UNIVERSITY OF CALIFORNIA SAN DIEGO

Uncovering the physiological function of drug transporters

A Dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Bioengineering with Specialization in Multi-Scale Biology

by

Jeffry Carlos Granados

Committee in charge:

    Professor Sanjay K. Nigam, Chair
    Professor Adam J. Engler Co-Chair
    Professor Trey Ideker
    Professor Nathan E. Lewis
    Professor Andrew D. McCulloch

2023

DISSERTATION APPROVAL PAGE

The Dissertation of Jeffry Carlos Granados is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2023

DEDICATION

This dissertation is dedicated to my mother, father, and to all those who have sacrificed.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my primary advisor, Dr. Sanjay K. Nigam, for your mentorship. I feel fortunate to have had you guide me during my PhD. Your approach to science was enlightening and introduced me to a new way of thinking about complex problems. I am grateful for your mentoring style, as you always seemed to know when I needed to be encouraged or pressed. I want to thank you for the responsibilities you gave me, which helped me mature as a person and a scientist. I would also like to acknowledge the members of the Nigam lab, including Darcy Engelhart, Vladimir Ermakov, Patrick Zhang, Kian Falah and Wilma Mclaughlin, for the hundreds of conversations we have had about transporters, kidneys, and everything in between.

I would like to thank my committee members for their support throughout my PhD. Dr. Adam J. Engler, thank you for serving as my co-advisor and guiding me through each major step of the graduate school process. Dr. Nathan E. Lewis, thank you for your helpful discussions on metabolic pathways, your perspective on career choices, and for being such a great collaborator on my first first-author paper. Dr. Andrew D. McCulloch, thank you for your work with the Multi-Scale Biology Graduate Interfaces Training Program, which welcomed me to a community where I could think about science from the organism down to the atom. Dr. Trey Ideker, thank you for your invaluable insights and brilliant questions, which were crucial in assembling my dissertation.

I would like to thank my friends for the sense of community that they provided during my time in San Diego. I am particularly grateful for your company and support during a global pandemic, when friends became more important than ever. Without each of you, I am sure that I would have never completed my dissertation. Looking forward to spending time with you all

after work and on weekends went a long way and helped more than you could ever imagine. I feel extremely lucky to have met so many brilliant and caring people.

I would like to thank my family, Morena, Carlos, and Jason Granados for providing me with the foundation that I needed to complete this dissertation. Being hundreds of miles away from my family was always taxing on me, but we did a great job of staying touch, and I always felt you with me. Though you lamented that you could not provide me career advice, you instilled in me values that were far more important than anything you could have read in a textbook. The courage and dedication that I learned from watching you each day continues to inspire me, and I hope to make you proud.

Lastly, I would like to thank my girlfriend, Sara Mirza, and our dog, Charlotte. We started a little family during our time in San Diego that always encourages me to be the best version of myself. I cannot count how many times you had to hear me ramble aimlessly about my research, but you never complained when I did. Your love through a difficult time in my life was what held me above water and made me feel like I could overcome any obstacle. Truly, this, and anything I ever achieve, is our joint accomplishment. I guarantee that this dissertation would not be possible without you and Charlotte.

Pharmacology and Therapeutics, 2022, 112.3, 653-664. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of the material as it appears in "Regulation of Human Endogenous Metabolites by Drug Transporters and Drug Metabolizing Enzymes: An Analysis of Targeted SNP-Metabolite Associations" by Jeffry C. Granados, Jeramie D. Watrous, Tao Long, Sara Brin Rosenthal, Susan Cheng, Mohit Jain, and Sanjay K. Nigam in *Metabolites*, 13.2, 171. The dissertation author was the primary investigator and author of this paper.

Chapter 5 is being prepared for publication. The dissertation author was the primary author of this chapter.

VITA

2017   Bachelor of Science in Biomedical Engineering, University of Arizona

2020   Master of Science in Bioengineering, University of California San Diego

2023   Doctor of Philosophy in Bioengineering, University of California San Diego

PUBLICATIONS

Granados, J.C., Ermakov, V., Maity, K., Vera, D.R., Chang, G., & Nigam, S.K. (2023). The kidney drug transporter OAT1 regulates gut microbiome-dependent host metabolism. *JCI Insight. 8(2).*

Granados, J.C., Watrous, J.D., Long, T., Rosenthal, S.B., Cheng, S., Jain, M., & Nigam, S.K. (2023). Regulation of Human endogenous metabolites by drug transporters and drug metabolizing enzymes: an analysis of targeted SNP-metabolite associations. *Metabolites. 13(2)*, 17.

Nigam, S. K.*, & **Granados, J. C.*** (2023). OAT, OATP, and MRP drug transporters and the remote sensing and signaling theory. *Annual Review of Pharmacology and Toxicology*, *63*, 637-660.

Granados, J. C., Falah, K., Koo, I., Morgan, E. W., Perdew, G. H., Patterson, A. D., ... & Nigam, S. K. (2022). AHR is a master regulator of diverse pathways in endogenous metabolism. *Scientific reports*, *12*(1), 1-16.

Nigam, S. K., & **Granados, J. C.** (2022). A Biological Basis for Pharmacokinetics: The Remote Sensing and Signaling Theory. *Clinical Pharmacology & Therapeutics, 112(3),* 456-460.

Granados, J. C., Bhatnagar, V., & Nigam, S. K. (2022). Blockade of organic anion transport in humans after treatment with the drug probenecid leads to major metabolic alterations in plasma and urine. *Clinical Pharmacology & Therapeutics, 112*(3), 653-664.

Granados, J. C., Zhang, J., You, G., & Nigam, S. K. (2022). ORGANIC ANION TRANSPORTERS (OATs). *Drug Transporters: Molecular Characterization and Role in Drug Disposition*, 57-77.

Torres, A. M., Dnyanmote, A. V., **Granados, J. C.,** & Nigam, S. K. (2021). Renal and non-renal response of ABC and SLC transporters in chronic kidney disease. *Expert Opinion on Drug Metabolism & Toxicology*, *17*(5), 515-542.

Granados, J. C., Richelle, A., Gutierrez, J. M., Zhang, P., Zhang, X., Bhatnagar, V., ... & Nigam, S. K. (2021). Coordinate regulation of systemic and kidney tryptophan metabolism by the drug transporters OAT1 and OAT3. *Journal of Biological Chemistry*, *296*, 100575.

Granados, J. C., Nigam, A. K., Bush, K. T., Jamshidi, N., & Nigam, S. K. (2021). A key role for the transporter OAT1 in systemic lipid metabolism. *Journal of Biological Chemistry*, *296*, 100603.

Engelhart, D. C., **Granados, J. C.**, Shi, D., Saier Jr, M. H., Baker, M. E., Abagyan, R., & Nigam, S. K. (2020). Systems biology analysis reveals eight SLC22 transporter subgroups, including OATs, OCTs, and OCTNs. *International journal of molecular sciences*, *21*(5), 1791.

Engelhart, D. C., Azad, P., Ali, S., **Granados, J. C.**, Haddad, G. G., & Nigam, S. K. (2020). Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress. *International journal of molecular sciences*, *21*(6), 2002.

*denotes co-first author

FIELD OF STUDY

Major Field: Bioengineering with Specialization in Multi-Scale Biology

ABSTRACT OF THE DISSERTATION

Uncovering the physiological function of drug transporters

by

Jeffry Carlos Granados

Doctor of Philosophy in Bioengineering with Specialization in Multi-Scale Biology

University of California San Diego, 2023

Professor Sanjay K. Nigam, Chair
Professor Adam J. Engler, Co-Chair

Drug transporters are membrane proteins known for their role in handling pharmaceutical products. However, recent work suggests they contribute to physiology by transporting endogenous metabolites. This understudied function may explain some aspects of organ crosstalk, inter-organismal communication, and adverse drug side effects. This dissertation focuses on determining the endogenous compounds that renal organic anion transporter 1 (OAT1/SLC22A6) interacts with using in vivo, in vitro, and in silico methodologies. Chapter 1 presents the current state of drug transporter research. Chapter 2 of this dissertation is a reprint of

published work that describes the role OAT1 plays in mediating the relationship between the host and the gut microbiome via regulation of circulating small molecule metabolites in a pre-clinical model. These in vivo results were supported with in vitro binding and transport assays. Chapter 3 of this dissertation is a reprint of published work that explores the role that a common OAT1-inhibiting drug, probenecid, has on the plasma and urine metabolomes of healthy humans. This revealed dozens of plasma and urine drug-metabolite interactions caused by short-term exposure, including multiple likely occurring at OAT1. Chapter 4 of this dissertation is a reprint of published work that combines genomic and metabolomic data to determine associations between single nucleotide polymorphisms and circulating small, polar, bioactive molecules. This work reveals that numerous drug transporter and drug metabolizing enzyme genes play important individual and combined roles in physiology. Chapter 5 describes in silico models used to better characterize the nature of protein-ligand interactions involving OAT1. By analyzing molecular descriptors of compounds known to interact with OAT1, we generated predictive ligand-based models. We then explored potential binding mechanisms for different classes of compounds using a predicted OAT1 structure. Chapter 6 is a summary of the contributions to the field and future directions in drug transporter research.

CHAPTER 1: INTRODUCTION

1.1 DRUG TRANSPORTERS

1.1.1. ABC and SLC Membrane Transporters

Membrane transporters are proteins localized to the plasma membrane that are involved in the uptake and efflux of various small molecules. These proteins are either ATP-binding cassette (ABC) transporters, which are usually efflux transporters, or solute carrier (SLC) transporters, which are usually uptake transporters. Together, these families consist of over 400 unique proteins with a wide variety of functions in transport of both endogenous and xenobiotic compounds [1, 2].

1.1.2. Drug Transporters

Drug transporters are a subset of SLC or ABC proteins that are best known for their role in the absorption distribution, metabolism, and excretion (ADME) of small molecule drugs. These transporters are characterized by a broad range of substrates and inhibitors, which encompass several distinct drug classes, such as antibiotics, antivirals, chemotherapeutics, and many others [3]. Each drug transporter has its own unique array of compounds it interacts with, thought there is some overlap between them with specific small molecules. This ability to handle dozens of unique compounds can be described as multi-specificity and distinguishes drug transporters from other transporters, which typically have a smaller range of compounds they interact with. These functionally limited transporters can be described as either oligo-specific or mono-specific, depending on the number of distinct compounds they interact with.

The main drug transporters are ABCB1 (P-gp), ABCG2 (BCRP), SLCO1B1 (OATP1B1), SLCO1B3 (OATP1B3), SLC22A6 (OAT1), SLC22A8, (OAT3), and SLC22A2 (OCT2). While these have received the majority of research attention, other drug transporters, such as ABCC2 (MRP2), ABCC4 (MRP4), ABCB11 (BSEP), SLC47A1 (MATE1), have begun

to receive research attention, as well. Because of the multi-specific nature of these transporters, they are potential sites for drug-drug interactions [4]. Drug-drug interactions occur when two or more drugs compete for access to a protein, typically transporters or enzymes. In the case of transporters, the drugs compete at the site of the transporter for either entry into the cell or efflux out of the cell, leading to altered extracellular and intracellular concentrations of one or more of the drugs. These changes can lead to negative consequences, such as decreased efficacy, increased toxicity, or adverse side effects. Global regulatory agencies (e.g., Food and Drug Administration, European Medicines Agency) have recommended that all novel drug entities be screened for interaction with the main drug transporters (P-gp, BCRP, OATP1B1, OATP1B3, OAT1, OAT3, OCT2) to avoid potential drug-drug interactions.

In large part due to regulatory interest, there is a plethora of data on the pharmaceutical function of drug transporters. These studies have revealed overlap in the function of these drug transporters. While some drugs may be associated with a single transporter, many have been shown to interact with more than one transporter. For example, OAT1 and OAT3 have considerable overlap with respect to drugs [5] and also share a number of substrates with the renal efflux transporters MATE-1 and MATE-2K. The shared function suggests that drug transporters are part of a larger system that may help protect the body from xenobiotics, especially given their enriched expression in excretory organs.

1.1.3. Drug Transporter Expression Patterns

Drug transporters are primarily expressed in barrier epithelial tissues, such as the kidney, liver, and intestine [6]. These barrier tissue separate body fluids, which is particularly important in drug clearance, as many compounds are excreted via the urine, feces, or bile (**Figure 1.1**). For example, the kidney and more specifically, the proximal tubule, separates the blood from the urine, the hepatocyte in the liver separates the blood from the bile, and the enterocyte in the

intestine separates the blood from the gut lumen. In addition to these organs, the brain (blood-brain barrier, choroid plexus) and the placenta (mother-fetus) also express multiple drug transporters.

Beyond the tissue and cell expression patterns, the subcellular localization for drug transporters is important when considering the individual and combined role of these proteins. The plasma membrane of polarized cells leads to two distinct sides of the membrane: the basolateral and apical membranes. Each of these membranes has their own distinct profile of transporters expressed, with most drug transporters exclusively expressed on a single side. For example, OAT1 and OAT3 are only expressed on the basolateral (blood-facing) side of the proximal tubule, meaning that they have a pronounced role in the uptake of compounds from the blood into the cell. On the apical side of the proximal tubule, MRP2 and MRP4 are efflux transporters that are responsible for excreting compounds from the cell into the urine and are exclusively expressed on the apical membrane. When taken with the different mechanisms of action (uptake, efflux), the combinatorial possibilities across multiple organs represents a robust network of transporters that can protect the body from numerous drugs and toxins and potentially mediate the intracellular and extracellular levels of other types of compounds.

1.1.4. Drug Transporter Function

With respect to clinically relevant drugs, there is structural and functional diversity among these compounds, as the hundreds of drugs that are substrates or inhibitors of drug transporters have unique chemical structures, mechanisms of actions, and ADME properties. The ability to handle so many distinct compounds implies that drug transporters may also handle other classes of compounds [6]. For example, OAT1 has been shown to be inhibited by or transport different antibiotics, antivirals, NSAIDs, natural products, uremic toxins, endogenous metabolites, as well as other classes of small molecules (**Figure 1.2**).

While most drug transporter research emphasizes their role in the handling of pharmaceutical products, roles in the handling of other non-pharmaceutical products, such as endogenous metabolites, natural products, and toxins, are have received comparatively little research attention. This role could lead to other types of competitive interactions, such as drug-metabolite and drug-natural product interactions, which can have negative health consequences. These understudied functions are pivotal, as the endogenous function may be the 'primary' role of drug transporters. Drug transporters are not suddenly expressed in response to the introduction of drugs and are always expressed, suggesting they may have an alternative role. One of the primary goals of this dissertation is to help uncover the physiological role of OAT1 by determining the endogenous compounds it interacts with and develop strategies for determining transporter function that can be applied to other drug transporters.

1.2 REMOTE SENSING AND SIGNALING THEORY

1.2.1 ADME Proteins in Endogenous Metabolism

The Remote Sensing and Signaling Theory posits that the primary function of proteins involved in the ADME of drugs is to simultaneously optimize the levels of hundreds of endogenous metabolites and signaling molecules [7, 8]. The Remote Sensing and Signaling System consists of drug transporters, drug metabolizing enzymes, ligand-activated transcriptional regulators, and related proteins that all play important roles in maintaining and returning the body to homeostasis following perturbation (**Figure 1.3**) [9, 10]. This system can be understood as a three-tier system where small molecules with important signaling roles are on the first layer, transporters and enzymes that transport or modify the compounds are on the second layer, and ligand-activated transcriptional regulators, which respond to the small molecules by changing the expression profiles of the relevant transporters/enzymes, are on the third level.

The Remote Sensing and Signaling Theory was initially formulated with the SLC22 family in mind, but has recently been expanded to include roughly 500 relevant genes/proteins [11]. Given the expression and subcellular localization of the proteins in the Remote Sensing and Signaling System in important epithelial tissues, it is well-equipped to serve as a robust network with numerous potential compensatory mechanisms. The shared function across multiple organs and cell types suggests that these transporters and enzymes may serve as key conduits to organ crosstalk (**Figure 1.3**) [11]. Indeed, it has been shown that when renal urate excretion is diminished due to kidney disease, ABCG2 expression/function in the intestine is upregulated to excrete urate into the gut lumen [12]. We expect other similar compensatory mechanisms to exist across multiple organ axes, further emphasizing the importance of determining the endogenous role of these drug transporters.

In addition to organ crosstalk, the system may also serve as a mediator of inter-organismal communication (**Figure 1.3**). One clear example is between the mother and the fetus. The communication between these organisms is largely mediated by the placenta, which is responsible for both the nutrition and the protection of the fetus from potential toxins [13]. The placenta expresses several drug transporters, suggesting that it is an organ with a strong, adaptive system for handling a myriad of small molecule compounds [14]. The function and regulation of numerous drug transporters, including BCRP, MRP2, and others, during different phases of pregnancy has been covered in other work [15].

Beyond the role of drug transporters in the mother-fetus relationship, there is evidence of drug transporters playing an important role in the inter-organismal communication between the host and the gut microbiota, which consists of hundreds of distinct bacterial species. The communication between the host and the gut microbiome is largely mediated by the small

molecules produced jointly by these two organisms. In this symbiotic relationship, the host consumes dietary products that are acted upon by the transporters/enzymes expressed by diverse microbial populations to produce metabolites that cannot be produced by the host alone [16]. These compounds often have important signaling roles in distal organs, such as the brain and heart [17, 18]. While the signaling roles of these metabolites have been studied, the ADME properties of these metabolites, such as the proteins involved in their clearance, has been largely ignored. The work presented in this dissertation addresses this shortcoming and indicates that the Remote Sensing and Signaling System plays a crucial role in regulating the circulating levels of these compounds, and as such, is important in mediating inter-organismal communication. Chapter 2 of this dissertation explores the relationship between OAT1 and the gut microbiome via the regulation of small molecule endogenous metabolites.

1.2.2 Evolutionary Conservation of ADME Genes

Understanding why drug transporters, drug-metabolizing enzymes, and related proteins are expressed in multiple species may seem obvious, but this is only in the context of today's society, where pharmaceutical products and environmental toxins are highly prevalent. A key question arises: Why are these genes evolutionarily conserved? Evolutionary analyses have shown that some of these genes, particularly ABC transporters and CYP450 enzymes, are highly conserved in model organisms, such as mice, flies, zebrafish, and bacterial species [19, 20]. This suggests that there is some essential, physiological role they play. While this essential role could be related to detoxification, another perspective is that these genes play crucial roles in endogenous metabolism. Increasing evidence of their individual and combined function in the movement of endogenous metabolites suggests further study, which is in part, addressed in this dissertation.

1.2.3 Potential Consequences of Perturbing Endogenous System

Drug-drug interactions are the primary competitive interactions that occur at drug transporters, but there are likely many other competitive interactions between different kinds of compounds [21]. Drug-metabolite interactions are often leveraged in drug development as the primary mechanism of action. Probenecid, a drug used to treat gout and hyperuricemia, participates in a drug-metabolite interaction with urate at the site of URAT1 on the apical membrane [22]. While the ADME system normally operates efficiently, there are some situations in which the perturbation of the endogenous function of drug-handling proteins can lead to negative consequences. For example, adverse drug reactions and drug side effects are often unexplained and can manifest in a number of distinct ways [23]. These can range from nausea and headaches to more serious cases, such as renal failure and internal bleeding. These cases often occur in older patients, who may be taking multiple drugs simultaneously for a variety of conditions for extended periods of time, leading to multiple possible competitive interactions at numerous transporters and enzymes [24].

In addition to short-term drug reactions, there is evidence of drug-induced metabolic diseases caused by long-term treatment of drugs [25]. HIV patients take antiviral cocktails and organ transplant patients take immunosuppressants for the rest of their lives. While these drugs are effective in suppressing the virus and the immune system, respectively, they will also impact the function of drug metabolizing enzymes and drug transporters consistently, which could be the root cause of the metabolic dysfunction [26, 27]. However, it is difficult to separate the disease state from the unintended perturbations to the Remote Sensing and Signaling System. The work we describe in Chapter 3 of this dissertation demonstrates a potential mechanism for these effects, as even the short-term impact of non-toxic drug administration can have a major

impact on the plasma and urine metabolome. The constant stressing of the system with drugs may impact its ability to return the body to homeostasis in an efficient manner.

1.3 CURRENT STRATEGIES FOR PREDICTING TRANSPORTER FUNCTION

Having established the broad and important role that drug transporters can play in several aspects of xenobiotic and endogenous small molecule handling, it is imperative to develop and utilize strategies to determine the small molecule function of these proteins. Several methodologies have been applied to determine transporter function, as described below. The work presented in Chapters 2-5 of this dissertation represent novel systems and approaches to understanding transporter function. Prior to presenting the advances in drug transporter research, the current state of the field is summarized.

1.3.1. In Vitro Assays

Historically, determining mechanistic interactions between transporters and small molecule compounds have been done using two dimensional in vitro assays with cells overexpressing the transporter of interest [28]. These assays have used transfected cell lines expressing transporters from different species, but recent work has established human cells with human transporters as the preferred choice [29]. These assays are designed to measure either inhibition or transport. In the inhibition assays, cells that overexpress the transporter are plated and exposed to the compound of interest along with either a radiolabeled or fluorescent probe for transporter function. These assays, when coupled with known transport properties of the probe, produce K_i values for the compound of interest and are run in replicates to ensure reproducibility. Transport assays, which are more time-consuming and expensive, measure the uptake or efflux of a radiolabeled compound and produce K_m values. Alternatively, with advances in metabolomics, specific compounds can be measured either within the cell lysate or the cell media to determine the rate of uptake/efflux [30, 31].

There are varying levels of complexity with respect to 2-D in vitro assays, with some assays attempting to recapture some aspects of polarization by using Transwell filters [32]. Furthermore, the cell types used can have a profound impact on the results, with HepG2 cells being used for liver-related studies [33], and HEK293 being used for renal studies [34]. Current studies are aiming to develop more sophisticated and reproducible systems that consider physiological factors. 3-D organoids, organs on a chip, and other strategies across multiple organs are rapidly advancing [35-40], but 2-D assays remain the gold standard. Despite their position as the gold standard, these assays are an imperfect system, as many phenomena observed in vitro are not reflected in vivo in humans, rodents, or other model organisms [41].

1.3.2. In Vivo Models

Two-dimensional in vitro assays are currently the gold standard for determining mechanistic transporter-compound interactions because they are relatively inexpensive, replicable, and use human proteins, however, these models lack physiological relevance. Despite efforts in organoid and organ-on-a-chip research to improve model systems, rodents remain the best in vivo models for understanding transporter function.

Several in vivo transporter knockout rodent models have been generated to study both the pharmaceutical and endogenous role of these proteins, with OATs, OATPs, MRPs, and other clinically relevant drug transporters being studied [42-49]. The most useful models are those with clear orthologs in humans. If there are not clear human orthologs, humanized mouse models have been used to explore some drug transporters, such as OATP1B1 and OATP1B3 [50, 51].

The mouse and human proteins for OAT1 are similar with respect to function, leading to useful clinical predictions. These mice have been used to understand drug clearance, but recent work by our group has emphasized the profound endogenous role that OAT1 has on physiology using serum metabolomics, as OAT1's expression on the basolateral side of the proximal tubule

9

makes the serum a valuable biofluid in understanding OAT1 function [42, 52-55]. Despite the wealth of knowledge gained from these *Oat1* knockout mice, these in vivo models have never been metabolically stressed in any capacity. The work presented in Chapter 2 of this dissertation addresses this shortcoming in the literature by introducing a cocktail of antibiotics to deplete the gut microbiome.

In addition to mouse models, rats have also been used to study drug transporter function, such as OAT1, OAT3, BCRP, and P-gp [56, 57]. Flies have also proved useful in understanding the function of understudied transporters, though there is no clear ortholog for OAT1 or some other drug transporters [58, 59].

Ultimately, each in vivo model organism needs to be analyzed in the context of the transporter. For example, renal drug transporters, such as OAT1, have been studied by using serum/plasma and urine as functional readouts of the transporter. Similarly, liver transporters separate the blood and the bile, and transporters in enterocytes separate the blood and the gut lumen, so these biological samples, respectively, would serve as useful information. These samples can be studied with developing tools like metabolomics or more targeted screens for specific compounds. Metabolomics, coupled with chromatography, allows for the quantitative or semi-quantitative measurement of any small molecule compound whose chemical standard exists within a pre-defined library. Metabolomics can be applied to bodily fluids, tissues, or cells and is a powerful tool. Organs in which the transporters of interest are expressed may also be probed, as other relevant genes/proteins in the tissue may alter their expression to compensate for the putative function of the transporter. These organs can be analyzed using transcriptomic and proteomic studies.

1.3.3. Genome Wide Association Studies (GWAS)

As genomic data becomes more readily available and easily interpretable, more studies have linked single nucleotide polymorphisms (SNPs) in or near transporter genes with phenotypes [60]. These genome wide association studies (GWAS) have been used to link genes to complex phenotypes, such as diseases [61]. For example, SNPs in the *SLC22A5* transporter gene have been associated with systemic carnitine deficiency [62]. While these studies have been useful, they are often difficult to unpack, as complex diseases are the result of many simultaneous processes and may involve multiple genes. To identify more specific interactions between transporters and small molecules, GWAS can focus on intermediate phenotypes with metabolomic and proteomic technology as the primary result. Considering that we are primarily interested in the small molecule related function of transporters, using metabolomics on bodily fluids, like the plasma and urine, can reveal important potential roles of transporters and related genes [63-65]. While these studies alone are insufficient in determining protein-metabolite interactions, they can provide high confidence predictions that have clear clinical implications. Chapter 4 of this dissertation focuses on the associations between SNPs in drug-handling genes and small, polar bioactive molecules with important signaling roles. The presented study emphasizes the understudied, endogenous role of what are best known as ADME genes.

1.3.4. Ligand-based Models

Wet lab experiments have largely driven transporter research, but the increased accessibility and knowledge of in silico approaches have helped advance computational approaches for analyzing transporter function. Within each chemical structure of a small molecule, there is a tremendous amount of information. The properties extracted from a single compound can be a simple as counts, such as number of hydrogens, but can also cover more complex features, such as total polar surface area or many others. Overall, there are 1-D, 2-D,

11

and 3-D chemical properties that can be calculated for any small molecule compound, with 1-D properties consisting of different fingerprints [66], 2-D properties consisting of properties calculated from a flat chemical structure [67], and 3-D properties, which are molecular descriptors calculated from the more complex 3-D structure. This is an active area of research, and more properties and combinations of properties are being discovered to better characterize molecules. This long list of features can then be used in the generation of mathematical ligand-based models. In Chapter 5 of this dissertation, chemical features of small molecules that are transporter-mediated are used to better understand transporter function. These ligand-based approaches can also be used to determine the small molecule specificity of enzymes or receptors or to determine what compounds with similar phenotypes have in common [68].

One of the main goals in transporter research is to develop high accuracy predictive models that can determine whether specific compounds will interact with the transporter of interest. This requires sufficient data, which only exists for drug transporters that have received major research attention. Even then, the majority of these data come from in vitro assays from multiple groups, which can be both time consuming, suffer from variability between assays, and misrepresent in vivo interactions. In spite of these limitations, multiple groups have developed useful models based on high throughput in vitro data to build quantitative structure activity relationship (QSAR) machine learning models for drug transporters, like OATP1B1, P-gp, and BCRP [69-71]. In these models, the target variable is either a quantitative value, such as K_m or K_i or a binary classification, such as substrate, non-substrate, inhibitor, or non-inhibitor. These models continue to improve with higher quality data, and advances in machine learning algorithms, but again, they primarily rely upon in vitro data. Chapter 5 of this dissertation

addresses this concern by developing a framework for predicting transporter-interacting molecules with in vivo metabolomics data from plasma and urine.

In vivo data has been scarce for most transporters, but studies by our group have generated metabolomics data for OAT1 and OAT3 that can be used in machine learning models [54, 72, 73]. Machine learning models were generated with the goal of differentiating OAT1-interacting compounds from OAT3-interactiong compounds [74]. A later study focused on separating OAT-interacting, OATP-interacting and OCT-interacting compounds [75, 76]. While distinguishing transporter specificity is important in drug design and drug dosing, in this dissertation, rather than classifying between transporters, we build models that determine whether a compound is or is not a substrate/inhibitor of relevant transporters with a flexible framework that can be applied to multiple metabolomics datasets.

In addition to machine learning models, pharmacophore models have also been used to understand transporter function [77, 78]. These pharmacophore models utilize the 3-D structures for compounds that have been established as either substrates or inhibitors of specific transporters. By clustering and transposing the compounds in the same clusters, representative pharmacophore models can be applied to novel compounds. This dissertation also applies unsupervised learning approaches to better understand the multiple distinct functions of a single transporter.

1.3.5. Protein-based Models

Ligand-based models depend on prior knowledge of function, so it can be difficult to identify novel substrates or inhibitors of transporters with these methods. An alternative and complementary approach is to use protein-based approaches. Protein-based approaches encompass methodologies where the protein structure is used with molecular modeling software to make predictions on protein-ligand interactions [79]. Because transporters are membrane-

bound, it has been historically difficult to isolate a protein structure [80, 81], though recent cryo-EM work has determined the structure for some human SLC and ABC transporters. Very few drug transporters, which includes P-gp and BCRP, have experimentally derived protein structures, with or without ligands bound [82-85].

To circumvent this problem, multiple strategies have been applied to better understand transporter function at the molecular level. For one, bacterial homologs have been used, which share some similarities with the human structure and can then be used to generate predictions on active sites and mechanisms [86]. However, these models can be difficult to translate to humans because of the different biological factors in bacteria. More recently, protein structure prediction has improved at a rapid rate and can be partly validated by existing cryo-EM results for related proteins. This has led to an influx of software, AlphaFold among others, that can take protein sequences and produce high confidence protein structures that can be probed for more specific virtual experiments [87]. Multiple SLC transporters (OATPs, OATs), including OAT1, have been studied using some of these novel tools in a variety of ways, which describe mechanism of action, conformations, and ligand interactions [88-91]. This dissertation focuses on OAT1/SLC22A6 and combines known ligand information with the predicted human protein structure to identify key residues in binding.

1.3.6. Systems Biology Models

Genome-scale metabolic reconstructions are well-established models to portray complex metabolic interactions in bacterial and mammalian species using transcriptomic data [92, 93]. To understand transporter function in the context of a complete metabolic reaction map, transcriptomic data from *Oat1* and *Oat3* knockout mouse kidneys has previously been used to predict endogenous metabolites that the transporters interact with [77, 94, 95]. Recon3D represents the latest version of the genome scale metabolic reconstruction, with several

metabolites and reactions added to the previous iteration [96]. Recon3D was applied to kidney transcriptomic data and constrained with serum metabolomics from *Oat1* knockout mice to better understand the functional role of OAT1 in vivo, revealing an important role in lipid metabolism that was supported with in vitro assays [54]. In addition to traditional metabolic reconstructions, there are also novel methods being developed to better gauge the consequences of transporter knockout. Metabolic task analysis was applied to *Oat1* knockout transcriptomic data and revealed major changes in intracellular tryptophan metabolism, which was supported by both in vivo and in vitro data demonstrating a key role for OAT1 in tryptophan metabolism [73]. Overall, systems biology models are helpful in analyzing transporters from a broad perspective, rather than from a mechanistic perspective, though they have only been applied to renal drug transporters. Additional studies have not specifically identified transporter knockout transcriptomic data, but they have emphasized the metabolic pathways that are perturbed by common drugs, suggesting that adverse drug reactions may indeed be caused in part by inhibition of drug transporter function [97].

1.4 SCOPE OF DISSERTATION

Since the majority of research on drug transporters has emphasized their role in the handling of small molecule pharmaceutical products, comparatively little attention has been paid their other potential functions. Prior to this dissertation, there was little evidence to support the idea that drug transporters interact with specific classes of endogenous metabolites. The goal of this work was to provide evidence and methodologies to show that drug transporters, and specifically OAT1, are involved in the handling of diverse small molecules. We explore this function using preclinical and human in vivo studies, in vitro cell assays, and predictive in silico models. Establishing frameworks with which to study specific transporters can lead to future

studies analyzing the non-pharmaceutical roles of other ADME-related proteins, such as drug transporters, drug metabolizing enzymes, and receptors. While the introduction has discussed drug transporters broadly, it is not realistic to explore the function of all drug transporters in a single study. Hence, the primary focus in this work is on the renal drug transporter OAT1 (SLC22A6), which was discovered by our group and handles over 150 unique drugs [98].

Chapter 2 describes how the renal drug transporter OAT1 plays a key role in regulating the inter-organismal communication between the host and the gut microbiome via the regulation of gut microbe-derived metabolites in the blood. We used a genetically engineered knockout mouse model coupled with microbe depletion via a broad-spectrum antibiotic cocktail to explore the overlap between OAT1 handling and microbial production of metabolites. We collected serum from the animals and performed global metabolomic profiling for endogenous metabolites on the samples. We found that 40 metabolites were significantly altered by both loss of *Oat1* and depletion of the gut microbiome. Among these, we were particularly interested in those whose circulating levels were increased by loss of OAT1, as this suggests they are normally cleared by an OAT1-mediated mechanism, and whose circulating levels were decreased by microbial depletion, as this indicates that the bacterial species involved in the generation of the compounds have been diminished. The compounds satisfying these criteria were generally characterized by number of sulfate groups and aromatic bonds. In addition to identifying gut-derived metabolites handled by OAT1 in vivo, we also supported our results with novel in vitro magnetic bead binding assays and traditional in vitro transport assays, which revealed that some gut-derived compounds that did not reach statistical significance or were not measured on the metabolomics platform interact with OAT1 in vitro.

Chapter 3 builds upon the preclinical results from Chapter 2 and explores the short-term clinical impact that an OAT-inhibiting drug has on both the plasma and urine metabolomes. We collected plasma and urine from healthy human participants and then administered an oral dose of probenecid, an OAT1-inhibiting drug used to treat gout. Following 5 hours, which falls within the expected half-life or probenecid, we collected plasma and urine again from the participants. By comparing the post-treatment samples to pre-treatment samples, we were able to identify simultaneous drug-metabolite interactions caused by probenecid. We were particularly interested in compounds that were both elevated in the plasma and decreased in the urine, as these compounds are reflective of potential OAT1 and OAT3 function. We found 97 compounds that were likely OAT-mediated drug-metabolite interactions, including ~25 that had either in vivo support from knockout mice or in vitro support. In addition to analyzing OAT1 and OAT3-mediated functions, we also explored the consequences of URAT1 inhibition. Since URAT1 is an uptake transporter expressed on the apical side of the membrane, we aimed to analyze the overlap of compounds that were increased in the plasma and decreased in the urine. Interestingly, we found that among ~600 measured compounds, only urate, the principal metabolite in gout, was altered in both. Probenecid's primary mechanism of action is by inhibition of urate reuptake from the urine, revealing that probenecid and urate both participate in a hyper-specific drug-metabolite interaction. Overall, we demonstrated that a clinically relevant drug has a profound short-term impact on the circulating and urine levels of endogenous metabolites via inhibition of OAT1 and OAT3.

Upon establishing the specific in vivo role of OAT1 with genetic engineering and inhibition by a drug using metabolomics in Chapters 2 and 3, we then aimed to understand how genetic factors in humans, such as SNPs present in SLC22 and other drug-handling genes, may

individually and jointly regulate the circulating levels of small, polar bioactive molecules. In Chapter 4, we combined genomic data with untargeted plasma metabolomic data to identify single nucleotide polymorphisms in transporter, enzyme, and related genes that are associated with the circulating level of hundreds of endogenous metabolites, such as eicosanoids, fatty acids, bile acids, and other compounds with established signaling roles. The majority of the metabolites were unidentified, so we focused on the specific compounds that were associated with multiple distinct genomic loci, suggesting coordinated regulation across multiple transporters (SLC22s, SLCOs, ABCCs) and enzymes (CYPs, SULTs) expressed in different organs. We found five metabolites that were associated with four distinct genomic loci and dozens that were associated with three distinct genomic loci, indicating that multiple proteins combine to mediate the levels of a particular metabolite. We then aimed to focus on the ~100 compounds with known identities. We identified 18 genomic loci associated with at least one of these compounds, with multiple relationships that have been previously established with in vivo and in vitro results, as well as multiple novel associations that require further study. These results showed that SNPs in or near the *SLCO1B1/3/7* genomic region were associated with dozens of unique metabolites. Many of these SNPs have also previously been linked to altered drug handling, showing that certain SNPs can exacerbate drug-metabolite interactions in particular populations. Overall, Chapter 4 demonstrates the profound individual and combined role that drug-handling genes play in regulating key signaling molecules.

Finally, in Chapter 5, we leverage the in vivo small molecule metabolomics data we have acquired and the existing in vitro data for OAT1 to develop in silico models for predicting and better understanding transporter function. We generated binary classification machine learning models that used calculated chemical properties as features to predict whether a compound

would be altered in vivo in the plasma and urine following administration of probenecid. The resulting models were then compared, and the best performing model was then tested on a novel dataset consisting of drugs. The results from these predictions demonstrated that endogenous metabolites did not provide enough information to predict transporter-mediated drugs. This led us to explore whether there are distinct mechanisms of action for different classes of compounds. We first collected and characterized all the small molecules known to interact with OAT1 in any capacity and identified distinct clusters based on their molecular properties. Given the structural diversity, we then aimed to understand how these molecules interact with the OAT1 protein. To this end, we developed protein-based models using the predicted human protein structure from AlphaFold2, as no experimental structure exists. We found that OAT1-interacting molecules have distinct residues within a proposed intracellular ligand binding site, suggesting multiple mechanisms of transport. Overall, this dissertation represents a step forward in developing methodology to determine the individual and combined role of transporters in endogenous metabolism.
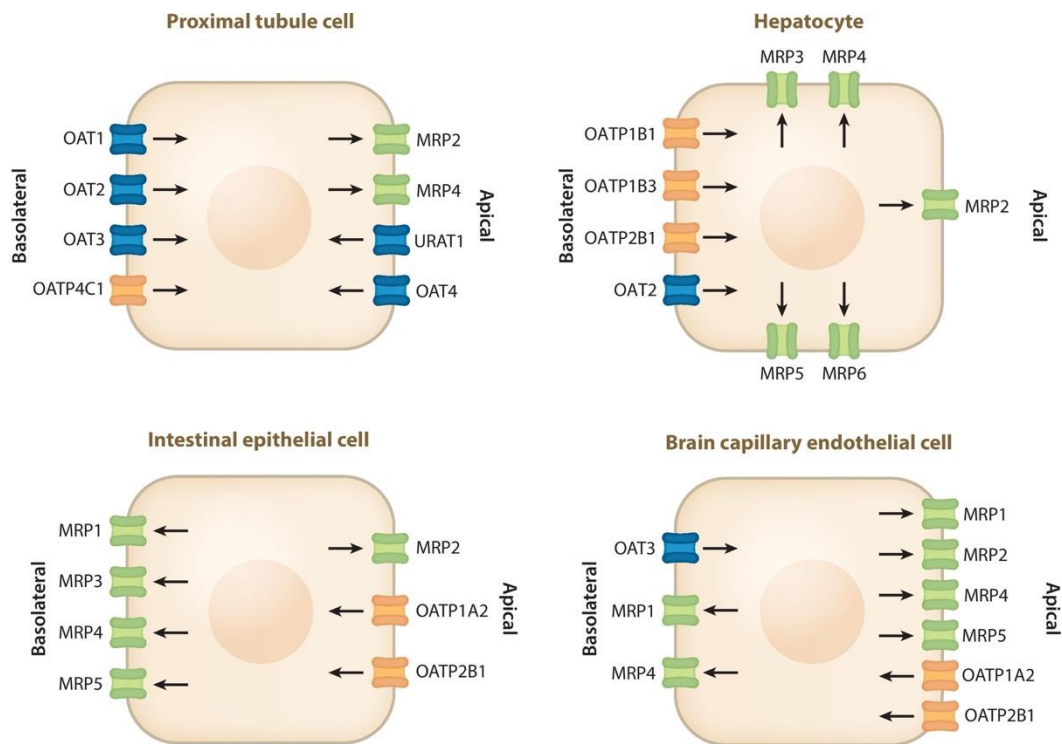
Figure 1.1 OATs, OATPs, and MRPs are expressed in polarized barrier epithelial tissues.

Organic anion handling proteins are differentially expressed at the basolateral and apical membranes of cells in the kidney, liver, intestine, and brain. When considered in light of their uptake or efflux mechanisms, the combinatorial possibilities for interaction between these multi-, oligo-, or monospecific transporters on both surfaces and their substrates create a robust, multiorgan network for handling small molecules and affecting net flux between tissues and body fluids. We emphasize that these are not the only multi-, oligo-, and monospecific transporters expressed on these surfaces involved in handling small molecules. The multispecific drug transporters OAT1, OAT3, OATP1B1, OATP1B3, MRP2, MRP3, and MRP4 are probably the best studied of these transporters from a functional perspective; much less is known about some of the other transporters that are shown. Abbreviations: MRP, multidrug resistance protein; OAT, organic anion transporter; OATP, organic anion transporter polypeptide. Adapted from Nigam and Granados, *Annual Reviews of Toxicology and Pharmacology,* 2023 [10].
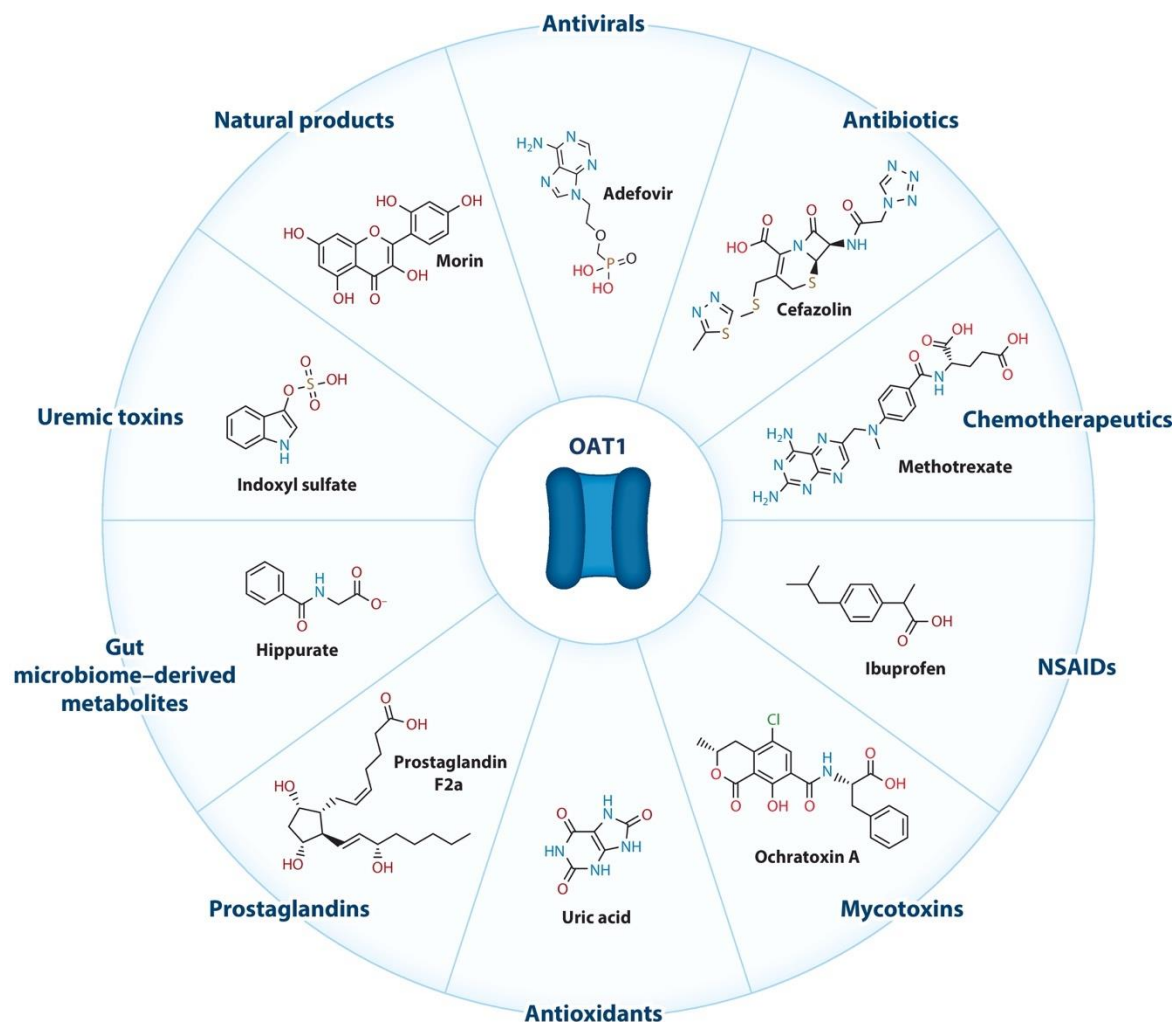
Figure 1.2 OAT1 is a multispecific transporter that interacts with several classes of small-molecule compounds.

OAT1 interacts in vitro or in vivo with numerous drugs, endogenous metabolites, natural products, antioxidants, nutrients, gut microbiome–derived compounds, and toxins. Other transporters in the OAT (OAT3), OATP (OATP1B1/OATP1B3), and MRP (MRP2/MRP4) families handle similar classes of small molecules. Abbreviations: MRP, multidrug resistance protein; NSAID, nonsteroidal anti-inflammatory drug; OAT, organic anion transporters; OATP, organic anion transporter polypeptide. Adapted from Nigam and Granados, *Annual Reviews of Toxicology and Pharmacology,* 2023 [10].

21

Figure 1.3 The Remote Sensing and Signaling Network can be represented by overlaying interactions at multiple scales.

(1) Protein interaction network (blue and teal icons). (2) Multi-specific drug and other transporters interacting with endogenous small molecules (magenta and pink icons), (3) Organs interacting via endogenous small molecules (blue arrows), and organisms interacting with the same (4) or different (5) species (blue arrows), OATs, OATPs, MRPs, and other multi-, oligo-, or monospecific transporters, such as P-gp (MDR) and BCRP (ABCG2), transport a wide array of endogenous small molecule substrates. Some of the main drug transporters have overlapping substrate specificity, potentially adding robustness to the regulation of systemic and local levels of organic anions. Remote organs communicate via transporters and enzymes (including DMEs) that regulate the movement and optimization of these high informational content metabolites. Each organ expresses a distinct set of transporters and enzymes that preferentially facilitate the movement of specific endogenous of specific endogenous small molecules (signals) that interact with regulatory proteins (sensors) and enable numerous feedback loops. Together, this helps optimize levels of endogenous small molecules across multiple scales (e.g., organelles, cells, organs, multi-organ systems, multiple organisms). The Remote Sensing and Signaling Network enables two types of interorganismal communication via these transporters and drug metabolizing enzymes. Within the same species, as in the mother-fetus connection, the placenta expresses many multi-, oligo-, and monospecific transporters that aid in the nutrition of the fetus by transporting metabolites and signaling molecules. Host-gut microbiome interactions are also mediated by transporters and enzymes in different organs (e.g., gut-liver-kidney-brain) as indicated, for example, by the effects of numerous small molecules arising in the gut microbiome (tryptophan-derived metabolites, secondary bile acids, fatty acids, and other signaling molecules) upon nuclear receptors, GPCRs, and kinases in remote organs. Not shown is how the Remote Sensing and Signaling Network works alongside other homeostatic systems, such as the autonomic nervous, neuroendocrine, immune, and growth factor and cytokine systems. Abbreviations: ABC, ATP-binding cassette; BCRP, breast cancer resistance protein; CNS, central nervous system; DME, drug metabolizing enzyme; GPCR, G protein-coupled receptor; MRP, multidrug resistance protein; OAT, organic anion transporter; OATP, organic anion transporter polypeptide; P-gp, P-glycoprotein; SLC, solute carrier. Adapted from Nigam and Granados, *Annual Reviews of Toxicology and Pharmacology,* 2023 [10].

1 Remote Sensing and Signaling Network of approximately 600 multi-, oligo-, and monospecific transporters, DMEs, nuclear receptors and other proteins

- ▌▌ Transporter
- ■ DME
- ⬠ Nuclear receptor
- ● Other protein
- — Protein interaction network

2 Multispecific drug transporters interacting with endogenous small molecules at many scales (organelles, cells, organs, organisms)

OATP1B1 OATP1B3 MRP2 MRP4

OAT1 OAT3

Pgp BCRP

Brain/CNS

3 Organs interacting via endogenous small molecules

Liver

Kidney

Gut

Placenta

- ● Metabolite
- ▲ Signaling molecule
- ◆ Antioxidant or other small molecule with high informational content

Host–gut microbiome

5 Organisms interacting with **different** species via small molecules

Mother–fetus

4 Organisms interacting with the **same** species via small molecules

Nigam SK, Granados JC. 2023
Annu. Rev. Pharmacol. Toxicol. 63:637–60

CHAPTER 2: THE KIDNEY DRUG TRANSPORTER OAT1 REGULATES GUT MICROBIOME-DEPENDENT HOST METABOLISM

2.1 ABSTRACT

Organic anion transporter 1 (OAT1/SLC22A6, NKT) is a multi-specific drug transporter in the kidney with numerous substrates, including pharmaceuticals, endogenous metabolites, natural products, and uremic toxins. Here, we show that OAT1 regulates levels of gut microbiome-derived metabolites. We depleted the gut microbiome of *Oat1* knockout and wildtype mice and performed metabolomics to analyze the effects of genotype (knockout vs wildtype) and microbiome depletion. OAT1 is an in vivo intermediary between the host and the microbes, with 40 of the 162 metabolites dependent on the gut microbiome also impacted by loss of *Oat1*. Chemoinformatic analysis revealed the altered metabolites (e.g., indoxyl sulfate, p-cresol sulfate, deoxycholate) had more ring structures and sulfate groups. This indicates a pathway from gut microbes to liver Phase II metabolism, to renal OAT1-mediated transport. The idea that multiple gut-derived metabolites directly interact with OAT1 was confirmed by in vitro transport and magnetic bead binding assays. We show that gut microbiome-derived metabolites dependent on OAT1 are impacted in a chronic kidney disease (CKD) model and human drug-metabolite interactions. Consistent with the Remote Sensing and Signaling Theory, our results support the view that drug transporters (e.g., OAT1, OAT3, OATP1B1, OATP1B3, MRP2, MRP4, ABCG2) play a central role in regulating gut microbe-dependent metabolism, as well as inter-organismal communication between the host and microbiome.

2.2 INTRODUCTION

Organic anion transporter 1 (OAT1/SLC22A6, originally described as NKT) is a multispecific drug transporter localized to the basolateral membrane of the kidney proximal tubule. OAT1 is involved in the uptake of multiple classes of drugs (e.g., antibiotics, antivirals, NSAIDs, diuretics), endogenous metabolites, toxins, antioxidants, and natural products from the blood into the proximal tubule cell, where they can then be excreted into the urine by apical efflux transporters, re-introduced to the bloodstream, or metabolized by the cell [9, 52, 77, 98-103]. While OAT1 favors the transport of organic anions, it can also handle several structurally different small molecules, including some cations and zwitterions [77]. To date, the overwhelming majority of research interest in OAT1 has been related to its role in the clearance of drugs, as the Food and Drug Administration (FDA) and other global regulatory agencies have recommended that novel drug entities be tested for OAT1 interaction due to potential drug-drug interactions (DDIs) occurring at the site of the transporter [3, 104]. Despite the pharmaceutical relevance of this transporter, recent studies have highlighted a further role of OAT1 and other drug transporters in endogenous metabolism in the context of several observations [6, 54, 73, 105]. Many of the metabolites that have received clinical attention-- such as 4-ethylphenyl sulfate, p-cresol sulfate, and indoxyl sulfate—are organic anions and well known OAT1 substrates and have been associated with the gut microbiome.

The gut microbiome plays an important role in the endogenous host metabolism by producing several important metabolites. While they are generated within the host, they are often the products of complex interactions between the host and the commensal microbes residing in the gut. Gut microbiome-derived metabolites include short chain fatty acids, secondary bile acids, aromatic amino acid derivatives, polyamines, and several others [16, 106, 107]. The full repertoire of gut microbiome-derived metabolites and how they are generated remains

25

incomplete, but it is clear that these metabolites play important signaling roles in both healthy and disease states [108]. For example, gut bacterial metabolites have been shown to influence the immune system [109]. Furthermore, inflammatory bowel disease (IBD), cardiovascular disease, and chronic kidney disease (CKD) are associated with microbial dysbiosis, which can lead to abnormal levels of serum metabolites [17, 110-113]. CKD, in particular, is associated with increases in circulating gut-derived uremic toxins due to diminished glomerular and tubular renal function [114-116]. OAT1 is known to be critical for the transport of many of these metabolites into the proximal tubule. These gut microbiome-derived metabolites include indoxyl sulfate, p-cresol sulfate, hippurate, and other metabolites and signaling molecules that are organic anions, suggesting there may be an important role for OAT1 in the mediation of host-microbiome interaction [52, 53, 114]. Considering the wide array of substrates transported by OAT1, the competition between drugs, toxins, and gut-derived metabolites at the site of the transporter could also lead to drug-metabolite interactions (DMI)--especially in patients with CKD, who are likely to be taking multiple drugs to treat symptoms associated with comorbidities.

The Remote Sensing and Signaling Theory (RSST) addresses how OAT1, along with other solute carrier (SLC) and ATP-binding cassette (ABC) "drug" transporters, plays a major role in homeostasis of many small molecules, including rate-limiting metabolites, signaling molecules, antioxidants, gut microbe-derived products, vitamins and cofactors [11]. These other SLC and ABC transporters include organic anion transporter 3 (OAT3, SLC22A8), organic anion transporting polypeptide 1B1 (OATP1B1, SLCO1B1), organic anion transporting polypeptide 1B3 (OATP1B3, SLCO1B3), multidrug resistance protein 2 (MRP2, ABCC2), multidrug resistance protein (MRP4, ABCC4), ATP-binding cassette G subfamily 2 (ABCG2, BCRP). The RSST proposes a complex adaptive system of drug transporters, drug metabolizing

enzymes, nuclear receptors and kinases that regulates endogenous metabolism through transport, metabolism and conjugation of small molecules with signaling roles between remote organs (e.g., gut, liver, kidney, brain) and multi-organismal systems (e.g., gut microbe-host, mother-fetus) [7, 8]. Drug transporters in the SLC22, solute carrier organic (SLCO), and ATP-binding cassette C family (ABCC) families and their multi-specificity (ability to handle structurally different organic anions) are central to the RSST and have been identified as important hubs in a cross-tissue co-expression network, suggesting a principal role in endogenous metabolism at multiple scales (organism to organ to organelle) [6, 10, 11]. The RSST has mainly been explored through the lens of interorgan communication (e.g., organ crosstalk), but an important, understudied aspect is interorganismal communication between the host and the bacterial species in the gut, which is possibly mediated by the role of kidney OAT1 and potentially other drugs transporters in modulating gut microbiome host interactions [117].

While OAT1 and OAT3 in the kidney are hypothesized to be central to interorganismal communication via gut-derived metabolites, the hepatic transporters, OATP1B1 and OATP1B3 along with various ABC transporters, such as ABCC2, ABCC4, and ABCG2 are also thought to contribute to the transport of gut microbiome derived compounds [118, 119]. Furthermore, drug-metabolizing enzymes are central to the metabolism and conjugation of these compounds along the gut-liver-kidney axis [120-122]. Nevertheless, much of the data supporting these notions is from in vitro rather than in vivo studies.

In this work, we focused on the in vivo role of a single multispecific kidney "drug" transporter, OAT1, in the regulation of gut-derived metabolites and the metabolic pathways involving these molecules. We first established the efficacy of the *Oat1* knockout mouse model by demonstrating in vivo alterations in the handling of a well-studied OAT1 substrate, as

evidenced by changes in levels in the blood and urine. We then focused on depleting the gut

microbiome of these mice and their wildtype counterparts. Gnotobiotic mice have frequently

been used as a model to understand the impact of gut microbes, but while they are generally

healthy, they are difficult to compare to wildtype mice due to differences in their genetic

backgrounds, as well as issues related to development, immune defects, and energy metabolism

[123, 124]. This is a key concern in the context of OAT1, since OAT1 is an alpha-ketoglutarate

antiporter and thus directly linked to aerobic metabolism, upon which the kidney proximal tubule

almost exclusively depends [53, 125]. Thus, like many in the field, we chose to employ antibiotic

treatment to deplete the gut microbes [126]. We then assessed the impact of loss of the *Oat1*

gene (*Oat1* knockout versus wildtype) and microbiome depletion on biochemical pathways and

applied chemoinformatics approaches to characterize the altered metabolites. To support our in

vivo findings, we performed in vitro transport assays and employed a magnetic bead binding

assay to evaluate mechanistic relationships between gut-derived metabolites and OAT1.

Furthermore, we established clinical and disease relevance of our results by showing that the gut

microbiome-derived metabolites that are OAT1-dependent are significantly affected in a clear

example of human DMI and in a rodent CKD model. Our results indicate that OAT1 plays a

surprisingly important in the handling of a number of gut-derived metabolites and, consistent

with the RSST, mediates inter-organismal communication between the host and gut microbes, in

large part by regulating the circulating levels of these compounds **(Figure 2.1)**.

2.3 RESULTS
2.3.1. Clearance of OAT1-interacting compounds altered in vivo in knockout mice
We first characterized our *Oat1* knockout mice and their wildtype counterparts by

measuring the levels of Tc-99m-mercaptoacetyl-triglycine (MAG3) in the urine (bladder) and the

blood. Tc-99m MAG3 is a probe compound used in the assessment of renal function that is

nearly entirely eliminated by tubular secretion. Previous studies have demonstrated that Tc-99m MAG3 is a rOAT1 substrate in vitro and that its uptake is inhibited by classic OAT1 inhibitors, such as PAH and probenecid [127]. Clinically, results in humans have shown that MAG3 levels in the blood were elevated following treatment with PAH and probenecid [128]. These observations were supported by assays using HEK293 cells expressing human OAT1, which showed that the protein is involved in the uptake of Tc-99m MAG3 [128].

We evaluated whether the *Oat1* knockout and wildtype mice had different clearance patterns with this well-established OAT1 substrate (**Figure 2.2A**). Tc-99m MAG3 was administered to the mice via tail-vein injection and its levels were monitored over the course of 30 minutes using a gamma camera. We mainly focused on the bladder, as Tc-99m MAG3 quickly passes through the kidneys into the urine. We found that the wildtype bladders reached their maximal levels of Tc-99m MAG3 more quickly than the *Oat1* knockout mice following direct injection of the probe compound (**Figure 2.2B**). These results were further supported by post-mortem gamma counts scaled to weight, which showed that for 4/5 pairs, the bladder levels of Tc-99m MAG3 were higher in the wildtype mice (**Figure 2.2C**), and the blood levels were higher in the knockout mice (**Figure 2.2D**). Given that OAT1 is considered the rate-limiting step for excretion of many organic anions into the urine, our results support the usefulness of the knockout mice as in vivo models for analyzing OAT1-related function.

2.3.2. Gut microbiome was depleted in Oat1 knockout and wildtype mice

Previous in vivo and in vitro experiments have shown that OAT1 has several putative gut-derived substrates, such as indoxyl sulfate, p-cresol sulfate, and hippurate [53, 73, 129]. While these metabolites are useful in understanding a potential role for OAT1 in regulating circulating levels of gut microbe-derived metabolites, the results were collated from multiple past experiments performed under a variety of conditions and not designed to evaluate the in

vivo contribution of gut microbiome and renal OAT1 to host systemic metabolism. In this work, we depleted the gut microbiome in both wildtype and knockout mice through the administration of an antibiotic cocktail (ampicillin, vancomycin, neomycin, and metronidazole [AVNM]). The AVNM antibiotic cocktail has been established as an effective method of depleting the gut microbiome and, in contrast to germ-free mice which can develop metabolic problems that can lead to obesity [124], seemed less likely to confound the essential role of OAT1 in kidney aerobic metabolism [53, 125] and the tendency to hepatic steatosis seen in approximately 24 month-old *Oat1* knockout mice [54]. The AVNM cocktail was administered through a vehicle control in the drinking water for 4 weeks [126]. Following the administration of the cocktail, depletion of the gut microbiome was confirmed via metagenomic analysis of the feces, which showed a significant decrease in the number of operational taxonomic units (OTUs) for AVNM-treated mice groups (**Figure 2.3A).** The global metabolic profiles of all animals were separable by linear discriminant analysis (**Figure 2.3B)**, and several well-established gut-derived metabolites were significantly decreased in the serum of AVNM-treated animals, regardless of genetic background **(Figure 2.3, C-F).** Furthermore, quantitative PCR (qPCR) using 16S primers for *Eubacteria* also showed a significant decrease in gut microbes (**Supplementary Figure 2.S1**).

2.3.3. Loss of Oat1 and microbiome depletion significantly affect the levels of over 200 metabolites

Since OAT1 is localized to the basolateral (blood-facing) side of the proximal tubule, its absence directly affects the circulating levels of metabolites in the serum. To identify these compounds, we performed a two-way ANOVA to determine the individual impact of genotype (*Oat1* KO vs WT), where a metabolite was considered altered if it had an FDR-corrected p-value below 0.05. Although metabolomics of the *Oat1* KO has been previously performed, this is the

first time nearly a thousand compounds were measured, and here, apart from including the effects of microbiome depletion, we also focus on all altered metabolites, not just elevated metabolites [42, 52-54, 73]. Global metabolic profiling detected a total of 964 metabolites in the volume-adjusted serum samples collected from these mice. Based on the Metabolon grouping of these metabolites, this analysis covered 10 biochemical superpathways (e.g., Lipid, Amino Acid, Xenobiotic, etc.) and 109 biochemical subpathways (e.g., Primary Bile Acid Metabolism, Tryptophan Metabolism, Benzoate Metabolism), with each metabolite belonging to one superpathway and one subpathway. We identified 103 significantly altered metabolites due to the absence of OAT1 (**Supplementary Table S1**), including several metabolites that are known to directly interact with OAT1, such as pyridoxate, indoxyl sulfate, and p-cresol sulfate [130]. We then performed an enrichment analysis of these metabolites and found that over twenty subpathways were altered, with enrichment values of 1 or greater, indicating an outsized effect. Benzoate Metabolism and Fatty Acid Metabolism (Acyl Glycine) were among the most significantly affected pathways (**Figure 2.4A**).

Having established depletion of the gut microbiome with AVNM treatment (by decreased OTUs and qPCR), we then analyzed how this impacted the serum metabolome of the microbiome-depleted mice. We found that 162 metabolites were significantly altered in the serum of the microbiome-depleted mice (**Supplementary Table 2.S2**). Thus, microbe depletion has a direct impact on over a hundred metabolites, including several metabolites that have previously been established as gut-derived, like cinnamoylglycine, indolepropionate and others [16]. Subpathway analysis revealed that Benzoate Metabolism, Phospholipid Metabolism, and Tyrosine Metabolism were among the most altered subpathways, and over twenty subpathways had enrichment values of 1 or greater (**Figure 2.4B**). Given that there is no current consensus on

the full range of commensal gut bacteria-derived metabolites that enter the host circulation, we interpreted these metabolites to be products of gut microbiome-associated metabolism with the understanding that for some metabolites, the levels in the serum may be due to complex interactions between bacterial species themselves and, upon entry into the host circulation, indirect effects on host metabolic pathways that likely include complex feedback and/or feedforward loops, with secondary bile acid metabolism being a good example [131].

2.3.4. Deoxycholate levels depend on an interaction between OAT1 and the gut microbiome

We then explored the statistical interaction between the two independent variables: loss of *Oat1* and microbiome depletion. Only 3 metabolites (2-amino-p-cresol sulfate, deoxycholate, docosahexaenoylcarnitine [C22:6]) were impacted by the interaction between the variables, which implies that genotype and treatment together influence only few metabolites compared to the individual effects of genotype versus treatment (**Figure 2.4C**). Docosahexaenoylcarnitine (C22:6) and 2-amino-p-cresol sulfate are poorly characterized, but deoxycholate has a well-established signaling role, suggesting an important role for OAT1, together with gut microbes, in regulation of this important bile acid signaling molecule [132-134].

2.3.5. Pathway and chemoinformatic analyses of the 40 metabolites affected by both loss of Oat1 and microbiome depletion

We then aimed to identify the overlap between the 103 metabolites altered by loss of *Oat1* and the 162 metabolites affected by microbiome depletion. We found 40 metabolites (**Figure 2.5A**) that satisfied both criteria and found that some subpathways, particularly Benzoate Metabolism with 11 compounds and Food Component/Plant with 5 compounds, were markedly affected in the overlap (**Figure 2.5B**). These 40 compounds could be separated into 4 distinct groups: Group 1 (elevated by loss of *Oat1* AND elevated by microbiome depletion); Group 2 (elevated by loss of *Oat1* AND decreased by microbiome depletion); Group 3

32

(decreased by loss of *Oat1* AND elevated by microbiome depletion); Group 4 (decreased by loss of *Oat1* AND decreased by microbiome depletion) (**Figure 2.5C**).

The metabolites we were most interested in were those that were in Group 2 (elevated in the *Oat1* knockout mice and decreased due to microbiome depletion), as these are likely OAT1 substrates that are generated by the gut microbiome. Of the 40 metabolites, 22 fell into this group, including indoxyl sulfate and p-cresol sulfate. We were also interested in the 9 metabolites in Group 1 (increased in the *Oat1* knockout mice and increased due to microbiome depletion), as microbiome depletion can also lead to increases in specific metabolites by reducing the species that metabolize these compounds. Finally, the last two groups were more difficult to interpret from the OAT1 perspective, as there is no clear renal physiological mechanism for their decreases in circulation; nonetheless, there were 7 metabolites in Group 3 (decreased in the *Oat1* knockout and decreased by microbiome depletion) and 5 metabolites in Group 4 (decreased in the *Oat1* knockout and increased by microbiome depletion). We then aimed to structurally characterize the 31 metabolites with known chemical structures.

Chemoinformatics analyses can shed light on sets of molecular properties that help define particular groups of metabolites altered by a biological experiment. In this case, for example, we were most interested in Group 2 (the metabolites that were both elevated due to loss of OAT1 and decreased after microbiome depletion) –as these were not only the largest group but also most likely to be gut microbe-derived organic anion metabolites transported by OAT1 in vivo. This could yield a kind of "signature" of metabolites that originate in the gut microbiome and then follow the gut-kidney or gut-liver-kidney axis to OAT1 in the renal proximal tubule cells.

To this end, we first calculated molecular properties for the 783 compounds with valid chemical structures. We performed linear discriminant analysis of the 31 compounds (Group 1: 5 metabolites with structures; Group 2: 17 metabolites with structures; Group 3: 3 metabolites with structures, Group 4: 6 metabolites with structures) that had available chemical structures and observed clear separation between the groups (**Figure 2.5D**). We then analyzed the weights of the top two linear discriminant analysis axes, which together, explained over 95% of the variance. Among the most influential variables were number of aromatic bonds and number of sulfate groups (**Figure 2.5, E and F**). Again, we were most interested in the 17 compounds with chemical structures that were elevated by loss of *Oat1* and decreased by microbiome depletion (Group 2), and these compounds tended to have a higher number of aromatic bonds. The gut microbiome is known to handle a number of aromatic compounds, such as tryptophan and tyrosine derivatives. With respect to sulfation, 8 of the 17 metabolites in Group 2 featured a sulfate group. This was especially interesting because only one sulfated metabolite was present in the other three groups combined (**Figure 5F**). Nevertheless, this is consistent with the notion that certain diet-derived compounds are metabolized by the gut microbiome before being "tagged" via sulfation by the host for excretion, primarily through the urine [135, 136].

2.3.6. Gut-derived metabolites interact with human OAT1 in cell-based transport assays

While the knockout mouse model is critical for establishing potential in vivo OAT1 substrates, the complex physiology in vivo could lead to alterations caused by a factor other than loss of OAT1 function. To evaluate a mechanistic interaction between metabolites and OAT1, we performed in vitro cell-based transport assays. Competitive inhibition assays were carried out for deoxycholate, indolepropionate, 4-hydroxycinnamate, 2-hydroxyphenylacetic acid, and 5-hydroxyindoleacetate, which are all thought to be gut-derived metabolites. In the statistical analysis of the serum metabolome, deoxycholate was affected by genotype-treatment interaction,

2-hydroxyphenylacetic acid was affected by genotype, indolepropionate was affected by

treatment. Although 4-hydroxycinnamate and 5-hydroxyindoleacetate were not impacted by

either variable, for experimental evaluation, they were included because they are derived from

cinnamate and indole, respectively. Each metabolite showed comparable inhibition when OAT1

was treated with probenecid, the prototypical inhibitor of OAT1 activity. The relatively low

IC50s (<115 uM) suggest that these metabolites interact with OAT1 with high affinity (**Figure

2.6**).

2.3.7. A magnetic bead binding assay shows direct physical OAT1 interaction with gut-derived metabolites

To further evaluate our results, we also employed a potentially novel magnetic bead

binding assay to analyze 20 gut-derived metabolites, a number of which were measured in the

serum metabolomics (**Figure 2.7A**). The strength of this is that the assay requires relatively low

amounts of OAT1 protein compared to that used for other methods (e.g., fluorescence

polarization technique). Using this method, we surveyed metabolites mainly known to derive

from tryptophan and tyrosine, some of which were measured in our in vivo experiments and have

been evaluated in transport assays (**Table 2.1**). Overall, we found that 15 of the metabolites

resulted in a significant shift, further supporting a direct interaction between the gut microbe-

derived metabolites and OAT1 (**Figure 2.7B**). Taken together with the cell-based in vitro

transport assays that also support a direct interaction of the gut-derived metabolites with OAT1,

as well as the fact that many of the interacting metabolites were from Group 2 (elevated in *Oat1*

knockout and decreased by microbiome depletion), the data support the in vivo involvement of

OAT1 in the regulation of this group of gut microbe-derived metabolites.

2.3.8. The in vivo OAT1-dependent, gut microbe-derived metabolites overlap with those impacted by a chronic kidney disease model (5/6th nephrectomy)

We then aimed to contextualize our findings by comparing our results to metabolomics data previously generated by our group in a rodent 5/6 nephrectomy model of CKD [114]. This model is thought to capture aspects of progressive CKD, and renal capacity is dramatically reduced over time [114]. In these experiments, plasma was collected from animals who had undergone a 5/6 nephrectomy and their healthy controls, and the relative levels of hundreds of metabolites were measured. In the comparison between the nephrectomized and healthy animals, many of the elevated metabolites have been shown to include numerous uremic solutes or uremic toxins [114]. However, the nature of their diminished clearance remains unclear since this model of renal insufficiency impacts both tubular and glomerular function. By comparing the metabolites elevated in the 5/6 nephrectomy to the 40 metabolites that are gut-derived and altered in the *Oat1* KO, we were able to identify metabolites that are likely impacted by diminished proximal tubule function, as that is where OAT1 is primarily expressed. Thus, seven metabolites (indoxyl sulfate, p-cresol sulfate, phenylacetylglycine, 4-ethylphenyl sulfate, 3-methylhistidine, N-acetylserine, 2-isopropylmalate) are likely uremic solutes or uremic toxins transported by OAT1 and produced by the gut microbiome (**Table 2.2**).

2.3.9. Gut-derived metabolites transported by OAT1 are involved in human drug-metabolite interactions

Having established that 40 metabolites are likely OAT1-mediated and produced by the gut microbiome in a mouse model, we then aimed to understand the clinical relevance of these results. A recent metabolomics study by our group analyzed the plasma and urine of healthy volunteers before and after probenecid treatment and identified dozens of unique short-term drug-metabolite interactions [137]. While that study did not concern itself with gut microbe-derived metabolites, we were able to reanalyze that data in the context of the new data in this

study. Thus, we performed an overlap of the metabolites implicated in the present *Oat1* KO microbe-depleted mouse study with those significantly elevated in the plasma and significantly decreased in the urine of humans treated with probenecid (**Figure 2.8, Supplementary Table 2.S3**). Over a quarter of the OAT1-transported gut-derived metabolites from the current mouse study (11/40) were significantly elevated in the plasma of the probenecid treated humans, including indoxyl sulfate, p-cresol sulfate, and other compounds. When we compared the 40 metabolites to those significantly decreased in the urine of probenecid treated humans, we found that half the metabolites (20/40) were present in both lists. Interestingly, 4-ethyphenyl sulfate, a compound associated with autism, was present in this overlap, along with others [138]. When all three lists were overlapped, eight metabolites were present, with most being sulfated organic anions (**Figure 2.8**). These gut microbiome-derived metabolites are dependent on OAT1 and are also involved in DMI.

2.4 DISCUSSION

The loss of *Oat1* and the depletion of the gut microbiome (**Figure 2.3**), separately and together, have major effects on systemic metabolism (**Figure 2.4, Supplementary Table 2.S1, Supplementary Table 2.S2**). The genetic knockout of *Oat1* primarily leads to elevated metabolites, presumably because they are no longer able to enter the proximal tubule and must remain in the blood, whereas the depletion of the gut microbiome mainly leads to lower circulating levels of gut-derived metabolites by eliminating the species that synthesize or modify the compounds. Both conditions together--knockout of *Oat1* and gut microbe depletion--provide perhaps the deepest glimpse to date of how the renal organic anion transport system works together with the gut microbiome to regulate systemic levels of many well-known metabolites and signaling molecules (**Figure 2.1**).

Most impressive were the effects of gut microbiome depletion on metabolites elevated due to loss of *Oat1* (compared to the wild type with a normal microbiome). Many of these metabolites elevated in the *Oat1* knockout were significantly, and sometimes markedly, decreased after gut microbiome depletion (**Figure 2.5**). Based on this in vivo data, and the evidence presented here for a direct in vitro interaction between a number of these metabolites and OAT1 in transport and binding assays (**Figures 2.6 and 2.7**), it is highly likely that the largest fraction of these metabolites derived from the gut microbes and are regulated by OAT1, which is important given the important roles these metabolites play in the immune response and other key physiological systems [109].

That said, it should be noted that there were also instances of metabolites significantly decreased by loss of *Oat1* as well as metabolites significantly increased by microbiome depletion. The interpretation of these changes is less clear but may be due to indirect effects such as elevation in the knockout of another metabolite transported by OAT1 that inhibits the synthesis of the decreased metabolite.

Overall, we identified 40 metabolites (31 with chemical structures) (**Supplementary Table 2.S3**) that were significantly impacted by both loss of *Oat1* and gut microbiome depletion, with 22 of those compounds being elevated due to knockout and decreased due to antibiotic treatment. These included derivatives of tryptophan and tyrosine. Chemoinformatics analyses of this group of metabolites (e.g., elevated in the *Oat1* knockout and decreased after gut microbiome depletion) revealed that these metabolites tended to have aromatic rings and more sulfate groups, which is generally consistent with known molecular properties of OAT1 substrates (**Figure 2.5**) [74, 76]. The high fraction of sulfated metabolites is particularly interesting and likely indicative of interaction of gut-derived metabolites with sulfotransferases

in the liver before transport into the kidney proximal tubule by OAT1--examples, similar to that

of indoxyl sulfate, of the conjunction of remote inter-organismal communication and organ

crosstalk [139]. Indeed, 13 of the 22 putative gut-derived OAT1 substrates in Group 2 are

sulfated compounds. However, since other OATs, such as OAT3 [72], have a strong preference

for steroid sulfates, thus, it may be the context in which the sulfated context are presented that

determines OAT1 interaction. Our data suggests that a one- or two-ringed structure with a sulfate

may be preferred by OAT1.

The strongest candidates for in vivo OAT1-transported gut microbiome-derived

metabolites would seem to be those that are: 1) elevated in the plasma of *Oat1* knockout mouse;

2) decreased by gut microbiome depletion; and 3) shown to directly interact with the transporter.

Thus, to further analyze the in vivo metabolomics results, a number of the identified metabolites

(and others that have been suggested to be gut-derived) were tested in vitro for interaction with

human OAT1 over-expressed in cells. These compounds displayed IC50 values in transport

assays that indicate a strong interaction with the transporter. Additional support was gained from

a magnetic bead binding assay demonstrating that a fluorescent prototypical OAT1 substrate was

displaced by a number of gut-derived metabolites (**Table 2.1**). While the transport assays are

traditionally used to determine interaction, the magnetic bead binding assay provides further

context for the nature of the interaction and allows for more rapid screening of small molecules.

Our results are also highly relevant to disease and clinical settings. We found that many

of the metabolites implicated in our study were elevated in 5/6 nephrectomy rodent models of

CKD, indicating that these gut-derived and OAT1-mediated metabolites can be altered in the

setting of diminished renal function [114]. Among the metabolites present in both studies were

indoxyl sulfate, p-cresol sulfate, and 4-ethylphenyl sulfate, further supporting the view that OAT1 and the gut microbiome are jointly involved in the generation/handling of uremic toxins.

We also addressed the important clinical issue of drug-metabolite interactions. A previous study by our group analyzed the drug-metabolite interactions caused by the drug, probenecid [137]. While that clinical study did not focus on the gut microbiome, given the results from the present mouse study, we investigated whether drug-metabolite interactions with an OAT-inhibiting drug (probenecid) had a major impact on the disposition of gut-derived metabolites. Indeed, eight metabolites were elevated in the plasma of probenecid-treated humans, decreased in the urine of probenecid-treated humans, and altered by both loss of *Oat1* and microbiome depletion (**Figure 2.8**). Furthermore, 20 of the 40 metabolites were decreased in the urine, while 11 of the 40 were elevated in the plasma. OAT1 is a major transporter of antibiotics, antivirals, NSAIDs, diuretics, and other common drugs [5]. Our analysis suggests that gut microbiome dependent drug-metabolite interactions (DMI) at the level of OAT1 could be quite widespread [140]. This requires further study.

Taken together, our results indicate that OAT1 is a crucial intermediary between the host and the microbes, with 40 metabolites of the 162 metabolites decreased by gut microbiome depletion being presumably influenced by OAT1. This suggests that as much as ~25% of microbiome-influenced metabolism may be modulated by OAT1. However, we must also note that the metabolomics platform we used is biased towards compounds that have already been identified and are likely relevant in clinical or research settings. Consistent with previous work, we too observed dozens of metabolites decreased by microbe depletion that were the products of hepatic metabolism [53, 107].

The communication between the host and the gut microbes is complex, but it is clear that these two entities have co-evolved over time to develop a symbiotic relationship from the perspective of metabolism. This is evidenced by the gut-derived metabolites, which cannot be produced by the host alone, that have important signaling roles, such as nuclear receptor and G-protein coupled receptor activation [141, 142]. Among the implicated metabolites, deoxycholate, indoxyl sulfate, indolepropionate, and others have been shown to have important signaling roles [143-146]. While the signaling roles of gut-derived metabolites with respect to target proteins is an important field of research, our results indicate that much more attention needs to be paid to the proteins that regulate their levels in biofluids, as well as in tissues and along organ axes (e.g., gut-liver-kidney, gut-brain). These proteins, which include OAT1, are important avenues by which the host and its gut microbes interact, as they control the bioavailability of signaling molecules.

SLC22 family members (e.g., OATs, organic cation transporters (OCTs), organic cation and zwitterion transporters (OCTNs)) and other multi-specific drug transport proteins (e.g., ABCG2, ABCC2), are hubs in a recently proposed Remote Sensing and Signaling Network [11]. The Remote Sensing and Signaling Theory emphasizes the importance of the adaptive network of multi-specific transporters, enzymes, and nuclear receptors--working together with oligo-specific and monospecific proteins--in the optimization of the levels of numerous metabolites in cells, tissues, organs and bodily fluids, such as blood, cerebrospinal fluid and urine [6-9, 11, 122]. These proteins have been extensively studied from the pharmaceutical perspective, but their ability to handle structurally diverse molecules is perhaps most important from the perspective of endogenous and gut-derived metabolites. The theory mainly serves as a framework to describe inter-organ communication, such as between the gut-liver-kidney (where

most drug transporters and DMEs are highly expressed) to maintain and re-establish homeostasis of key metabolites, signaling molecules, antioxidants and other small molecules with "high informational content"[9]. While we focused on the individual role of OAT1 in this work, it is likely that the combined network of transporters and enzymes, including CYPs, SULTs, UGTs, ABCCs, and SLCOs, also contribute to the handling of gut-derived products. Indeed, it has been shown that the SLCOs transport the gut-derived secondary bile acids [48].

One aspect of the Remote Sensing and Signaling Theory that has received less attention is inter-organismal communication between the host and the gut microbes [147]. In depth study is likely to have important clinical ramifications--for instance, in understanding the role of gut microbe-derived uremic toxins in the aberrant metabolism of chronic kidney disease [114, 116, 139]. If we treat the gut microbiome as an independent organ, it can be considered to express thousands of transporters and enzymes, many of which exhibit broad substrate specificity and are commonly involved in movement of metabolites, and the synthesis of small molecules as well as their hydrolysis, reduction, or removal of conjugated groups [148]. It is established that the substrates and products of the enzymatic reactions occurring in bacterial species overlap with that of the transporters and enzymes in the host, enabling inter-organismal communication via multi-, oligo-, and mono-specific enzymes that are part of the Remote Sensing and Signaling network. Furthermore, there is evidence that the gut microbiome and its metabolites can impact transporter and enzyme expression. In renal disease, it has been shown that gut-derived metabolites have an impact on the expression of several drug-metabolizing enzymes in the kidney [149]. The presence of gut microbes has been shown to alter the expression of hepatic DME genes in mice [150]. Indoxyl sulfate, an important gut-derived uremic toxin, has also been shown to regulate the level of OAT1 expression through AHR activation, and along with other

42

aspects of this pathway, has been interpreted as an experimentally verified example of the Remote Sensing and Signaling Theory [144, 151].

Establishing that the communication between the host and the gut microbes is so strongly mediated by renal OAT1 opens up many questions. For example, it has been shown that several drugs--including statins and ACE inhibitors--have an impact on the composition of the gut microbiome and, by implication, the levels and composition of gut microbe-derived metabolites in the host. However, the mechanism(s) are unclear. Since these and other drugs impacting the gut microbiota are OAT1 substrates, it is possible that these changes could alter competition of the drug with metabolites at the level of the transporter. There is now good evidence in humans to support this kind of drug-metabolite interaction at the level of OAT1 [137].

In summary, our studies indicate that multispecific transporters and enzymes combine with the gut microbiome to regulate circulating levels of key metabolites, including those with signaling capabilities. As these effects need not be limited to OAT1 in the kidney, it is worthwhile to perform similar analyses with multi-specific transporters (e.g., SLCO and ABCC families) expressed in the kidney, liver, intestine, and other organs. A much more complex and clinically actionable picture of the regulation of microbiome-dependent host metabolism is likely to emerge. This can be particularly useful for studying drug-metabolite interactions.

2.5 METHODS AND MATERIALS

2.5.1. Animals

Adult male knockout and wild type mice were housed in a 12-hour light-dark cycle and allowed ad libitum access to food and water. *Oat1* knockout mice were generated and maintained as previously described [42]. Feces were collected from mice weekly in sterile 2 mL centrifuge tubes, flash frozen, and stored at -80º C. Animals were sacrificed by CO2 inhalation and blood

samples were extracted from mice by cardiac puncture. Serum was extracted and samples were flash frozen and stored at -80º C until further analysis.

2.5.2. Tc-99m-MAG3 imaging

Live imaging experiments were performed with the In Vivo Imaging Shared Resource at the Moore's Cancer Center at the University of California San Diego. Mice were transported from the vivarium to the In Vivo Imaging Shared Resource. 100 uCi of Tc-99m-mercaptoacetyl-triglycine (MAG3) was injected by tail vein into mice prior to imaging. Adult, male knockout (n=5) and wildtype mice (n=5) were imaged in five, separate pairs containing one knockout mouse and one WT mouse each. Mice were initially weighed and anesthetized (2% isoflurane, 200 mL/min flow of O_2). Mice were placed on their backs on the high-resolution gamma imager (gamma imager; Biospace) fitted with a high-resolution low-energy collimator. For thirty minutes, time-activity curves were collected from the kidneys, liver, and bladder. Following imaging, blood, urine, kidneys, liver, and spleen were isolated and weighed from each mouse for radioactive assessment using a gamma counter.

2.5.3. Microbiome depletion protocol

Over a four-week period, mice were given a 125 mL antibiotic cocktail or vehicle control in place of drinking water. The cocktail consisted of 1 mg/mL of neomycin sulfate (Fisher Scientific, BP-2669-25), ampicillin (Sigma-Aldrich, A9518-100G), metronidazole (Alfa Aesar, H60258), 0.5 mg/mL of vancomycin (Alfa Aesar, J62790) and 3.75 mg/mL of Kool Aid grape drink powder (Kraft-Heinz Foods Company). The Kool Aid encouraged consumption of the cocktail. Antibiotic cocktails were replenished every other day on Monday, Wednesday, and Friday. New solutions were passed through a 0.22μm cellulose acetate sterilizing filter (Corning, 430517). Bottles of antibiotic cocktail were also wrapped in foil to prevent light damage. The weights of mice and the amount of antibiotic cocktail consumed were monitored over the

treatment period as markers of consumption. Following a one week decrease in both weight and consumption, mice returned to near their original weights (**Supplementary Figure 2.S1**).

2.5.4. Assessment of gut microbiome depletion

For 16S variable region sequencing, murine fecal samples from Week 0 and Week 4 (end of treatment timepoint) were extracted using the the MagMax™ Microbiome Ultra Nucleic Acid Isolation Kit (Thermo, A42357) according to the manufacturer's instructions. Variable (V4) regions of 16S SSU rRNA were amplified using 515F-806R primers according to the protocol described in (http://earthmicrobiome.ucsd.edu/protocols-and-standards/16s/). 16S sequencing was performed by the Institute for Genomic Medicine (IGM) UC San Diego. Resulting files were analyzed using the web-based QIITA tool [152].

For the qPCR, feces total RNA was extracted using the RNEasy PowerMicrobiome Kit (Qiagen, 26000-50) according to the manufacturer's instructions. A cDNA library of the RNA extracted was created using the SUPERSCRIPT III kit (Invitrogen, 18080-044) with random hexamers (Invitrogen, 48190011) as primers according to the guidelines of the manufacturer. RNA and cDNA were quantified using a Nanodrop 1000 (Thermo Scientific 2353-30-0010) and were subsequently used to load equal amounts of cDNA for the qPCR performed. Each well in the qPCR plate contained 20 ng of cDNA from fecal RNA. Duplicates of each sample containing *Eubacteria* primers were utilized, targeting the universal 16S rRNA gene that captures a majority of bacteria [153]. 1.KAPA SYBR FAST® Universal kit was utilized with an accompanying protocol (Roche, KK4608).

A default 16S metagenomics workflow was run in Qiita under Qiime version 1.9. Raw reads were demultiplexed and trimmed, and Operational Taxonomic Units (OTUs) were closed-reference picked using SortMeRNA v2.1 with a 97% sequence similarity minimum. OTUs were assigned taxonomies from the GreenGenes 16S rRNA database, version 13_8, and tabulated into

feature and reference tables. All analyses of the feature and reference tables were performed in the Qiita platform and graphed using the Python package, Seaborn.

2.5.5. Metabolomics analysis of wild type and knockout mice (untreated and treated)

Serum samples were shipped on dry ice to Metabolon (Durham, NC) for preparation and metabolomics. Per Metabolon, proteins were removed from the serum and five fractions were generated for different mass spectrometry methods. Each sample passed quality control compared to well-characterized controls and was analyzed by Ultrahigh Performance Liquid Chromatography-Tandem Mass Spectroscopy (UPLC-MS/MS). Peaks were identified and associated with defined compounds based on retention time, mass-to-charge ratio, and MS/MS spectral data. Quantification of peaks was performed using area-under-curve.

2.5.6. Physicochemical analysis of metabolites

Two-dimensional chemical structures were obtained from 783 of the measured metabolites by their Pubchem IDs. Seventy-seven one-dimensional descriptors were calculated for each metabolite using ICM Molsoft-Pro (San Diego, CA). These molecular properties were then trimmed to eliminate heavily correlated features using a Spearman's correlation cutoff of 0.90. Linear discriminant analysis was then applied and visualized using the Python packages Seaborn and sci-kit learn. Visualizations of the chemical structures were performed using RDKit.

2.5.7. In vitro OAT1 transport assays

Human embryonic kidney (HEK) cells stably overexpressing human OAT1 (SOLVO Biotechnology) were used for in vitro inhibition assays. Cells were maintained in DMEM supplemented with 10% FBS, 1% penicillin/streptomycin, and blasticidin, a selective marker for OAT1 expression. Cells were tested for mycoplasma contamination, and no contamination was observed. Prior to functional assays, cells were plated in 96 well plates and grown for 24 hours, or until confluent in media without blasticidin. Metabolites were added at either 1 or 2 mM and serial dilution was performed down all columns. A fixed concentration of 10 uM 6-

carboxyfluorescein was introduced to each well for 10 minutes. Cells were rinsed three times in ice-cold PBS, and the fluorescence was measured using a fluorescent plate reader. IC50 values were calculated using GraphPad Prism 9 (La Jolla, CA). Controls were carried out using probenecid, an established inhibitor of OAT1 function.

2.5.8. Magnetic bead binding assay

The OAT1 gene was cloned into a third-generation lentiviral vector system. The full-length protein was expressed as GFP fusion in HEK293 cells to check for protein expression. OAT1 protein was solubilized using n-Dodecyl-β-D-Maltoside (bDDM) detergent and then purified by immobilization on magnetic beads coupled to anti-flag antibody (MedChemExpress, HY-K0207). We used magnetic beads (5 microns) as the basis for a binding assay to screen small molecule compounds competing with a well-established OAT1 substrate, 6-carboxyfluorescein (6-CF), which was used as a fluorescent tracer. Loss of fluorescence when challenged with another substrate at a given concentration indicated potential competition for the same binding site. Multiple compounds and concentrations were assayed in a 96-well format using flow cytometry, gating directly on the scattering of the beads. Candidate compounds were initially screened at 3-10 times the concentration of 6-CF (kept constant at 6 uM, for example).

Fluorescent measurements were conducted using a Novacyte flow cytometer (Agilent) with sampler that reads one sample well at a time at regular time intervals. As these magnetic beads are denser than water, we used 50% glycerol to reduce bead sedimentation. Loss of fluorescence when challenged with candidate compounds at different concentrations was normalized against the maximum fluorescence due to 6-CF binding to OAT1, which was also measured periodically between every set of 10 samples. These check-point measurements (evaluating 6-CF binding to OAT1 in this case) spaced in time allowed us to monitor and to correct fluorescence due to bead sedimentation over time. Using this approach, we screened 20

compounds and categorized them (as OAT1 binder versus non-binder) based on their competitive efficiency against 6 uM 6-CF, selecting for loss of fluorescence signal compared to the relative error in repeated measurements of 6 uM 6-CF binding to OAT1 alone.

2.5.9. 5/6 nephrectomy model

In the metabolomics data from 5/6 nephrectomy model previously described by us [114], one kidney and 2/3 of the other were removed to model diminished renal function. A sham operation was performed on the healthy controls. After two weeks, plasma samples were collected from the animals following sacrifice. The samples were metabolically profiled, and 5/6 nephrectomy vs healthy were analyzed. We then compared these metabolites to those in **Supplementary Table 2.S3.**

2.5.10. Human drug-metabolite interactions

As previously described by us, plasma and urine samples were collected from 20 healthy participants before and 5 hours after an oral dose of probenecid [137]. These samples were metabolically profiled and pre-post comparisons were analyzed to determine compounds that were significantly altered. We analyzed compounds that were elevated in the plasma, decreased in the urine, and those that satisfied both criteria. We then compared these metabolites to those in **Supplementary Table 2.S3.**

2.5.11. Statistics

Scaled intensity for each metabolite was normalized to volume, and missing values were imputed with the lowest value for the compound. For all fold-change calculations, scaled intensities were averaged and compared to each other. Statistical comparisons were performed using the Python module, statsmodels (https://www.statsmodels.org/stable/index.html) and were made between groups by Two-way ANOVA following log transformation and false discovery rate (FDR) correction, with $p < 0.05$ being considered significant. Enrichment for each superpathway and subpathway was calculated using the number of total metabolites measured

and the number of metabolites in each respective superpathway and subpathway, as previously described [54].

2.5.12. Study approval

All experimental protocols were approved by the University of California San Diego Institutional Animal Care and Use Committee (IACUC), and the animals were handled in accordance with the Institutional Guidelines on the Use of Live Animals for Research. All the experiments described here follow the ARRIVE guidelines.

2.6 AUTHOR CONTRIBUTIONS

J.C.G. and S.K.N. wrote the manuscript. J.C.G., V.E., and K.M. conducted experiments and acquired data. J.C.G. and K.M. analyzed data. S.K.N. and G.A.C. provided reagents and resources. S.K.N. conceived of the project. S.K.N. and G.A.C. designed various aspects of the research studies. All authors reviewed the manuscript.

2.7 ACKNOWLEDGMENTS

2.8 SUPPLEMENTAL TABLES

Supplemental Table S2.1: 103 metabolites significantly altered by loss of *Oat1*.

Supplemental Table S2.2: 162 metabolites significantly altered by microbiome depletion.

Supplemental Table S2.3: 40 metabolites significantly altered by both loss of *Oat1* and microbiome depletion.

Table 2.1 Summary of gut-derived compounds measured for physical binding with OAT1.

Magnetic bead binding assays show results in this work. In vitro cell-based transport assays are collated from this work and other publications. In vivo mouse model *Oat1* KO refers to ANOVA results in this work from loss of *Oat1*. In vivo mouse model microbiome depletion refers to whether the ANOVA results in this work due to microbiome depletion. In vivo humans treated with probenecid refers to results from an in vivo analysis of these metabolites in human plasma and urine treated with probenecid [137]. ✓*:* evidence of interaction. N/A: compound not measured.

Compound	Magnetic bead binding assay	In vitro cell-based transport assay	In vivo mouse model *Oat1 KO* only	In vivo mouse model microbiome depletion only
2-hydroxphenylacetate		✓(this work)	✓	✓
Indolin-2-one	✓		✓	✓
Indoleacetate	✓	✓[73]		
Indolepropionate	✓	✓(this work)	✓	
3-phenylpropionate			✓	✓
4-hydroxyphenylacetate	✓		✓	
4-hydroxyphenylpyruvate	✓	✓[154]		
Cinnamate	✓		✓	✓
Deoxycholate	✓	✓(this work)	✓	✓
Imidazole propionate	✓			
Indole-3 acetamide			N/A	N/A
Indole-3-carboxaldehyde	✓		N/A	N/A
Indoleacrylate	✓		N/A	N/A
3-indoxyl sulfate	✓	✓[53]	✓	✓
4-hydroxycinnamate	✓	✓(this work)		
Phenylpyruvate	✓			
Serotonin		✓[73]		
Skatole			N/A	N/A
Trimethylamine N-oxide				✓
Tyramine	✓		N/A	N/A

Table 2.2 Overlap of metabolites impacted by loss of Oat1 and microbiome depletion in this study and metabolites affected by 5/6 nephrectomy in a chronic kidney disease model.

Metabolite	Impacted by loss of *Oat1*	Impacted by microbiome depletion	Elevated by 5/6 nephrectomy model [114]
3-indoxyl sulfate	✓	✓	✓
phenylacetylglycine	✓	✓	✓
p-cresol sulfate	✓	✓	✓
3-methylhistidine	✓	✓	✓
N-acetylserine	✓	✓	✓
2-isopropylmalate	✓	✓	✓
4-ethylphenyl sulfate	✓	✓	✓
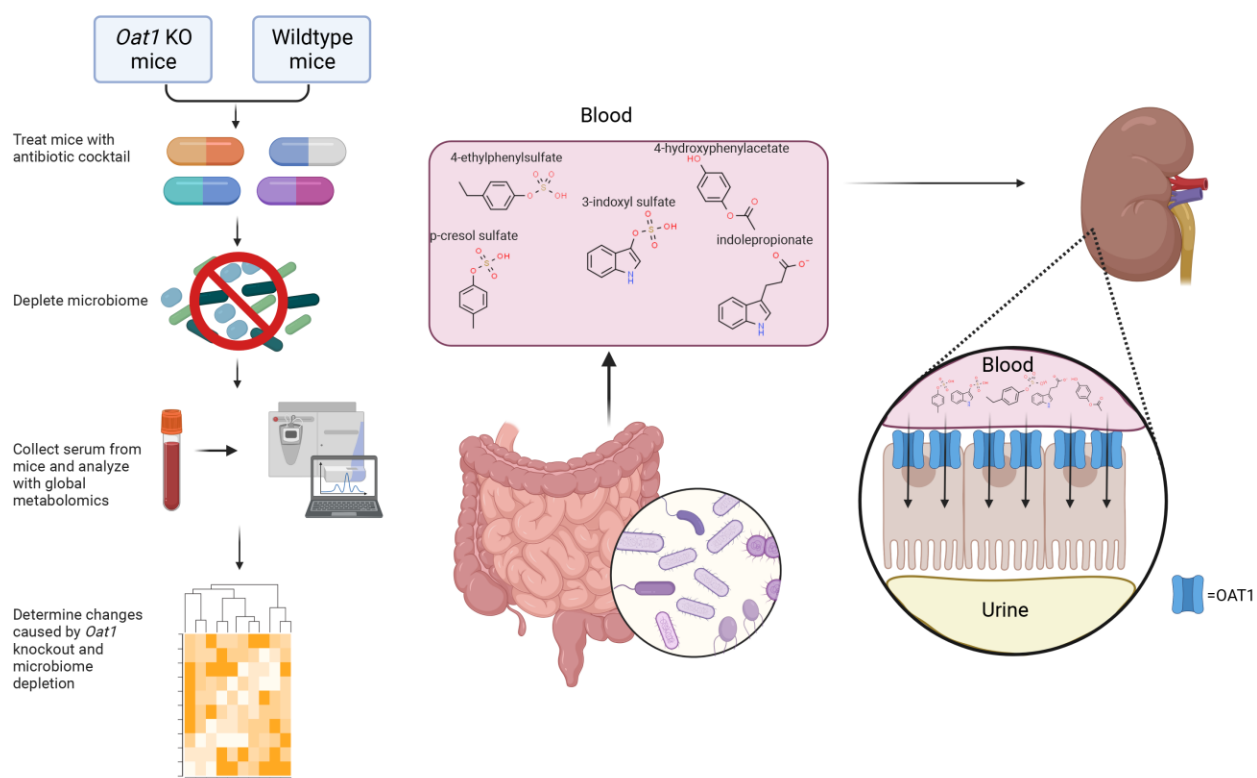
Figure 2.1 Workflow of experiment and schematic of gut-derived metabolite transport by OAT1.

Oat1 knockout and wildtype mice were treated with an antibiotic cocktail to deplete the gut microbes. We then assessed the resulting changes on the serum metabolome and determined that many metabolites produced by commensal bacteria in the gut enter the blood stream, where their systemic levels are regulated in vivo by OAT1 in the kidney.

Figure 2.2: Tc-99m MAG3, an OAT1 substrate, has different clearance patterns in wildtype and knockout mice.

A) Schematic for measurement of Tc-99m MAG3 in the urine in the bladder and the blood. **B)** Over the course of 30 minutes, the bladders of wildtype mice (n=5) reached their maximal levels of Tc-99m MAG3 more quickly than knockout mice (n=5). The central line represents the mean, while the error bands represent a standard deviation at each timepoint. **C)** In weight-scaled post-mortem gamma counts, 4 out of 5 pairs of mice showed higher levels of Tc-99m MAG3. Given that OAT1-mediated transport is often the rate limiting step for clearance into the urine, this pattern demonstrates the functional usefulness of genetic knockout for the studies that follow. **D)** In weight scaled post-mortem gamma counts, Tc-99m MAG3 levels in the blood were higher in the *Oat1* knockout mice in 4 out of 5 pairs. Since OAT1 is expressed at the basolateral membrane of the proximal tubule, it follows that blood levels of a substrate would be elevated in the *Oat1* knockout mice. %ID/g: percent injected dose over gram.
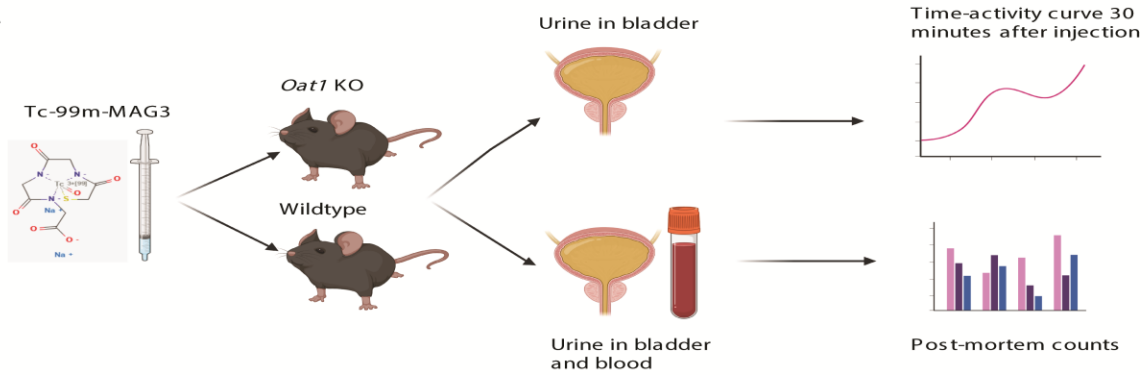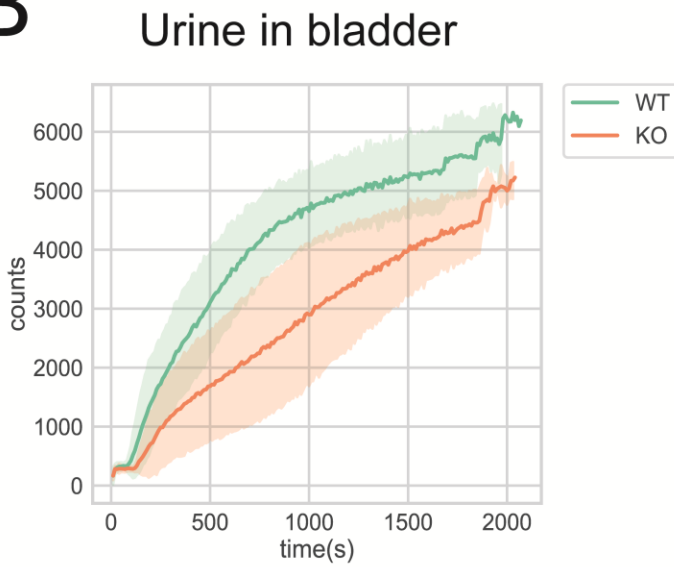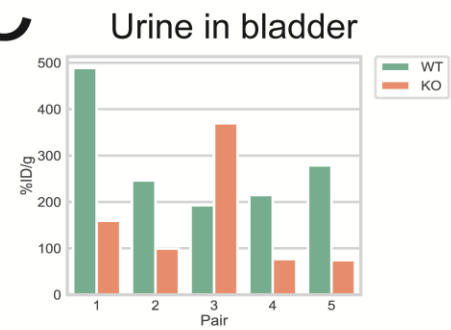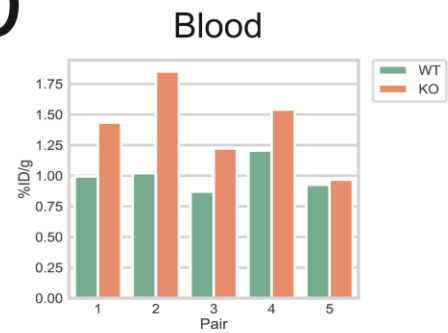
Figure 2.3 An antibiotic cocktail depleted the gut microbes in *Oat1* knockout and wildtype mice and decreased the circulating levels of gut-derived metabolites.

A) Metagenomic analysis of mice showed that observed OTUs are decreased in mice (n=12 for WT untreated, n = 11 for *Oat1* KO untreated, n = 5 for *Oat1* KO treated, n=4 for WT treated) treated with antibiotic cocktail. **B)** Linear discriminant analysis reveals separation between the metabolomic profile of the four groups (n=4 for all groups). **C–F)** The serum abundance of well-established gut-derived metabolites (cinnamoylglycine [2.31×10^{-10}], indolepropionate [5.66×10^{-08}], hippurate [7.99×10^{-08}], and trimethylamine N-oxide [3.85×10^{-04}] with different origins is significantly decreased in treated groups, as determined by corrected 2-way ANOVA. (n = 4 for all groups). Box plots include the median as the central line, the lower quartile as the lower limit of the box, the upper quartile as the upper limit of the box, the max value as the upper limit of the whisker, and the minimum value as the lower limit of the whisker. Diamonds indicate a value that falls outside of the interquartile range.
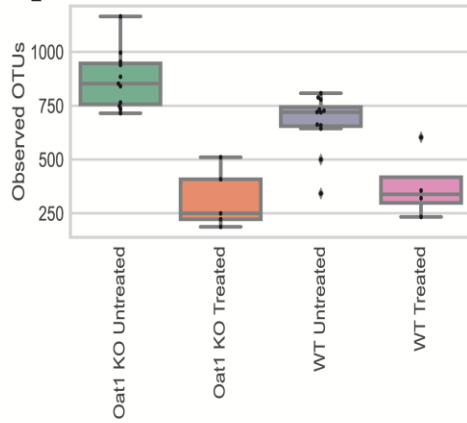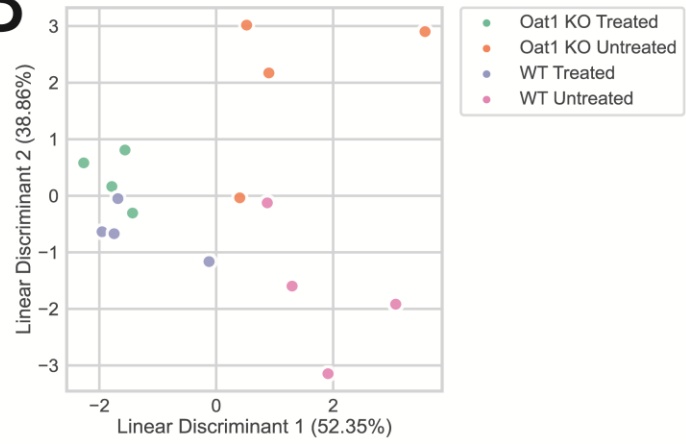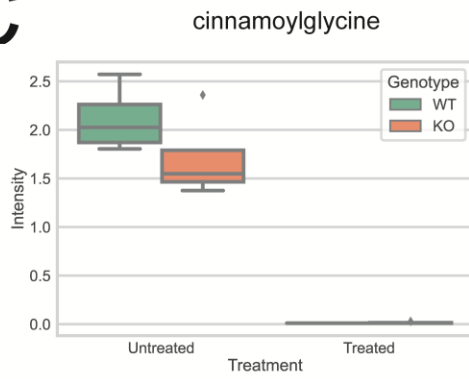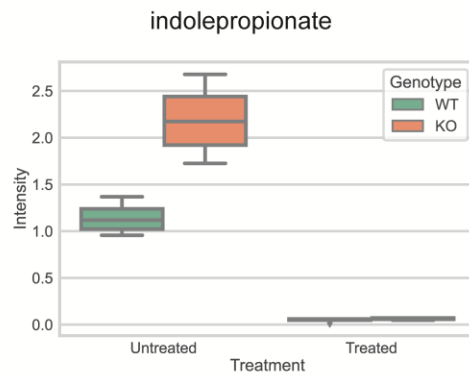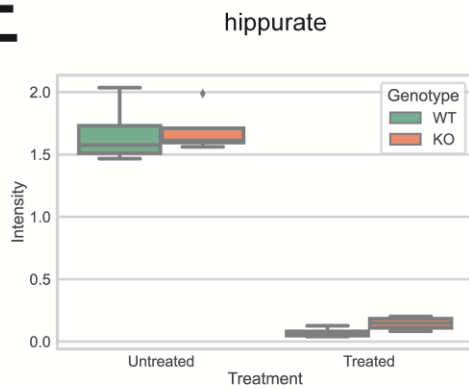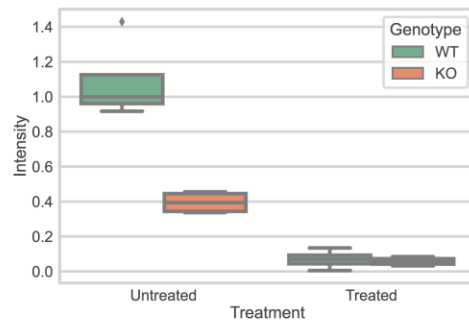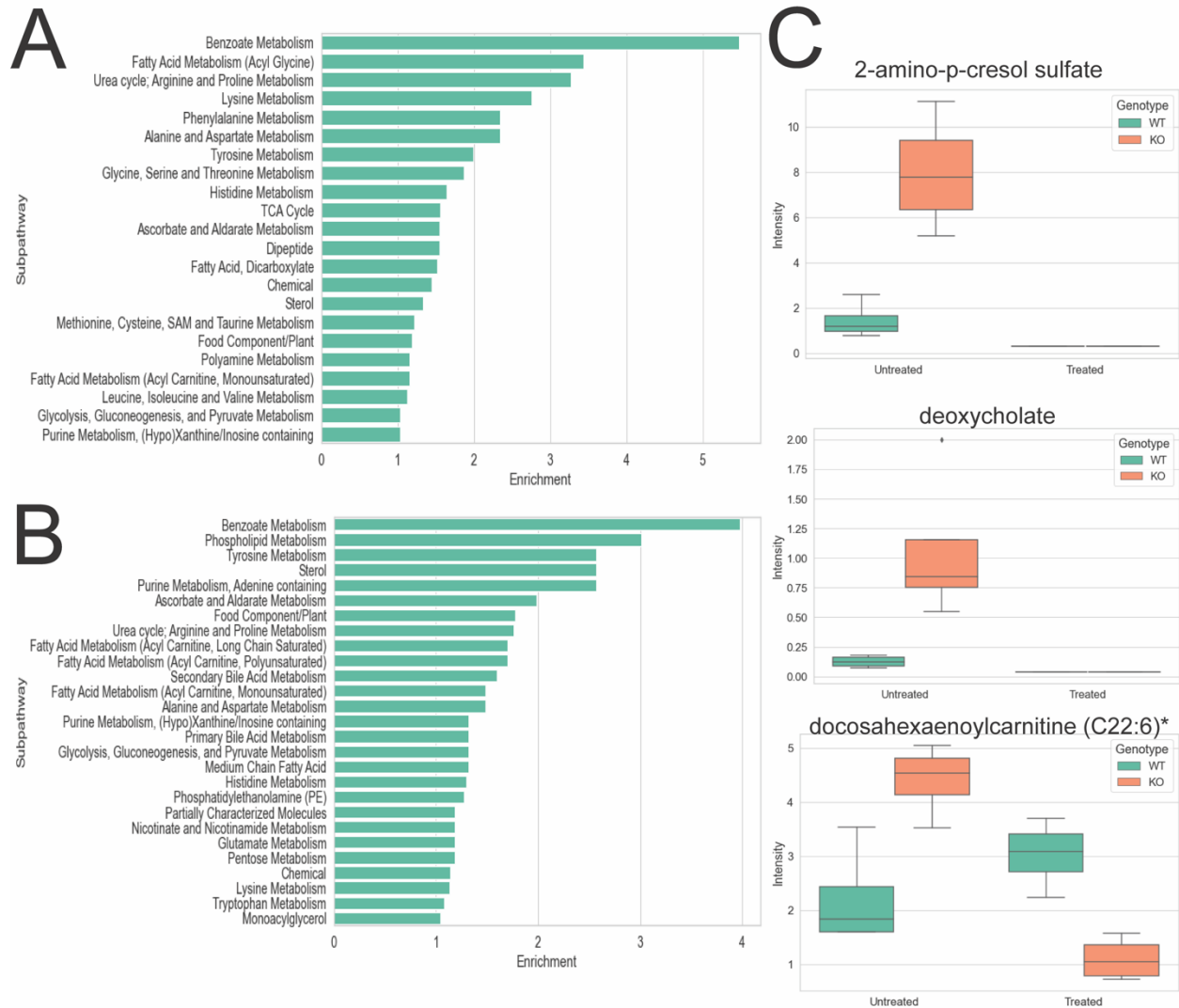
Figure 2.4 Genetic knockout of *Oat1* and antibiotic treatment lead to multiple altered biochemical pathways.

A) Enrichment results from the 103 metabolites significantly altered by loss of Oat1 show that Benzoate Metabolism, Fatty Acid Metabolism, and others are among the most affected subpathways. **B)** Enrichment results from the 162 metabolites significantly altered by microbiome depletion reveal that Benzoate Metabolism, Phospholipid Metabolism, Tyrosine Metabolism, and others are among the most affected subpathways. **C)** 2-Amino–p-cresol sulfate, deoxycholate, and docosahexaenoylcarnitine (C22:6)* were significantly affected by the interaction between genotype and treatment, as determined by 2-way ANOVA. Asterisk denotes the identity has not yet been confirmed based on a chemical standard, but there is high confidence in its identity. Box plots include the median as the central line, the lower quartile as the lower limit of the box, the upper quartile as the upper limit of the box, the max value as the upper limit of the whisker, and the minimum value as the lower limit of the whisker.

Figure 2.5 Treatment and genotype overlap in their impact on 40 circulating metabolites.

A) Of the 103 metabolites affected by loss of *Oat1* and the 162 metabolites affected by microbiome depletion, 40 metabolites overlapped. **B)** The 40 metabolites belonged to 19 unique subpathways, with 11 belonging to the Benzoate Metabolism subpathway. **C)** The levels of the 40 metabolites are shown, with the scaled intensity of each metabolite being scaled to range from 0 to 3 for improved visualization, though there are some metabolites that have higher scaled intensities. The individual groups of mice are shown on the y axis, and the individual metabolites are shown on the x axis. KOU, *Oat1*-KO untreated; KOT, *Oat1*-KO treated; WTU, WT untreated; WTT, WT treated. **D)** Linear discriminant analysis (LDA) shows clear separation between the 4 groups of metabolites altered by both loss of *Oat1* and microbiome depletion. Thirty-one of the 40 overlapping metabolites have available chemical structures (Group 1, 5 metabolites with structures; Group 2, 17 metabolites with structures; Group 3, 3 metabolites with structures; Group 4, 6 metabolites with structures). We were unable to find clearcut information for 2-amino–p-cresol sulfate, N2-acetyl, N6,N6-dimethyllysine, 4-allylcatechol sulfate, 4-ethylcatechol sulfate, 4-methoxyphenol sulfate, N-acetylhomocitrulline, succinoyltaurine, 3-methoxycatechol sulfate (2), or 1-methyl-5-imidazolelactate. **E)** Thirteen of the 22 compounds in Group 2 (elevated by KO and decreased by microbiome depletion) were sulfated. Only 1 other compound in the other groups was sulfated. **F)** Number of aromatic bonds was one of the features that most separates the 4 groups of compounds affected by *Oat1* KO and microbiome depletion. Group 2 had higher numbers of aromatic bonds than the other 3 groups.

A

B

C

BIOCHEMICAL

D

E number of sulfate groups

F number of aromatic bonds
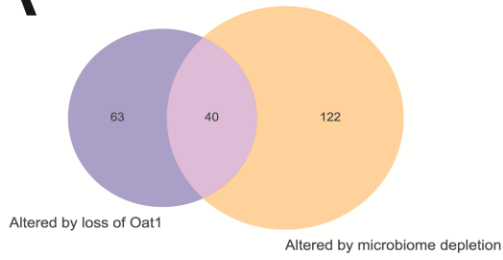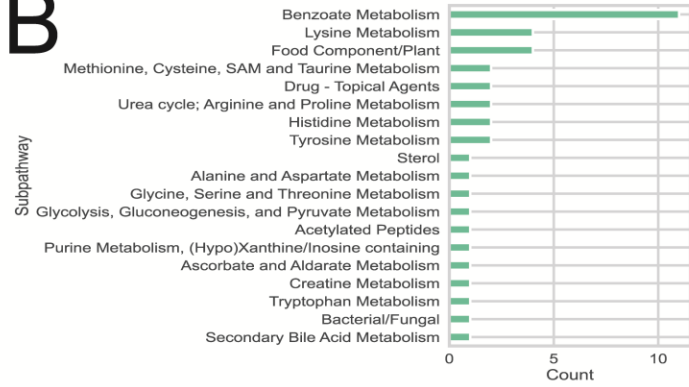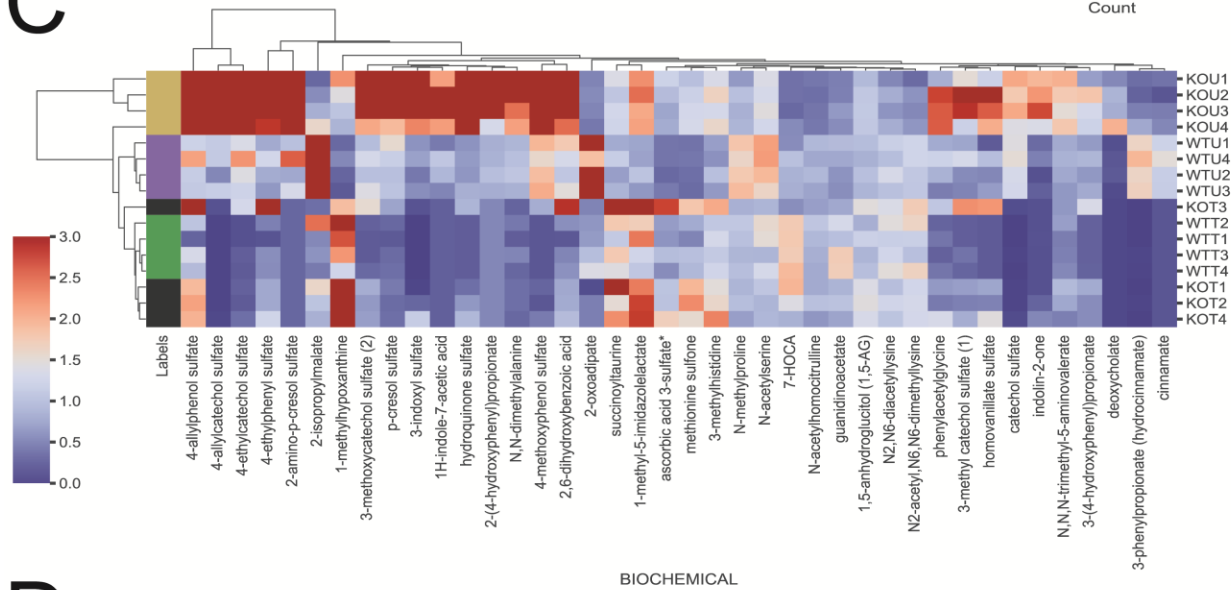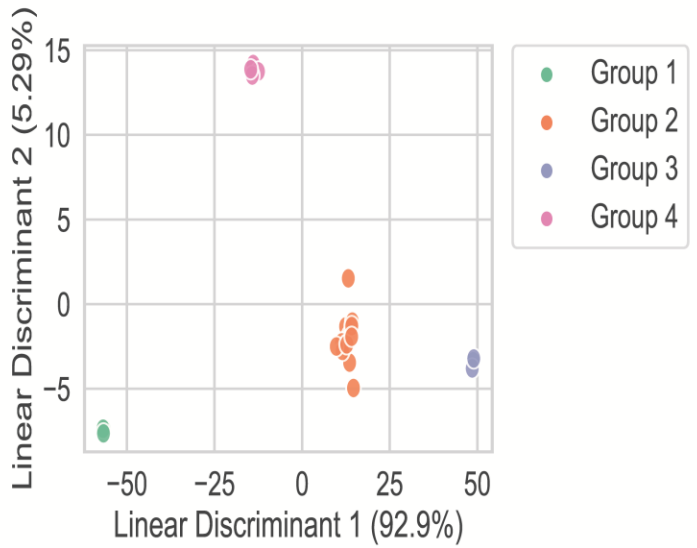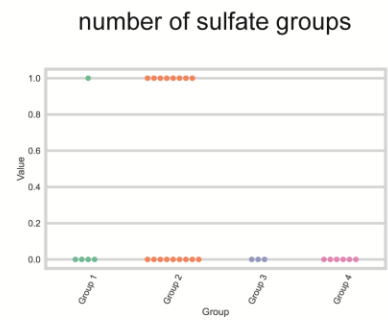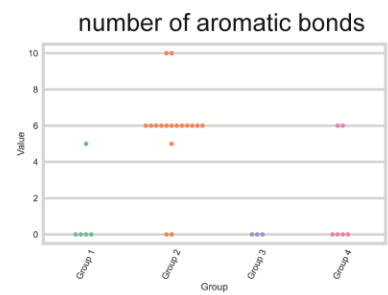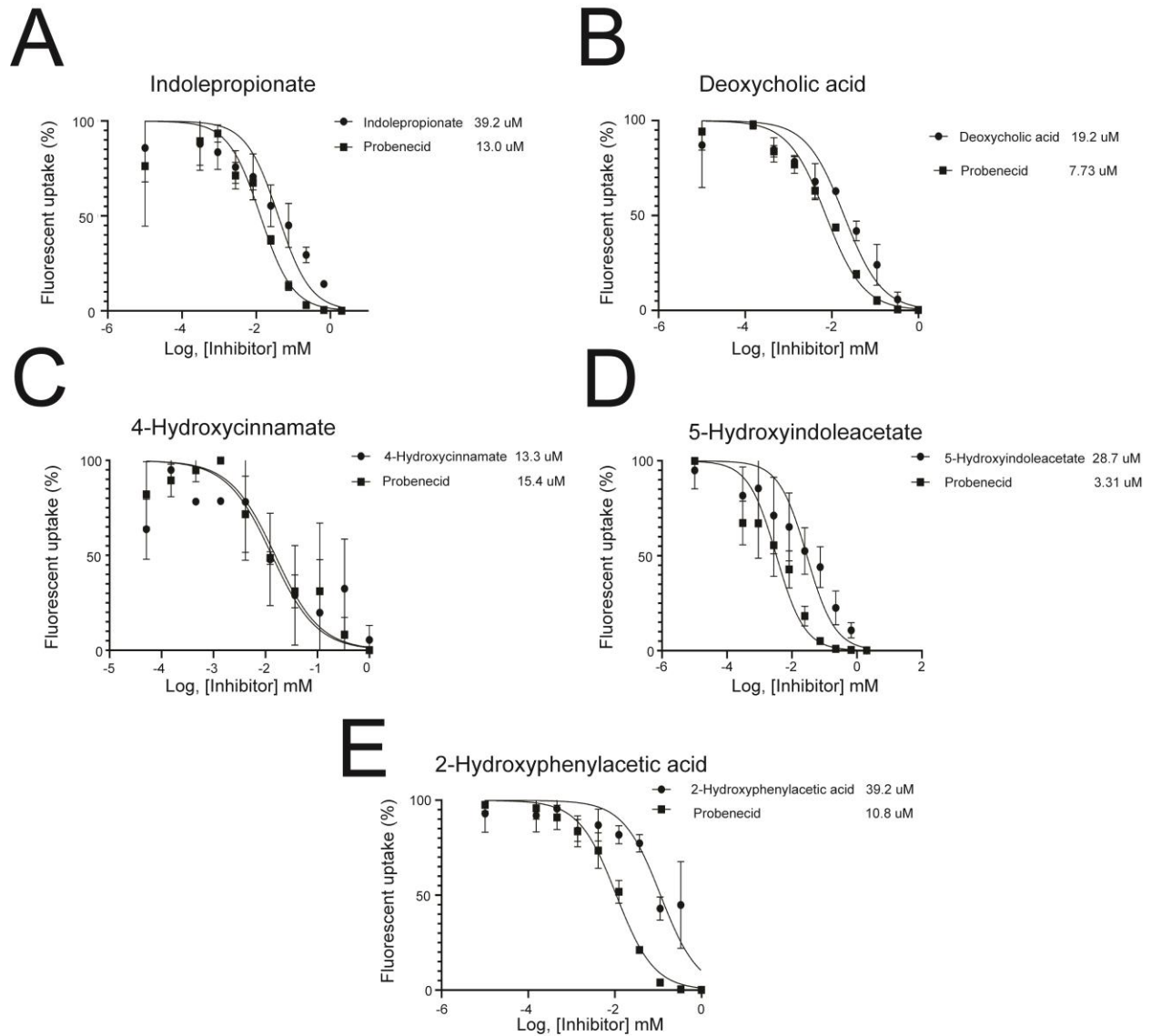
Figure 2.6 OAT1 transport is inhibited by gut-derived compounds in vitro.

A-E) Indolepropionate (n=4), deoxycholic acid (n=4), 4-hydroxycinnamate (n=2), 5-hydroxyindoleacetate (n=5), 2-hydroxyphenylacetic acid (n=2) inhibited the transport of 6-carboxyfluorescein in OAT1-expressing HEK293 cells**.** Controls were performed with the prototypical OAT1 inhibitor, probenecid, and IC50 values from at least n=2 assays are shown.
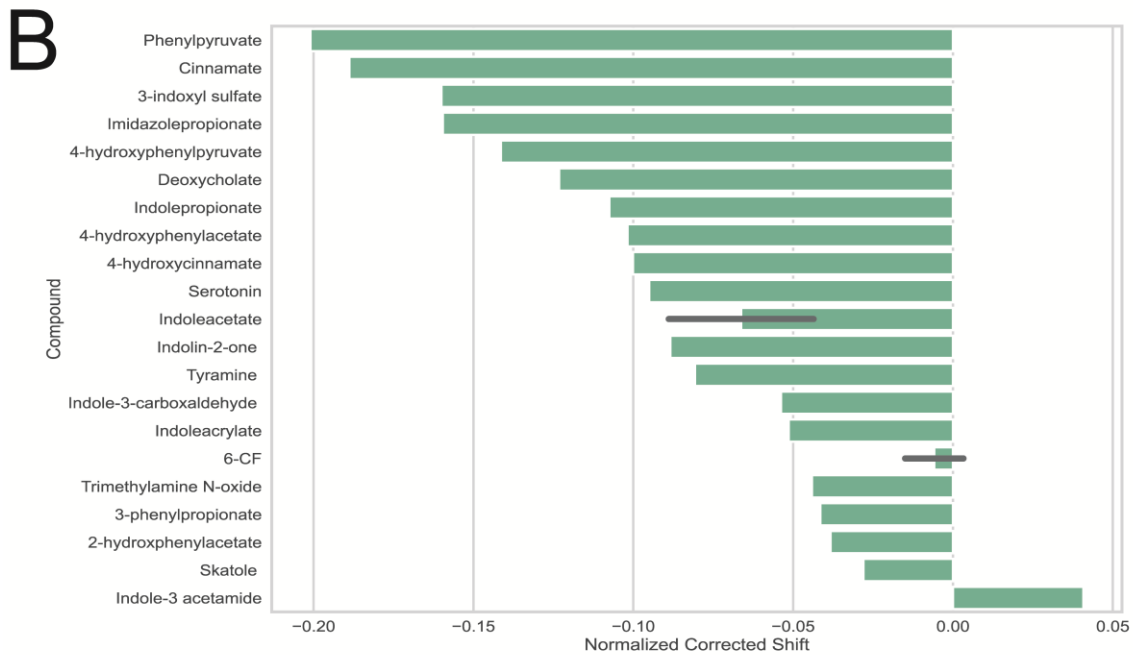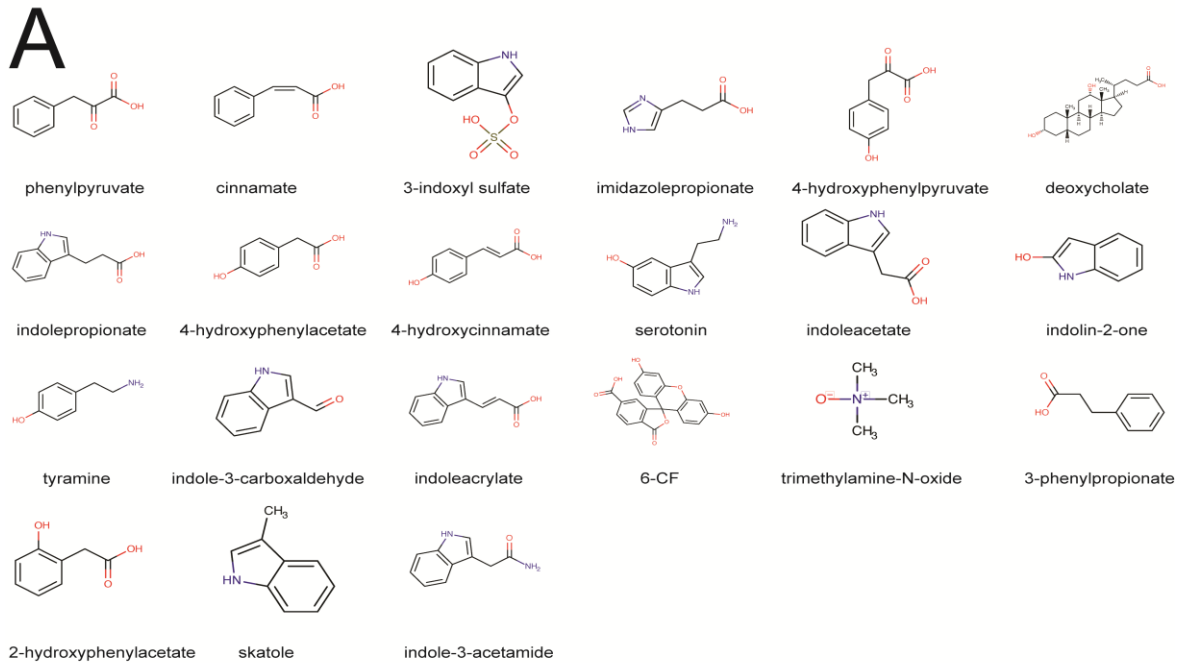
Figure 2.7 OAT1 binds gut-derived compounds in vitro.

A) 20 gut-derived metabolites and a control compound (6-carboxyfluorescein) were measured using the magnetic bead binding assay. **B)** All but 5 of the compounds (15 of 20) showed a significant normalized corrected shift in 6-carboxyflourescein signal, indicating that these compounds bind to OAT1.
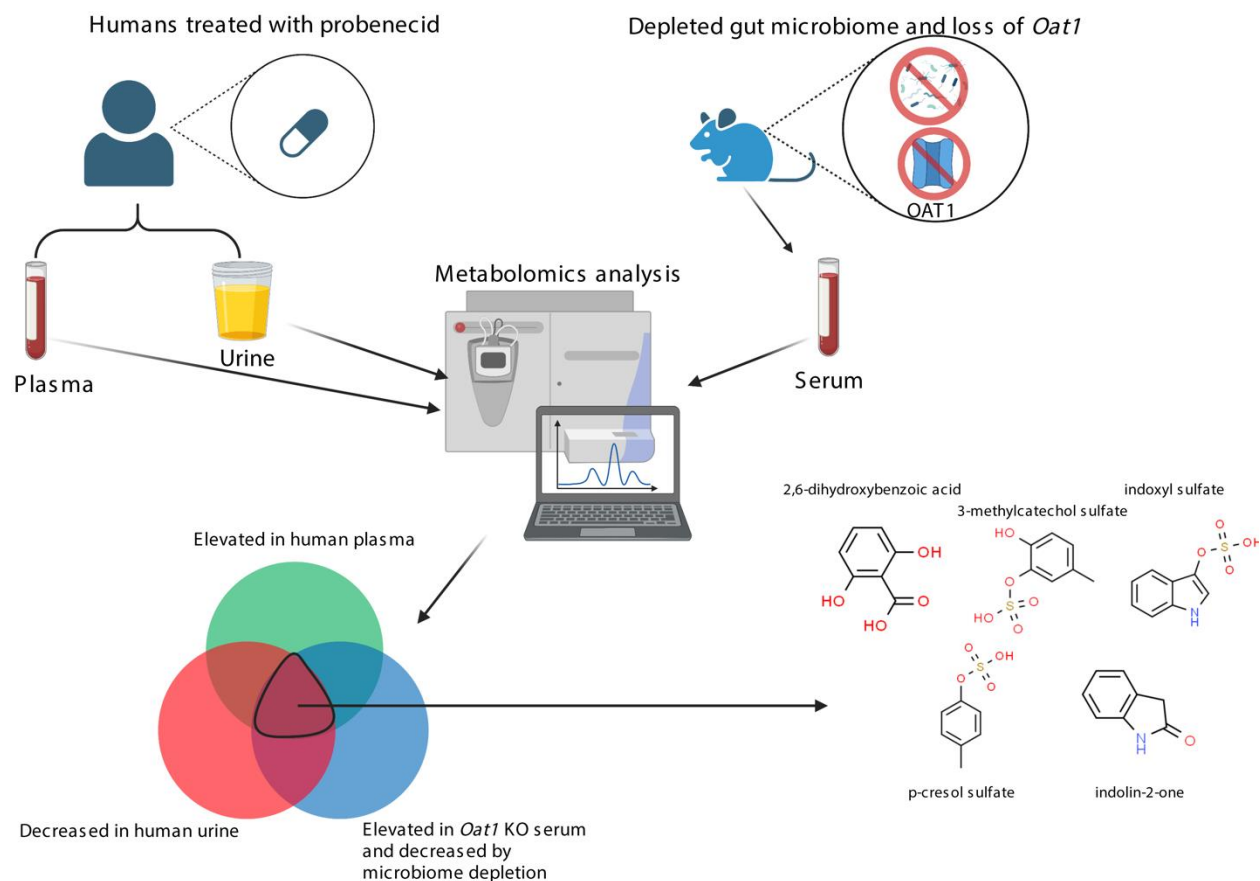
Figure 2.8 Gut-derived metabolites mediated by OAT1 kidney function are involved in clinical drug-metabolite interactions.

Of the 40 metabolites significantly affected by both loss of *Oat1* and microbiome depletion, 8 were also implicated in clinical drug-metabolite interactions with the drug, probenecid. These compounds were both elevated in the plasma and decreased in the urine, indicating that OAT1-mediated movement is the rate limiting step. Of the 8 metabolites, 5 had chemical structures and are shown in the figure. The remaining 3 are 4-ethylcatechol sulfate, 4-methoxyphenol sulfate, and 4-allylcatechol sulfate. There were also 20 of 40 metabolites decreased in the urine, and 11 of 40 increased in the plasma.

Chapter 2, in full, is a reprint of the material as it appears in "The kidney drug transporter OAT1 regulates gut microbiome-dependent host metabolism" by Jeffry C. Granados, Vladimir Ermakov, Koustav Maity, David R. Vera, Geoffrey Chang, and Sanjay K. Nigam in *JCI Insight*, 2023, 8.2. The dissertation author was the primary investigator and author of this paper.

CHAPTER 3: BLOCKADE OF ORGANIC ANION TRANSPORT IN HUMANS AFTER TREATMENT WITH THE DRUG PROBENECID LEADS TO MAJOR METABOLIC ALTERATIONS IN PLASMA AND URINE

3.1 ABSTRACT

Probenecid is used to treat gout and hyperuricemia as well as increase plasma levels of antiviral drugs and antibiotics. *In vivo,* probenecid mainly inhibits the renal SLC22 organic anion transporters OAT1 (SLC22A6), OAT3 (SLC22A8), and URAT1 (SLC22A12). To understand the endogenous role of these transporters in humans, we administered probenecid to 20 healthy participants and metabolically profiled the plasma and urine before and after dosage. Hundreds of metabolites were significantly altered, indicating numerous drug-metabolite interactions. We focused on potential OAT1 substrates by identifying 97 metabolites that were significantly elevated in the plasma and decreased in the urine, indicating OAT-mediated clearance. These included signaling molecules, antioxidants, and gut microbiome products. In contrast, urate was the only metabolite significantly decreased in the plasma and elevated in the urine, consistent with an effect on renal reuptake by URAT1. Additional support comes from metabolomics analyses of our *Oat1* and *Oat3* knockout mice, where over 50% of the metabolites that were likely OAT substrates in humans were elevated in the serum of the mice. Fifteen of these compounds were elevated in both knockout mice, while 6 were exclusive to the *Oat1* knockout and 4 to the *Oat3* knockout. These may be endogenous biomarkers of OAT function. We also propose a probenecid stress test to evaluate kidney proximal tubule organic anion transport function in kidney disease. Consistent with the Remote Sensing and Signaling Theory, the profound changes in metabolites following probenecid treatment support the view that SLC22 transporters are hubs in regulation of systemic human metabolism.

3.2 INTRODUCTION

Probenecid is a US Food and Drug Administration (FDA)-approved drug that has historically been used to slow the clearance of drugs in short supply [22]. In World War II, probenecid was co-administered with penicillin to increase the half-life of the antibiotic to treat infections in wounded soldiers [22]. Probenecid has also been used with several other drugs in this manner, exploiting what are now referred to as drug-drug interactions (DDI) [155]. Although probenecid is used to increase the half-life or drugs excreted by the kidney, it is also used in the treatment of gout and hyperuricemia, conditions associated with disordered urate homeostasis [156-158]. Probenecid increases renal excretion of urate by inhibiting its reabsorption from the proximal tubule lumen back into the blood, exploiting a drug-metabolite interaction [157, 159-161]. In addition to clinical usage, probenecid is also used as an inhibitor of several transporters of organic anions for *in vitro* studies in research settings [22, 162].

The early uses of probenecid were discovered without specific knowledge of its molecular targets [163]. Since then, many of the proteins that participate in the renal organic anion secretory system have been identified in mice and humans, and several have been shown to have direct interactions with probenecid [98]. Probenecid has three widely accepted *in vivo* renal targets primarily expressed in the proximal tubule: organic anion transporters 1 and 3 (SLC22A6/OAT1, SLC22A8/OAT3) and uric acid transporter 1 (SLC22A12/URAT1). Following oral administration, probenecid rapidly enters the bloodstream, where it is highly bound to albumin [163]. Protein-bound probenecid is carried through the blood and ultimately the peritubular capillaries, where it inhibits the function of OAT1 and OAT3 (**Figure 3.1**). These SLC22 uptake drug transporters are localized to the basolateral membrane of the proximal tubule and transport a wide array of substrates including drugs, endogenous metabolites, natural products, and toxins, as evidenced by the characterization of the *Oat1* and *Oat3* knockout mice,

as well as numerous *in vitro* studies [52, 54, 72-74, 77, 99, 164]. Free, unbound probenecid is filtered by the glomerulus and enters the urinary filtrate, where it acts by inhibiting URAT1, an apical transporter that is involved in the reabsorption of urate from the urine into the cell [165, 166] (**Figure 3.1**).

These transporters, as well as several other multi-specific, oligo-specific, and mono-specific transporters and enzymes are important regulators of endogenous metabolism, as proposed in the Remote Sensing and Signaling Theory (RSST) [6, 8]. The RSST describes the combined role of drug transporters and drug metabolizing enzymes in inter-organ and intra-organ communication through the movement of small molecules [6-8, 122]. Many of these proteins are best known for their role in the absorption, distribution, metabolism, and excretion (ADME) of drugs, but their tissue expression patterns and shared substrate specificity allow them to collaboratively handle many other classes of small molecules, such as endogenous metabolites, natural products, toxins, gut microbiome products and nutrients. One of the key tenets of the RSST is that the primary function of these "drug" ADME proteins is to regulate endogenous metabolism, and that drugs are effectively probes for the endogenous Remote Sensing and Signaling system [6]. This is supported by the evolutionary conservation of these gene families across several species, including worms, flies, and sea urchins [58, 167]. The OATs and SLC22s in general, are important hubs in a proposed network of multi-specific proteins that aid in returning the body to homeostasis following perturbations, and there is now considerable experimental and human genetic support for this view [11, 164]. Given that the primary expression of known *in vivo* probenecid targets are OAT1, OAT3, and URAT1 in the proximal tubule of the kidney, we hypothesized that the inhibition of these proteins by probenecid would markedly alter the levels of key metabolites and signaling molecules in the blood and urine.

Here, we show that the administration of a clinically prescribed drug, probenecid, has a major impact on the human plasma and urine metabolomes, likely through the direct inhibition of physiologically important transporters, OAT1, OAT3, and URAT1 expressed in the kidney. This broad effect on metabolism is consistent with predictions of the RSST. Although DDIs receive the majority of drug transporter-related research attention, the wide substrate specificity of these OATs and other multi-specific transporters raises the possibility of several other competitive interactions, including drug-metabolite, drug-nutrient, and drug-toxin interactions [21, 168, 169]. In brief, we performed global metabolomics on the plasma and urine of 20 healthy participants collected before and 5 hours after an oral dose of probenecid and analyzed numerous small molecules reflective of potential drug-metabolite interactions (DMI), drug-nutrient interactions (DNI), and other competitive interactions. Hundreds of metabolites were significantly altered in each medium, including 124 that were significantly elevated in the plasma and decreased in the urine, indicating likely OAT-mediated transport. We then compared the altered metabolites to those impacted in the serum of *Oat1* and *Oat3* knockout mice to identify potential biomarkers for drugs that are handled by the organic anion transport system, which is the main elimination pathway for protein-bound drugs in the kidney. The data support the value of the *Oat* knockouts for understanding human physiology. They also suggest that the body experiences a metabolic shift following administration of a drug that inhibits organic anion transporters, hubs in human systemic metabolism. We also propose the "probenecid stress test", which can provide a functional readout on tubular function in healthy and diseased states.

3.3 METHODS
3.3.1. Participants
All experimental protocols were reviewed and approved by the Institutional Review Board and abide by the Declaration of Helsinki Ethical Principles. Blood and urine were

collected from 20 healthy participants (14 women and six men) before and 5 hours after an oral dose of 1 gram of probenecid. Based on clinical settings (e.g., treatment of gonorrhea with antibiotics), one gram of probenecid was the dose used [170]. Participants had an average age of 30.85±10.93 and an average body mass index of 24.18±3.52. The protocol was developed in consultation with clinical researchers at University of California San Diego Altman Clinical and Translational Research Institute (ACTRI). According to the protocol, all participants were asked to be on a diet of no meat, fish, or eggs for 3 days prior to their visit and the day of their visit. Participants were also asked to not take any medications, vitamin C tablets, nutritional supplements, caffeine, chocolates, or cruciferous vegetables the day before and the day of the visit. Participants were also asked to consume an extra liter of water the morning of the visit. There were no additional restrictions on diet between blood/urine collection. All participant data was deidentified. Samples were stored at -80 C until metabolomic analysis.

3.3.2. Metabolomic analysis

Samples were shipped on dry ice to Metabolon Inc. (Durham, NC). Plasma and urine samples were processed separately. Protein was removed from samples, and samples were separated into fractions. These fractions were analyzed by either reverse phase UPLC-MS/MS with positive ion mode electrospray ionization, RP/UPLC-MS/MS with negative ion mode ESI, or HILIC/UPLC-MS/MS with negative ion mode ESI. Following quality control and accounting for instrument and process variability, data was extracted, peak-identified, and assessed for quality by Metabolon. Plasma data was normalized to volume, and urine data was normalized to volume and osmolality. *Oat1* and *Oat3* knockout plasma metabolomics data, which has been previously extensively analyzed by us, were collated from prior publications and other available data from our lab and followed previous protocols with respect to acquisition and analysis [54,

72, 73]. Recent serum metabolomics data on newer platforms is generally consistent with previously published data, and relevant information is included in Supplementary Information.

3.3.3. Metabolic pathway analysis

Plasma and urine metabolomics data were analyzed separately. Missing values were replaced with the minimum observed value for each compound. Plasma fold changes were calculated using volume normalized data, and urine fold changes were calculated using volume and osmolality normalized data by Metabolon. For statistical comparisons, data were log transformed, and a pairwise t-test was used to calculate p-values in plasma and urine. Enrichment was calculated as previously described [73]. Principal component analysis was performed using the sci-kit learn package in Python 3.8. Visualizations were performed using the Seaborn package in Python 3.8.

3.4 RESULTS

3.4.1. Short-term probenecid treatment alters the levels of hundreds of circulating metabolites

We first focused our attention on the plasma of the participants, as OAT1 and OAT3 have been shown to impact circulating levels of small molecules in human and rodent models due to their basolateral localization (blood-facing) and have been associated with far more small molecule compounds than URAT1 [5]. Thus, the combined metabolic roles of OAT1 and OAT3 could best be determined by analyzing the changes in plasma levels of endogenous metabolites and other compounds. We measured the levels of 1,234 unique metabolites, including probenecid and multiple unidentified compounds, spanning numerous biochemical pathways in 20 healthy participants (**Figure 3.2A, Supplementary Table 3.S1**). The global metabolic profiles before and after treatment were largely separable by principal component analysis (PCA), showing that oral dosage of probenecid impacts the plasma metabolome in a consistent fashion (**Figure 3.2B**). We analyzed the levels of metabolites that were elevated, as they are likely due to the inhibition of OAT1 and OAT3 in the kidney, and this can potentially be

validated by considering metabolites elevated in the *Oat1* and *Oat3* KO mice previously described by our group [52-54, 72, 73]. We found 354 metabolites that were significantly elevated ($p < 0.05$, fold change > 1) (**Figure 3.2C**) and that many of them were present in biochemical pathways implicated in rodent models, such as tryptophan, tyrosine, phenylalanine, and bile acid metabolism (**Figure 3.2D**) [52, 53, 72-74, 140, 171]. In addition to determining the metabolic roles of OAT1 and OAT3, we were also interested in how URAT1 inhibition might impact endogenous metabolism. Since probenecid inhibits URAT1 and potentially other organic anion reuptake transporters in the kidney, it is possible that certain metabolites decreased in the plasma are a result of small molecules not being reabsorbed back into the blood from the urine. We identified 230 metabolites that were significantly decreased in the plasma, suggesting that URAT1 may play an important role in the circulating levels of compounds in the blood (**Supplementary Figure 3.S1**). However, considering the relatively limited substrate specificity of URAT1, it is unlikely that many of these compounds are elevated due to URAT1 inhibition, and some of these changes may be due to temporal variations.

3.4.2. Metabolites in the urine are mainly decreased following probenecid treatment

Our initial focus was on the plasma of the participants because of the clear implication of OAT1/OAT3 function, but we were also interested in the urine, as OAT1/OAT3 uptake is often the rate limiting step for excretion into the urine. Hence, we assumed that OAT1/OAT3 substrates would have lower levels in the urine. In contrast, URAT1 acts by reabsorbing urate and other compounds from the urine into the proximal tubule, so we also analyzed the metabolites that were elevated in the urine. We measured 1,315 unique metabolites in the urine of the same 20 healthy participants before and after probenecid administration (**Supplementary Table 3.S2**). High levels of probenecid were detected in the urine 5 hours after the oral dosage in all the participants (**Figure 3.3A**). PCA revealed a separation between pre and post treatment,

demonstrating the effect on excretion (**Figure 3.3B**). Due to the number of metabolites that were

elevated in the plasma, we focused on the metabolites that were decreased in the urine under the

assumption that probenecid also prevents these metabolites from passing through the tubular

secretion system and entering the urine. Six hundred twenty-two metabolites were significantly

decreased in the urine (p < 0.05, FC < 1), including multiple metabolites that overlapped with

those altered in the plasma. Indeed, some of the subpathways containing elevated metabolites in

the plasma also had decreased metabolites in the urine, suggesting that their levels of these

metabolites were mediated by probenecid targets (**Figure 3.3, C and D**). Although the exact

metabolites and subpathways altered in the plasma and urine slightly differed, tryptophan,

tyrosine, and bile acid pathways were all enriched for decreased metabolites, which reflected the

changes observed in the plasma. In addition, 113 metabolites were elevated in the urine, some

presumably due to the inhibition of URAT1 and other less well characterized organic anion

reuptake transporters, which covered a different set of metabolic subpathways (**Supplementary

Figure 3.S2**).

3.4.3. Ninety-seven compounds are likely human OAT substrates based on plasma and urine metabolomics

Of the nearly 2,000 metabolites measured across both experiments, 124 were elevated in

the plasma and decreased in the urine, indicating physiologically relevant inhibition of

OAT1/OAT3, as both uptake and excretion were altered in the expected manner. To better

determine which metabolites may be relevant as endogenous biomarkers, we applied fold change

criteria used by regulatory agencies, which generally indicate that for DDI studies the safety

margins are between 80 and 125% [4]. Hence, we limited our overlap to metabolites with fold

changes over 1.25 in the plasma and under 0.8 in the urine, resulting in 97 metabolites, 40 of

which have known chemical structures, including 14 with *in vitro* support (**Figure 3.4,**

Supplementary Table 3.S3). Many of these metabolites come from a subset of biochemical pathways, with 7 from tryptophan metabolism, 6 from tyrosine metabolism, 9 from either primary or secondary bile acid metabolism, and 4 from androgenic steroids, which have all previously been associated with OAT function. Several of the altered compounds were characterized by ring structures and negative charges, and many are linked to Phase II drug metabolism**.** Sulfation and glucuronidation are among the most common conjugations that improve renal excretion, and they are well represented in our subset, with 27 sulfated compounds and 7 glucuronidated compounds, indicating the close association between the OATs and drug metabolizing enzymes. Thirty unidentified metabolites were also present, but these have not been linked to any chemical structure or known biochemical role.

3.4.4. Probenecid has a specific drug-metabolite interaction with urate

Whereas the nature of the compounds that were elevated in the plasma and decreased in the urine suggested inhibition of OAT1/OAT3, we also aimed to understand the potential effect of URAT1 inhibition. Along with OAT1 and OAT3, URAT1 is considered a primary *in vivo* target of probenecid, and its mechanism of action is blocking reabsorption of urate from the urine into the proximal tubule and back into the blood. We compared the metabolites decreased in the plasma and elevated in the urine and found that three compounds (urate, quinate, N-acetylglycine) satisfied both criteria. However, when the more stringent fold changes (FC < 0.8, FC >1.25) were applied only one compound, urate, satisfied both conditions (**Figure 3.5**). Because probenecid was highly elevated in the urine, it was able to exert its inhibitory effect, which led to a specific drug-metabolite interaction between probenecid and urate at URAT1. Again, we find it remarkable that of the hundreds of metabolites analyzed in both plasma and urine, only one was elevated in the urine and decreased in the plasma, as might be inferred from clinical, knockout, and *in vitro* data.

While OAT1, OAT3, and URAT1 are widely accepted as the main *in vivo* targets of

probenecid, there is the possibility of effects on other transporters (SLC and ABC) which are

inhibited *in vitro* by probenecid. That said, *in vivo* evidence is lacking for a major role for these

other transporters in probenecid-sensitive organic anion transport. Thus, we focused on OAT1,

OAT3, and URAT1, for which considerable *in vitro, in vivo*, and *ex vivo* support exists [42, 52-

54, 72-74, 98]. Previous work by our group has focused on the roles of these transporters in *in

vivo* endogenous metabolism using genetically engineered mice. We analyzed global metabolic

profiling data from the serum of *Oat1* and *Oat3* knockout mice from multiple studies by our

group and found changes that support the important physiological role of the OATs [52-54, 73,

74, 114, 172]. Though there are differences in general physiology, gene expression patterns, and

microbiome composition between mice and humans, it was expected that there would be overlap

between the knockout mice and humans treated with probenecid. Indeed, there were multiple

metabolites that were elevated in the human plasma, the knockout mouse serum, and decreased

in the urine (indicative of OAT-mediated transport).

When using the 124 metabolites elevated in the plasma and decreased in the urine, we

found that 52 and 48 of these metabolites were measured in the *Oat1* and *Oat3* knockout mice,

respectively. We focused on the metabolites that were elevated in each knockout mouse and 52%

(27/52) for the *Oat1* and 48% (23/48) for the *Oat3* mice were altered. However, we mainly

focused on the 97 metabolites that fit the fold change criteria previously described. For this

subset, 55% (21/38) of the metabolites measured in humans and *Oat1* mice had *in vivo* knockout

mouse support. With respect to OAT3, 56% (19/34) had *in vivo* knockout mouse support. Fifteen

metabolites were elevated in both knockout mice, while 6 were unique to OAT1, and 4 were

unique to OAT3 (**Figure 3.6**). This work demonstrates that *Oat* knockout mice can be used to predict potential DMIs at the site of transporters and produce candidates for endogenous biomarkers. Although these compounds would need to be further characterized (rate of synthesis, metabolic breakdown, other routes of clearance, etc.) to prioritize for usefulness as endogenous biomarkers, many of them have been shown to be impacted by OAT perturbation *in vivo* (humans and mice) and *in vitro*.

3.5 DISCUSSION

We found that probenecid, a drug used to treat gout and hyperuricemia and increase levels of OAT-transported drugs (e.g., antibiotics, antivirals), had a major impact on the levels of endogenous metabolites and diet-derived compounds in the plasma and urine. The altered pathways spanned several biochemical pathways and functional clusters based on chemical structures, such as bile acids and aromatic amino acids. These biochemical pathways were generally similar to those altered in *Oat1* and *Oat3* knockout mice [52, 72, 73, 94, 95]. Among the surveyed metabolites in the plasma and urine, we noted that many metabolites were elevated in the plasma and decreased in the urine, indicating that inhibition of OAT1 and OAT3 by probenecid leads to a much more pronounced systemic impact than URAT1 inhibition. However, use of probenecid's known mechanism of action (inhibition of urate reabsorption at URAT1) was also apparent, as urate alone was both elevated in the urine and decreased in the plasma.

Overall, the profound metabolite alterations support the view that OAT1 and OAT3 are "hubs" in a Remote Sensing and Signaling Network of transporters and enzymes regulating metabolism; in particular, the RSST emphasizes the roles of these multi-specific "drug" transporters in inter-organ communication mediated by endogenous small organic compounds such as key metabolites, antioxidants, gut microbiome products and signaling molecules that activate GPCRs and nuclear receptors [6, 122, 164].

Many of the impacted metabolites were also altered in the serum of *Oat1* and *Oat3* knockout mice, supporting the view that the resulting changes are due to the inhibition of OATs. Thus, in the process of analyzing human data, we further validated the usefulness of previous mouse knockout models. In particular, the *Oat* knockout mice may be useful in predicting drug-metabolite interactions at the site of these transporters and identifying potential endogenous biomarkers of transporter function. Since probenecid is known to target both OAT1 and OAT3, the endogenous biomarkers implicated here would be for general OAT function rather than for one specific transporter. Nonetheless, by comparison with the knockouts, we identified metabolites that may be OAT1 or OAT3 selective, as previously described [74].

Although probenecid is not known to have notable short or long term side effects, it is striking that within hours of taking probenecid so many metabolites are elevated in the plasma, such as bile acids and indole derivatives, which have important signaling roles in the body [173, 174]. Their increases in circulation can lead to the activation of signaling cascades in organs that interface with the blood, as their bioavailability is elevated. For example, bile acids activate bile acid-specific receptors (i.e., TGR5) and the nuclear receptor FXR (expressed in liver and kidney), which govern many key processes in the gut-liver-kidney axis [175, 176]. Consistent with the Remote Sensing and Signaling Theory, the elevated levels of tryptophan and tyrosine metabolites could also play be physiologically important, as they each have distinct signaling roles in different organs [177, 178].

It is important to point out that this increase in OAT-regulated metabolites may also occur with other drugs that inhibit the OATs at varying levels, which could lead to a wide range of metabolic side effects via drug-metabolite interactions at the level of the transporter. While the main targets of probenecid are the renal transporters (OAT1, OAT3, URAT1), several other

77

proteins are known to interact with the drug *in vitro*, including multi-specific SLC (often SLC22) and ABC transporters (MRPs) and enzymes that also play important roles in key aspects of endogenous metabolism [179]. Whether these other proteins are *in vivo* targets of probenecid in humans is far from clear, but it is conceivable that the inhibition of these proteins may lead to drug-metabolite interactions that are reported in this work.

In the kidney, many metabolites must be salvaged from the urinary filtrate and reabsorbed into the blood. Hundreds of measured compounds in the urine were decreased by probenecid, likely due to the lack of tubular secretion by the OATs, the rate-limiting step in urinary clearance for many organic anions. These metabolites, which include purine derivatives, aromatic amino acid derivatives, and others are primarily excreted through the urine. Steroids were also decreased in the urine, which is consistent with the use of probenecid to mask levels of androgenic steroids in urine samples [180]. Like the plasma, it appears that the urine metabolome is more influenced by the inhibition of the OATs than URAT1. In the context of OAT1/3 (basolateral uptake transporters) vs URAT1 (apical uptake transporter), metabolites elevated in the plasma and decreased in the urine are linked to OAT1/3, while metabolites decreased in the plasma and elevated in the urine are linked to URAT1. However, we must also consider the possibility that probenecid may inhibit efflux and retro-transporters expressed on the apical membrane of the proximal tubule, although this remains to be established *in vivo*.

Recent work by our group and others has highlighted the endogenous role of OAT1 and OAT3 in regulating uremic toxins, amino acid derivatives, lipids, and several other classes of metabolites using genetically engineered knockout mice [52, 54, 72-74]. We hypothesized that the inhibition of these physiologically important proteins with probenecid in humans would lead to similar changes and found that while not all metabolite alterations were reproduced, several

classes of metabolites were consistently altered in both the plasma and the urine, namely the metabolites known to interact with OAT1 and OAT3. URAT1, on the other hand, is more specific, with only a few known unique interacting small molecules. While knockout mice appear to be useful models in determining potential drug-metabolite interactions, the inter-species differences in gene expression, diet, and gut microbiome should also be considered, particularly considering that many of the metabolites altered in the mice and humans originate from the gut microbiome. Nonetheless, some of these compounds are strong candidates for endogenous biomarkers used in predicting drug-drug interactions at the site of the OATs.

Of the 25 compounds altered across human and knockout mouse experiments, 6 of these (2-hydroxyphenylacetate, 3-acetylphenol sulfate, 4-acetamidobutanoate, 4-methylcatechol sulfate, N-formylmethionine, 4-methoxyphenol sulfate) are unique to OAT1 and they differ from the compounds currently being tested as endogenous biomarkers for OAT1 [181]. As for OAT3, 3 of 4 unique compounds (gentisate, N-acetylphenylalanine, N-acetyltryptophan) altered in humans and *Oat3* knockout mice are novel--except for indoleacetate, which has been shown to be a potential endogenous biomarker for OAT function [182]. We also compared our results to other potential OAT biomarkers identified by other groups, and we found that our results supported the use of p-cresol sulfate and pyridoxate, in that both of these compounds were elevated in the plasma and decreased in the urine [130, 182]. Further criteria, including consistent levels throughout the day, stable production independent of diet, minimal interactions with other proteins, and *in vitro* support for specificity to OAT1/3 would further support these compounds as strong endogenous biomarkers. The altered compounds could potentially be used as biomarkers for organic anion-related tubular transport, which is largely mediated by OATs and is gaining more attention in the context of renal diseases [183].

Our results suggest that drugs handled by the OATs produce several simultaneous drug-metabolite interactions. The main mechanism of action of probenecid is as an inhibitor of key renal transporters; thus, its impact on the plasma and urine metabolomes is likely stronger than drugs with other targets [22, 157]. However, the absorption, distribution, metabolism, and excretion (ADME) of nearly every drug and xenobiotic is handled by a subset of drug transporters and drug metabolizing enzymes [184]. It is possible that other drugs cleared by OAT1 or OAT3 could lead to similar consequences in humans if they inhibit the transporter strongly enough. OAT1 and OAT3, in particular, are among the most multi-specific drug transporters, with each interacting with over 100 unique drugs, including NSAIDs, antivirals, antibiotics, and others [99]. Chronic treatment with drugs that interfere with the function of transporters could lead to long term metabolic side effects. In HIV patients taking anti-retroviral therapy drugs (many of which are OAT1/3 substrates), it is common to see metabolic side effects relating to lipids after several months [185]. Similar situations have been reported with NSAIDs and antibiotics, many of which are OAT substrates [186, 187].

Finally, we note again that tubular secretion and glomerular filtration both contribute to overall kidney function but in many contexts, only glomerular filtration rate is considered. In recent years, there has been a new emphasis on assessing tubular secretion, which is, however, complicated by the fact that renal disease state (CKD, AKI, etc.), genetics and other factors all influence tubular secretion. In light of this, and the data presented in this article, we propose a "probenecid stress test" to evaluate the transport capacity of the proximal tubule, as there is a need to assess organic anion-related tubular function, preferably without the administration of furosemide, a strong diuretic [188].

The probenecid stress test, as we currently envision it, is similar in design to the studies performed here. It would measure the plasma and urine levels of the ~100 metabolites identified here as potential human OAT substrates before and after a single dose of probenecid. With this, a quantitative measure of how compromised organic anion-related tubular function is can be calculated to influence drug dosing and help assess disease state. It can also be used to follow progression of disease. From the pharmaceutical perspective, the test can be used determine how similar a novel drug entity is to probenecid (a complete inhibitor of OATs). Patients may also need to avoid (or be differently dosed) drugs that are primarily secreted through OAT1 and OAT3. Indeed, many drugs that are prescribed in the setting of renal disease (e.g., antibiotics, antivirals, antihypertensives, diuretics) are substrates of OAT1 and OAT3. The use of probenecid in late stage renal disease to isolate tubular function has previously been demonstrated in animal models [114]. From the disease perspective, the proposed OAT substrates can be used as biomarkers to assess renal OAT1 and OAT3 tubular function, which comprise the bulk of organic anion transport in the proximal tubule of the kidney and serve as indicators of interacting drugs, toxins, endogenous metabolites, natural products, and several other classes of small molecule compounds. By focusing on endogenous metabolites, it is possible to assess the physiological function of the organic anion handling proteins in the proximal tubule without (or before) subjecting the tubular system to one or more drugs that may pose risks, especially in the context of declining renal function. The probenecid stress test can also be used as a novel clinical test to measure tubular secretion relative to glomerular filtration rate. Finally, administration of probenecid can also be useful for determining and following organic anion tubular function responsiveness as CKD progresses.

3.6 STUDY HIGHLIGHTS
3.6.1. What is the current knowledge on the topic?

Probenecid is believed to inhibit three kidney proximal tubule transporters *in vivo*. URAT1 mainly transports urate from the urine into the cell, while the basolateral uptake transporters, OAT1 and OAT3, regulate endogenous metabolism as evidenced by alterations in *Oat1* and *Oat3* knockout mice.

3.6.2. What question did this study address?

What alterations in metabolism does probenecid cause, presumably by drug-metabolite interactions (DMI), and are they occurring at the level of OAT1, OAT3, or URAT1?

3.6.3. What does this study add to our knowledge?

We have identified *in vivo* DMI potential endogenous biomarkers for OAT1 and OAT3, as well as potential biomarkers for more general tubular function. We also propose a "probenecid stress test" to assess tubular function.

3.6.4. How might this change clinical pharmacology or translational science?

These findings could prove helpful in drug development, and in assessing a patient's tubular function in acute kidney injury and chronic kidney disease. We have also shown that a drug that is considered quite safe nonetheless leads to hundreds of simultaneous drug-metabolite interactions and suspect this to be the case for other drugs, which could explain some drug side effects or adverse drug reactions.

3.7 ACKNOWLEDGEMENTS

3.8 AUTHOR CONTRIBUTIONS

J. C. G. performed the research, analyzed the data, and wrote the manuscript. V. B. designed the research. S. K. N. designed the research and wrote the manuscript.

3.9 SUPPLEMENTARY INFORMATION

Supplementary Figure S1: Subpathways enriched for decreased metabolites in the plasma.

Supplementary Figure S2: Subpathway enrichment for metabolites increased in the urine.

Supplementary Table S1: Metabolites significantly altered in the plasma of probenecid-treated humans.

Supplementary Table S2: Metabolites significantly altered in the plasma of probenecid-treated humans.

Supplementary Table S3: Metabolites elevated in the plasma and decreased in the urine that are putative OAT1/3 substrates with relevant *in vitro* support.

Supplementary Table S4: Metabolites from *Oat1* and *Oat3* serum metabolomics experiments that were elevated in the knockout mice, elevated in the probenecid-treated human plasma, and decreased in the probenecid-treated human urine.
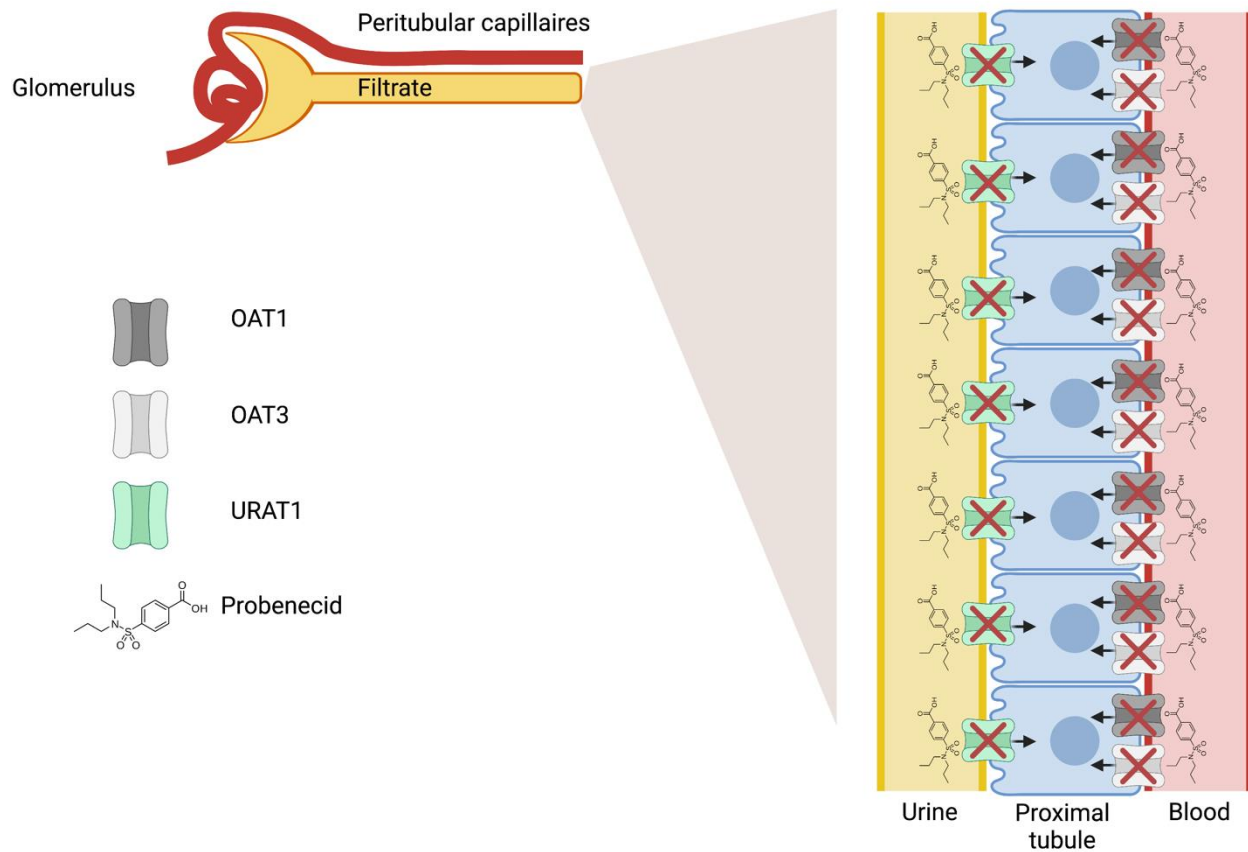
Figure 3.1 Probenecid effect on the kidney.

Probenecid inhibits the function of URAT1 on the apical membrane of the proximal tubule. Unfiltered probenecid goes through the peritubular capillaries and inhibits the function of OAT1/OAT3.
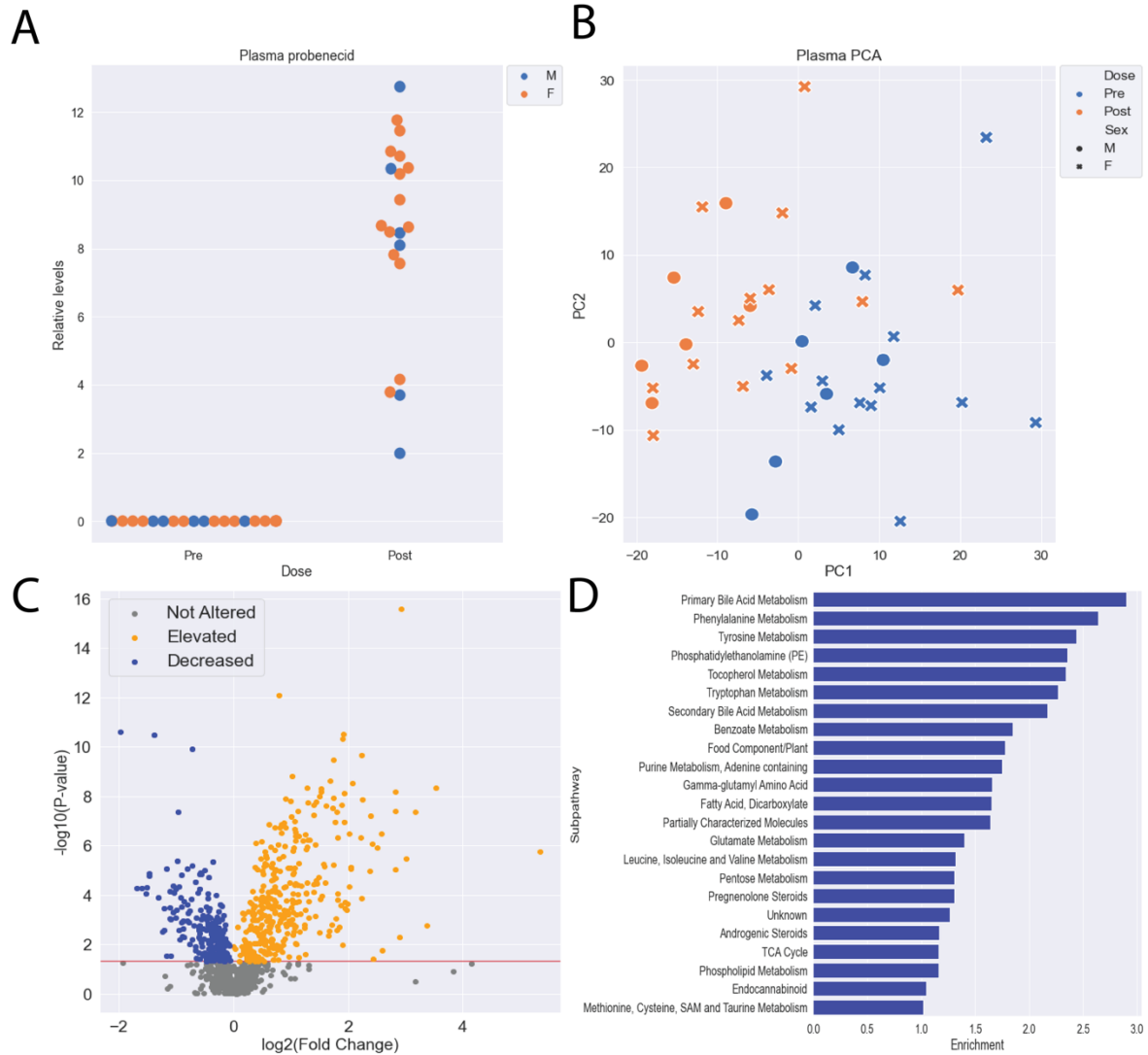
Figure 3.2 Probenecid treatment alters the plasma metabolome.

A) Probenecid levels in the plasma were significantly elevated 5 hours after oral dosage in all participants. **B)** Principal component analysis (PCA) reveals separation between pre and post treatment with probenecid plasma metabolomes. **C)** Hundreds of metabolites were significantly altered (elevated and decreased) following treatment with probenecid. **D)** Among the significantly elevated metabolites, 21 subpathways with at least 5 metabolites were enriched, including subpathways traditionally associated with OAT-mediated transport (Primary Bile Acid Metabolism, Phenylalanine Metabolism, Tyrosine Metabolism, Tryptophan Metabolism, etc.).
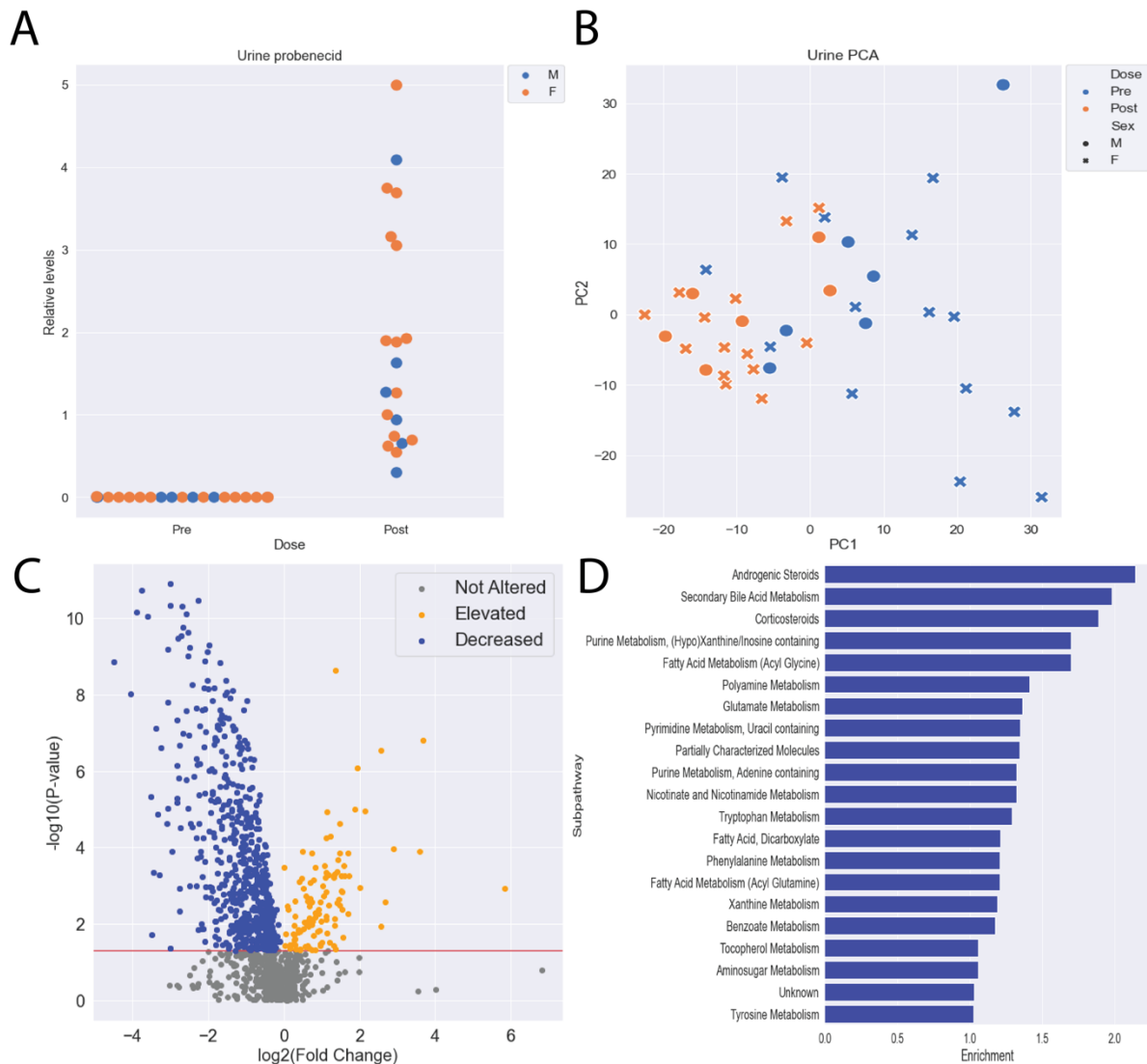
Figure 3.3 Probenecid treatment alters the urine metabolome.

A) Probenecid levels in the urine were significantly elevated 5 hours after oral dosage in all participants. **B)** Principal component analysis reveals separation between pre and post treatment with probenecid urine metabolomes. **C)** Hundreds of metabolites were significantly altered following treatment with probenecid, with most being decreased. **D)** Among the significantly decreased metabolites, 23 subpathways with at least 5 metabolites were enriched, including subpathways traditionally associated with OAT-mediated transport (Secondary Bile Acid Metabolism, Tryptophan Metabolism, Phenylalanine Metabolism, etc.).

Row 1: 1-methylurate | 1,3-dimethylurate | 2-hydroxyphenylacetate | 2,6-dihydroxybenzoic acid | 3-(4-hydroxyphenyl)lactate (HPLA)

Row 2: 3-formylindole | 3-indoxyl sulfate | 3-ureidopropionate | 4-acetamidobutanoate | 4-hydroxyphenylpyruvate

Row 3: 6-hydroxyindole sulfate | alliin | allo-threonine | androstenediol (3beta,17beta) monosulfate (1) | androsterone glucuronide

Row 4: dihydrocaffeate sulfate (2) | etiocholanolone glucuronide | gentisate | glycolithocholate sulfate* | glycoursodeoxycholate

Row 5: guaiacol sulfate | indoleacetate | indoleacetylglutamine | indolelactate | indolin-2-one

Row 6: kynurenine | methionine sulfoxide | N-acetyl-2-aminooctanoate* | N-acetyl-S-allyl-L-cysteine | N-acetylcitrulline

Row 7: N-acetylleucine | N-acetylphenylalanine | N-acetyltryptophan | N-formylanthranilic acid | N-formylmethionine

Row 8: p-cresol glucuronide* | p-cresol sulfate | phenol sulfate | pyridoxate | taurolithocholate 3-sulfate

Figure 3.4 Metabolites elevated in the plasma and decreased in the urine are likely OAT1/3 substrates.

Metabolites were further filtered by fold change criteria, with only plasma metabolites with fold changes over 1.25 and urine metabolites with fold changes under 0.80 were included. Overall, 97 metabolites fit these criteria, with 40 having known chemical structures. Thirty-four of these 40 compounds had a total negative charge, and many were also supported by existing *in vitro* data (**Supplementary Table 3.S3**).

87
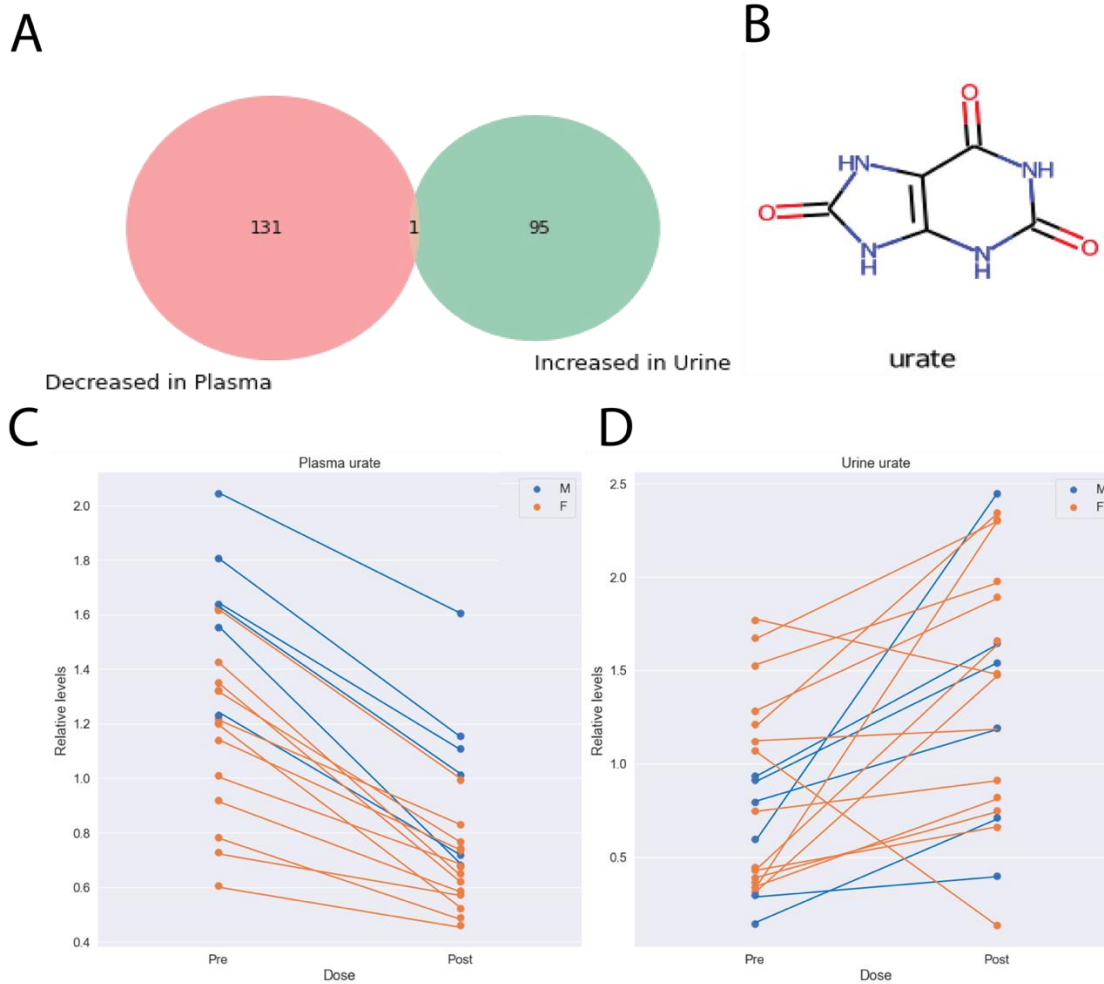
Figure 3.5 Presumed inhibition of urate reuptake transporters such as URAT1 led to a specific drug-metabolite interaction between probenecid and urate.

A) Urate was the only metabolite to be significantly decreased in the plasma (fold change < 0.8) and increased in the urine (fold change > 1.25) with more selective fold change criteria. **B)** The chemical structure for urate. **C)** Urate levels were significantly decreased in the plasma following treatment with probenecid (p-value: 1.20E-10, fold change: 0.606). **D)** Urate levels were significantly increased in the urine following treatment with probenecid (p-value: 0.008, fold change: 1.705).

Both

2,6-dihydroxybenzoic acid   3-indoxyl sulfate   3-ureidopropionate   6-hydroxyindole sulfate   guaiacol sulfate

indolelactate   indolin-2-one   kynurenine   N-formylanthranilic acid   p-cresol glucuronide*

p-cresol sulfate   phenol sulfate   pyridoxate

OAT1 only

2-hydroxyphenylacetate   4-acetamidobutanoate   N-formylmethionine

OAT3 only

gentisate   indoleacetate   N-acetylphenylalanine   N-acetyltryptophan

Figure 3.6 Multiple metabolites suggested to be OAT substrates are supported by in vivo *Oat1* and *Oat3* knockout mice.

Twenty-five metabolites elevated in the probenecid-treated human plasma, decreased in the probenecid-treated urine, and elevated in one or both knockout mice. Twenty of these metabolites had associated chemical structures. Fifteen (13 with chemical structures) were common to both knockout mice, while 6 (3 with chemical structures) were unique to the *Oat1* knockout mice, and 4 were unique to the *Oat3* knockout mice.

Chapter 3, in full, is a reprint of the material as it appears in "Blockade of organic anion transport in humans after treatment with the drug probenecid leads to major metabolic alterations in plasma and urine" by Jeffry C. Granados, Vibha Bhatnagar, and Sanjay K. Nigam in *Clinical Pharmacology and Therapeutics*, 2022, 112.3, 653-664. The dissertation author was the primary investigator and author of this paper.

CHAPTER 4: Regulation of Human Endogenous Metabolites by Drug Transporters and Drug Metabolizing Enzymes: An Analysis of Targeted SNP-Metabolite Associations

4.1 ABSTRACT

Drug transporters and drug metabolizing enzymes are primarily known for their role in the absorption, distribution, metabolism, and excretion (ADME) of small molecule drugs, but they also play a key role in handling endogenous metabolites. Recent cross-tissue co-expression network analyses have revealed a "Remote Sensing and Signaling Network" of multi-specific, oligo-specific, and monospecific transporters and enzymes involved in endogenous metabolism. This includes many proteins from families involved in ADME (e.g., SLC22, SLCO, ABCC, CYP, UGT). Focusing on the gut-liver-kidney axis, we identified the endogenous metabolites potentially regulated by this network of ~1000 proteins by associating SNPs in these genes with the circulating levels of thousands of small, polar, bioactive metabolites, including free fatty acids, eicosanoids, bile acids, and other signaling metabolites that act in part via G-protein coupled receptors (GPCRs), nuclear receptors and kinases. We identified 77 genomic loci associated with 7,236 unique metabolites. This included metabolites that were associated with multiple, distinct loci, indicating coordinated regulation between multiple genes (including drug transporters and drug metabolizing enzymes) of specific metabolites. We analyzed existing pharmacogenomic data and noted SNPs implicated in endogenous metabolite handling (e.g., rs4149056 in *SLCO1B1*) also affecting drug ADME. The overall results support the existence of close relationships, via interactions with signaling metabolites, between drug transporters and drug metabolizing enzymes that are part of the Remote Sensing and Signaling Network, along with GPCRs and nuclear receptors. These analyses highlight the potential for drug-metabolite

interactions at the interfaces of the Remote Sensing and Signaling Network and the ADME protein network.

Keywords: transporters, enzymes, ADME, metabolomics, SNPs, pharmacogenomics, fatty acids, eicosanoids, homeostasis

4.2 INTRODUCTION

Genome-wide association studies (GWAS) have been used to identify single nucleotide polymorphisms (SNPs) that are linked to phenotypes [189]. The phenotypic traits examined include disease states, drug efficacy and many others, indicating that GWAS can be used to gain further insight on the genetic causes of many conditions [190-193].With the increased generation of large omics datasets, GWAS have also been used to link SNPs to multiple intermediate phenotypes with metabolomics and proteomics [194, 195].

While much of the research in this area has focused on identifying differences caused by disease states or other lifestyle factors, GWAS on healthy patients can elucidate the endogenous role of genes by associating specific SNPs to levels of endogenous metabolites. Recent studies have combined GWAS and metabolomics on plasma and urine of participants to identify potential interactions between proteins and metabolites [63, 64, 196]. Here, we focused on SNPs in genes of multi-, oligo- and mono-specific transporters and "drug" metabolizing enzymes (DMEs), many of which are best known for their handling of pharmaceutical products, and their associations with thousands of circulating endogenous metabolites. The choice of genes was partly influenced by recent data indicating that these multi-, oligo- and mono-specific transporters and enzymes are found in or near hubs in co-expression networks, especially along the gut-liver-kidney axis, suggesting an important endogenous role [11].

Drug transporters and DMEs are among the most studied proteins in pharmacology because of their roles in the ADME (absorption, distribution, metabolism, excretion) of

pharmaceutical products [184]. Many of the multi-specific transporters and DMEs have the capacity to handle structurally diverse drugs, while their more oligo- and (relatively) mono-specific counterparts may transport or modify as few as one or two endogenous substrates [5, 10, 197]. GWAS have linked SNPs in these genes to changes in drug toxicity, efficacy, and distribution [198, 199].

However, the multi-specific nature of these proteins is not limited to pharmaceutical products [9]. Mainly in model organisms, but also in humans, endogenous metabolites, including those with well-defined signaling roles, have also been identified as likely *in vivo* substrates of these proteins, often supported by *in vitro* studies [6]. In GWAS, other results have demonstrated that SNPs in transporter and enzyme genes are associated with endogenous metabolites participating in biochemical pathways, like amino acid catabolism, glycolysis, ketone body metabolism, and others [63-65, 200, 201]

Understanding the full range of endogenous substrates of drug transporters and DMEs can help uncover the physiological metabolic processes that are perturbed when a patient takes drugs. In drug-metabolite interactions (DMI), a drug competes with a metabolite for access to a transporter or enzyme, and thus shifts metabolism by impacting the intracellular and extracellular concentration of the endogenous substrate [137]. Pharmacogenomic studies have focused on the implications of polymorphisms in these genes with respect to the drug-handling, but the "natural" function of these genes and its potential impact in drug-induced diseases or drug side-effects has received comparatively little attention.

The Remote Sensing and Signaling Theory proposes that the primary function of drug transporters and DMEs, together with closely related genes, is to help optimize levels of endogenous metabolites in bodily fluids and tissues by mediating inter-organ and inter-

organismal (e.g., gut microbe-host) communication through small molecule metabolites and signaling molecules [11]. This mechanism, while now experimentally supported in model organisms [73, 77, 95, 144], is also supported in humans [12]. Many endogenous metabolites have signaling capabilities that contribute to the regulation of the expression and/or function of other membrane transporters and enzymes by activating nuclear receptors, creating feedback loops [202].

Furthermore, many of these proteins share substrates with one another and are expressed in multiple epithelial tissues, suggesting the possibility of communication between them via these proteins, thereby mediating organ crosstalk [139]. Transporters are regulators of entry (uptake) and exit (efflux) of compounds into the epithelial tissues and body fluids they separate. For example, solute carrier organic (SLCO), solute carrier 22 (SLC22), and ATP-binding cassette subfamily C (ABCC) transporters are expressed in many of the same barrier epithelia tissues, like the proximal tubule (blood-urine), hepatocyte (blood-bile), and choroid plexus (blood-cerebrospinal fluid), and share many common pharmaceutical and endogenous substrates, suggesting that they may be jointly involved in the regulation of these substrates across multiple organs [122]. Indeed, SLC22 and ABCC proteins are among the many "drug transporter" families that were identified as hubs in the aforementioned co-expression gut-liver-kidney network of ~600 proteins--largely consisting of multi-, oligo- and mono-specific transporters, enzymes and nuclear receptors (including many ADME proteins)--and presented as a preliminary "Remote Sensing and Signaling Network" [11].

The focus of this study is to identify the metabolites and metabolic pathways regulated by these and related proteins in this Remote Sensing and Signaling Network. An additional focus is to determine if evidence for drug-metabolite interactions (DMI) can be found--given the overlap

in proteins of the Remote Sensing and Signaling Network (mediating endogenous small molecule homeostasis) and the ADME protein network (mediating the metabolism and elimination of drugs). The scale of potential DMI at the level of the major human drug transporters, like organic anion transporter 1(OAT1) and organic anion transporter 3 (OAT3), has recently become evident as well [137].

While a limited number of *in vitro* cell-based assays and *in vivo* rodent experiments have been performed to uncover the role of drug handling proteins in metabolic processes, these experiments can be technically challenging, time consuming, and labor intensive—and each has limitations in their application to humans. Virtual screening can aid in this process, but for many proteins, particularly membrane bound human transporters, determining substrate-transporter interactions has proven to be challenging due to the lack of crystal structures [80]. Though lacking in specific mechanism of action, by using SNP associations with metabolomics data, it is possible to prioritize potential protein-metabolite interactions in humans to further evaluate the possible physiological role of hundreds of genes.

Here, we combined genomic data targeting SNPs in drug transporter, DME, and related genes with non-targeted plasma metabolomics of over 2500 patients from the Framingham Offspring Cohort Exam 8 to link SNPs in these genes to the levels of circulating endogenous metabolites. Because the majority of circulating molecules are unknowns, we performed directed, non-targeted LC-MS approaches to specifically capture and assay small, polar, bioactive metabolites, including free fatty acids, eicosanoids and oxylipins, bile acids, fatty acid esters of hydroxy fatty acids, and other related metabolites of known and unknown chemistries. These types of metabolites have been shown to signal via cell surface G-protein coupled

receptors (GPCRs) and nuclear receptors and be critical for a host of physiologic processes [203, 204].

This work represents a step forward in understanding the individual and combined roles of ADME and other genes in endogenous metabolic processes. Of the several interactions reported here, some have been confirmed by independent *in vivo* or *in vitro* experiments, indicating that many novel SNP-metabolite associations likely have a functional protein-ligand relationship. We found metabolites that were linked to multiple SNPs on distinct genomic loci containing genes expressed in different cells and tissues, which raises the possibility of transporter and/or DME mediated remote communication via small molecule metabolites. We also analyzed the existing pharmaceutical GWAS to determine DMIs that may occur in patients with genes harboring certain SNPs involved in metabolism. The results indicate that a wide range of DMI can result at the interfaces of the Remote Sensing and Signaling (protein) network and the ADME (protein) network.

4.3 MATERIALS AND METHODS

4.3.1. Sample Population

Genotyping was performed on the Framingham Heart Study (FHS) Offspring Cohort Exam 8 (**Table 4.1**) [205].

4.3.2. Gene List

The initial gene list for targeted SNP-metabolite associations was constructed based on the Remote Sensing and Signaling Network reported previously based on a co-expression analysis [11]. This network included solute carrier (SLC), ATP-binding cassette (ABC), and several DME families. Among the DME families were subfamilies such as cytochrome P450s (CYPs), uridine 5'-diphospho-glucuronosyltransferases (UGTs), and sulfotransferases (SULTs). This list of multi-specific, oligo-specific and mono-specific proteins overlaps considerably with genes known to be involved in the absorption, distribution, metabolism, and excretion (ADME)

of drugs. The list was enlarged by considering other transporters and DMEs involved in ADME, as well as related transporters and enzymes based on their roles in ADME, handling of endogenous small molecules, or sequence homologies. In total, this resulted in consideration of 1131 genes (**Supplementary Table 4.S1**).

4.3.3. SNP identification

SNPs were included in the analysis if they were mapped to genes from **Supplementary Table 4.S1** using SnpEff [206]. Each SNP was associated with a reference SNP cluster ID (rsID) or a position on a chromosome. Those SNPs with an rsID were present in dbSNP version 151. All SNPs are in hg19 allele reference format.

4.3.4. Imputation

Several genotyping arrays (Affymetrix) were used to identify SNPs for the population. SNPs were imputed using Minimac3. SNPs associated with genes within an initial list of 1131 genes containing drug transporters and drug metabolizing enzymes (**Supplementary Table 4.S1**) were queried.

4.3.5. Metabolomics analysis

Metabolomic studies were performed using directed, non-targeted liquid chromatography-mass spectrometry (LC-MS) approaches to specifically capture and assay small polar 'bioactive' metabolites. These were deemed to have a higher likelihood of interacting with cell surface receptors involved in signaling. These include free fatty acids, eicosanoids and oxylipins, bile acids, fatty acid esters of hydroxy fatty acids, among hundreds of unidentified related metabolites [203, 204]. Metabolite levels are used as continuous traits with a mean of 0 and standard deviation of 1. Identified metabolites were confirmed through internal standards.

4.3.6. Statistical analysis

SNP-metabolite p-values were determined using linear mixed models (LMM) with an additive genetics model, where 0,1, and 2 indicate the number of effect alleles for each SNP in

the targeted set. The BOLT-LMM algorithm was used to account for age, gender, and other

factors [207]. Each metabolite-SNP association had a p-value and only statistically significant

associations are reported here. Overall, 673,141 statistically significant SNP-metabolite

associations were detected. The p-value cutoff for significance was set at $4.9*10^{-12}$.

4.3.7. Genomic loci

Genomic loci were defined by grouping SNPs that were located within 250,000 base pair

windows. SNPs on the same chromosome more than 250,000 base pairs apart were considered to

be on different genomic loci. Genes were then mapped to genomic loci if any portion of the gene

was within 10,000 base pairs of the genomic locus. For loci without any genes within 10,000

base pairs, the nearest gene was associated with the locus. The GRCh37 build was used for all

mapping. Genomic plots were generated using FUMA [208]. Manhattan plots were generated

using assocplots [209].

4.3.8. Tissue-specific enrichment

Tissue-specific enrichment for genes within genomic loci was calculated using the

TissueEnrich web-based tool (https://tissueenrich.gdcb.iastate.edu/) [210]. **Supplementary**

Table 4.S1 was used as the background gene list, and the 178 genes within the genomic loci was

used as the input gene list. Tissue expression was determined using the Human Protein Atlas.

4.3.9. Disease and pharmaceutical variant associations

The 'Variant and Clinical Annotations', 'Variant, Gene, and Drug Relationship Data',

and 'Clinical Variant Data' files were downloaded from the PharmGKB database

(https://www.pharmgkb.org/downloads, accessed on 23 August 2022) [211].

4.4 RESULTS

4.4.1. 77 genomic loci are linked to circulating levels of small, polar bioactive molecules

Plasma and DNA from each participant in the Framingham Offspring Exam 8 Cohort

were analyzed to identify relationships between specific genes and endogenous metabolites

(**Figure 4.1**). We focused on bioactive small, polar molecules, aiming to capture endogenous

small molecules that bind receptors involved in signaling. These include eicosanoids, fatty acids, and sex steroids that are known to interact with G-protein coupled receptors (GPCRs) and nuclear receptors (NRs) [212, 213].

We analyzed the plasma levels of thousands of unique metabolites and their associations with SNPs contained within a set of 1131 genes. Many of these genes were taken from a previously constructed co-expression network that is believed to reflect their roles mediating endogenous small molecule inter-organ communication, as described in the Remote Sensing and Signaling Theory (RSST). The original list of genes consisted of solute carrier (SLC) transporters, ATP binding cassette (ABC) transporters, DMEs, including CYPs, SULTs, and UGTs, and other drug-related genes that have multi-, oligo-, and mono-specific substrate specificity and are expressed in the gut, liver, kidney, and other tissues (**Supplementary Table 4.S1**). Their known substrates include a wide range of metabolites, signaling molecules, antioxidants, vitamins and cofactors, and gut microbe-derived metabolites.

In total, 673,141 statistically significant SNP-metabolite associations were reported, covering 8,634 unique SNPs and 7,326 unique metabolites (**Figure 4.2A**). The surveyed metabolites ranged from mass to charge ratios (M/Z) of 225.110 to 649.3938 and retention time (RT) values ranged from 0.6690834 to 6.988375 seconds. Each SNP was mapped to a genomic locus based on its position in the GRCh37 build of the human genome, as described in the Methods. We identified 77 distinct genomic loci, which covered 284 unique genes (**Supplementary Table 4.S2**). Each genomic locus was associated with a different set of genes, SNPs, and metabolites (**Figure 4.2B-G**). Genomic locus 15, containing the *UGT1A* genes, and genomic locus 54, containing *SLCO1B1, SLCO1B3, SLCO1B7,* and *SLCO1A2*, were associated with 42% and 33% of the total interactions, respectively. This is consistent with the functions of

these genes, as they are known, largely from *in vitro* work, to be among the most multi-specific transporters and enzymes of xenobiotics and metabolites, with dozens of unique substrates. The significant associations are reported in **Supplementary Table 4.S3**.

4.4.2. Tissue-specific enrichment of genes with SNPs shows overrepresentation of liver genes

Of the ~1000 surveyed genes, 178 contained SNPs that were present in our study and significantly associated with at least one endogenous metabolite. Tissue-specific enrichment revealed that the liver, breast, kidney, gallbladder, duodenum, and small intestine were over-represented within these genes (**Figure 4.3A**), using the 1131 genes from **Supplementary Table 4.S1** as the background set. The liver was the most highly enriched organ (adjusted -log-scaled p-value = 17.1) with 61 tissue-specific genes (**Figure 4.3B**). In the pharmaceutical literature, the tissues enriched with these genes are traditionally associated with ADME, with some exceptions. While not traditionally associated with drug ADME, the breast plays a role in the regulation of small molecule metabolites as it contains epithelial tissue that separates the blood and milk and expresses important DMEs of the glutathione S-transferase (GST) and UGT2B families. The next most enriched tissue was the kidney, followed by the gall bladder and intestinal tissues. The 178 metabolite-associated genes were largely enriched in the gut, liver, and kidney, consistent with previously identified roles in remote sensing and signaling of small, polar, bioactive metabolites and signaling molecules across the gut-liver-kidney axis [11] (**Figure 4.3C,D**).

4.4.3. Unidentified metabolites are potentially regulated by distinct genomic loci

Considering that most of the metabolites surveyed were unidentified (chemical identity unknown, but unique mass/charge ratio (MZ) and retention time (RT) combination), we aimed to understand which genomic loci worked collaboratively to regulate or modulate the levels of metabolites rather than focus on the metabolic role of the compound. Depending on the tissue expression and cellular localization of the implicated genes, these could be useful examples in

determining potential cases of inter and/or intra-organ communication and lead to a more mechanistic view. We identified five metabolites that were associated with four distinct genomic loci (**Table 4.2**). Metabolite 1116529 was the only metabolite not associated with both genomic loci 28 or 29 and was uniquely associated with genomic loci 15, 19, 48, 54 (**Figure 4.4**). Thus, we presented the genomic regional plots with the implicated SNPs associated with metabolite 1116529 as an example (**Figure 4.5**). Metabolite 1116529 was associated with loci containing the *UGT1A, UGT2B7, ABCC2,* and *SLCO1B1* genes, which are all multi-specific hepatic proteins known to handle metabolites and drugs. Genomic locus 15 includes several genes (*UGT1A6/7/8/9/10*), so it is difficult to associate any single gene with the resulting changes. Nonetheless, the *UGT1A* genes are primarily expressed in the kidney and liver. Two other implicated genes, *UGT2B7* and *ABCC2*, are also mainly expressed in the kidney and liver, whereas *SLCO1B1* is expressed only in the liver.

Even if unidentified, it is still possible to glean some hints of a metabolite's potential physiological role(s). The function of these proteins (uptake transporter, efflux transporter, and glucuronidation enzymes) and their different sites of expression within the liver (apical plasma membrane, basolateral plasma membrane, and cytosol) and kidney (apical plasma membrane, cytosol) support the view that these proteins work together to regulate the levels of this metabolite along the liver-kidney axis. We also investigated phenotypes related to SNPs in this genomic locus from dbSNP and the GWAS catalog and found that the SNPs associated with this metabolite were also linked to disorders of bilirubin excretion, serum 25-hydroxyvitamin-D levels, and testosterone levels [60, 214]. In addition to this metabolite, there are several other examples of unidentified metabolites associated with multiple genomic loci. With respect to these, 79 metabolites were linked to three distinct loci, including 25 unique combinations; 606

101

metabolites were linked to two distinct loci; and 6,636 metabolites were associated with only one genomic locus (**Supplementary Figure 4.S1**).

4.4.4. Circulating eicosanoids, fatty acids, and bile acids are impacted by SNPs in 18 genomic loci

While all 7326 measured metabolites had a unique metabolite ID, most had not had their chemical identity confirmed. However, 98 metabolites were identified by name, including eicosanoids, fatty acids, and several other signaling molecules. Even in this subset, some metabolites have not been unambiguously identified, but their general class is known. For example, EIC_45 represents a putative eicosanoid [204]. By limiting our analysis to the associations involving these identified metabolites, 762 SNP-metabolite associations were analyzed (**Supplementary Table 4.S4**). These associations spanned 18 genomic loci, with genomic locus 54 being associated with 62 identified metabolites, the most of any genomic loci surveyed here (**Figure 4.6**).

4.4.5. A putative eicosanoid is independently associated with SNPs in Phase I and II drug metabolism and transporter genes

As mentioned in previous sections, we were interested in those metabolites that were associated with multiple genomic loci, as they may be examples of genes involved in inter-organ or intra-organ communication contributing to the systemic levels of particular metabolites. The eicosanoid EIC_311 was the only identified metabolite associated with three unique genomic loci (Genomic loci 41, 54, 72) (**Figure 4.7**). These loci contained SNPs in the *CYP3A5, SLCO1B1/SLCO1A2,* and *SULT2A1* genetic regions, respectively. These proteins are primarily expressed in the liver and serve critical roles in drug metabolism. CYP3A5 is a Phase I drug metabolizing enzyme, SULT2A1 is a Phase II drug metabolizing enzyme, and SLCO1B1/SLCO1A2 are drug transporters (Phase III drug handling), suggesting that these genes may have a combined role in regulating this eicosanoid. Some of these genomic regions have

also been linked to blood metabolite levels, urine metabolite levels in chronic kidney disease, and cholelithiasis/cholecystitis [65, 215, 216]. In addition to EIC_311, eight identified metabolites were also associated with two distinct genomic loci (**Table 4.3**).

4.4.6. Conjugated sex steroids are strongly associated with SLC22 genes

While our main focus was on the potential shared function of genomic loci in regulating circulating metabolites, the associations between genomic loci and identified metabolites represented potential physiological roles for the implicated genes. The strongest associations in our study were between genomic locus 53 and conjugated sex steroids (**Figure 4.8**). This genomic locus contains a cluster of genes in the SLC22 family that are best known for their role in organic anion transport [167]. Recent functional studies have shown that five conjugated sex steroids directly interact with SLC22A24 *in vitro*, as well as in GWAS [217]. Here, we report that four similar metabolites (Putative_5a-Androstan-17b-ol-3-one glucosiduronate, Putative_Androstan-3-ol-17-one 3-glucuronide, Putative_Androstan-3-ol-17-one 3-glucuronide, and Putative_4-Androsten-17b-ol-3-one glucosiduronate) are associated with the genomic locus containing *SLC22A6, A8, A9, A10, A24,* and *A25*. The strongest associations involve Putative_5a-Androstan-17b-ol-3-one glucosiduronate. The SNPs rs78176967, rs142131421, rs113939203, and rs113497640, which had log-scaled p-values between -225 and 2-80, suggesting a strong functional relationship between one or many of the genes expressed on this locus and this metabolite.

4.4.7. SNPs in drug transporter and DME genes are pleiotropic and linked to multiple identified metabolites

Within our subset of surveyed genes were several that are known to be functionally related to multiple classes of drugs. For example, CYP3A4 is among the most promiscuous of the DMEs, with hundreds of drug substrates and dozens of endogenous substrates [218]. Among endogenous molecules detected by our metabolomics approach, we found that genomic locus 41,

which contained *CYP3A4*, *CYP3A5*, and others in the *CYP3A* family, was associated with

multiple identified metabolites, including bile acids, sex steroids, eicosanoids, and

prostaglandins. Furthermore, the multi-specific *SLCO* drug transporters in genomic locus 54

were associated with 62 metabolites, mainly eicosanoids and fatty acids. Genomic locus 70

harbored 16 genes, mainly in the *CYP4F* family, and was associated with 15 metabolites, mostly

eicosanoids and fatty acids. Although multiple types of genes are included in locus 70, such as

GPCRs, we expect these associations to be due to functional changes in the CYP4F family. The

CYP4F family is heavily involved in metabolism of fatty acids and their derivatives [219]. The

fact that many of the associations in this work have been validated in other studies, suggests that

the novel associations will prove useful in determining potential metabolic roles for the

implicated genes.

4.4.8. Implicated SNPs in endogenous metabolism have been reported to impact drug handling

As mentioned, many of the SNPs linked to the metabolites in our study have been

previously associated with the efficacy or toxicity of different drugs. This begets the question of

potential drug-metabolite interactions (DMI). This might also be expected because many of the

aforementioned genes are at the interfaces of the Remote Sensing and Signaling (protein)

Network and the overlapping ADME protein network. Variant-drug relationships were

downloaded from the PharmGKB database and compared to our data to predict potential DMI.

Ten SNPs were present in both our study and the PharmGKB database and were associated with

at least one drug (**Figure 4.9**). The most common SNP was rs4149056, which is present within

SLCO1B1. In addition to being linked to 50 unique identified metabolites (**Supplementary

Table 4.S4**), this SNP is also associated with affecting 21 unique drugs, including simvastatin,

lopinavir, and doxorubicin. Most of the SNP-drug pairs associated with SNPs in our study were

present in genomic locus 54, which is consistent with its role in the regulation of endogenous

metabolites. Indeed, SLCO1B1/SLCO1A2 are well known as multi-specific drug transporters with a wide array of both xenobiotic and endogenous substrates [220]. In addition to genomic locus 54, genomic locus 70 (containing the CYP4F genes) had the second most associations, with 14 unique metabolites and 6 unique drugs. As we discuss below, the use of SNPs linked with both drug handling and endogenous metabolism is likely to be useful for predicting clinically relevant drug-metabolite interactions.

4.5 DISCUSSION

The Remote Sensing and Signaling Theory emphasizes the role of multi-, oligo-, and mono-specific transporters, enzymes, and regulatory proteins in the homeostasis of endogenous metabolites, signaling molecules, antioxidants and other small molecules with "high informational content" in bodily fluids and tissues by mediating inter-organ and inter-organismal (gut microbe-host) communication [11]. These transporters and enzymes lead to the availability of these metabolites and signaling molecules in specific tissues and body fluids, often "setting up" the classical signaling events by GPCRs, nuclear receptors and kinases. Since many of the molecules involved in signaling via cell surface and nuclear receptors are small, polar, bioactive metabolites (e.g., free fatty acids, eicosanoids, bile acids, fatty acid esters of hydroxy fatty acids), we utilized non-targeted LC-MS methods that specifically capture these and other physiologically-important molecules [203, 204]. This approach also allowed us to explore both known and unknown chemistries of circulating molecules.

Many aspects of Remote Sensing and Signaling Theory are supported in model organisms, including mouse and fly [54, 58, 59, 72, 73, 77, 95, 144] and beginning to be supported in human studies [12, 137]. Key to the theory is the development of as comprehensive a parts list as possible—consisting, for instance, of interacting transporters and enzymes with their metabolite substrates. One approach to identifying the Remote Sensing and Signaling

(protein) Network has been through the creation and analysis of cross-tissue co-expression networks of multi-, oligo- and mono-specific transporters, enzymes and nuclear receptors [11]. This led to a preliminary gut-liver-kidney Remote Sensing and Signaling (protein) Network involved in endogenous metabolism that included, as hubs, many well-known SLC and ABC "drug" transporters and DMEs among its ~600 nodes. Thus, it was not surprising that there was similarity and overlap with a smaller network that specifically integrated ADME proteins [11]. However, it is important to keep in mind that the apparent physiological objective of the Remote Sensing and Signaling Network is the mediation of endogenous small molecule homeostasis while a large part of what the ADME network is presumed to mediate is the metabolism and elimination of drugs.

That said, a major goal here was to define the metabolites and signaling molecules regulated or modulated by multi-, oligo- and mono-specific transporters and enzymes in this Remote Sensing and Signaling Network. However, because of the considerable overlap in proteins of the Remote Sensing and Signaling Network and the ADME protein network, it was possible to consider whether drug-metabolite interactions might occur at the interfaces of the two networks [137].

Determining substrates of transporters or enzymes is typically done with *in vitro* assays or *in vivo* animal experiments [29]. *In silico* methods using experimental or predicted protein structures have also been used to predict potential substrates, most notably for enzymes [221-223]. Unfortunately, for membrane bound transporters, there are comparatively few crystal structures available, so protein-based predictions are more difficult to generate [80, 224]. GWAS or targeted SNP association studies in tandem with metabolomics represent another method for determining potential small molecules that may interact with proteins in a direct or indirect way

106

and can suggest a physiological role for these proteins in the modulation of plasma metabolite levels (**Figure 4.1**) [61]. Although *in vitro* or *in vivo* experiments are required to confirm the interactions, these results can, as described in this study, help broaden the list of potential *in vivo* interactions of endogenous metabolites with human transporters and drug metabolizing enzymes. Treating metabolite levels themselves as phenotypes can provide insight on the endogenous metabolic roles of genes and the intermediate processes they may participate in [63, 64, 195, 200].

By uncovering the molecular mechanisms of these proteins in physiological processes, we can improve our understanding of the roles of the hundreds of genes conventionally associated with drug ADME (absorption, distribution, metabolism, elimination), as well as others involved in broader aspects of small molecule homeostasis. We argue their role in endogenous small molecule homeostasis is their major role in humans and other organisms [59, 167], but because of the tremendous pharmaceutical and toxicological relevance of these genes, their role in endogenous physiology has largely been neglected. Here, we identified 77 genomic loci containing 284 unique genes (**Figure 4.2**) that were associated with the circulating levels of at least one endogenous, polar, bioactive molecule of the kind known to bind signaling receptors on the cell surface and in the nucleus.

Many of the surveyed genes are known to play a major role in drug metabolism and work together along the gut-liver-kidney axis (**Figure 4.3**) [72, 225, 226]. Typically, drugs are absorbed and enter the bloodstream via intestinal transporters. They then enter the liver through hepatic transporters, where the majority of enzymatic drug metabolism occurs. The modified compounds are then cleared or re-introduced to the bloodstream by efflux transporters. If the modified compounds re-enter the bloodstream, they are taken up, metabolized by DMEs in the

kidney, and ultimately cleared into the urine by renal transporters or re-introduced into the bloodstream. The same occurs for many small polar metabolites, signaling molecules, antioxidants, nutrients, natural products, gut microbe-derived metabolites, and vitamins. Thus, the remote communication between proteins expressed across these and other organs via small molecules is crucial to the regulation of endogenous metabolism and crosstalk along organ axes or organ systems, as is evident in bile acid and urate homeostasis [10, 158]. Defective inter-organ communication involving metabolite transporters as in the case of *OCTN2*, also considered a drug transporter, can lead to potentially lethal diseases such as Systemic Carnitine Deficiency [227].

While mainly studied for their roles in the ADME of drugs, here we show a number of examples of many of the same ADME proteins jointly contributing to the regulation of a single endogenous metabolite or multiple metabolites. As we have shown, this could involve as many as four transporters and/or enzymes (**Figure 4.4, Figure 4.5, Table 4.1**) of the Remote Sensing and Signaling Network potentially overlapping with drug handling proteins in the ADME network regulating a single metabolite. For example, among unidentified, unique metabolites, five metabolites were associated with four distinct loci. In addition, there were 79 metabolites associated with three distinct loci, including 25 distinct combinations of loci. Although most of these metabolites have yet to be fully defined in terms of chemical identity, the loci that influence their circulating levels include multi, oligo, and mono-specific transporters and enzymes, including well-known drug handling proteins. For instance, among the identified metabolites, the eicosanoid EIC_311 was associated with SNPs near *SLCO1B1, CYP3A5,* and *SULT2A1*, which are, respectively, a Food and Drug Administration (FDA) highlighted transporter, a Phase I DME and a Phase II DME, all on separate chromosomes (**Figure 4.7**).

These genes are heavily involved in ADME and also implicated in remote sensing and signaling via co-expression analysis and in vitro interactions with drugs and metabolites [11]. Understanding the full extent of the role of these genes can also help better understand drug-metabolite interactions (DMI). DMIs are often ignored in reference to drug side effects and adverse drug reactions, which can potentially be mitigated through better dosing of drugs, so as to not overly perturb the Remote Sensing and Signaling Network involved in small molecule homeostasis across cells, tissues, organs, and organ systems.

While most of the SNP-metabolite associations involved unidentified metabolites, the 98 identified metabolites and their associations with specific SNPs include well-known physiological protein-metabolite interactions (**Figure 4.6**). For example, the *UGT1A* locus, which encodes multi-specific enzymes involved in Phase II drug metabolism, is also known to modify bilirubin and mutated in human Gilbert's Syndrome, and that interaction is reflected in our results [228]. Likewise, *SLC22A9/10/24/25*, which appear to be relatively mono-specific or oligo-specific in one of the SLC22 transporter subgroup [167], were associated with conjugated sex steroids (**Figure 4.8**). The role of SLC22A24 in human steroid metabolism and disease has been previously reported [217]. The *CYP3A, CYP4F, CYP2C* genes, including multi-specific and oligo-specific enzymes, are known to generate and degrade signaling eicosanoids and fatty acids, which is reflected in our results here [229-232]. The multi-specific hepatic "drug" transporter OATP1B1 (SLCO1B1), associated with statin myopathy [233], also had several associations with a wide array of small molecules, including eicosanoids, bile acid conjugates, and fatty acids, which is consistent with its known function [234]. FAAH, an enzyme that might be considered oligo-specific, is known to modulate the levels of endocannabinoids in tissues, and in this study, we show it also influences the levels of endocannabinoids in plasma [235].

These existing relationships suggest that many unexplored associations between SNPs and identified metabolites may be of great physiological and clinical importance. Among the unexplored relationships with no existing literature to date are those between steroid 5-alpha reductase 2 (*SRD5A2*) and xanthine dehydrogenase (*XDH*) with Allo_Tetrayhydrocortisol in genomic locus 11, the *SLC17* family (transporters of phosphate and other organic anions) with acetyltryptophan in genomic locus 32, and several others. Although we are able to associate a gene family with a class of metabolites, more in depth studies would be required to confirm the mechanistic relationship between these proteins and metabolites, as well as their joint role in regulating certain metabolic pathways. The identification of the very large number of unnamed metabolites will also allow the design of more functional assays to better define the metabolic role of drug transporters and DMEs and their potential role in DMIs.

The SNPs in genes that are not known to be functionally related to the ADME of drugs or the handling of endogenous metabolites indicate that certain SNPs can indirectly impact the levels of plasma metabolites independent of transport and enzymatic activity. Of the three genomic loci (8, 49, and 56) that do not contain any transporters or enzymes, each has a different potential mechanism for regulating the levels of circulating compounds. The polycystic kidney disease 2-like 1 *(PKD2L1)* gene in genomic locus 49 was linked to 91 metabolites, including three named eicosanoids. This gene codes for a calcium channel that is involved in signaling, development, and taste; yet its direct association with any polar bioactive molecules has yet to be reported [236]. It is expressed in numerous tissues, and the relatively large number of unique metabolites it is associated with suggests that general calcium signaling can have important consequences on the plasma metabolome. Genomic locus 56 contained *HNF1A* (hepatocyte nuclear factor 1 alpha), a nuclear receptor activated by signaling ligands, *HNF1A-AS1,* and

C12orf43. The open reading frame gene is understudied, but HNF1A and HNF1A-AS1 play roles in transcriptional regulation. Indeed, HNF1A regulates many ADME related genes in metabolically active organs and thus, can impact circulating metabolite levels (amino acids, bicarbonate, sugars) [237, 238]. Genomic locus 8 contains *NOS1AP*, a gene that binds to NOS1 for signaling purposes [239]. We examined SNPs in the NOS1 gene but found no significant metabolite associations. This suggests that NOS1AP, perhaps through the regulation of NOS1-mediated signaling, can modulate more complex interactions that ultimately lead to altered levels of plasma metabolites.

The field of pharmacogenomics is expected to play a major role in personalized medicine in the future, as drug administration and dosage can be more appropriately determined with knowledge of a patient's genome [240, 241]. Many drugs are taken up into the liver by drug transporters (e.g., SLCO family) and then metabolized by Phase I and Phase II DMEs before being eliminated through drug transporter-mediated mechanisms, such as members of the SLC22 family in the kidney. It is important, however, to understand the potential metabolic dysregulations that can stem from existing drugs and entities in the drug development pipeline. Drug targets differ depending on the intended function, but the proteins involved in the ADME processes overlap greatly with those regulating key processes in endogenous physiology (e.g., bilirubin metabolism, eicosanoids, bioenergetics). Indeed, the Remote Sensing and Signaling Theory argues that drugs often "hijack" endogenous pathways involved in remote organ communication and gut microbe-host communication. Thus, common adverse drug reactions, drug side effects, and drug-induced metabolic diseases may be caused by the competition between drugs and metabolites at the level of so-called drug transporters and drug metabolizing enzymes involved in key biochemical pathways. By comparing the previously determined role of

SNPs via the PharmGKB database, we related our analysis to potential drug-metabolite interactions. For example, the rs4149056 SNP in the *SLCO1B1* gene affects drug-response, as well as several bioactive molecules. If a patient has this SNP, treatment with a drug impacted by this SNP may exacerbate the metabolic consequences. Within our dataset, we identified 10 SNPs with evidence of potential drug-metabolite interactions (**Figure 4.9**).

As knowledge on the role of ADME genes in endogenous metabolic processes increases, more will likely be identified. It is worth emphasizing again that the untargeted metabolomics approach used here focused on small, polar, bioactive metabolites, both identified and unidentified, likely to interact with GPCRs, nuclear receptors and other signaling protein—and that they were significantly associated with SNPs in multiple distinct genomic loci. We have also presented Remote Sensing and Signaling Theory as a framework for understanding communication between organs through the regulated expression and function of multi-specific, oligo-specific, and (relatively) mono-specific proteins, such as drug transporters, drug metabolizing enzymes, and their relatives [7, 8, 11]. The broad substrate specificity of "drug" transporters and "drug" metabolizing enzymes mainly refers to pharmaceutical products—often with very different structures and mechanisms of actions, but this multi-specificity likely also applies to endogenous metabolites, as is clear with the organic anion transporters (OATs), SLC22A6 and SLC22A8 [6]. It is useful to note here that oligo-specific and mono-specific close relatives of the well-known drug transporters (OATs and organic cation transporters (OCTs)) are strongly implicated in the handling of metabolites like urate (SLC22A12) and carnitine (SLC22A5, SLC22A15/16). This fact emphasizes a main concept in the Remote Sensing and Signaling Theory—that multi, oligo, and mono-specific transporters and enzymes work within

and between organs to optimize endogenous metabolism in cells, tissues, organs, and multi-organ systems [6, 9, 140].

4.6 SUPPLEMENTARY MATERIAL

The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/metabo13020171/s1.

Supplementary Figure S4.1: Five unique metabolites were associated with 4 unique loci, 79 unique metabolites were associated with 3 unique loci, 606 unique metabolites were associated with 2 unique loci, and 6636 metabolites were associated with one locus.

Supplementary Table S4.1: Original list of genes selected for targeted SNP associations. These genes include transporters, enzymes, and related proteins that are known to handle small molecules or related to proteins that do.

Supplementary Table S4.2: Genomic locus and gene assignments that the detected SNPs associated with circulating metabolites map to. Genes are listed by their HUGO gene nomenclature committee (hgnc) symbols. ENSEMBL gene IDs are also listed in a separate column. The start and end position of each gene on the chromosome is listed. All positions come from the GRCh37 build.

Supplementary Table S4.3: All unique statistically significant SNP-metabolite associations detected are listed. Snp: Single nucleotide polymorphism. Snp_cpra: Single nucleotide polymorphism with chromosome position and reference and alternate alleles. mtb: Metabolite id. MZ: Mass to charge ratio. RT: Retention time. chr: Chromosome. pos: Position. ref: Reference allele. alt: Alternative allele. pvalue: P-value indicating statistical strength of association between SNP and metabolite. beta: Beta coefficient for fit. se: Standard error. alt_freq: Alternative frequency.

Supplementary Table S4.4: All unique statistically significant SNP-metabolite associations involving identified metabolites. Identity: Putative endogenous metabolite with the unique combination of MZ and RT. Genomic Locus: Assigned genomic locus listed in Supplementary Table S2. Snp: Single nucleotide polymorphism. Snp_cpra: Single nucleotide polymorphism with chromosome position and reference and alternate alleles. mtb: Metabolite id. MZ: Mass to charge ratio. RT: Retention time. chr: Chromosome. pos: Position. ref: Reference allele. alt: Alternative allele. pvalue: P-value indicating statistical strength of association between SNP and metabolite. LogP: Logarithmic value of p-value. beta: Beta coefficient for fit. se: Standard error. alt_freq: Alternative frequency. Mass_error: Error bars for accuracy of mass to charge ratio. RT_error: Error bars for accuracy of retention time. SMILES (if available): For each entry in identity, the isomeric SMILES sequence is listed. For metabolites with unknown identity (e.g., EIC_311), NA is listed.

4.7 AUTHOR CONTRIBUTIONS

Conceptualization, Sanjay Nigam; Data curation, Jeramie Watrous and Tao Long; Formal analysis, Jeffry Granados; Funding acquisition, Sanjay Nigam; Methodology, Jeramie Watrous, Tao Long and Susan Cheng; Project administration, Susan Cheng and Mohit Jain; Resources, Mohit Jain and Sanjay Nigam; Supervision, Mohit Jain and Sanjay Nigam; Visualization, Jeffry Granados; Writing – original draft, Jeffry Granados; Writing – review & editing, Jeffry Granados, Sara Rosenthal, Mohit Jain and Sanjay Nigam.

4.8 FUNDING

4.9 INSTITUTIONAL REVIEW BOARD STATEMENT

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review of boards at Boston University Medical Center and the University of California San Diego (Protocol 2015P001069).

4.10 INFORMED CONSENT STATEMENT

Participants gave informed consent to participate in the study before taking part.

4.11 DATA AVAILABILITY STATEMENT

The data presented in this study are available in Supplementary Materials.

4.12 ACKNOWLEDGMENTS

The authors would like to thank Sapient. Some figures were generated using Biorender.

4.13 CONFLICT OF INTEREST

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Table 4.1 Summary statistics of surveyed participants.

Participants are from the Framingham Offspring Cohort Exam 8.

Category	Value
Participants	2,886
Men	1,315
Women	1,571 (54.4%)
Age	66 ± 9 years
Body Mass Index (BMI)	28.3 ± 5.4 kg/m^2

Table 4.2 Five of 7326 unidentified metabolites are associated with four unique combinations of genomic loci.

These metabolites are associated with 4 unique combinations of genomic loci. The mass to charge ratio (MZ) and the retention time (RT) of each the 5 metabolites are listed and indicate that metabolites 1272586 and 1291919 are likely to be very similar compounds. Genomic loci 28 (*ACSL6*), 29 (*SLC22A4/5*), 46 (*SLC16A9*), and 54 (*SLCO1B1/3/7, SLCO1A2*) are associated with more than one metabolite, and genomic loci 4 (*SLC44A5*), 15 (*UGT1A6/7/8/9/10*), 19 (*UGT2B*), and 31 (*ECI2*) appear only once. The full list of genes associated with each locus are present in **Supplementary Table 4.S2.**

Genomic Locus 1	Genomic Locus 2	Genomic Locus 3	Genomic Locus 4	mtb	MZ	RT
4	28	29	54	1380594	284.2233	4.248833
15	19	48	54	1116529	607.3553	3.428667
28	29	46	54	1272586	282.2076	3.878833
28	29	31	46	1291919	282.2085	3.959
28	29	46	54	1592026	310.2399	5.155334

Table 4.3 Eight identified metabolites are associated with two distinct genomic loci.

Only one identified metabolite, an eicosanoid identified as EIC_311, was associated with 3 distinct loci, but eight others were associated with two, suggesting a more specific regulation. Some metabolites are presented twice because their identities are expected to be the same despite minor differences in their MZ or RT values.

Identity	Genomic Locus 1	Genomic Locus 2	MZ	RT
Eicosanoid_13,14-dihydro-15-keto-tetranor-PGE2 [M-H]	3	32	297.1744	1.816083
Eicosanoid_12-HHTrE [M-H+Acetate]	47	54	339.2178	3.669167
Eicosanoid_12-HHTrE [M-H+Acetate]	47	54	339.2197	3.766292
Putative_N-Oleoyl-L-serine	2	54	368.2847	6.353979
Putative_N-Oleoyl-L-serine	2	54	368.286	6.253
Endocannabinoid_Oleoyl Ethanolamide [M-H+Acetate]	2	70	384.3096	6.401
Endocannabinoid_Oleoyl Ethanolamide [M-H+Acetate]	2	70	384.3174	6.479625
Putative_Androstan-3-ol-17-one 3-glucuronide	41	53	465.2491	2.152167
Putative_Androstan-3-ol-17-one 3-glucuronide	41	53	465.2492	2.0535
Putative_Androstan-3-ol-17-one 3-glucuronide	41	53	465.2498	2.2015
Putative_5a-Androstan-3a,17b-diol-17b-glucuronide	19	53	467.2574	2.035
Putative_5a-Androstan-3a,17b-diol-17b-glucuronide	19	53	467.2634	1.89625
Putative_Chenodeoxycholic acid 24-acyl-b-D-glucuronide	19	54	567.3177	2.890625
Putative_Chenodeoxycholic acid 24-acyl-b-D-glucuronide	19	54	567.3182	2.76575
Putative_1,3,5(10)-Estratrien-3,17b-diol diglucosiduronate	41	54	623.3406	2.713333
Putative_1,3,5(10)-Estratrien-3,17b-diol diglucosiduronate	41	54	623.342	2.58075
Putative_1,3,5(10)-Estratrien-3,17b-diol diglucosiduronate	41	54	623.3441	2.540667
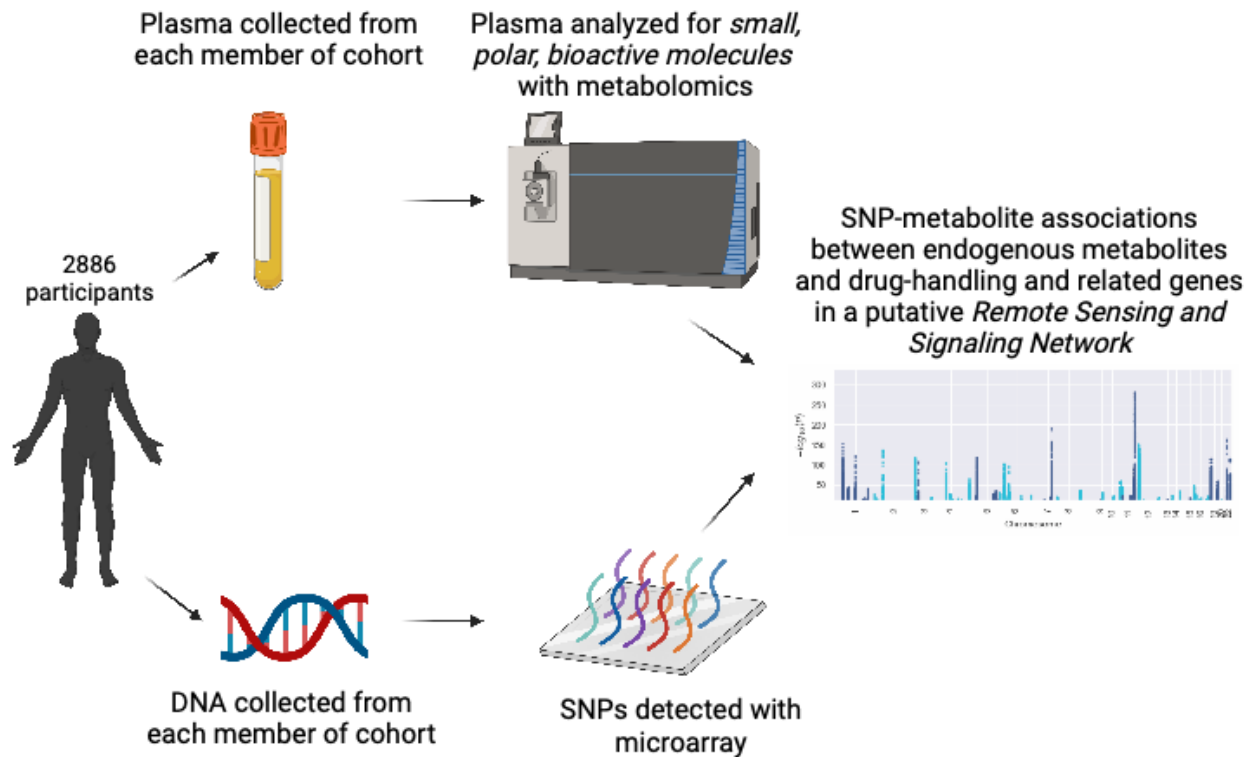
Figure 4.1 Schematic for data acquisition and subsequent analysis.

Plasma was collected from each patient and analyzed by liquid chromatography/mass spectrometry, which sought to capture small, polar, bioactive molecules presumed most likely to be involved in signaling via cell surface and other receptors. DNA was collected from participants and SNPs within a subset of genes were associated with the levels of plasma metabolites.

Figure 4.2 The targeted SNP association study linked SNPs in drug-related genes, like enzymes and transporters, to the circulating levels of small, polar, bioactive molecules.

A) Manhattan plot showing the targeted SNP-metabolite associations on the genome. The chromosome and relative genomic location are marked on the x-axis, and the log-scaled p-value is marked on the y-axis. Each point in the plot represents an association between the SNP and a measured metabolite. **B)** Schematic showing the unique genes associated with genomic locus 70, where an edge represents a gene located within or near the genomic locus, as described in the Methods. **C)** The 77 distinct loci were associated with 284 unique genes, including pseudogenes. Loci were associated with different numbers of unique genes ranging from 19 to 1. **D)** Schematic showing the unique SNPs associated with genomic locus 9, where an edge represents an SNP located within the genomic locus. **E)** With respect to SNPs, 8634 unique SNPs were detected, with nearly each loci containing unique SNPs. Some genomic loci, such as genomic loci 47 and 54, were associated with over 1000 distinct SNPs. The number of unique SNPs for each locus ranged from 1149 to 1. **F)** Schematic showing the unique metabolites associated with genomic locus 5, where an edge represents a statistically significant association. **G)** 7326 unique, small, polar, bioactive metabolites were measured in the plasma of the participants. Genomic loci 15 and 54 were associated with the highest number of unique metabolites (2059 and 3014, respectively). Unique metabolites associated with each locus ranged from 3014 to 1.
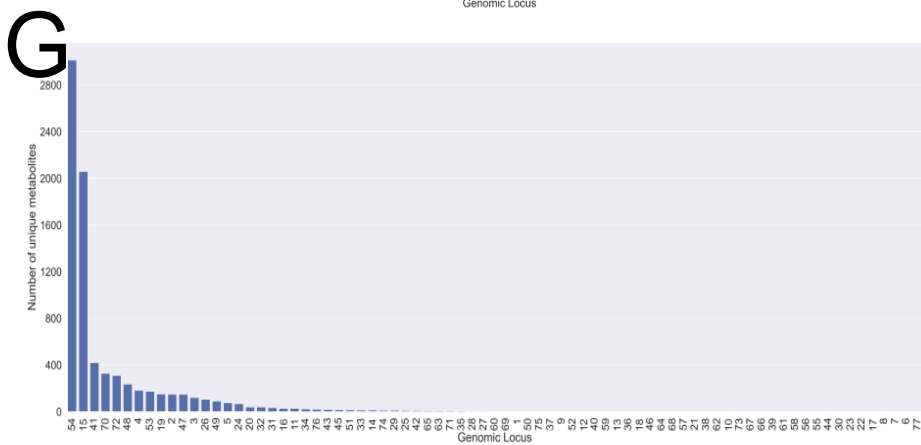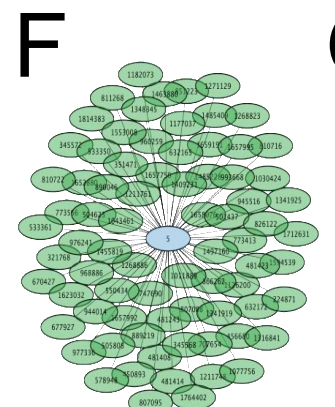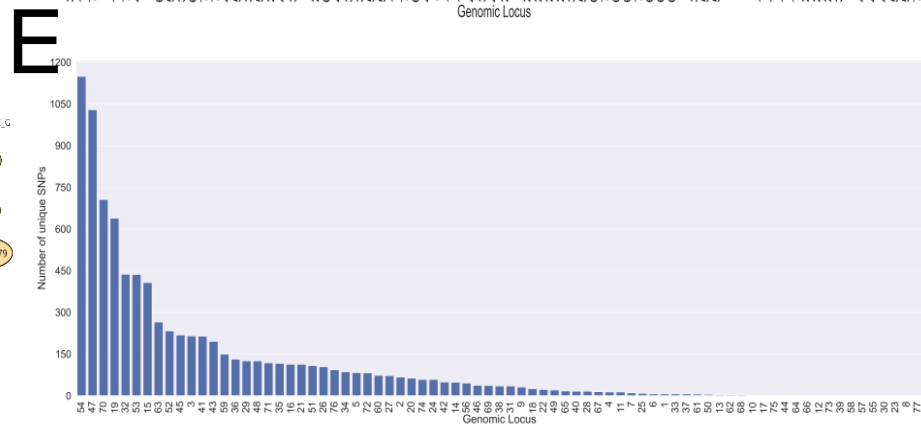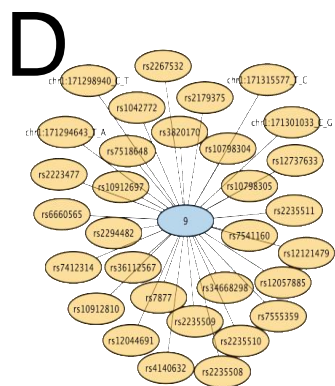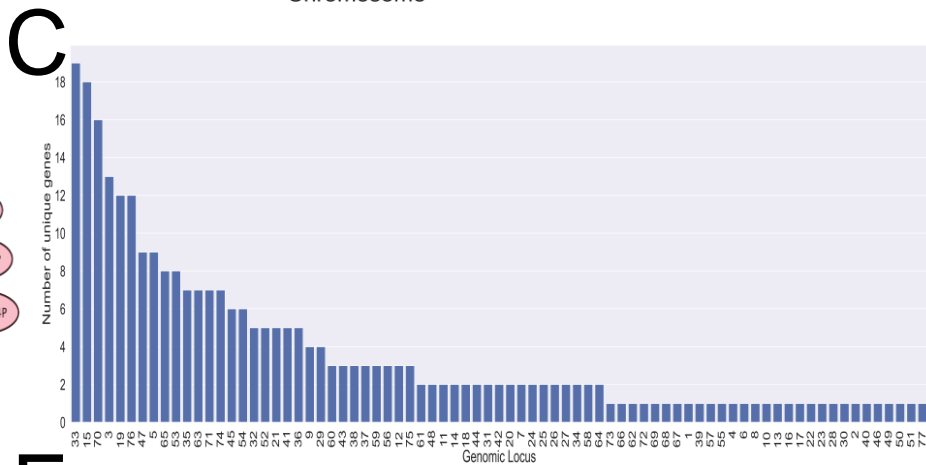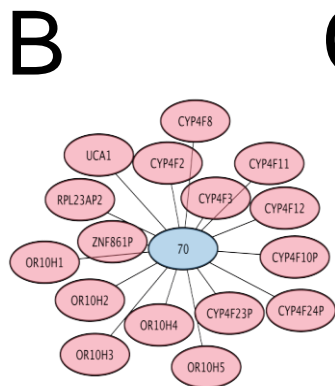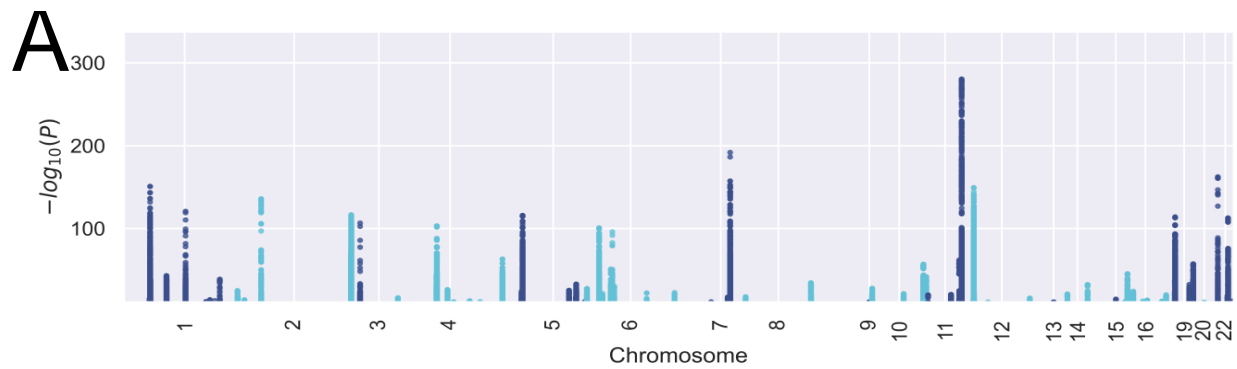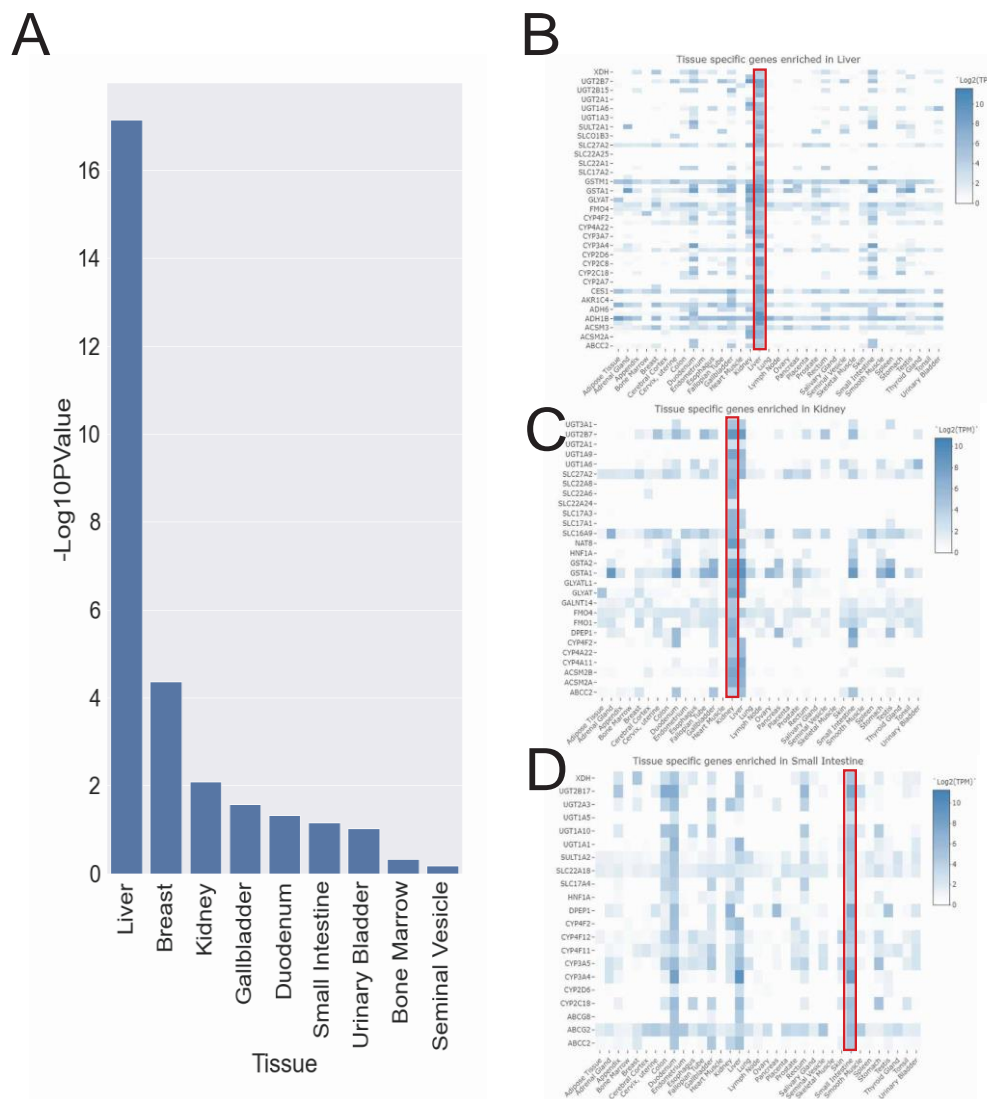
Figure 4.3 Tissue specific enrichment reveals over-representation of organs active in ADME.

A) The genes with significant associations with metabolites were compared to the reference genes listed in **Supplementary Table 4.S1**. The liver, breast, kidney, gallbladder, duodenum, small intestine, urinary bladder, bone marrow, and seminal vesicle all had significant enrichment. **B)** Genes enriched in the liver. **C)** Genes enriched in the kidney. **D)** Genes enriched in the small intestine.
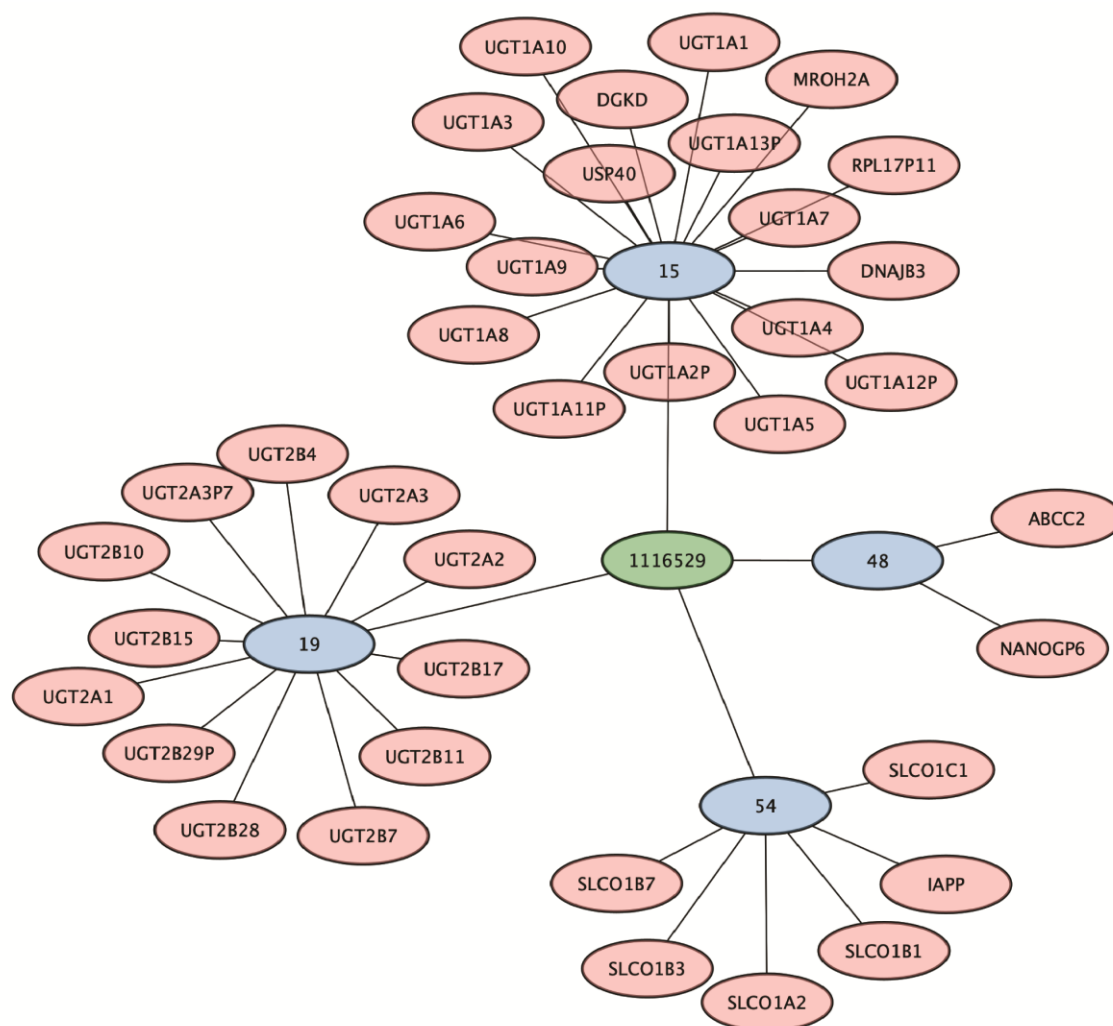
Figure 4.4 Metabolite 1116529 (MZ: 607.3553, RT: 3.428667) is associated with genomic loci 15, 19, 48, and 54.

SNPs within the aforementioned genomic loci are all linked to metabolite 1116529. Genomic locus 15 contains regions relating to the *UGT1A* genes, genomic locus 19 contains regions relating to the *UGT2B7* gene, genomic locus 48 contains regions relating to the *ABCC2* gene, and genomic locus 54 contains regions relating to the *SLCO* genes.
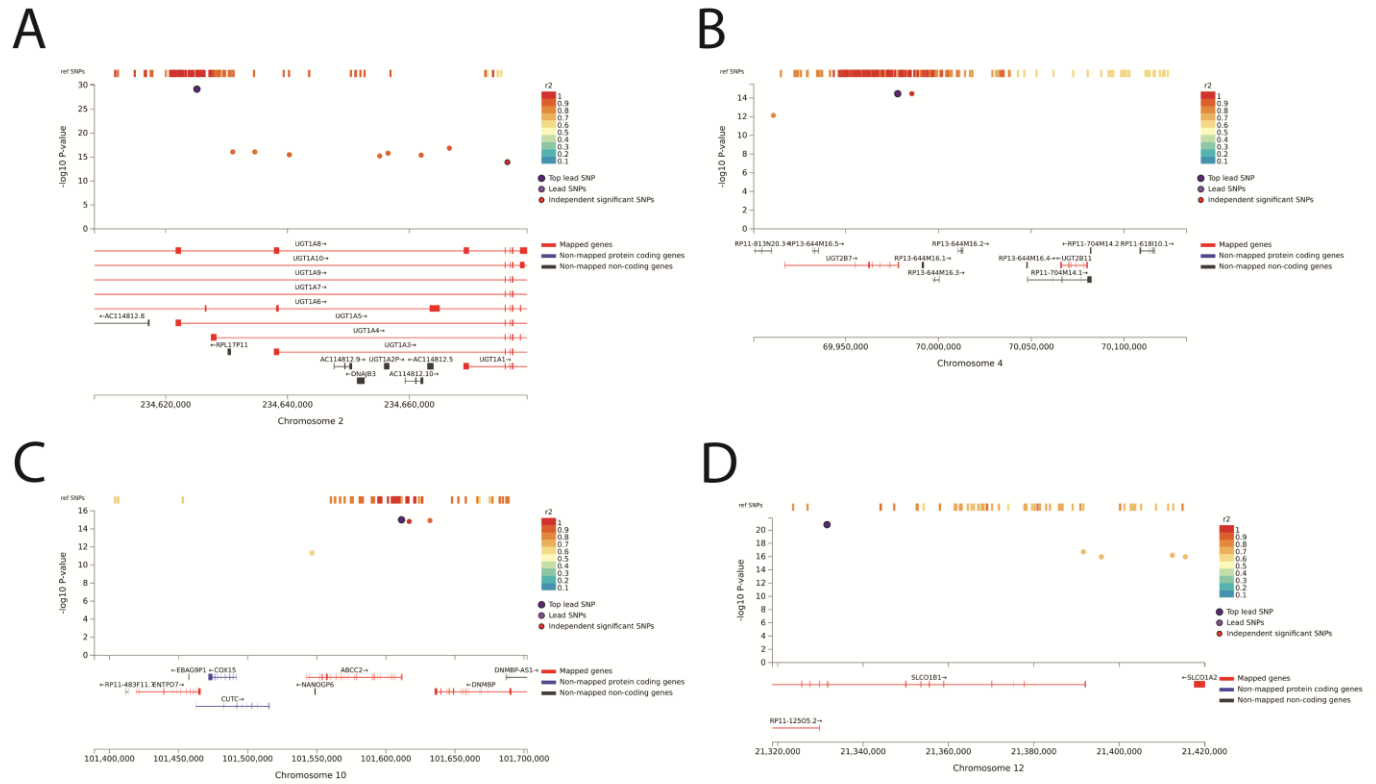
Figure 4.5 Genomic regions associated with metabolite 1116529, which is associated with 4 distinct loci.

A) Genomic locus 15. **B)** Genomic locus 19. **C)** Genomic locus 48. **D)** Genomic locus 54. In all panels of the figure, the SNPs associated with the levels of circulating metabolite 1116529 are represented by points, where purple points refer to the top lead SNP, and other SNPs are represented by points colored by their r2 value. The r2 value, which represents phenotypic variation, is high in these regions and reference SNPs that have previously been analyzed. The nearest mapped genes are shown below each plot.

Figure 4.6 A total of 762 SNP-metabolite associations with identified metabolites were reported, with 18 unique genomic loci and 98 identified metabolites.

Values in the heatmap represent the log of the p-value of each SNP-metabolite association, with a lower value indicating a stronger relationship between the genomic locus and the metabolite. For improved visualization, the identity of the metabolites in the x-axis has been shortened to include only the first 50 characters. The colorbar showing log-scaled p-values has been adjusted to improve visualization, and the highest values go beyond -100. The full names for each identified metabolite are present in **Supplementary Table 4.S4**.

Figure 4.7 Several identified metabolites were associated with multiple distinct loci.

Genomic loci are indicated by blue nodes. All green and purple nodes are identified metabolites. For improved visualization, only those associated with multiple distinct loci are shown in purple and have their identities shown in the figure. The identities for all the green metabolite nodes are listed in Supplementary Table S4. Notably, the eicosanoid EIC_311 is associated with 3 distinct genomic loci, consisting of genomic locus 41 (containing regions relating to the CYP3A genes), genomic locus 54 (containing regions relating to the SLCO genes), and genomic locus 72 (containing regions relating to the SULT2A1 gene).
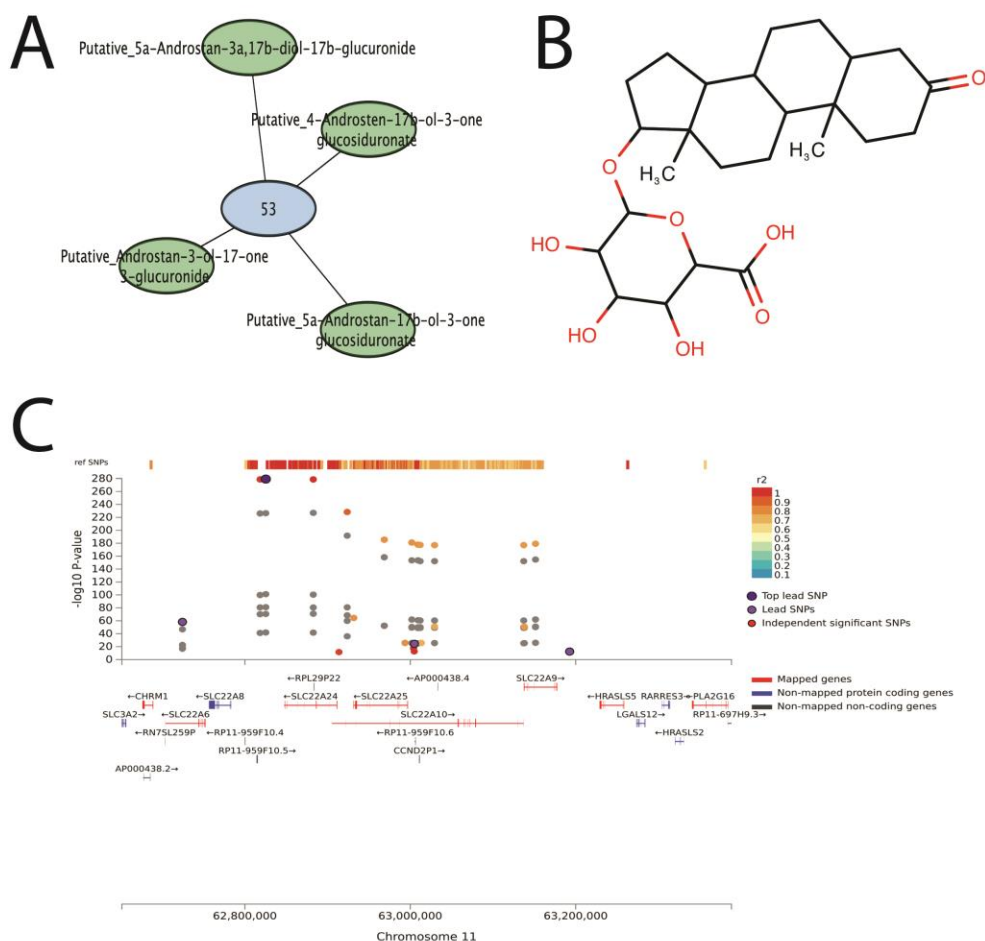
Figure 4.8 Genomic locus 53, containing SLC22 genes, is associated with conjugated sex steroid hormones.

A) The SLC22 genes, *SLC22A6*, *SLC22A8*, *SLC22A9*, *SLC22A10*, *SLC22A24*, and *SLC22A25* are highly associated with circulating levels of 5a-Androstan-17b-ol-3-one glucosiduronate, 4aAndrostan-17b-ol-3-one glucosiduronate, Androstan-3-ol-17-one 3-glucuronide, and 5a-Androstan3a,17b-diol-17b-glucuronide. The specific associations between SNPs and identified metabolites are listed in **Supplementary Table 4.S4**. **B)** The chemical structure of 5a-Androstan-17b-ol-3-one glucosiduronate is shown as a representative example of the metabolites potentially regulated by these transporter genes. **C)** The SNPs shown are associated with the levels of any implicated metabolites, where purple points refer to the top lead SNP, and other SNPs are represented by points colored by their r2 value. The r2 value, which represents phenotypic variation, is high in these regions and reference SNPs that have previously been analyzed. The nearest mapped genes are shown below each plot. SNPs, which are not in linkage disequilibrium of any significant independent lead SNPs in the selected region, are colored grey.

127
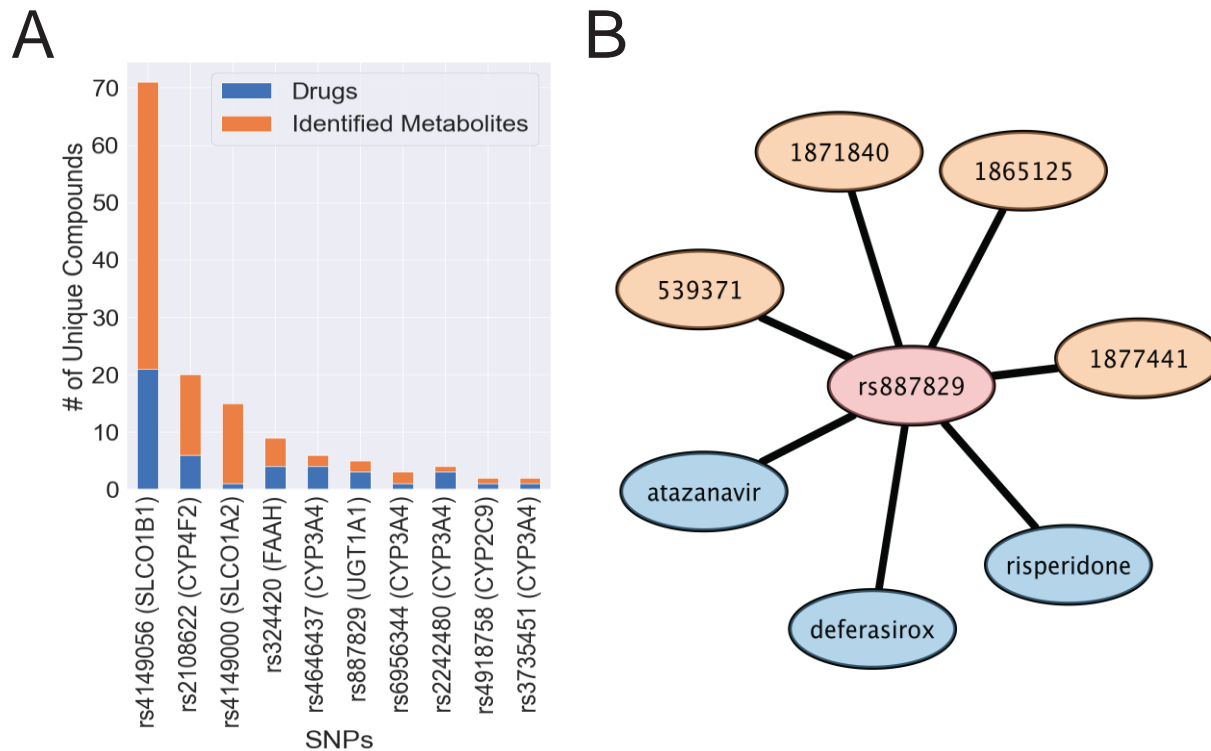
Figure 4.9 Some SNPs in ADME genes are involved in the regulation of both drugs and endogenous metabolites.

A) Certain SNPs are associated with levels of multiple identified metabolites in our study and with several drugs from other independent studies. Rs4149056, a SNP that impacts the function of SLCO1B1, is the most frequent SNP with respect to metabolites and drugs. **B)** rs887829 (*UGT1A1*) is an example of a SNP that is associated with three different drugs and identified metabolites that are all related to bilirubin (**Supplementary Table 4.S4).** The SNP is shown as a pink node, the drugs are blue nodes, and the identified metabolite IDs are shown as orange nodes.

CHAPTER 5: IN SILICO APPROACHES FOR DETERMINING TRANSPORTER FUNCTION

5.1 ABSTRACT

Organic anion transporter 1 (OAT1) is a renal drug transporter known for its role in the clearance of pharmaceutical products. However, recent studies have shown that OAT1 handles multiple other classes of small molecule compounds, such as endogenous metabolites and natural products. In this work, we analyze the full extent of OAT1 function using ligand and protein-based models. Determining transporter function has largely relied upon in vivo and in vitro experiments, which can be expensive, time-consuming, and suffer from inter-lab variability. As such, there is a need for computational models. We first use clinical in vivo plasma and urine metabolomics from participants treated with probenecid, an OAT-inhibiting drug, to develop binary classification machine learning models to determine likely OAT-mediated compounds. These models were trained on endogenous metabolite data and were tested on a set of drugs. While these models were largely ineffective in determining OAT-mediated drugs, they led us to characterize the total chemical space covered by OAT1. We curated a list of OAT1-interacting molecules and clustered them based on molecular descriptors. Upon revealing multiple distinct clusters, including some containing combinations of drugs, endogenous metabolites, and natural products, we aimed to analyze potential different binding mechanism using the predicted protein structure of human OAT1. We performed ligand docking studies to analyze protein-ligand interaction fingerprints, revealing specific sets of residues implicated in the handling of traditional OAT1 probes. Overall, this work represents a step forward in analyzing the chemical space associated with a single multispecific drug transporter from a multi-scale perspective.

5.2 INTRODUCTION

Organic anion transporter 1 (OAT1/SLC22A6) is a renal transporter primarily expressed on the basolateral side of the proximal tubule in the kidney. OAT1 has been mainly studied for its role in the excretion of dozens of small molecule drugs, including antibiotics, antivirals, NSAIDs, and many other classes of drugs [3, 99, 242, 243]. Recent work by our group and others, however, has shown that the specificity of OAT1 with respect to interacting molecules is not limited to drugs [5, 6, 21, 140, 164, 197]. Endogenous metabolites, natural products, toxins, and other classes of small molecules interact with OAT1 in vitro and in vivo [8, 42, 77, 95, 100, 101, 154, 171]. To date, the majority of these results have come from in vitro transport assays, but in vivo metabolomics results from humans treated with OAT1-inhibiting drugs and *Oat1* knockout mice have recently provided novel data with increased physiological relevance to these results [54, 73, 137]. The dozens of compounds associated with OAT1 are characteristic of some multi-specific proteins, such as drug transporters and drug-metabolizing enzymes.

Understanding the molecular basis for transporter interaction with a variety of small molecules is one of the active questions in the drug transporter field. Experimental evidence indicates broad function for drug transporters like OAT1, but the specific nature of these interactions, like the molecular descriptors or the protein residues, are relatively unexplored. Several types of models, including ligand-based, protein-based, and systems biology-based models, have been applied to better understand the complete function of OAT1 [76, 77, 244, 245]. Ligand-based models have been useful, but the bulk of these models rely upon results from in vitro assays, where dozens to hundreds of experiments must be conducted to generate sufficient data with which to train quantitative structure activity relationship (QSAR) or machine learning models. Regression models can be built to predict the K_i or K_m of a particular compound or classification models can be built to predict whether a compound is an inhibitor/substrate.

Pharmacophore models, which combine structurally similar compounds to create superimposed structures, have also been applied to understand OAT1/3 function [76, 77]. While these types of models are useful, they are dependent on the results of in vitro assays, which do not recapitulate key physiological factors.

In addition to ligand-based models, protein-based models have been used to better understand protein-ligand interactions. These kinds of models have been difficult to generate for drug transporters because of their localization to the plasma membrane makes determining experimental structures very difficult [80]. Previous work with OAT1 has emphasized predictions based on bacterial homologs [86]. However, advances in artificial intelligence have generated high quality structure predictions for transmembrane proteins based on existing structures, enabling a new kind of analysis [87]. The structure of OAT1 has been analyzed in part, but not extensively in the context of protein-ligand interactions [89, 91]. It has long been thought that multi-specific drug transporters have multiple modes of action, given the diversity of the small molecule structures, including compounds with a variety of origins and mechanisms of action, the transporters interact with, but this has yet to be fully explored for OAT1.

In this work, we first build ligand-based machine learning models based on endogenous metabolite data from the plasma and urine of a clinical probenecid administration study to predict whether other small molecule endogenous metabolites can be predicted to be OAT-mediated. Most machine learning models for transporter function have been developed using in vitro data. Here, we leverage existing in vivo metabolomics data to develop binary classification models that can predict whether compounds will be altered in vivo. We focused on clinically relevant data using plasma and urine data from humans treated with probenecid, as our model's predictions could be validated with clinical drug-drug interactions involving probenecid. The

combination of blood and urine data with knowledge of probenecid's mechanism of action implies that the compounds have been impacted by inhibition of OAT1 or OAT3. After testing the models with small molecule drugs, the results suggested that endogenous metabolite data were not sufficient to predict drug-related function.

To address this, we then aimed to understand the entire chemical space covered by OAT1 by focusing on all compounds with in vivo or in vitro evidence. We applied unsupervised learning techniques to better explore the chemical space associated with a single transporter, OAT1. These methods revealed multiple diverse clusters, suggesting different binding mechanisms for OAT1. To explore the possibility of multiple mechanisms, we used the predicted human OAT1 protein structure by AlphaFold2. This protein structure was embedded in a lipid bilayer plasma membrane for physiological relevance and docking experiments were performed to better understand how structurally diverse compounds interact with putative binding pockets. This work provides a multi-scale framework for understanding the full extent of drug transporter function.

5.3 RESULTS

5.3.1 Machine learning binary classification models are built using in vivo metabolomics data

Our first step to developing clinically relevant machine learning models was to leverage the plasma and urine metabolomics data our group had previously acquired (**Figure 5.1A**) [137]. In this experiment, healthy human participants received an oral dose of probenecid, an OAT-inhibiting drug. Over 1,000 metabolites were measured, including 400 with known chemical structures measured in both the plasma and urine. We were particularly interested in those that were significantly elevated in the plasma and significantly decreased in the urine because they are likely OAT-mediated due to inhibition of the blood-facing proteins OAT1 and OAT3 by probenecid. With these data, we aimed to generate models that can predict probenecid-sensitive

drug-drug interactions in vivo, furthering our understanding of OAT function. These models represent an advance in three distinct ways: 1) they do not require a large number of laborious in vitro assays, as metabolomics simultaneously measures the circulating levels of hundreds of metabolites, 2) they predict in vivo alterations, which are more physiologically relevant than the results stemming from 2-dimensional cell culture assays, and 3) they leverage data on endogenous metabolites to make predictions for drugs, natural products, and other classes of compounds.

We collected a list of 400 metabolites with valid structures that were measured in both the plasma and the urine of the probenecid-treated human participants. We first generated a binary target variable, with a compound being considered 'Altered' (OAT-mediated) if the compound had a p-value below 0.1 in both plasma and urine, a positive fold change in the plasma over 1.1, and a negative fold change in the urine below 0.9, with p-values and fold changes determined by comparing post and pre-treatment metabolomes in the initial study [137]. These criteria led to 64 metabolites that were considered 'Altered', with 'Altered' suggesting OAT mediation. For the negative cases, which we labeled as 'Unaltered', we focused on compounds that were not significantly altered in the expected way in either the plasma or urine. After applying this criteria, 64 metabolites were in the 'Altered' class and 122 in the 'Unaltered' class (**Figure 5.1B**). With target variables assigned, we then calculated molecular descriptors for each compound. After scaling, we trimmed the molecular descriptors to a list of 22 interpretable features, including molecular weight, total polar surface area, and max partial charge, using variance and correlation thresholds. Using iterative feature selection, we further filtered these 22 to 10 molecular descriptors.

5.3.2. XGBoost and Random Forest models performs best in predicting OAT-mediated
endogenous metabolites

With the trimmed features and the target variables, we then aimed to build binary

classification machine learning models. The data were split into training and test sets, with the

'Altered' cases equally represented in both sets. We also accounted for the imbalanced dataset

using oversampling techniques on the training data. We then applied multiple algorithms (e.g.,

Random Forest, XGBoost) to build binary classification models, and their hyperparameters were

tuned. Classification reports were then generated with the test data, and each model was ranked

using a combination of metrics, including accuracy, f1, and roc_auc. Overall, the XGBoost and

Random Forest models performed the best (**Figure 5.1C**). These XGBoost and Random Forest

models predicted the 'Altered' class correctly 63% and 59% of the time, respectively. They

performed better in predicting the 'Unaltered' class with an accuracy of 82% for XGBoost and

83% for the Random Forest. The classification report for the Random Forest model is shown as a

representative example (**Figure 5.1D**)**.** We also generated ROC curves to compare all the

models, with XGBoost performing the best with an area under the curve of 0.806 (**Figure 5.1E**).

5.3.3. Endogenous metabolite data are insufficient in predicting drug-related function

We applied our trained XGBoost model to a novel data consisting of FDA-approved

drugs. We compared our prediction on these drugs to a list of known clinically relevant drug-

drug interactions involving probenecid to compare the results from our models. We found 155 of

the drugs implicated in probenecid-sensitive drug-drug interactions were present in our list of

FDA-approved drugs. Of the 155 drugs, which should be considered 'Altered' by our XGBoost

model, only 67 were predicted to be 'Altered', while 88 were predicted to be 'Unaltered'. Closer

examination of these predictions shows that some OAT-interacting compounds were considered

'Altered', such as antibiotics (e.g., cefazolin), NSAIDs (e.g., indomethacin), and other drugs

(e.g., oseltamivir). However, the drugs predicted to be 'Unaltered' included many classic OAT-

interacting compounds, such as furosemide, acyclovir, and ibuprofen (**Supplementary Table S5.1**). These results suggest that while some drug-drug interactions likely occurring at the site of OATs can be predicted using data from in vivo metabolomics data, endogenous metabolite data alone are not sufficient in predicting drug-related function. As such, we decided to further analyze the potential structural differences between drugs and metabolites for a specific transporter, OAT1. We focused on the wide array of compounds that have been associated with OAT1 in vitro and in vivo experimentally.

5.3.4. Compounds interacting with OAT1 are structurally distinct

OAT1 (SLC22A6, initially NKT) was discovered in 1997 and identified as a transporter [98]. Since then, several model systems have been used to study its function in vitro and in vivo. To maximize the data available, we curated results from literature reporting OAT1 interaction regardless of the model system used. This resulted in a list of 485 unique compounds, consisting of drugs and metabolites, but also natural products, which have recently gained research attention (**Supplementary Table S5.2**). We first aimed to see how these compounds fit in with the library of known small molecules from the Human Metabolome Database (HMDB), which consisted of over 200,000 unique small molecules. We calculated molecular properties and performed PCA on the initial 22 molecular descriptors to better visualize the data, which revealed that OAT1 covers a reasonable area of the total known small molecule chemical space (**Figure 5.2A**). We then aimed to see if these compounds known to interact with OAT1 were structurally similar to one another by calculating Morgan fingerprints for each compound, which enabled comparison**.** Tanimoto similarity coefficients between each of the compounds revealed that very few pairs of molecules were similar (Tanimoto similarity > 0.70), indicating diversity among the compounds (**Figure 5.2B**). This led us to further explore the set of OAT1-interacting molecules from a structural basis.

5.3.5. Supervised and unsupervised learning methods reveal diverse sets of OAT1-interacting compounds

Given that our initial interest in OAT1 function was on distinguishing between the types of small molecules OAT1 handles, we initially labelled each of the OAT1-interacting compounds as either a drug, metabolite, or natural product. Because our goal was to identify the features that separated these compounds, we performed linear discriminant analysis with a subset of 22 uncorrelated molecular descriptors. The two linear discriminant axes combined to explain nearly all the variance and demonstrated a separation between the categorical variables (metabolite, drug, natural product) that were assigned to each compound (**Figure 5.3A**). Linear discriminant analysis maximizes the variance between the data points based on their categorical variable, using a weighted combination of the features. We analyzed the molecular descriptors with the highest weights in the axes and found that some, such as molecular weight (MolWt), octanol-water partition coefficient (MolLogP), TPSA/SArea (Total Polar Surface Area divided by Labute Accessible Surface Area), and fraction of carbon atoms that are SP3 hybridized contributed to the separation of these compounds (**Figure 5.3B**).

Though the separation with our assigned groups was effective, we also wanted to explore how many clusters form within the dataset with unsupervised learning techniques. We applied multiple clustering algorithms (Affinity Propagation, OPTICs, hierarchical clustering, etc.) to the dataset and found either too few or too many clusters. As such, we used K-Means clustering, which required an a priori number of clusters to be determined. We applied the silhouette score and Davies-Bouldin index evaluation metrics to determine the optimal number of clusters and found that 5 or 6 were appropriate selections based on each metric, respectively (**Figure 5.4A, B, Supplementary Table S5.3**). To better visualize the data, the clusters were mapped to a dimension reduction performed with PCA, which shows some separation between the clusters

(**Figure 5.4C**). We were particularly interested in clusters that had similar counts of drugs, endogenous metabolites, and natural products. For example, in K-Means = 6, Cluster 2 contained 44 drugs, 47 metabolites, and 29 natural products (**Figure 5.4D**). These particular clusters suggest that competitive interactions between different classes of compounds could be occurring, indicating that drug-metabolite and drug-natural product interactions at OAT1 should be considered. We then aimed to understand how these different clusters may interact with the OAT1 protein structure.

5.3.6. Predicted OAT1 structure has two distinct binding sites

Given the structurally diverse compounds that have been shown to interact with OAT1 in vivo and in vitro, we investigated how these different clusters may interact with OAT1 at the molecular level, as it is possible that there are distinct binding mechanisms. OAT1, along with most other membrane transporters, has no experimentally derived protein structure. However, recent advances in protein structure prediction have generated high confidence templates for these proteins. We utilized the AlphaFold2 protein structure for OAT1 to analyze potential interactions between the protein and its hundreds of interacting compounds. We first analyzed the protein structure to assess whether the prediction agreed with previous knowledge on OAT1. OAT1, and SLC22 transporters in general, are characterized by 12 transmembrane domains, a large extracellular loop and two smaller intracellular loops. For the canonical OAT1 structure, residues 31-135 comprise the extracellular loop, while residues 270-337 and residues 506-563 make up two distinct intracellular loops, respectively. The predicted AlphaFold2 structure aligns with the location of these loops and indicates an inward-facing conformation.

To add physiological relevance, we embedded the protein structure in a mammalian membrane and prepared the combined structure using MOE's QuickPrep feature (**Figure 5.5A**). We then focused on identifying likely binding sites for the structure. MOE's SiteFinder function

uses a calculated metric, protein-ligand binding index (PLB), to determine the most likely sites for protein-ligand interactions. This analysis yielded two binding sites, with the highest-ranking site situated on the intracellular side of the protein (**Figure 5.5B**), while the second highest ranking was located on the extracellular side of the membrane (**Figure 5.5C**). Site 1 was much larger and consisted of 82 residues, while Site 2 had only 30 residues. Given that this structure is unvalidated experimentally, we then reviewed existing mutagenesis studies to see if any of the implicated residues were present in either site (**Table 5.1**) [91]. Site 1 contained nine residues that had been associated with reduced transport, while Site 2 contained only one. Most of these studies have been conducted with PAH as the probe substrate, so we investigated whether other compounds may interact with similar residues.

5.3.7. Ligand docking reveals potential different binding mechanisms for Site 1

We prepared a library of ligands that have been shown to interact with OAT1 in any in vivo or in vitro experiment for docking experiments. In these two separate experiments, each ligand was screened against both identified ligand binding sites and the top 3 poses for each compound were retained for further analysis. After docking, we used the protein-ligand interaction fingerprints (PLIFs) calculated by MOE. Not every compound input for docking resulted in a valid protein-ligand docking interaction, which led to 1220 PLIFs for Site 1 and 1132 PLIFs for Site 2. These PLIFs show which residues are involved in the virtual docking experiments. Our results for Site 1 show 55 fingerprints across 23 residue positions. The most common residue is Lysine at position 321 with any interaction occurring. Arginine 273 and Glutamine 455 were also involved in numerous PLIFs (**Figure 5.6A**). Interestingly, neither of these residues has been previously implicated in any mutagenesis studies. Among those that had been shown to be relevant to in vitro transport were Tyrosine 354, Arginine 454, and Arginine

466. Site 2 has 42 distinct fingerprints across 17 residues, with Arginine 423 involved in 37.3% of the total interactions (**Figure 5.6B**). None of the residues implicated in the PLIFs had been previously associated with altered transport activity.

We next wanted to understand our PLIF residues in each site in the context of structural diversity. For a more interpretable result, we decided to focus on molecules that have previously been used as probes for OAT1 function that were grouped into distinct clusters in the unsupervised learning processes, using the results from $K = 6$ clusters. We chose to emphasize the interactions with para-aminohippurate (PAH) (Cluster 3), 6-carboxyfluorescein (Cluster 1), tenofovir (Cluster 6), and probenecid (Cluster 2), as each of these molecules has either been used in the context of OAT1 transport or inhibition.

Because mutagenesis studies had implicated some residues in Site 1, we chose to focus on the protein-ligand interactions occurring within that putative binding site. Interestingly, the preferred poses of each ligand within the pocket interacted with different residues. Para-aminohippurate (PAH) interacted with Valine 211 and Glutamine 455, 6-carboxyfluourescein interacted with Serine 277 and Lysine 321, probenecid interacted with Aspartate 157, and tenofovir interacted with Arginine 161 and Serine 278 (**Figure 5.7A, B, C, D**). We also noted that other classes of compounds, such as the natural product quercetin-3-glucuronide (Cluster 4), which interacted with Threonine 208, Glutamate 326, Glycine 322, and Threonine 451 (**Figure 5.7E**). We also analyzed the gut-derived uremic toxin indoxyl sulfate (Cluster 6), which interacted with Methionine 452 and Glutamine 455. Though this specific combination is unique among the surveyed compounds, Glutamine 455 interacted with both PAH and indoxyl sulfate, despite their being in different clusters (**Figure 5.7F**). None of the residues implicated in these fingerprints were among those that had been shown to have an in vitro impact on the transport of

139

probe substrates, indicating that more residues may be relevant in transport or inhibition processes. While our studies are limited, they suggest that for structurally diverse compounds, there may be multiple binding mechanisms in the OAT1 protein that are dependent upon specific residues. The total results, including all compounds in each docking pose, are shown in **Supplementary Table S5.4** for Site 1 and **Supplementary Table S5.5** for Site 2.

5.4 DISCUSSION

In this work, we explored OAT1 function using in silico approaches, including ligand-based and protein-based strategies. We first developed binary classification machine learning models based on in vivo metabolomics data from the plasma and urine of probenecid-treated humans as a proxy to OAT-mediated handling of small molecules. We calculated molecular descriptors for each molecule measured to use as features and used a binary target variable defined by p-values and fold changes in both plasma and urine to build predictive models. These models were entirely trained on limited endogenous metabolite data, and while they were able to predict some probenecid-sensitive drug-drug interactions, they were ineffective in predicting OAT-mediated drugs. We then analyzed the chemical space associated with OAT1 by compiling all experimental data and found that OAT1 interacts with multiple structurally distinct clusters of small molecule compounds. This diversity suggested different binding mechanisms for certain ligands, which we explored by docking each compound against the predicted structure for human OAT1. These analyses revealed residues in a putative intracellular binding pocket that may be important in mediating specific protein-ligand interactions.

Though most drug transporter machine learning models use in vitro data [69], previous models generated by our group have used in vivo metabolomics to distinguish OAT1 and OAT3 function [74]. Our work here has a different goal and focuses on compounds that are either OAT-

mediated or not OAT-mediated. Another advantage is that the models developed were trained and tested using in vivo metabolomics data from the same experiment. However, due to limitations in metabolomics data available, these models only contained 192 total metabolites, with 64 classified as OAT-mediated and 128 metabolites serving as negative cases. This led to models that performed relatively well on the test set consisting of endogenous metabolites. However, when the trained model was exposed to a novel dataset consisting of FDA-approved drugs, the model incorrectly predicted that numerous drugs would not be altered, when in vivo probenecid drug-drug interactions provided evidence that the drug would be altered. It should be noted, however, that probenecid-sensitive drug-drug interactions may not be a direct result of OAT1/3 inhibition, as probenecid may have roles as clinical inhibitors of other renal and extrarenal transporters and even enzymes [22, 179]. Acquisition of a validation set would require multiple, expensive experiments, so we relied upon existing data, despite the limitations.

While these models were relatively unsuccessful, the framework can be applied when higher quality data is obtained. As metabolomics experiments become cheaper and more accessible, there will be an increasing need to analyze the data from a structural perspective. The ligand-based approach we employ here is flexible and can be useful in interpreting several metabolomics experiments with different goals, as the binary target variable can be modified depending on the specifics of the experiment. For example, depending on treatment, genetic knockout, or another physiologically relevant factor, a different biological sample or combination of biological samples may be more appropriate. The expected change (e.g., elevated in plasma, decreased in urine) in these types of experiments could be refined to identify either a binary target variable or even multiple classes for a target variable.

141

Though we noted the limitations of our machine learning models, the indication that endogenous metabolite data cannot be used to predict drug-related function merited further exploration. By collecting all the compounds associated with OAT1 by in vivo or in vitro experiments, we were able to probe the full extent of the function of OAT1. The list of OAT1-interacting compounds consisted of drugs, endogenous metabolites, and natural products. To simplify the problem, we included small molecules from a variety of different experiments and did not distinguish between substrates and inhibitors. We first separated these using linear discriminant analysis with their assigned categorical variables, which demonstrated that a combination of interpretable molecular descriptors was able to separate these distinct OAT1-interacting compounds with different origins and mechanisms of actions. While this may be the case for OAT1, future studies can explore whether other drug transporters have similar characteristics.

While this analysis showed a general separation between classes, there did appear to be some compounds on the edges of their pre-defined categorical variables (drugs, metabolites, natural products) suggesting that there may be organic clusters that are based on molecular descriptors rather than categorical assignments. We performed K-Means clustering with the calculated optimal number of clusters and found that 5 or 6 distinct clusters were present in the dataset. These results yielded clusters that mainly consisted of different types of molecules, suggesting that some competitive interactions between similar compounds may be more likely than others. Drug-drug interactions have dominated research about competitive interactions at transporters, but our work suggests drug-metabolite and drug-natural product interactions should also be considered [21, 137, 246]. Future studies can emphasize the differences between

142

substrates and inhibitors, as it is possible that certain molecular descriptors may be characteristic of one or the other.

To explore the possibility of structurally distinct molecules having different binding modes, we used the recently predicted AlphaFold2 prediction for OAT1 [87]. AlphaFold2 has been shown to be effective in predicting the structure of transmembrane proteins and cautious application of the predicted structures could yield valuable insights [247]. The protein structure prediction, while limited in some capacity, did recapitulate many of the features that OAT1 and SLC22 transporters possess, allowing for protein-ligand interaction analysis that had not previously been possible. Recent results have made use of predicted OAT1 structures in the context of the empty structure, or the structure bound with a specific ligand [89, 91]. Here, we focus on the potential interactions between a putative binding site and multiple different experimentally confirmed interacting compounds. Our mixed results show that it is possible that different residues contribute to the binding and potentially the transport of structurally diverse compounds. While these studies can be improved with more conformations of the transporter (outward occluding, inward open, etc.) and long molecular dynamic simulations, they provide a basis for understanding the broad range of interacting molecules with OAT1 from the perspective of the protein. These results can be supported by in vitro mutagenesis studies altering the implicated residues.

Overall, the studies presented in this work elucidate the molecular descriptors and potential OAT1 amino acid residues that are important in protein-ligand interactions. OAT1 has a wide array of interacting molecules, including drugs, endogenous metabolites, and natural products, with distinct chemical features. When analyzed from the perspective of the protein, though the evidence is not absolute, it appears that certain residues may be more important than

others for specific small molecule compounds with a subset of molecular descriptors. With the increasing accessibility of metabolomics data and experimentally derived or predicted protein structure, a similar approach can also be applied to other membrane-bound transporters. This work could help detail the specific nature of important drug-drug, drug-metabolite, drug-natural product, or metabolite-natural product interactions at the site of multi-specific drug transporters.

5.5 MATERIALS AND METHODS

5.5.1. Metabolomics data collection and preparation

In vivo plasma and urine metabolomics data from humans treated with probenecid [137] were collected with the associated fold changes, p-values and SMILES sequences. Only compounds measured in both the plasma and the urine were included to isolate kidney-specific probenecid-sensitive transporter function, which included 400 metabolites. Binary classifications for altered and unaltered compounds were added using p-value and fold change cutoffs in the plasma and the urine. Positive cases were both elevated in the plasma and decreased in the urine, while negative cases were unaltered in both the plasma and the urine. SMILES sequences were input into the MolFromSMILES package in RDKit to generalte .mol files, which were then used as inputs into the RDKit Molecular Descriptors package to generate 208 molecular descriptors for each molecule, with some dropped due to poor interpretability. Most data science techniques were applied using a combination of the sklearn, numpy, and pandas packages in Python. Visualizations were performed with the seaborn package. Data was scaled using sklearn's StandardScaler function. Following pre-processing with variation (0.90 or higher) and correlation thresholds (0.75 correlation), 22 interpretable molecular features ('qed', 'MolWt', 'MaxPartialCharge', 'MinPartialCharge', 'FpDensityMorgan1', 'fr_Ar_NH', 'fr_C_O', 'fr_NH1', 'fr_NH2', 'TPSA', 'FractionCSP3', 'NumAliphaticHeterocycles', 'NumAliphaticRings',' NumAromaticCarbocycles', 'NumAromaticHeterocycles', 'NumRotatableBonds', 'MolLogP',

'fr_Al_COO', 'fr_Al_OH', 'fr_C_O_noCOO', 'fr_ether', 'TPSA/SArea') were retained. Molecules with any missing values were dropped. Following iterative feature selection, we retained 10 molecular descriptors ('MolLogP', 'MolWt', 'TPSA/SArea', 'qed', 'MaxPartialCharge', 'FpDensityMorgan1', 'TPSA', 'NumAromaticCarbocycles', 'FractionCSP3', 'MinPartialCharge'), including the engineered TPSA/SArea.

5.5.2. Machine learning models

The observations and the remaining features were then split into training and test sets with a 70/30 split, with the binary target variable being stratified to ensure equal representation in the training and test sets. Because of the imbalanced dataset, synthetic minority oversampling technique (SMOTE) in the imblearn package was used to generate synthetic data points for the minority class. Other balancing methods were also explored. Multiple algorithms (Logisitic Regression, K-Nearest Neighbors, Support Vector Machines, Decision Tree, Random Forest, XGBoost) were assessed following on the training data. Hyperparameters for each model were determined using GridSearchCV with cross validation over a customized set of parameters. The StratifiedShuffleSplit function in sklearn was used in the hyperparameter optimization for cross validation. Multiple different scoring functions (f1 score, roc_auc, precision, recall, etc.) for the hyperparameter optimization were used, but f1_weighted was ultimately used. Evaluation of the models was assessed with multiple metrics (accuracy, balanced_accuracy, roc_auc, f1, etc.) with a combination of the factors being used to determine the best performing model on the training data. The resulting models were then used on a novel dataset consisting of drugs with existing clinical literature being used to validate the probenecid-sensitive interactions.

To test the machine learning models on novel data, we downloaded a list of FDA-approved drugs (https://chemoinfo.ipmc.cnrs.fr/MOLDB/index.php). The drugs in this dataset

were preprocessed identically to the endogenous metabolites as previously described, leading to 10 molecular descriptors for each drug. We collected a list of known clinically relevant drug-drug interactions involving probenecid to compare the results from our models (https://www.drugs.com/drug-interactions/probenecid-index.html) to serve as the validation set.

5.5.3. OAT1-interacting compound literature search and dimensionality reduction

We curated a list of OAT1-interacting compounds from the literature including all human, mouse, and rat in vivo and in vitro OAT1 experiments (**Supplementary Table S5.3**). All data reported were included with no specific threshold for interaction. Some compounds from in vivo knockout mouse experiments were removed due to their outlier status with respect to chemical structure. Compounds were labeled as drugs, endogenous metabolites, and natural products based on expert opinion. 485 compounds were collected in total. Molecular fingerprints (Morgan) were calculated for 483 compounds to determine Tanimoto similarity coefficients between all compounds. Molecular descriptors were calculated for 478 compounds as described for the machine learning, with 22 uncorrelated features being retained.

We then applied different dimensionality reduction techniques for improved visualization and analysis. We used sklearn's principal component analysis (PCA), linear discriminant analysis (LDA), and tSNE to separate the compounds based on their assigned class (metabolite, drug, natural product). We then analyzed the axes for the linear dimensionality reduction methods to determine which molecular features contribute to the individual axes.

We also calculated molecular descriptors for all compounds in HMDB (https://hmdb.ca/downloads), including all OAT1-interacting molecules (accessed in January 2023). PCA was calculated and two components were retained to show the chemical space covered by OAT1 vs the known chemical space.

146

5.5.4. Unsupervised Learning

For clustering purposes, to reduce the dimensions, we used the results from the PCA in visualizations. Sklearn clustering algorithms were applied, including but not limited to K-Means, hierarchical, affinity propagation, and others. To determine the optimal number of clusters for K-Means, we used clustering evaluation metrics in sklearn, including the silhouette score and Davies-Bouldin score.

5.5.5. Ligand Preparation

PubChem IDs were used to download molecular structures for each compound. They were then uploaded into a structural database within Molecular Operating Environment 2022 (MOE). Chemical fingerprints, pharmacophores, and atomic force energy fields were calculated for each compound.

5.5.6. Protein Structure Preparation

Human OAT1 currently does not have a published crystal structure that can be used for molecular dynamics simulations and docking, so we used the predicted protein structure. The human protein structure for SLC22A6 was downloaded directly from the AlphaFold2 database (https://alphafold.ebi.ac.uk/entry/Q4U2R8) on December 3, 2022. The structure displayed high confidence, with 485 of the 563 residues having a per-residue confidence score (pLDDT) of greater than 70, which we considered suitable for further simulations and analysis.

The resulting structure was prepared with CHARMM-GUI for simulation in a lipid bilayer membrane [248]. The membrane was composed of POPC with a dimension of 90 x 90, and a water bounding box of 155 Angstroms. Potassium and chlorine ions were placed in the solvent using Monte Carlo placement at 150 mM. We used AMBERFF19, Lipid17, and the OPC water model to parameterize the system before simulating with OpenMM. We followed the default CHARMM-GUI minimization and equilibration steps, which consisted of 5000 steps of

minimization, followed by six stages of equilibration with decreasing levels of constraints on system. We ran the production MD simulation for 50 ns, without any constraints at a 2 fs simulation timestep. We expect that the final frame of the simulation will be sufficiently equilibrated and more physically accurate compared to the structure produced directly from AlphaFold. Furthermore, because OAT1 is membrane bound, it is essential to model the phospholipid membrane prior to running any docking simulations, as this will influence the regions of the protein available to ligand binding. The combined structure was then corrected using MOE's QuickPrep feature. MOE's SiteFinder tool was then used to identify high likelihood binding sites on the protein structure, which revealed two putative binding sites on the intracellular and extracellular side of the protein. These were then used as the sites for two separate docking studies.

5.5.7. Ligand Docking Analysis

Ligand docking was performed in MOE. Default docking parameters were used. The relationship between the ligands and the putative binding sites were analyzed using the protein-ligand interaction fingerprints (PLIFs) as calculated by MOE. The ligand interaction diagrams were used to visualize the PLIFs.

5.6 SUPPLEMENTARY MATERIAL

Supplementary Table S5.1: Predicted classes of drugs involved in probenecid-sensitive drug-drug interactions.

Supplementary Table S5.2: All OAT1-interacting molecules from in vitro and in vivo experiments.

Supplementary Table S5.3: Clustering assignments for K = 5 and K = 6 in KMeans clustering.

Supplementary Table S5.4: Protein-ligand interaction fingerprints occurring at Site 1 with molecules identified by Pubchem ID. -: bit not set, D: sidechain hydrogen bond donor, A:

sidechain hydrogen bond acceptor, d: backbone hydrogen bond donor, a: backbone hydrogen bond acceptor, O: solvent hydrogen bond, I: ionic attraction, M: metal ligation, R: arene attraction, H: hydrophobic surface contact, Q: charged surface contact, P: partial hydrophobic contact, C: total surface contact.

Supplementary Table S5.5: Protein-ligand interaction fingerprints occurring at Site 2 with molecules identified by Pubchem ID. -: bit not set, D: sidechain hydrogen bond donor, A: sidechain hydrogen bond acceptor, d: backbone hydrogen bond donor, a: backbone hydrogen bond acceptor, O: solvent hydrogen bond, I: ionic attraction, M: metal ligation, R: arene attraction, H: hydrophobic surface contact, Q: charged surface contact, P: partial hydrophobic contact, C: total surface contact.

Table 5.1 OAT1 Mutagenesis sites associated with reduced transporter function.

Substrate	Residue	Present in Site 1 (82 residues)	Present in Site 2 (30 residues)
PAH	30		
PAH	36		
PAH	39		
PAH	56		
PAH	86		
PAH	92		
PAH	97		
PAH	113		
PAH	189		
Adefovir, cidofovir, tenofovir	203	YES	
PAH	230	YES	
PAH	335		
PAH	341		
PAH	353	YES	
PAH	354	YES	
PAH	379		
PAH	394	YES	
PAH	427		
PAH	431		
PAH	434		
Cidofovir	438	YES	
PAH	440		
PAH	454	YES	
PAH	466	YES	
PAH	478		
PAH	490		YES
PAH	503		
PAH	504		
PAH	506		
PAH	512	YES	

Figure 5.1 XGBoost and Random Forest binary classifiers trained on molecular descriptors are able to predict OAT-mediated endogenous metabolites.

A) Plasma and urine metabolomics data were combined, and molecular descriptors were calculated for each compound measured in both experiments. **B)** Each compound was classified as altered or unaltered based on p-value and fold change criteria in both plasma and urine. Overall, 64 were considered altered and 128 were considered unaltered. **C)** Of the surveyed binary classification algorithms, XGBoost and Random Forest performed the best in roc_auc, f1, balanced accuracy, and accuracy on the test data. **D)** Classification report for the Random Forest model. **E)** ROC curves show that XGBoost performs the best with respect to area under the curve (AUC). The dashed line where true positive rate is equal to false positive rate represents an untrained model.
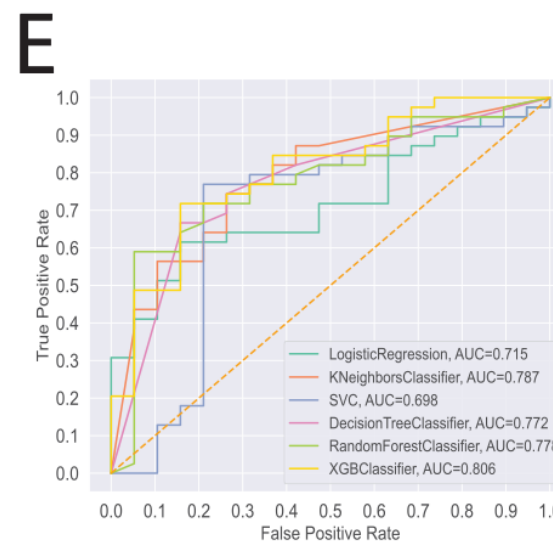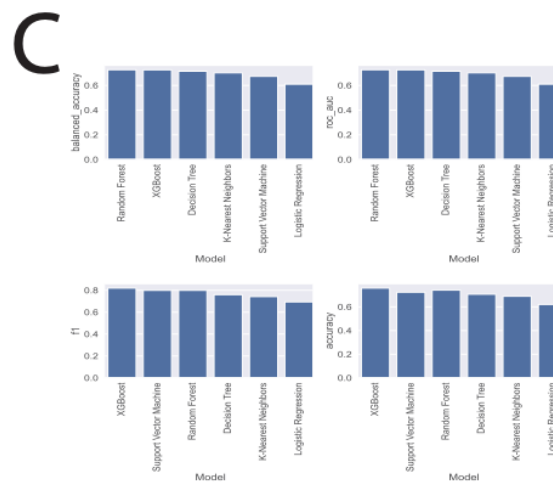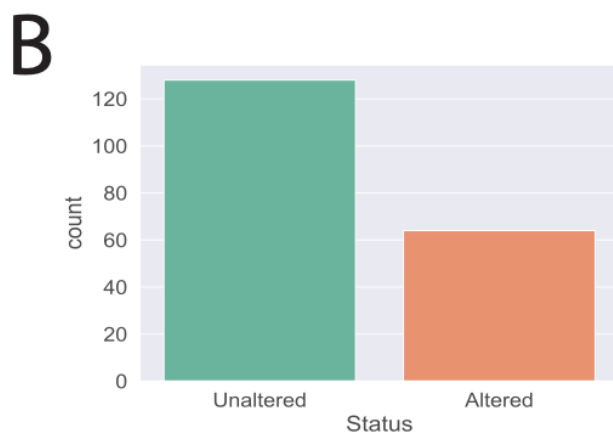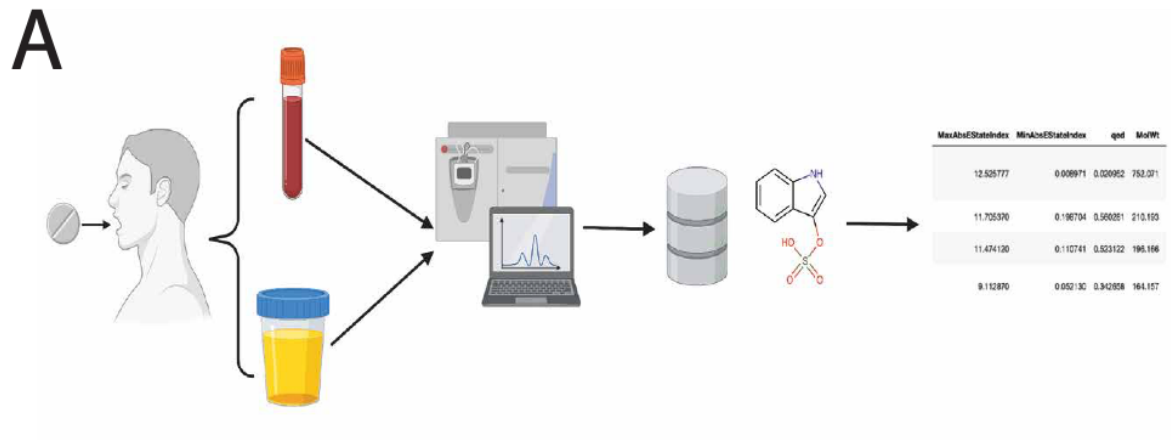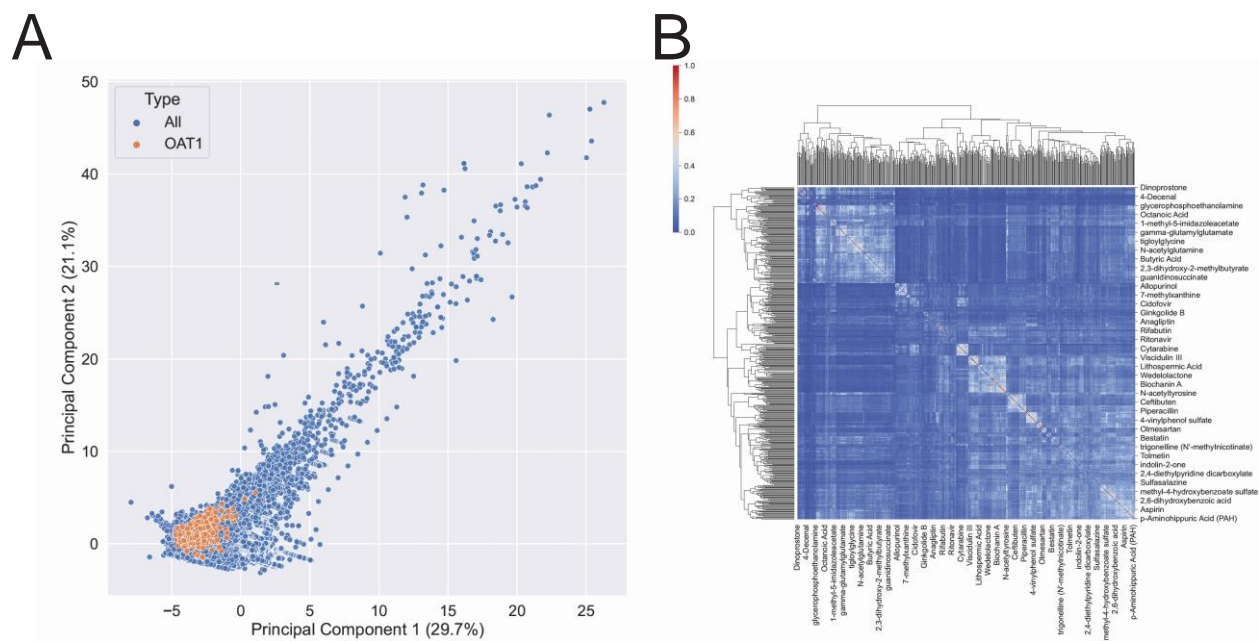
Figure 5.2 OAT1 covers a relatively large range of known chemical space, but the compounds are not similar.

A). Principal component analysis of the molecular descriptors all molecules in the Human Metabolome Database, including the collected OAT1-interacting molecules, reveals that OAT1 covers a wide range of the known chemical space. **B)** Morgan fingerprints were calculated for each chemical structure, which enabled the calculation of Tanimoto similarity. The compounds were all compared to one another, with the majority having low similarity metrics.

Figure 5.3 Linear discriminant analysis can separate OAT1-interacting molecules based on categorical groupings of drugs, metabolites, or natural products.

A) Linear discriminant analysis of the OAT1-interacting compounds reveals separation between the different categorical variables. **B)** Eight features contributed most to linear discriminants 1 and 2, which optimize for the variance between the categorical variables.

A

Linear Discriminant Analysis Based on Class

B

MolWt

MolLogP

TPSA/SArea

FractionCSP3
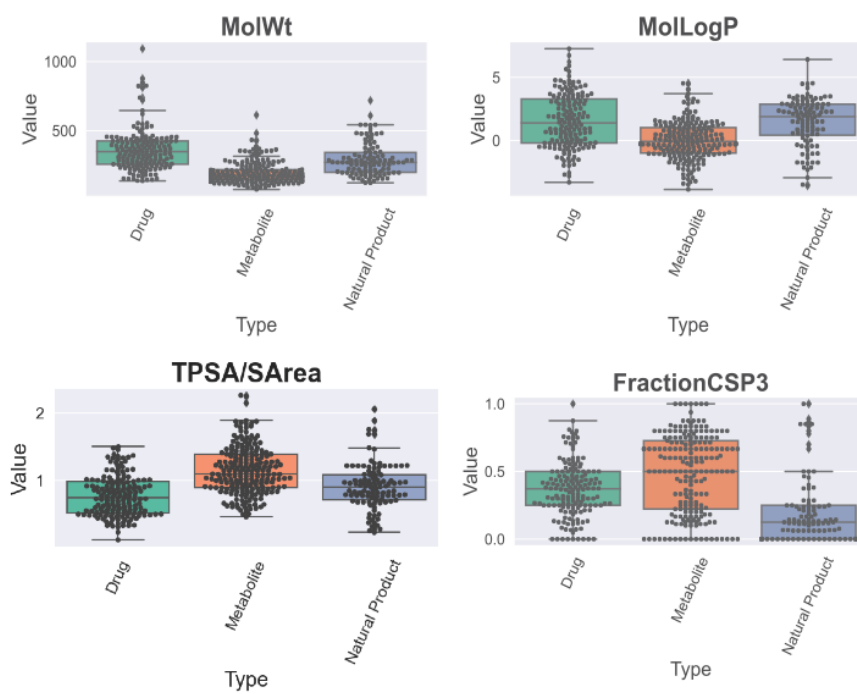
Figure 5.4 OAT1-interacting molecules can be split into distinct clusters based on molecular descriptors.

A) The silhouette score evaluation metric was used to determine the optimal number of clusters for K-Means Clustering and indicated that 5 was the best score using this metric **B)** The Davies-Bouldin index evaluation metric was used to determine the optimal number of clusters for K-Means Clustering and indicated that 6 clusters was the best score using this metric. When taken with the silhouette score, we decided to choose either 5 or 6 clusters for K-Means clustering. **C)** Principal component analysis was performed using the molecular descriptors for the OAT1-interacting compounds to improve visualization. Each point represents an individual small molecule, with the color representing the cluster it belongs to and the shape of the point representing its categorical assignment. **D)** The counts for each categorical assignment are presented for each individual cluster, showing that some clusters contain different classes of small molecules.

A

B

C

OAT1-interacting molecules clustered

D
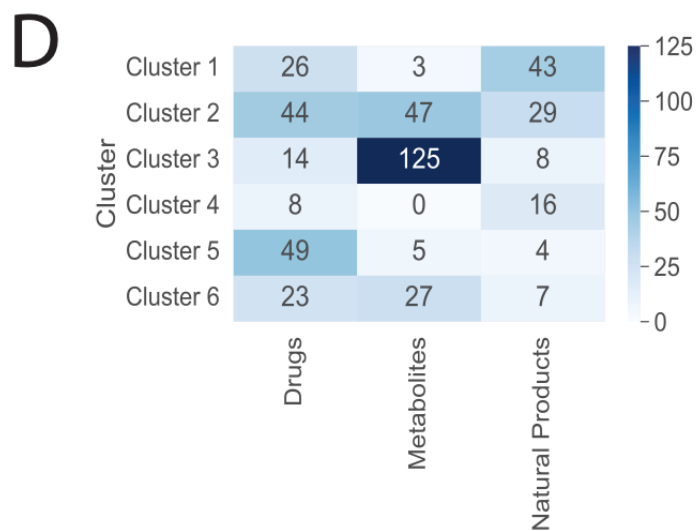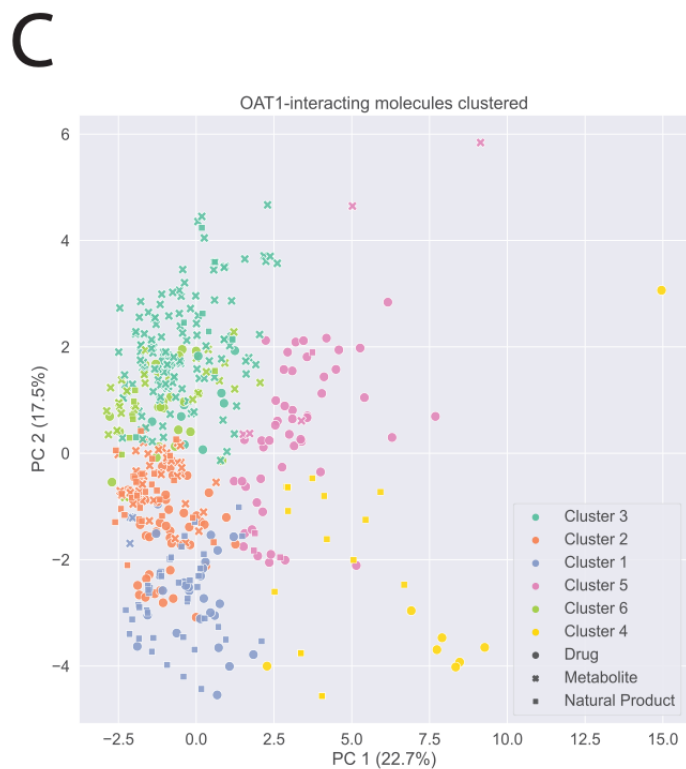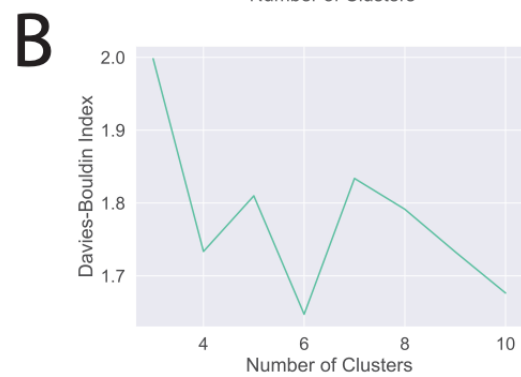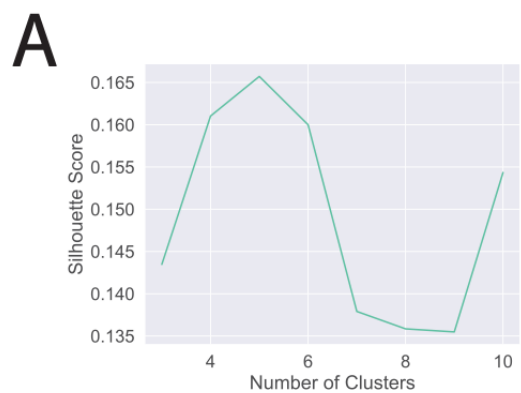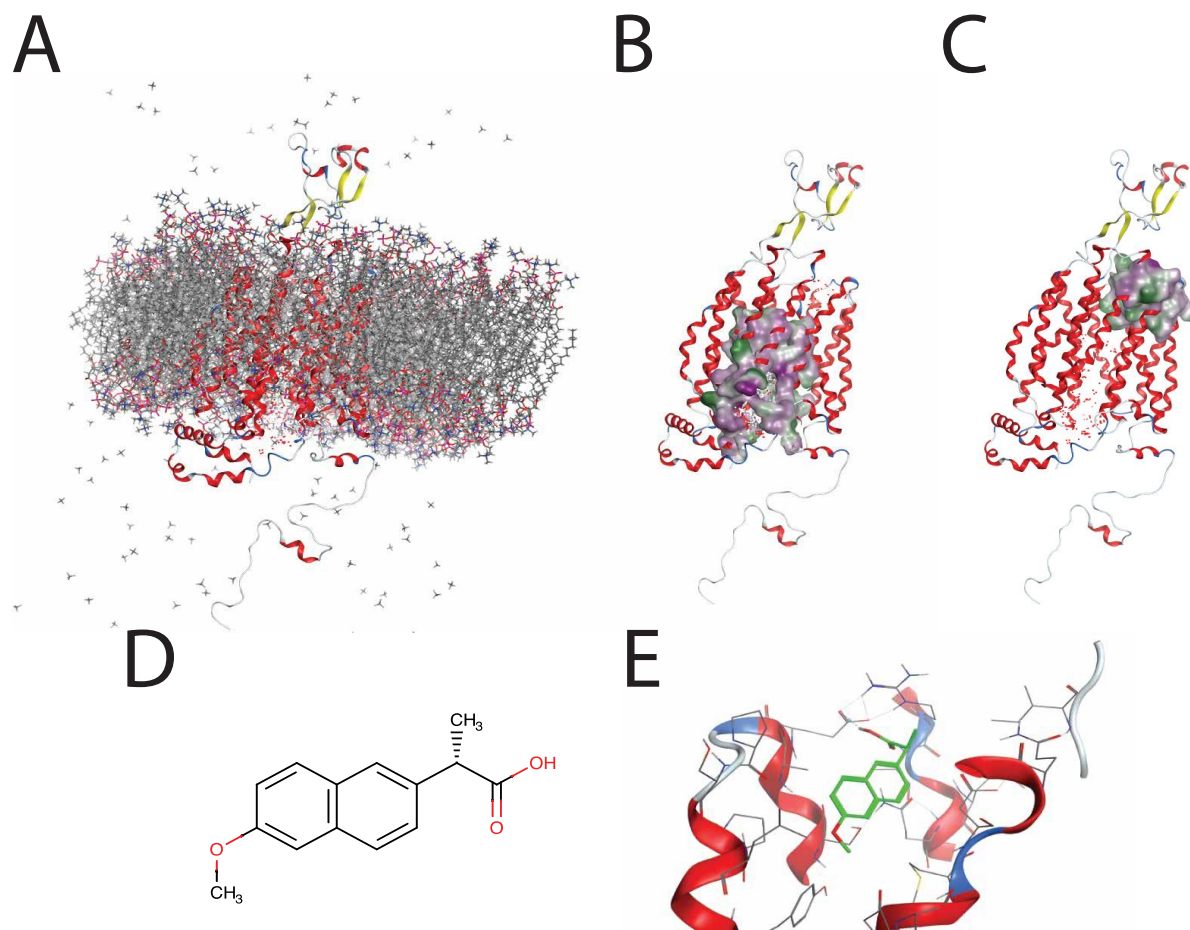
Figure 5.5 Predicted OAT1 structure was embedded in a mammalian plasma membrane and used for docking studies.

A) OAT1 predicted protein structure was embedded in a mammalian cell membrane. **B)** The most likely binding sites on the protein were identified. Site 1 is present on the intracellular (bottom) side of the protein. **C)** Site 2 is located on the extracellular (top) side of the protein. The binding sites are colored differently from the protein, where hydrophilic regions are purple, neutral regions are gray, and lipophilic regions are green. **D)** Naproxen, a non-steroidal anti-inflammatory drug, was one of the compounds that has been shown to interact with OAT1 experimentally. **E)** All compounds were docked against each binding site. Naproxen is shown as a representative example here.
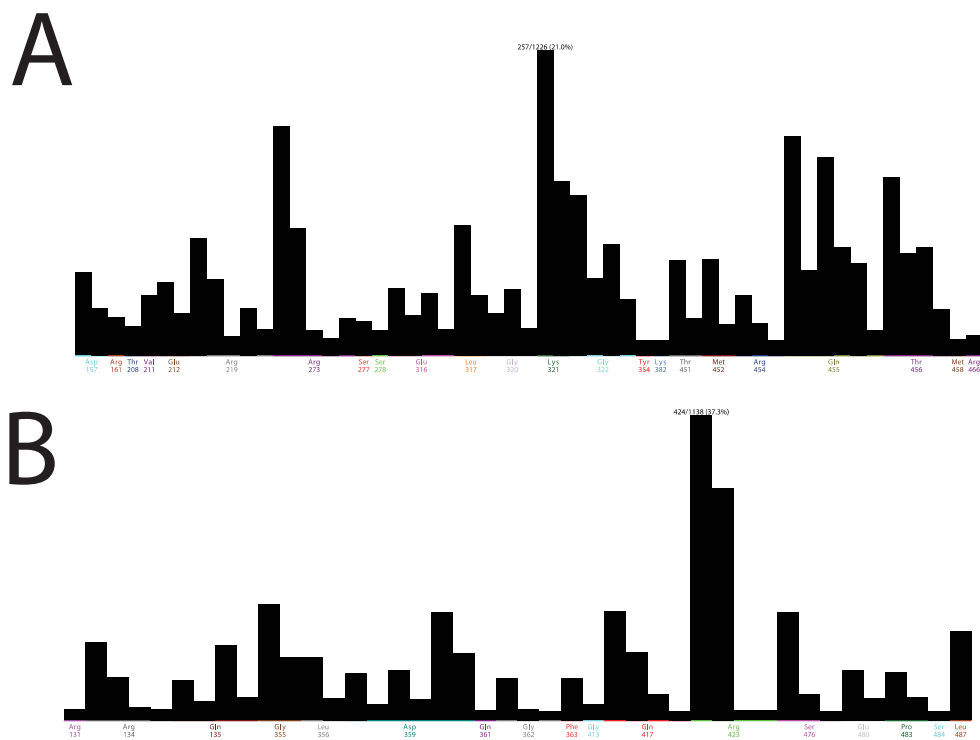
Figure 5.6 Protein-ligand interaction fingerprints describe the residues implicated in binding.

A) Site 1, consisting of 80 residues, is located on the intracellular side of the transporter and has 55 unique fingerprints covering 23 residues. **B)** Site 2, consisting of 30 residues, is located on the extracellular side of the transporter and has 42 unique fingerprints covering 17 residues.

Figure 5.7 Prototypical OAT1 probe substrates and inhibitors seem to interact with different residues in Site 1.

A) PAH (para-aminohippurate) is a probe traditionally used in tandem with radiolabeling to measure OAT1 transport function. The top-ranked docked pose interacts with Valine 211 and Glutamine 455. **B)** 6-carboxyfluorescein is a fluorescent probe used to measure OAT1 function in vitro. The top-ranked docked pose interacts with Serine 277 and Lysine 321. **C)** Probenecid is a drug used to inhibit OAT1 in vitro and in vivo. The top-ranked docked pose interacts with Aspartate 157. **D)** Tenofovir is an antiviral drug primarily cleared via OAT1. The top-ranked docked pose interacts with Arginine 161 and Serine 278. **E)** Quercetin-3-glucuronide is a natural product shown to inhibit OAT1 in vitro. The top-ranked docked pose interacts with Threonine 208, Glutamate 316, Glycine 322, and Threonine 451. **F)** Indoxyl sulfate is a uremic toxin transported by OAT1 in vivo and in vitro. The top-ranked docked pose interacts with Methionine 452 and Glutamine 455.

Chapter 5 is being prepared for publication. The dissertation author was the primary author of this chapter.

CHAPTER 6: CONCLUSION

6.1 DISCUSSION

This dissertation performs a multi-scale analysis using in vivo, in vitro, and in silico tools of a single drug transporter, OAT1 (SLC22A6), and extends to preliminary analyses for other drug-handling genes, such as transporters and enzymes. We demonstrated OAT1 plays an important role in mediating the levels of various endogenous metabolites, notably multiple gut-derived compounds with signaling roles. While this had been partially probed in the past, the work presented here provides novel and robust evidence for this role in a single experiment supported by in vivo mouse and human studies. Furthermore, we combined available data for OAT1 to develop both ligand and protein-based models to better characterize the full extent of drug transporter function (endogenous metabolites, drugs, natural products, etc.).

Considering the association between OAT1 and some microbiome associated uremic toxins present in chronic kidney disease [249], we first aimed to understand to which extent OAT1 handled gut microbiome derived endogenous metabolites. The gut microbiome has been a major research topic in recent years because of the impact that it can play in healthy and disease states [250]. The tools to explore the bacterial species and communities in the gut have made analyzing the metabolic capabilities of the microbiome more accessible and have improved our understanding of this nuanced inter-organismal communication. While simplifying the microbiome to a single 'organism' is reductive because of the hundreds of unique species present, framing the interactions between the host (human, mouse, etc.) as a relationship between two organisms provides a clearer picture. There may be some physical interactions between the gut microbes and the intestine, but one of the main ways that the gut microbiome can interact with the host is via the production of small molecules with signaling roles in the host.

162

Gut microbiome-derived small molecules are compounds that cannot be produced by the host alone. The genes (enzymes, transporters) expressed by the diverse bacterial species in the gut provide machinery for generating compounds that can have important roles within the host. For example, some gut-microbiome derived metabolites have been shown to have signaling roles in the brain, the heart, and other organs [16]. While this signaling role is crucial, the ADME properties of these compounds have been largely ignored. It is imperative to understand the proteins that are responsible for the generation and clearance of these compounds. The work presented in Chapter 2 shows that OAT1, a renal drug transporter, mediates the circulating levels of 40 gut microbiome-derived metabolites, as evidenced by in vivo serum metabolomics of a gut microbiome-depleted knockout mouse model [251]. These results were also supported by two types of in vitro assays, with one assay demonstrating inhibition of traditional OAT1 transport and the other assay showing competitive binding to the OAT1 protein. OAT1 has a broad array of substrates/inhibitors, but it is likely that other multi-specific transporters, such as OATPs and MRPs, also contribute to the ADME of gut microbiome-derived compounds and inter-organismal communication. Deeper analyses could also focus on the specific bacterial species involved in the generation of transporter-specific substrates.

The vast majority of drug transporter-related studies have used in vitro cell assays or in vivo knockout mouse models. While these have generally been useful, there is still a gap between preclinical and clinical studies. In Chapter 3 of this dissertation, we analyze the short-term effect of an oral dosage of an OAT-inhibiting drug, probenecid, on the plasma and urine metabolomes healthy human participants after 5 hours [137]. Probenecid, a drug used to treat gout via URAT1 inhibition, has been shown to inhibit OAT1 and OAT3 in vivo, making it the closest analog to a chemical OAT1/OAT3 human knockout. Though probenecid has been shown

to inhibit other transporters expressed in diverse tissues, this has primarily been shown in vitro. By focusing on the compounds elevated in the plasma and decreased in the urine, we were able to isolate the impact of renal drug transporter inhibition.

For multiple drug transporters, including OAT1 and OAT3, it has long been known that they have some endogenous substrates, but the extent to which these transporters contribute to physiology had yet to be determined. The clinical probenecid study described in Chapter 3 of this dissertation identifies hundreds of simultaneous drug-metabolite interactions in the plasma in the urine. For uptake transporters expressed on the basolateral (blood-facing) side of cells, inhibition would lead to the accumulation of compounds in the blood, as they can no longer enter the cell and must remain in circulation. This inhibition is likely to have downstream effects, such as decreased levels of substrates in the urine, because uptake from the blood is often the rate limiting step for clearance of compounds.

Of the compounds measured in both the plasma and the urine, 97 were significantly elevated in the plasma and significantly decreased in the urine, suggesting they are substrates of OAT1 or OAT3. To ensure the specific interaction at the site of the transporters of interest, we overlapped our results with the serum metabolomics from our knockout mouse studies, and identified 25 metabolites, including many gut microbiome-derived metabolites, including some that were specific to OAT1 or OAT3, indicating that these transporters have shared and distinct functions. We also identified a hyper-specific drug-metabolite interaction at the site of URAT1, with urate being the only compound that was significantly elevated in the urine and significantly decreased in the plasma. This is consistent with the known function of URAT1, as it is an uptake transporter on the apical membrane with high specificity for urate. This dissertation focuses on multi-specific transporters, like OAT1, with dozens of unique substrates, but it is also important

164

to note that proteins within the same family as multi-specific transporters can have very limited substrate specificity with clinical relevance. Overall, Chapter 3 shows that taking common, non-toxic drugs has a profound impact on the levels of endogenous metabolites that may be substrates of transporters. Future studies could focus on the impact of other common individual drugs or drug combinations, especially those that inhibit other important drug transporters, on endogenous metabolism in both the short-term and the long-term, as these could explain metabolic side effects stemming from drug administration.

Having established that OAT1 played an important role in endogenous metabolite handling with a genetic knockout mouse model and a human inhibition study, we then aimed to understand how single nucleotide polymorphisms (SNPs) in drug transporter and drug metabolizing enzyme genes can influence the circulating levels of small, bioactive polar molecules, such as fatty acids, eicosanoids, and bile acids in Chapter 4 of this dissertation. By combining genomic and untargeted plasma metabolomics data from a large human cohort, we were able to identify statistically significant associations between SNPs and hundreds of metabolites. While one our initial goals was to see if OAT1 was strongly linked to any metabolites, an advantage of this methodology is that several other genes were surveyed. Due to potential mechanistic interactions, we limited our study to ~1000 genes including transporters, enzymes, and other related genes.

The targeted-SNP study led to over 600,000 statistically significant associations. To simplify this, we grouped SNPs within a certain number of base pairs to each other into genomic loci, as it is often difficult to determine whether a single SNP is the causative SNP for the observed phenotype. We found 77 distinct genomic loci associated with 7,326 metabolites, with the vast majority of these metabolites being unidentified but having a unique combination of

mass to charge ratio (M/Z) and retention time (RT). One of our first goals was to understand which genomic loci combined to regulate specific metabolites. We found five metabolites (all unidentified) that were associated with four distinct loci, suggesting that the circulating levels of these compounds are dependent upon the endogenous function of multiple drug-related genes expressed in key metabolic organs, such as the liver, kidney and intestine. In addition to the few examples of four distinct loci associated with a single metabolite, we also identified 79 metabolites associated with three loci, and 606 associated with two distinct loci. Taken together, these results support the belief that drug-handling genes, such as drug transporters and drug-metabolizing enzymes, combine to perform crucial physiological functions. This also suggests that shared substrate specificity may be a key function of the network of ADME genes.

We then turned our attention to the ~100 metabolites that were identified and found dozens of associations that had been previously established, such as *SLCO* transporter genes with bilirubin and *SLC22* transporter genes with conjugated sex steroids. That many of these associations had been confirmed by independent in vivo and in vitro experiments demonstrates the strength of the approach and implies that other unexplored associations may prove to be true. While not all relationships between SNPs and metabolites will lead to physiologically relevant interactions, these data generate strong hypotheses for potential roles of drug-handling genes. We also remarked that some SNPs identified in our study, such as rs4149056 in the *SLCO1B1* gene, had previously been associated with several drugs and can impact the toxicity, efficacy, or clearance. This is consistent with the theme of this dissertation, as it provides evidence that a drug transporter can have a wide array of interacting molecules including endogenous metabolites and drugs. This also suggests that certain populations may be susceptible to drug-metabolite interactions. Our approach of using genomic data also opens the door for use in

personalized medicine. By understanding the consequences of specific SNPs on circulating endogenous metabolites, we can better understand the 'baseline' physiology of individuals and apply this in various treatments.

The bulk of this dissertation uses in vivo metabolomics data to gain insight on the physiological role of OAT1 and some other drug-handling genes. While these data have increased our knowledge on the endogenous function of these proteins, these experiments are relatively expensive, as well as difficult to design and execute. Being able to generate high confidence predictions would aid in better understanding not only the endogenous function of drug transporters, but the complete role, including interactions with drugs, toxins, and natural products. In Chapter 5 of this dissertation, we developed ligand-based and protein-based models to better understand transporter-ligand interactions.

Our first approach was to leverage the data acquired from clinically relevant in vivo metabolomics data to build machine learning models. Typically, machine learning models require high throughput in vitro assays for sufficient data. However, in our case, since multiple metabolites are measured in metabolomics experiments, we were able to develop binary classification models. Though our number of observations was relatively low compared to traditional machine learning models, this work is clinically relevant and generated in the same experiment. The models also provide a framework upon which future models focused on different problems can be improved.

We chose to focus on the combined human plasma and urine data from Chapter 3 of this dissertation. Through the use of both samples, we assigned a binary classification to each endogenous metabolite of 'Altered' (if it was significantly elevated in the plasma and significantly decreased in the urine) or 'Unaltered' (if it was not significantly altered in both

plasma and urine). These classifications allowed us to focus on the compounds that are likely

OAT-mediated, as probenecid has been shown to inhibit both OAT1 and OAT3 in vivo. For the

features in these models, we calculated molecular descriptors for each chemical structure and

trimmed down to a subset of interpretable features, such as molecular weight. We then trained

multiple binary classification algorithms on the training data and tuned the hyperparameters for

optimal performance. Upon achieving reasonable accuracy on the test data consisting of

endogenous metabolites, we applied our XGBoost model to novel data that only included drugs.

The results could be, in part, validated by accessible clinical drug-drug interaction data involving

probenecid. Ultimately, though our models predicted some OAT-mediated, probenecid-sensitive

drug-drug interactions, the models were unable to predict OAT-mediated drugs from endogenous

metabolite data alone, suggesting diverse mechanisms for different kinds of compounds.

To simplify our problem, we focused on OAT1 because of the available data describing

its function, though probenecid inhibits both OAT1 and OAT3 at comparable levels. We

collected data from all relevant in vivo and in vitro experiments to better understand the full

extent of OAT1 interacting compounds. This literature search yielded nearly 500 compounds and

included drugs, endogenous metabolites, and natural products. We then explored the chemical

space covered by these using dimensionality reduction techniques on the molecular descriptors

and found that based on a categorical assignment (drug, endogenous metabolite, natural product)

the compounds could be largely separated with linear discriminant analysis, though the

separation was not perfect for some cases. We further explored this by performing K-Means

clustering using the molecular descriptors on the compounds and identified 5 or 6 as the optimal

number of clusters. Most of these clusters included compounds from different classes (drug,

endogenous metabolite, natural product), raising the possibility that features inherent to the chemical structures may be more important than the known function of the compounds.

Given the structural and functional diversity among these compounds, we then analyzed the protein-ligand interactions between the predicted human OAT1 structure and the wide array of compounds identified in our literature search with molecular docking studies. We first added physiological relevance to our studies by embedding the human OAT1 structure in a mammalian plasma membrane and identifying the most likely binding sites. These studies suggested that traditional probes of OAT1 function interacted with distinct sets of residues within a putative binding pocket on the intracellular region of the transporter. Though they must be supported by in vitro assays with the implicated residues altered, it is possible that there are multiple, distinct binding mechanisms for a specific drug transporter that contributes to its diverse function.

Chapters 2-5 of this dissertation represent a major advance in our understanding the physiological role of OAT1 and some other drug-handling genes (e.g., OAT3, OATP1B1). The strategies employed here can be adapted to focus on other multi-specific, oligo-specific, or mono-specific transporters to better characterize the seemingly important role in helping the regulation of endogenous metabolism.

6.2 FUTURE DIRECTIONS

While this dissertation represents a major step forward in understanding the function of OAT1, it is by no means a complete exploration of drug transporter function. The body of work described here can be built upon with some of the following future research directions.

The endogenous role of OAT1 presented in this dissertation makes use of in vivo serum and urine metabolomics data to establish the impact that a renal drug transporter can have on the circulating levels of dozens of molecules not previously associated with OAT1. These studies are

inherently limited by the libraries of small molecule compounds available with current metabolomics technologies. In the studies described in this dissertation, roughly ~1300 identified, unique metabolites are detected in either human plasma, human urine, or mouse plasma. The libraries of available compounds will only expand in the coming years as metabolomics technologies advance. Furthermore, some metabolomics platforms are able to measure exact concentrations of specific compounds rather than semiquantitative values that can only be compared across samples. Specific values for the levels of compounds in different conditions, such as those in this dissertation, can provide insight on the capacity of OAT1 and other drug transporters. Other genetic knockout mouse models can be analyzed using these approaches.

Improved metabolomics can also be coupled with human studies, such as those described in this dissertation. The probenecid experiment described in Chapter 3 demonstrates the impact that a non-toxic drug has on the plasma and urine metabolome after a single oral dosage. This striking effect on endogenous metabolism may be, in part, recapitulated by other drugs that inhibit OAT1/3. Probenecid's mechanism of action is the inhibition of transporters, but other drugs (e.g., antivirals, antibiotics, NSAIDs) may lead to similar outcomes given that they are transported by or inhibit OAT1. In a similar fashion, common combinations of drugs may exacerbate the individual effect by further inhibiting OAT1 or perturbing the function of other transporters/enzymes. Given the increasing rate of drug prescriptions, it is important to account for multiple xenobiotics stressing the system and the potential drug-metabolite interactions that occur.

Beyond OAT1, drugs that interact with other transporters (e.g., statins with OATPs) can also be administered to healthy patients and the appropriate body fluids can be collected for

metabolomics analysis. In this dissertation, as we were focused on renal drug transporters, the plasma and urine provided us with information on the putative function of OAT1 and OAT3. For hepatic and intestinal transporters, bile and feces samples, respectively, would be more indicative of function. Collecting these biological samples and others would also support studies such as the genomic work described in Chapter 4 of this dissertation. While in our study we focus on the plasma, other biological samples, such as the urine, bile, feces, or cerebrospinal fluid, would be useful in showing the total effect of SNPs in genes that are primarily expressed in excretory organs. Future experiments could also factor in disease states, as it has been shown that chronic kidney disease has an impact on the expression of renal and extrarenal drug transporters [252]. Whether SNPs in these genes would exacerbate the impaired function is an important factor in personalized treatment.

The work presented in Chapter 5 of this dissertation describes the use of ligand-based machine learning models to predict transporter function. While this work combines newly available in vivo metabolomics data with machine learning algorithms, there is still room for improvement. For one, the quality of the data that is input into the models will only improve over time. As previously described, the current metabolomics technology in the plasma and the urine limited us to 400 metabolites measured in the same experiment. This is by no means the full repertoire of metabolites that are likely to exist in physiology, and future work will likely discover more metabolites in the blood, urine, bile, or feces, which could uncover further roles for drug transporters.

In addition to in vivo metabolomics data, higher quality in vitro data can also lead to different kinds of models that can even be combined with the in vivo data. Building ligand-based models from in vitro data is difficult because of the various systems, conditions, and probes used

across multiple groups. Furthermore, it is common for only positive results to be published (i.e., significant inhibition, transport), making it challenging to populate a dataset with sufficient positive and negative outcomes. High throughput data from the same lab and experimental conditions are also very scarce. As these datasets become easier to generate and publicly available, more models will be built, leading to increased interpretation of differing results, and ultimately, a better understanding of transporter function.

A shortcoming of the machine learning framework that we developed to better understand changes in in vivo metabolomics data is that all compounds in the altered class are treated the same. Though we did apply a fold change and p-value threshold to identify these altered compounds, there does appear to be some chemical diversity between the compounds present. Separating them into more distinct groups based on their features may also lead to multiple models that focus on a smaller subset of compounds. The splitting of altered compounds into multiple groups could be performed based on chemical diversity but would likely be better informed by expertise with respect to a specific transporter and its known function. Machine learning technologies have also advanced in recent years. For example, XGBoost and LightGBM are novel algorithms that have been developed within the last decade. Future algorithms could lead to more robust results.

Chapter 5 of this dissertation makes use of a high confidence human protein structure that proved useful in learning more about OAT1 function at the atomic level. However, these models can be much improved. Experimentally derived structures would likely yield more accurate representations of the protein's folding with and without specific ligands bound. Observing transporters with bound ligands from the different clusters identified in this work would clarify binding mechanisms and whether competition for particular binding sites is occurring. Another

172

shortcoming of this dissertation's study is that we focus on only a single conformation of the transporter. Transport and inhibition are complex processes where the transporter goes through multiple conformations, such as outward open, inward open, outward occluded, and inward occluded. For a protein-based model to fully capture the biological process, molecular dynamics must be able to capture all these steps with each experimentally tested compound. Docking and transport results will also benefit from advances in algorithms and computing power that allow for the exploration of more demanding simulations.

From the physiological perspective, a model that incorporates the other transporters (sodium-potassium pump, NaDC3) required for tertiary OAT1 function would more accurately represent the process, as these membrane proteins could have an important effect on the overall rate of transport. The sodium potassium pump creates an electrochemical gradient that supports the sodium dicarboxylate exchange occurring through NaDC3, which ultimately contributes to the intracellular level of dicarboxylates. OAT1 operates via exchange of an extracellular organic anion for an intracellular dicarboxylate. These mechanistic processes have not been factored into any model and are likely necessary for a complete understanding of how the transporter operates.

REFERENCES

[1] P. Borst, R.O. Elferink. Mammalian ABC transporters in health and disease. *Annu Rev Biochem* **71** (2002) 537-592. https://doi.org/10.1146/annurev.biochem.71.102301.093055.

[2] M.D. Pizzagalli, A. Bensimon, G. Superti-Furga. A guide to plasma membrane solute carrier proteins. *FEBS J* **288** (2021) 2784-2835. https://doi.org/10.1111/febs.15531.

[3] K.M. Giacomini, S.M. Huang, D.J. Tweedie, L.Z. Benet, K.L.R. Brouwer, X.Y. Chu, A. Dahlin, R. Evers, V. Fischer, K.M. Hillgren, K.A. Hoffmaster, T. Ishikawa, D. Keppler, R.B. Kim, C.A. Lee, M. Niemi, J.W. Polli, Y. Sugiyama, P.W. Swaan, J.A. Ware, S.H. Wright, S.W. Yee, M.J. Zamek-Gliszczynski, L. Zhang, I. Transporter. Membrane transporters in drug development. *Nature Reviews Drug Discovery* **9** (2010) 215-236. https://doi.org/10.1038/nrd3028.

[4] S. Sudsakorn, P. Bahadduri, J. Fretland, C. Lu. 2020 FDA Drug-drug Interaction Guidance: A Comparison Analysis and Action Plan by Pharmaceutical Industrial Scientists. *Curr Drug Metab* **21** (2020) 403-426. https://doi.org/10.2174/1389200221666200620210522.

[5] S.K. Nigam, K.T. Bush, G. Martovetsky, S.Y. Ahn, H.C. Liu, E. Richard, V. Bhatnagar, W. Wu. The organic anion transporter (OAT) family: a systems biology perspective. *Physiol Rev* **95** (2015) 83-123. https://doi.org/10.1152/physrev.00025.2013.

[6] S.K. Nigam. What do drug transporters really do? *Nature Reviews Drug Discovery* **14** (2015) 29-44. https://doi.org/10.1038/nrd4461.

[7] W. Wu, A.V. Dnyanmote, S.K. Nigam. Remote Communication through Solute Carriers and ATP Binding Cassette Drug Transporter Pathways: An Update on the Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology* **79** (2011) 795-805. https://doi.org/10.1124/mol.110.070607.

[8] S.Y. Ahn, S.K. Nigam. Toward a Systems Level Understanding of Organic Anion and Other Multispecific Drug Transporters: A Remote Sensing and Signaling Hypothesis. *Molecular Pharmacology* **76** (2009) 481-490. https://doi.org/10.1124/mol.109.056564.

[9] S.K. Nigam, J.C. Granados. A Biological Basis for Pharmacokinetics: The Remote Sensing and Signaling Theory. *Clin Pharmacol Ther* (2022). https://doi.org/10.1002/cpt.2634.

[10] S.K. Nigam, J.C. Granados. OAT, OATP, and MRP Drug Transporters and the Remote Sensing and Signaling Theory. *Annu Rev Pharmacol Toxicol* **63** (2023) 637-660. https://doi.org/10.1146/annurev-pharmtox-030322-084058.

[11] S.B. Rosenthal, K.T. Bush, S.K. Nigam. A Network of SLC and ABC Transporter and DME Genes Involved in Remote Sensing and Signaling in the Gut-Liver-Kidney Axis. *Scientific Reports* **9** (2019). https://doi.org/10.1038/s41598-019-47798-x.

[12] V. Bhatnagar, E.L. Richard, W. Wu, C.M. Nievergelt, M.S. Lipkowitz, J. Jeff, A.X. Maihofer, S.K. Nigam. Analysis of ABCG2 and other urate transporters in uric acid homeostasis in chronic kidney disease: potential role of remote sensing and signaling. *Clinical Kidney Journal* **9** (2016) 444-453. https://doi.org/10.1093/ckj/sfw010.

[13] M. Iqbal, M.C. Audette, S. Petropoulos, W. Gibb, S.G. Matthews. Placental drug transporters and their role in fetal protection. *Placenta* **33** (2012) 137-142. https://doi.org/10.1016/j.placenta.2012.01.008.

[14] M. Yamashita, U.R. Markert. Overview of Drug Transporters in Human Placenta. *Int J Mol Sci* **22** (2021). https://doi.org/10.3390/ijms222313149.

[15] D. Kozlosky, E. Barrett, L.M. Aleksunes. Regulation of Placental Efflux Transporters during Pregnancy Complications. *Drug Metab Dispos* **50** (2022) 1364-1375. https://doi.org/10.1124/dmd.121.000449.

[16] K.A. Krautkramer, J. Fan, F. Backhed. Gut microbial metabolites as multi-kingdom intermediates. *Nat Rev Microbiol* **19** (2021) 77-94. https://doi.org/10.1038/s41579-020-0438-4.

[17] Z.N. Wang, Y.Z. Zhao. Gut microbiota derived metabolites in cardiovascular health and disease. *Protein & Cell* **9** (2018) 416-431. https://doi.org/10.1007/s13238-018-0549-0.

[18] A. Parker, S. Fonseca, S.R. Carding. Gut microbes and metabolites as modulators of blood-brain barrier integrity and brain health. *Gut Microbes* **11** (2020) 135-157. https://doi.org/10.1080/19490976.2019.1638722.

[19] H. Sezutsu, G. Le Goff, R. Feyereisen. Origins of P450 diversity. *Philos Trans R Soc Lond B Biol Sci* **368** (2013) 20120428. https://doi.org/10.1098/rstb.2012.0428.

[20] M. Dean, T. Annilo. Evolution of the ATP-binding cassette (ABC) transporter superfamily in vertebrates. *Annu Rev Genomics Hum Genet* **6** (2005) 123-142. https://doi.org/10.1146/annurev.genom.6.080604.162122.

[21] E.I. Lepist, A.S. Ray. Beyond drug-drug interactions: effects of transporter inhibition on endobiotics, nutrients and toxins. *Expert Opinion on Drug Metabolism & Toxicology* **13** (2017) 1075-1087. https://doi.org/10.1080/17425255.2017.1372425.

[22] N. Robbins, S.E. Koch, M. Tranter, J. Rubinstein. The history and future of probenecid. *Cardiovasc Toxicol* **12** (2012) 1-9. https://doi.org/10.1007/s12012-011-9145-8.

[23] J.J. Coleman, S.K. Pontefract. Adverse drug reactions. *Clin Med (Lond)* **16** (2016) 481-485. https://doi.org/10.7861/clinmedicine.16-5-481.

[24] A.D. Halli-Tierney, C. Scarbrough, D. Carroll. Polypharmacy: Evaluating Risks and Deprescribing. *Am Fam Physician* **100** (2019) 32-38.

[25] M.R. Wofford, D.S. King, T.K. Harrell. Drug-induced metabolic syndrome. *J Clin Hypertens (Greenwich)* **8** (2006) 114-119. https://doi.org/10.1111/j.1524-6175.2006.04751.x.

[26] O. Bamgbola. Metabolic consequences of modern immunosuppressive agents in solid organ transplantation. *Ther Adv Endocrinol Metab* **7** (2016) 110-127. https://doi.org/10.1177/2042018816641580.

[27] H.E. Ergin, E.E. Inga, T.Z. Maung, M. Javed, S. Khan. HIV, Antiretroviral Therapy and Metabolic Alterations: A Review. *Cureus* **12** (2020) e8059. https://doi.org/10.7759/cureus.8059.

[28] D.A. Volpe. Transporter assays as useful in vitro tools in drug discovery and development. *Expert Opin Drug Discov* **11** (2016) 91-103. https://doi.org/10.1517/17460441.2016.1101064.

[29] D.A. Volpe, P.V. Balimane. Application of in vitro CYP and transporter assays to predict clinical drug-drug interactions. *Bioanalysis* **10** (2018) 619-623. https://doi.org/10.4155/bio-2017-0266.

[30] J.M. Malinowska, T. Palosaari, J. Sund, D. Carpi, M. Bouhifd, R.J.M. Weber, M. Whelan, M.R. Viant. Integrating in vitro metabolomics with a 96-well high-throughput screening platform. *Metabolomics* **18** (2022) 11. https://doi.org/10.1007/s11306-021-01867-3.

[31] S.W. Yee, D. Buitrago, A. Stecula, H.X. Ngo, H.C. Chien, L. Zou, M.L. Koleske, K.M. Giacomini. Deorphaning a solute carrier 22 family member, SLC22A15, through functional genomic studies. *FASEB J* **34** (2020) 15734-15752. https://doi.org/10.1096/fj.202001497R.

[32] R. Ghaffarian, S. Muro. Models and methods to evaluate transport of drug delivery systems across cellular barriers. *J Vis Exp* (2013) e50638. https://doi.org/10.3791/50638.

[33] M. Louisa, F.D. Suyatna, S.I. Wanandi, P.B. Asih, D. Syafruddin. Differential expression of several drug transporter genes in HepG2 and Huh-7 cell lines. *Adv Biomed Res* **5** (2016) 104. https://doi.org/10.4103/2277-9175.183664.

[34] G. Ahlin, C. Hilgendorf, J. Karlsson, C.A. Szigyarto, M. Uhlen, P. Artursson. Endogenous gene and protein expression of drug-transporting proteins in cell lines routinely used in drug discovery programs. *Drug Metab Dispos* **37** (2009) 2275-2283. https://doi.org/10.1124/dmd.109.028654.

[35] B.S. Freedman. Physiology assays in human kidney organoids. *Am J Physiol Renal Physiol* **322** (2022) F625-F638. https://doi.org/10.1152/ajprenal.00400.2021.

[36] T. Zietek, P. Giesbertz, M. Ewers, F. Reichart, M. Weinmuller, E. Urbauer, D. Haller, I.E. Demir, G.O. Ceyhan, H. Kessler, E. Rath. Organoids to Study Intestinal Nutrient Transport, Drug Uptake and Metabolism - Update to the Human Model and Expansion of Applications. *Front Bioeng Biotechnol* **8** (2020) 577656. https://doi.org/10.3389/fbioe.2020.577656.

[37] S. Akbari, N. Arslan, S. Senturk, E. Erdal. Next-Generation Liver Medicine Using Organoid Models. *Front Cell Dev Biol* **7** (2019) 345. https://doi.org/10.3389/fcell.2019.00345.

[38] J. Deng, W. Wei, Z. Chen, B. Lin, W. Zhao, Y. Luo, X. Zhang. Engineered Liver-on-a-Chip Platform to Mimic Liver Functions and Its Biomedical Applications: A Review. *Micromachines (Basel)* **10** (2019). https://doi.org/10.3390/mi10100676.

[39] A. Bein, W. Shin, S. Jalili-Firoozinezhad, M.H. Park, A. Sontheimer-Phelps, A. Tovaglieri, A. Chalkiadaki, H.J. Kim, D.E. Ingber. Microfluidic Organ-on-a-Chip Models of Human Intestine. *Cell Mol Gastroenterol Hepatol* **5** (2018) 659-668. https://doi.org/10.1016/j.jcmgh.2017.12.010.

[40] S. Kim, S. Takayama. Organ-on-a-chip and the kidney. *Kidney Res Clin Pract* **34** (2015) 165-169. https://doi.org/10.1016/j.krcp.2015.08.001.

[41] Y. Lu, S. Kim, K. Park. In vitro-in vivo correlation: perspectives on model development. *Int J Pharm* **418** (2011) 142-148. https://doi.org/10.1016/j.ijpharm.2011.01.010.

[42] S.A. Eraly, V. Vallon, D.A. Vaughn, J.A. Gangoiti, K. Richter, M. Nagle, J.C. Monte, T. Rieg, D.M. Truong, J.M. Long, B.A. Barshop, G. Kaler, S.K. Nigam. Decreased renal organic anion secretion and plasma accumulation of endogenous organic anions in OAT1 knock-out mice. *Journal of Biological Chemistry* **281** (2006) 5072-5083. https://doi.org/10.1074/jbc.M508050200.

[43] X.Y. Chu, J.R. Strauss, M.A. Mariano, J. Li, D.J. Newton, X. Cai, R.W. Wang, J. Yabut, D.P. Hartley, D.C. Evans, R. Evers. Characterization of mice lacking the multidrug resistance protein MRP2 (ABCC2). *J Pharmacol Exp Ther* **317** (2006) 579-589. https://doi.org/10.1124/jpet.105.098665.

[44] M.G. Belinsky, P.A. Dawson, I. Shchaveleva, L.J. Bain, R. Wang, V. Ling, Z.S. Chen, A. Grinberg, H. Westphal, A. Klein-Szanto, A. Lerro, G.D. Kruh. Analysis of the in vivo functions of Mrp3. *Mol Pharmacol* **68** (2005) 160-168. https://doi.org/10.1124/mol.104.010587.

[45] D.H. Sweet, D.S. Miller, J.B. Pritchard, Y. Fujiwara, D.R. Beier, S.K. Nigam. Impaired organic anion transport in kidney and choroid plexus of organic anion transporter 3 (Oat3 (Slc22a8)) knockout mice. *Journal of Biological Chemistry* **277** (2002) 26934-26943. https://doi.org/10.1074/jbc.M203803200.

[46] Y. Zhang, I.L. Csanaky, L.D. Lehman-McKeeman, C.D. Klaassen. Loss of organic anion transporting polypeptide 1a1 increases deoxycholic acid absorption in mice by increasing intestinal permeability. *Toxicol Sci* **124** (2011) 251-260. https://doi.org/10.1093/toxsci/kfr236.

[47] K.B. Scheidweiler, B. Ladenheim, J.L. Cadet, M.A. Huestis. Mice lacking multidrug resistance protein 1a show altered dopaminergic responses to methylenedioxymethamphetamine (MDMA) in striatum. *Neurotox Res* **18** (2010) 200-209. https://doi.org/10.1007/s12640-009-9124-z.

[48] E. van de Steeg, E. Wagenaar, C.M. van der Kruijssen, J.E. Burggraaff, D.R. de Waart, R.P. Elferink, K.E. Kenworthy, A.H. Schinkel. Organic anion transporting polypeptide 1a/1b-knockout mice provide insights into hepatic handling of bilirubin, bile acids, and drugs. *J Clin Invest* **120** (2010) 2942-2952. https://doi.org/10.1172/JCI42168.

[49] M.L. Vlaming, A. van Esch, Z. Pala, E. Wagenaar, K. van de Wetering, O. van Tellingen, A.H. Schinkel. Abcc2 (Mrp2), Abcc3 (Mrp3), and Abcg2 (Bcrp1) are the main determinants for rapid elimination of methotrexate and its toxic metabolite 7-hydroxymethotrexate in vivo. *Mol Cancer Ther* **8** (2009) 3350-3359. https://doi.org/10.1158/1535-7163.MCT-09-0668.

[50] J.W. Higgins, J.Q. Bao, A.B. Ke, J.R. Manro, J.K. Fallon, P.C. Smith, M.J. Zamek-Gliszczynski. Utility of Oatp1a/1b-knockout and OATP1B1/3-humanized mice in the study of OATP-mediated pharmacokinetics and tissue distribution: case studies with pravastatin, atorvastatin, simvastatin, and carboxydichlorofluorescein. *Drug Metab Dispos* **42** (2014) 182-192. https://doi.org/10.1124/dmd.113.054783.

[51] D. Iusuf, M. Ludwig, A. Elbatsh, A. van Esch, E. van de Steeg, E. Wagenaar, M. van der Valk, F. Lin, O. van Tellingen, A.H. Schinkel. OATP1A/1B transporters affect irinotecan and SN-38 pharmacokinetics and carboxylesterase expression in knockout and humanized transgenic mice. *Mol Cancer Ther* **13** (2014) 492-503. https://doi.org/10.1158/1535-7163.MCT-13-0541.

[52] W. Wu, K.T. Bush, S.K. Nigam. Key Role for the Organic Anion Transporters, OAT1 and OAT3, in the in vivo Handling of Uremic Toxins and Solutes. *Sci Rep* **7** (2017) 4939. https://doi.org/10.1038/s41598-017-04949-2.

[53] W.R. Wikoff, M.A. Nagle, V.L. Kouznetsova, I.F. Tsigelny, S.K. Nigam. Untargeted metabolomics identifies enterobiome metabolites and putative uremic toxins as substrates of organic anion transporter 1 (Oat1). *J Proteome Res* **10** (2011) 2842-2851. https://doi.org/10.1021/pr200093w.

[54] J.C. Granados, A.K. Nigam, K.T. Bush, N. Jamshidi, S.K. Nigam. A key role for the transporter OAT1 in systemic lipid metabolism. *J Biol Chem* (2021) 100603. https://doi.org/10.1016/j.jbc.2021.100603.

[55] V. Vallon, T. Rieg, S.Y. Ahn, W. Wu, S.A. Eraly, S.K. Nigam. Overlapping in vitro and in vivo specificities of the organic anion transporters OAT1 and OAT3 for loop and thiazide diuretics. *American Journal of Physiology-Renal Physiology* **294** (2008) F867-F873. https://doi.org/10.1152/ajprenal.00528.2007.

[56] S. Ganguly, D. Finkelstein, T.I. Shaw, R.D. Michalek, K.M. Zorn, S. Ekins, K. Yasuda, Y. Fukuda, J.D. Schuetz, K. Mukherjee, E.G. Schuetz. Metabolomic and transcriptomic analysis reveals endogenous substrates and metabolic adaptation in rats lacking Abcg2 and Abcb1a transporters. *PLoS One* **16** (2021) e0253852. https://doi.org/10.1371/journal.pone.0253852.

[57] X. Gou, F. Ran, J. Yang, Y. Ma, X. Wu. Construction and Evaluation of a Novel Organic Anion Transporter 1/3 CRISPR/Cas9 Double-Knockout Rat Model. *Pharmaceutics* **14** (2022). https://doi.org/10.3390/pharmaceutics14112307.

[58] P. Zhang, P. Azad, D.C. Engelhart, G.G. Haddad, S.K. Nigam. SLC22 Transporters in the Fly Renal System Regulate Response to Oxidative Stress In Vivo. *Int J Mol Sci* **22** (2021). https://doi.org/10.3390/ijms222413407.

[59] D.C. Engelhart, P. Azad, S. Ali, J.C. Granados, G.G. Haddad, S.K. Nigam. Drosophila SLC22 Orthologs Related to OATs, OCTs, and OCTNs Regulate Development and Responsiveness to Oxidative Stress. *Int J Mol Sci* **21** (2020). https://doi.org/10.3390/ijms21062002.

[60] A. Buniello, J.A.L. MacArthur, M. Cerezo, L.W. Harris, J. Hayhurst, C. Malangone, A. McMahon, J. Morales, E. Mountjoy, E. Sollis, D. Suveges, O. Vrousgou, P.L. Whetzel, R. Amode, J.A. Guillen, H.S. Riat, S.J. Trevanion, P. Hall, H. Junkins, P. Flicek, T. Burdett, L.A. Hindorff, F. Cunningham, H. Parkinson. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47** (2019) D1005-D1012. https://doi.org/10.1093/nar/gky1120.

[61] E. Cano-Gamez, G. Trynka. From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. *Frontiers in Genetics* **11** (2020). https://doi.org/10.3389/fgene.2020.00424.

[62] F.Y. Li, A.W. El-Hattab, E.V. Bawle, R.G. Boles, E.S. Schmitt, F. Scaglia, L.J. Wong. Molecular Spectrum of SLC22A5 (OCTN2) Gene Mutations Detected in 143 Subjects Evaluated for Systemic Carnitine Deficiency. *Human Mutation* **31** (2010) E1632-E1651. https://doi.org/10.1002/humu.21311.

[63] K. Suhre, S.Y. Shin, A.K. Petersen, R.P. Mohney, D. Meredith, B. Wagele, E. Altmaier, P. Deloukas, J. Erdmann, E. Grundberg, C.J. Hammond, M.H. de Angelis, G. Kastenmuller, A. Kottgen, F. Kronenberg, M. Mangino, C. Meisinger, T. Meitinger, H.W. Mewes, M.V. Milburn, C. Prehn, J. Raffler, J.S. Ried, W. Romisch-Margl, N.J. Samani, K.S. Small, H.E. Wichmann, G.J. Zhai, T. Illig, T.D. Spector, J. Adamski, N. Soranzo, C. Gieger, CARDIoGRAM. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **477** (2011) 54-U60. https://doi.org/10.1038/nature10354.

[64] S.Y. Shin, E.B. Fauman, A.K. Petersen, J. Krumsiek, R. Santos, J. Huang, M. Arnold, I. Erte, V. Forgetta, T.P. Yang, K. Walter, C. Menni, L. Chen, L. Vasquez, A.M. Valdes, C.L. Hyde, V. Wang, D. Ziemek, P. Roberts, L. Xi, E. Grundberg, M. Waldenberger, J.B. Richards, R.P. Mohney, M.V. Milburn, S.L. John, J. Trimmer, F.J. Theis, J.P. Overington, K. Suhre, M.J. Brosnan, C. Gieger, G. Kastenmuller, T.D. Spector, N. Soranzo, M.T.H.E. R. An atlas of genetic influences on human blood metabolites. *Nature Genetics* **46** (2014) 543-550. https://doi.org/10.1038/ng.2982.

[65] P. Schlosser, Y. Li, P. Sekula, J. Raffler, F. Grundner-Culemann, M. Pietzner, Y. Cheng, M. Wuttke, I. Steinbrenner, U.T. Schultheiss, F. Kotsis, T. Kacprowski, L. Forer, B. Hausknecht, A.B. Ekici, M. Nauck, U. Volker, G. Investigators, G. Walz, P.J. Oefner, F. Kronenberg, R.P. Mohney, M. Kottgen, K. Suhre, K.U. Eckardt, G. Kastenmuller, A. Kottgen. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. *Nat Genet* **52** (2020) 167-176. https://doi.org/10.1038/s41588-019-0567-8.

[66] B. Zagidullin, Z. Wang, Y. Guan, E. Pitkanen, J. Tang. Comparative analysis of molecular fingerprints in prediction of drug combination effects. *Brief Bioinform* **22** (2021). https://doi.org/10.1093/bib/bbab291.

[67] K. Roy, R.N. Das. A review on principles, theory and practices of 2D-QSAR. *Curr Drug Metab* **15** (2014) 346-379. https://doi.org/10.2174/1389200215666140908102230.

[68] Y.C. Lo, S.E. Rensi, W. Torng, R.B. Altman. Machine learning in chemoinformatics and drug discovery. *Drug Discov Today* **23** (2018) 1538-1546. https://doi.org/10.1016/j.drudis.2018.05.010.

[69] T.R. Lane, F. Urbina, X. Zhang, M. Fye, J. Gerlach, S.H. Wright, S. Ekins. Machine Learning Models Identify New Inhibitors for Human OATP1B1. *Mol Pharm* **19** (2022) 4320-4332. https://doi.org/10.1021/acs.molpharmaceut.2c00662.

[70] V. Belekar, K. Lingineni, P. Garg. Classification of Breast Cancer Resistant Protein (BCRP) Inhibitors and Non-Inhibitors Using Machine Learning Approaches. *Comb Chem High Throughput Screen* **18** (2015) 476-485. https://doi.org/10.2174/1386207318666150525094503.

[71] P.H. Wang, Y.S. Tu, Y.J. Tseng. PgpRules: a decision tree based prediction server for P-glycoprotein substrates and inhibitors. *Bioinformatics* **35** (2019) 4193-4195. https://doi.org/10.1093/bioinformatics/btz213.

[72] K.T. Bush, W. Wu, C. Lun, S.K. Nigam. The drug transporter OAT3 (SLC22A8) and endogenous metabolite communication via the gut-liver-kidney axis. *J Biol Chem* **292** (2017) 15789-15803. https://doi.org/10.1074/jbc.M117.796516.

[73] J.C. Granados, A. Richelle, J.M. Gutierrez, P. Zhang, X. Zhang, V. Bhatnagar, N.E. Lewis, S.K. Nigam. Coordinate regulation of systemic and kidney tryptophan metabolism by the drug transporters OAT1 and OAT3. *J Biol Chem* (2021) 100575. https://doi.org/10.1016/j.jbc.2021.100575.

[74] A.K. Nigam, J.G. Li, K. Lall, D. Shi, K.T. Bush, V. Bhatnagar, R. Abagyan, S.K. Nigam. Unique metabolite preferences of the drug transporters OAT1 and OAT3 analyzed by machine learning. *J Biol Chem* **295** (2020) 1829-1842. https://doi.org/10.1074/jbc.RA119.010729.

[75] A.K. Nigam, A.A. Ojha, J.G. Li, D. Shi, V. Bhatnagar, K.B. Nigam, R. Abagyan, S.K. Nigam. Molecular Properties of Drugs Handled by Kidney OATs and Liver OATPs Revealed by

Chemoinformatics and Machine Learning: Implications for Kidney and Liver Disease. *Pharmaceutics* **13** (2021). https://doi.org/10.3390/pharmaceutics13101720.

[76] H.C. Liu, A. Goldenberg, Y. Chen, C. Lun, W. Wu, K.T. Bush, N. Balac, P. Rodriguez, R. Abagyan, S.K. Nigam. Molecular Properties of Drugs Interacting with SLC22 Transporters OAT1, OAT3, OCT1, and OCT2: A Machine-Learning Approach. *J Pharmacol Exp Ther* **359** (2016) 215-229. https://doi.org/10.1124/jpet.116.232660.

[77] H.C. Liu, N. Jamshidi, Y.C. Chen, S.A. Eraly, S.Y. Cho, V. Bhatnagar, W. Wu, K.T. Bush, R. Abagyan, B.O. Palsson, S.K. Nigam. An Organic Anion Transporter 1 (OAT1)-centered Metabolic Network. *Journal of Biological Chemistry* **291** (2016) 19474-19486. https://doi.org/10.1074/jbc.M116.745216.

[78] C. Chang, S. Ekins, P. Bahadduri, P.W. Swaan. Pharmacophore-based discovery of ligands for drug transporters. *Adv Drug Deliv Rev* **58** (2006) 1431-1450. https://doi.org/10.1016/j.addr.2006.09.006.

[79] J. Zhao, Y. Cao, L. Zhang. Exploring the computational methods for protein-ligand binding site prediction. *Comput Struct Biotechnol J* **18** (2020) 417-426. https://doi.org/10.1016/j.csbj.2020.02.008.

[80] A. Schlessinger, M.A. Welch, H. van Vlijmen, K. Korzekwa, P.W. Swaan, P. Matsson. Molecular Modeling of Drug-Transporter Interactions-An International Transporter Consortium Perspective. *Clin Pharmacol Ther* **104** (2018) 818-835. https://doi.org/10.1002/cpt.1174.

[81] A. Schlessinger, N. Khuri, K.M. Giacomini, A. Sali. Molecular modeling and ligand docking for solute carrier (SLC) transporters. *Curr Top Med Chem* **13** (2013) 843-856. https://doi.org/10.2174/1568026611313070007.

[82] A.G. Roberts. The Structure and Mechanism of Drug Transporters. *Methods Mol Biol* **2342** (2021) 193-234. https://doi.org/10.1007/978-1-0716-1554-6_8.

[83] Y. Kim, J. Chen. Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science* **359** (2018) 915-919. https://doi.org/10.1126/science.aar7389.

[84] A. Rasouli, Q. Yu, S. Dehghani-Ghahnaviyeh, P.C. Wen, J. Kowal, K.P. Locher, E. Tajkhorshid. Differential dynamics and direct interaction of bound ligands with lipids in multidrug transporter ABCG2. *Proc Natl Acad Sci U S A* **120** (2023) e2213437120. https://doi.org/10.1073/pnas.2213437120.

[85] N.M.I. Taylor, I. Manolaridis, S.M. Jackson, J. Kowal, H. Stahlberg, K.P. Locher. Structure of the human multidrug transporter ABCG2. *Nature* **546** (2017) 504-509. https://doi.org/10.1038/nature22345.

[86] I.F. Tsigelny, D. Kovalskyy, V.L. Kouznetsova, O. Balinskyi, Y. Sharikov, V. Bhatnagar, S.K. Nigam. Conformational changes of the multispecific transporter organic anion transporter 1

(OAT1/SLC22A6) suggests a molecular mechanism for initial stages of drug and metabolite transport. *Cell Biochem Biophys* **61** (2011) 251-259. https://doi.org/10.1007/s12013-011-9191-7.

[87] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A.A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis. Highly accurate protein structure prediction with AlphaFold. *Nature* **596** (2021) 583-589. https://doi.org/10.1038/s41586-021-03819-2.

[88] A. Tuerkova, O. Ungvari, R. Laczko-Rigo, E. Mernyak, G. Szakacs, C. Ozvegy-Laczka, B. Zdrazil. Data-Driven Ensemble Docking to Map Molecular Interactions of Steroid Analogs with Hepatic Organic Anion Transporting Polypeptides. *J Chem Inf Model* **61** (2021) 3109-3127. https://doi.org/10.1021/acs.jcim.1c00362.

[89] A. Janaszkiewicz, A. Toth, Q. Faucher, M. Martin, B. Chantemargue, C. Barin-Le Guellec, P. Marquet, F. Di Meo. Insights into the structure and function of the human organic anion transporter 1 in lipid bilayer membranes. *Sci Rep* **12** (2022) 7057. https://doi.org/10.1038/s41598-022-10755-2.

[90] Y. Suo, N.J. Wright, H. Guterres, J.G. Fedor, K.J. Butay, M.J. Borgnia, W. Im, S.Y. Lee. Molecular basis of polyspecific drug binding and transport by OCT1 and OCT2. *bioRxiv* (2023). https://doi.org/10.1101/2023.03.15.532610.

[91] A. Janaszkiewicz, A. Toth, Q. Faucher, H. Arnion, N. Vedrenne, C. Barin-Le Guellec, P. Marquet, F. Di Meo. Substrate binding and lipid-mediated allostery in the human organic anion transporter 1 at the atomic-scale. *Biomed Pharmacother* **160** (2023) 114342. https://doi.org/10.1016/j.biopha.2023.114342.

[92] J. Schellenberger, R. Que, R.M. Fleming, I. Thiele, J.D. Orth, A.M. Feist, D.C. Zielinski, A. Bordbar, N.E. Lewis, S. Rahmanian, J. Kang, D.R. Hyduke, B.O. Palsson. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat Protoc* **6** (2011) 1290-1307. https://doi.org/10.1038/nprot.2011.308.

[93] S.A. Becker, B.O. Palsson. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol* **4** (2008) e1000082. https://doi.org/10.1371/journal.pcbi.1000082.

[94] W. Wu, N. Jamshidi, S.A. Eraly, H.C. Liu, K.T. Bush, B.O. Palsson, S.K. Nigam. Multispecific drug transporter Slc22a8 (Oat3) regulates multiple metabolic and signaling pathways. *Drug Metab Dispos* **41** (2013) 1825-1834. https://doi.org/10.1124/dmd.113.052647.

[95] S.Y. Ahn, N. Jamshidi, M.L. Mo, W. Wu, S.A. Eraly, A. Dnyanmote, K.T. Bush, T.F. Gallegos, D.H. Sweet, B.O. Palsson, S.K. Nigam. Linkage of Organic Anion Transporter-1 to Metabolic Pathways through Integrated "Omics"-driven Network and Functional Analysis.

Journal of Biological Chemistry **286** (2011) 31522-31531. https://doi.org/10.1074/jbc.M111.272534.

[96] E. Brunk, S. Sahoo, D.C. Zielinski, A. Altunkaya, A. Drager, N. Mih, F. Gatto, A. Nilsson, G.A. Preciat Gonzalez, M.K. Aurich, A. Prlic, A. Sastry, A.D. Danielsdottir, A. Heinken, A. Noronha, P.W. Rose, S.K. Burley, R.M.T. Fleming, J. Nielsen, I. Thiele, B.O. Palsson. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat Biotechnol* **36** (2018) 272-281. https://doi.org/10.1038/nbt.4072.

[97] S. Sahoo, H.S. Haraldsdottir, R.M. Fleming, I. Thiele. Modeling the effects of commonly used drugs on human metabolism. *FEBS J* **282** (2015) 297-317. https://doi.org/10.1111/febs.13128.

[98] C.E. Lopez-Nieto, G. You, K.T. Bush, E.J. Barros, D.R. Beier, S.K. Nigam. Molecular cloning and characterization of NKT, a gene product related to the organic cation transporter family that is almost exclusively expressed in the kidney. *J Biol Chem* **272** (1997) 6471-6478. https://doi.org/10.1074/jbc.272.10.6471.

[99] G. Burckhardt. Drug transport by Organic Anion Transporters (OATs). *Pharmacol Ther* **136** (2012) 106-130. https://doi.org/10.1016/j.pharmthera.2012.07.010.

[100] X. Wang, L.F. Han, G.T. Li, W. Peng, X.M. Gao, C.D. Klaassen, G.W. Fan, Y.C. Zhang. Identification of Natural Products as Inhibitors of Human Organic Anion Transporters (OAT1 and OAT3) and Their Protective Effect on Mercury-Induced Toxicity. *Toxicological Sciences* **161** (2018) 321-334. https://doi.org/10.1093/toxsci/kfx216.

[101] H. Lu, Z. Lu, X. Li, G. Li, Y. Qiao, R.P. Borris, Y. Zhang. Interactions of 172 plant extracts with human organic anion transporter 1 (SLC22A6) and 3 (SLC22A8): a study on herb-drug interactions. *PeerJ* **5** (2017) e3333. https://doi.org/10.7717/peerj.3333.

[102] A.M. Torres, A.V. Dnyanmote, K.T. Bush, W. Wu, S.K. Nigam. Deletion of multispecific organic anion transporter Oat1/Slc22a6 protects against mercury-induced kidney injury. *J Biol Chem* **286** (2011) 26391-26395. https://doi.org/10.1074/jbc.M111.249292.

[103] J. Zhang, H. Wang, Y. Fan, Z. Yu, G. You. Regulation of organic anion transporters: Role in physiology, pathophysiology, and drug elimination. *Pharmacol Ther* **217** (2021) 107647. https://doi.org/10.1016/j.pharmthera.2020.107647.

[104] M.F. Fromm. Transporters and drug-drug interactions: Important determinants of drug disposition and effects. *Toxicology Letters* **238** (2015) S49-S49. https://doi.org/10.1016/j.toxlet.2015.08.136.

[105] Y. Liang, S.Q. Li, L.G. Chen. The physiological role of drug transporters. *Protein & Cell* **6** (2015) 334-350. https://doi.org/10.1007/s13238-015-0148-2.

[106] K.B. Martinez, V. Leone, E.B. Chang. Microbial metabolites in health and disease: Navigating the unknown in search of function. *Journal of Biological Chemistry* **292** (2017) 8553-8559. https://doi.org/10.1074/jbc.R116.752899.

[107] W.R. Wikoff, A.T. Anfora, J. Liu, P.G. Schultz, S.A. Lesley, E.C. Peters, G. Siuzdak. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America* **106** (2009) 3698-3703. https://doi.org/10.1073/pnas.0812874106.

[108] G. Sharon, N. Garg, J. Debelius, R. Knight, P.C. Dorrestein, S.K. Mazmanian. Specialized Metabolites from the Microbiome in Health and Disease. *Cell Metabolism* **20** (2014) 719-730. https://doi.org/10.1016/j.cmet.2014.10.016.

[109] F. Hosseinkhani, A. Heinken, I. Thiele, P.W. Lindenburg, A.C. Harms, T. Hankemeier. The contribution of gut bacterial metabolites in the human immune signaling pathway of non-communicable diseases. *Gut Microbes* **13** (2021) 1-22. https://doi.org/10.1080/19490976.2021.1882927.

[110] W.H.W. Tang, T. Kitai, S.L. Hazen. Gut Microbiota in Cardiovascular Health and Disease. *Circulation Research* **120** (2017) 1183-1196. https://doi.org/10.1161/Circresaha.117.309715.

[111] K.L. Glassner, B.P. Abraham, E.M.M. Quigley. The microbiome and inflammatory bowel disease. *Journal of Allergy and Clinical Immunology* **145** (2020) 16-27. https://doi.org/10.1016/j.jaci.2019.11.003.

[112] A. Nallu, S. Sharma, A. Ramezani, J. Muralidharan, D. Raj. Gut microbiome in chronic kidney disease: challenges and opportunities. *Translational Research* **179** (2017) 24-37. https://doi.org/10.1016/j.trsl.2016.04.007.

[113] G.P. Hobby, O. Karaduta, G.F. Dusio, M. Singh, B.L. Zybailov, J.M. Arthur. Chronic kidney disease and the gut microbiome. *Am J Physiol Renal Physiol* **316** (2019) F1211-F1217. https://doi.org/10.1152/ajprenal.00298.2018.

[114] K.T. Bush, P. Singh, S.K. Nigam. Gut-derived uremic toxin handling in vivo requires OAT-mediated tubular secretion in chronic kidney disease. *Jci Insight* **5** (2020). https://doi.org/10.1172/jci.insight.133817.

[115] F. Duranton, G. Cohen, R. De Smet, M. Rodriguez, J. Jankowski, R. Vanholder, A. Argiles, E.U.T.W. Grp. Normal and Pathologic Concentrations of Uremic Toxins. *Journal of the American Society of Nephrology* **23** (2012) 1258-1270. https://doi.org/10.1681/Asn.2011121175.

[116] S.K. Nigam, K.T. Bush. Uraemic syndrome of chronic kidney disease: altered remote sensing and signalling. *Nature Reviews Nephrology* **15** (2019) 301-316. https://doi.org/10.1038/s41581-019-0111-1.

[117] M.X. Byndloss, A.J. Baumler. The germ-organ theory of non-communicable diseases. *Nat Rev Microbiol* **16** (2018) 103-110. https://doi.org/10.1038/nrmicro.2017.158.

[118] T. Suga, H. Yamaguchi, T. Sato, M. Maekawa, J. Goto, N. Mano. Preference of Conjugated Bile Acids over Unconjugated Bile Acids as Substrates for OATP1B1 and OATP1B3. *PLoS One* **12** (2017). https://doi.org/10.1371/journal.pone.0169719.

[119] T. Toyohara, T. Suzuki, R. Morimoto, Y. Akiyama, T. Souma, H.O. Shiwaku, Y. Takeuchi, E. Mishima, M. Abe, M. Tanemoto, S. Masuda, H. Kawano, K. Maemura, M. Nakayama, H. Sato, T. Mikkaichi, H. Yamaguchi, S. Fukui, Y. Fukumoto, H. Shimokawa, K. Inui, T. Terasaki, J. Goto, S. Ito, T. Hishinuma, I. Rubera, M. Tauc, Y. Fujii-Kuriyama, H. Yabuuchi, Y. Moriyama, T. Soga, T. Abe. SLCO4C1 transporter eliminates uremic toxins and attenuates hypertension and renal inflammation. *J Am Soc Nephrol* **20** (2009) 2546-2555. https://doi.org/10.1681/ASN.2009070696.

[120] H.I. Swanson. Drug Metabolism by the Host and Gut Microbiota: A Partnership or Rivalry? *Drug Metabolism and Disposition* **43** (2015) 1499-1504. https://doi.org/10.1124/dmd.115.065714.

[121] R.G. Nichols, J.M. Peters, A.D. Patterson. Interplay Between the Host, the Human Microbiome, and Drug Metabolism. *Human Genomics* **13** (2019). https://doi.org/10.1186/s40246-019-0211-9.

[122] S.K. Nigam, K.T. Bush, V. Bhatnagar, S.M. Poloyac, J.D. Momper. The Systems Biology of Drug Metabolizing Enzymes and Transporters: Relevance to Quantitative Systems Pharmacology. *Clinical Pharmacology & Therapeutics* **108** (2020) 40-53. https://doi.org/10.1002/cpt.1818.

[123] E.A. Kennedy, K.Y. King, M.T. Baldridge. Mouse Microbiota Models: Comparing Germ-Free Mice and Antibiotics Treatment as Tools for Modifying Gut Bacteria. *Frontiers in Physiology* **9** (2018). https://doi.org/10.3389/fphys.2018.01534.

[124] M. Rosenbaum, R. Knight, R.L. Leibel. The gut microbiota in human energy homeostasis and obesity. *Trends Endocrinol Metab* **26** (2015) 493-501. https://doi.org/10.1016/j.tem.2015.07.002.

[125] J. Vriend, C.A. Hoogstraten, K.R. Venrooij, B.T. van den Berge, L.P. Govers, A. van Rooij, M. Huigen, T.J.J. Schirris, F.G.M. Russel, R. Masereeuw, M.J. Wilmer. Organic anion transporters 1 and 3 influence cellular energy metabolism in renal proximal tubule cells. *Biol Chem* **400** (2019) 1347-1358. https://doi.org/10.1515/hsz-2018-0446.

[126] P. Tirelle, J. Breton, G. Riou, P. Dechelotte, M. Coeffier, D. Ribet. Comparison of different modes of antibiotic delivery on gut microbiota depletion efficiency and body composition in mouse. *BMC Microbiol* **20** (2020) 340. https://doi.org/10.1186/s12866-020-02018-9.

[127] N. Shikano, Y. Kanai, K. Kawai, N. Ishikawa, H. Endou. Transport of 99mTc-MAG3 via rat renal organic anion transporter 1. *J Nucl Med* **45** (2004) 80-85.

[128] N. Takahara, T. Saga, M. Inubushi, H. Kusuhara, C. Seki, S. Ito, N. Oyama, O. Yokoyama, Y. Sugiyama, Y. Fujibayashi. Drugs interacting with organic anion transporter-1 affect uptake of Tc-99m-mercaptoacetyl-triglycine (MAG3) in the human kidney: therapeutic drug interaction in Tc-99m-MAG3 diagnosis of renal function and possible application of Tc-99m-MAG3 for drug development. *Nucl Med Biol* **40** (2013) 643-650. https://doi.org/10.1016/j.nucmedbio.2013.03.006.

[129] C.H. Hsueh, K. Yoshida, P. Zhao, T.W. Meyer, L. Zhang, S.M. Huang, K.M. Giacomini. Identification and Quantitative Assessment of Uremic Solutes as Inhibitors of Renal Organic Anion Transporters, OAT1 and OAT3. *Molecular Pharmaceutics* **13** (2016) 3130-3140. https://doi.org/10.1021/acs.molpharmaceut.6b00332.

[130] H. Shen, V.K. Holenarsipur, T.T. Mariappan, D.M. Drexler, J.L. Cantone, P. Rajanna, S. Singh Gautam, Y. Zhang, J. Gan, P.A. Shipkova, P. Marathe, W.G. Humphreys. Evidence for the Validity of Pyridoxic Acid (PDA) as a Plasma-Based Endogenous Probe for OAT1 and OAT3 Function in Healthy Subjects. *J Pharmacol Exp Ther* **368** (2019) 136-145. https://doi.org/10.1124/jpet.118.252643.

[131] A. Visconti, C.I. Le Roy, F. Rosa, N. Rossi, T.C. Martin, R.P. Mohney, W.Z. Li, E. de Rinaldis, J.T. Bell, J.C. Venter, K.E. Nelson, T.D. Spector, M. Falchi. Interplay between the human gut microbiome and host metabolism. *Nature Communications* **10** (2019). https://doi.org/10.1038/s41467-019-12476-z.

[132] Y. Yu, E.C. Villalobos-Hernandez, S. Pradhananga, C.C. Baker, C. Keating, D. Grundy, A.E. Lomax, D.E. Reed. Deoxycholic acid activates colonic afferent nerves via 5-HT3 receptor-dependent and -independent mechanisms. *Am J Physiol Gastrointest Liver Physiol* **317** (2019) G275-G284. https://doi.org/10.1152/ajpgi.00016.2019.

[133] X. Liu, B. Chen, W. You, S. Xue, H. Qin, H. Jiang. The membrane bile acid receptor TGR5 drives cell growth and migration via activation of the JAK2/STAT3 signaling pathway in non-small cell lung cancer. *Cancer Lett* **412** (2018) 194-207. https://doi.org/10.1016/j.canlet.2017.10.017.

[134] D.D. Jensen, C.B. Godfrey, C. Niklas, M. Canals, M. Kocan, D.P. Poole, J.E. Murphy, F. Alemi, G.S. Cottrell, C. Korbmacher, N.A. Lambert, N.W. Bunnett, C.U. Corvera. The bile acid receptor TGR5 does not interact with beta-arrestins or traffic to endosomes but transmits sustained signals from plasma membrane rafts. *J Biol Chem* **288** (2013) 22942-22960. https://doi.org/10.1074/jbc.M113.455774.

[135] M.S.P. Correia, A. Jain, W. Alotaibi, P. Young Tie Yang, A. Rodriguez-Mateos, D. Globisch. Comparative dietary sulfated metabolome analysis reveals unknown metabolic interactions of the gut microbiome and the human host. *Free Radic Biol Med* **160** (2020) 745-754. https://doi.org/10.1016/j.freeradbiomed.2020.09.006.

[136] N. Koppel, V. Maini Rekdal, E.P. Balskus. Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356** (2017). https://doi.org/10.1126/science.aag2770.

[137] J.C. Granados, V. Bhatnagar, S.K. Nigam. Blockade of organic anion transport in humans after treatment with the drug probenecid leads to major metabolic alterations in plasma and urine. *Clin Pharmacol Ther* (2022). https://doi.org/10.1002/cpt.2630.

[138] B.D. Needham, M.D. Adame, G. Serena, D.R. Rose, G.M. Preston, M.C. Conrad, A.S. Campbell, D.H. Donabedian, A. Fasano, P. Ashwood, S.K. Mazmanian. Plasma and Fecal Metabolite Profiles in Autism Spectrum Disorder. *Biol Psychiatry* **89** (2021) 451-462. https://doi.org/10.1016/j.biopsych.2020.09.025.

[139] J. Lowenstein, S.K. Nigam. Uremic Toxins in Organ Crosstalk. *Front Med (Lausanne)* **8** (2021) 592602. https://doi.org/10.3389/fmed.2021.592602.

[140] S.K. Nigam, W. Wu, K.T. Bush, M.P. Hoenig, R.C. Blantz, V. Bhatnagar. Handling of Drugs, Metabolites, and Uremic Toxins by Kidney Proximal Tubule Drug Transporters. *Clin J Am Soc Nephrol* **10** (2015) 2039-2049. https://doi.org/10.2215/CJN.02440314.

[141] J.K. Tan, C. McKenzie, E. Marino, L. Macia, C.R. Mackay. Metabolite-Sensing G Protein-Coupled Receptors-Facilitators of Diet-Related Immune Regulation. *Annu Rev Immunol* **35** (2017) 371-402. https://doi.org/10.1146/annurev-immunol-051116-052235.

[142] D.A. Colosimo, J.A. Kohn, P.M. Luo, F.J. Piscotta, S.M. Han, A.J. Pickard, A. Rao, J.R. Cross, L.J. Cohen, S.F. Brady. Mapping Interactions of Microbial Metabolites with Human G-Protein-Coupled Receptors. *Cell Host Microbe* **26** (2019) 273-282 e277. https://doi.org/10.1016/j.chom.2019.07.002.

[143] A.L. Graboski, M.R. Redinbo. Gut-Derived Protein-Bound Uremic Toxins. *Toxins* **12** (2020). https://doi.org/10.3390/toxins12090590.

[144] J. Jansen, K. Jansen, E. Neven, R. Poesen, A. Othman, A. van Mil, J. Sluijter, J.S. Torano, E.A. Zaal, C.R. Berkers, D. Esser, H.J. Wichers, K. van Ede, M. van Duursen, S. Burtey, M.C. Verhaar, B. Meijers, R. Masereeuw. Remote sensing and signaling in kidney proximal tubules stimulates gut microbiome-derived organic anion secretion. *Proceedings of the National Academy of Sciences of the United States of America* **116** (2019) 16105-16110. https://doi.org/10.1073/pnas.1821809116.

[145] J.C. Schroeder, B.C. DiNatale, I.A. Murray, C.A. Flaveny, Q. Liu, E.M. Laurenzana, J.M. Lin, S.C. Strom, C.J. Omiecinski, S. Amin, G.H. Perdew. The Uremic Toxin 3-Indoxyl Sulfate Is a Potent Endogenous Agonist for the Human Aryl Hydrocarbon Receptor. *Biochemistry* **49** (2010) 393-400. https://doi.org/10.1021/bi901786x.

[146] H. Yasuda, S. Hirata, K. Inoue, H. Mashima, H. Ohnishi, M. Yoshiba. Involvement of membrane-type bile acid receptor M-BAR/TGR5 in bile acid-induced activation of epidermal

growth factor receptor and mitogen-activated protein kinases in gastric carcinoma cells. *Biochem Biophys Res Commun* **354** (2007) 154-159. https://doi.org/10.1016/j.bbrc.2006.12.168.

[147] N. Jamshidi, S.K. Nigam. Drug transporters OAT1 and OAT3 have specific effects on multiple organs and gut microbiome as revealed by contextualized metabolic network reconstructions. *Sci Rep* **12** (2022) 18308. https://doi.org/10.1038/s41598-022-21091-w.

[148] G. Clarke, K.V. Sandhu, B.T. Griffin, T.G. Dinan, J.F. Cryan, N.P. Hyland. Gut Reactions: Breaking Down Xenobiotic-Microbiome Interactions. *Pharmacol Rev* **71** (2019) 198-224. https://doi.org/10.1124/pr.118.015768.

[149] A.J. Prokopienko, T.D. Nolin. Microbiota-derived uremic retention solutes: perpetrators of altered nonrenal drug clearance in kidney disease. *Expert Rev Clin Pharmacol* **11** (2018) 71-82. https://doi.org/10.1080/17512433.2018.1378095.

[150] F.P. Selwyn, J.Y. Cui, C.D. Klaassen. RNA-Seq Quantification of Hepatic Drug Processing Genes in Germ-Free Mice. *Drug Metabolism and Disposition* **43** (2015) 1572-1580. https://doi.org/10.1124/dmd.115.063545.

[151] J.C. Granados, K. Falah, I. Koo, E.W. Morgan, G.H. Perdew, A.D. Patterson, N. Jamshidi, S.K. Nigam. AHR is a master regulator of diverse pathways in endogenous metabolism. *Sci Rep* **12** (2022) 16625. https://doi.org/10.1038/s41598-022-20572-2.

[152] A. Gonzalez, J.A. Navas-Molina, T. Kosciolek, D. McDonald, Y. Vazquez-Baeza, G. Ackermann, J. DeReus, S. Janssen, A.D. Swafford, S.B. Orchanian, J.G. Sanders, J. Shorenstein, H. Holste, S. Petrus, A. Robbins-Pianka, C.J. Brislawn, M. Wang, J.R. Rideout, E. Bolyen, M. Dillon, J.G. Caporaso, P.C. Dorrestein, R. Knight. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* **15** (2018) 796-798. https://doi.org/10.1038/s41592-018-0141-9.

[153] S.J. Robertson, J.Y. Zhou, K. Geddes, S.J. Rubino, J.H. Cho, S.E. Girardin, D.J. Philpott. Nod1 and Nod2 signaling does not alter the composition of intestinal bacterial communities at homeostasis. *Gut Microbes* **4** (2013) 222-231. https://doi.org/10.4161/gmic.24373.

[154] G. Kaler, D.M. Truong, A. Khandelwal, M. Nagle, S.A. Eraly, P.W. Swaan, S.K. Nigam. Structural variation governs substrate specificity for organic anion transporter (OAT) homologs. Potential remote sensing by OAT family members. *J Biol Chem* **282** (2007) 23841-23853. https://doi.org/10.1074/jbc.M703467200.

[155] D. Butler. Wartime tactic doubles power of scarce bird-flu drug. *Nature* **438** (2005) 6. https://doi.org/10.1038/438006a.

[156] E. Pea. Pharmacology of drugs for hyperuricemia - Mechanisms, kinetics and interactions. *Hyperuricemic Syndromes: Pathophysiology and Therapy* **147** (2005) 35-46.

[157] K. Pui, P.J. Gow, N. Dalbeth. Efficacy and Tolerability of Probenecid as Urate-lowering Therapy in Gout; Clinical Experience in High-prevalence Population. *Journal of Rheumatology* **40** (2013) 872-876. https://doi.org/10.3899/jrheum.121301.

[158] S.K. Nigam, V. Bhatnagar. The systems biology of uric acid transporters: the role of remote sensing and signaling. *Current Opinion in Nephrology and Hypertension* **27** (2018) 305-313. https://doi.org/10.1097/Mnh.0000000000000427.

[159] A. So, B. Thorens. Uric acid transport and disease. *J Clin Invest* **120** (2010) 1791-1799. https://doi.org/10.1172/JCI42344.

[160] G. Capasso, P. Jaeger, W.G. Robertson, R.J. Unwin. Uric acid and the kidney: urate transport, stone disease and progressive renal failure. *Curr Pharm Des* **11** (2005) 4153-4159. https://doi.org/10.2174/138161205774913219.

[161] Y. Sato, D.I. Feig, A.G. Stack, D.H. Kang, M.A. Lanaspa, A.A. Ejaz, L.G. Sanchez-Lozada, M. Kuwabara, C. Borghi, R.J. Johnson. The case for uric acid-lowering treatment in patients with hyperuricaemia and CKD. *Nat Rev Nephrol* **15** (2019) 767-775. https://doi.org/10.1038/s41581-019-0174-z.

[162] G. Ragab, M. Elshahaly, T. Bardin. Gout: An old disease in new perspective - A review. *J Adv Res* **8** (2017) 495-511. https://doi.org/10.1016/j.jare.2017.04.008.

[163] R.F. Cunningham, Z.H. Israili, P.G. Dayton. Clinical pharmacokinetics of probenecid. *Clin Pharmacokinet* **6** (1981) 135-151. https://doi.org/10.2165/00003088-198106020-00004.

[164] S.K. Nigam. The SLC22 Transporter Family: A Paradigm for the Impact of Drug Transporters on Metabolic Pathways, Signaling, and Disease. *Annual Review of Pharmacology and Toxicology, Vol 58* **58** (2018) 663-687. https://doi.org/10.1146/annurev-pharmtox-010617-052713.

[165] P.K. Tan, T.M. Ostertag, J.N. Miner. Mechanism of high affinity inhibition of the human urate transporter URAT1. *Sci Rep* **6** (2016) 34995. https://doi.org/10.1038/srep34995.

[166] H.J. Shin, M. Takeda, A. Enomoto, M. Fujimura, H. Miyazaki, N. Anzai, H. Endou. Interactions of urate transporter URAT1 in human kidney with uricosuric drugs. *Nephrology (Carlton)* **16** (2011) 156-162. https://doi.org/10.1111/j.1440-1797.2010.01368.x.

[167] D.C. Engelhart, J.C. Granados, D. Shi, M.H. Saier Jr, Jr., M.E. Baker, R. Abagyan, S.K. Nigam. Systems Biology Analysis Reveals Eight SLC22 Transporter Subgroups, Including OATs, OCTs, and OCTNs. *Int J Mol Sci* **21** (2020). https://doi.org/10.3390/ijms21051791.

[168] J. Konig, F. Muller, M.F. Fromm. Transporters and drug-drug interactions: important determinants of drug disposition and effects. *Pharmacol Rev* **65** (2013) 944-966. https://doi.org/10.1124/pr.113.007518.

[169] R.E. Lai, C.E. Jay, D.H. Sweet. Organic solute carrier 22 (SLC22) family: Potential for interactions with food, herbal/dietary supplements, endogenous compounds, and drugs. *Journal of Food and Drug Analysis* **26** (2018) S45-S60. https://doi.org/10.1016/j.jfda.2018.03.002.

[170] A.Z. Meheus, V. Mubiligi, G. Vandenberghe. Treatment of gonorrhoea with a combination of probenecid and procaine penicillin in Rwanda. *Afr J Med Med Sci* **5** (1976) 209-212.

[171] A. Bahn, M. Ljubojevic, H. Lorenz, C. Schultz, E. Ghebremedhin, B. Ugele, I. Sabolic, G. Burckhardt, Y. Hagos. Murine renal organic anion transporters mOAT1 and mOAT3 facilitate the transport of neuroactive tryptophan metabolites. *Am J Physiol Cell Physiol* **289** (2005) C1075-1084. https://doi.org/10.1152/ajpcell.00619.2004.

[172] V. Vallon, S.A. Eraly, W.R. Wikoff, T. Rieg, G. Kaler, D.M. Truong, S.Y. Ahn, N.R. Mahapatra, S.K. Mahata, J.A. Gangoiti, W. Wu, B.A. Barshop, G. Siuzdak, S.K. Nigam. Organic anion transporter 3 contributes to the regulation of blood pressure. *Journal of the American Society of Nephrology* **19** (2008) 1732-1740. https://doi.org/10.1681/Asn.2008020180.

[173] K.L. Mertens, A. Kalsbeek, M.R. Soeters, H.M. Eggink. Bile Acid Signaling Pathways from the Enterohepatic Circulation to the Central Nervous System. *Front Neurosci* **11** (2017) 617. https://doi.org/10.3389/fnins.2017.00617.

[174] H.M. Roager, T.R. Licht. Microbial tryptophan catabolites in health and disease. *Nature Communications* **9** (2018). https://doi.org/10.1038/s41467-018-05470-4.

[175] T.W. Pols, L.G. Noriega, M. Nomura, J. Auwerx, K. Schoonjans. The bile acid membrane receptor TGR5 as an emerging target in metabolism and inflammation. *J Hepatol* **54** (2011) 1263-1272. https://doi.org/10.1016/j.jhep.2010.12.004.

[176] G. Rizzo, B. Renga, A. Mencarelli, R. Pellicciari, S. Fiorucci. Role of FXR in regulating bile acid homeostasis and relevance for human diseases. *Curr Drug Targets Immune Endocr Metabol Disord* **5** (2005) 289-303. https://doi.org/10.2174/1568008054863781.

[177] H. Deng, H. Hu, Y. Fang. Multiple tyrosine metabolites are GPR35 agonists. *Sci Rep* **2** (2012) 373. https://doi.org/10.1038/srep00373.

[178] I. Cervenka, L.Z. Agudelo, J.L. Ruas. Kynurenines: Tryptophan's metabolites in exercise, inflammation, and mental health. *Science* **357** (2017). https://doi.org/10.1126/science.aaf9794.

[179] Y. Zhang, V.K. Holenarsipur, H. Kandoussi, J. Zeng, T.T. Mariappan, M. Sinz, H. Shen. Detection of Weak Organic Anion-Transporting Polypeptide 1B Inhibition by Probenecid with Plasma-Based Coproporphyrin in Humans. *Drug Metab Dispos* **48** (2020) 841-848. https://doi.org/10.1124/dmd.120.000076.

[180] P. Hemmersbach. The Probenecid-story - A success in the fight against doping through out-of-competition testing. *Drug Test Anal* **12** (2020) 589-594. https://doi.org/10.1002/dta.2727.

[181] M.E. Willemin, T.K. Van der Made, I. Pijpers, L. Dillen, A. Kunze, S. Jonkers, K. Steemans, A. Tuytelaars, F. Jacobs, M. Monshouwer, D. Scotcher, A. Rostami-Hodjegan, A. Galetin, J. Snoeys. Clinical Investigation on Endogenous Biomarkers to Predict Strong OAT-Mediated Drug-Drug Interactions. *Clinical Pharmacokinetics* **60** (2021) 1187-1199. https://doi.org/10.1007/s40262-021-01004-2.

[182] J. Tang, H. Shen, X. Zhao, V.K. Holenarsipur, T.T. Mariappan, Y. Zhang, E. Panfen, J. Zheng, W.G. Humphreys, Y. Lai. Endogenous Plasma Kynurenic Acid in Human: A Newly Discovered Biomarker for Drug-Drug Interactions Involving Organic Anion Transporter 1 and 3 Inhibition. *Drug Metab Dispos* **49** (2021) 1063-1069. https://doi.org/10.1124/dmd.121.000486.
[183] A.L. Bullen, P.S. Garimella. Beyond the Glomerulus-Kidney Tubule Markers and Diabetic Kidney Disease Progression. *Kidney Int Rep* **6** (2021) 1200-1202. https://doi.org/10.1016/j.ekir.2021.03.879.

[184] S.J. Shi, Y.Q. Li. Interplay of Drug-Metabolizing Enzymes and Transporters in Drug Absorption and Disposition. *Current Drug Metabolism* **15** (2014) 915-941.

[185] E.R. Feeney, P.W. Mallon. HIV and HAART-Associated Dyslipidemia. *Open Cardiovasc Med J* **5** (2011) 49-63. https://doi.org/10.2174/1874192401105010049.

[186] S.B. Meropol, K.A. Chan, Z. Chen, J.A. Finkelstein, S. Hennessy, E. Lautenbach, R. Platt, S.D. Schech, D. Shatin, J.P. Metlay. Adverse events associated with prolonged antibiotic use. *Pharmacoepidemiol Drug Saf* **17** (2008) 523-532. https://doi.org/10.1002/pds.1547.

[187] Z.A. Marcum, J.T. Hanlon. Recognizing the Risks of Chronic Nonsteroidal Anti-Inflammatory Drug Use in Older Adults. *Ann Longterm Care* **18** (2010) 24-27.

[188] A. Armenta, M. Madero, B. Rodriguez-Iturbe. Functional Reserve of the Kidney. *Clin J Am Soc Nephrol* **17** (2022) 458-466. https://doi.org/10.2215/CJN.11070821.

[189] P.M. Visscher, N.R. Wray, Q. Zhang, P. Sklar, M.I. McCarthy, M.A. Brown, J. Yang. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am J Hum Genet* **101** (2017) 5-22. https://doi.org/10.1016/j.ajhg.2017.06.005.

[190] C. Kooperberg, M. LeBlanc, V. Obenchain. Risk Prediction Using Genome-Wide Association Studies. *Genetic Epidemiology* **34** (2010) 643-652. https://doi.org/10.1002/gepi.20509.

[191] V. Tam, N. Patel, M. Turcotte, Y. Bosse, G. Pare, D. Meyre. Benefits and limitations of genome-wide association studies. *Nature Reviews Genetics* **20** (2019) 467-484. https://doi.org/10.1038/s41576-019-0127-1.

[192] G. Breen, Q. Li, B.L. Roth, P. O'Donnell, M. Didriksen, R. Dolmetsch, P.F. O'Reilly, H.A. Gaspar, H. Manji, C. Huebel, J.R. Kelsoe, D. Malhotra, A. Bertolino, D. Posthuma, P. Sklar, S. Kapur, P.F. Sullivan, D.A. Collier, H.J. Edenberg. Translating genome-wide association findings

into new therapeutics for psychiatry. *Nat Neurosci* **19** (2016) 1392-1396. https://doi.org/10.1038/nn.4411.

[193] M.R. van der Sijde, A. Ng, J.Y. Fu. Systems genetics: From GWAS to disease pathways. *Biochimica Et Biophysica Acta-Molecular Basis of Disease* **1842** (2014) 1903-1909. https://doi.org/10.1016/j.bbadis.2014.04.025.

[194] E.P. Rhee, J.E. Ho, M.H. Chen, D.X. Shen, S. Cheng, M.G. Larson, A. Ghorbani, X. Shi, I.T. Helenius, C.J. O'Donnell, A.L. Souza, A. Deik, K.A. Pierce, K. Bullock, G.A. Walford, R.S. Vasan, J.C. Florez, C. Clish, J.R.J. Yeh, T.J. Wang, R.E. Gerszten. A Genome-wide Association Study of the Human Metabolome in a Community-Based Cohort. *Cell Metabolism* **18** (2013) 130-143. https://doi.org/10.1016/j.cmet.2013.06.013.

[195] K. Suhre, M. Arnold, A.M. Bhagwat, R.J. Cotton, R. Engelke, J. Raffler, H. Sarwath, G. Thareja, A. Wahl, R.K. DeLisle, L. Gold, M. Pezer, G. Lauc, M.A.E. Selim, D.O. Mook-Kanamori, E.K. Al-Dous, Y.A. Mohamoud, J. Malek, K. Strauch, H. Grallert, A. Peters, G. Kastenmuller, C. Gieger, J. Graumann. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature Communications* **8** (2017). https://doi.org/10.1038/ncomms14357.

[196] J. Raffler, N. Friedrich, M. Arnold, T. Kacprowski, R. Rueedi, E. Altmaier, S. Bergmann, K. Budde, C. Gieger, G. Homuth, M. Pietzner, W. Romisch-Margl, K. Strauch, H. Volzke, M. Waldenberger, H. Wallaschofski, M. Nauck, U. Volker, G. Kastenmuller, K. Suhre. Genome-Wide Association Study with Targeted and Non-targeted NMR Metabolomics Identifies 15 Novel Loci of Urinary Human Metabolic Individuality. *Plos Genetics* **11** (2015). https://doi.org/10.1371/journal.pgen.1005487.

[197] S.W. Yee, K.M. Giacomini. Emerging Roles of the Human Solute Carrier 22 Family. *Drug Metab Dispos* (2021). https://doi.org/10.1124/dmd.121.000702.

[198] J. Li, M.H. Bluth. Pharmacogenomics of drug metabolizing enzymes and transporters: implications for cancer therapy. *Pharmacogenomics & Personalized Medicine* **4** (2011) 11-33. https://doi.org/10.2147/Pgpm.S18861.

[199] S. Ahmed, Z. Zhou, J. Zhou, S.Q. Chen. Pharmacogenomics of Drug Metabolizing Enzymes and Transporters: Relevance to Precision Medicine. *Genomics Proteomics Bioinformatics* **14** (2016) 298-313. https://doi.org/10.1016/j.gpb.2016.03.008.

[200] C.J. Smith, N. Sinnott-Armstrong, A. Cichonska, H. Julkunen, E.B. Fauman, P. Wurtz, J.K. Pritchard. Integrative analysis of metabolite GWAS illuminates the molecular basis of pleiotropy and genetic correlation. *Elife* **11** (2022). https://doi.org/10.7554/eLife.79348.

[201] X. Yin, L.S. Chan, D. Bose, A.U. Jackson, P. VandeHaar, A.E. Locke, C. Fuchsberger, H.M. Stringham, R. Welch, K. Yu, L. Fernandes Silva, S.K. Service, D. Zhang, E.C. Hector, E. Young, L. Ganel, I. Das, H. Abel, M.R. Erdos, L.L. Bonnycastle, J. Kuusisto, N.O. Stitziel, I.M. Hall, G.R. Wagner, FinnGen, J. Kang, J. Morrison, C.F. Burant, F.S. Collins, S. Ripatti, A.

Palotie, N.B. Freimer, K.L. Mohlke, L.J. Scott, X. Wen, E.B. Fauman, M. Laakso, M. Boehnke. Genome-wide association studies of metabolites in Finnish men identify disease-relevant loci. *Nat Commun* **13** (2022) 1644. https://doi.org/10.1038/s41467-022-29143-5.

[202] R. Sever, C.K. Glass. Signaling by Nuclear Receptors. *Cold Spring Harbor Perspectives in Biology* **5** (2013). https://doi.org/10.1101/cshperspect.a016709.

[203] K.A. Lagerborg, J.D. Watrous, S. Cheng, M. Jain. High-Throughput Measure of Bioactive Lipids Using Non-targeted Mass Spectrometry. *Methods Mol Biol* **1862** (2019) 17-35. https://doi.org/10.1007/978-1-4939-8769-6_2.

[204] J.D. Watrous, T.J. Niiranen, K.A. Lagerborg, M. Henglin, Y.J. Xu, J. Rong, S. Sharma, R.S. Vasan, M.G. Larson, A. Armando, S. Mora, O. Quehenberger, E.A. Dennis, S. Cheng, M. Jain. Directed Non-targeted Mass Spectrometry and Chemical Networking for Discovery of Eicosanoids and Related Oxylipins. *Cell Chem Biol* **26** (2019) 433-442 e434. https://doi.org/10.1016/j.chembiol.2018.11.015.

[205] C.W. Tsao, R.S. Vasan. Cohort Profile: The Framingham Heart Study (FHS): overview of milestones in cardiovascular epidemiology. *Int J Epidemiol* **44** (2015) 1800-1813. https://doi.org/10.1093/ije/dyv337.

[206] P. Cingolani, A. Platts, L. Wang le, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, D.M. Ruden. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6** (2012) 80-92. https://doi.org/10.4161/fly.19695.

[207] P.R. Loh, G. Tucker, B.K. Bulik-Sullivan, B.J. Vilhjalmsson, H.K. Finucane, R.M. Salem, D.I. Chasman, P.M. Ridker, B.M. Neale, B. Berger, N. Patterson, A.L. Price. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nature Genetics* **47** (2015) 284-+. https://doi.org/10.1038/ng.3190.

[208] K. Watanabe, E. Taskesen, A. van Bochoven, D. Posthuma. Functional mapping and annotation of genetic associations with FUMA. *Nature Communications* **8** (2017). https://doi.org/10.1038/s41467-017-01261-5.

[209] E.A. Khramtsova, B.E. Stranger. Assocplots: a Python package for static and interactive visualization of multiple-group GWAS results. *Bioinformatics* **33** (2017) 432-434. https://doi.org/10.1093/bioinformatics/btw641.

[210] A. Jain, G. Tuteja. TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics* **35** (2019) 1966-1967. https://doi.org/10.1093/bioinformatics/bty890.

[211] M. Whirl-Carrillo, E.M. McDonagh, J.M. Hebert, L. Gong, K. Sangkuhl, C.F. Thorn, R.B. Altman, T.E. Klein. Pharmacogenomics Knowledge for Personalized Medicine. *Clinical Pharmacology & Therapeutics* **92** (2012) 414-417. https://doi.org/10.1038/clpt.2012.96.

[212] F.M. Sladek. What are nuclear receptor ligands? *Mol Cell Endocrinol* **334** (2011) 3-13. https://doi.org/10.1016/j.mce.2010.06.018.

[213] D. Wacker, R.C. Stevens, B.L. Roth. How Ligands Illuminate GPCR Molecular Pharmacology. *Cell* **170** (2017) 414-427. https://doi.org/10.1016/j.cell.2017.07.009.

[214] S.T. Sherry, M.H. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, K. Sirotkin. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29** (2001) 308-311. https://doi.org/10.1093/nar/29.1.308.

[215] A.D. Joshi, C. Andersson, S. Buch, S. Stender, R. Noordam, L.C. Weng, P.E. Weeke, P.L. Auer, B. Boehm, C. Chen, H. Choi, G. Curhan, J.C. Denny, I. De Vivo, J.D. Eicher, D. Ellinghaus, A.R. Folsom, C. Fuchs, M. Gala, J. Haessler, A. Hofman, F. Hu, D.J. Hunter, H.L. Janssen, J.H. Kang, C. Kooperberg, P. Kraft, W. Kratzer, W. Lieb, P.L. Lutsey, S. Darwish Murad, B.G. Nordestgaard, L.R. Pasquale, A.P. Reiner, P.M. Ridker, E. Rimm, L.M. Rose, C.M. Shaffer, C. Schafmayer, R.M. Tamimi, A.G. Uitterlinden, U. Volker, H. Volzke, Y. Wakabayashi, J.L. Wiggs, J. Zhu, D.M. Roden, B.H. Stricker, W. Tang, A. Teumer, J. Hampe, A. Tybjaerg-Hansen, D.I. Chasman, A.T. Chan, A.D. Johnson. Four Susceptibility Loci for Gallstone Disease Identified in a Meta-analysis of Genome-Wide Association Studies. *Gastroenterology* **151** (2016) 351-363 e328. https://doi.org/10.1053/j.gastro.2016.04.007.

[216] S.W. Yee, M.M. Giacomini, C.H. Hsueh, D. Weitz, X. Liang, S. Goswami, J.M. Kinchen, A. Coelho, A.A. Zur, K. Mertsch, W. Brian, D.L. Kroetz, K.M. Giacomini. Metabolomic and Genome-wide Association Studies Reveal Potential Endogenous Biomarkers for OATP1B1. *Clin Pharmacol Ther* **100** (2016) 524-536. https://doi.org/10.1002/cpt.434.

[217] S. Yee, A. Stecula, H. Chien, L. Zou, E. Feofanova, M. van Borselen, V. Artegoitia, J. Newman, K. Cheung, N. Yousri, K. Suhre, R. Irannejad, B. Yu, K. Giacomini, J. Kinchen, X. Zhou. Unraveling the Functional Role of Orphan Transporter Slc22a24 as an Anion Transporter Selective for Steroid Conjugates. *Clinical Pharmacology & Therapeutics* **107** (2020) S118-S118.

[218] A. Galetin, K. Ito, D. Hallifax, J.B. Houston. CYP3A4 substrate selection and substitution in the prediction of potential drug-drug interactions. *J Pharmacol Exp Ther* **314** (2005) 180-190. https://doi.org/10.1124/jpet.104.082826.

[219] A. Kalsotra, H.W. Strobel. Cytochrome P4504F subfamily: At the crossroads of eicosanoid and drug metabolism. *Pharmacology & Therapeutics* **112** (2006) 589-611. https://doi.org/10.1016/j.pharmthera.2006.03.008.

[220] M. Roth, A. Obaidat, B. Hagenbuch. OATPs, OATs and OCTs: the organic anion and cation transporters of the SLCO and SLC22A gene superfamilies. *British Journal of Pharmacology* **165** (2012) 1260-1287. https://doi.org/10.1111/j.1476-5381.2011.01724.x.

[221] M.A. Demel, R. Schwaha, O. Kramer, P. Ettmayer, E.E. Haaksma, G.F. Ecker. In silico prediction of substrate properties for ABC-multidrug transporters. *Expert Opin Drug Metab Toxicol* **4** (2008) 1167-1180. https://doi.org/10.1517/17425255.4.9.1167.

[222] M. Hohne, S. Schatzle, H. Jochens, K. Robins, U.T. Bornscheuer. Rational assignment of key motifs for function guides in silico enzyme identification. *Nat Chem Biol* **6** (2010) 807-813. https://doi.org/10.1038/nchembio.447.

[223] M.S. Yu, H.M. Lee, A. Park, C. Park, H. Ceong, K.H. Rhee, D. Na. In silico prediction of potential chemical reactions mediated by human enzymes. *BMC Bioinformatics* **19** (2018) 207. https://doi.org/10.1186/s12859-018-2194-2.

[224] Y.Z. Guo. Be Cautious with Crystal Structures of Membrane Proteins or Complexes Prepared in Detergents. *Crystals* **10** (2020). https://doi.org/10.3390/cryst10020086.

[225] D. Raj, B. Tomar, A. Lahiri, S.R. Mulay. The gut-liver-kidney axis: Novel regulator of fatty liver associated chronic kidney disease. *Pharmacol Res* **152** (2020) 104617. https://doi.org/10.1016/j.phrs.2019.104617.

[226] M. Schwenk. Drug transport in intestine, liver and kidney. *Arch Toxicol* **60** (1987) 37-42. https://doi.org/10.1007/BF00296944.

[227] L. Pochini, M. Galluccio, M. Scalise, L. Console, C. Indiveri. OCTN: A Small Transporter Subfamily with Great Relevance to Human Pathophysiology, Drug Discovery, and Diagnostics. *SLAS Discov* **24** (2019) 89-110. https://doi.org/10.1177/2472555218812821.

[228] P.J. Bosma, J. Seppen, B. Goldhoorn, C. Bakker, R.P.J.O. Elferink, J.R. Chowdhury, N.R. Chowdhury, P.L.M. Jansen. Bilirubin Udp-Glucuronosyltransferase-1 Is the Only Relevant Bilirubin Glucuronidating Isoform in Man. *Journal of Biological Chemistry* **269** (1994) 17960-17964.

[229] K.N. Theken, Y.M. Deng, M.A. Kannon, T.M. Miller, S.M. Poloyac, C.R. Lee. Activation of the Acute Inflammatory Response Alters Cytochrome P450 Expression and Eicosanoid Metabolism. *Drug Metabolism and Disposition* **39** (2011) 22-29. https://doi.org/10.1124/dmd.110.035287.

[230] N. Sehgal, V. Agarwal, R.K. Valli, S.D. Joshi, L. Antonovic, H.W. Strobel, V. Ravindranath. Cytochrome P4504f, a potential therapeutic target limiting neuroinflammation. *Biochemical Pharmacology* **82** (2011) 53-64. https://doi.org/10.1016/j.bcp.2011.03.025.

[231] K.Z. Edson, A.E. Rettie. CYP4 Enzymes As Potential Drug Targets: Focus on Enzyme Multiplicity, Inducers and Inhibitors, and Therapeutic Modulation of 20-Hydroxyeicosatetraenoic Acid (20-HETE) Synthase and Fatty Acid omega-Hydroxylase Activities. *Current Topics in Medicinal Chemistry* **13** (2013) 1429-1440.

[232] J.P. Hardwick, K. Eckman, Y.K. Lee, M.A. Abdelmegeed, A. Esterle, W.M. Chilian, J.Y. Chiang, B.J. Song. Eicosanoids in Metabolic Syndrome. *Immunopharmacology* **66** (2013) 157-266. https://doi.org/10.1016/B978-0-12-404717-4.00005-6.

[233] A. Stewart. SLCO1B1 Polymorphisms and Statin-Induced Myopathy. *PLoS Curr* **5** (2013). https://doi.org/10.1371/currents.eogt.d21e7f0c58463571bb0d9d3a19b82203.

[234] C. Oshiro, L. Mangravite, T. Klein, R. Altman. PharmGKB very important pharmacogene: SLCO1B1. *Pharmacogenetics and Genomics* **20** (2010) 211-216. https://doi.org/10.1097/FPC.0b013e328333b99c.

[235] E. Dainese, S. Oddi, M. Simonetti, A. Sabatucci, C.B. Angelucci, A. Ballone, B. Dufrusine, F. Fezza, G. De Fabritiis, M. Maccarrone. The endocannabinoid hydrolase FAAH is an allosteric enzyme. *Scientific Reports* **10** (2020). https://doi.org/10.1038/s41598-020-59120-1.

[236] W. Zheng, S. Hussein, J. Yang, J. Huang, F. Zhang, S. Hernandez-Anzaldo, C. Fernandez-Patron, Y. Cao, H.B. Zeng, J.F. Tang, X.Z. Chen. A novel PKD2L1 C-terminal domain critical for trimerization and channel function. *Scientific Reports* **5** (2015). https://doi.org/10.1038/srep09460.

[237] H.H. Lau, N.H.J. Ng, L.S.W. Loo, J.B. Jasmen, A.K.K. Teo. The molecular functions of hepatocyte nuclear factors - In and beyond the liver. *Journal of Hepatology* **68** (2018) 1033-1048. https://doi.org/10.1016/j.jhep.2017.11.026.

[238] G. Martovetsky, J.B. Tee, S.K. Nigam. Hepatocyte nuclear factors 4alpha and 1alpha regulate kidney developmental expression of drug-metabolizing enzymes and drug transporters. *Mol Pharmacol* **84** (2013) 808-823. https://doi.org/10.1124/mol.113.088229.

[239] S.R. Jaffrey, A.M. Snowman, M.J.L. Eliasson, N.A. Cohen, S.H. Snyder. CAPON: A protein associated with neuronal nitric oxide synthase that regulates its interactions with PSD95. *Neuron* **20** (1998) 115-124. https://doi.org/10.1016/S0896-6273(00)80439-0.

[240] A.K. Daly. Pharmacogenetics: a general review on progress to date. *British Medical Bulletin* **124** (2017) 65-79. https://doi.org/10.1093/bmb/ldx035.

[241] M. Pirmohamed. Pharmacogenomics: Relevance to Personalized Medicine. *Drug Metabolism and Pharmacokinetics* **32** (2017) S10-S10.

[242] R. Govindarajan, A. Sparreboom. Drug Transporters: Advances and Opportunities. *Clin Pharmacol Ther* **100** (2016) 398-403. https://doi.org/10.1002/cpt.454.

[243] J.P. Keogh. Membrane Transporters in Drug Development. *Current Concepts in Drug Metabolism and Toxicology* **63** (2012) 1-42. https://doi.org/10.1016/B978-0-12-398339-8.00001-X.

[244] P. Duan, S.S. Li, N. Ai, L.Q. Hu, W.J. Welsh, G.F. You. Potent Inhibitors of Human Organic Anion Transporters 1 and 3 from Clinical Drug Libraries: Discovery and Molecular Characterization. *Molecular Pharmaceutics* **9** (2012) 3340-3346. https://doi.org/10.1021/mp300365t.

[245] J.L. Perry, N. Dembla-Rajpal, L.A. Hall, J.B. Pritchard. A three-dimensional model of human organic anion transporter 1: aromatic amino acids required for substrate transport. *J Biol Chem* **281** (2006) 38071-38079. https://doi.org/10.1074/jbc.M608834200.

[246] E.J. Cox, A.E. Rettie, J.D. Unadkat, K.E. Thummel, J.S. McCune, M.F. Paine. Adapting regulatory drug-drug interaction guidance to design clinical pharmacokinetic natural product-drug interaction studies: A NaPDI Center recommended approach. *Clin Transl Sci* **15** (2022) 322-329. https://doi.org/10.1111/cts.13172.

[247] T. Hegedus, M. Geisler, G.L. Lukacs, B. Farkas. Ins and outs of AlphaFold2 transmembrane protein structure predictions. *Cell Mol Life Sci* **79** (2022) 73. https://doi.org/10.1007/s00018-021-04112-1.

[248] S. Jo, T. Kim, V.G. Iyer, W. Im. CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* **29** (2008) 1859-1865. https://doi.org/10.1002/jcc.20945.

[249] T. Gryp, K. De Paepe, R. Vanholder, F.M. Kerckhof, W. Van Biesen, T. Van de Wiele, F. Verbeke, M. Speeckaert, M. Joossens, M.M. Couttenye, M. Vaneechoutte, G. Glorieux. Gut microbiota generation of protein-bound uremic toxins and related metabolites is not altered at different stages of chronic kidney disease. *Kidney Int* **97** (2020) 1230-1242. https://doi.org/10.1016/j.kint.2020.01.028.

[250] A. Agus, K. Clement, H. Sokol. Gut microbiota-derived metabolites as central regulators in metabolic disorders. *Gut* **70** (2021) 1174-1182. https://doi.org/10.1136/gutjnl-2020-323071.

[251] J.C. Granados, V. Ermakov, K. Maity, D.R. Vera, G. Chang, S.K. Nigam. The kidney drug transporter OAT1 regulates gut microbiome-dependent host metabolism. *JCI Insight* **8** (2023). https://doi.org/10.1172/jci.insight.160437.

[252] A.M. Torres, A.V. Dnyanmote, J.C. Granados, S.K. Nigam. Renal and non-renal response of ABC and SLC transporters in chronic kidney disease. *Expert Opin Drug Metab Toxicol* (2021). https://doi.org/10.1080/17425255.2021.1899159.