Applications of Genome-Scale Metabolic Network Reconstructions to Characterize Drug-Induced Toxicity

---

A Dissertation

Presented to

the faculty of the School of Engineering and Applied Science

University of Virginia

---

in partial fulfillment
of the requirements for the degree

Doctor of Philosophy

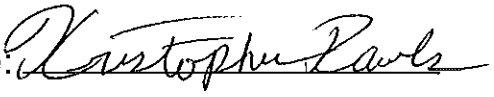
by

Kristopher D Rawls

December 2019

APPROVAL SHEET

This Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author Signature: 

This Dissertation has been read and approved by the examining committee:

Advisor: Jason Papin

Committee Member: Shayn Peirce-Cottler

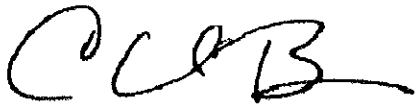
Committee Member: Jeffrey Saucerman

Committee Member: Christopher Holstege

Committee Member: Mete Civelek

Committee Member: _____

Accepted for the School of Engineering and Applied Science:



Craig H. Benson, School of Engineering and Applied Science

December 2019

Applications of genome-scale metabolic network reconstructions for drug-induced toxicity

A Dissertation

presented to the faculty of the School of Engineering and Applied Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

by

Kristopher D. Rawls

December 2019

Department of Biomedical Engineering
University of Virginia

Abstract

Metabolism is the process of deriving energy by breaking down compounds and using that same energy to build new compounds, which is necessary for other biological processes. One way we can model metabolism computationally is with GEnome-Scale metabolic Network REconstructions (GENREs), which use linear algebra to analyze an underdetermined system under steady state assumptions. GENREs allow us to simulate how growth of an organism is affected by its surrounding system, as well as how genetic mutations can inhibit normal function of an organism. In the context of human metabolism, GENREs can be used to evaluate how drugs or diseases can impact the function of healthy cells, and potentially lead to the discovery of novel drug targets. In this dissertation I describe what GENREs are and create a simplified GENRE, *iSIM*, to explain how to make various types of predictions. To make GENREs more accessible to biologists without a programming background, source code is provided on how to run these analyses. For drug-induced hepatotoxicity and nephrotoxicity, I developed a paired transcriptomics data and metabolomics data pipeline to use an existing GENRE, *iRno*, to make predictions using experimental data and create validation data for model predictions from the same biological sample. This approach identifies why metabolite production levels are changed, and can lead to the identification of biomarkers of drug-induced toxicity. Additionally, I predict changes in metabolite level changes with *iRno*, and compared results to metabolomics data and existing literature. This dissertation provides new tools for biologists who want to learn to program, and gives toxicologist and other scientists interested in drug discovery a tool and pipeline to identify potential impacts of new drugs.

Acknowledgments

As a first-generation college student, navigating undergrad and now graduate school has been difficult but fortunately, I was able to succeed with the help of an amazing support system. I would not have gotten through any of it without the help of others so I want to take the time to acknowledge those that have helped me get to this point.

To my family, you all have been the best consistent source of support and I am forever grateful. To my mother, Juanita Williams-Davis, I want to thank you for being there for me since the very beginning. I appreciate all the sacrifices you have made on my behalf, and for encouraging me to chase happiness in life, rather than a specific career or lifestyle. To Rodney Davis Sr., I appreciate you being the father I needed and for helping me grow into the man I am today. I appreciate your love and hard work, to continue to provide for our family, and to help me even when I don't always ask. To my younger brother Rodney Davis Jr., I appreciate the joy and laughter you consistently bring into my life. You serve as my reminder to enjoy life and to take a breath when needed. You guys have been with me throughout my journey through higher education and even before, and have always believed that I could achieve this dream, even when I couldn't sometimes. To my youngest brother Jaheim Rawls I thank you for the love you have given, during our hard times over the past few years.

Thank you to my grandparents and extended family, including aunts, uncles, cousins, etc., for also cheering me on and providing me with the love I needed to make it through life. To the friends outside of UVA that I have met, either through NC State or the National

Society of Black Engineers (NSBE), I thank all of you for helping me experience life outside of graduate school and encouraging me along the way.

There are a few people who I want to thank that unfortunately didn't live to see me finish this journey. My great grandmother Annie-Ruth Evans is someone I spent a lot of time with growing up, and she has always expressed how proud of me she was. I thank you for all you have done for me. To my grandfather Louis Rawls and my uncle Michael Williams, I appreciate your support and for being there with me when I graduated from NC State. I know you would have been here if you were able, and I thank you for all you have done. Lastly to my dad Prince Rawls, I'm thankful you made sure that I knew that you were proud of what I was doing and that you wanted me to keep going in school.

To the host of UVA BME faculty, staff, and students I thank you for your continued support of my development as a scientist, and for the training, I have received thus far. Specifically, I would like to thank Jeff Saucerman, and Jeff Holmes for allowing me to rotate in their labs during my first semester, and for continuing to provide feedback and guidance even after I left their labs. I also want to thank my dissertation committee Shayn Pierce-Cottler, Chris Holstege, Mete Civelek, and Jeff Saucerman for the feedback, and direction you have provided over the course of my Ph.D.

The Papin lab has been an amazing experience and I am honored to have worked with all of my past colleagues. Thank you, Jennie Bartell, Matthew Biggs, Phil Yen, Maureen Carey, Anna Blazier, and Greg Medlock for your candid conversations about your

experiences, feedback on research presentations, and guidance with helping me craft my next steps after graduate school. To the current lab members, Laura Dunphy, Matthew Jenior, and Tom Moutinho thank you for continuing to make the Papin lab a great place to keep working, and for giving feedback on research methods, and the end of my Ph.D. This dissertation would not have been possible without the help and direction of my fellow ratcon members as well. Edik Blais, thank you for laying a strong foundation and for the help you have given me throughout this project. To Bonnie Dougherty, I appreciate your scientific insight and the genuine friendship that you have given. We have worked on almost every project I have published on together and I am thankful for all of your assistance. Glynis, it's amazing how much you have helped me since my first year. I had zero wet lab experience, and you have been so patient and kind with teaching me how to do everything in the lab. Thank you for your mentorship.

Lastly, to my Ph.D. advisor Jason Papin. You have been an amazing mentor and have set the bar very high as far as bosses/mentors go and I truly feel spoiled. I thank you for allowing me to join the lab, showing me kindness, and welcoming me here at UVA. I also appreciate your continued advocacy for me inside and outside of research. You have shown me how to keep the big picture in mind, and how to persist when things aren't working. Thank you for allowing me space to succeed, and learn from my mistakes while also encouraging me to keep going. You've never questioned my ability to make it through this program and I will be forever grateful for you.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Foreword

In this chapter, I discuss the power and application of computational modeling for representing biological systems, with Genome-Scale Metabolic Network Reconstructions (GENREs). More specifically, I will highlight how these models have been used to make advances in understanding human metabolism concerning health and disease. As an undergraduate biomedical engineering student, I was permitted to enroll in a class on computational models and the power of using mathematics to represent complex biology. Naturally, this interested me, and my interest in computational modeling is what led me to pursue graduate-level training. During graduate school, I was able to explore my fascination with computational modeling and found a new interest in combining these techniques with big data, data science, and bioinformatics. In this chapter I discuss the evolution of GENREs, starting from their initial creation for representing the metabolism of different microbial species, and how GENRE-based analyses were extended over time to create the first human reconstruction. I also discuss how context-specific models were adapted from the first human reconstruction, followed by a summary of the application of GENREs in toxicology.

1.2 Introduction

One biological concept that is important to measure and understand is metabolism, as it is a key process that contributes to other biological processes. Metabolism refers to both the catabolism (breakdown) and anabolism (synthesis) of distinct compounds, which can generate energy for cells. Different metabolic functionalities include how dietary

components are processed, as well as how bioactive molecules like drugs are handled. Metabolism also plays a role in critical physiological functions such as how the kidneys regulate pH, or how the heart is less efficient under specific forms of stress. With the complexity that is presented by the metabolism of the human body, it is important to understand how metabolic changes can affect downstream physiological functions. One way to capture the intricacies of metabolism is with constraint-based modeling, which allows one to simulate systems that aren't directly measurable in controlled laboratory settings. One type of constraint-based model used for metabolism is a Genome-Scale Metabolic Network Reconstructions (GENREs). GENREs are under-constrained systems that require linear algebra to model changes at the gene and reaction level for an organism. GENREs are well suited for answering questions about metabolism because they connect how changes at the gene level affect downstream reactions and metabolites in conditions such as cancer or diabetes.

1.3 History of GENREs

GENREs were developed in the late 1990s to understand microorganisms such as *Haemophilus influenzae* Rd (Edwards and Palsson 1999), *Saccharomyces cerevisiae* (yeast) (Forster 2003), and *Escherichia coli* (Orth et al. 2011). These GENREs contained anywhere from 500 to over 2000 reactions, and between 300 and roughly 1100 metabolites. Since their creation, GENREs have been used to explore the metabolic function of microbial species with applications in human health. Most notably, models for *Pseudomonas aeruginosa* (Bartell et al. 2017), the altered Schaedler flora (ASF) (Biggs et al. 2017), and *Plasmodium falciparum* (Carey et al. 2017) have also been developed for researching questions in the fields of antibiotic resistance, the gut microbiome, and

malaria, respectively. In addition to microorganisms, GENREs have more recently been built to represent mammalian species to more directly address questions relating to human health such as cancer, and drug toxicity.

In 2007, the first human GENRE, *Homo sapiens* Recon 1, was created by Duarte *et al (Duarte et al. 2007)*. This reconstruction was larger than previous microbial GENREs as it contained 2766 metabolites, 3311 biochemical reactions, and captured 288 metabolic functions from different cells and tissues present in humans. Recon 1 was an instrumental first step in moving the field toward effectively studying human metabolism. The authors of Recon 1 summarize what was known about human metabolism at the time, and highlighted important gaps in knowledge, such as the lack of information on the intracellular transport of metabolites. Since the creation of Recon 1, this reconstruction was vastly expanded to include 5324 metabolites and 7785 biochemical reactions in Recon 2 (Thiele et al. 2013) and subsequently updated to Recon 2.2 (Swainston et al. 2016) to correct errors in the curation of Recon 2. The updates from Recon 1 to Recon 2 included new metabolic reactions that better-captured drug metabolism, lipid metabolism, and steroid metabolism, making it a more effective platform for analyzing the metabolic ramifications of an increased array of disorders. Additionally, a large fraction of metabolites and enzymes were annotated with information from DrugBank (Wishart, Yannick D. Feunang, et al. 2018) and other third-party resources. The authors went on to note that Recon 2.2 was the first mammalian model to correctly predict free energy production from different carbon sources (Swainston et al. 2016). Recon 3D was recently created from Recon 2.2, where protein and metabolite structures were added to give a

three-dimensional view of human metabolism (Brunk et al. 2018), further expanding the field to include additional information to more accurately capture metabolism.

In addition to helping to advance human systems biology, the creation of Recon 1 also opened up a pathway for other eukaryotic reconstructions, both human (Mardinoglu et al. 2014; Pornputtapong et al. 2015) and non-human (Sigurdsson et al. 2010; Mardinoglu et al. 2015), to be generated. However, only one set of models has been produced in parallel to include a human and a rodent model to capture the species-specific differences between the two organisms (Blais et al. 2017). These paired rat and human reconstructions provide comprehensive coverage of human and rat metabolism and help to bridge the gap between animal experiments (*in vitro* and *in vivo*) and clinical findings. This connection affords researchers the ability to more rapidly generate and test hypotheses for differential drug treatment responses between humans and rats.

1.4 Current State of the Field

Since Recon 1's initial creation, many research groups have adopted their own versions of human reconstructions to answer different questions in human metabolism. Within the metabolic modeling community, it is common practice to integrate different omics datasets into GENREs to create context-specific models. The goal of this process is to create models that are maximally reflective of the occurring biology within the context from which the datasets were collected. These models can then be used to predict drug targets (Agren et al. 2014; Ghaffari et al. 2015), clinical biomarkers (Jerby and Ruppin

2012; Asgari et al. 2018), or phenotypic switches at the cellular level (Yizhak, Le Dévédec, et al. 2014; Marín de Mas et al. 2018).

One example of how human reconstructions have been used is by predicting drug targets for cancer. One research group generated metabolic models of healthy vs cancer cells by developing a new algorithm, the Personalized Reconstruction of Metabolic models (PRIME) (Yizhak, Gaude, et al. 2014), that combined the Recon 1 reconstruction with gene expression data from the HapMap project (International HapMap Consortium 2005) and NCI-60. The models generated in this study were assessed at different proliferation rates and helped identify a potential drug target, Malonyl-CoA Decarboxylase (MYLCD), which inhibits cancer cell growth while avoiding attacking healthy cells. This enzyme was found by systematically removing reactions from the network to see which reactions inhibited growth in cancer cells, but not healthy cells. This example provides a direct application of how metabolic models can be used in combination with experimental data to identify how we can improve human health. As our knowledge of human metabolism continues to expand, the utility of metabolic models will commensurately increase, and we can apply these models to address questions in areas such as personalized medicine, or drug discovery.

Another example of a context-specific model is HepatoNet1, which has been used to simulate different functions of the liver (Gille et al. 2010). HepatoNet1 was created from a combination of genes and reactions with evidence of enzyme activity in either human hepatocytes or other mammalian hepatocytes from three sources: Recon 1 (Duarte et al.

2007), the Edinburgh human metabolic network reconstruction (EHMN) (Ma et al. 2007), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Kanehisa and Goto 2000; Kanehisa et al. 2019; Kanehisa 2019). This model was used to explore the synthesis of bile acids and the detoxification of ammonia. Additionally, the model was tested for 123 other liver functions and compared the completion of these functions against Recon 1. Specific functions that were captured include gluconeogenesis, glycogenesis, and ureagenesis. The creation of HepatoNet1 helped to capture what was known about liver metabolism and was even used as a resource for creating Recon 2 (Thiele et al. 2013). Hepatonet1 provides one example of how context-specific models are used as tools to simulate known physiology at the tissue or cellular level.

GENREs have also been used to look at the off-target effects of drugs on the kidney (Chang et al. 2010). A kidney-specific model (the Chang *et al*. model) was created by applying relevant physiological constraints from the Human Metabolome DataBase (HMDB) (Wishart, Yannick Djoumbou Feunang, et al. 2018) to a reduced version of Recon 1. These kidney-specific constraints helped determine the inclusion of specific reactions and metabolites relevant to the kidney. The authors then applied these constraints to the kidney model after using the GIMME algorithm (Becker and Palsson 2008) to identify genes and reactions that were active based on microarray data. The authors then modeled blood pressure control as a metabolic reaction by reviewing the literature and generating a list of metabolites that are reabsorbed and secreted by the kidney. The absorption and secretion of these metabolites combined with ATP maintenance represented the kidney's ability to regulate blood pressure. To simulate drug

off-target effects, the authors eliminated any flux through a list of specific reactions that corresponded with a known drug target to simulate a genetic knockout and subsequent impacts on renal function. From this approach, the authors generated a list of how drug targets would impact renal function. Ultimately the authors make note of this strategy being used as a tool for testing drug off-target effects computationally.

In addition to the Chang *et al*. model of kidney metabolism, there have been three additional models of kidney metabolism published (These models are compared in size in **Table 1.1)**. The human kidney-specific metabolic network from Zhang *et al*. (Zhang et al. 2013) was built from Recon 1 and pruned with the Model Building Algorithm (Jerby et al. 2010). This model explored the importance of the control that certain housekeeping genes had over kidney-specific genes for a range of kidney functions. While this model was highly suited for exploring genes important for general kidney function, the authors were focused on genes in the model and elected not to explore changes to kidney metabolism as a whole. With a larger or more well-curated model, more kidney functions could be explored. A more recent model of kidney metabolism was based on Recon 2 (termed Reduced Recon 2 model) and targeted how internal fluxes were altered under a variety of growth conditions (Quek et al. 2014). This study emphasized the model reduction process itself and specifically measured differences resulting from growth conditions, but not kidney function outside of growth. In the latest kidney-specific metabolic model, there was a “merged model” of kidney metabolism published two years later (Sohrabi-Jahromi et al. 2016), based on two of the previous models (Chang et al. 2010; Zhang et al. 2013). This merged model specifically emphasized focal segmental

glomerulosclerosis and explored how flux values through reactions and pathways changed during this disease. The existing kidney models either addressed different questions about kidney biology or were built to summarize current knowledge of kidney metabolism. However, these existing models lack curation around drug metabolism or don't include specific kidney-specific functionalities, which makes it difficult to study different forms of kidney injury. The limitations of existing models highlight the necessity for the creation of a new kidney model over the course of my dissertation work.

1.5 GENREs in the field of toxicology

GENREs have been used to gain an increased understanding of the toxicology of different organs. Toxicity is an unintended effect of many compounds, resulting in significant health complications. The liver, kidney, and heart are often subject to adverse, potentially toxic effects because of their role in drug metabolism (Albini et al. 2010; Chen et al. 2015; Awdishu and Mehta 2017). Understanding the metabolic changes to these organs can facilitate understanding the mechanisms associated with toxicity, thereby guiding the development of novel strategies to counterbalance any toxic effects. Because GENREs capture gene and reaction level information, these platforms are ideal for providing insight into the mechanisms of toxicity, and how acute toxicity progresses towards long-term damage. Additionally, omics data has frequently been used with GENREs to make predictions on how metabolism is altered after compound exposure (Agren et al. 2014; Stempler et al. 2014; O'Brien et al. 2015; Sawada et al. 2018), further allowing for the discovery of potential mechanisms associated with toxicity. With such mechanistic interrogation of metabolism, we can identify potential biomarkers associated with toxicity and potential intervention points involved with toxicological processes.

To further enable interrogation of metabolism, I have contributed to the development of Transcriptionally Inferred Metabolite Biomarker Response (TIMBR) (Blais et al. 2017), which predicts changes in metabolite levels due to chemical exposure. Since TIMBR's publication, the algorithm has been used to study acetaminophen-induced injury for the liver and the kidney using *in vivo* omics data (Pannala et al. 2018; Pannala et al. 2019). With my research, I further validated TIMBR predictions by measuring *in vitro* experimental data and validation data from the same biological sample. I discuss this further in chapters three and four. This is one way in which I have used my research to advance the field of toxicology using GENREs.

In this dissertation, I discuss how GENREs work, and how I have used metabolic models to characterize toxicity in the liver and kidney as well as predict how metabolism in these organs is impacted by exposure to different compounds. With this work, I have provided tools to advance the fields of systems biology, metabolism, and toxicology, which can lead to other discoveries to improve human health.

Models	Authors	Reactions	Metabolites
Reduced Kidney Model	Chang 2010	443	340
Human kidney-specific metabolic model	Zhang 2013	2904	1898
Reduced Recon2	Quek 2014	7327	4962
Merged model	Sohrabi-Jahromi 2016	3034	1996

Table 1.1. Summary of existing metabolic networks of kidney metabolism
The four metabolic network models of kidney metabolism published before this dissertation. Quantities of reactions and metabolites included in each model are listed, along with the name of each.

Chapter 2: A simplified metabolic network reconstruction to promote understanding and development of flux balance analysis tools

2.1 Foreword

In this chapter, I discuss the full details on how Genome-scale metabolic Network REconstructions (GENREs) work and describe some of the constraint-based analyses used to simulate biological systems. This chapter allowed me to combine my interest in teaching, with science to educate other scientists on how to use the research tools I use in this dissertation. The concept for this chapter started with a previous graduate student Edik Blais, in his dissertation and has since evolved into a published article titled “A simplified metabolic network reconstruction to promote understanding and development of flux balance analysis tools” (Rawls, Dougherty, et al. 2019). With that, I would like to thank my co-first authors Bonnie V. Dougherty, and Edik M. Blais, along with the other co-authors Ethan Stancliffe, Venkat R. Pannala, Kalyan Vinnakota, Glynis L. Kolling and Jason A. Papin.

2.2 Introduction

Genome-scale metabolic network reconstructions (GENREs) have emerged as powerful tools for the contextualization of high-throughput data, to guide discovery in biological systems, and to simulate the effect of genetic and environmental perturbations(Oberhardt et al. 2009; Mardinoglu, Gatto, et al. 2013). While sophisticated computational tools have

been developed to facilitate the understanding and use of these network models(Becker et al. 2007; Agren et al. 2013), there remains a significant challenge for the development of intuition for how associated modeling analyses are applied and can be used. While several “simple” networks have been developed (Orth et al. 2010; Hädicke and Klamt 2017), these networks often fail to capture key features of energy metabolism and are too large (>100 reactions) to readily develop intuition for associated modeling analyses.

GENREs represent metabolic reactions and corresponding genes and capture the stoichiometric relationships between metabolites and associated chemical transformations. With GENREs as a foundation, constraint-based reconstruction and analysis (COBRA) methods allow for computational predictions of metabolic phenotypes(Becker et al. 2007). Constraints placed on the GENRE are based on genetic, environmental, or thermodynamic factors and reduce the space of possible phenotypes of the system. To constrain the network, experimental data such as transcriptomics or metabolomics can be integrated into GENREs, adding specificity to the network representing a particular environment or cell state. Recently developed methods(Jerby et al. 2010; Zur et al. 2010; Wang et al. 2012; Agren et al. 2014; Schultz and Qutub 2016) allow for the integration of context-specific ‘omics datasets, constraining general models of metabolism for more context-specific computational predictions.

Applying constraints to the model allows the generation of novel scientific hypotheses about the underlying biology of the system studied; therefore, it is important for the reconstructions to maintain mass balance and consider thermodynamic constraints

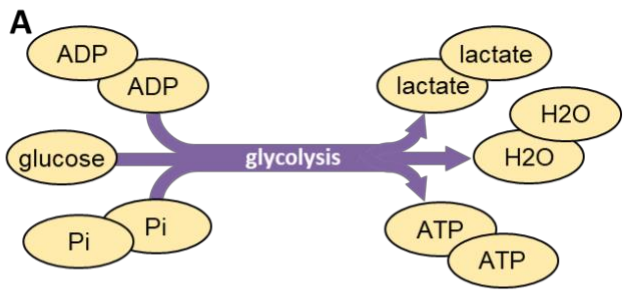
appropriately. For example, a well-documented problem for several published reconstructions is the infinite synthesis of ATP due to thermodynamically infeasible loops (Swainston et al. 2016; Fritzemeier et al. 2017). A key challenge for the development and understanding of methods for the analysis of metabolic networks is that they are inherently complex and difficult to interrogate. Such complexity makes it a challenge to understand the impact of particular characteristics of a new method. Simple networks can be a crucial tool for new method development. To date, simple metabolic networks have captured representative catabolic and anabolic reactions, but other key characteristics of real metabolic networks are often neglected.

Here we present *iSIM*, a simplified metabolic network that captures key features of energy metabolism that are frequently neglected in other similar simple network models. We demonstrate the application of several constraint-based techniques to this simplified network reconstruction, including how ATP production and metabolic flux ranges for each reaction in the network are altered through simulations of single and double gene deletions. Finally, we illustrate how *iSIM* can be used to help understand how thermodynamic errors can arise in metabolic networks and affect computational predictions. We provide source code for analyses in Python, R, and MATLAB. In summary, *iSIM* serves as a tool both for understanding and benchmarking GENREs and their associated methods. *iSIM* also demonstrates the need for quality control measures that are important to consider when reconstructing and analyzing genome-scale metabolic networks and in the development of constraint-based methods.

2.3 Methods

2.3.1 Stoichiometric matrix

The stoichiometric matrix (*S* matrix) captures the stoichiometric coefficients for metabolic reactions in the GENRE with each metabolite represented in a row and each reaction in a column. For example, the simplified representation of glycolysis in **Figure 2.1A** is represented as one column in the S matrix, and the metabolites in **Figure 2.1B** are represented as rows in the matrix. The elements of the matrix are the stoichiometric coefficients of the metabolite consumed (negative) or produced (positive) in the reaction. **Figure 2.1C** is a visual representation of an S matrix for the human GENRE (Blais et al. 2017) which contains over 8000 reactions, demonstrating how individual metabolites participate in many reactions in a network. Of note, there are several reactions in the general human metabolic network reconstruction that contain many metabolites (indicated by a large number of colors in one column) which represent lumped reactions or reactions that represent a large number of metabolites such as lipid synthesis and degradation.



B

Metabolite	Abbreviation	Coefficient
Adenosine diphosphate	ADP	-2
Phosphate	Pi	-2
Glucose	Glucose	-1
Lactate	Lactate	2
Water	H20	2
Adenosine triphosphate	ATP	2

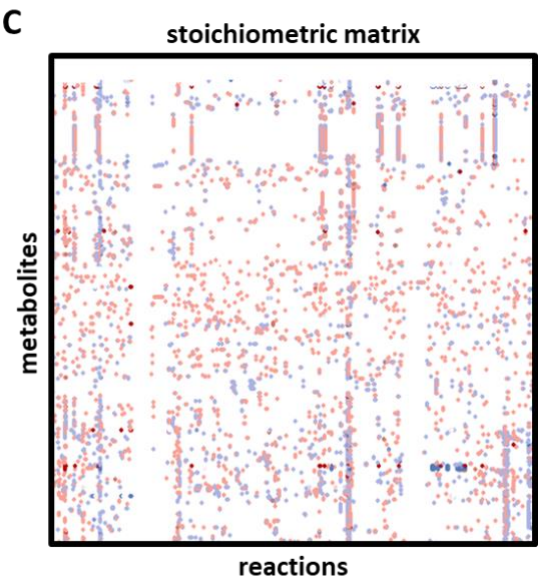


Figure 2.1 – The stoichiometric matrix captures stoichiometric relationships between metabolites in reactions. (**A**) Representation of a simplified glycolysis reaction, catalyzing the conversion of one unit of glucose and two units of adenosine diphosphate and two units of phosphate into two water molecules, two units of ATP and two units of lactate. (**B**) Stoichiometric coefficients describing the amounts of metabolites consumed (blue) and produced (red) by this simplified representation of glycolysis. (**C**) The stoichiometric matrix (S matrix) is a mathematical representation of reactions like the example in **A** formatted as a sparse matrix where each column represents a reaction and each row represents a metabolite. Each point in the sparse matrix accounts for the stoichiometric coefficient of each reaction-metabolite pair, where the color represents whether the metabolite was consumed (blue) or produced (red). The S matrix in C represents thousands of reactions within the human metabolic network (Blais et al. 2017).

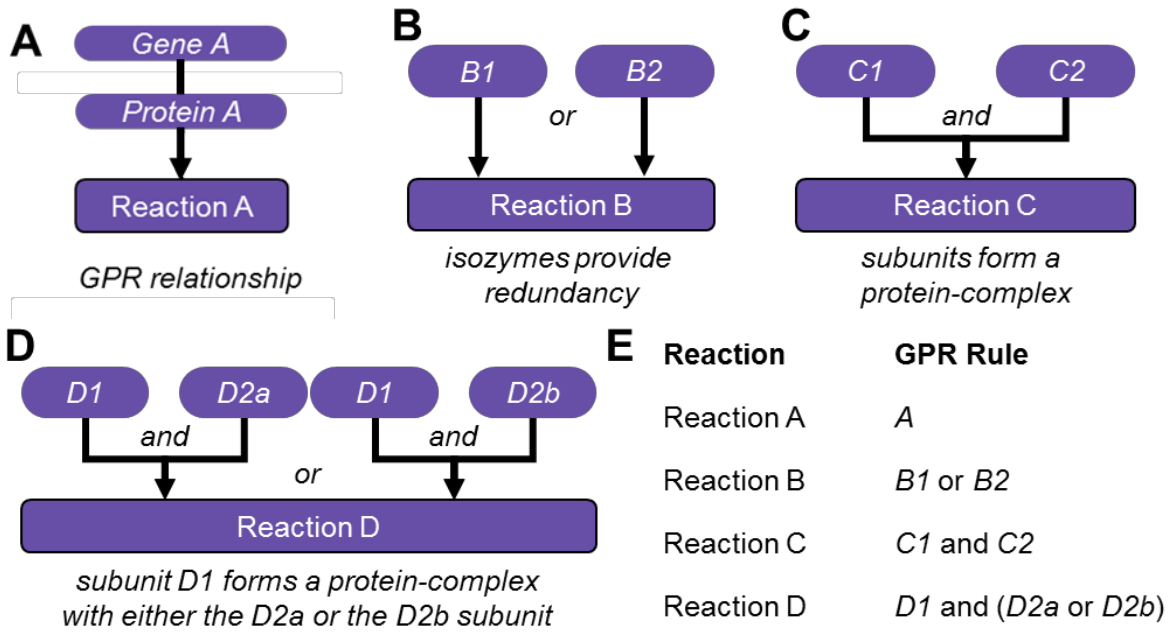


Figure 2.2 – Gene-protein-reaction (GPR) rules describe the relationship between genotype and phenotype. (**A**) Example of a GPR rule representing an enzymatic reaction catalyzed by the protein product of a single gene. Genes within GPR rules are often represented as Entrez gene identifiers, Ensembl transcripts, UniProt proteins, or Enzyme Commission numbers. (**B**) Example of a redundant GPR rule where either protein *B1* or protein *B2* can independently catalyze the same function. In this case, these isozymes are separated by an "or" statement in the GPR rule. (**C**) Example of a complex GPR rule where both C1 and C2 are required for the catalytic reaction to occur. In this case, two non-redundant subunits that form a protein complex are separated by an "and" statement. (**D**) Example of a complex GPR with redundancies where D1 can form a protein complex with either D2a or D2b. In this case, the GPR rule can be separated by unique protein complexes or first by subunits then by redundancies as represented in E. (**E**) Table summarizing genotype-phenotype relationships from A-D as Boolean GPR rules.

2.3.2 Gene-Protein-Reaction Relationship Rules

GENREs contain gene-protein-reaction (GPR) rules that describe a gene's relationship to a protein and the reaction catalyzed by the protein. These GPR rules allow for the generation of tissue-specific or organism specific models for further analyses. GPR rules are organized with Boolean logic relationships between genes, proteins, and reactions.

Figure 2.2 gives examples of different Boolean relationships between genes and reactions. Using these GPR rules, simulations of gene deletions can systematically account for the removal of specific genes and their associated reactions from the network to determine the effect on the system. For example, in **Figure 2.2C**, if gene C2 is removed, reaction C is also removed because both genes C1 and C2 are essential for reaction C to function. This is not the case in the reaction illustrated in **Figure 2.2B**, as either gene B1 or B2 catalyze reaction B. Therefore, if gene B1 is removed, reaction B can still occur. GPR rules also allow for the integration of gene or protein expression data with GENREs, where genes or proteins can be turned "on" or "off" based on expression data. **Figure 2.2D** shows a complex GPR rule, where the gene D1 is needed, and either D2a or D2b. Each of the different types of GPR rules and their corresponding reaction is summarized in **Figure 2.2E.**

2.3.3 Flux Balance Analysis

The *S*-matrix is often an underdetermined system with more reactions than metabolites, meaning there are many potential solutions to the system; consequently, constraints are used to further reduce the solution space. Optimization methods are used to find a solution that satisfies the constraints and is of particular interest, for example, the flux distribution that corresponds to maximum biomass yield. . Flux Balance Analysis (FBA) is a constraint-based method that calculates the maximum possible flux through a specified reaction, known as the objective function, subject to constraints on reaction fluxes. Given a stoichiometric matrix (S), lower and upper bounds (v_{lb} and v_{ub}) on reaction

fluxes (v), and an objective (v_{obj}), FBA uses linear programming to solve the following optimization problem:

max v

s.t. $S \cdot v = \frac{dC}{dt} = 0$ Equation 1

$v_{lb} \leq v \leq v_{ub}$ Equation 2

Equation 1 describes the mass balance constraints, where S is the stoichiometric matrix multiplied by a flux vector (v). This equation represents the steady-state assumption inherent in many FBA simulations. Equation 2 describes reaction constraints applied to the FBA problem. These constraints are often applied to simulate conditions relevant to a biological question. For example, these constraints may be defined to represent a particular media condition (presence of glucose in the media, absence of fructose in the media) where the bounds are set to zero for metabolites absent in the media. Furthermore, these constraints are also often used to capture irreversibility of reactions as known (i.e., one of the bounds is set to zero so that flux through the corresponding reaction can only go in one direction).

2.3.4 Flux Variability Analysis

Flux variability analysis (FVA) is a constraint-based modeling technique that calculates the range of flux through each reaction that are feasible given a set of constraints. Unlike FBA, which is often used to calculate a single flux distribution to solve the associated

optimization problem, FVA results in the calculation of a range of feasible fluxes through each reaction which still results in a defined value of the objective function. FVA is implemented in two steps; first, FVA constrains a metabolic network to require a minimum amount of flux through an objective function, which is usually a percentage of the maximum value determined by FBA. In the second step, FVA optimizes for the flux through each reaction in the network. This analysis enables the identification of the minimum and maximum flux values for each reaction in the network, providing feasible flux ranges that are consistent with the constraints imposed in the first step of FVA.

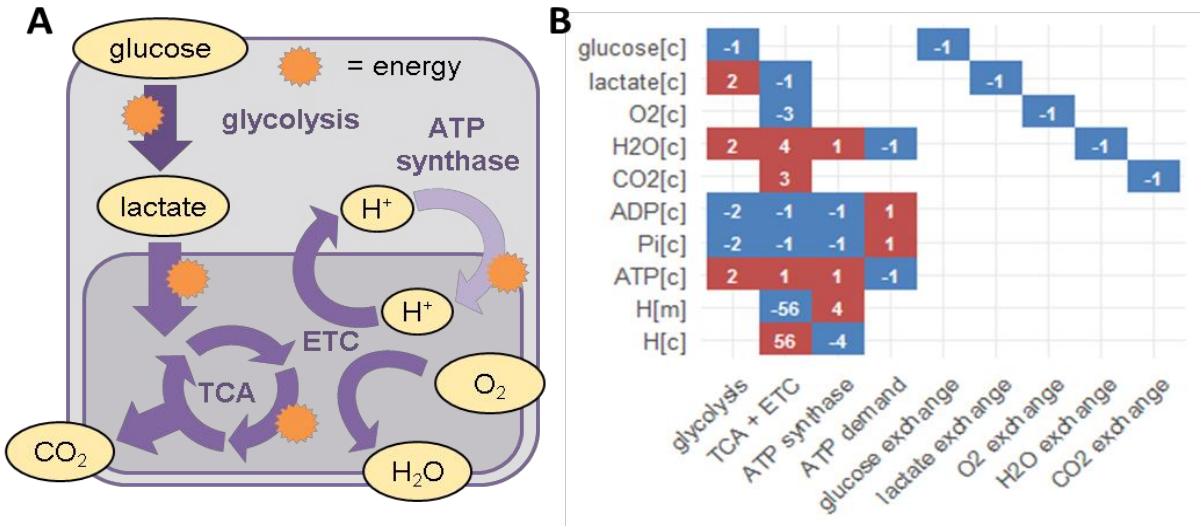


Figure 2.3. iSIM, a prototypic metabolic network that represents simplified energy metabolism. (A) *iSIM* represents two major catabolic pathways for glucose that generate cellular energy via the addition of a phosphate group to ADP to generate ATP, indicated here in orange. Glucose is broken down into lactate while producing ATP and lactate is broken down to carbon dioxide through the tricarboxylic acid (TCA) cycle and the electron transport chain (ETC). The proton concentration gradient created by the ETC is then utilized to generate more ATP. **(B)** The stoichiometric matrix of *iSIM*. Each column represents a reaction and each row represents a metabolite, where the brackets indicate the cellular localization of metabolite: [c], cytosol; [m], mitochondria. Negative values (blue) indicate a metabolite is being consumed in the reaction and positive values (red) indicate a metabolite is produced. The first 4 columns represent biochemical reactions while the last 5 columns represent exchange reactions that allow metabolites to enter or leave the network.

2.4 Results

2.4.1 iSIM: a simple metabolic network

We have created a prototypic GENRE, *iSIM,* to represent central energy metabolism (**Figure 3**). *iSIM* includes 9 unique metabolites in the cytosolic and mitochondrial compartments and five exchange reactions for the consumption or secretion of glucose, lactate, O_2, H_2O, and CO_2 (**Figure 2.3B**). To simplify the network, select linear metabolic reactions are aggregated as lumped reactions that do not explicitly include all the intermediate steps within a pathway. For example, the first metabolic reaction, 'glycolysis' (**Figure 2.3**), represents the reactions involved in conversion of glucose to lactate and the generation of ATP, which includes multiple intermediates not considered here. The second metabolic reaction, 'TCA + ETC,' represents the tricarboxylic acid (TCA) cycle which produces CO_2 and H_2O coupled with the generation of the mitochondrial proton gradient via the electron transport chain (ETC). As a result of the simplification of reactions, there is only one metabolite in the mitochondrial compartment, H+, represented as H[m] **(Figure 2.3B)**. The third reaction, catalyzed by the protein complex ATP synthase, utilizes the proton concentration gradient generated by the ETC to regenerate ATP from ADP. The fourth reaction, ATP demand, represents the consumption of energy through the hydrolysis of ATP to ADP through an ATPase. To explore basic concepts of bioenergetics, maximization of flux through the ATP demand reaction was used as an objective function. The other three metabolic reactions described above facilitate the regeneration of ATP from ADP to fuel the ATP demand reaction. The remaining five reactions, glucose exchange, lactate exchange, O2 exchange, H2O exchange, and CO2 exchange constrain the uptake and secretion of these metabolites in the network. Here,

we use *iSIM* to demonstrate the utility of computational methods and GENREs, including the use of metabolic tasks to demonstrate network functionality, *in silico* genetic perturbations to identify genes of interest, and flux variability analysis to analyze possible flux distributions through a metabolic network.

2.4.2 Recapitulating biological functions with metabolic tasks

To test model functionality, organism-specific biological metabolic processes are summarized as metabolic tasks or metabolic objectives that the model should be able to perform. Metabolic tasks are formulated by specifying a set of input and output metabolites, such as an input of glucose and output of ATP, with constraints for upper and lower bounds of flux. If the model can produce a feasible flux distribution, then the model is capable of completing the task (Thiele et al. 2013). Flux in the model is defined as 1 unit = 1 fmol/cell/hour. For example, given that catabolism of glucose to lactate through the glycolytic pathway is a function often present in metabolic networks, we can formulate a metabolic task that allows uptake of 1 unit of glucose and requires secretion of 2 units of lactate and 2 units of water. This metabolic task could be tested in a given model by setting a lower bound of -1 on the glucose exchange reaction and a lower bound of +2 on the lactate and water exchange reactions. If there is a feasible solution for the model given the constraints, the task “passes”. If not, we have identified an area for model curation or improvement. For each metabolic task, exchange metabolites in the network for which upper and lower bounds are not explicitly stated are set to 0, unless a parameter is included as an output which allows for the excretion of all metabolites. This parameter

should be included in a metabolic task when not all potential metabolic byproducts are known.

To demonstrate the functionality of *iSIM*, we have curated three metabolic tasks representing glucose catabolism with ATP demand, the requirement of O_2 for large ATP production, and oxidative phosphorylation **(Table 2.1)**. Building on the glucose catabolism metabolic task from above, we can require minimum activity through the ATP demand reaction, representing production and then consumption of ATP (Task 1). This metabolic task uses a lower bound of -1 on the glucose exchange reaction, a lower bound of +2 on the lactate and water exchange reactions, and a lower bound of +2 on the ATP demand reaction. To capture the necessity of O_2 for increased ATP production per unit of glucose (Task 2), we can formulate metabolic task that should not pass. If the ATP demand reaction from above has a lower bound of +10, the task should fail since glycolysis can only produce 2 ATP per 1 unit of glucose. However, if we remove the requirement for the secretion of lactate (lower bound to 0) and allow O_2 uptake (lower bound less than 0) (Task 3), the metabolic network will use the TCA + ETC reaction to meet the +10 lower bound on ATP flux. **Table 2.1** provides a summary of lower and upper bounds for these metabolic tasks; the tasks marked as infeasible could not be completed or failed. These three metabolic tasks demonstrate the functionality of *iSIM* in capturing a simplified representation of central energy metabolism.

Metabolic tasks provide a way to demonstrate the functionality of the model and identify areas for improvement. Numerous reconstructions are published with a list of metabolic

tasks or metabolic objectives that the model can perform, demonstrating the validity and applicability of the reconstruction(Gille et al. 2010; Karlstädt et al. 2012; Blais et al. 2017). If a task is unable to pass when the biological function is known to be present, the task identifies an area where the model may be misrepresenting or missing a metabolic reaction, allowing for further expansion and/or curation of the model.

Table 2.1 – Metabolic tasks for *iSIM* *iSIM* metabolic tasks simulate cellular functions by specifying lower bound values (v_{lb}) and upper bound values (v_{ub}) for individual reactions. (*) Minimum required ATP yields for feasible metabolic tasks are sub-optimal (less than maximum). (**) Infeasible metabolic tasks require unrealistic flux through ATP demand that should fail when simulated.

Metabolic Task	Reaction	v_{lb}	v_{ub}
(1) glucose catabolism	glucose exchange	-1	∞
	lactate exchange	2	∞
	H_2O exchange	0	∞
	ATP demand*	1	∞
(2) glucose catabolism without oxygen	glucose exchange	-1	∞
	lactate exchange	2	∞
	H_2O exchange	0	∞
	ATP demand**	10	∞
(3) glucose catabolism with oxygen	glucose exchange	-1	∞
	O_2 exchange	-∞	∞
	CO_2 exchange	0	∞
	H_2O exchange	0	∞
	ATP demand*	10	∞

2.4.3 Genetic perturbations and gene essentiality

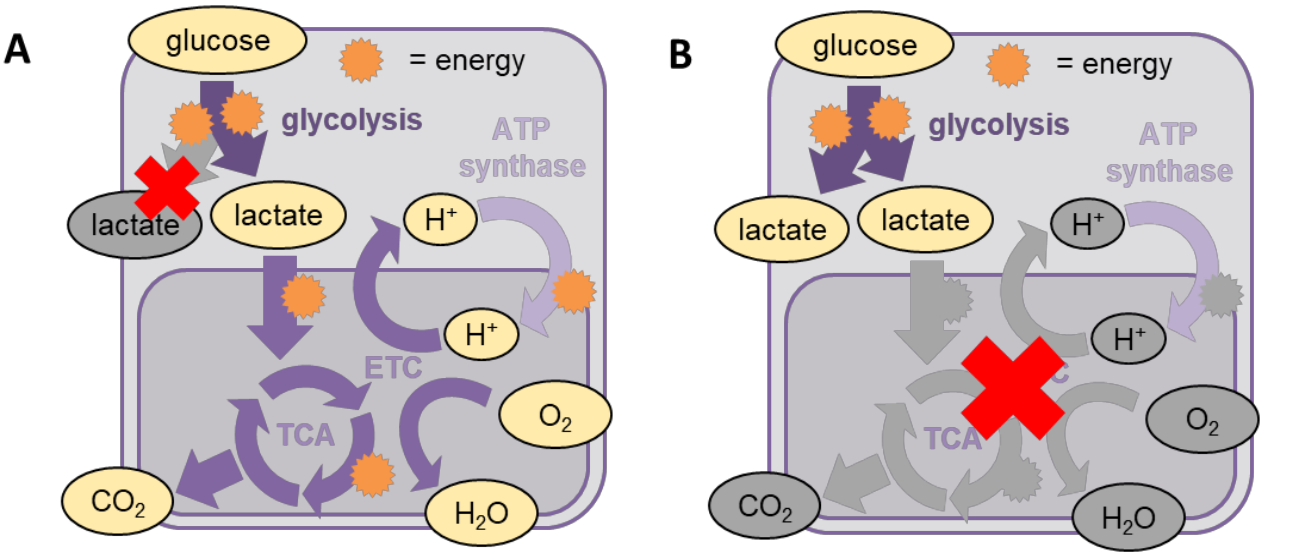
To identify gene targets of interest in a metabolic network, GENREs are analyzed with constraint-based methods to simulate the effects of gene knockouts using the gene-protein-reaction (GPR) relationships. In an *in silico* gene knockout screen, each gene in the model is removed through the identification of reactions that are catalyzed by that gene through the GPR relationships and these associated reactions are disabled by setting the lower and upper bounds to 0. In *iSIM*, for simplicity, each reaction is associated with only one gene. Therefore, a gene knockout screen will remove each reaction systematically and evaluate the effect on the objective function of ATP demand. **Table 2.2** shows the maximum values through the objective function (ATP demand), after disabling each gene and the reaction associated with that gene. Removing the gene associated with glycolysis (phosphofructokinase (PFK)), disables the glycolysis reaction which in turn completely blocks ATP production since glucose is the only carbon source entering the system. The same holds for removing the gene for glucose exchange, glucose transporter (GLUT), which blocks the import of the sole carbon substrate into the model. Removing the gene associated with the TCA + ETC reaction, citrate synthase (CS), disables the TCA + ETC reaction and decreases the flux through ATP demand from 32 to 2. Finally, removing the gene associated with ATP demand (with a key ATPase in muscle contraction bioenergetics (heavy chain 2 (MYH2)) as representative of several other ATPases) completely removes flux through the network since there is no demand reaction.

Table 2.2 – Single gene deletion results for *iSIM*. Maximum possible flux through the ATP demand reaction predicted by FBA after deleting individual reactions from *iSIM*.

Gene Deletion	Reaction Deletion	ATP demand (fmol/cell/hr)
GLUT	glucose exchange	0
LDH	lactate exchange	32
ETC	O_2 exchange	2
AQP	H_2O exchange	2
CO2	CO_2 exchange	2
PFK	glycolysis	0
CS	TCA + ETC	2
ATPV	ATP synthase	2
MYH2	ATP demand	0

Double or triple gene knockout *in silico* simulations can reveal potential combinatorial effects that are not apparent with single gene knockout simulations. With the exception of disabling the glucose exchange and glycolysis reactions, removing individual reactions in *iSIM* did not reduce ATP demand flux to 0 (**Table 2.2**). To identify potential knockout combinations that inhibit flux through the ATP demand reaction, we used *iSIM* to simulate a system-wide double-knockout screen for 36 unique pairs of genes. By comparing double versus single knockout predictions, we identified five double gene knockout combinations that reduced ATP production more than any of the individual simulated gene knockouts (**Figure 2.4C**). For example, removing only the lactate exchange reaction maintained a maximal ATP demand since lactate was utilized by the TCA + ETC reaction to generate ATP **(Figure 2.4A)**. However, if the lactate exchange reaction is removed in combination with any of the reactions associated with oxidative phosphorylation (O2 exchange, CO2 exchange, H20 exchange, ATP synthase, or ETC + TCA), the flux through the ATP demand reaction drops to zero where removing any of these reactions

on their own (**Figure 2.4B)** did not completely reduce ATP production. Double knockouts provide an indication of the redundant nature of the network, which corresponds to multiple pathways for ATP synthesis.



C

Gene A	Reaction A	Flux A (fmol/cell/hr)	Gene B	Reaction B	Flux B (fmol/cell/hr)	Flux A and B (fmol/cell/hr)
LDH	lactate exchange	32	ATPV	ATP synthase	2	0
LDH	lactate exchange	32	CO2	CO_2 exchange	2	0
LDH	lactate exchange	32	AQP	H_2O exchange	2	0
LDH	lactate exchange	32	ETC	O_2 exchange	2	0
LDH	lactate exchange	32	CS	TCA + ETC	2	0

Figure 2.4. Maximum possible flux through the ATP demand reaction predicted by FBA after performing a pairwise gene deletion screen with iSIM. (**A**) Network schematic representing flux through the network with the removal of lactate exchange from the network. (**B**) Network schematic representing the flux through the network with removal of the TCA + ETC reaction from the network. (**C**) Of the 36 possible reaction pairs, 5 double gene knockouts that reduced ATP production more than either single gene knockouts are shown.

In silico gene knockout studies provide insight into the flexibility and potential genetic targets of the network. For *iSIM*, we identified multiple single genes (**Table 2.2**) and double genes (**Figure 2.4C**) that could be targeted to reduce the amount of ATP flux

through the system. *iSIM* demonstrates that double knockouts can identify pairs of genes that do not additively inhibit the objective function of the model, identifying non-obvious targets. In real biological networks, single and double gene knockouts are used to identify gene targets to prevent biomass synthesis which could, for example, serve as potential drug targets for microbial pathogens or cancer cells.

2.4.4 Flux variability analysis

As described above, flux variability analysis (FVA) is a method to calculate the range of possible fluxes through each reaction in a network given a specific value for the objective function. This analysis enables a quantitative assessment of the flexibility in the network as well as identifying essential, non-essential, flexible, and non-flexible reactions for a

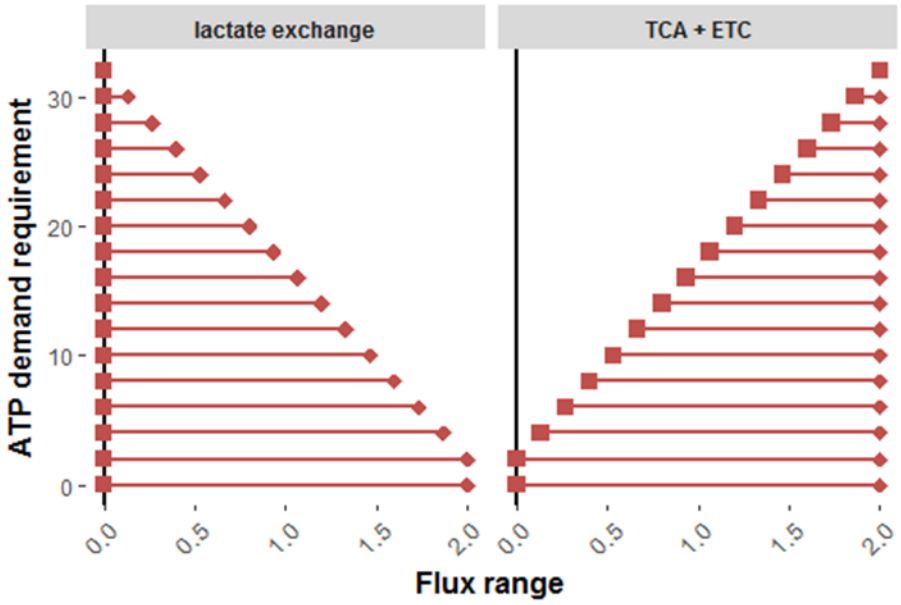


Figure 2.5. Flux Variability Analysis in *iSIM*. Flux variability analysis (FVA) was performed requiring increasing amounts of ATP demand flux, as shown along the y-axis. Maximum (♦), minimum (■), and the range of flux (-) are shown for two reactions in the network, lactate exchange and TCA + ETC. As the ATP demand requirement increases, maximum possible flux values for lactate exchange decrease, indicating a decrease in lactate fermentation, while the minimum required flux values for glucose oxidation increase. The box indicates the flux values for an objective function with 50% of maximal ATP production.

defined objective function. Essential reactions are reactions with a non-zero lower limit on flux, indicating that the reaction must carry flux in order to meet the objective function, whereas non-essential reactions have a flux range that includes zero, meaning the reaction can but does not have to carry flux in order to meet the objective function. Flexible reactions are reactions where the upper and lower range on flux are not equal, indicating that the reaction can carry a range of fluxes in order to meet the objective function, whereas non-flexible reactions have upper and lower range that are the same, indicating that the reaction must carry a specific flux through the network in order to meet the objective function. As above, FBA can be used to calculate a maximum flux of 32 units through the ATP demand reaction per 1 unit of glucose entering the *iSIM* network. To determine which reactions in the network are essential, non-essential, flexible and non-flexible, FVA can be used to determine feasible flux ranges for the remaining eight reactions. For this example, the lower bound on the flux through the objective function (ATP demand) was set at 50% of the maximum (16 flux units). Given that the lactate exchange reaction was involved in all double gene knockouts which completely inhibited flux through the ATP demand reaction, we have chosen to illustrate the flux range through this reaction. When requiring 50% of maximal flux through the ATP demand reaction, the lactate exchange reaction has a range of 0.000 to 1.066 units of flux while the glycolysis reaction has a range of 0.933 to 2.000 units of flux (**Figure 2.5**). Therefore, lactate exchange is a non-essential (lower bound of zero), flexible (upper and lower bounds are not equal) reaction in order to maintain a minimum of 16 units of flux through the ATP demand reaction while the glycolysis reaction is an essential (positive lower bound), flexible (upper and lower bounds are not equal) reaction.

To better understand the relationship between ATP production and glucose breakdown, FVA was performed using incremental requirements of flux through the ATP demand reaction, from 0 to 100% of maximal flux or 0 to 32 units of flux (**Figure 2.5**). Flux variability, or the range of possible flux values, decreased for both the lactate exchange reaction and the TCA + ETC reaction when the minimum flux through the ATP demand reaction was greater than 2. As the minimum flux through the ATP demand reaction increased to 32, the flux through the lactate exchange reaction decreased to 0, completely disabling flux through the reaction.

FVA is a useful approach to determine ranges of fluxes through individual reactions in the network, demonstrating the flexibility of metabolic networks to meet specific objective functions. With *iSIM*, we can see a tradeoff in the values of flux through the reactions involved in either anaerobic glycolysis or oxidative phosphorylation, based on the ATP needs of the cell. Combined with other approaches described above, FVA can also be used to identify how networks adapt to changes in the environment or genetic perturbations.

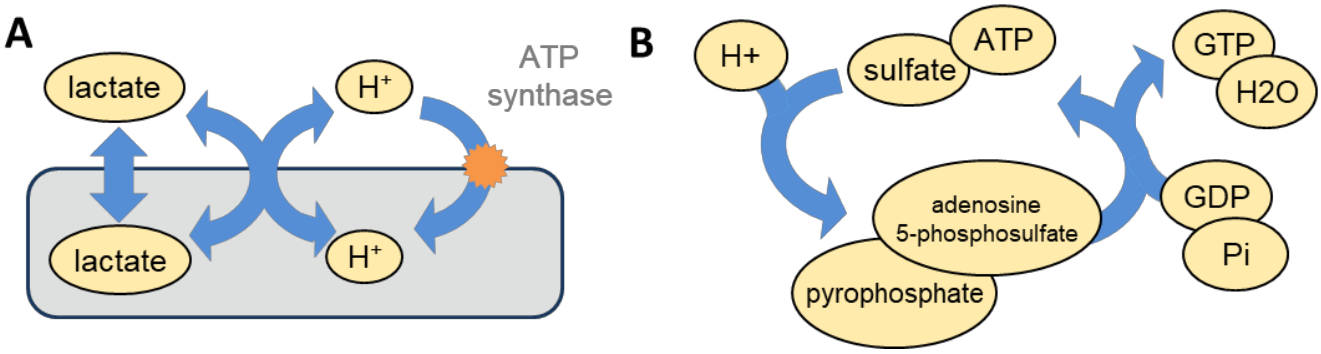


Figure 2.6. Introducing errors into metabolic network reconstructions. (**A**) Including two new reactions in *iSIM* creates a thermodynamically infeasible loop that creates infinite amounts of ATP. Lactate is transported into the mitochondrial with no energy cost but is transported out of the mitochondria with a proton which is then used to generate ATP using the ATP synthase reaction. (**B**) As identified by Fritzemeier et al (Agren et al. 2014), erroneous energy generating cycles often include groups of reactions which produce energy metabolites without an input. In this case, GTP is produced for "free" through each cycle of these two reactions.

2.4.5 The importance of balancing thermodynamics in reactions

Errors can be unintentionally incorporated into metabolic network reconstructions through thermodynamically infeasible reactions and stoichiometrically unbalanced reactions that can affect computational predictions. Since all living organisms rely on external nutrients for energy to fuel biological processes, it is important that 1 unit of glucose in mammalian network reconstructions generates 32-36 ATP under ideal conditions with an unlimited supply of oxygen and at least 1 unit of ATP in the absence of oxygen([CSL STYLE ERROR: reference with no printed form.]). In HMR2 (Mardinoglu et al. 2014) and other human GENREs (Duarte et al. 2007; Thiele et al. 2013), 1 unit of glucose yields an infinite amount of ATP with an unlimited supply of oxygen due to thermodynamically infeasible loops, an error which has been corrected in Recon 2.2 (Swainston et al. 2016). *iSIM*, a prototypic metabolic network consisting of nine reactions, captures theoretical ATP yields but the addition of two reactions can generate unrealistic ATP yields.

The regeneration of ATP from ADP, catalyzed by ATP synthase and driven by the proton gradient between the mitochondria and the cytosol, is an important reaction in central energy metabolism. In *iSIM*, ATP synthase catalyzes the reaction ADP[c] + P_i [c] + 4 H[c] → ATP[c] + H_2O[c] + 4 H[m] to generate ATP by the movement of protons between the mitochondrial and cytosolic compartments. This reaction represents a simplification of the 2.7 protons necessary for the formation of one ATP in mammalian mitochondria(Watt et al. 2010) with the one proton necessary to transport lactate into the mitochondria. To demonstrate the creation of a thermodynamically infeasible loop, two reactions were added to *iSIM*: a hypothetical reversible transport reaction representing the passive

diffusion of lactate across the mitochondrial membrane (lactate[c] ↔ lactate[m]) and a transport reaction representing the symport of lactate and a proton across the mitochondrial membrane (lactate[m] + H[m] ↔ lactate[c] + H[c]). These two reactions create a loop where lactate is transported into the mitochondria through the first reaction and then back out of the mitochondria with a proton through the second reaction (**Figure 2.6A**). ATP synthase then utilizes this H[c] to generate ATP. The addition of these two reactions represents two errors: the addition of an infeasible but stoichiometrically balanced reaction and the addition of a thermodynamically infeasible reaction. First, lactate cannot freely diffuse through the cell membrane due to its positive charge and therefore must be transported with an H. Secondly, although the transporter responsible for transporting lactate with an H can operate reversibly, the transporter only transports down a concentration gradient, which in this case is from lactate produced through glycolysis in the cytosol to the mitochondria. Together, these two reactions can produce ATP with no energy source. To resolve this issue, lactate should only be allowed to enter the mitochondria through irreversible active transport which requires a proton for entry.

To avoid introducing thermodynamically infeasible loops during the network reconciliation and manual curation process, it is important to check tasks regularly to ensure realistic ATP production. To avoid proton movement problems as described above for the lactate transporter, we included a metabolic task that should fail: regeneration of ATP from ADP without a carbon-based fuel source. Second, excess oxygen and inorganic ions should not increase ATP yields above 36 units of flux (Salway 2017) for 1 unit of glucose, a result which was observed with multiple human GENREs. Various algorithms exist to include

thermodynamic constraints with metabolic network reconstructions and are discussed in detail elsewhere (Ataman and Hatzimanikatis 2015).

2.4.6 The importance of maintaining stoichiometric balance in reactions

As demonstrated, infinite loops can be introduced into networks through thermodynamically infeasible reactions. Stoichiometrically unbalanced reactions can produce metabolites without the necessary inputs, which can then be used to maximize the objective function, or as a source for ATP generation (van Heck et al. 2016; Fritzemeier et al. 2017). As noted by Fritzemerier et al, an example of a pair of non-stoichiometrically balanced reactions are the reactions catalyzed by sulfate adenyltransferase which were included in multiple models generated by ModelSEED **(Figure 2.6B)**. Together, this pair of reactions can generate GTP without consuming energy.

Here, we demonstrate the importance of the fundamental assumptions of thermodynamically and stoichiometrically balanced reactions which form metabolic network reconstructions. Failures in these assumptions, as seen with the reversible lactate transporter or the sulfate adenyltransferase reactions, can lead to infeasible ATP or energy production and inaccurate results.

2.5 Discussion

GENREs allow for the system-wide integration of genetic and metabolic information in a mathematical formalism, enabling the prediction of phenotypes with constraint-based

analytical methods. Prototypic networks are often created to illustrate new modeling method concepts. While additional small networks (Hadicke and Klamt, 2017; Orth et al, 2010) allow for predictions that aren't immediately apparent, the larger number of reactions (> 100 reactions) makes it difficult to trace pathways in the network and understand basic COBRA concepts. In addition, most such systems capture catabolic and anabolic functions while neglecting energy-generating metabolism. We created *iSIM*, a simplified metabolic network, to illustrate key considerations for metabolic network reconstructions and associated modeling methods.

In particular, *iSIM* highlights the importance of thermodynamically and stoichiometrically balanced reactions in a metabolic network reconstruction. Various algorithms have been proposed to automate the identification of thermodynamically infeasible reactions (Schellenberger, Lewis, et al. 2011) and these methods can be used with metabolic tasks in network reconstructions to ensure feasible ATP yields. Recent work highlights energy generating cycles that were removed from the most recent human metabolic network reconstructions to produce feasible ATP yields for a variety of carbon sources (Fritzemeier et al. 2017).

In summary, *iSIM* serves as a tool for understanding constraint-based methods commonly used with GENREs. With access to the network reconstruction in several commonly used languages for constraint-based modeling methods (Matlab, R, Python), *iSIM* can be a useful didactic tool for illustrating new methods and promoting understanding of key concepts.

Chapter 3: Genome-Scale Characterization of Toxicity-Induced Metabolic Alterations in Primary Hepatocytes

3.1 Foreword

In this chapter, I present a new paired transcriptomics data and metabolomics data approach to use GENREs with experimental data to predict changes in metabolite levels and validate those predictions. The paired omics data approach is used to identify potential biomarkers of drug-induced hepatotoxicity. My experience using GENREs and interest in studying liver metabolism led to my pursuit of an internship at Hemoshear Therapeutics in the summer of 2018, where I did similar work. This internship helped me get hired full time after graduate school at Hemoshear. This chapter has been published as a research article titled “Genome-Scale Characterization of Toxicity-Induced Metabolic Alterations in Primary Hepatocytes” (Rawls, Blais, et al. 2019). With that, I would like to thank my co-authors Edik M. Blais, Bonnie V. Dougherty, Venkat R. Pannala, Kalyan Vinnakota, Glynis L. Kolling and Jason A. Papin.

3.2 Introduction

Toxicity is an unintended effect of many compounds, resulting in significant health complications. The liver, kidney, and heart are often subject to adverse, potentially toxic effects because of their role in drug metabolism (Albini et al. 2010; Chen et al. 2015; Awdishu and Mehta 2017). Hepatotoxicity is of particular concern (Zimmerman 1999;

Church and Watkins 2017; Rueda-Zárate et al. 2017), highlighting the need to understand how liver metabolism is altered as a result of toxicity. Understanding the metabolic changes to the liver can facilitate understanding the mechanisms associated with toxicity, thereby guiding development of novel strategies to counterbalance any toxic effects. Furthermore, with such mechanistic interrogation of liver metabolism, we can identify potential biomarkers associated with toxicity and potential intervention points involved with toxicological processes.

Genome-scale metabolic network reconstructions (GENREs) have emerged as useful tools for the study of cellular metabolism (Gille et al. 2010; Karlstädt et al. 2012; Mardinoglu, Agren, et al. 2013; Väremo et al. 2015) . GENREs represent metabolic reactions in a stoichiometric matrix that accounts for the stoichiometric coefficients of chemical transformations and the associated metabolites. GENREs also account for gene-protein-reaction (GPR) rules that map relationships between genes, the proteins they encode, and the reactions they catalyze in the network. With the GPR mappings and stoichiometric matrix to account for associated metabolic reactions, GENREs can be used to predict gene essentiality, changes in metabolites secreted, and the ability of a cell to catabolize particular carbon substrates; because of these characteristics, GENREs are increasingly applied to tackle questions about cellular toxicological responses (Bartell et al. 2014; Gatto et al. 2015; Carbonell et al. 2017; Brunk et al. 2018; Pannala et al. 2018).

The incorporation of omics data into GENREs allows for cell-type specific interrogation of metabolism. Transcriptomics and proteomics data are frequently integrated into GENREs

to create cell-type specific models. Several algorithms to integrate omics data into GENREs have been developed (Shlomi et al. 2008; Zur et al. 2010). Often with such methods, the integration of omics data constrains the GENRE by turning “on” and “off” genes and their associated reactions, reflecting gene expression in different conditions. These expression integration algorithms help to contextualize these omics data and improve predictions of cellular metabolic functions.

Biomarkers are currently used in the diagnosis of cancer, cardiac function, and renal function (Shlipak et al. 2012; Jungbauer et al. 2016; Pan et al. 2018; Lotan et al. 2018 Mar 31) among other pathologies, often associating the presence or absence of a molecule with a specific diagnosis. For example, alanine aminotransferase (AST) is a protein that is used frequently as a biomarker of liver function (Zimmerman 1999; Dufour et al. 2000); high levels of this protein indicate that the liver has been damaged. A recently developed computational method for predicting biomarkers is called Transcriptionally Inferred Biomarker Response (TIMBR) algorithm (Blais et al. 2017), which uses gene expression data contextualized in a GENRE to estimate relative changes in secreted metabolite levels. In a previous study (Blais et al. 2017), TIMBR predicted changes in extracellular metabolite levels based on gene expression data for cells exposed to various chemical compounds. Predictions of a limited number of metabolite biomarkers for one chemical were validated, but a global evaluation of how well the biomarker predictions matched experimental data was missing. In this study, predictions from TIMBR are compared with paired metabolomics data to observe the differences between computational predictions and experimental data. Agreement between predictions and

experimental data can be illustrative of mechanism behind an observed biomarker; disagreements between the computational model and experimental data can facilitate the development of specific testable hypotheses.

Here, we exposed primary rat hepatocytes to four chemical compounds and characterized their acute metabolic response (**Figure 3.1**). After exposure, transcriptomics and metabolomics data were collected from the same sample. We characterized the response of the hepatocytes to the compounds through changes in gene expression and metabolite levels, and evaluated similarities and differences between the cell's responses across all conditions. The transcriptomics data was integrated into a GENRE of rat metabolism via iMAT (Zur et al. 2010) to create a hepatocyte-specific network model, then the TIMBR algorithm was used to predict changes in the secreted metabolite profile. We compared these predictions with the coupled metabolomics data. With this methodology, we present a comprehensive strategy to characterize the toxicological response of hepatocytes to compounds of interest, and provide a framework to identify further areas of study in hepatocyte drug and toxicity metabolism.

3.3 Methods

3.3.1 Hepatocyte growth conditions

Frozen, primary rat hepatocytes (male, Sprague-Dawley) were purchased from Thermo Fisher Scientific and cultured according to manufacturer's directions. Briefly, cells were rapidly thawed in a water bath (37°C), resuspended in plating media (William's E media base supplemented with FBS, dexamethasone, penicillin/streptomycin, insulin,

GlutaMAX, and HEPES; Gibco #CM3000), pelleted (50 x *g*, 5 min), and plated at ~85% confluence in 12-well tissue culture plates. After 24 hours, plating media was replaced with maintenance media (William's E media base supplemented with dexamethasone, penicillin/streptomycin, ITS+, GlutaMAX and HEPES; Gibco #CM4000) and cells were incubated at 37°C under 5% CO_2 for the remainder of the experiment.

3.3.2 Hepatocyte exposure to compounds

Hepatocytes were exposed to a hepatotoxicant and general toxicants at sub-toxic levels. Sub-toxic levels were defined as concentrations that resulted in minimal cell death but observed phenotypic changes (e.g., decreases in albumin production, ATP levels, increases in cytochrome p450 activity). The compounds were acetaminophen (APAP) at 3mM, carbon tetrachloride (CCl_4) at 10mM, trichloroethylene (TCE) at 1mM, and 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) at 1nM. APAP and CCl_4 are known hepatotoxicants, while TCDD and TCE are not considered primary hepatotoxicants typically, but do indeed induce hepatotoxicity. APAP, TCDD, and TCE conditions have 4 replicates, while CCl_4 and the DMSO controls have 3 replicates. Solutions were made in WEM containing 0.1% DMSO with 0.1% DMSO as a control. Cells were exposed to the compounds for six hours. Concentrations and the 6 hour time point were selected based off literature evidence of comparable studies and conditions (Mitchell et al. 1985; Cai et al. 2005; Aly and Domènech 2009; Kienhuis et al. 2009; Uehara et al. 2010; Dere et al. 2011; Xu et al. 2012; Forgacs et al. 2013).

3.3.3 RNA isolation, sequencing, and analysis

After supernatants were collected, cells from each condition were treated with TRIzol® and then scraped and collected into tubes. Chloroform was added to each tube and after shaking, cells were poured into pre-spun phase-lock gel tubes (5PRIME). Tubes were then spun in a cold room, the upper phase was collected, and isopropanol and glycogen were added to each tube followed by gentle inversion. Supernatants were again spun in a cold room and the resulting pellet was washed twice with 75% ethanol. The pellet was semi-dried and then dissolved in nuclease-free H_2O. RNA samples were treated with DNA-*free* DNA removal kit (Ambion/Invitrogen), according to manufacturer's instructions, to remove any remaining DNA. RNA was quantified using the Qubit RNA broad range kit and sample integrity assessed using Agilent. RNA samples were subjected to rRNA depletion prior to library construction and sequencing; all services were performed by GENEWIZ. Libraries were sequenced using the Illumina HiSeq2500 platform in a 2x100bp pair-end (PE) configuration in High Output mode (V4 chemistry). The Unix-based program Kallisto v. 0.43.0 (Bray et al. 2016) was used to process RNA sequence data in fastq format and quantify transcript abundances. Normalized transcript abundance values (TPM, Transcripts Per Million) were calculated by Kallisto, using default settings, and imported to R for differential analysis. To quantify transcript abundances and aggregate toward the gene level, the package tximport in R was used (Soneson et al. 2015). Differential gene expression was then performed with the standard DESeq2 R package (Love et al. 2014) to obtain a list of differentially expressed genes with their log2 fold change values.

3.3.4 Metabolomics

After hepatocytes were exposed to the different compounds, supernatants were collected and stored at -20°C. Supernatants were then shipped to West Coast Metabolomics (http://metabolomics.ucdavis.edu/) at the University of California, Davis and untargeted analysis of primary metabolites, complex lipids, and biogenic amines was conducted on each sample, DMSO controls, and on blank media. An extraction solvent of 3:3:2 acetonitrile/isopropanol/water was prepared to use with the collected samples for Gas Chromatography Mass Spectrometry (GC-MS) to analyze primary metabolites. External and internal standards for quality control were also prepared along with the samples. Raw results were reported as peak heights for quantification ion at the specific retention index. A full description of the protocol was outlined previously (Fiehn 2016). Lipidomics analysis was performed by preparing samples with methanol, methyl tert-butyl ether (MTBE), and water before running Liquid Chromatography Mass Spectrometry (LC-MS). LipidBlast was used to identify and annotate lipids, and peak heights were reported according to the published protocol (Cajka and Fiehn 2017).

Biogenic amine peak heights were quantified using Hydrophilic Interaction Chromatography Quadrupole Time of Flight (HILIC-QTOF) Mass Spectrometry, and peak heights were calculated followed methods previously described (Meissen et al. 2015). Samples were processed and analyzed according to West Coast Metabolomics protocols. Proteins and small polar hydrophilic small molecules were separated from lipids according to the protocol published by Matyash et al (Matyash et al. 2008). Data was acquired using the following chromatographic parameters. Ultrapure water with 10mM ammonium formate and 0.125% formic acid (pH 3) for mobile phase A, and 95:5 v/v

acetonitrile:ultrapure water with 10mM ammonium formate with 0.125% formic acid (pH 3) for mobile phase B. A column temperature of 40˚C, with the flow rate of 0.4 mL/min and injection volume of 3μL for ESI (+) and temperature of 4˚C was used. The ESI Capillary voltage was +4.5kV for ESI (+), the scan range was m/z 60-1200 Da, and the mass resolution was 10,000 for ESI (+) on an Agilent 6530 QTOF MS. After raw peaks were obtained, they were processed by mzMine 2.0 software to find peaks in up to 300 chromatograms. Relative peak intensities of both identified and unidentified metabolites were generated and used for further analyses.

3.3.5 Data analysis

Before differential expression analysis, genes with no counts were removed from analysis to avoid skewing the results. A gene was considered significantly differentially expressed if the False Discovery Rate (FDR) corrected p-value was < 0.1. Standard Euclidean hierarchical clustering was performed on all the gene expression data and clustering was done by each individual gene. For the metabolomics dataset, primary metabolites, lipidomics, and biogenic amines were read in and combined into one data frame to analyze the data similarly. Experimental replicates were averaged together and fold changes were calculated from the cell samples and the fresh media samples. Significance of metabolite differences were determined with a p-value < 0.05 using the Mann-Whitney U test. All statistical analyses were performed using R version 3.4.0.

3.3.6 Gene enrichment analysis

To perform gene enrichment analysis, the Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics Resource was used with a list of differentially expressed genes for each compound (Huang et al. 2009a; Huang et al. 2009b). The Functional Annotation Tool was used to determine which Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were overrepresented, or enriched. Entrez gene IDs were submitted to the DAVID Bioinformatics Resource website and the *Rattus norvegicus* species was selected. The category "KEGG pathways" and functional annotation clustering were selected. KEGG pathway terms were considered significantly enriched if the FDR corrected *p*-value was less than 0.1.

3.3.7 Flux balance analysis and the creation of tissue-specific models

The stoichiometric matrix (S matrix) was analyzed using the COBRA toolbox v. 2.0.6 (Schellenberger, Que, et al. 2011). The *iRno* reconstruction of rat metabolism, which accounts for the function of 5620 metabolites, 2324 genes, and 8268 reactions, was used to make computational predictions (Blais et al. 2017). *iRno* has been curated to perform liver-specific metabolic tasks, making it appropriate as a base model of liver metabolism. Flux Balance Analysis (FBA) was performed using the optimizeCBmodel function in the COBRA toolbox in MATLAB v. R2016b. Condition-specific models were then created using the iMAT algorithm in the COBRA toolbox. The createTissueSpecificModel function in the COBRA toolbox was used, with iMAT set as the method for expression data integration, using reactions associated with differentially expressed genes and exchange reactions as high confidence reactions to include in the model. Log-fold changes for

differentially expressed genes were supplied as inputs along with a model with genes created for exchange reactions, while the hepatocyte-specific model was provided as an output.

3.3.8 TIMBR Algorithm

The TIMBR algorithm combines the transcriptomics data with the *iRno* network reconstruction to determine production scores for each exchangeable metabolite relative to a control as previously described (Blais et al. 2017). The transcriptomics data was used to generate weights for a control case and a treatment case on each reaction in the reconstruction. Next, for each metabolite, the weighted flux through each reaction was minimized while maintaining positive flux through that metabolite's exchange reaction for the control and treatment conditions. Production scores are normalized using the previously described formula (Blais et al. 2017) to determine whether a metabolite has increased or decreased production relative to the control and used for further downstream analysis. The scripts used to generate each of the datasets can be found on the github site (www.github.com/csbl) published with the TIMBR algorithm.

3.4 Results

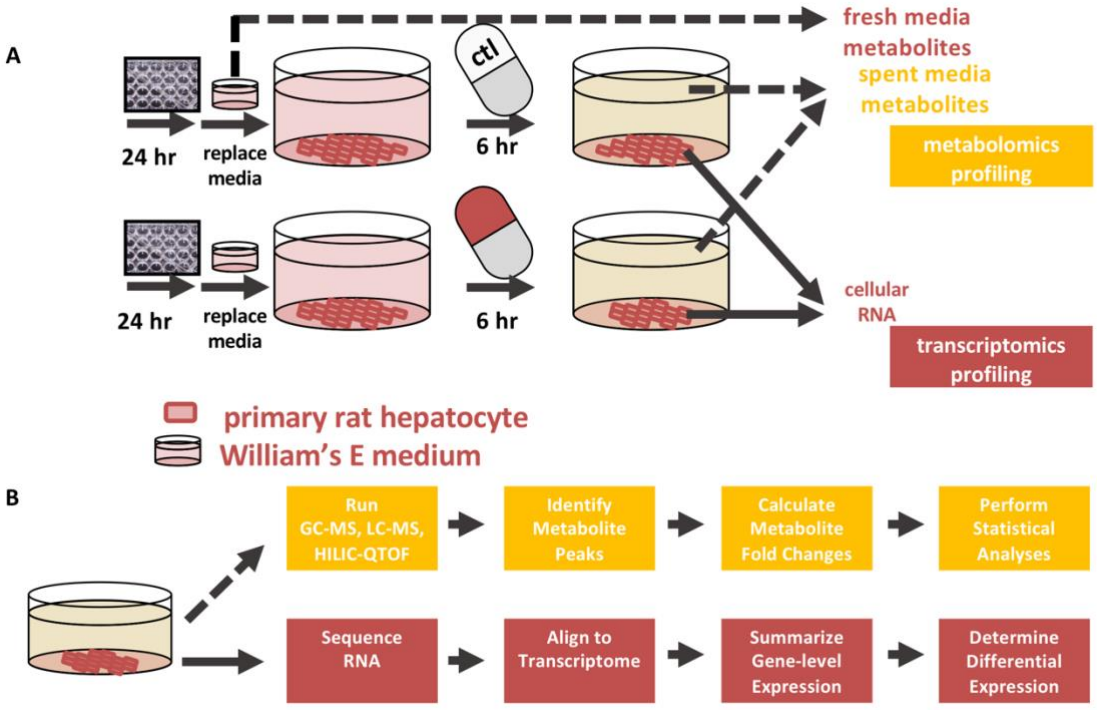


Figure 3.1. Schematic of the experimental set up
(A) Primary rat hepatocytes were plated in 12-well format and exposed to acetaminophen, carbon tetrachloride, 2,3,7,8-Tetrachlorodibenzodioxin (TCDD), or trichloroethylene for six hours. After compound exposure, supernatants were collected and sent for metabolomics analysis. Hepatocytes were lysed and RNA was collected for sequencing. (B) Cellular RNA was isolated and sequenced by Genewiz. With the raw sequencing reads as an input, the program kallisto was used to align sequencing reads to a reference transcriptome. The R packages TxImport and DESeq2 were used to summarize transcript counts to the gene level and to perform differential gene analysis respectively. Spent media from the hepatocytes were collected and sent for GC-MS, LC-MS, and HILIC-QTOF metabolomics at West Coast Metabolomics. After receiving metabolite peak intensities, the data was processed in R to generate a list of differentially abundant metabolites in each condition.

3.4.1 Transcriptomics data reveal compound-specific responses of hepatocytes

Hepatocytes were exposed to acetaminophen (APAP), carbon tetrachloride (CCl_4) 2,3,7,8-tetrachlorodibenzodioxin (TCDD), or trichloroethylene (TCE) for six hours to characterize the differential toxicity-induced metabolic response. **Figure 3.1** shows the experimental layout; after hepatocytes were exposed to each compound, supernatants were collected for metabolomics analyses and RNA was isolated for transcriptomics analysis. DMSO was used as a non-drug control. The number of differentially expressed genes (DEGs) for each condition and time point were determined (**Table 3.1**) and a list of genes from the differential gene analysis was produced. APAP induced the most DEGs in the hepatocytes, while TCE induced the least number of DEGs. To further analyze the genes that were differentially expressed, we used the DAVID Bioinformatics platform to identify enriched KEGG pathways for each compound. **Figure 3.2A** shows the enrichment results of the differentially expressed genes for APAP, CCl_4, TCDD, and TCE. APAP at six hours showed an enrichment for metabolic pathways, while CCl_4, TCDD and TCE at six hours did not (**Figure 3.2**), suggesting that the hepatocyte's metabolism was more altered globally in response to APAP compared to the other three compounds. As evidenced in the enrichment analysis, APAP exposure induced a wide variety of gene expression changes, while gene expression changes after CCl_4, TCDD, and TCE exposure appeared focused towards RNA and protein processing. After investigating the broad effects of the compounds, we then focused on metabolic genes to evaluate how each compound perturbed hepatocyte metabolism.

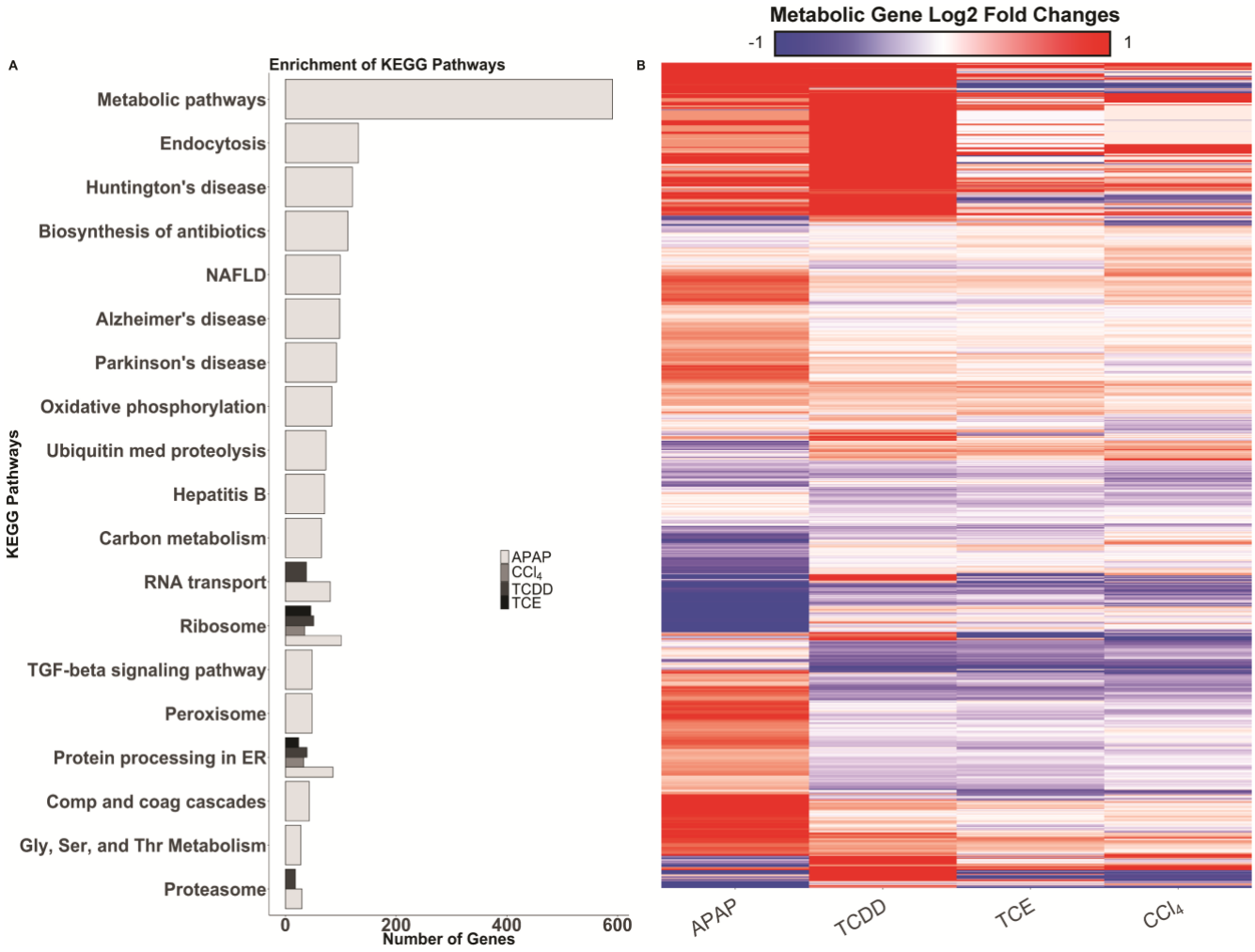


Figure 3.2 Gene enrichment and metabolic gene expression data
(**A**) DAVID enrichment of KEGG Pathways for six hours in APAP-, CCl_4-,TCDD-, and TCE-induced toxicity conditions. The heat map above shows the log2 fold changes of the metabolic genes from sequencing (**B**). Each condition is listed on the x-axis, and the individual genes are listed on the y-axis. Genes that are upregulated are shown in red, while downregulated genes are shown in blue. Genes on the x-axis are clustered by Euclidean distance, using complete linkage.

Figure 3.2B shows a heat map of the log2 fold changes of all the metabolic DEGs with a Benjamini-Hochberg adjusted p-value of less than 0.1 in at least one condition. This heatmap shows that CCl_4 and TCE elicit similar changes in gene expression. We observed changes in expression for the Cyp450 family of genes, often associated with metabolizing drugs (Guengerich 2008). We saw a decrease in *Cyp3a4* in APAP but no

changes in the other compounds, likely because other Cyp450 genes play a role in rat metabolism of compounds (Tran et al. 2001; Zuber et al. 2002). Specifically, *Cyp2e1* is induced in hepatotoxicity (Jaeschke et al. 2002; McGill et al. 2012). We saw upregulation of *Cyp2e1* in APAP-induced toxicity, but not for TCDD- and TCE-induced toxicity; however, there were other genes in the Cyp450 family that were differentially expressed in these other conditions. In APAP- and TCDD-induced toxicity, the Cyp450 gene *Cyp2d4,* also associated with the metabolism of drugs (Mizuno et al. 2003), was upregulated. TCDD-induced toxicity resulted in upregulation of most other Cyp450-related genes, while TCE-induced toxicity resulted in downregulation for many of the same genes. This result highlights that even though there are common pathways of toxicity associated with the liver, these compounds ultimately result in different specific effects on the hepatocytes. In an effort to identify potential biomarkers specific to each compound, we next interrogated the metabolomics data to identify differential effects of each compound.

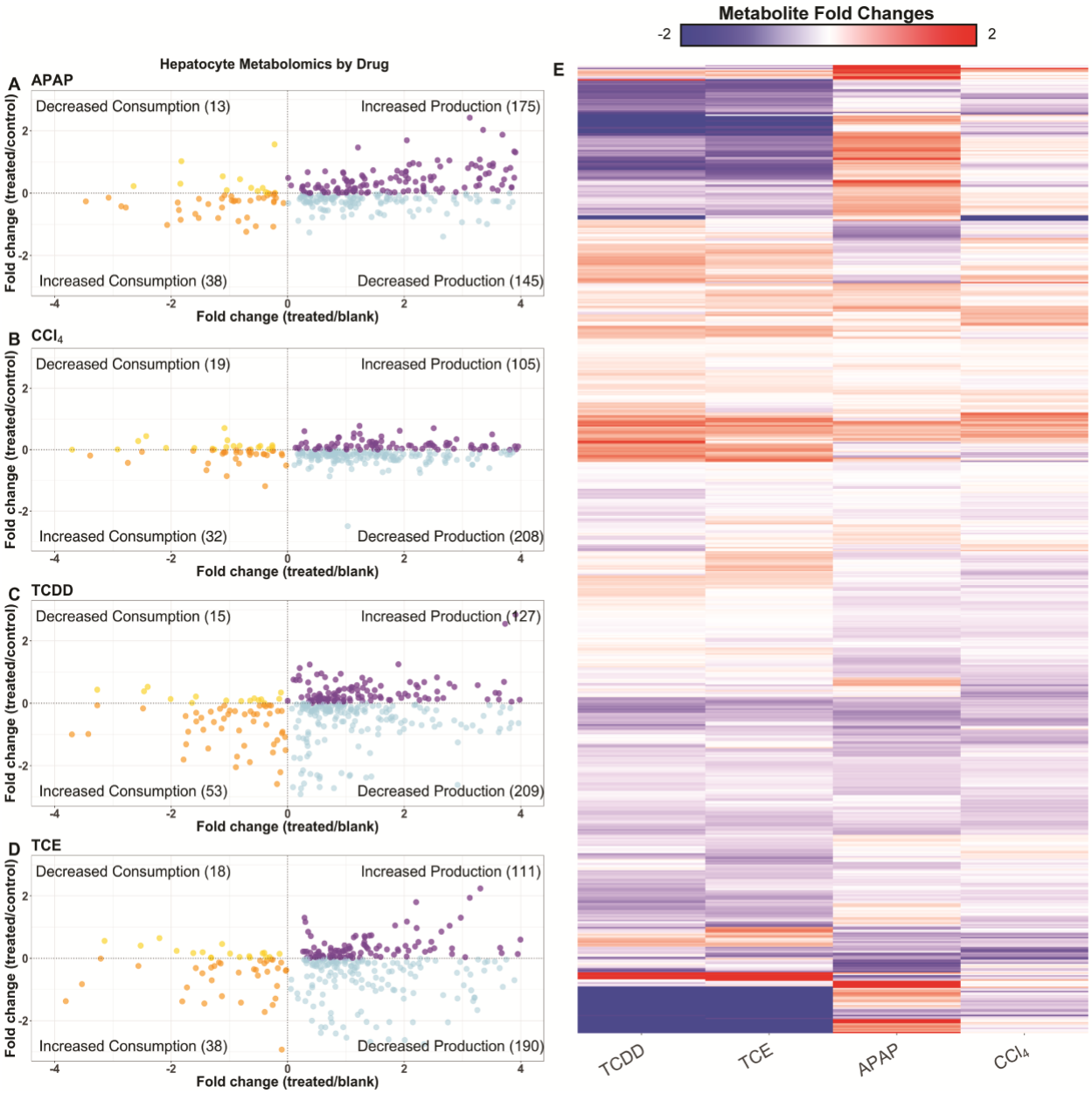


Figure 3.3 Overview of the metabolomics data

The scatter plots show the distribution of metabolites that are significantly ($p < 0.05$) changed when compared to either the control media or blank media, and colored according to their levels when compared to both sets of media. Metabolites in gold have decreased overall consumption, metabolites in orange have increased overall consumption, light blue indicates decreased overall production, while purple shows increased overall production, all with respect to the control media. Plots are shown for APAP- (**A**), CCl_4 – (**B**), TCDD- (**C**), and TCE- (**D**) induced toxicity conditions at six hours. The heat map above shows the log2 fold changes for metabolites compared to their respective controls (**E**). Each condition is listed on the x-axis, and the metabolites are listed on the y-axis. Metabolites that are elevated in production with respect to the control condition are shown in red, while metabolites reduced in production, compared to the control condition, are shown in blue. Metabolites on the x-axis are clustered by Euclidean distance, using complete linkage.

3.4.2 Metabolomic data discriminates the response of the primary hepatocytes specific to each treatment.

APAP produces the most distinct signature of the three compounds, while TCDD and TCE display a similar profile. The metabolomics data are illustrated in scatter plots for APAP (**Figure 3.3A**), CCl_4 (**Figure 3.3B**), TCDD (**Figure 3.3C**), and TCE (**Figure 3.3D**) exposure conditions. The scatter plots show each metabolite, with the fold change of average relative metabolite peak intensity compared to blank medium on the x-axis, and compared to the DMSO controls on the y-axis. With this arrangement, metabolites are classified as having increased or decreased production if the fold change relative to blank is positive, or increased or decreased consumption if the fold change relative to blank is negative. Metabolites are also color coordinated, to help distinguish metabolites that were increased or decreased in their production or consumption. Only metabolites that were significantly changed in either the treated vs. control, or treated vs. blank cases are displayed. From these data, we see that APAP induces the greatest number of metabolites with an increase in production, while the other compounds induced a decrease in production of most measured metabolites. There is a trend for metabolites to either be increased in production (upper right) or decreased in production (lower right). This trend is clear in each condition, as these were the two categories with the most metabolites, although many of these metabolites are not yet identified. In APAP-induced toxicity, there were several amino acids that decreased in production compared to the control case. This result indicates that hepatocytes consumed more amino acids after being exposed to APAP. TCDD and TCE both caused hepatocytes to decrease production of fatty acids, while APAP triggered an increased production of fatty acids. The results from the metabolomics data suggests a clear metabolic difference in the

hepatocytes treated with different compounds, and that the mechanism of action or off target effects of the toxicants may be the likely cause of this shift.

We next decided to interrogate the total metabolic response of the hepatocytes to further discriminate treatment conditions. **Figure 3.3E** shows a heatmap of the individual metabolite levels, and whether or not the amount of the metabolites increased or decreased with respect to the control condition. Again, we noticed that TCE and TCDD showed a similar but distinct pattern of changes in metabolite levels. Valine and leucine were uniquely increased in TCE, while tryptophan, serine, and glutamate were uniquely decreased in TCDD. Between both compounds, nicotinate, glucose-1-phosphate, and aminomalonate all decreased. There were only seven metabolites that increased for both TCDD and TCE, 1,3-diheptadecanoyl-2-(10Z-heptadecenoyl)-glycerol d5 and six unidentified metabolites. There were 80 metabolites that decreased between both compounds including both identified and unidentified metabolites. In the heatmap in **Figure 3.3D** there are a few prominent clusters of metabolites. There was a small cluster of unidentified metabolites in the TCDD and TCE condition whose levels were decreased when compared to the control condition. APAP did not follow this trend, as a number of those same metabolites were increased. Within this large cluster the only identified metabolite was nicotinate. Of the 559 metabolites we were able to detect, only 115 could be identified. Of the identified metabolites, we then looked at the unique metabolites altered by each condition to compare and contrast each compound's effect on the hepatocytes. Common metabolites that consistently decreased across all conditions were L-lactate, glycerate, and alpha-ketoglutarate (AKG), which have been shown to decrease

in other toxicity studies (Kim and Moon 2012). Other studies have shown decreases in citrate and AKG (Ishihara et al. 2006), which have been attributed to disruptions of the TCA cycle. Finally, there were increased lipid levels in TCE and TCDD compared to their controls, suggesting a strong alteration in lipid metabolism in response to these compounds.

3.4.3 TIMBR predictions suggest unique responses to each toxicant

To make predictions on metabolite production levels relative to control from the gene expression data, we created a hepatocyte-specific metabolic model from the unconstrained *iRno* GENRE using iMAT (Zur et al. 2010) along with the gene expression data described earlier. The Transcriptionally Inferred Metabolic Biomarker Response (TIMBR) response algorithm (Blais et al. 2017) was used to create normalized production scores for each metabolite that could be secreted by the model and we compared these values with the fold changes calculated from the metabolomics data above. **Figure 3.4A** shows a distribution of normalized TIMBR production scores by compound, with the median indicated by the notches and black line and the mean represented by the white diamond in the middle of the box plot. From **Figure 3.4A** we see that each group has its mean at about zero, however the median for each group is different. The APAP and TCE conditions show that more scores have positive TIMBR scores while CCl_4 has slightly more negative TIMBR scores. This result suggests that hepatocytes are predicted to produce more metabolites in response to APAP and TCE exposure compared to other toxicant conditions.

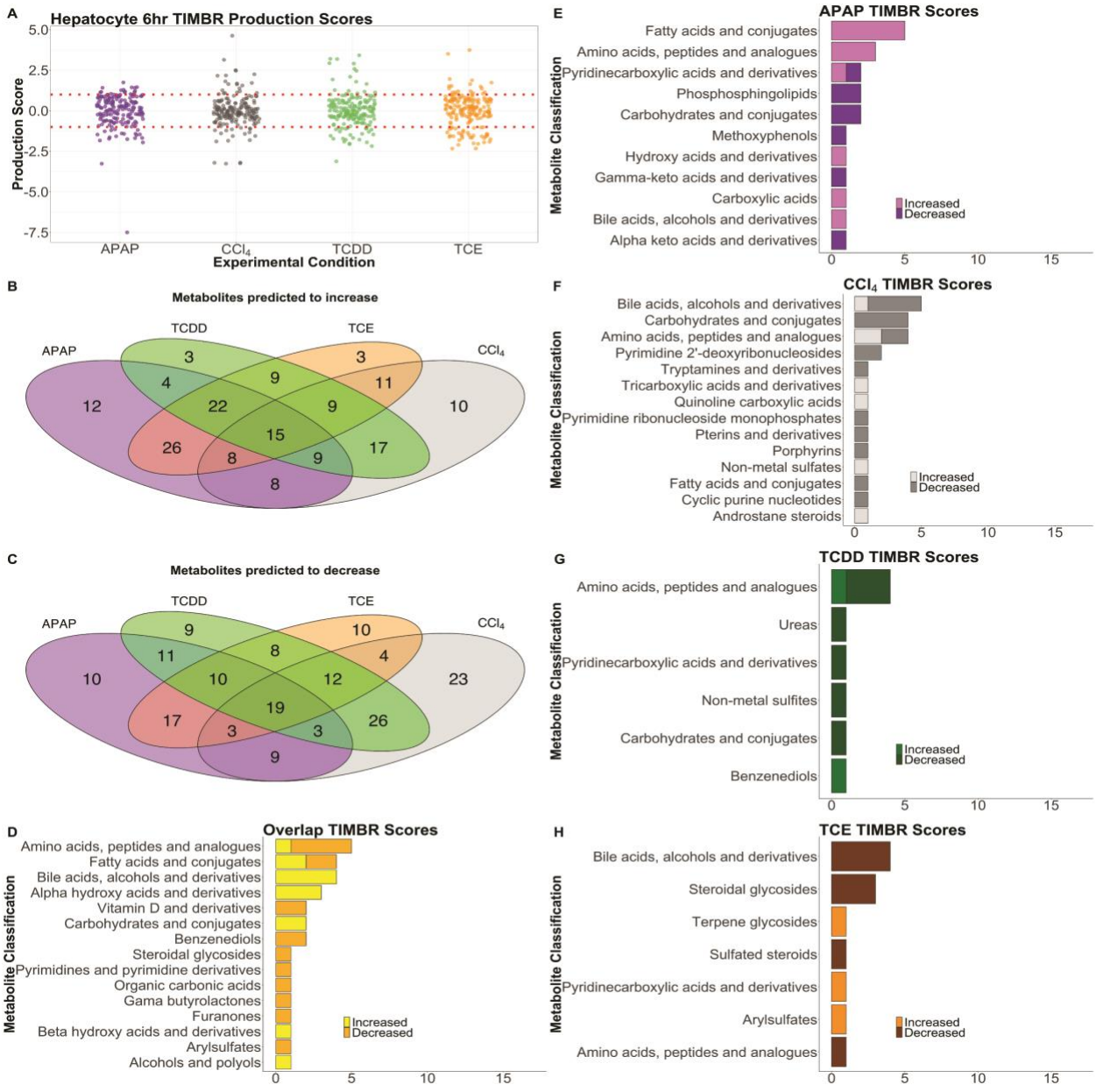


Figure 3.4. Summary and Distribution of TIMBR production scores
The distribution of TIMBR production scores are shown (**A**) indicating that the ranges are similar, but scores have a slight skew according to their condition. The APAP condition results in more negative production scores, while TCE results in more positive production scores. Red lines mark y = 1 and y = -1. Venn Diagrams compare all positive (**B**) and negative (**C**) production scores for each compound, and the overlap between the three conditions. TIMBR scores that are common across all conditions (**D**), and unique to APAP (**E**), CCl_4 (**F**), TCDD (**G**), TCE (**H**) are illustrated. Here, metabolites are classified into categories taken from the subclass names from the Human Metabolome DataBase (HMDB) if available. Metabolite in a category that increases were given a light color, while metabolites in a category that decrease were given a darker color.

We then compared common metabolites that were predicted to increase or decrease after all treatments, which indicate common metabolic shifts in response to drug treatment. **Figures 3.4B** and **3.4C** shows Venn diagrams of the metabolites that were predicted to commonly increase or decrease in production, or uniquely increase or decrease in response to APAP, CCl_4, TCDD, or TCE at six hours, respectively. Similar to the trend noted in all of the TIMBR production scores (**Figure 3.4A**), CCl_4 exposure was predicted to decrease a higher number of metabolite production scores (**Figure 3.4C**) that do not also decrease in other conditions. However, APAP exposure was predicted to increase more metabolite production scores that were not increased in other conditions (**Figure 3.4B**), which is consistent with the prediction of more positive production scores. We then classified metabolites according to their Human Metabolome Database (HMDB) sub-classification that were changing in each condition from the results shown in **Figures 3.4B** and **3.4C**. The bar charts in **Figures 3.4D-3.4H** indicate the number of metabolites uniquely predicted to increase or decrease production after toxicant exposure according to their sub-classification. For shared metabolites across all conditions (**Figure 3.4D**), a small number of amino acids are predicted to decrease, while bile acids are predicted to increase. APAP exposure (**Figure 3.4E**) resulted in the highest number of fatty acids predicted to increase in production, followed by amino acids. The increase in bile acids and amino acids suggests alterations in these pathways in response to liver injury, and has been observed in literature (Kumar et al. 2012; Sun et al. 2013). In CCl_4 exposure (**Figure 3.4F**), carbohydrate compounds are predicted to decrease in production, while these same metabolites were predicted to increase in the other three conditions. With TCDD exposure (**Figure 3.4G**), amino acids are predicted to decrease in production while

with TCE exposure bile acids are predicted to decrease (**Figure 3.4H**) which is similar to CCl_4. Overall the TIMBR predictions illustrate that the response of the hepatocytes to each compound is primarily due to carbohydrate and amino acid metabolism, which could represent a generic response towards toxic compounds. However, predictions from APAP exposure indicate a distinct response in fatty acid metabolism, with CCl_4 and TCE eliciting more of a change in bile acid metabolism, suggesting that we can predict compound-specific effects on the hepatocytes.

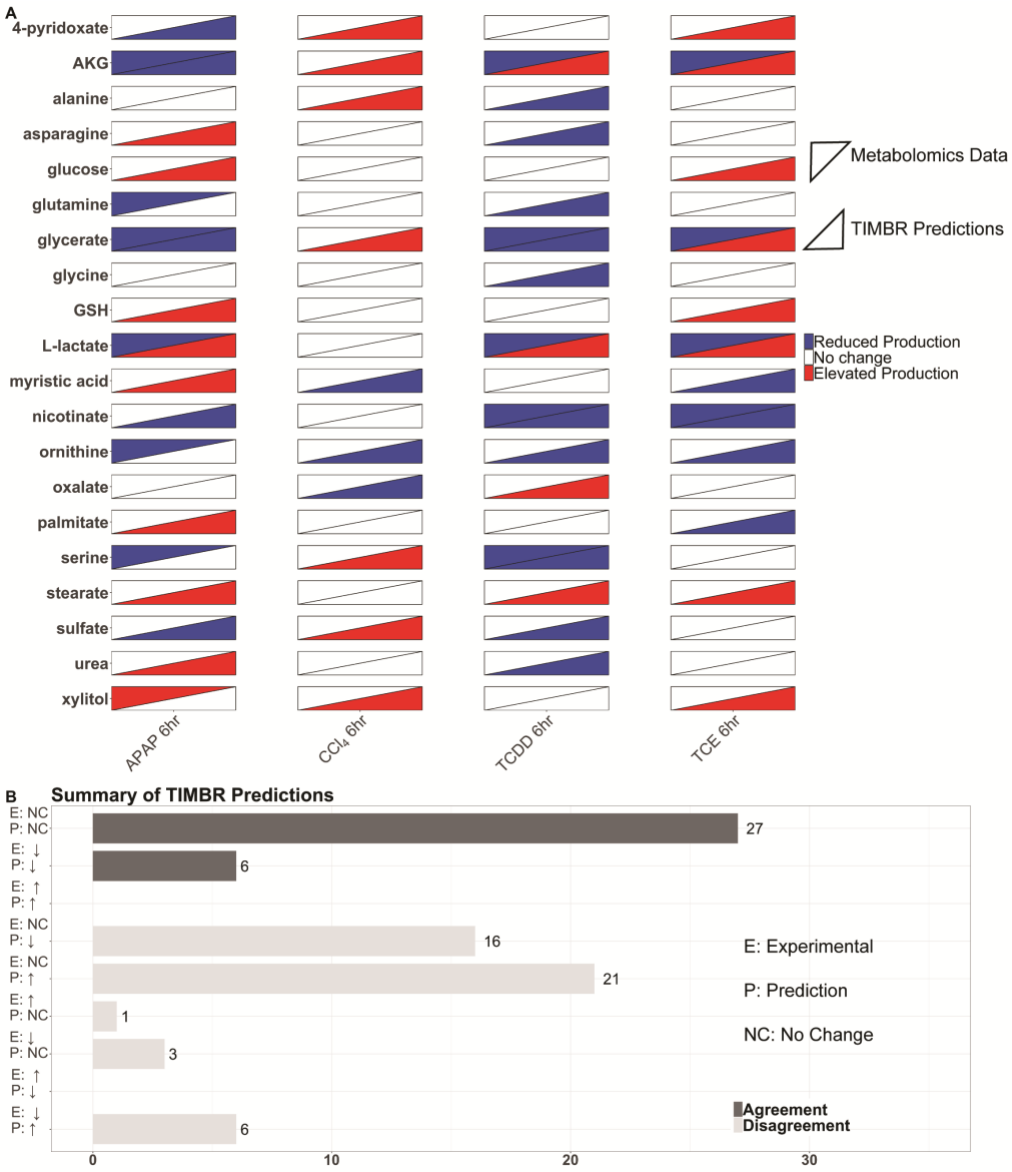


Figure 3.5. Validation of TIMBR production scores using Metabolomics Data
The heat map (**A**) shows the results from the metabolomics data, and the TIMBR production scores for each metabolite we were able to make a prediction for and validate. Each condition is listed on the x-axis, and the metabolites are listed on the y-axis. Metabolomics data are shown in the upper left triangle, and TIMBR Production scores are shown in the bottom right triangle. Red indicates a metabolite is elevated, or predicted to be elevated in production, while blue indicates a metabolite is decreased, or predicted to decrease, in production. The bar chart (**B**) shows the categories a prediction can fall into on the y-axis ranging from increase, decrease, or no change for both the experimental data and the TIMBR predictions. The x-axis contains the number of predictions that fall into the category on the y-axis. Predictions that agree with the experimental data are colored with green bars, while disagreement between the data shows red bars.

3.4.4 Comparing TIMBR predictions and metabolomics data

We next wanted to quantify the similarity and dissimilarity between the predictions and metabolomics data, to determine how indicative gene expression changes were in predicting metabolite levels. In addition to the fold changes calculated from the metabolomics data, the Mann-Whitney U test was used to determine statistical significance at the $p < 0.05$ level. A change in metabolite levels with $p > 0.05$ when compared to the control condition, was classified as “no change”, and represented with a fold change of zero. We then took the subset of secreted, identified metabolites and compared this list with our TIMBR predictions which resulted in 20 metabolites we could validate for each experimental condition. For the TIMBR predictions, metabolite production scores were ranked, and metabolites in the middle 50% of the list were classified as no change and given a value of zero for comparing with the metabolomics data. **Figure 3.5A** shows a heatmap of the metabolomics data and the production scores for metabolites on the y-axis with each condition on the x-axis. Of the 20 metabolites in each condition, we predicted five correctly in the APAP condition, 10 correctly in the CCl_4 condition, nine correctly in the TCDD condition, and nine correctly in the TCE condition. From the list of successful predictions nicotinate, and glycine were correct in three of the conditions, while nine metabolites were correct in two of the conditions. There were six amino acids in the set of 20 that we could make predictions for, and of those six, we correctly predicted two in the APAP condition, three in the CCl_4 condition, while only one prediction was correct in the TCDD condition and five in the TCE condition. While we were able to predict broad changes in carbohydrate and energy metabolism from the TIMBR predictions as described above, the data were too limited to draw the same conclusions from the subset of experimental data that we were able to validate.

Figure 3.5B quantifies our validation results, and shows exactly where predictions were right and where predictions were wrong. The bulk of the correct predictions came from identifying no change in both the experimental condition and the computational prediction. Overall, our accuracy for our predictions was 41%. We made no correct predictions on metabolites that were measured as increase or predicted to increase. Thirty-seven of the predictions were incorrect from detecting a change and predicting there was none, or vice versa. Our sensitivity for detecting no change was 42% and lower for predicting an increase (0%) or decrease (40%). Our specificity for no change or decrease was high at 75%, but was lower (65%) for the no change condition.

3.5 Discussion

There is limited information on biomarkers of toxicity; therefore, novel approaches to elucidate and validate relevant biomarkers are needed. A promise of metabolomics as an approach for identifying biomarkers is its connection to cell phenotype as a change in metabolite levels may represent changes in the functional state of the cells. Here, we present the first use of paired transcriptomics and metabolomics with GENREs to study hepatocytes exposed to different compounds and to integrate these data with metabolic network models to provide insight into the changes that are occurring. At the concentrations and timepoint we selected for exposure, we used standard measures of hepatocyte function (albumin production, Cyp450 activity, etc.) to ensure we were not killing the cells. While we observed minimal changes in traditional measurements of toxicity after six-hour exposure, we observed changes in metabolism as indicated by the transcriptomics and metabolomics data. Additionally, connecting transcriptomic changes

to secreted metabolites even at low-toxic compound concentration can be useful in clinical settings, as these secreted metabolites can be measured to gain an early indication of hepatic injury. Secreted metabolites can then be connected back to transcriptional changes using metabolic network models, which allows us to generate mechanistic insight into observed changes in transcript or metabolite levels.

From this study, we observed a number of transcriptional changes in metabolic genes of hepatocytes following exposure to compounds. Analyses of transcriptional changes highlighted that APAP produced the largest change in hepatocyte gene expression, as expected (Ben-Shachar et al. 2012; McGill and Jaeschke 2013; Sjogren et al. 2014; Taguchi et al. 2015). There were 31 differentially expressed metabolic genes that were shared across all four compounds. There were five genes that were upregulated in each of the four treatment conditions. Among the group of upregulated genes includes two glutathione S-transferase genes, indicative of detoxification mechanisms given glutathione is used to conjugate toxic metabolites (Monks et al. 1990; Guengerich 2008). Furthermore, glutathione S-transferase is responsible for the detoxification of NAPQI, a toxic metabolite that is generated from metabolizing APAP (Henderson et al. 2000). In the metabolomics data, glutathione production was decreased in APAP (albeit $p = 0.34$). For the TIMBR predictions, in the APAP condition we did predict to see decreased production of glutathione, which is attributed towards glutathione detoxifying NAPQI in the hepatocytes.

Twelve differentially expressed metabolic genes that were shared across all compounds were downregulated. Among this group was isocitrate dehydrogenase 3 (*Idh3a)*, which is responsible for the NAD+ dependent conversion of isocitrate to alpha-ketoglutarate. In the metabolomics data, we observed decreased production of alpha-ketoglutarate in response to APAP, TCDD and TCE compared to their respective controls. We also computationally predicted this decrease in AKG in the APAP condition (**Figure 3.5**). These examples suggest that there are some transcriptional changes that are indicative of downstream metabolite changes.

Glycolysis and the TCA cycle were disrupted as a result of compound exposure. In the metabolomics data, we observed that glycerate was decreased in response to exposure to APAP, TCDD and TCE, and glucose-1-phosphate was decreased after treatment with TCDD and TCE. Both glucose-1-phosphate and glycerate can feed into glycolysis and then progress to the TCA cycle. Decreases in these metabolites indicate that the hepatocytes are inefficiently producing ATP via the TCA cycle. This observation is also further supported by the measured decrease in alpha-ketoglutarate in most of the conditions as well as the decrease in citrate in response to APAP. Carbohydrates also feed into glycolysis, and a decrease in carbohydrates can also decrease TCA activity. In **Figure 3.4,** we observe that carbohydrates were predicted to decrease in production after exposure to APAP and CCl_4, which are both hepatotoxicants. While we were not able to correctly predict changes in glycerate production in every condition (**Figure 3.5**), we were able to predict this shift in metabolism via the carbohydrates, which is supported by the

metabolomics data. Thus, TIMBR predictions can be useful for suggesting pathway level differences of a treatment that can be experimentally validated.

We compared our *in vitro* and computational results with other *in vivo* toxicity studies that have been done. Across the different studies, lipid metabolism, amino acid metabolism, and energy metabolism (TCA Cycle) were all affected by exposure to different compounds. One study that focused on TCDD-induced transcriptomic changes identified several genes associated with these pathways that were both upregulated and down regulated (Boverhof et al. 2006). From a metabolomics perspective, TCA cycle intermediates were down regulated in response to APAP (Sun et al. 2008), which agreed with our data. These same pathways came up in common with our TIMBR predictions (**Figure 3.4**) which are based on our measured transcriptional changes. One study noted that in response to APAP-induced toxicity, metabolite levels for glycerol and kynurenine were increased, while threonine, serine, ornithine, lysine, glycerate, and glutathione were reduced (Pannala et al. 2018). The authors also observed enrichment in the glycine, serine, and threonine metabolic pathway (Pannala et al. 2018). While we did not observe the decrease in glutathione levels, we did note enrichment in the glycine, serine, and threonine pathway in the APAP condition (**Figure 3.2A**). The decrease in glutathione in APAP was shown to occur at later time points, due to increasing progression of liver injury as noted by the authors (Pannala et al. 2018). Lastly, CCl_4 is known to cause hepatocytes to increase urinary bile acid levels (Yang et al. 2008). We observed that there were a few bile acids predicted to increase (**Figure 3.4D**), but unique to CCl_4 was the observation that most of the bile acids were predicted to decrease (**Figure 3.4F**). Since *in vitro*

conditions do not fully capture *in vivo* conditions due to differences in time-scales, actual exposure concentrations, among other variables, there is not complete agreement between the *in vitro* and *in vivo* results as expected. However, our *in vitro* experiment provides a means to study changes in hepatocyte metabolism without the variability of an *in vivo* experiment. While results may not fully match, general trends in metabolic changes do agree, as indicated by the shift in fatty acid metabolism from TCA cycle and amino acid breakdown noted earlier, which highlights the utility of *in vitro* systems for interrogating toxicological responses.

One limitation of this study that affected the ability to make predictions was the lack of overlap between the metabolomics data, and metabolites for which we were able to make TIMBR predictions. For the primary metabolites in the metabolomics dataset, only 115 out of 559 were identifiable. Of these 115, there were only 21 metabolites in the subset that were secreted and that were accounted for in our current network reconstruction, as shown in **Figure 3.5**. While the number of correct predictions was limited, we were still able to make predictions on glycolysis, the TCA cycle, and amino acid metabolism which were supported by either the metabolomics data or literature from other toxicity studies (Beger et al. 2010; Kumar et al. 2012). There are opportunities for further curation of the network reconstruction to account for more metabolites and metabolic reactions, as well as further curation of the metabolomics data.

This study used transcriptomics data paired with metabolomics data to provide insight into the changes induced by these toxicants on hepatocytes. Protein fold changes could

be used in place of gene expression data and ultimately could have been used for TIMBR predictions because we can map such data to the metabolic reactions accounted for in the metabolic network reconstruction. As large data sets are made accessible or easy to collect, the use of multi-omic datasets to predict and validate modeling results becomes critical in interrogating specific phenotypes of interest for a chosen system. Our paired experimental and computational approach is one step towards characterizing the cellular response to a compound and identifying potential biomarkers indicative of cell state.

3.6 Tables

Table 3.1 Differentially expressed genes from chemical exposure of hepatocytes – Comparison of the number of differentially expressed genes in response to each chemical compound at the six-hour time point, compared to their respective controls.

Chemical Compound	Number of differentially expressed genes (FDR < 0.1)	Number of differentially expressed genes (FDR < 0.1) in the *iRno* model
APAP – 6 hours	7370	1009
CCl_4 – 6 hours	824	131
TCDD – 6 hours	2493	304
TCE – 6 hours	907	151

Table 3.2 Differentially changed metabolites from chemical exposure of hepatocytes – Comparison of the number of differentially changed metabolites for each chemical compound at the six-hour time point, compared to their respective controls.

Chemical Compound	Statistically significant metabolites changed	Subset of statistically changed metabolites identified
APAP – 6 hours	82	8
CCl_4 – 6 hours	11	3
TCDD – 6 hours	102	6
TCE – 6 hours	84	6

Chapter 4: Predicting Changes in Renal Metabolism after Chemical Exposure with a Genome-Scale Metabolic Model

4.1 Foreword

In this chapter, I present an updated version of the iRno reconstruction, that has expanded to capture more kidney functionality. Additionally, I also use iRno to along with the paired transcriptomics data and metabolomics data approach from chapter 3 to make predictions of changes in metabolism from drug-induced nephrotoxicity. This chapter is currently formatted as a manuscript titled “Predicting Changes in Renal Metabolism after Chemical Exposure with a Genome-Scale Metabolic Model” and to be submitted soon. With that, I would like to thank my co-authors Bonnie V. Dougherty, Venkat R. Pannala, Kalyan Vinnakota, Glynis L. Kolling and Jason A. Papin.

4.2 Introduction

The kidneys are vital organs responsible for many functions such as filtering blood, regulating water and electrolyte balance, and filtering waste from the body (Onopiuk et al. 2015; Scott and Quaggin 2015). Metabolism is a key biochemical process in the performance of these functions by (1) generating energy for filtering and reabsorbing metabolites back into the blood, and (2) breaking down fatty acids, amino acids, and other metabolites to be used by other organs (Weidemann and Krebs 1969; Bobulescu 2010). The kidneys can sustain consistent injury before a loss of function is observed, typically measured by increases in creatinine or urea to diagnose damage has occurred (Bellomo

et al. 2012). Understanding the genesis and progression of these diseases could be useful in investigating treatment strategies to help restore kidney function.

Computational models can be used to interrogate how a disease or specific condition affects kidney function. Computational models can simulate a biological system of interest to understand how perturbations can change system dynamics (Waikar and Bonventre 2009; Layton 2013; Sgouralis and Layton 2015). Genome-Scale Network Reconstructions (GENREs) provide a mathematical framework to represent the biochemical reactions and metabolites of a cell or organism to depict its metabolism (Rawls et al. 2018). GENREs have been used to represent bacterial species (Orth et al. 2011; Bartell et al. 2017; Carey et al. 2017), and more recently been to represent global changes in human and rat metabolism (Mardinoglu et al. 2014; Blais et al. 2017). These network reconstructions can be adapted to capture tissue or cell-type specificity to address key questions about underlying biological mechanisms or changes in phenotypes. To date, there have been a small number of models developed to represent kidney metabolism (Chang et al. 2010; Zhang et al. 2013; Sohrabi-Jahromi et al. 2016). However, these models are based on a previous reconstruction that has since been updated to capture more metabolic pathways and include more metabolites and metabolic reactions (Thiele et al. 2013; Swainston et al. 2016). Also, these models were focused on predicting drug effects, or focal segmental glomerulosclerosis (FSGS), and doesn't directly address drug toxicity, thus motivating a new application of a model of kidney metabolism.

It is possible to experimentally measure the changes in levels of metabolites in the blood or urine to assess the degree of damage that has occurred in damaged kidney tissue. Traditional measures of declining kidney function include measuring serum creatinine clearance (Himmelfarb and Ikizler 2007) and blood urea nitrogen (Gowda et al. 2010); however, these markers change under many conditions and aren't limited to just kidney-specific injury, highlighting the need for new markers of kidney function, or biomarkers, are needed to properly assess damage (Kim and Moon 2012). New protein biomarkers have been discovered (Dieterle et al. 2010; Adiyanti and Loho 2012; Bonventre 2014), but have shown inconsistent results in human studies (Endre et al. 2011; de Geus et al. 2012). One way to discover potential biomarkers is with the use of omics profiling data (Connor et al. 2010; Blanchet et al. 2011; Matheis et al. 2011; Zierer et al. 2015). Transcriptomics and metabolomics data are useful for characterizing global changes in mRNA expression and metabolite levels. Omics data has been used with GENREs to make predictions on how metabolism is altered after compound exposure (Agren et al. 2014; Stempler et al. 2014; O'Brien et al. 2015; Sawada et al. 2018).

Here, we present an updated network reconstruction (*iRno*) that has been expanded to include kidney function, collect and analyze paired transcriptomics and metabolomics data from compound treated kidney cells, and make predictions of changes in metabolite levels after compound exposure. The updated *iRno* reconstruction contains changes to existing biochemical reactions and the addition of several hundred new reactions. We also exposed primary rat renal proximal tubule epithelial cells (RPTECs) to five different compounds with varying nephrotoxic effects for six and twenty-four hours. From this

experiment, we collected paired transcriptomic and metabolomic data to characterize the response of the RPTECs to the different compounds. Lastly, we combined the transcriptomics data with *iRno* using the Transcriptionally Inferred Metabolite Biomarker Response (TIMBR) algorithm (Blais et al. 2017) to predict changes in metabolite levels based on the control and treatment conditions. Metabolites that change between the two groups could potentially serve as biomarkers of kidney injury. With this approach, we provide an updated GENRE expanded to include kidney-specific functionality, as well as a framework for determining novel biomarkers of nephrotoxicity.

4.3 Methods

4.3.1 Creation of Kidney-Specific Metabolic Tasks

Metabolic tasks are reactions or pathways that ensure the conversion of one metabolite to another, representing the known biological function of the organism of interest. Here metabolic tasks were created by reviewing the literature on rat kidney function, and more specifically RPTECs. Tasks were first taken from the existing *iRno* model that represented overlapping functions between the liver and the kidney. Next, literature was reviewed and tasks were added that came from published data on functions that occur in the rat kidney. Overall, a list of 155 tasks was created, and have been used in a previous publication (Pannala et al. 2019).

4.3.2 Flux Balance Analysis and the Expansion of iRno

To expand the *iRno* network reconstruction to more completely capture kidney metabolic function, literature was searched to find evidence of reactions known to occur but not

previously captured in the model. Additionally, metabolites and reactions were added to *iRno* that allow the model to secrete metabolites detectable in plasma. The list of changes resulted in the addition of two new metabolites, 87 metabolites that were newly assigned to different compartments, four new reactions, 89 transport reactions to bring metabolites into their newly assigned compartments, and 193 exchange reactions. *iRno* now accounts for the function of 5716 metabolites and 8532 reactions. A new metabolic objective function was created based on the previous one used (Blais et al. 2017) excluding bile acids. Gluconeogenesis was also used as an objective function since the kidneys synthesize a considerable amount of glucose for the body (Gerich et al. 2001). To simulate gluconeogenesis, glutamine, lactate, and glycerol are used as inputs and glucose was used as output was explored. These two objective functions were explored independently to interrogate kidney function. Flux balance analysis was used to simulate the flux through individual reactions using the cobra toolbox v.2.0.6 (Schellenberger, Que, et al. 2011) for MATLAB 2016b.

4.3.3 Renal Proximal Tubule Epithelial Cell Growth Conditions

Primary Renal Proximal Tubule Epithelial Cells (RPTECs) isolated from Sprague-Dawley rats purchased from Cell Biologics were grown on gelatin-coated wells using DMEM:F12 supplemented with penicillin/streptomycin, L-glutamine, 5% FSB, ITS (insulin-transferrin-selenium) and epidermal growth factor (EGF; 100µg/mL) without phenol red. RPTECs were plated at a density of 200,000 cells/well in a 12-well plate and cultured overnight before chemical exposure.

4.3.4 RPTEC Exposure Conditions

RPTECs were exposed to five chemical compounds at sub-toxic concentrations. The compounds and concentrations selected were acetaminophen (APAP) at 10mM, carbon-tetrachloride (CCl4) at 10mM, gentamicin (GENT) at 10mM, 2,3,7,8-tetrachlorodibenzodioxin (TCDD) at 1nM and trichloroethylene at 1mM. The concentrations and time points were selected based on previous concentrations for similar studies in rat, human, and mouse kidney cells (Smith 1988; Boogaard et al. 1989; Mugford 1997; Lash et al. 2001; Robbiano et al. 2004; Dong et al. 2010; Vrbová et al. 2016). Gentamicin is a primary nephrotoxicant. There is literature evidence to suggest that nephrotoxicity can occur in APAP, CCl4, and TCE (Newton et al. 1983; Cojocel et al. 1989; Mazer and Perrone 2008; Khan et al. 2009). TCDD has been shown to induce hepatotoxicity (Boverhof et al. 2006; Bentli et al. 2013); it was included in our experiments for comparison to previously published data and as a reference for the nephrotoxic-specific compounds described above.

4.3.5 RNA Isolation, Sequencing, and Analysis

RPTECs were exposed to the chemicals mentioned above. After exposure, the cells were lysed with Trizol to begin RNA extraction. Cell lysates were spun with chloroform in phase-lock gel tubes inside a cold room and the upper phase was then decanted into new tubes. Isopropanol and glycogen were added to the mixture and spun again resulting in an RNA pellet, which was washed with 75% ethanol twice. DNA was removed with a kit (Ambion/Invitrogen) and then RNA was quantified. rRNA was then depleted and mRNA was sent to a core facility (med.virginia.edu/gatc/about/) for library construction and

sequencing. RNA was sequenced in a 2x125bp pair-end (PE) configuration and fastq files were generated. Kallisto v 0.43.0 (Bray et al. 2016) was used to process raw fastq files to quantify transcript abundances in transcripts per million (TPM) under default settings. Transcript abundances were then aggregated to the gene level in R v. 3.5.1 with the package tximport (Soneson et al. 2015), and differential gene expression was calculated with DESeq2 (Love et al. 2014) with a significance threshold at FDR < 0.1.

4.3.6 Gene Enrichment Analysis

To further analyze the differentially expressed genes, we used the R package clusterProfiler (Yu et al. 2012) to find KEGG pathways that were enriched. Differentially expressed genes for each condition and time point were first sorted into lists of Entrez Gene ID numbers. Next, the list of Entrez Gene ID numbers was passed through the command enrichKEGG to find the relevant pathways in the KEGG database that were enriched. Pathways were considered enriched with an adjusted $p < 0.05$.

4.3.7 Metabolomics

Once RPTECs were treated, cell supernatants were collected and shipped to West Coast Metabolomics at the University of California, Davis (http://metabolomics.ucdavis.edu/) for metabolomics analysis by their Core Facilities. At West Coast Metabolomics, samples were processed for untargeted analysis of primary metabolites by Gas Chromatography Mass Spectrometry (GC-MS), analysis of complex lipids via Liquid Chromatography Mass Spectrometry (LC-MS), and biogenic amines through Hydrophilic Interaction Chromatography Quadrupole Time of Flight (HILIC-QTOF) Mass Spectrometry. Both external and internal standards for quality control were prepared and analyzed along with individual samples.

Primary metabolites were analyzed using a previously published protocol (Fiehn 2016) and results were reported by the relative peak intensities at the specified mass/charge retention index. For lipid analysis, samples were prepared with methanol, methyl tert-butyl ether (MTBE), and water before running LC-MS; peak intensities were reported following a published protocol (Cajka and Fiehn 2017). Biogenic amines were prepared by separating polar hydrophilic small molecules from lipids, according to a previously established method (Matyash et al. 2008), and raw peak intensities were then reported from analyzed samples using previously published protocols (Meissen et al. 2015). Relative peak intensities of identified and unidentified metabolites were generated together. Peak intensities were normalized and then analyzed by background subtraction, log transformation, centering the data around zero, and then Pareto-scaled within each metabolite. Normalized values were then subtracted from the fresh media samples to give a direction of change, either consumption or production. Data analysis was performed using R v 3.5.1. Metabolites were considered statistically changed from control if BH-adjusted p-values from the Kruskal-Wallis test, run in the FSA package (Ogle et al. 2019), were below a threshold ($p < 0.1$).

4.3.8 TIMBR Algorithm

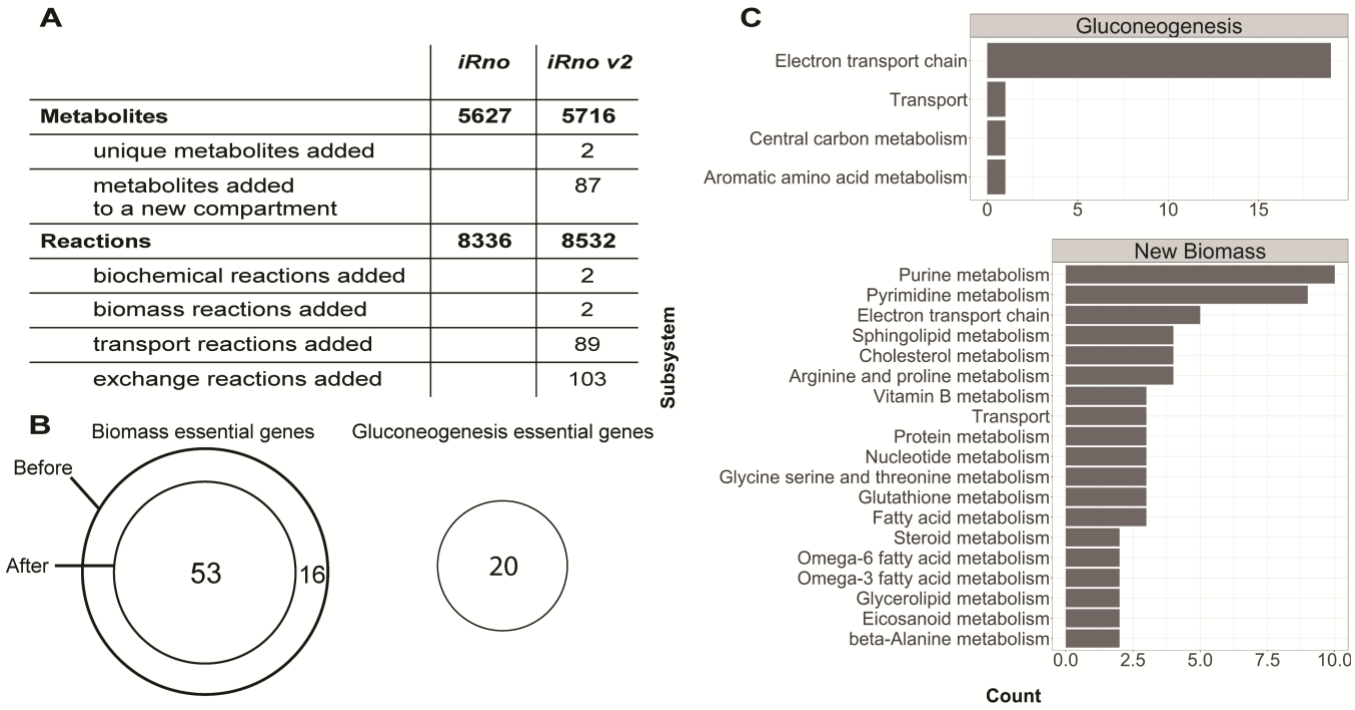
The TIMBR algorithm (Blais et al. 2017) uses transcriptomics data and a GENRE to make predictions on the relative production levels of metabolites. The default weight is assigned to each reaction based on the type of reaction it is (biochemical, boundary, transport, etc.). For reactions associated with differentially expressed genes, the log2fold changes are then multiplied by default reaction weights to get the final reaction weights. Raw

production scores for control and treatment conditions were calculated by minimizing the sum of the product of the final reaction weights and flux through each reaction, across all reactions. Production scores for the control condition and the treatment condition were then combined using the previously described formula (Blais et al. 2017) to determine the relative production of a metabolite. These production scores were then z-transformed and used for downstream analyses. A full description of the method is available (Blais et al. 2017) and the source code to run the algorithm can be found on GitHub (www.github.com/csbl/ratcon1).

4.4 Results

4.4.1 Updating iRno to Reflect Kidney-Specific Metabolic Function

We expanded *iRno* by reviewing literature for evidence of metabolic reactions we had not yet captured and added metabolites that participated in these reactions. One reaction was the L-glutamate:2-aminobutanoate gamma-ligase reaction, converting glutamate and (S)-2-aminobutanoate to γ-L-glutamyl-L-alpha-aminobutyrate. Another reaction was the conversion of gamma-L-glutamyl-L-alpha-aminobutyrate to opthalmate, which was previously confirmed as a by-product of glutathione metabolism (Soga et al. 2006). Additionally, we also found evidence of metabolites that were detected in the plasma, so we updated *iRno* to reflect this change. To incorporate the secretion of the new metabolites, we added transport reactions to move metabolites from the cytosol to the extracellular compartment, as well as exchange reactions to move the metabolites from the extracellular compartment to the external environment. This curation effort resulted in the addition of 89 metabolites, 196 reactions, and eleven changes to existing reactions (**Figure 4.1A**).



	iRno	*iRno v2*
Metabolites	**5627**	**5716**
unique metabolites added		2
metabolites added to a new compartment		87
Reactions	**8336**	**8532**
biochemical reactions added		2
biomass reactions added		2
transport reactions added		89
exchange reactions added		103

Figure 4.1. Summary of changes to *iRno*
A table of model statistics from the first version of *iRno* to this update termed *iRno* v2 (**A**). The number of additions and changes for both metabolites and reactions are displayed. The Venn diagram shows changes in the number of genes that are required to produce biomass before the update, after the update, and with the addition of a gluconeogenesis objective function (**B**). Bar charts show the subsystem classification from the model of the genes that are required for gluconeogenesis and biomass.

We created a new biomass equation that was derived from the previously published biomass equation for *iRno* (Blais et al. 2017) with the removal of bile acid related terms since given their specificity to hepatocyte function which was the focus of the previous work. For kidney-specific functionality, we also tested gluconeogenesis from lactate, glutamine, and glycerol as objective functions using previously defined physiological constraints (Elhamri et al. 1993; Pannala et al. 2019). After exploring biomass and gluconeogenesis as objective functions, we simulated gene knockouts to determine

genes that are required for flux through both of these reactions. This analysis generated a list of genes that are necessary to be active in order for biomass and glucose to be produced. Removing bile acids from the biomass equation resulted in the removal of 16 genes from the list of genes required to achieve flux through the biomass equation (**Figure 4.1**). For gluconeogenesis, 20 genes were required to be active for glucose to be synthesized from lactate, glutamine, and glycerol (**Figure 4.1)**. It was interesting to note that none of the reactions associated with these genes were required to carry flux to produce biomass. The difference in required active genes for biomass and gluconeogenesis demonstrates the range of function captured with the reconstruction. Once we finished expanding *iRno*, we used the model to predict how kidney metabolism would be altered after compound exposure.

4.4.2 Transcriptomics Reveal Kidney-Specific Response to Injury

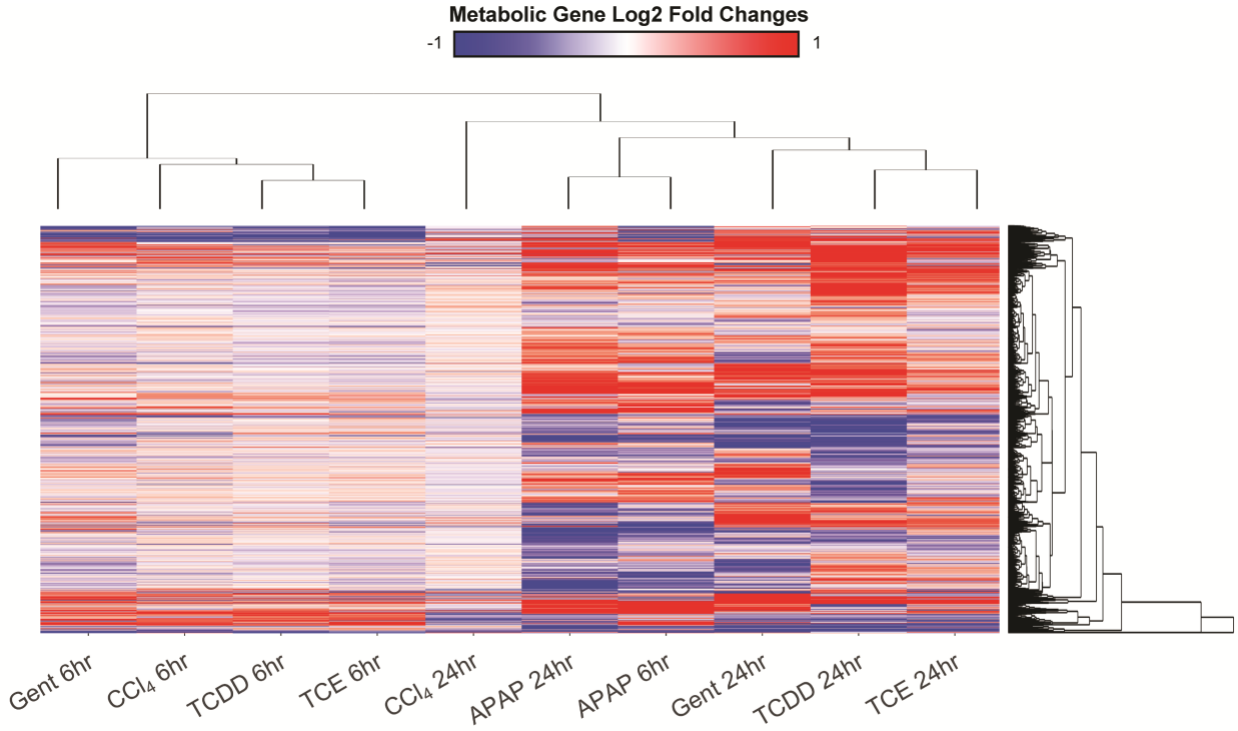


Figure 4.2. Heatmap of gene expression changes
The heatmap above shows the fold changes for differentially expressed genes that can be found within the *iRno* model for any gene that was significant in at least 1 of the 10 conditions. Blue shows log2 fold changes less than 0 (downregulated) while red shows log2 fold changes greater than 0 (upregulated). Genes are clustered using Euclidean distance and with complete linkage, while conditions are clustered using the distance of Spearman correlations, with complete linkage.

After exposing renal proximal tubule epithelial cells (RPTECs) to Acetaminophen (APAP), Carbon tetrachloride (CCl_4), Gentamicin (Gent), 2,3,7,8-tetrachlorodibenzodioxin (TCDD), and Trichloroethylene (TCE) for six and twenty-four hours, we measured the cell's response at the mRNA level. **Table 4.1** provides a list of the number of differentially expressed genes from each condition and time point. The third column shows the number of differentially expressed genes that we consider to be metabolic, mapping to genes accounted for in *iRno*. A summary of changes for metabolic genes is displayed in **Figure 4.2.** Within each condition, the highest number of DEGs was

seen at 24 hours except for the CCl_4-exposed condition. We also noticed several genes changed similarly in the 24-hour condition (**Figure 4.2**) so we decided to take a further look. There were 370 genes that were upregulated, and 258 genes that were downregulated across the APAP, TCDD, Gent, and TCE conditions. Among this group were genes relating to amino acid synthesis such as *Asns*, and *Thnsl1*, as well as amino acid transport into the cell like *Slc7a5*, *Slc16a10*, and *Slc16a17*. In addition to amino acid metabolism, genes relating to fatty acid metabolism were also increased. This includes genes relating to fatty acid elongation including *Elovl4*, *Elovl5*, and *Acsl1*. For the downregulated genes, a number of them are involved in mitochondrial processes, such as cytochrome c oxidase (*Cox6c*, *Cox6b1*), and NADH dehydrogenase (*Mt-nd5*, *Mt-nd3*, *Mt-nd1*), which suggests an impairment in energy production in the cell. The upregulation of fatty acids and amino acid processes highlights the importance of these two pathways in kidney metabolism.

While exploring the results from the transcriptomics data, we also looked at specific genes that map to enzymes known to be markers of kidney injury to validate that we see intoxication at our chosen concentrations. The first gene we looked at was for the enzyme clusterin (*Clu*), which has been shown to play a role in oxidative stress in kidney cells (Trougakos and Gonos 2006). We saw an increase in differential expression in the 24-hour conditions for APAP, Gent, TCDD, and TCE, as well as the six-hour condition for APAP. This increase suggests that the RPTECs are undergoing some form of oxidative stress after compound exposure at the concentrations we tested. We also observed increases in the *Slc22a5* gene, another marker of kidney injury, responsible for the tubular

reabsorption of carnitine, an important molecule for beta-oxidation of fatty acids in the mitochondria in RPTECs (Kato et al. 2006; Sugiura et al. 2008: 1). We saw increases in this gene's expression in the APAP and Gent conditions at both time points, as well as the TCDD 24-hour condition. Lastly, we saw an increase in Lipocalin 2 (*Lcn2*), which is an early marker of kidney injury in TCE and TCDD at 24 hours, but a decrease in the APAP and Gent conditions. From the changes in *Clu*, *Slc22a5* and some of the changes in *Lcn2*, we were able to identify trends consistent with the literature on the expression of particular biomarkers of kidney injury and confirm that exposure to these compounds significantly perturbs metabolism of the RPTECs. We next wanted to identify common pathways enriched as a result of chemical exposure.

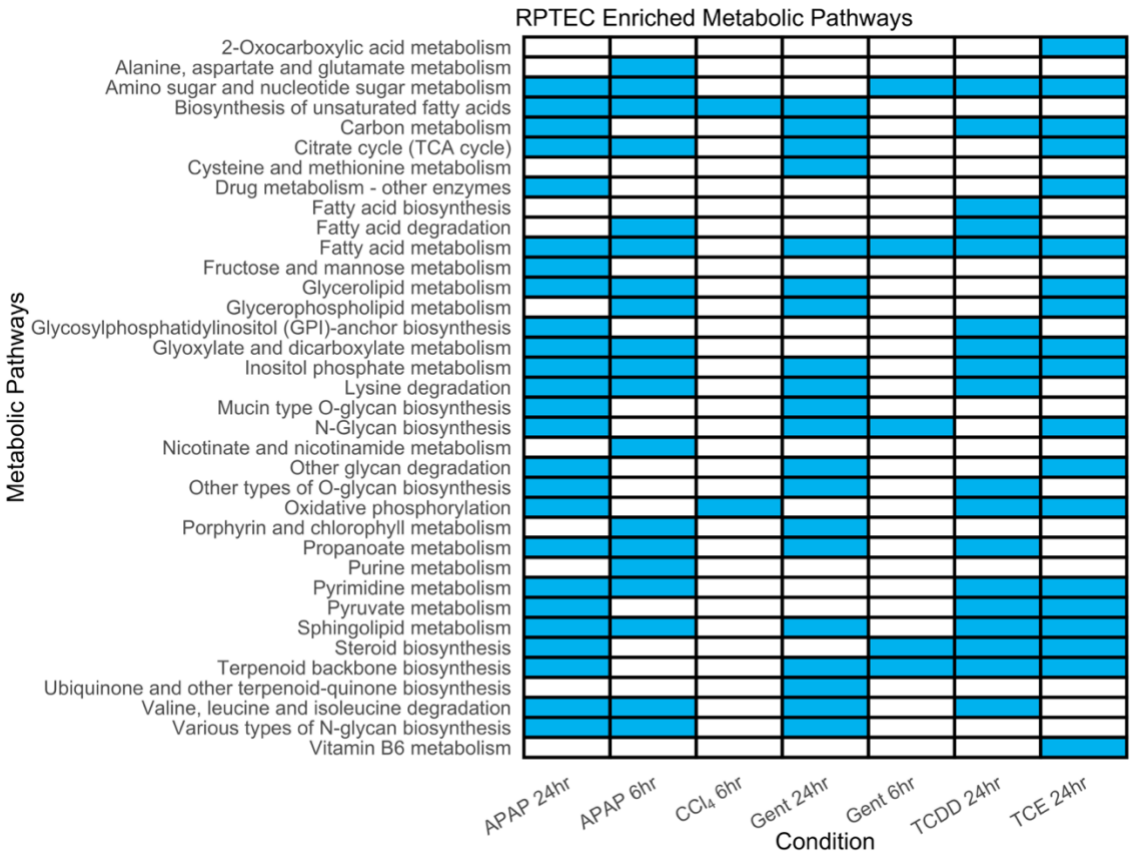


Figure 4.3. Heatmap of Enriched Metabolic Pathways
For the conditions for which there were enough differentially expressed genes, KEGG pathway enrichment is shown with blue indicating that a pathway is enriched, and white showing that pathways were not enriched for the condition. A select number of KEGG metabolic pathways are displayed on the y-axis, with experimental conditions listed on the x-axis.

4.4.3 Pathways Enriched from Compound Exposure

Using the data for metabolic genes differentially expressed in each condition, we examined enriched pathways using the clusterProflier R package (Yu et al. 2012) to identify any general or condition-specific responses of the RPTECs to treatment**.** Oxidative phosphorylation was enriched in CCl_4 at six hours, as well as APAP, TCDD, and TCE at 24 hours. In the CCl_4 six-hour and TCE 24-hour conditions, genes in this pathway were decreased, suggesting an impairment in the production of energy. In the

APAP and TCDD 24-hour conditions, there was a mix of responses with some portions of the pathway upregulated, and other portions downregulated. While oxidative phosphorylation is necessary for energy production given the demands of active transport of the kidney, reactive oxygen species are produced, and this increase in ROS production contributes to the oxidative stress of the RPTECs after exposure (Baud and Ardaillou 1986; Ratliff et al. 2016). Pathways associated with lipid and fatty acid metabolism were enriched in most conditions (**Figure 4.3**), signifying that alteration of fatty acid metabolism could be a common response of the kidney cells to various chemical exposure(s). Pathways associated with branched-chain amino acid metabolism came up as enriched in the APAP condition at both time points, as well as Gent and TCDD at 24 hours. However, these pathways were not enriched in the CCl_4 conditions, indicating that branched-chain amino acid metabolism could be critical in the different responses of RPTECs to the tested compounds. In the APAP 24-hour condition, the aldosterone-regulated sodium regulation pathway was enriched, while in Gent the ATP-binding Cassette transporter pathway was enriched. Both of these pathways are important to kidney metabolism and transport function, as ATP is needed to transport metabolites across the membrane, while aldosterone can trigger the reabsorption and excretion of sodium and water (Spitzer 1982). At the pathway level, we observed that energy metabolism, amino acid metabolism, and fatty acid/lipid metabolism were enriched across all categories suggesting that these are key metabolic processes altered by compound exposure. The ability of renal cells to uptake or secrete amino acids, glucose, or fatty acids could be indicative of kidney dysfunction. From this result, we next wanted to independently look at the metabolomics data and see which metabolites and pathways

were altered and how the pathways identified as different with the metabolomics data compare to pathways that were enriched as identified in the transcriptomics data.

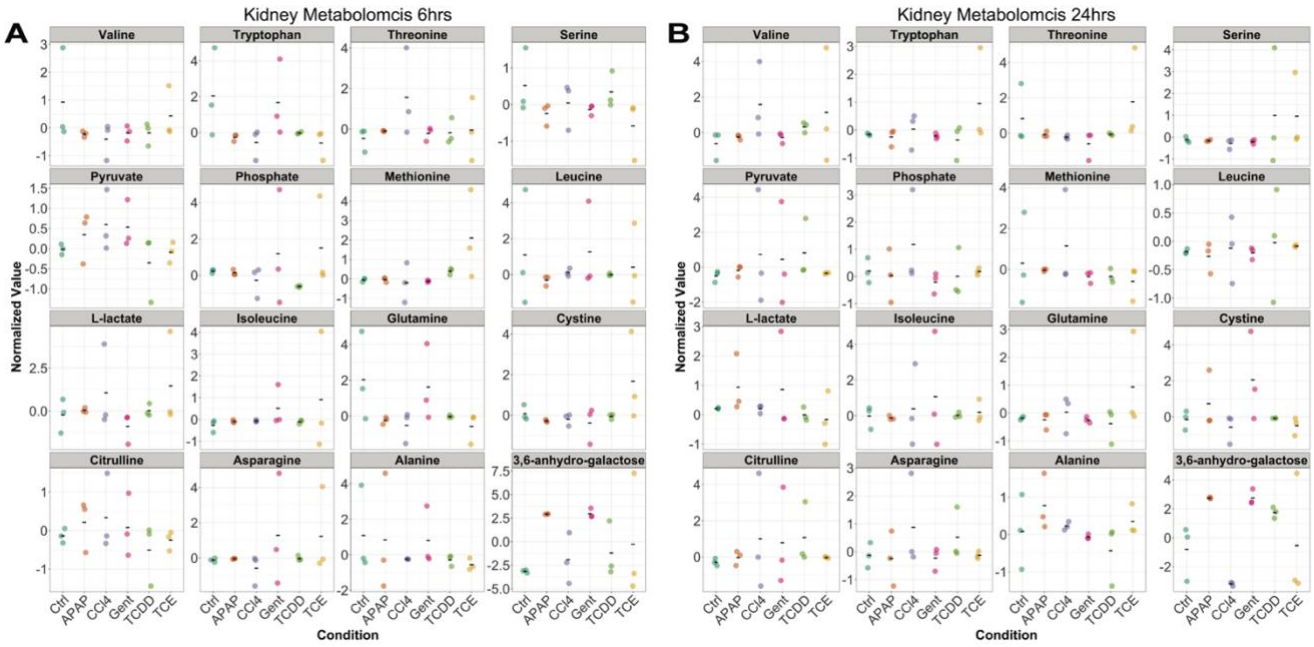


Figure 4.4: Dot plots of Metabolomics Data
The dot plots above show the results from the metabolomics data after normalization for both six (**A**) and 24-hr (**B**) hours. Each condition at both time points was done with n=3, and data has been normalized to blank media to show positive values meaning a metabolite was produced, and negative values meaning a metabolite was consumed. The sixteen metabolites shown are those that were statistically changed in one of the five experimental conditions compared to blank media or the DMSO control samples. Median values are indicated with (**-**)

4.4.4 Metabolomics Data Reveals Compounds Cause Oxidative Stress on Renal Cells

We computed the total number of detected metabolites that were statistically changed for each condition (**Table 4.2**), comparing all conditions against each other and discovered that TCDD exposure at 24 hours resulted in the most changes of the five compounds, while APAP exposure at 24 hours produced the next largest change. Metabolite data were

normalized and summarized as presented in **Figure 4.4**. Each identified metabolite, for which there was a statistically significant change for blank or control, is shown in the plot, with the RPTECs exposure condition on the x-axis, and the normalized metabolite levels on the y-axis. The levels of each metabolite in the media are also normalized to the "Blank" media condition in **Figure 4.4,** where positive values indicate a metabolite has been produced, while negative values indicate a metabolite was consumed.

Across all conditions, 16 identified metabolite levels changed with statistical significance from the control condition or the blank media condition. Of these 16 metabolites, 12 are amino acids, and two are related to central carbon metabolism. There is a larger number of amino acid changes in the TCDD condition, primarily where there is an increase in these metabolites (**Figure 4.4**). Branched-chain amino acids are a major source of energy for the kidneys through oxidation during a fasted state (Adibi 1976) indicating that metabolite levels should be lower. However, we saw increases in the levels of these metabolites after TCDD exposure (**Figure 4.4**), which is consistent with existing literature on compound exposure of kidney cells to different chemicals (Macpherson et al. 1991; Ghiculescu and Kubler 2006; Portilla et al. 2006; Niemann and Serkova 2007; Xu et al. 2008) potentially due to decreased reabsorption by the RPTECs.

3,6-anhydro-D-galactose is a metabolite produced from D-galactose, which is in the polysaccharide porphyran. Porphyran has been shown to have antioxidant effects that protect the kidneys from oxidative stress (Wang et al. 2017). Given that we see increases in 3,6-anhydro-D-galactose in the supernatant, we can hypothesize that there was a

breakdown of galactose in the kidneys, and that they are trying to maintain homeostasis by protecting against the oxidative stress that comes from acetaminophen-induced and gentamicin-induced nephrotoxicity (Weinberg et al. 1980; Weinberg and Humes 1980; Banday et al. 2008; Narayana 2008; Canayakin et al. 2016). Overall, the metabolomics data show the changes in metabolism associated with oxidative stress and amino acid metabolism. To investigate further, we looked at the transcriptomics data and metabolomics data together to see if there were consistent changes between the datasets we could identify.

4.4.5 Transcriptomics and metabolomics data suggest different responses to chemical exposure.

After separately analyzing the transcriptomics data and the metabolomics data, we then interrogated the data together to identify consistencies and inconsistencies. In the TCDD and TCE conditions at 24 hours, the number of amino acids was increased as measured by the metabolomics data. To connect this observation back to the gene expression data, we identified reactions from *iRno* that contain these metabolites and have genes associated with them. We then filtered this list of reactions based on significant differentially expressed genes from the transcriptomics analysis. Increased levels of the branched-chain amino acids (valine, leucine, and isoleucine), were observed in the media after treatment, with an increase in gene expression contributing to the breakdown of these metabolites. The gene *Bcat1*, which is the branched-chain amino acid transaminase, is responsible for converting alpha-ketoglutarate and valine, leucine, or isoleucine, to their respective branched-chain keto acid and glutamate (Bixel et al. 1997).

The increased expression of this gene signifies that these amino acids, in particular, are broken down by the cells. Other genes in this pathway (*Bckdha*, *Acads*, *Echs1*, *Auh*, *Hadh*) were also increased in expression, further agreeing that the kidney cells are breaking down these amino acids. Conversely, several genes associated with branched-chain amino acid transport were increased in expression, which explains the increased levels of valine, leucine, and isoleucine in the spent media.

The argininosuccinate synthetase gene *Ass1* is responsible for argininosuccinate production from citrulline in the pathway that converts citrulline to arginine in the proximal tubules (Dhanakoti et al. 1990). In the transcriptomics data, we saw increased expression of this gene but measured a statistically significant increase in citrulline as a result of TCDD exposure. This observation suggests that the kidneys were appropriately consuming citrulline, however, it was not clear why there was an increase in citrulline measured in the media. Another gene, nitric oxide synthase 3 (*Nos3*), was increased in response to treatment and is responsible for the conversion of 2-N-(omega)-hydroxyarginine to citrulline and nitric oxide, which is captured by our model. The increased expression of *Nos3* could be responsible for the increase in citrulline found after TCDD exposure for 24 hours. Additionally, nitric oxide has been shown to have an inhibitory effect on sodium reabsorption (Mount and Power 2006), which contributes to the idea of kidney function declining after compound exposure likely due to stress. With both supporting and contrasting results from the transcriptomics and metabolomics data, we then wanted to use our model to make predictions of changes in metabolite levels, to

see if the model could explain some of the metabolite changes from the transcriptomics data.

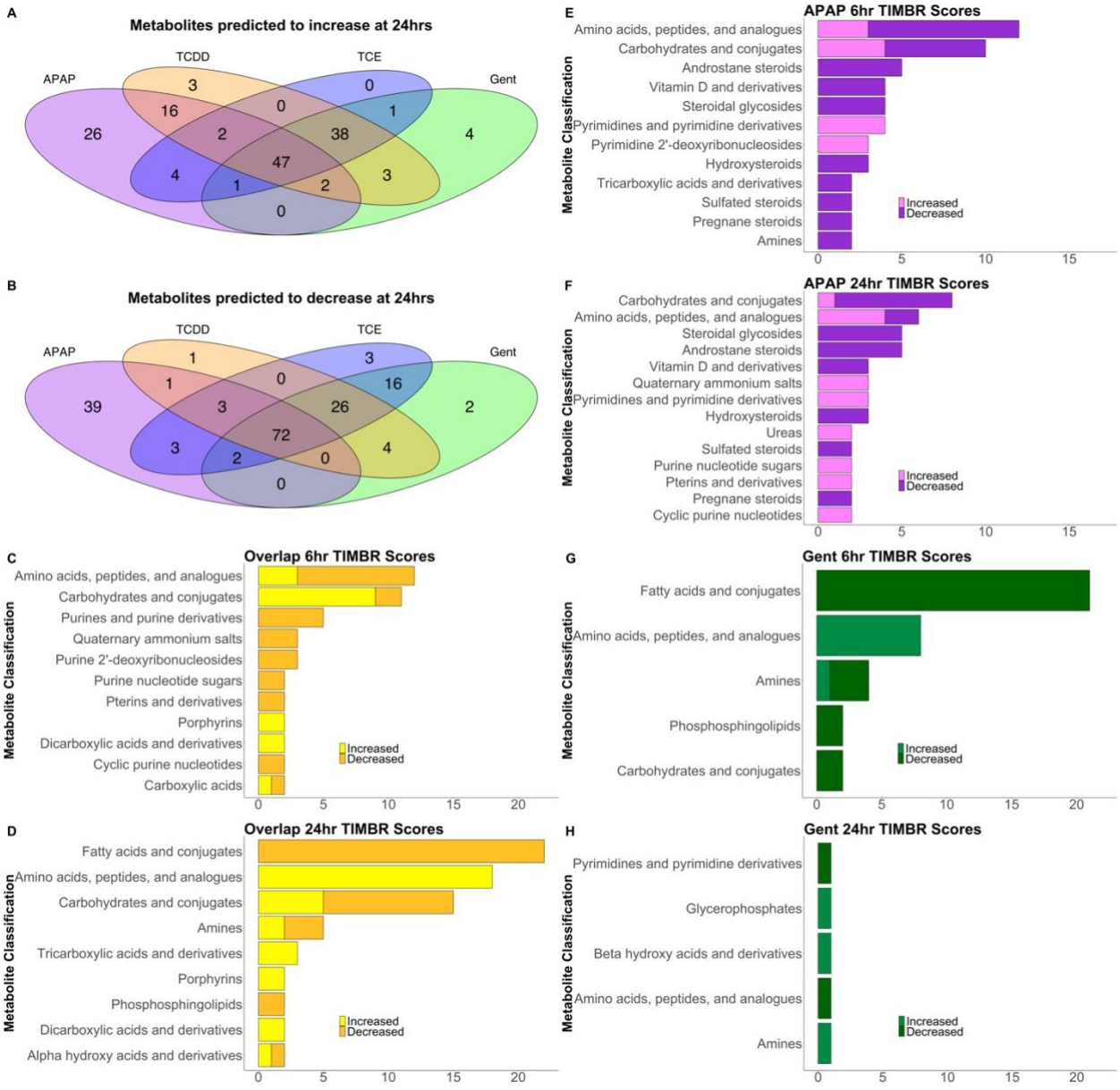


Figure 4.5: Summary of TIMBR production scores
Venn diagrams show the number of metabolites that were predicted to increase (A) or decrease (B) after exposure to APAP, Gent, TCDD, or TCE after 24 hours. Bar charts show the Human Metabolome DataBase classifications for metabolites that overlap, APAP, and Gent for both six hours (C, E, G) and 24-hours (D, F, H).

4.4.6 iRno Predicts Changes in Metabolites Observed by the Omics Data

To make predictions on metabolite level changes, we used our updated *iRno* model and the transcriptomics data along with the Transcriptionally Inferred Metabolite Biomarker Response (TIMBR) algorithm (Blais et al. 2017). TIMBR overlays gene expression fold changes and applies corresponding weights to metabolic reactions in the model. For each metabolite that can be produced in the model, TIMBR then minimizes the sum of fluxes for a treatment and a control case, to determine a metabolite's availability to be produced from the model. Overall, we were able to make TIMBR predictions for APAP, CCl_4, and Gent at the six-hour timepoint, and APAP, Gent, TCDD, and TCE at the 24-hour timepoint for a total of 245 metabolites, based on differentially expressed genes at each of these conditions.

Figure 4.5 shows the number of metabolites predicted to increase or decrease at the 24-hour timepoint**.** There were 24 metabolites predicted to increase in production and 45 metabolites predicted to decrease in production across APAP, CCl_4, and Gent at six hours including the production of fructose, glucose-1,6-bisphosphate, and 3-phospho-D-glycerate. These three metabolites are intermediates in the glycolysis and gluconeogenesis pathways, and an increase in these intermediates suggests that metabolites are not being converted across the full pathway, or the metabolites are being secreted. This observation could indicate a common response of kidney cells to these compounds to either lose energy production due to a breakdown in glycolysis or inhibit glucose synthesis, which is important for proximal tubule epithelial cells (Burch et al. 1978; Mather and Pollock 2011). Of the group of 45 metabolites that were predicted to

decrease in production at six hours, amino acids and analogs such as arginine, argininosuccinate, creatinine, and serine were included. Additionally, some nucleosides and nucleotides were predicted to decrease in production, signifying that changes in these pathways could be important in response to injury.

At the 24-hour timepoint, increases in 47 metabolites and decreases in 72 metabolites were predicted to decrease in production across all conditions (APAP, Gent, TCDD, and TCE). For this timepoint, amino acids were predicted to increase in response to treatment, which was contradictory to what we saw at the six-hour timepoint. One example of this difference is with serine, as it was predicted to increase at 24 hours and was predicted to decrease at six hours. The increase in serine could support the idea that gluconeogenesis is inhibited, where the increase in serine could be due to the conversion of glutamine and glutamate to serine (van de Poll et al. 2004). At the 24-hour timepoint, fatty acids and conjugates were the largest groups predicted to decrease across all compounds. The decrease in fatty acids and conjugates could potentially point to the use of fatty acids as an energy source during stress, instead of amino acids which were predicted to increase across all compounds at 24 hours. The proximal tubules mostly reabsorb filtered amino acids (Dantzler and Silbernagl 1988), so this predicted increase in amino acid production could be an indication that the kidney is not properly breaking down or taking up amino acids.

We next looked at compound-specific responses to see if there were other aspects of kidney metabolism that were altered. At the six-hour timepoint, APAP was predicted to

cause a decrease in the production of amino acids and analogs and an increase in the production in carbohydrate compounds contrary to what we predicted at the 24-hour timepoint. For Gent at the six-hour time-point, amino acids and analogs were predicted to increase, similar to the general trend we predicted at the 24-hour timepoint. Fatty acids were predicted to decrease, which also agrees with the general trend we noticed at the 24-hour timepoint. Fatty acid oxidation inhibits glycolysis and increases enzyme activity of gluconeogenesis (Owen et al. 1969), so this decrease further supports the breakdown of the production of glucose by allowing glycolysis to proceed. These results highlight the utility of using the model to further investigate the changes in metabolism measured by transcriptomics data and predicting how metabolism proceeds in proximal tubule cells in response to xenobiotics.

4.5 Discussion

The kidneys are highly metabolically active (Gallagher et al. 2006). While we have a general understanding of kidney function, more information is needed on exactly how they lose function as well as indicators of declining function. One advantage of computational models is that they can be used to investigate biological changes or emergent phenomena of a particular biological system that could arise from perturbations to the surrounding environment. GENREs have emerged as useful tools to help point specifically to genes or proteins of interest associated with observed phenotypic changes or to characterize the overall response to an altered state of the system. To date, GENREs that represent kidney function have been limited in size and coverage of

metabolic pathways, and some networks have been created to study a particular disease. Here, we present an update to an existing model of rat metabolism (Blais et al. 2017) that is validated for kidney function. Additionally, we profiled the metabolic responses of RPTECs exposed to a broad range of chemicals of different classifications and used our model to explain some of the measured metabolic changes. This analysis demonstrated the utility of our model, by predicting changes in metabolism that were compared with experimental data.

We have added new objective functions to *iRno* to include a generic biomass function for non-liver cells and to capture functions specific to kidney metabolism. First, we used the model to simulate gluconeogenesis from lactate, glutamine, and glycerol precursors. Additionally, we looked at genes necessary for gluconeogenesis from the available precursors and found that 22 genes related to either ATP synthase or cytochrome c oxidase were necessary to produce glucose. Since gluconeogenesis requires ATP for completion (Ross et al. 1986), the cytochrome c oxidase genes would also be necessary as they are involved in the generation of ATP (Fontanesi et al. 2006). For the new biomass equation, we found that sixteen genes were no longer required to synthesize biomass. These genes were from the Cyp450 family (e.g., *Cyp27a1*, *Cyp3a18*) and also included genes that are involved in the reduction or oxidation of steroids and fatty acids (e.g., *Acaa1*, *Akr1c14*, *Akr1d1*, *Amacr*), which are necessary for bile acid synthesis (Chiang 2013; Šarenac and Mikov 2018).

Among the differentially expressed genes for each condition, the gene *Slc22a5 (Octn2)* was increased in the APAP and Gent conditions at both time points, as well as the TCDD condition at 24 hours. *Slc22a5* is a solute carrier that mediates both sodium independent and dependent carnitine transport. Carnitine is important because it is required to transport fatty acids into the mitochondria for beta-oxidation (Longo et al. 2006; Tamai 2013), which is a primary energy source for kidney cells (Le Hir and Dubach 1982; Braissant et al. 1996). The increased expression of this gene likely indicates increased oxidation of fatty acids, to provide relief to the kidney cells from oxidative stress (Nakamura et al. 2008).

From the metabolomics data, we noted increases in amino acid levels in response to TCDD and TCE treatment after 24 hours, but smaller changes after APAP or Gent exposure. This observation contrasts with existing literature that shows large changes in amino acid levels for both APAP and Gent. One *in vivo* study found increases in serine, threonine, valine, and leucine measured from urine samples in response to gentamicin exposure in rats (Boudonck et al. 2009). This study also measured increases in lipids from the rat urine samples as well (Boudonck et al. 2009). In addition to these changes, other *in vivo* studies using different compounds have noted a decrease in extracellular TCA cycle metabolites, and increases in lactate and amino acids (Lenz et al. 2004; Park et al. 2009). While the metabolomics data in our study, does not directly agree with these results, the differences seen here could be attributed to differences between *in vivo* and *in vitro* experimental systems. Amino acids were consistently predicted to be increased as a result of treatment at the 24-hour timepoint across all conditions, and in Gent at six

hours. For APAP, TCA cycle metabolites increase at the six-hour timepoint. This difference could be due to other studies not looking at the early exposure of compounds to the renal system, or differences between compounds used to assess nephrotoxicity.

One limitation of this study is that there is a small number of metabolites that overlap between the metabolomics data and the model. While we can still learn about general trends in metabolism from the paired omics datasets, this study could go further by validating predictions of changes in the metabolite levels at six hours to support studying changes in kidney metabolism earlier than 24 hours. Another limitation of this study is that *iRno* can be further curated to include more genes and metabolic reactions to capture more areas of metabolism. As our knowledge of kidney biological function expands, the model will also expand to better account for rat kidney metabolism. Currently, we are successfully able to capture some changes in kidney metabolism and to use the model to make predictions on the changes in metabolism associated with exposure to nephrotoxic compounds. The ability to capture theses changes in kidney metabolism is useful for understanding indicators of declining kidney function, but could also be further developed as a method to identify candidate drugs with potential nephrotoxic effects. Our approach for using a GENRE with paired omics data provides a holistic approach to identify and understand declining kidney function.

4.6 Tables

Table 4.1 Differentially expressed genes from chemical exposure of renal proximal tubule epithelial cells – Differentially expressed gene counts from proximal tubule epithelial cells after exposure to APAP, CCl_4, Gent, TCDD, and TCE for six and twenty-four hours. Each treatment condition is compared to DMSO controls.

Chemical Compound	Number of differentially expressed genes (FDR < 0.1)	Number of differentially expressed genes (FDR < 0.1) accounted for in the *iRno* reconstruction
APAP – 6 hours	7446	1003
APAP – 24 hours	8797	1212
CCl_4 – 6 hours	810	83
CCl_4 – 24 hours	0	0
Gent – 6 hours	2667	361
Gent – 24 hours	8770	1157
TCDD – 6 hours	3	0
TCDD – 24 hours	9548	1269
TCE – 6 hours	7	1
TCE – 24 hours	6543	924

Table 4.2 Differentially changed metabolites from chemical exposure of renal proximal tubule epithelial cells – Table of the number of metabolite levels that were differentially changed compared to DMSO controls, and how many of those metabolites are accounted for in the *iRno* network reconstruction. Changes in metabolite levels are considered statistically significant if the normalized value was different than DMSO controls at the BH-adjusted p-value < 0.1.

Chemical Compound	Statistically significant changes in metabolite levels	Subset of statistically significant changes in metabolite levels that map *to iRno*
APAP – 6 hours	2	0
APAP – 24 hours	7	0
CCl_4 – 6 hours	0	0
CCl_4 – 24 hours	0	0
Gent – 6 hours	1	0
Gent – 24 hours	0	0
TCDD – 6 hours	2	1
TCDD – 24 hours	11	8
TCE – 6 hours	0	0
TCE – 24 hours	3	3

Chapter 5: Conclusions and Future Directions

5.1 Conclusions

Completion of this dissertation research has resulted in the creation of several public resources and tools to advance the field of metabolic modeling and drug metabolism. First, I published a simple metabolic network, *iSIM*, as a teaching tool for those learning how to work with Genome-Scale Metabolic Network Reconstructions (GENREs). The paper for *iSIM* explains how to use metabolic models and how to run basic analyses to interpret how biological processes change across environments (Rawls et al. 2018). In addition to *iSIM*, we also provided code in MATLAB, R, and Python that has been well commented so that a novice may be able to step into computational biology. In chapters three and four, I provided a pipeline for analyzing paired transcriptomics data and metabolomics data, and combine this approach with an existing algorithm to make predictions of biomarkers of liver and kidney toxicity. This approach provides a method to not only make predictions of metabolite changes but also provides a means to explore reasons metabolite production levels have shifted. Using the paired metabolomics data, predicted changes in metabolite levels can be validated from the same samples from which the experimental data was generated. Lastly, we expanded the *iRno* reconstruction to capture kidney metabolism in addition to liver metabolism. This included adding new metabolic reactions and transport reactions as well as simulating different kidney functions with the network. The creation of *iSIM*, the updated *iRno* reconstruction, and the paired omics data approach for predicting biomarkers are the contributions I have made to the field to overall make an impact on the scientific community.

This work has made an impact not only on the metabolic modeling field but also in the broader scientific community. The *iSIM* paper exists to create interest among those in the programming field in biology and vice versa. It is important to me to make science accessible to those who may not have access or those who are unaware that this kind of work exists. That was the motivation for constructing this manuscript as a tutorial, and lead readers step by step through the biology with freely available code. Additionally, I provided a paired transcriptomics data and metabolomics data approach with metabolic models for the scientists in the toxicology field. For those who are more familiar with drug metabolism, my second study allows for more accessible entry to the field of computational biology. As the metabolic modeling community moves toward capturing the metabolism of the entire human body, my third publication extends a reconstruction that has been validated for liver metabolism (*iRno*), to cover kidney metabolism. This reconstruction is unique as it has a companion model of human metabolism, *iHsa*, to look at species-specific differences between the two organisms. My curation of *iRno* moves the field closer to having full multi-compartment, species-specific models of metabolism that includes all metabolic functionality in an organism, which can then be used for a broad range of applications. In general, my dissertation work has made significant contributions to disparate areas of human-associated systems biology and allows for greater dissemination of these principles to new groups of scientists.

5.2 Future Directions

In chapter 4 we successfully characterized the kidney cells' response to different toxicants and showed that the proximal tubule cells experienced some form of oxidative stress. Another way the kidneys can experience oxidative stress is through the generation of

reactive oxygen species (ROS), which wasn't fully captured through our experiments. One of the unanswered questions about kidney function is the effect of ROS on the reabsorption of sodium in the proximal tubules (Gonzalez-Vicente et al. 2019). Sodium reabsorption is important because it helps drive the uptake of other solutes in the kidney and can contribute to changes in blood pressure (Wang et al. 2009). Conflicting information has been presented on whether or not ROS contributes to or inhibits sodium reabsorption. This idea has been primarily explored in the context of toxicity, but not much information is presented on the effects of ROS on healthy proximal tubule cells. One potential future direction of this project would be to explore if ROS inhibit or increase sodium reabsorption from proximal tubule cells *in vitro* and *in silico*. An *in vitro* approach would be useful for collecting proteomics or transcriptomics data to assess how the expression of many sodium transporter enzymes changes from exposure to varying levels of ROS. Additionally, we can use the *iRno* model to explore how increased ROS production affects other kidney functions. Computationally, further curation of the *iRno* model would be required to ensure ROS production is captured before exploring this problem.

After measuring the impact of ROS on kidney cells, the next step would be to look at the effects of antioxidants on mediating oxidative stress in kidney cells with the model. It has been shown *in vitro* and in animal models that antioxidants can protect the kidney from the oxidative stress-induced decline in function (Wolf et al. 2001; Abdelrahman et al. 2010). In Chapter 4, I described how oxidative stress was increased in proximal tubule epithelial cells in response to chemical exposure. With paired omics data approach, future

scientists could explore whether *in vitro* exposure to antioxidants before chemical treatment could prevent oxidative stress from occurring. Post-injury, antioxidants have exhibited therapeutic effects in animal studies (Chander et al. 2005; Arozal et al. 2009), but have not had the same level of benefit in human patients (Ratliff et al. 2016; Dennis and Witting 2017). This presents an opportunity for further exploration of different treatment strategies to mitigate the harmful effects of ROS.

In addition to the projects listed above, there are two additional directions that this work may lead. The first is to extend the *iRno* model further than was done in chapter 4, to be a multi-compartment model. This idea would help us combine what we learned from the toxicity studies from the liver (Chapter 3) and the kidneys (Chapter 4) with a gut compartment to create a model that encompasses the interplay of these organs in drug metabolism. With the development of tissue-specific metabolic network reconstructions over the last few years, the scope of metabolic network reconstructions has expanded to cover multiple tissue types at once. This can be seen by our collaborators who investigated how liver and kidney metabolism changed as a result of drug exposure (Pannala et al. 2019). In addition to liver and kidney segments of this multi-scale model, the authors also included compartments for the blood and urine and allowed for metabolites to be exchanged across organs by way of the blood. The next step with this research would be to expand *iRno* to capture the gut epithelium due to its involvement in drug metabolism and its relation to the liver and kidney (Krishna and Klotz 1994; Wacher et al. 2001; Michell et al. 2008; Bush et al. 2017). The gut hosts many different microbial species that interact with each other and their metabolism changes with what we

consume. In the context of drug metabolism, bacteria in the gut chemically modify the compounds we ingest (Nicholson et al. 2005; Wilson and Nicholson 2017). These compounds then go on the be metabolized in the liver and excreted either in the liver or kidney. Understanding how these compounds are fully broken down can allow us to capture all by-products and necessary drug conjugates that may be excreted.

Another potential research direction is modeling whole-body metabolism. Within the metabolic modeling field, Recon 3D (Brunk et al. 2018) is an example of a reconstruction that was turned into a model that looks at whole-body metabolism, capturing metabolic function spanning different organs. Another modeling group used whole-body metabolism by creating two models called Harvey and Harvetta (Thiele et al. 2018 Jan 29). With *iRno* and our human model companion *iHsa*, more extensive curation to capture whole rat and human metabolism, similar to Harvey and Harvetta, would be useful to the metabolic modeling field, because of the inclusion of species-specific differences between the two organisms. We have previously used *in vitro* data to characterize toxicity data, however, with a whole-body model we could potentially use *in vivo* data to contextualize models and further represent the species-specific differences between rat and human metabolism. One limitation of these whole-body metabolic models is that while these models are larger in scope, they are still new to the field, and have not yet had broad usability. The accuracy of these models and the amount of data needed to constrain them to make useful predictions could also prove to be a challenge, highlighting the utility of single-cell or tissue-specific metabolic models. While computational systems biologists may take different approaches to advance the field, there is common agreement that

more steps must be taken before fully understanding human metabolism. My current research has made strides on capturing and predicting *in vitro* responses to drug-induced toxicity, and these next steps aim at getting closer to predicting how human metabolism is altered from potential exposure to chemical compounds.

6. References

Abdelrahman AM, Al Salam S, AlMahruqi AS, Al husseni IS, Mansour MA, Ali BH. 2010. N-acetylcysteine improves renal hemodynamics in rats with cisplatin-induced nephrotoxicity. J Appl Toxicol JAT. 30(1):15–21. doi:10.1002/jat.1465.

Adibi SA. 1976. Metabolism of branched-chain amino acids in altered nutrition. Metabolism. 25(11):1287–1302. doi:10.1016/S0026-0495(76)80012-1.

Adiyanti SS, Loho T. 2012. Acute Kidney Injury (AKI) biomarker. Acta Medica Indones. 44(3):246–255.

Agren R, Liu L, Shoaie S, Vongsangnak W, Nookaew I, Nielsen J. 2013. The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for Penicillium chrysogenum. PLOS Comput Biol. 9(3):e1002980. doi:10.1371/journal.pcbi.1002980.

Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. 2014. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. Mol Syst Biol. 10:721. doi:10.1002/msb.145122.

Albini A, Pennesi G, Donatelli F, Cammarota R, De Flora S, Noonan DM. 2010. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. J Natl Cancer Inst. 102(1):14–25. doi:10.1093/jnci/djp440.

Aly HAA, Domènech Ò. 2009. Cytotoxicity and mitochondrial dysfunction of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in isolated rat hepatocytes. Toxicol Lett. 191(1):79–87. doi:10.1016/j.toxlet.2009.08.008.

Arozal W, Watanabe K, Veeraveedu PT, Ma M, Thandavarayan RA, Suzuki K, Tachikawa H, Kodama M, Aizawa Y. 2009. Effects of angiotensin receptor blocker on oxidative stress and cardio-renal function in streptozotocin-induced diabetic rats. Biol Pharm Bull. 32(8):1411–1416. doi:10.1248/bpb.32.1411.

Asgari Y, Khosravi P, Zabihinpour Z, Habibi M. 2018. Exploring candidate biomarkers for lung and prostate cancers using gene expression and flux variability analysis. Integr Biol Quant Biosci Nano Macro. 10(2):113–120. doi:10.1039/c7ib00135e.

Ataman M, Hatzimanikatis V. 2015. Heading in the right direction: thermodynamics-based network analysis and pathway engineering. Curr Opin Biotechnol. 36:176–182. doi:10.1016/j.copbio.2015.08.021.

Awdishu L, Mehta RL. 2017. The 6R's of drug induced nephrotoxicity. BMC Nephrol. 18(1):124. doi:10.1186/s12882-017-0536-3.

Banday AA, Farooq N, Priyamvada S, Yusufi ANK, Khan F. 2008. Time dependent effects of gentamicin on the enzymes of carbohydrate metabolism, brush border membrane and oxidative stress in rat kidney tissues. Life Sci. 82(9–10):450–459. doi:10.1016/j.lfs.2007.11.014.

Bartell JA, Blazier AS, Yen P, Thøgersen JC, Jelsbak L, Goldberg JB, Papin JA. 2017. Reconstruction of the metabolic network of Pseudomonas aeruginosa to interrogate virulence factor synthesis. Nat Commun. 8:14631. doi:10.1038/ncomms14631.

Bartell JA, Yen P, Varga JJ, Goldberg JB, Papin JA. 2014. Comparative metabolic systems analysis of pathogenic Burkholderia. J Bacteriol. 196(2):210–226. doi:10.1128/JB.00997-13.

Baud L, Ardaillou R. 1986. Reactive oxygen species: production and role in the kidney. Am J Physiol. 251(5 Pt 2):F765-776. doi:10.1152/ajprenal.1986.251.5.F765.

Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ. 2007. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. Nat Protoc. 2(3):727–738. doi:10.1038/nprot.2007.99.

Becker SA, Palsson BO. 2008. Context-specific metabolic networks are consistent with experiments. PLoS Comput Biol. 4(5):e1000082. doi:10.1371/journal.pcbi.1000082.

Beger RD, Sun J, Schnackenberg LK. 2010. Metabolomics approaches for discovering biomarkers of drug-induced hepatotoxicity and nephrotoxicity. Toxicol Appl Pharmacol. 243(2):154–166. doi:10.1016/j.taap.2009.11.019.

Bellomo R, Kellum JA, Ronco C. 2012. Acute kidney injury. The Lancet. 380(9843):756–766. doi:10.1016/S0140-6736(11)61454-2.

Ben-Shachar R, Chen Y, Luo S, Hartman C, Reed M, Nijhout HF. 2012. The biochemistry of acetaminophen hepatotoxicity and rescue: a mathematical model. Theor Biol Med Model. 9:55. doi:10.1186/1742-4682-9-55.

Bentli R, Ciftci O, Cetin A, Unlu M, Basak N, Çay M. 2013. Oral administration of hesperidin, a citrus flavonone, in rats counteracts the oxidative stress, the inflammatory cytokine production, and the hepatotoxicity induced by the ingestion of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Eur Cytokine Netw. 24(2):91–96. doi:10.1684/ecn.2013.0337.

Biggs MB, Medlock GL, Moutinho TJ, Lees HJ, Swann JR, Kolling GL, Papin JA. 2017. Systems-level metabolism of the altered Schaedler flora, a complete gut microbiota. ISME J. 11(2):426–438. doi:10.1038/ismej.2016.130.

Bixel MG, Hutson SM, Hamprecht B. 1997. Cellular distribution of branched-chain amino acid aminotransferase isoenzymes among rat brain glial cells in culture. J Histochem Cytochem Off J Histochem Soc. 45(5):685–694. doi:10.1177/002215549704500506.

Blais EM, Rawls KD, Dougherty BV, Li ZI, Kolling GL, Ye P, Wallqvist A, Papin JA. 2017. Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions. Nat Commun. 8:14250. doi:10.1038/ncomms14250.

Blanchet L, Smolinska A, Attali A, Stoop MP, Ampt KA, van Aken H, Suidgeest E, Tuinstra T, Wijmenga SS, Luider T, et al. 2011. Fusion of metabolomics and proteomics data for biomarkers discovery: case study on the experimental autoimmune encephalomyelitis. BMC Bioinformatics. 12(1). doi:10.1186/1471-2105-12-254. [accessed 2019 Oct 29]. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-254.

Bobulescu IA. 2010. Renal lipid metabolism and lipotoxicity: Curr Opin Nephrol Hypertens. 19(4):393–402. doi:10.1097/MNH.0b013e32833aa4ac.

Bonventre JV. 2014. Kidney injury molecule-1: a translational journey. Trans Am Clin Climatol Assoc. 125:293–299; discussion 299.

Boogaard PJ, Mulder GJ, Nagelkerke JF. 1989. Isolated proximal tubular cells from rat kidney as an in vitro model for studies on nephrotoxicity. Toxicol Appl Pharmacol. 101(1):144–157. doi:10.1016/0041-008X(89)90220-2.

Boudonck KJ, Mitchell MW, Német L, Keresztes L, Nyska A, Shinar D, Rosenstock M. 2009. Discovery of metabolomics biomarkers for early detection of nephrotoxicity. Toxicol Pathol. 37(3):280–292. doi:10.1177/0192623309332992.

Boverhof DR, Burgoon LD, Tashiro C, Sharratt B, Chittim B, Harkema JR, Mendrick DL, Zacharewski TR. 2006. Comparative Toxicogenomic Analysis of the Hepatotoxic Effects of TCDD in Sprague Dawley Rats and C57BL/6 Mice. Toxicol Sci. 94(2):398–416. doi:10.1093/toxsci/kfl100.

Braissant O, Foufelle F, Scotto C, Dauça M, Wahli W. 1996. Differential expression of peroxisome proliferator-activated receptors (PPARs): tissue distribution of PPAR-alpha, -beta, and -gamma in the adult rat. Endocrinology. 137(1):354–366. doi:10.1210/endo.137.1.8536636.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 34(5):525–527. doi:10.1038/nbt.3519.

Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, Gatto F, Nilsson A, Preciat Gonzalez GA, Aurich MK, et al. 2018. Recon3D enables a three-dimensional view of gene variation in human metabolism. Nat Biotechnol. 36(3):272–281. doi:10.1038/nbt.4072.

Burch HB, Narins RG, Chu C, Fagioli S, Choi S, McCarthy W, Lowry OH. 1978. Distribution along the rat nephron of three enzymes of gluconeogenesis in acidosis and starvation. Am J Physiol. 235(3):F246-253. doi:10.1152/ajprenal.1978.235.3.F246.

Bush KT, Wu W, Lun C, Nigam SK. 2017. The drug transporter OAT3 (SLC22A8) and endogenous metabolite communication via the gut–liver–kidney axis. J Biol Chem. 292(38):15789–15803. doi:10.1074/jbc.M117.796516.

Cai Y, Gong L, Qi X, Li X, Ren J. 2005. Apoptosis initiated by carbon tetrachloride in mitochondria of rat primary cultured hepatocytes. Acta Pharmacol Sin. 26(8):969–975. doi:10.1111/j.1745-7254.2005.00143.x.

Cajka T, Fiehn O. 2017. LC-MS-Based Lipidomics and Automated Identification of Lipids Using the LipidBlast In-Silico MS/MS Library. Methods Mol Biol Clifton NJ. 1609:149–170. doi:10.1007/978-1-4939-6996-8_14.

Canayakin D, Bayir Y, Kilic Baygutalp N, Sezen Karaoglan E, Atmaca HT, Kocak Ozgeris FB, Keles MS, Halici Z. 2016. Paracetamol-induced nephrotoxicity and oxidative stress in rats: the protective role of Nigella sativa. Pharm Biol. 54(10):2082–2091. doi:10.3109/13880209.2016.1145701.

Carbonell P, Lopez O, Amberg A, Pastor M, Sanz F. 2017. Hepatotoxicity prediction by systems biology modeling of disturbed metabolic pathways using gene expression data. ALTEX. 34(2):219–234. doi:10.14573/altex.1602071.

Carey MA, Papin JA, Guler JL. 2017. Novel Plasmodium falciparum metabolic network reconstruction identifies shifts associated with clinical antimalarial resistance. BMC Genomics. 18(1):543. doi:10.1186/s12864-017-3905-1.

Chander V, Tirkey N, Chopra K. 2005. Resveratrol, a polyphenolic phytoalexin protects against cyclosporine-induced nephrotoxicity through nitric oxide dependent mechanism. Toxicology. 210(1):55–64. doi:10.1016/j.tox.2005.01.011.

Chang RL, Xie Li, Xie Lei, Bourne PE, Palsson BØ. 2010. Drug Off-Target Effects Predicted Using Structural Analysis in the Context of a Metabolic Network Model. Dunbrack RL, editor. PLoS Comput Biol. 6(9):e1000938. doi:10.1371/journal.pcbi.1000938.

Chen M, Suzuki A, Borlak J, Andrade RJ, Lucena MI. 2015. Drug-induced liver injury: Interactions between drug properties and host factors. J Hepatol. 63(2):503–514. doi:10.1016/j.jhep.2015.04.016.

Chiang JYL. 2013. Bile acid metabolism and signaling. Compr Physiol. 3(3):1191–1212. doi:10.1002/cphy.c120023.

Church RJ, Watkins PB. 2017. The transformation in biomarker detection and management of drug-induced liver injury. Liver Int. 37(11):1582–1590. doi:10.1111/liv.13441.

Cojocel C, Beuter W, Müller W, Mayer D. 1989. Lipid peroxidation: A possible mechanism of trichloroethylene-induced nephrotoxicity. Toxicology. 55(1–2):131–141. doi:10.1016/0300-483X(89)90180-7.

Connor SC, Hansen MK, Corner A, Smith RF, Ryan TE. 2010. Integration of metabolomics and transcriptomics data to aid biomarker discovery in type 2 diabetes. Mol Biosyst. 6(5):909. doi:10.1039/b914182k.

Dantzler WH, Silbernagl S. 1988. Amino acid transport by juxtamedullary nephrons: distal reabsorption and recycling. Am J Physiol. 255(3 Pt 2):F397-407. doi:10.1152/ajprenal.1988.255.3.F397.

Dennis JM, Witting PK. 2017. Protective Role for Antioxidants in Acute Kidney Disease. Nutrients. 9(7). doi:10.3390/nu9070718.

Dere E, Lee AW, Burgoon LD, Zacharewski TR. 2011. Differences in TCDD-elicited gene expression profiles in human HepG2, mouse Hepa1c1c7 and rat H4IIE hepatoma cells. BMC Genomics. 12(1). doi:10.1186/1471-2164-12-193. [accessed 2019 Jun 28]. http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-12-193.

Dhanakoti SN, Brosnan JT, Herzberg GR, Brosnan ME. 1990. Renal arginine synthesis: studies in vitro and in vivo. Am J Physiol. 259(3 Pt 1):E437-442. doi:10.1152/ajpendo.1990.259.3.E437.

Dieterle F, Perentes E, Cordier A, Roth DR, Verdes P, Grenet O, Pantano S, Moulin P, Wahl D, Mahl A, et al. 2010. Urinary clusterin, cystatin C, beta2-microglobulin and total protein as markers to detect drug-induced kidney injury. Nat Biotechnol. 28(5):463–469. doi:10.1038/nbt.1622.

Dong B, Nishimura N, Vogel CF, Tohyama C, Matsumura F. 2010. TCDD-induced cyclooxygenase-2 expression is mediated by the nongenomic pathway in mouse MMDD1 macula densa cells and kidneys. Biochem Pharmacol. 79(3):487–497. doi:10.1016/j.bcp.2009.08.031.

Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, Srivas R, Palsson BO. 2007. Global reconstruction of the human metabolic network based on genomic and bibliomic data. Proc Natl Acad Sci. 104(6):1777–1782. doi:10.1073/pnas.0610772104.

Dufour DR, Lott JA, Nolte FS, Gretch DR, Koff RS, Seeff LB. 2000. Diagnosis and monitoring of hepatic injury. I. Performance characteristics of laboratory tests. Clin Chem. 46(12):2027–2049.

Edwards JS, Palsson BO. 1999. Systems properties of the Haemophilus influenzae Rd metabolic genotype. J Biol Chem. 274(25):17410–17416. doi:10.1074/jbc.274.25.17410.

Elhamri M, Martin M, Ferrier B, Baverel G. 1993. Substrate uptake and utilization by the kidney of fed and starved rats in vivo. Ren Physiol Biochem. 16(6):311–324.

Endre ZH, Pickering JW, Walker RJ, Devarajan P, Edelstein CL, Bonventre JV, Frampton CM, Bennett MR, Ma Q, Sabbisetti VS, et al. 2011. Improved performance of urinary biomarkers of acute kidney injury in the critically ill by stratification for injury duration and baseline renal function. Kidney Int. 79(10):1119–1130. doi:10.1038/ki.2010.555.

Fiehn O. 2016. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. Curr Protoc Mol Biol. 114:30.4.1-30.4.32. doi:10.1002/0471142727.mb3004s114.

Fontanesi F, Soto IC, Horn D, Barrientos A. 2006. Assembly of mitochondrial cytochrome *c* -oxidase, a complicated and highly regulated cellular process. Am J Physiol-Cell Physiol. 291(6):C1129–C1147. doi:10.1152/ajpcell.00233.2006.

Forgacs AL, Dere E, Angrish MM, Zacharewski TR. 2013. Comparative Analysis of Temporal and Dose-Dependent TCDD-Elicited Gene Expression in Human, Mouse, and Rat Primary Hepatocytes. Toxicol Sci. 133(1):54–66. doi:10.1093/toxsci/kft028.

Forster J. 2003. Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network. Genome Res. 13(2):244–253. doi:10.1101/gr.234503.

Fritzemeier CJ, Hartleb D, Szappanos B, Papp B, Lercher MJ. 2017. Erroneous energy-generating cycles in published genome scale metabolic networks: Identification and removal. PLOS Comput Biol. 13(4):e1005494. doi:10.1371/journal.pcbi.1005494.

Gallagher D, Albu J, He Q, Heshka S, Boxt L, Krasnow N, Elia M. 2006. Small organs with a high metabolic rate explain lower resting energy expenditure in African American than in white adults. Am J Clin Nutr. 83(5):1062–1067. doi:10.1093/ajcn/83.5.1062.

Gatto F, Miess H, Schulze A, Nielsen J. 2015. Flux balance analysis predicts essential genes in clear cell renal cell carcinoma metabolism. Sci Rep. 5:10738. doi:10.1038/srep10738.

Gerich JE, Meyer C, Woerle HJ, Stumvoll M. 2001. Renal Gluconeogenesis: Its importance in human glucose homeostasis. Diabetes Care. 24(2):382–391. doi:10.2337/diacare.24.2.382.

de Geus HRH, Betjes MG, Bakker J. 2012. Biomarkers for the prediction of acute kidney injury: a narrative review on current status and future challenges. Clin Kidney J. 5(2):102–108. doi:10.1093/ckj/sfs008.

Ghaffari P, Mardinoglu A, Asplund A, Shoaie S, Kampf C, Uhlen M, Nielsen J. 2015. Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling. Sci Rep. 5(1). doi:10.1038/srep08183. [accessed 2019 Oct 30]. http://www.nature.com/articles/srep08183.

Ghiculescu RA, Kubler PA. 2006. Aminoglycoside-Associated Fanconi Syndrome. Am J Kidney Dis. 48(6):e89–e93. doi:10.1053/j.ajkd.2006.08.009.

Gille C, Bölling C, Hoppe A, Bulik S, Hoffmann S, Hübner K, Karlstädt A, Ganeshan R, König M, Rother K, et al. 2010. HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. Mol Syst Biol. 6:411. doi:10.1038/msb.2010.62.

Gonzalez-Vicente A, Hong N, Garvin JL. 2019. Effects of reactive oxygen species on renal tubular transport. Am J Physiol-Ren Physiol. 317(2):F444–F455. doi:10.1152/ajprenal.00604.2018.

Gowda S, Desai PB, Kulkarni SS, Hull VV, Math AAK, Vernekar SN. 2010. Markers of renal function tests. North Am J Med Sci. 2(4):170–173.

Guengerich FP. 2008. Cytochrome P450 and Chemical Toxicology. Chem Res Toxicol. 21(1):70–83. doi:10.1021/tx700079z.

Hädicke O, Klamt S. 2017. EColiCore2: a reference network model of the central metabolism of Escherichia coli and relationships to its genome-scale parent model. Sci Rep. 7:39647–39647. doi:10.1038/srep39647.

van Heck RGA, Ganter M, Martins dos Santos VAP, Stelling J. 2016. Efficient Reconstruction of Predictive Consensus Metabolic Network Models. PLoS Comput Biol. 12(8). doi:10.1371/journal.pcbi.1005085. [accessed 2018 Nov 13]. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5001716/.

Henderson CJ, Wolf CR, Kitteringham N, Powell H, Otto D, Park BK. 2000. Increased resistance to acetaminophen hepatotoxicity in mice lacking glutathione S-transferase Pi. Proc Natl Acad Sci U S A. 97(23):12741–12745. doi:10.1073/pnas.220176997.

Himmelfarb J, Ikizler TA. 2007. Acute kidney injury: changing lexicography, definitions, and epidemiology. Kidney Int. 71(10):971–976. doi:10.1038/sj.ki.5002224.

Huang DW, Sherman BT, Lempicki RA. 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 4(1):44–57. doi:10.1038/nprot.2008.211.

Huang DW, Sherman BT, Lempicki RA. 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37(1):1–13. doi:10.1093/nar/gkn923.

International HapMap Consortium. 2005. A haplotype map of the human genome. Nature. 437(7063):1299–1320. doi:10.1038/nature04226.

Ishihara K, Katsutani N, Aoki T. 2006. A Metabonomics Study of the Hepatotoxicants Galactosamine, Methylene Dianiline and Clofibrate in Rats. Basic Clin Pharmacol Toxicol. 99(3):251–260. doi:10.1111/j.1742-7843.2006.pto_455.x.

Jaeschke H, Gores GJ, Cederbaum AI, Hinson JA, Pessayre D, Lemasters JJ. 2002. Mechanisms of hepatotoxicity. Toxicol Sci Off J Soc Toxicol. 65(2):166–176.

Jerby L, Ruppin E. 2012. Predicting drug targets and biomarkers of cancer via genome-scale metabolic modeling. Clin Cancer Res Off J Am Assoc Cancer Res. 18(20):5572–5584. doi:10.1158/1078-0432.CCR-12-1856.

Jerby L, Shlomi T, Ruppin E. 2010. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. Mol Syst Biol. 6:401. doi:10.1038/msb.2010.56.

Jungbauer CG, Uecer E, Stadler S, Birner C, Buchner S, Maier LS, Luchner A. 2016. N-acteyl-ß-D-glucosaminidase and kidney injury molecule-1: New predictors for long-term progression of chronic kidney disease in patients with heart failure. Nephrol Carlton Vic. 21(6):490–498. doi:10.1111/nep.12632.

Kanehisa M. 2019. Toward understanding the origin and evolution of cellular organisms. Protein Sci Publ Protein Soc. 28(11):1947–1951. doi:10.1002/pro.3715.

Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28(1):27–30. doi:10.1093/nar/28.1.27.

Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. 2019. New approach for understanding genome variations in KEGG. Nucleic Acids Res. 47(D1):D590–D595. doi:10.1093/nar/gky962.

Karlstädt A, Fliegner D, Kararigas G, Ruderisch HS, Regitz-Zagrosek V, Holzhütter H-G. 2012. CardioNet: a human metabolic network suited for the study of cardiomyocyte metabolism. BMC Syst Biol. 6:114. doi:10.1186/1752-0509-6-114.

Kato Y, Sugiura M, Sugiura T, Wakayama T, Kubo Y, Kobayashi D, Sai Y, Tamai I, Iseki S, Tsuji A. 2006. Organic cation/carnitine transporter OCTN2 (Slc22a5) is responsible for carnitine transport across apical membranes of small intestinal epithelial cells in mouse. Mol Pharmacol. 70(3):829–837. doi:10.1124/mol.106.024158.

Khan MR, Rizvi W, Khan GN, Khan RA, Shaheen S. 2009. Carbon tetrachloride-induced nephrotoxicity in rats: Protective role of Digera muricata. J Ethnopharmacol. 122(1):91–99. doi:10.1016/j.jep.2008.12.006.

Kienhuis AS, van de Poll MCG, Wortelboer H, van Herwijnen M, Gottschalk R, Dejong CHC, Boorsma A, Paules RS, Kleinjans JCS, Stierum RH, et al. 2009. Parallelogram Approach Using Rat-Human In Vitro and Rat In Vivo Toxicogenomics Predicts Acetaminophen-induced Hepatotoxicity in Humans. Toxicol Sci. 107(2):544–552. doi:10.1093/toxsci/kfn237.

Kim SY, Moon A. 2012. Drug-induced nephrotoxicity and its biomarkers. Biomol Ther. 20(3):268–272. doi:10.4062/biomolther.2012.20.3.268.

Krishna DR, Klotz U. 1994. Extrahepatic Metabolism of Drugs in Humans: Clin Pharmacokinet. 26(2):144–160. doi:10.2165/00003088-199426020-00007.

Kumar BS, Chung BC, Kwon O-S, Jung BH. 2012. Discovery of common urinary biomarkers for hepatotoxicity induced by carbon tetrachloride, acetaminophen and methotrexate by mass spectrometry-based metabolomics. J Appl Toxicol JAT. 32(7):505–520. doi:10.1002/jat.1746.

Lash LH, Qian W, Putt DA, Hueni SE, Elfarra AA, Krause RJ, Parker JC. 2001. Renal and hepatic toxicity of trichloroethylene and its glutathione-derived metabolites in rats and mice: sex-, species-, and tissue-dependent differences. J Pharmacol Exp Ther. 297(1):155–164.

Layton AT. 2013. Mathematical modeling of kidney transport. Wiley Interdiscip Rev Syst Biol Med. 5(5):557–573. doi:10.1002/wsbm.1232.

Le Hir M, Dubach UC. 1982. Peroxisomal and mitochondrial beta-oxidation in the rat kidney: distribution of fatty acyl-coenzyme A oxidase and 3-hydroxyacyl-coenzyme A dehydrogenase activities along the nephron. J Histochem Cytochem. 30(5):441–444. doi:10.1177/30.5.7200500.

Lenz EM, Bright J, Knight R, Wilson ID, Major H. 2004. Cyclosporin A-induced changes in endogenous metabolites in rat urine: a metabonomic investigation using high field 1H NMR spectroscopy, HPLC-TOF/MS and chemometrics. J Pharm Biomed Anal. 35(3):599–608. doi:10.1016/j.jpba.2004.02.013.

Longo N, Amat di San Filippo C, Pasquali M. 2006. Disorders of carnitine transport and the carnitine cycle. Am J Med Genet C Semin Med Genet. 142C(2):77–85. doi:10.1002/ajmg.c.30087.

Lotan Y, Woldu SL, Sanli O, Black P, Milowsky MI. 2018 Mar 31. Modeling Cost-Effectiveness of a Biomarker-Based Approach to Neoadjuvant Chemotherapy for Muscle-Invasive Bladder Cancer. BJU Int. doi:10.1111/bju.14220.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550. doi:10.1186/s13059-014-0550-8.

Ma H, Sorokin A, Mazein A, Selkov A, Selkov E, Demin O, Goryanin I. 2007. The Edinburgh human metabolic network reconstruction and its functional analysis. Mol Syst Biol. 3:135. doi:10.1038/msb4100177.

Macpherson NA, Moscarello MA, Goldberg DM. 1991. Aminoaciduria is an earlier index of renal tubular damage than conventional renal disease markers in the gentamicin-rat model of acute renal failure. Clin Investig Med Med Clin Exp. 14(2):101–110.

Mardinoglu A, Agren R, Kampf C, Asplund A, Nookaew I, Jacobson P, Walley AJ, Froguel P, Carlsson LM, Uhlen M, et al. 2013. Integration of clinical data with a genome-scale metabolic model of the human adipocyte. Mol Syst Biol. 9:649. doi:10.1038/msb.2013.5.

Mardinoglu A, Agren R, Kampf C, Asplund A, Uhlen M, Nielsen J. 2014. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. Nat Commun. 5:3083. doi:10.1038/ncomms4083.

Mardinoglu A, Gatto F, Nielsen J. 2013. Genome-scale modeling of human metabolism – a systems biology approach. Biotechnol J. 8(9):985–996. doi:10.1002/biot.201200275.

Mardinoglu A, Shoaie S, Bergentall M, Ghaffari P, Zhang C, Larsson E, Bäckhed F, Nielsen J. 2015. The gut microbiota modulates host amino acid and glutathione metabolism in mice. Mol Syst Biol. 11(10):834. doi:10.15252/msb.20156487.

Marín de Mas I, Aguilar E, Zodda E, Balcells C, Marin S, Dallmann G, Thomson TM, Papp B, Cascante M. 2018. Model-driven discovery of long-chain fatty acid metabolic reprogramming in heterogeneous prostate cancer cells. PLoS Comput Biol. 14(1):e1005914. doi:10.1371/journal.pcbi.1005914.

Matheis KA, Com E, Gautier J-C, Guerreiro N, Brandenburg A, Gmuender H, Sposny A, Hewitt P, Amberg A, Boernsen O, et al. 2011. Cross-study and cross-omics comparisons of three nephrotoxic compounds reveal mechanistic insights and new candidate biomarkers. Toxicol Appl Pharmacol. 252(2):112–122. doi:10.1016/j.taap.2010.11.006.

Mather A, Pollock C. 2011. Glucose handling by the kidney. Kidney Int. 79:S1–S6. doi:10.1038/ki.2010.509.

Matyash V, Liebisch G, Kurzchalia TV, Shevchenko A, Schwudke D. 2008. Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. J Lipid Res. 49(5):1137–1146. doi:10.1194/jlr.D700041-JLR200.

Mazer M, Perrone J. 2008. Acetaminophen-induced nephrotoxicity: Pathophysiology, clinical manifestations, and management. J Med Toxicol. 4(1):2–6. doi:10.1007/BF03160941.

McGill MR, Jaeschke H. 2013. Metabolism and disposition of acetaminophen: recent advances in relation to hepatotoxicity and diagnosis. Pharm Res. 30(9):2174–2187. doi:10.1007/s11095-013-1007-6.

McGill MR, Sharpe MR, Williams CD, Taha M, Curry SC, Jaeschke H. 2012. The mechanism underlying acetaminophen-induced hepatotoxicity in humans and mice involves mitochondrial damage and nuclear DNA fragmentation. J Clin Invest. 122(4):1574–1583. doi:10.1172/JCI59755.

Meissen JK, Hirahatake KM, Adams SH, Fiehn O. 2015. Temporal metabolomic responses of cultured HepG2 liver cells to high fructose and high glucose exposures. Metabolomics Off J Metabolomic Soc. 11(3):707–721. doi:10.1007/s11306-014-0729-8.

Metabolism at a Glance, 3rd Edition. Wiley.com. [accessed 2018 Jul 25]. https://www.wiley.com/en-us/Metabolism+at+a+Glance%2C+3rd+Edition-p-9781118682074.

Michell AR, Debnam ES, Unwin RJ. 2008. Regulation of Renal Function by the Gastrointestinal Tract: Potential Role of Gut-Derived Peptides and Hormones. Annu Rev Physiol. 70(1):379–403. doi:10.1146/annurev.physiol.69.040705.141330.

Mitchell DB, Acosta D, Bruckner JV. 1985. Role of glutathione depletion in the cytotoxicity of acetaminophen in a primary culture system of rat hepatocytes. Toxicology. 37(1–2):127–146. doi:10.1016/0300-483X(85)90119-2.

Mizuno D, Takahashi Y, Hiroi T, Imaoka S, Kamataki T, Funae Y. 2003. A novel transcriptional element which regulates expression of the CYP2D4 gene by Oct-1 and YY-1 binding. Biochim Biophys Acta. 1627(2–3):121–128.

Monks TJ, Anders MW, Dekant W, Stevens JL, Lau SS, van Bladeren PJ. 1990. Glutathione conjugate mediated toxicities. Toxicol Appl Pharmacol. 106(1):1–19.

Mount PF, Power DA. 2006. Nitric oxide in the kidney: functions and regulation of synthesis. Acta Physiol Oxf Engl. 187(4):433–446. doi:10.1111/j.1748-1716.2006.01582.x.

Mugford C. 1997. The contribution of oxidation and deacetylation to acetaminophen nephrotoxicity in female Sprague-Dawley rats. Toxicol Lett. 93(1):15–22. doi:10.1016/S0378-4274(97)00063-5.

Nakamura T, Yoshida K, Yabuuchi H, Maeda T, Tamai I. 2008. Functional characterization of ergothioneine transport by rat organic cation/carnitine transporter Octn1 (slc22a4). Biol Pharm Bull. 31(8):1580–1584. doi:10.1248/bpb.31.1580.

Narayana K. 2008. An aminoglycoside antibiotic gentamycin induces oxidative stress, reduces antioxidant reserve and impairs spermatogenesis in rats. J Toxicol Sci. 33(1):85–96. doi:10.2131/jts.33.85.

Newton JF, Yoshimoto M, Bernstein J, Rush GF, Hook JB. 1983. Acetaminophen nephrotoxicity in the rat. Toxicol Appl Pharmacol. 69(2):291–306. doi:10.1016/0041-008X(83)90311-3.

Nicholson JK, Holmes E, Wilson ID. 2005. Gut microorganisms, mammalian metabolism and personalized health care. Nat Rev Microbiol. 3(5):431–438. doi:10.1038/nrmicro1152.

Niemann CU, Serkova NJ. 2007. Biochemical mechanisms of nephrotoxicity: application for metabolomics. Expert Opin Drug Metab Toxicol. 3(4):527–544. doi:10.1517/17425255.3.4.527.

Oberhardt MA, Palsson BØ, Papin JA. 2009. Applications of genome-scale metabolic reconstructions. Mol Syst Biol. 5:320. doi:10.1038/msb.2009.77.

O'Brien EJ, Monk JM, Palsson BO. 2015. Using Genome-scale Models to Predict Biological Capabilities. Cell. 161(5):971–987. doi:10.1016/j.cell.2015.05.019.

Ogle DH, Wheeler P, Dinno A. 2019. FSA: Fisheries Stock Analysis. https://github.com/droglenc/FSA.

Onopiuk A, Tokarzewicz A, Gorodkiewicz E. 2015. Cystatin C. In: Advances in Clinical Chemistry. Vol. 68. Elsevier. p. 57–69. [accessed 2019 Oct 29]. https://linkinghub.elsevier.com/retrieve/pii/S0065242314000389.

Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ. 2011. A comprehensive genome-scale reconstruction of Escherichia coli metabolism--2011. Mol Syst Biol. 7:535. doi:10.1038/msb.2011.65.

Orth JD, Fleming RMT, Palsson BØ. 2010. Reconstruction and Use of Microbial Metabolic Networks: the Core Escherichia coli Metabolic Model as an Educational Guide. EcoSal Plus. 4(1). doi:10.1128/ecosalplus.10.2.1. [accessed 2018 Nov 7]. http://www.asmscience.org/content/journal/ecosalplus/10.1128/ecosalplus.10.2.1.

Owen OE, Felig P, Morgan AP, Wahren J, Cahill GF. 1969. Liver and kidney metabolism during prolonged starvation. J Clin Invest. 48(3):574–583. doi:10.1172/JCI106016.

Pan X, Quan J, Li Z, Zhao L, Zhou L, Jinling X, Weijie X, Guan X, Li H, Yang S, et al. 2018. miR-566 functions as an oncogene and a potential biomarker for prognosis in renal cell carcinoma. Biomed Pharmacother Biomedecine Pharmacother. 102:718–727. doi:10.1016/j.biopha.2018.03.072.

Pannala VR, Vinnakota KC, Rawls KD, Estes SK, O'Brien TP, Printz RL, Papin JA, Reifman J, Shiota M, Young JD, et al. 2019. Mechanistic identification of biofluid metabolite changes as markers of acetaminophen-induced liver toxicity in rats. Toxicol Appl Pharmacol. 372:19–32. doi:10.1016/j.taap.2019.04.001.

Pannala VR, Wall ML, Estes SK, Trenary I, O'Brien TP, Printz RL, Vinnakota KC, Reifman J, Shiota M, Young JD, et al. 2018. Metabolic network-based predictions of toxicant-induced metabolite changes in the laboratory rat. Sci Rep. 8(1):11678. doi:10.1038/s41598-018-30149-7.

Park J-C, Hong Y-S, Joo Kim Y, Yang J-Y, Kim E-Y, Jun Kwack S, Hyun Ryu D, Hwang G-S, Mu Lee B. 2009. A Metabonomic Study on the Biochemical Effects of Doxorubicin in Rats Using ^{1}H-NMR Spectroscopy. J Toxicol Environ Health A. 72(6):374–384. doi:10.1080/15287390802647195.

van de Poll MC, Soeters PB, Deutz NE, Fearon KC, Dejong CH. 2004. Renal metabolism of amino acids: its role in interorgan amino acid exchange. Am J Clin Nutr. 79(2):185–197. doi:10.1093/ajcn/79.2.185.

Pornputtapong N, Nookaew I, Nielsen J. 2015. Human metabolic atlas: an online resource for human metabolism. Database J Biol Databases Curation. 2015:bav068. doi:10.1093/database/bav068.

Portilla D, Li S, Nagothu KK, Megyesi J, Kaissling B, Schnackenberg L, Safirstein RL, Beger RD. 2006. Metabolomic study of cisplatin-induced nephrotoxicity. Kidney Int. 69(12):2194–2204. doi:10.1038/sj.ki.5000433.

Quek L-E, Dietmair S, Hanscho M, Martínez VS, Borth N, Nielsen LK. 2014. Reducing Recon 2 for steady-state flux analysis of HEK cell culture. J Biotechnol. 184:172–178. doi:10.1016/j.jbiotec.2014.05.021.

Ratliff BB, Abdulmahdi W, Pawar R, Wolin MS. 2016. Oxidant Mechanisms in Renal Injury and Disease. Antioxid Redox Signal. 25(3):119–146. doi:10.1089/ars.2016.6665.

Rawls KD, Blais EM, Dougherty BV, Pannala VR, Vinnakota KC, Wallqvist A, Kolling GL, Papin JA. 2019. Genome-Scale Characterization of Toxicity-Induced Metabolic Alterations in Primary Hepatocytes. doi:10.5061/dryad.04vk390.

Rawls KD, Dougherty BV, Blais EM, Stancliffe E, Kolling GL, Vinnakota K, Pannala VR, Wallqvist A, Papin JA. 2018. A simplified metabolic network reconstruction to promote understanding and development of flux balance analysis tools. Comput Biol Med. 105:64–71. doi:10.1016/j.compbiomed.2018.12.010.

Rawls KD, Dougherty BV, Blais EM, Stancliffe E, Kolling GL, Vinnakota K, Pannala VR, Wallqvist A, Papin JA. 2019. A simplified metabolic network reconstruction to promote understanding and development of flux balance analysis tools. Comput Biol Med. 105:64–71. doi:10.1016/j.compbiomed.2018.12.010.

Robbiano L, Baroni D, Carrozzino R, Mereto E, Brambilla G. 2004. DNA damage and micronuclei induced in rat and human kidney cells by six chemicals carcinogenic to the rat kidney. Toxicology. 204(2–3):187–195. doi:10.1016/j.tox.2004.06.057.

Ross BD, Espinal J, Silva P. 1986. Glucose metabolism in renal tubular function. Kidney Int. 29(1):54–67. doi:10.1038/ki.1986.8.

Rueda-Zárate HA, Imaz-Rosshandler I, Cárdenas-Ovando RA, Castillo-Fernández JE, Noguez-Monroy J, Rangel-Escareño C. 2017. A computational toxicogenomics approach identifies a list of highly hepatotoxic compounds from a large microarray database. Deoraj A, editor. PLOS ONE. 12(4):e0176284. doi:10.1371/journal.pone.0176284.

Salway JG. 2017. Metabolism at a glance. Fourth edition. Chichester, West Sussex ; Hoboken, NJ: John Wiley & Sons Inc (At a glance series).

Šarenac TM, Mikov M. 2018. Bile Acid Synthesis: From Nature to the Chemical Modification and Synthesis and Their Applications as Drugs and Nutrients. Front Pharmacol. 9. doi:10.3389/fphar.2018.00939. [accessed 2019 Oct 30]. https://www.frontiersin.org/article/10.3389/fphar.2018.00939/full.

Sawada R, Iwata M, Tabei Y, Yamato H, Yamanishi Y. 2018. Predicting inhibitory and activatory drug targets by chemically and genetically perturbed transcriptome signatures. Sci Rep. 8(1). doi:10.1038/s41598-017-18315-9. [accessed 2019 Oct 29]. http://www.nature.com/articles/s41598-017-18315-9.

Schellenberger J, Lewis NE, Palsson BØ. 2011. Elimination of Thermodynamically Infeasible Loops in Steady-State Metabolic Models. Biophys J. 100(3):544–553. doi:10.1016/j.bpj.2010.12.3707.

Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, et al. 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat Protoc. 6(9):1290–1307. doi:10.1038/nprot.2011.308.

Schultz A, Qutub AA. 2016. Reconstruction of Tissue-Specific Metabolic Networks Using CORDA. PLOS Comput Biol. 12(3):e1004808. doi:10.1371/journal.pcbi.1004808.

Scott RP, Quaggin SE. 2015. The cell biology of renal filtration. J Cell Biol. 209(2):199–210. doi:10.1083/jcb.201410017.

Sgouralis I, Layton AT. 2015. Mathematical modeling of renal hemodynamics in physiology and pathophysiology. Math Biosci. 264:8–20. doi:10.1016/j.mbs.2015.02.016.

Shlipak MG, Scherzer R, Abraham A, Tien PC, Grunfeld C, Peralta CA, Devarajan P, Bennett M, Butch AW, Anastos K, et al. 2012. Urinary markers of kidney injury and kidney function decline in HIV-infected women. J Acquir Immune Defic Syndr 1999. 61(5):565–573. doi:10.1097/QAI.0b013e3182737706.

Shlomi T, Cabili MN, Herrgård MJ, Palsson BØ, Ruppin E. 2008. Network-based prediction of human tissue-specific metabolism. Nat Biotechnol. 26(9):1003–1010. doi:10.1038/nbt.1487.

Sigurdsson MI, Jamshidi N, Steingrimsson E, Thiele I, Palsson BØ. 2010. A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. BMC Syst Biol. 4(1):140. doi:10.1186/1752-0509-4-140.

Sjogren A-KM, Liljevald M, Glinghammar B, Sagemark J, Li X-Q, Jonebring A, Cotgreave I, Brolén G, Andersson TB. 2014. Critical differences in toxicity mechanisms in induced pluripotent stem cell-derived hepatocytes, hepatic cell lines and primary hepatocytes. Arch Toxicol. 88(7):1427–1437. doi:10.1007/s00204-014-1265-z.

Smith JH. 1988. The use of renal cortical slices from the Fischer 344 rat as an in vitro model to evaluate nephrotoxicity. Fundam Appl Toxicol Off J Soc Toxicol. 11(1):132–142. doi:10.1016/0272-0590(88)90277-1.

Soga T, Baran R, Suematsu M, Ueno Y, Ikeda S, Sakurakawa T, Kakazu Y, Ishikawa T, Robert M, Nishioka T, et al. 2006. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. J Biol Chem. 281(24):16768–16776. doi:10.1074/jbc.M601876200.

Sohrabi-Jahromi S, Marashi S-A, Kalantari S. 2016. A kidney-specific genome-scale metabolic network model for analyzing focal segmental glomerulosclerosis. Mamm Genome Off J Int Mamm Genome Soc. 27(3–4):158–167. doi:10.1007/s00335-016-9622-2.

Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research. 4:1521. doi:10.12688/f1000research.7563.2.

Spitzer A. 1982. The role of the kidney in sodium homeostasis during maturation. Kidney Int. 21(4):539–545. doi:10.1038/ki.1982.60.

Stempler S, Yizhak K, Ruppin E. 2014. Integrating Transcriptomics with Metabolic Modeling Predicts Biomarkers and Drug Targets for Alzheimer's Disease. Fong SS, editor. PLoS ONE. 9(8):e105383. doi:10.1371/journal.pone.0105383.

Sugiura T, Kato Y, Wakayama T, Silver DL, Kubo Y, Iseki S, Tsuji A. 2008. PDZK1 regulates two intestinal solute carriers (Slc15a1 and Slc22a5) in mice. Drug Metab Dispos Biol Fate Chem. 36(6):1181–1188. doi:10.1124/dmd.107.020321.

Sun J, Ando Y, Ahlbory-Dieker D, Schnackenberg LK, Yang X, Greenhaw J, Pence L, Qian F, Salminen W, Mendrick DL, et al. 2013. Systems Biology Investigation to Discover Metabolic Biomarkers of Acetaminophen-Induced Hepatic Injury Using Integrated Transcriptomics and

Metabolomics. J Mol Biomark Diagn. s1. doi:10.4172/2155-9929.S1-002. [accessed 2019 Apr 1]. https://www.omicsonline.org/systems-biology-investigation-to-discover-metabolic-biomarkers-of-acetaminophen-induced-hepatic-injury-using-integrated-transcriptomics-and-metabolomics-2155-9929.S1-002.php?aid=11503.

Sun J, Schnackenberg LK, Holland RD, Schmitt TC, Cantor GH, Dragan YP, Beger RD. 2008. Metabonomics evaluation of urine from rats given acute and chronic doses of acetaminophen using NMR and UPLC/MS. J Chromatogr B. 871(2):328–340. doi:10.1016/j.jchromb.2008.04.008.

Swainston N, Smallbone K, Hefzi H, Dobson PD, Brewer J, Hanscho M, Zielinski DC, Ang KS, Gardiner NJ, Gutierrez JM, et al. 2016. Recon 2.2: from reconstruction to model of human metabolism. Metabolomics Off J Metabolomic Soc. 12:109. doi:10.1007/s11306-016-1051-4.

Taguchi K, Tokuno M, Yamasaki K, Kadowaki D, Seo H, Otagiri M. 2015. Establishment of a model of acetaminophen-induced hepatotoxicity in different weekly-aged ICR mice. Lab Anim. 49(4):294–301. doi:10.1177/0023677215573041.

Tamai I. 2013. Pharmacological and pathophysiological roles of carnitine/organic cation transporters (OCTNs: SLC22A4, SLC22A5 and Slc22a21): ROLE OF OCTN TRANSPORTERS. Biopharm Drug Dispos. 34(1):29–44. doi:10.1002/bdd.1816.

Thiele I, Sahoo S, Heinken A, Heirendt L, Aurich MK, Noronha A, Fleming RMT. 2018 Jan 29. When metabolism meets physiology: Harvey and Harvetta. bioRxiv. doi:10.1101/255885. [accessed 2019 Oct 31]. http://biorxiv.org/lookup/doi/10.1101/255885.

Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, et al. 2013. A community-driven global reconstruction of human metabolism. Nat Biotechnol. 31(5):419–425. doi:10.1038/nbt.2488.

Tran A, Tréluyer J-M, Rey E, Barbet J, Ferracci G, d'Athis P, Vincent J, Pons G. 2001. Protective Effect of Stiripentol on Acetaminophen-Induced Hepatotoxicity in Rat. Toxicol Appl Pharmacol. 170(3):145–152. doi:10.1006/taap.2000.9091.

Trougakos IP, Gonos ES. 2006. Regulation of clusterin/apolipoprotein J, a functional homologue to the small heat shock proteins, by oxidative stress in ageing and age-related diseases. Free Radic Res. 40(12):1324–1334. doi:10.1080/10715760600902310.

Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y, Urushidani T. 2010. The Japanese toxicogenomics project: Application of toxicogenomics. Mol Nutr Food Res. 54(2):218–227. doi:10.1002/mnfr.200900169.

Väremo L, Scheele C, Broholm C, Mardinoglu A, Kampf C, Asplund A, Nookaew I, Uhlén M, Pedersen BK, Nielsen J. 2015. Proteome- and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes. Cell Rep. 11(6):921–933. doi:10.1016/j.celrep.2015.04.010.

Vrbová M, Roušarová E, Brůčková L, Česla P, Roušar T. 2016. Characterization of acetaminophen toxicity in human kidney HK-2 cells. Physiol Res. 65(4):627–635.

Wacher VJ, Salphati L, Benet LZ. 2001. Active secretion and enterocytic drug metabolism barriers to drug absorption1PII of original article: S0169-409X(96)003304. The article was originally published in Advanced Drug Delivery Reviews 20 (1996) 99–112.1. Adv Drug Deliv Rev. 46(1–3):89–102. doi:10.1016/S0169-409X(00)00126-5.

Waikar SS, Bonventre JV. 2009. Creatinine Kinetics and the Definition of Acute Kidney Injury. J Am Soc Nephrol. 20(3):672–679. doi:10.1681/ASN.2008070669.

Wang J, Hou Y, Duan D, Zhang Q. 2017. The Structure and Nephroprotective Activity of Oligo-Porphyran on Glycerol-Induced Acute Renal Failure in Rats. Mar Drugs. 15(5). doi:10.3390/md15050135.

Wang X, Armando I, Upadhyay K, Pascua A, Jose PA. 2009. The regulation of proximal tubular salt transport in hypertension: an update. Curr Opin Nephrol Hypertens. 18(5):412–420. doi:10.1097/MNH.0b013e32832f5775.

Wang Y, Eddy JA, Price ND. 2012. Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. BMC Syst Biol. 6(1):153. doi:10.1186/1752-0509-6-153.

Watt IN, Montgomery MG, Runswick MJ, Leslie AGW, Walker JE. 2010. Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria. Proc Natl Acad Sci. 107(39):16823–16827. doi:10.1073/pnas.1011099107.

Weidemann MJ, Krebs HA. 1969. The fuel of respiration of rat kidney cortex. Biochem J. 112(2):149–166. doi:10.1042/bj1120149.

Weinberg JM, Humes HD. 1980. Mechanisms of gentamicin-induced dysfunction of renal cortical mitochondria. I. Effects on mitochondrial respiration. Arch Biochem Biophys. 205(1):222–231. doi:10.1016/0003-9861(80)90102-2.

Weinberg JM, Simmons F, Humes HD. 1980. Alterations of mitochondrial respiration induced by aminoglycoside antibiotics. Res Commun Chem Pathol Pharmacol. 27(3):521–531.

Wilson ID, Nicholson JK. 2017. Gut microbiome interactions with drug metabolism, efficacy, and toxicity. Transl Res. 179:204–222. doi:10.1016/j.trsl.2016.08.002.

Wishart DS, Feunang Yannick D., Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, et al. 2018. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 46(D1):D1074–D1082. doi:10.1093/nar/gkx1037.

Wishart DS, Feunang Yannick Djoumbou, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. 2018. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res. 46(D1):D608–D617. doi:10.1093/nar/gkx1089.

Wolf G, Hannken T, Schroeder R, Zahner G, Ziyadeh FN, Stahl RA. 2001. Antioxidant treatment induces transcription and expression of transforming growth factor beta in cultured renal proximal tubular cells. FEBS Lett. 488(3):154–159. doi:10.1016/s0014-5793(00)02403-0.

Xu F, Papanayotou I, Putt DA, Wang J, Lash LH. 2008. Role of mitochondrial dysfunction in cellular responses to S-(1,2-dichlorovinyl)-l-cysteine in primary cultures of human proximal tubular cells. Biochem Pharmacol. 76(4):552–567. doi:10.1016/j.bcp.2008.05.016.

Xu X, Liu Y, Lu L, Ke Y, Mao J, Mao K. 2012. Altered expression of hepatic metabolic enzyme and apoptosis-related gene transcripts in human hepatocytes treated with trichloroethylene. Hum Exp Toxicol. 31(9):861–867. doi:10.1177/0960327112444935.

Yang Li, Xiong A, He Y, Wang Zaiyong, Wang C, Wang Zhengtao, Li W, Yang Ling, Hu Z. 2008. Bile Acids Metabonomic Study on the CCl_4 - and α-Naphthylisothiocyanate-Induced Animal Models: Quantitative Analysis of 22 Bile Acids by Ultraperformance Liquid Chromatography−Mass Spectrometry. Chem Res Toxicol. 21(12):2280–2288. doi:10.1021/tx800225q.

Yizhak K, Gaude E, Le Dévédec S, Waldman YY, Stein GY, van de Water B, Frezza C, Ruppin E. 2014. Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer. eLife. 3. doi:10.7554/eLife.03641.

Yizhak K, Le Dévédec SE, Rogkoti VM, Baenke F, de Boer VC, Frezza C, Schulze A, van de Water B, Ruppin E. 2014. A computational study of the Warburg effect identifies metabolic targets inhibiting cancer migration. Mol Syst Biol. 10:744. doi:10.15252/msb.20134993.

Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. OMICS J Integr Biol. 16(5):284–287. doi:10.1089/omi.2011.0118.

Zhang A-D, Dai S-X, Huang J-F. 2013. Reconstruction and Analysis of Human Kidney-Specific Metabolic Network Based on Omics Data. BioMed Res Int. 2013:1–11. doi:10.1155/2013/187509.

Zierer J, Menni C, Kastenmüller G, Spector TD. 2015. Integration of ‘omics’ data in aging research: from biomarkers to systems biology. Aging Cell. 14(6):933–944. doi:10.1111/acel.12386.

Zimmerman HJ. 1999. Hepatotoxicity: the adverse effects of drugs and other chemicals on the liver. 2. ed. Philadelphia: Lippincott Williams & Wilkins.

Zuber R, Anzenbacherová E, Anzenbacher P. 2002. Cytochromes P450 and experimental models of drug metabolism. J Cell Mol Med. 6(2):189–198. doi:10.1111/j.1582-4934.2002.tb00186.x.

Zur H, Ruppin E, Shlomi T. 2010. iMAT: an integrative metabolic analysis tool. Bioinforma Oxf Engl. 26(24):3140–3142. doi:10.1093/bioinformatics/btq602.