNPs for inclusion in the final mathematical model. In doing this, 8 SNPs were removed due to allelic effects that did not fulfill the defined criteria (see Methods). The 9 remaining variants were statistically significant in the model structure and collectively accounted for approximately one-third of the variability in the disease progression model parameter. Of the 9 variants, rs12907856 (*VPS13C*), rs2954625 (*CSMD1*), and rs3160009 (*SLC22A2*) individually accounted for approximately 6%, 5%, and 8% of the variability, respectively. The characteristics of each SNP are shown in Table 4.3.

**Table 4.3** Summary of top genetic variants included in final mathematical model of metformin

| SNP | Chr | Position | Gene | Minor Allele | OFV_change | Estimate | MAF |
|---|---|---|---|---|---|---|---|
| rs12907856 | 15 | 59897770 | *VPS13C* | G | -15.99 | -0.79 | 0.30 |
| rs2617102 | 8 | 4451617 | *CSMD1* | C | -7.77 | 1.03 | 0.13 |
| rs2815022 | 6 | 39347243 | *KCNK16* | G | -9.32 | 0.99 | 0.42 |
| rs2954625 | 8 | 3757098 | *CSMD1* | T | -7.08 | 0.79 | 0.24 |
| rs316009 | 6 | 160595754 | *SLC22A2* | T | -5.15 | -0.85 | 0.08 |
| rs642887 | 18 | 2864408 | *EMILIN2* | A | -6.06 | -0.42 | 0.13 |
| rs6982250 | 8 | 70575004 | *SULF1* | T | -12.15 | -0.45 | 0.18 |
| rs7159552 | 14 | 89176069 | *FOXN3* | T | -9.42 | -0.61 | 0.31 |
| rs7500549 | 16 | 76800885 | *WWOX* | C | -7.07 | -0.75 | 0.59 |

OFV_change = Objective function value change from demographic corrected base model; MAF = minor allele frequency. Estimate = Model based effect size estimate of 2 minor alleles on the disease progression parameter.

In the final model, several simulations were performed to illustrate the potential clinical impact of each SNP on long-term HbA1c levels. Figure 4.4 quantitatively summarizes the predicted effects of final model genetic and non-genetic factors on HbA1c levels. Hypothetical gene/gene interactions were also explored and the combinatorial effects of high risk SNPs in the *CSMD1*, *WWOX*, and *SLC22A2* genes are shown in Figure 4.5.

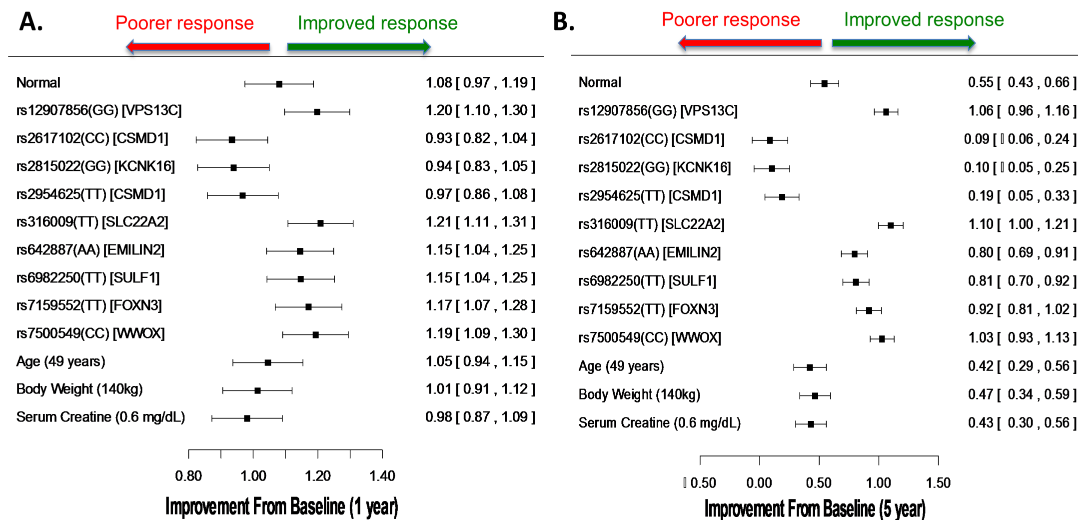



**Figure 4.4 Top genetic and demographic covariates on long term HbA1c levels.**

A. The effect of covariates on the simulated median (bands show 5$^{th}$ and 95$^{th}$ CI of simulated median) of HbA1c levels at the 1-year mark. B. The effect of covariates on the simulated median (5$^{th}$ and 95$^{th}$ CI of simulated median) of HbA1c levels at the 5-year mark. A normal individual here represents a hypothetical patient with no minor alleles of any of the identified variants with median age, body weight, and serum creatinine values.

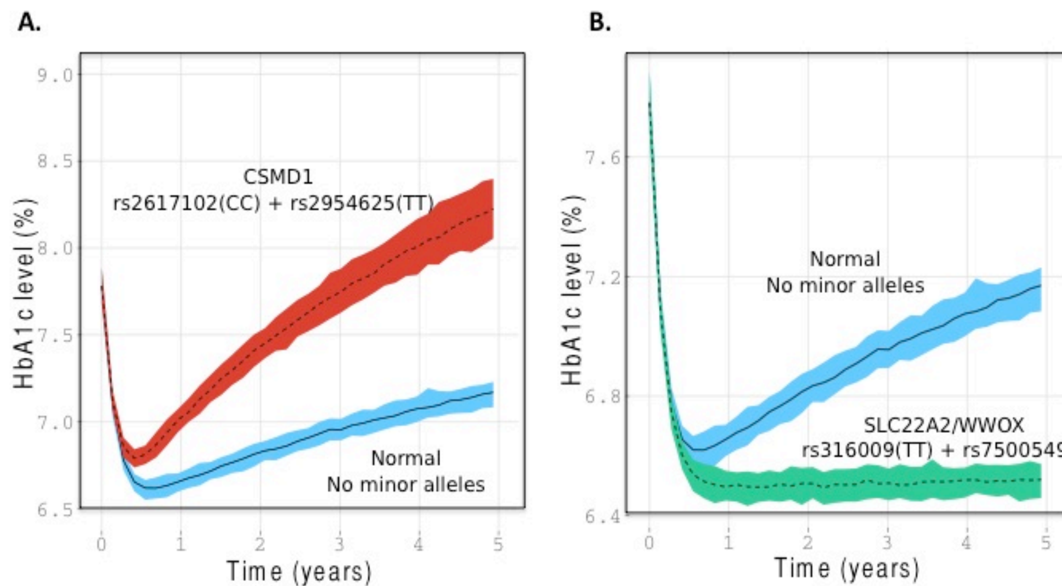



**Figure 4.5 Effect of SNP combinations in *CSMD1, SLC22A2, and WWOX* on the dynamics of HbA1c levels.**
A: Simulated median HbA1c levels (with 95% CI bands) over 5 years comparing carriers and non-carriers of CSMD1 minor (risk) alleles. B: Simulated median HbA1c levels over 5 years comparing carriers and non-carriers of *SLC22A1*/*WWOX* genes minor alleles. Blue shaded region with solid line: Simulated median for patients carrying no minor alleles with $5^{th}$ and $95^{th}$ confidence interval. Red/green shade with dashed line: Simulated median for patients carrying minor alleles of labeled gene(s) with $5^{th}$ and $95^{th}$ confidence interval of median.

In the explorative studies, patients carrying one or more minor alleles of the identified variants in the *CSMD1* gene (rs2617102 (C), rs2954625 (T)) were predicted to have significantly higher longitudinal HbA1c levels compared to patients not carrying any *CSMD1* minor alleles or patients with homozygous rs3160009 TT (*SLC22A2)* and/or homozygous rs7500549 CC (*WWOX*) genotypes.

### Functional annotation of top variants

Three out of the nine variants (rs12907856, rs316009, and rs7159552) are located in LD to a regulatory region based on an algorithmic prediction by RegulomeDB[28]. In particular, rs316009 and rs7159552 are located in a transcription factor binding motif as identified by the ENCODE project[29]. The rs316009 variant is in LD to the nonsynonymous variant of SLC22A2 - rs316019 - which is known to play a role in metformin pharmacokinetics[,30,31,32]. Another variant, rs6982250, is in an intronic region of *SULF1*. Several SNPs in SULF1 have been associated with many phenotypes in the GWAS Catalog[33] and one associated with fasting insulin-related traits[34].

## DISCUSSION

Previous pharmacogenetic studies of metformin response have focused on the effect of selected variants in relevant pharmacogenes on single-time point outcomes of metformin (i.e. HbA1c levels after 90 days, FPG levels, etc.)[15,35,36,22]. Long-term, time-dependent changes of HbA1c have been previously overlooked, resulting in a collapse of valuable information that may inform disease progression as well as temporal response patterns.

In this study, we developed a longitudinal HbA1c model by leveraging a large T2D dataset and subsequently investigating the role of genetic and non-genetic factors on the long-term dynamics of HbA1c following metformin initiation. Special focus was given to identifying the factors that are responsible for the long-term variance in HbA1c levels.

Three important findings emerged from this analysis: (i) a mathematical model incorporating disease progression and a reversible metformin effect best characterized the longitudinal HbA1c data in T2D Patients. (ii) The model presented herein predicted that the onset of disease progression for patients on metformin is approximately 321 days, at which point, levels to increase, on average, at a rate of 0.1% [0.04%-0.16%] HbA1c per year; HbA1c levels are expected to increase at a steady state rate of approximately 0.16% [0.08%-0.22%] per year in patients not being treated with metformin. (iii) Nine variants in 8 genes (of 267 genes interrogated) accounted for approximately one-third of the total estimated variability in the disease progression parameter. Variants in three of these genes (*CSMD1*, *WWOX*, and *SLC22A2*) were identified as significant influencers of disease progression on metformin therapy.

The development of the final mathematical model resulted from the exploration of several approaches with various empirical and semi-mechanistic considerations. The structural parameters from the model were estimated with high precision. The between-subject variability estimates of baseline, metformin effect, and disease progression were also estimated with relatively high precision (3%, 4%, and 17% relative standard error (RSE), respectively). The high degree of parameter confidence was due to the abundance of available HbA1c data, which ultimately allows for the reliable assessment of clinical, demographic, and genetic covariates on disease progression. Disease progression (upward trajectory of HbA1c levels) is a function of both the patient's underlying disease as well as the build up of metformin resistance. In order to differentiate between

the effects of a patient's biology and a reduction of metformin's reversible effect, it is necessary to model longitudinal HbA1c data prior to the administration of treatment; unfortunately, this was not possible in our analysis. The HbA1c model was however able to adequately predict the dynamics of HbA1c levels, capturing the long-term upward trend observed in this population. The ability to predict long-term HbA1c changes is especially valuable: the onset of disease progression and the rate of HbA1c increase were quantified for patients on metformin therapy (~0.1% increase per year for the first three years after 321 days) due to the richness of HbA1c data available. This finding was particularly interesting in relation to the study by Winter et al. where the authors noted a slight rise in patients' HbA1c levels during three preceding visits (between 200 and 400 days after metformin initiation); however, they were unable to quantify this upward trend through their simulations – a limitation which resulted from the lack of longitudinal data points available after 400 days[16]. In our analysis, the average length of time in the study was 1014 days and several HbA1c measurements were available to inform the progression of HbA1c up to 10 years so we were able to quantitate this trend with high precision. The robustness in the model also allowed for the simulation of patient-specific disease progression with the assumption of no drug on board (approximate increase of 0.16% in HbA1c per year). The ability to separate disease progression and metformin effect is based on early HbA1c data (up to 1 year following metformin initiation). The simulations of disease progression without drug on board were performed by removing metformin's estimated effect on the HbA1c synthesis rate within the

model structure. The simulations demonstrate that on average, the disease progression in patients who are metformin-naive will occur faster than in patients taking metformin for several months. Comparing this estimate to existing literature is problematic since T2D progression is a gradual process that typically takes place over several years and thus allows only a small trajectory of change within the limited time frame available for most studies. In the few studies reported, the rate of HbA1c increase was estimated to be approximately 0.2% per year, a value consistent with our observations[37].

A stepwise multivariate analysis was performed to identify statistically significant demographic and clinical covariates on model parameters. Average serum creatinine level surfaced as a significant factor that influenced the magnitude of metformin's effect. This finding was expected since average serum creatinine levels are considered a key surrogate for metformin exposure. Serum creatinine directly influences a patient's creatinine clearance, which ultimately influences the exposure to metformin by affecting the apparent clearance. The effect of age was also noted and an inverse relationship was observed between age and the magnitude of disease progression. It is important to note that although age was statistically significant through a stepwise analysis, the effect size was fairly small and a reproduction of these results is required for confirmation. Previously, in a study by Williams et al., lower HbA1c levels were reported in African Americans compared to European American individuals[38]. In our analysis, however, there was no significant effect of self-reported ethnicity on any of the model features, including the disease progression parameter.

We used a multi-pronged genetic approach to prioritize influential variants on disease progression. Hyperlasso methodology was selected over a stepwise procedure, as well as several other algorithms, because the hyperlasso approach has been shown to be robust when many covariates are correlated, which is the case here with strong LD patterns in the genotype data. The final selection of variants is based on the performance of individual variants within the model structure so that the correlation across various model parameters may be taken into consideration.

Nine of the variants investigated emerged as being potentially linked to the progression of HbA1c levels on metformin. Collectively, the variants accounted for approximately one-third of the variance in the disease progression model parameter. It was also observed that these genetic variants had larger effects on HbA1c levels than the demographic and clinical covariates identified from the stepwise analysis.

Of the top genes, minor alleles of two SNPs (rs2617102, rs2954625) in the *CSMD1* (CUB and Sushi multiple domains 1) gene had the strongest impact on disease progression. Although the pharmacological and biological mechanism remains unclear, *CSMD1* has been previously linked to insulin sensitivity and lipid levels[39,40]. From this analysis, the *CSMD1* variants may have a significant impact on longitudinal HbA1c levels, especially at the five-year mark when the simulated HbA1c improvement from baseline becomes quite low – especially for homozygous carriers (TT) of rs2617102. The simulated 5-year HbA1c level was very similar to baseline levels (Figure 4.3). Furthermore, the effect on HbA1c

levels at the 5-year mark is even more pronounced for hypothetical homozygous carriers of both *CSMD1* SNPs and HbA1c levels are predicted to be higher than baseline levels.

Minor alleles of SNPs in genes *SLC22A2*, *WWOX, EMILIN2,* and *FOXN3* were linked to more favorable trajectories (lower disease progression) of HbA1c levels compared to major allele carriers. Of these genes, *SLC22A2* and *WWOX* (rs316009 (T) and rs7500549 (C)) showed the strongest effect. In contrast to homozygous carriers of *CSMD1* risk alleles, homozygous carriers of both the *SLC22A2* and *WWOX* SNPs were predicted to have a favorable outcome and maintain their peak HbA1c level improvement from baseline through 5 years following metformin initiation. The rs316009 variant is in LD to a nonsynonymous variant of *SLC22A2* (rs316019), which is a SNP that has been previously shown to alter transporter function as well as modulate metformin pharmacokinetics.[31,32] Therefore, the clinical expectation that the reduced function rs316009 (T) allele would lead to a more favorable outcome is pharmacologically sound. OCT2 (*SLC22A2*) is predominantly expressed at the basolateral membrane in distal renal tubules and is responsible for the uptake of metformin from circulation into renal epithelial cells, working in concert with other renal transporters to excrete metformin. Though functional studies have been controversial[30], loss of transporter function is expected to increase plasma levels of metformin, potentially leading to a more favorable pharmacodynamic outcome and relatively lower HbA1c levels.

Also of clinical interest, the gene *WWOX* has been previously associated

with several T2D traits including body weight, C-reactive protein, insulin, obesity, and lipid levels[41]. *WWOX* encodes for an enzyme that is found in all eukaryotes and has been biologically shown to play an important role in the regulation of a wide variety of cellular functions such as protein degradation, transcription, and RNA splicing. Unlike *SLC22A2,* a pharmacological mechanism for *WWOX* is not clear. However, the clinical impact (if replicated) would mean that carriers of the rs7500549 (C) allele would respond more favorably to metformin therapy. Future studies should focus on elucidating the biology of *WWOX* and replicating the genetic findings on disease progression.

Overall, our study has successfully integrated powerful model-based approaches with genetic analyses to uncover genes linked to the progression of HbA1c on metformin therapy in a large T2D cohort. If replicated, these genetic findings may have a significant influence on T2D treatment strategy. Ultimately, the long-term goal of this research is to translate this refined mathematical model into clinical practice and enable clinicians to provide data-driven, personalized treatment advice to T2D patients based on genetics, demographics, and real-time HbA1c measurements.

**REFERENCES**


1. Scarpello, J. H. & Howlett, H. C. Metformin therapy and clinical uses. *Diab. Vasc. Dis. Res.* **5,** 157–67 (2008).

2. Gong, L., Goswami, S., Giacomini, K. M., Altman, R. B. & Klein, T. Metformin pathways: pharmacokinetics and pharmacodynamics. *Pharmacogenet Genomics* **22,** 820–27 (2013).

3. Viollet, B., Guigas, B., Sanz Garcia, N., Leclerc, J., Foretz, M. & Andreelli, F. Cellular and molecular mechanisms of metformin: an overview. *Clin. Sci. (Lond).* **122,** 253–70 (2012).

4. Graham, G. G., Punt, J., Arora, M., Day, R. O., Doogue, M. P., Duong, J. K., Furlong, T. J., Greenfield, J. R., Greenup, L. C., Kirkpatrick, C. M., Ray, J. E., Timmins, P. & Williams, K. M. Clinical pharmacokinetics of metformin. *Clin. Pharmacokinet.* **50,** 81–98 (2011).

5. Hamrén, B., Björk, E., Sunzel, M. & Karlsson, M. Models for plasma glucose, HbA1c, and hemoglobin interrelationships in patients with type 2 diabetes following tesaglitazar treatment. *Clin. Pharmacol. Ther.* **84,** 228–35 (2008).

6. Menon, V., Greene, T., Pereira, A. A., Wang, X., Beck, G. J., Kusek, J. W., Collins, A. J., Levey, A. S. & Sarnak, M. J. Glycosylated hemoglobin and mortality in patients with nondiabetic chronic kidney disease. *J. Am. Soc. Nephrol.* **16,** 3411–27 (2005).

7. Zoungas, S., Chalmers, J., Ninomiya, T., Li, Q., Cooper, M. E., Colagiuri, S., Fulcher, G., de Galan, B. E., Harrap, S., Hamet, P., Heller, S., MacMahon, S., Marre, M., Poulter, N., Travert, F., Patel, A., Neal, B. & Woodward, M. Association of HbA1c levels with vascular complications and death in patients with type 2 diabetes: evidence of glycaemic thresholds. *Diabetologia* **55,** 636–43 (2012).

8. Andersson, C., van Gaal, L., Caterson, I. D., Weeke, P., James, W. P. T., Coutinho, W., Couthino, W., Finer, N., Sharma, A. M., Maggioni, A. P. & Torp-Pedersen, C. Relationship between HbA1c levels and risk of cardiovascular adverse outcomes and all-cause mortality in overweight and obese cardiovascular high-risk women and men with type 2 diabetes. *Diabetologia* **55,** 2348–55 (2012).

9. Cook, M. N., Girman, C. J., Stein, P. P. & Alexander, C. M. Initial monotherapy with either metformin or sulphonylureas often fails to achieve or maintain current glycaemic goals in patients with type 2 diabetes in UK primary care. *Diabet. Med.* **24,** 350–58 (2007).

10. Goswami, S., Yee, S. W., Stocker, S., Mosley, J. D., Kubo, M., Castro, R., Mefford, J., Wen, C., Liang, X., Witte, J., Brett, C., Maeda, S., Simpson, M. D., Hedderson, M. M., Davis, R. L., Roden, D. M., Giacomini, K. M. & Savic, R. M. Genetic variants in transcription factors are associated with the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **96,** 370–90 (2014).

11. Shu, Y., Leabman, M. K., Feng, B., Mangravite, L. M., Huang, C. C., Stryke, D., Kawamoto, M., Johns, S. J., DeYoung, J., Carlson, E., Ferrin, T. E., Herskowitz, I. & Giacomini, K. M. Evolutionary conservation predicts function of variants of the human organic cation transporter, OCT1. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 5902–7 (2003).

12. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 ( OCT1 ) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

13. Stocker, S. L., Morrissey, K. M., Yee, S. W., Castro, R., Xu, L., Dahlin, A., Ramirez, H., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L., Brett, C. M. & Giacomini, K. M. The effect of novel promoter variants in MATE1 and MATE2 on the pharmacokinetics and pharmacodynamics of metformin. *Clin. Pharmacol. Ther.* **93,** 186–94 (2013).

14. Christensen, M. M. H., Brasch-Andersen, C., Green, H., Nielsen, F., Damkier, P., Beck-Nielsen, H. & Brosen, K. The pharmacogenetics of metformin and its impact on plasma metformin steady-state levels and glycosylated hemoglobin A1c. *Pharmacogenet. Genomics* **21,** 837–50 (2011).

15. Zhou, K., Bellenguez, C., Spencer, C. C., Bennett, A. J., Coleman, R. L., Tavendale, R., Hawley, S., Donnelly, L., Schofield, C., Groves, C. J., Burch, L., Carr, F., Strange, A., Freeman, C., Blackwell, J. M., Bramon, E., Brown, M., Casas, J. P., Corvin, A., Craddock, N., Deloukas, P., Dronov, S., Duncanson, A., Edkins, S., Gray, E., Hunt, S., Jankowski, J., Langford, C., Markus, H. S., Mathew, C. G., Plomin, R., Rautanen, A., Sawcer, S. J., Samani, N. J., Trembath, R., Viswanathan, A. C., Wood, N. W., Harries, L. W., Hattersley, A. T., Doney, A. S. F., Colhoun, H., Morris, A. D., Sutherland, C., Hardie, D. G., Peltonen, L., McCarthy, M. I., Holman, R. R., Palmer, C. N., Donnelly, P. & Pearson, E. R. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat. Genet.* **43,** 117–20 (2011).

16. De Winter, W., DeJongh, J., Post, T., Ploeger, B., Urquhart, R., Moules, I., Eckland, D. & Danhof, M. A mechanism-based disease progression model for comparison of long-term effects of pioglitazone, metformin and gliclazide on disease processes underlying type 2 diabetes mellitus. *J. Pharmacokinet. Pharmacodyn.* **33,** 313–43 (2006).

17. Ette, E. I. & Williams, P. J. Population pharmacokinetics I: background, concepts, and models. *Ann. Pharmacother.* **38,** 1702–16 (2004).

18. Ette, E. I., Williams, P. J. & Lane, J. R. Population pharmacokinetics III: design, analysis, and application of population pharmacokinetic studies. *Ann. Pharmacother.* **38,** 2136–44 (2004).

19. Pillai, G. C., Mentré, F. & Steimer, J.-L. Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science. *J. Pharmacokinet. Pharmacodyn.* **32,** 161–83 (2005).

20. Hong, Y., Rohatagi, S., Habtemariam, B., Walker, J. R., Schwartz, S. L. & Mager, D. E. Population exposure-response modeling of metformin in patients with type 2 diabetes mellitus. *J. Clin. Pharmacol.* **48,** 696–707 (2008).

21. Møller, J. B., Kristensen, N. R., Klim, S., Karlsson, M. O., Ingwersen, S. H. & Kjellsson, M. C. Methods for predicting diabetes phase III efficacy outcome from early data: superior performance obtained using longitudinal approaches. *CPT pharmacometrics Syst. Pharmacol.* **3,** 122–30 (2014).

22. Choi, J. H., Yee, S. W., Ramirez, H., Morrissey, K. M., Jang, G. H., Joski, P. J., Mefford, J., Hesselson, S. E., Schlessinger, A., Jenkins, G., Castro, R., Johns, S. J., Stryke, D., Sali, A., Ferrin, T. E., Witte, J. S., Kwok, P.-Y., Roden, D. M., Wilke, R., McCarty, C., Davis, R. L. & Giacomini, K. M. A common 5'-UTR variant in MATE2-K is associated with poor response to metformin. *Clin. Pharmacol. Ther.* **90,** 674–84 (2011).

23. Yu, W., Clyne, M., Khoury, M. J. & Gwinn, M. Phenopedia and Genopedia: disease-centered and gene-centered views of the evolving knowledge of human genetic associations. *Bioinformatics* **26,** 145–56 (2010).

24. Whirl-Carrillo, M., McDonagh, E. M., Hebert, J. M., Gong, L., Sangkuhl, K., Thorn, C. F., Altman, R. B. & Klein, T. E. Pharmacogenomics knowledge for personalized medicine. *Clin. Pharmacol. Ther.* **92,** 414–27 (2012).

25. Hoggart, C. J., Whittaker, J. C., De Iorio, M. & Balding, D. J. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. *PLoS Genet.* **4,** 1–8 (2008).

26. Bertrand, J. & Balding, D. J. Multiple single nucleotide polymorphism analysis using penalized regression in nonlinear mixed-effect pharmacokinetic models. *Pharmacogenet. Genomics* **23,** 167–74 (2013).

27. Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* **40,** 930–34 (2012).

28. 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. & Snyder, M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22,** 1790–97 (2012).

29. Myers, R. M., Stamatoyannopoulos, J., Snyder, M., Dunham, I., Hardison, R. C., Bernstein, B. E., Gingeras, T. R., Kent, W. J., Birney, E., Wold, B., Crawford, G. E., Bernstein, B. E., Epstein, C. B., Shoresh, N., Ernst, J., Mikkelsen, T. S., Kheradpour, P., Zhang, X., Wang, L., Issner, R., Coyne, M. J., Durham, T., Ku, M., Truong, T., Ward, L. D., Altshuler, R. C., Lin, M. F., Kellis, M., Gingeras, T. R., Davis, C. a., Kapranov, P., Dobin, A., Zaleski, C., Schlesinger, F., Batut, P., Chakraborty, S., Jha, S., Lin, W., Drenkow, J., Wang, H., Bell, K., Gao, H., Bell, I., Dumais, E., Dumais, J., Antonarakis, S. E., Ucla, C., Borel, C., Guigo, R., Djebali, S., Lagarde, J., Kingswood, C., Ribeca, P., Sammeth, M., Alioto, T., Merkel, A., Tilgner, H., Carninci, P., Hayashizaki, Y., Lassmann, T., Takahashi, H., Abdelhamid, R. F., Hannon, G., Fejes, K. T., Preall, J., Gordon, A., Sotirova, V., Reymond, A., Howald, C., Graison, E. A. Y., Chrast, J., Ruan, Y., Ruan, X., Shahab, A., Poh, W. T., Wei, C. L., Crawford, G. E., Furey, T. S., Boyle, A. P., Sheffield, N. C., Song, L., Shibata, Y., Vales, T., Winter, D., Zhang, Z., London, D., Wang, T., Birney, E., Keefe, D., Iyer, V. R., Lee, B. K., McDaniell, R. M., Liu, Z., Battenhouse, A., Bhinge, A., Lieb, J. D., Grasfeder, L. L., Showers, K., Giresi, P. G., Kim, S. K. C., Shestak, C., Myers, R. M., Pauli, F., Reddy, T. E., Gertz, J., Partridge, E. C., Jain, P., Sprouse, R. O., Bansal, A., Pusey, B., Muratet, M., Varley, K. E., Bowling, K. M., Newberry, K. M., Nesmith, A. S., Dilocker, J., Parker, S. L., Waite, L. L., Thibeault, K., Roberts, K., Absher, D. M., Wold, B., Mortazavi, A., Williams, B., Marinov, G., Trout, D., Pepke, S., King, B., McCue, K., Kirilusha, A., DeSalvo, G., Fisher, K. A., Amrhein, H., Vielmetter, J., Sherlock, G., Sidow, A., Batzoglou, S., Rauch, R., Kundaje, A., Libbrecht, M., Margulies, E. H., Parker, S. C. J., Elnitski, L., Green, E. D., Hubbard, T., Harrow, J., Searle, S., Kokocinski, F., Aken, B., Frankish, A., Hunt, T., Despacio-Reyes, G., Kay, M., Mukherjee, G., Bignell, A., Saunders, G., Boychenko, V., Brent, M., van Baren, M. J., Brown, R. H., Gerstein, M., Khurana, E., Balasubramanian, S., Zhang, Z., Lam, H., Cayting, P., Robilotto, R., Lu, Z., Guigo, R., Derrien, T., Tanzer, A., Knowles, D. G., Mariotti, M., Kent, W. J., Haussler, D., Harte, R., Diekhans, M., Kellis, M., Lin, M., Kheradpour, P., Ernst, J., Reymond, A., Howald, C., Graison, E. A. Y., Chrast, J., Valencia, A., Tress, M., Rodriguez, J. M., Snyder, M., Landt, S. G., Raha, D., Shi, M., Euskirchen, G., Grubert, F., Kasowski, M., Lian, J., Cayting, P., Lacroute, P., Xu, Y., Monahan, H., Patacsil, D., Slifer, T., Yang, X., Charos, A., Reed, B., Wu, L., Auerbach, R. K., Habegger, L., Hariharan, M., Rozowsky, J., Abyzov, A., Weissman, S. M., Gerstein, M., Struhl, K., Lamarre, N. V., Lindahl, M. A., Miotto, B., Moqtaderi, Z., Fleming, J. D., Newburger, P., Farnham, P. J., Frietze, S., O'Geen, H., Xu, X., Blahnik, K. R., Cao, A. R., Iyengar, S., Kaul, R., Thurman, R. E., Wang, H., Navas, P., Sandstrom, R., Sabo, P. J., Weaver,

M., Canfield, T., Lee, K., Neph, S., Roach, V., Reynolds, A., Johnson, A., Rynes, E., Giste, E., Vong, S., Neri, J., Frum, T., Johnson, E. M., Nguyen, E. D., Ebersol, A. K., Sanchez, M. E., Sheffer, H. H., Lotakis, D., Haugen, E., Humbert, R., Kutyavin, T., Shafer, T., Dekker, J., Lajoie, B. R., Sanyal, A., Kent, W. J., Rosenbloom, K. R., Dreszer, T. R., Raney, B. J., Barber, G. P., Meyer, L. R., Sloan, C., Malladi, V. S., Cline, M. S., Learned, K., Swing, V. K., Zweig, A. S., Rhead, B., Fujita, P., Roskin, K., Karolchik, D., Kuhn, R. M., Haussler, D., Birney, E., Dunham, I., Wilder, S. P., Keefe, D., Sobral, D., Herrero, J., Beal, K., Lukk, M., Brazma, A., Vaquerizas, J. M., Luscombe, N. M., Bickel, P. J., Boley, N., Brown, J. B., Li, Q., Huang, H., Gerstein, M., Habegger, L., Sboner, A., Rozowsky, J., Auerbach, R. K., Yip, K. Y., Cheng, C., Yan, K. K., Bhardwaj, N., Wang, J., Lochovsky, L., Jee, J., Gibson, T., Leng, J., Du, J., Hardison, R. C., Harris, R. S., Song, G., Miller, W., Haussler, D., Roskin, K., Suh, B., Wang, T., Paten, B., Noble, W. S., Hoffman, M. M., Buske, O. J., Weng, Z., Dong, X., Wang, J., Xi, H., Tenenbaum, S., Doyle, F., Penalva, L. O., Chittur, S., Tullius, T. D., Parker, S. C. J., White, K. P., Karmakar, S., Victorsen, A., Jameel, N., Bild, N., Grossman, R. L., Snyder, M., Landt, S. G., Yang, X., Patacsil, D., Slifer, T., Dekker, J., Lajoie, B. R., Sanyal, A., Weng, Z., Whitfield, T. W., Wang, J., Collins, P. J., Trinklein, N. D., Partridge, E. C., Myers, R. M., Giddings, M. C., Chen, X., Khatun, J., Maier, C., Yu, Y., Gunawardena, H., Risk, B., Feingold, E., Lowdon, R. F., Dillon, L. A. & Good, P. J. A user's guide to the Encyclopedia of DNA elements (ENCODE). *PLoS Biol.* **9,** 245–99 (2011).

30. Chen, Y., Li, S., Brown, C., Cheatham, S., Castro, R. A., Leabman, M. K., Urban, T. J., Chen, L., Yee, S. W., Choi, J. H., Huang, Y., Brett, C. M., Burchard, E. G. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 2 on the renal elimination of metformin. *Pharmacogenet. Genomics* **19,** 497–504 (2009).

31. Wang, Z.-J., Yin, O. Q. P., Tomlinson, B. & Chow, M. S. S. OCT2 polymorphisms and in-vivo renal functional consequence: studies with metformin and cimetidine. *Pharmacogenet. Genomics* **18,** 637–45 (2008).

32. Christensen, M. M. H., Pedersen, R. S., Stage, T. B., Brasch-andersen, C., Nielsen, F., Damkier, P., Beck-nielsen, H. & Brøsen, K. A gene – gene interaction between polymorphisms in the OCT2 and MATE1 genes influences the renal clearance of metformin. *Pharmacogenet. Genomics* **23,** 526–34 (2013).

33. Welter, D., MacArthur, J., Morales, J., Burdett, T., Hall, P., Junkins, H., Klemm, A., Flicek, P., Manolio, T., Hindorff, L. & Parkinson, H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42,** 1001–6 (2014).

34. Manning, A. K., Hivert, M.-F., Scott, R. A., Grimsby, J. L., Bouatia-Naji, N., Chen, H., Rybin, D., Liu, C.-T., Bielak, L. F., Prokopenko, I., Amin, N., Barnes, D., Cadby, G., Hottenga, J.-J., Ingelsson, E., Jackson, A. U., Johnson, T., Kanoni, S., Ladenvall, C., Lagou, V., Lahti, J., Lecoeur, C., Liu, Y., Martinez-Larrad, M. T., Montasser, M. E., Navarro, P., Perry, J. R. B., Rasmussen-Torvik, L. J., Salo, P., Sattar, N., Shungin, D., Strawbridge, R. J., Tanaka, T., van Duijn, C. M., An, P., de Andrade, M., Andrews, J. S., Aspelund, T., Atalay, M., Aulchenko, Y., Balkau, B., Bandinelli, S., Beckmann, J. S., Beilby, J. P., Bellis, C., Bergman, R. N., Blangero, J., Boban, M., Boehnke, M., Boerwinkle, E., Bonnycastle, L. L., Boomsma, D. I., Borecki, I. B., Böttcher, Y., Bouchard, C., Brunner, E., Budimir, D., Campbell, H., Carlson, O., Chines, P. S., Clarke, R., Collins, F. S., Corbatón-Anchuelo, A., Couper, D., de Faire, U., Dedoussis, G. V, Deloukas, P., Dimitriou, M., Egan, J. M., Eiriksdottir, G., Erdos, M. R., Eriksson, J. G., Eury, E., Ferrucci, L., Ford, I., Forouhi, N. G., Fox, C. S., Franzosi, M. G., Franks, P. W., Frayling, T. M., Froguel, P., Galan, P., de Geus, E., Gigante, B., Glazer, N. L., Goel, A., Groop, L., Gudnason, V., Hallmans, G., Hamsten, A., Hansson, O., Harris, T. B., Hayward, C., Heath, S., Hercberg, S., Hicks, A. A., Hingorani, A., Hofman, A., Hui, J., Hung, J., Jarvelin, M.-R., Jhun, M. A., Johnson, P. C. D., Jukema, J. W., Jula, A., Kao, W. H., Kaprio, J., Kardia, S. L. R., Keinanen-Kiukaanniemi, S., Kivimaki, M., Kolcic, I., Kovacs, P., Kumari, M., Kuusisto, J., Kyvik, K. O., Laakso, M., Lakka, T., Lannfelt, L., Lathrop, G. M., Launer, L. J., Leander, K., Li, G., Lind, L., Lindstrom, J., Lobbens, S., Loos, R. J. F., Luan, J., Lyssenko, V., Mägi, R., Magnusson, P. K. E., Marmot, M., Meneton, P., Mohlke, K. L., Mooser, V., Morken, M. A., Miljkovic, I., Narisu, N., O'Connell, J., Ong, K. K., Oostra, B. A., Palmer, L. J., Palotie, A., Pankow, J. S., Peden, J. F., Pedersen, N. L., Pehlic, M., Peltonen, L., Penninx, B., Pericic, M., Perola, M., Perusse, L., Peyser, P. A., Polasek, O., Pramstaller, P. P., Province, M. A., Räikkönen, K., Rauramaa, R., Rehnberg, E., Rice, K., Rotter, J. I., Rudan, I., Ruokonen, A., Saaristo, T., Sabater-Lleal, M., Salomaa, V., Savage, D. B., Saxena, R., Schwarz, P., Seedorf, U., Sennblad, B., Serrano-Rios, M., Shuldiner, A. R., Sijbrands, E. J. G., Siscovick, D. S., Smit, J. H., Small, K. S., Smith, N. L., Smith, A. V., Stančáková, A., Stirrups, K., Stumvoll, M., Sun, Y. V, Swift, A. J., Tönjes, A., Tuomilehto, J., Trompet, S., Uitterlinden, A. G., Uusitupa, M., Vikström, M., Vitart, V., Vohl, M.-C., Voight, B. F., Vollenweider, P., Waeber, G., Waterworth, D. M., Watkins, H., Wheeler, E., Widen, E., Wild, S. H., Willems, S. M., Willemsen, G., Wilson, J. F., Witteman, J. C. M., Wright, A. F., Yaghootkar, H., Zelenika, D., Zemunik, T., Zgaga, L., Wareham, N. J., McCarthy, M. I., Barroso, I., Watanabe, R. M., Florez, J. C., Dupuis, J., Meigs, J. B. & Langenberg, C. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **44,** 659–69 (2012).

35. Florez, J. C., Barrett-Connor, E., Jablonski, K., Knowler, W. C., Taylor, A., Shuldiner, A. R., Mather, K., Pollin, T. I., Horton, E. & White, N. H. The C allele of ATM rs11212617 does not associate with metformin response in the diabetes prevention program. *Diabetes Care* **35,** 1864–67 (2012).

36. Tzvetkov, M. V, Saadatmand, A. R., Bokelmann, K., Meineke, I., Kaiser, R. & Brockmöller, J. Effects of OCT1 polymorphisms on the cellular uptake, plasma concentrations and efficacy of the 5-HT3 antagonists tropisetron and ondansetron. *Pharmacogenomics J.* **12,** 22–29 (2012).

37. Turner, R. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet* **352,** 854–65 (1998).

38. Keoki Williams, L., Padhukasahasram, B., Ahmedani, B. K., Peterson, E. L., Wells, K. E., González Burchard, E. & Lanfear, D. E. Differing effects of metformin on glycemic control by race-ethnicity. *J. Clin. Endocrinol. Metab.* 39–55 (2014).

39. Irvin, M. R., Wineinger, N. E., Rice, T. K., Pajewski, N. M., Kabagambe, E. K., Gu, C. C., Pankow, J., North, K. E., Wilk, J. B., Freedman, B. I., Franceschini, N., Broeckel, U., Tiwari, H. K. & Arnett, D. K. Genome-wide detection of allele specific copy number variation associated with insulin resistance in african americans from the hyperGEN study. *PLoS One* **6,** 9–11 (2011).

40. Hong, K.-W., Go, M. J., Jin, H.-S., Lim, J.-E., Lee, J.-Y., Han, B. G., Hwang, S.-Y., Lee, S.-H., Park, H. K., Cho, Y. S. & Oh, B. Genetic variations in ATP2B1, CSK, ARSG and CSMD1 loci are related to blood pressure and/or hypertension in two Korean cohorts. *J. Hum. Hypertens.* **24,** 367–72 (2010).

41. Aldaz, C. M., Ferguson, B. W. & Abba, M. C. WWOX at the crossroads of cancer, metabolic syndrome related traits and CNS pathologies. *Biochim. Biophys. Acta* **1846,** 188–200 (2014).

# Chapter 5

## Summary and Conclusions

Despite recent progress in determining the role of various demographic, clinical and genetic factors in metformin pharmacology, such factors only account for a small fraction of the overall variation (i.e., in drug exposure, HbA1c levels, etc.). Metformin is not metabolized and elimination is primarily mediated by active tubular secretion from the kidney. Membrane transporters expressed in various tissues are therefore expected to significantly impact metformin pharmacokinetics and response; numerous studies have been conducted over the years to investigate this concept. Previous research has shown that individuals carrying any of the reduced function OCT1 alleles demonstrate a higher area under the concentration–time curve (AUC), a higher maximum plasma concentration ($C_{max}$), and an impaired glucose response compared to individuals carrying wild type alleles[1]. Genetic variants of *OCT2* (c.596C>T, c602C>T, and c.808G>T (rs316019)) were associated with differences in pharmacokinetics[2]. Although these studies have demonstrated significant associations with SNPs in transporters with metformin PK/PD, some of the data have not been replicated and each SNP only accounts for a small fraction of the variation in pharmacological phenotypes. This is not surprising considering metformin disposition is governed by a system of transporters rather than a single transporter. Furthermore, the pharmacological parameters (e.g., HbA1c levels) typically investigated are single-time point biomarkers that effectively

collapse critical information that may inform an individual's response patterns. Novel approaches with biological and pharmacological considerations are required to explain a larger portion of this variability in metformin pharmacology and response.

One such approach is to study gene expression modulators (transcription factors) that regulate the expression of multiple transporters involved in metformin disposition and elimination. This mechanistic concept is intuitive in that a loss or gain of transcription factor function may mediate the expression of many critical transporters that dictate the pharmacokinetics of the drug; therefore a single variant in a transcription factor may have a subsequent downstream clinical impact on metformin pharmacology. As a result, genetic polymorphisms in these genes should influence this system with potentially stronger effect sizes than transporter variants on pharmacological phenotypes.

Furthermore, studying the role of eQTLs on transcriptional regulation may lead to a deeper biological understanding of this mechanism. Genetic loci that mediate expression levels of crucial genes by altering either the expression or structure of a transcription factor - or by disrupting a DNA consensus motif that harbors preferential transcription factor binding - may provide a clear link between the underlying DNA mechanism and pharmacological outcome.

Finally, the role of eQTLs, transporter variants, disease-based genetic variants, and demographic factors has not been assessed on long-term patient HbA1c changes. Novel approaches to explore long-term response variation of HbA1c will make detection of responders and non-responders to metformin

easier[3]. The overall goal of this dissertation was to combine model-based approaches with various genetic analyses techniques to describe, quantify, and further our understanding of the mechanism that governs the pleiotropic nature of metformin pharmacokinetics and pharmacodynamics. Below is a chapter-by-chapter summary of the key findings and remaining challenges to be addressed in future studies.

**Chapter 2**

Studies described in chapter 2 focused on the roles of transcription factor variants in metformin pharmacokinetics and pharmacodynamics. As described, transcription factors may be considered as key nodes that modulate a system of transporters, leading to more pronounced changes in pharmacokinetics and pharmacodynamics. A list of transcription factor genes were prioritized accordingly based on the literature findings, which suggest that they regulate one or more metformin transporters. Genetic variants proximal to these transcription factors were then explored in T2D patients with genomic and HbA1c information. Variants found to be highly associated with relative changes in HbA1c levels were investigated for a pharmacokinetic mechanism using two approaches: first, an association with metformin secretory clearance in healthy volunteers and second, a pharmacokinetic model was developed in order to investigate the role of prioritized genetic variants on the overall exposure of metformin in the context of demographic variables. From this analysis, five variants in SP1 (a transcription factor that modulates the expression of *SLC47A1* and *SLC22A3*) were

significantly associated with changes in treatment HbA1c ($p < 0.01$) and metformin secretory clearance ($p < 0.05$). Population pharmacokinetic modeling further confirmed a 24% reduction in apparent clearance (increase in exposure) in homozygous carriers of one such variant, rs784888. Genetic variants in transcription factors *PPAR-alpha* and *HNF4-alpha* were significantly associated with HbA1c change only, but were not significantly associated with pharmacokinetics. This indicates that *PPAR-alpha* and *HNF4-alpha* transcription factors may be important in the pharmacodynamics of metformin and independent of a pharmacokinetic mechanism.

**Chapter 3**

In Chapter 3, the focus was shifted from exploring the clinical effects of transcription factor variants to investigating a plausible biological mechanism by which genetic variants may affect the pharmacological variation of metformin. The goal here was to profile SNPs linked to gene expression levels (eQTLs) of metformin transporter genes by performing a regression based analysis on profiled kidney tissue samples and subsequently investigating the clinical impact of the identified eQTLs in healthy volunteers with pharmacokinetic data.

From the *cis*-eQTL based analysis, a total of 6 variants were significantly associated with the expression levels of *SLC22A2*, *SLC22A3*, and *SLC47A1*. Three *cis* regions less than 50 kb from *SLC47A1* computationally demonstrated differential binding of transcription factors (e.g., Sox-5, GATA-1, SP1, and TATA). The variant rs2247436, which is located in an intron of the *SLC47A1* gene, was

predicted to be in a hot spot of regulatory activity, as multiple transcription factors (such as *SP1*, *FOXA1*, and *FOXA2*) have a strong potential to bind to this particular DNA region[4]. Of these transcription factors, SP1 is of particular interest because this transcription factor was previously shown to modulate expression levels of MATE1[5]. Furthermore, genetic variants in the SP1 locus were shown in chapter 2 to have a pharmacological impact on metformin pharmacokinetics and pharmacodynamics.

The *trans*-eQTL analysis highlighted the potential importance of *HNF4-alpha* on regulating the expression levels of *SLC22A1* and *SLC47A1*. A total of 2 *HNF4-alpha* variants were significantly associated with the expression level of *SLC22A1* and 3 variants were associated with the expression level of *SLC47A1*. The T allele of one such variant, rs2093248, was significantly associated with a lower metformin secretory clearance in an additive manner. This directionality of the minor allele may be mechanistically and pharmacologically intuitive since the T allele was biologically linked to lower expression levels of *SLC22A1*. Although the expression of *SLC22A1* on the proximal tubule of the kidney has not been confirmed, the analysis shows reasonable levels of this transporter in the kidney. Furthermore, the *trans*-eQTL based analysis provided a link between *HNF4-alpha* and *SLC47A1*, a previously well-profiled transporter in the kidney. *HNF4-alpha* is a multifunctional transcription factor primarily expressed in the liver but also expressed in the kidney[6]. Mutations in *HNF4-alpha* have also been linked to T2D; therefore, the findings described in this chapter align with prior knowledge of this transcription factor's clinical role[7]. Finally, results from chapter 3 support

findings from chapter 2, which suggest that SNPs in *HNF4-alpha* are linked to metformin pharmacodynamics. In summary, kidney specific eQTLs of metformin transporters is demonstrated as a novel method to highlight important genomic regions and transcription factors regulating gene expression levels that may affect the pharmacological outcomes of a patient on metformin therapy.

**Chapter 4**

This final chapter addressed a limitation that many genetic studies have: the focus on single time-point biomarkers collapses crucial information that may effectively inform a patient's response characteristics. Here, we developed a longitudinal mathematical model to characterize and quantify disease progression on metformin therapy using all HbA1c data available with the goal of explaining long-term HbA1c variability through the investigation of genetic, demographic, and clinical factors.

From this analysis, a turnover HbA1c model with a reversible metformin effect on the synthesis rate of HbA1c best characterized the longitudinal data from T2D patients. The model predicts that the onset of disease progression for patients on metformin is approximately 321 days, at which point, levels increase at a rate of approximately 0.1% [0.04%-0.16% 95% confidence interval] HbA1c per year; HbA1c levels are expected to increase at a steady state rate of approximately 0.16% [0.08%-0.22%] per year in patients not being treated with metformin. A set of candidate genes was studied for effects on disease progression of patients on metformin. Following the genetic analysis using

various approaches (e.g., HyperLasso, regression, model-based, etc.), the top 9 variants accounted for approximately one third of the total variability in the disease progression model parameter, which was markedly higher than the significant demographic predictors. Two SNPs in CSMD1 (rs2617102, rs2954625) and one SNP in *SLC22A2* (rs316009) surfaced as significantly influencing the long-term variance in HbA1c, with minor alleles leading to less and more favorable outcomes, respectively.

**Challenges and future direction**

The research presented herein leverages novel approaches to expand the current understanding of the factors that contribute to metformin response variability. This research furthers the current knowledge of the pharmacogenetic landscape by expanding the set of pharmacologically relevant genes from metformin transporters to transcription factors. Furthermore, a potential biological and pharmacokinetic basis for the identified associations is provided for transcription factors *SP1, HNF4-alpha*, and *PPAR-alpha*. To build on this work, future research should focus on providing a mechanistic link between the transcription factors and the DNA binding domains proximal to the pharmacological/disease gene of interest. Moreover, determining whether the variants are truly causative (not in LD to a variant that is) will lead to evidence-based pharmacogenetics with a strong biological rationale for the genetic factor at work. Through this research, a mathematical model that allows for the quantification of a robust phenotype of HbA1c and enables the exploration of

genetic and non-genetic factors on long-term response patterns in T2D patients was also produced. A disease gene (*CSMD1*) and a pharmacologically relevant transporter gene (*SLC22A2*) surfaced as significant influencers on long-term HbA1c variation. Future studies should focus on replication cohorts to recapitulate the genetic findings of *CSMD1* and *SLC22A2* and to characterize how the genes modulate response to metformin. Though the role of SLC22A2 in metformin pharmacokinetics seems established, the transporter may also have other roles in response to metformin. A mechanistic basis for *CSMD1* on the progression of HbA1c will be necessary to understand the pharmacological effect of this variant on response to metformin.

With interest in implementing true precision medicine growing, the research presented will prove beneficial for the T2D patient community. The ultimate goal of this dissertation research is to translate pharmacogenetic, model-based findings into clinical practice so that clinicians may provide data-driven, personalized treatment advice to patients based on genetics, demographics, and real-time HbA1c measurements.

## REFERENCES


1. Shu, Y., Sheardown, S. A., Brown, C., Owen, R. P., Zhang, S., Castro, R. A., Ianculescu, A. G., Yue, L., Lo, J. C., Burchard, E. G., Brett, C. M. & Giacomini, K. M. Effect of genetic variation in the organic cation transporter 1 ( OCT1 ) on metformin action. *J. Clin. Invest.* **117,** 1422–31 (2007).

2. Song, I. S., Shin, H. J., Shim, E. J., Jung, I. S., Kim, W. Y., Shon, J. H. & Shin, J. G. Genetic variants of the organic cation transporter 2 influence the disposition of metformin. *Clin. Pharmacol. Ther.* **84,** 559–62 (2008).

3. De Winter, W., DeJongh, J., Post, T., Ploeger, B., Urquhart, R., Moules, I., Eckland, D. & Danhof, M. A mechanism-based disease progression model for comparison of long-term effects of pioglitazone, metformin and gliclazide on disease processes underlying type 2 diabetes mellitus. *J. Pharmacokinet. Pharmacodyn.* **33,** 313–43 (2006).

4. Kent, W. J., Sugnet, C. W., Furey, T. S., Roskin, K. M., Pringle, T. H., Zahler, A. M. & Haussler, A. D. The human genome browser at UCSC. *Genome Res.* **12,** 996–1006 (2002).

5. Kajiwara, M., Terada, T., Asaka, J., Ogasawara, K., Katsura, T., Ogawa, O., Fukatsu, A., Doi, T. & Inui, K. Critical roles of Sp1 in gene expression of human and rat H+/organic cation antiporter MATE1. *Am. J. Physiol. Renal Physiol.* **293,** F1564–70 (2007).

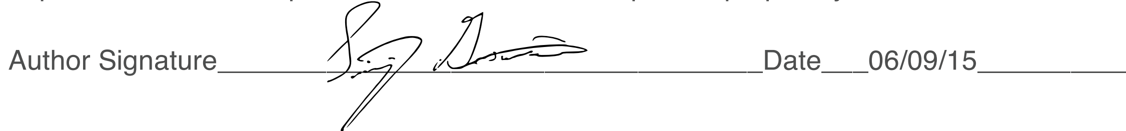
6. Saborowski, M., Kullak-Ublick, G. A. & Eloranta, J. J. The human organic cation transporter-1 gene is transactivated by hepatocyte nuclear factor-4alpha. *J. Pharmacol. Exp. Ther.* **317,** 778–85 (2006).

7. Mohlke, K. L. & Boehnke, M. The role of HNF4A variants in the risk of type 2 diabetes. *Curr. Diab. Rep.* **5,** 149–56 (2005).

**Publishing Agreement**

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature________________________________________Date___06/09/15__________