APPLICATIONS OF MASS SPECTROMETRY-BASED METABOLOMICS TO ADDRESS BIOMEDICAL PROBLEMS

A Dissertation

Presented to the Faculty of the Weill Cornell Graduate School

of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Benjamin I. Schwartz

May 2019

APPLICATIONS OF MASS SPECTROMETRY-BASED METABOLOMICS TO ADDRESS BIOMEDICAL PROBLEMS

Benjamin I. Schwartz, Ph.D

Cornell University 2019

The cadre of small molecule metabolites in a cell, aka the "metabolome," is the final output of the genome and ultimate determinant of cell phenotype (DNA → RNA → protein → metabolite → phenotype). In addition to a genetic contribution, the metabolome is conditioned by diet, lifestyle and composition of the microbiome - ultimately determining health vs. disease status. Therefore, the profiling of small molecule levels using mass spectrometry-based metabolomic analyses can contribute significantly to our understanding of disease processes, enable confident molecular diagnoses, monitor the efficacy of therapies, and serve many other biomedical applications. Here, we discuss applications and considerations for the use of targeted and untargeted metabolite profiling using mass spectrometry. This topic is presented in the context of three independent research studies, described herein. First, we describe the use of targeted metabolomics to characterize the endocannabinoid system in a mouse model of alcohol "binge" drinking. We show that anandamide, one of the primary endogenous agonists of the cannabinoid receptor, along with several other endocannabinoids, increase in response to acute alcohol withdrawal. Further, we consider the biological significance of these observations. Second, we describe the use of mass spectrometry-based untargeted metabolite profiling and stable isotope tracing for defining metabolic perturbations that occur in the setting of sporadic amyotrophic lateral sclerosis (sALS), considering discoveries made with patient-derived skin fibroblasts. Using a multiomic approach, this study identifies and characterizes a

distinct subgroup of sALS patients, possessing fibroblasts that are typified by enhanced transsulfuration pathway activity and glucose hypermetabolism. We speculate that this sALS patient subclass will prove to be selectively responsive to anti-oxidant therapies and therefore a recognition of this patient subclass may allow for the future establishment of personalized medicines. Finally, we discuss efforts to characterize solute carrier 25 (SLC25) family mitochondrial transporters using CRISPR knockout cell lines and a heterologous bacterial overexpression system-based protocol. These studies utilize mass spectrometry and stable isotope-labeled compounds in attempt to recognize potential biological substrates for these transporters.

BIOGRAPHICAL SKETCH

Benjamin Schwartz was born in Baltimore, Maryland. After completing his schoolwork at Yeshivat Rambam high school in 2007, he spent two years studying at the Kerem B'yavneh Yeshiva in Givat Washington, Israel. He attended Yeshiva University beginning in 2009 and received a Bachelor of Arts in Chemistry in 2012. In August 2013, he entered the Pharmacology Ph.D program at Weill Cornell Medicine in New York City.

To my parents, who are excited by everything I do.

ACKNOWLEDGEMENTS

Chapter 1

I would like to thank Yan Zhou for involving me in this project and for the extensive contributions he made to project design and mouse behavioral studies. I am grateful to Joanna Giza, Steven Gross, Francis Lee and Mary Jeanne Kreek for their contributions. YZ designed the study, conducted behavioral studies, wrote the protocol, managed the literature searches and analyses, undertook the statistical analysis, and wrote the manuscript; BIS, SSG designed and conducted the experiments on tissue eCB measurements; JG, FSL designed and conducted the experiments on FAAH Western blot; MJK contributed to the final versions of manuscript writing. Special thanks to Austin Kordic for providing the comments and corrections on the manuscript. I also thank the following funding sourcing for supporting this work: NIH AA021970 (YZ), NS095698 (BIS), NHLBI 528915035 (SSG), Pritzker Neuropsychiatric Disorders Research Consortium (FSL) and Dr. Miriam and Sheldon G. Adelson Medical Research Foundation (MJK).

Chapter 2

I would like to thank Csaba Konrad for the heroic amount of work he put into helping prepare the large number of samples involved in my work on this project. Thank you to Qiuying Chen and Davinder Sandhu for working towards making sense of all the data that was produced from those samples and for the design of follow-up experiments. I am grateful to Dipa Roychoudhury, Roger Cheng, Kirsten Bredvik, Hibiki Kawamata, Elizabeth Caulder, Lorenz Studer, Steven Fischer, Giovanni Manfredi and Steven Gross for their contributions. DS, BIS, RRC performed metabolite extraction and targeted metabolite analysis, DR and SMF analyzed the

transcriptomics data. CK, KB, and HK performed cell culture and the fibroblast bioenergetics analysis. ELC and LC contributed to the data interpretation and concepts presented; QC, GM and SSG conceived the experiments, analyzed the data and wrote the original manuscript draft for editing by co-authors. Thank you to Dr. Hiroshi Mitsumoto (Columbia University) and the COSMOS initiative for providing the fibroblast and plasma samples utilized in this work. I also thank the following funding sourcing for supporting this work: NIH R01NS093872 (GM, SG, and LS), R01NS062055 (GM), 1 F31 NS095698-01 (BS).

Chapter 3

I particularly thank Kristian Laursen for teaching me molecular biology from the ground up and for advising me in all aspects of this project. I could not have progressed as quickly as I did without his selfless help. I thank Leona Nease for her help getting the *Lactococcus Lactis* protein expression system working. I thank Qiuying Chen for her guidance with MS instrumentation and stable isotope tracing experiment data. Thank you to Kivanc Birsoy for providing *L. Lactis* bacteria and the pNZ8048 plasmid.

Generally, I thank past and current members of the Gross lab for their comradery and support: Darya Akimova, Fengli Zhang, Joshua Zuk, Changyuan Lu Qiuying Chen, Davinder Sandhu, Miriam Sindelar, Evangelia Zgonis, Ismail Ismailoglu, Eugene Koblik and Yuliang Ma. I hope their eardrums can heal once I am gone. Special thanks to Miriam and Qiuying for MS instrumentation assistance and to Josh for tending to my samples periodically.

TABLE OF CONTENTS

Page

Acknowledgments v

List of Figures xvi

List of Tables xxii

Introduction: Mass Spectrometry-based Metabolomics 1

Metabolomics: small molecules, big deal 1

Metabolite extraction and sample preparation for metabolomics analysis 3

Chromatography 9

Targeted vs. untargeted 11

Metabolite identification 13

Stable isotope tracing 15

Data processing, normalization, analysis, interpretation and visualization 17

New applications and methods 21

Conclusion 23

References 26

Chapter 1: Blockade of alcohol escalation and “relapse” drinking by pharmacological FAAH inhibition in male and female C57BL/6J mice 34

Introduction 34

Materials and methods....37
Animals....37
Materials....38
Procedures....38
Chronic intermittent access (IA)....38
Chronic (3-week) drinking-in-the-dark (DID)....41
Sucrose (caloric reinforcer) and saccharin (non-caloric reinforcer) drinking....43
The alcohol deprivation effect (ADE) after 1-week withdrawal from 3-week IA....44
N-Acyl Ethanolamide (NAE) extraction for mass spectrometry (MS) quantification....46
Western blot....50
Data analysis....50
Results....54
Single, acute administration of URB597 reduced alcohol (but not sucrose) intake and preference after 1 day of withdrawal from chronic IA....54
Single, acute URB597 at 0.5 mg/kg reduced 15% alcohol intake and preference in male and female mice....54
Dose-response of URB597 in male mice....55

Single, acute URB597 (0.5 mg/kg) reduced 30% (but not 7.5%) alcohol intake and preference in male mice 55

Pretreatment with selective CB1R antagonist AM251 (0.1 mg/kg) blocked the effect of acute URB597 (0.5 mg/kg) on reducing 15% alcohol drinking in male mice 58

No effects of single, acute URB597 (0.5 mg/kg) on sucrose intake and preference in male mice 60

Single, acute URB597 (0.5 mg/kg) increased saccharin preference ratio, but not intake in male mice 60

Effect of single, acute or repeated administration of URB597 on ADE in both male and female mice 62

Single, acute URB597 62

Repeated URB597 64

Acute (1 day) alcohol withdrawal transiently caused global brain eCB system dysregulation in male mice 66

Acute (1 day) withdrawal 66

Chronic IA and long-term withdrawal 66

FAAH levels in the prefrontal cortex after withdrawal from chronic IA in male mice 70

Discussion.....70

Supplementary information.....78

References 82

Chapter 2: Identification of a distinct metabolomic subtype of sporadic ALS patients....90

Introduction....90

Materials and methods....93

Reagents....93

Cell Culture....93

Metabolite extraction....94

Untargeted metabolite profiling by LC/MS....95

Metabolite Structure Specification....95

Stable isotope tracing of [U-13C] glucose and [2,3,3-2H] serine....96

Microarray analysis....96

Statistics....97

Results....97

Untargeted metabolite profiling of human skin fibroblasts identifies an sALS subgroup....97

The sALS-1 subgroup has increased serine incorporation into glutathione....101

The sALS-1 subgroup has increased glucose incorporation into glutathione and TCA cycle intermediates....104

sALS-1 subgroup transcriptomic changes support the metabolomic findings 107

sALS-1 subgroup plasma displays altered trans-sulfuration pathway metabolites 113

Discussion 116

Supplementary information 120

References 123

Chapter 3: Towards characterizing SLC25 family mitochondrial transporters with mass spectrometry-based metabolomics 129

Introduction 129

Materials and methods 135

Reagents 135

CRISPR knockout cell line production 136

Plasmid construction for CRISPR genome editing (eSpCas9) 136

Generation of knockout cell lines 136

Cut-site screening 137

Protein reinsertion cell line production 137

Plasmid construction for gene reinsertion 137

Generation of reinsertion cell lines 137

MDA-MB-468 derived cell line metabolite profiling 140

Cell culture 140

Metabolite extraction.....140

Untargeted metabolite profiling by LC/MS 141

Metabolite structure specification 141

Stable isotope tracing 142

Next-Generation Amplicon Sequencing of SLC25 knockout cell line mRNAs. 142

Transporter overexpression in *Lactococcus Lactis* 142

Electrocompetent *Lactococcus Lactis* production.....142

Plasmid construction for gene reinsertion 143

Transporter overexpression 144

Western blot and dot-blot for FLAG tagged transporters ...144

Metabolite profiling of *Lactococcus Lactis*.....145

Initial metabolite extraction protocol 145

Improved metabolite extraction protocol 145

Results 147

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A38* knockout cell line reveals modest metabolite changes compared to parental control 147

Metabolite profiling validation using a second *SLC25A38* knockout cell line shows a less significant decrease in N-acetylglycine 150

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A40* knockout cell line reveals modest metabolite changes compared to parental control 150

MS/MS fragmentation confirms betaine identity 153

Heavy isotope labeled (9 deuterated) choline tracing confirms decreased choline-derived betaine in the *SLC25A40* knockout cell line 153

Choline tracing of additional *SLC25A40* knockout cell lines shows unchanged choline-derived betaine abundance 161

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A39* knockout cell line reveals highly reproducible metabolite changes compared to parental control 161

Metabolite profiling of additional *SLC25A39* knockout cell lines show no profound shared metabolite changes and unchanged guanidinoacetate abundance.....166

Metabolite profiling of SLC25A39 EGFP tagged reinsertion cell lines do not rescue the observed metabolic perturbations 166

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A44* knockout cell line reveals highly reproducible unidentified metabolite changes compared to parental control.....170

Next-generation sequencing of SLC25 knockout cell line mRNAs shows no indication of expression of additional alleles but reveals presence of alternate transcript splice variants 170

Human SLC25A38, SLC25A39, SLC25A40 and SLC25A44 are successfully expressed in *Lactococcus Lactis* bacteria 178

Untargeted metabolite profiling of SLC25A38, SLC25A39, SLC25A40 and SLC25A44 (both FLAG and non-FLAG tagged) expressing *Lactococcus Lactis* shows metabolites with strongly increased abundance in specific bacterial lines, likely protein related peptides.....178

Most human SLC25A family carboxylate transporters are successfully expressed in *Lactococcus Lactis* bacteria 182

Transport assays with SLC25A11 expressing bacteria (with and without FLAG tag) fail to show increased 2-oxoglutarate transport with protein expression 182

Transport assays with FLAG tagged SLC25A21 expressing bacteria fail to show increased 2-oxoglutarate transport.....190

Transport assays with SLC25A38 expressing bacteria (with and without FLAG tag) fail to show increased glycine transport ...194

Transport assays with SLC25A40 expressing bacteria (with FLAG tag) fail to show increased glycine transport with protein expression.....194

Transport assays with SLC25A8 expressing bacteria (without FLAG tag) fail to show increased aspartate transport with protein expression....194

Discussion....199

SLC25A38....200

SLC25A40....201

SLC25A39....202

SLC25A44....203

Suitability of knockout cell lines as a system for untargeted mitochondrial transporter studies....204

Lactococcus Lactis transporter expression studies....211

References....222

Next Steps....229

Chapter 1....229

Chapter 2....230

Chapter 3....231

LIST OF FIGURES

Page

Figure I.1: Interest in and use of mass spectrometry metabolomics is growing rapidly....2

Figure I.2: Schematic overview of a typical mass spectrometry-based metabolite profiling workflow....4

Figure I.3: Two water washes of adherent cells (skin fibroblasts) are ideal to reduce background ion suppression....6

Figure I.4: Features produced in-source can be algorithmically identified....19

Figure 1.1: Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 15% alcohol intake and preference ratio in male (A and B, n=7-8) and female (C and D, n=8) mice after 1 day of withdrawal from 3-week chronic intermittent access alcohol drinking....56

Figure 1.2: Dose responses of single, acute administration of URB597 (URB, 0, 0.125-1 mg/kg) on reducing 15% alcohol intake (A) and preference ratio (B) in male mice (n=6-7) after 1 day of withdrawal from 3-week chronic intermittent access alcohol drinking....57

Figure 1.3: Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 0.1% or 0.2% saccharin intake (A) and preference ratio (B) in male mice (n=6) after 3-week chronic intermittent access alcohol drinking....61

Figure 1.4: Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on alcohol intake in an alcohol deprivation effect (ADE) model at 4

hours in male (A, n=7) and female (B, n=6-9) mice after 1 week of withdrawal from 3-week chronic intermittent access alcohol drinking 63

Figure 1.5: Effects of five repeated administration of URB597 (URB, 0.5 mg/kg) on alcohol intake in an alcohol deprivation effect (ADE) model at 4 hours in male (A, n=6) and female (B, n=7-8) mice after 1 week of withdrawal from 3-week chronic intermittent access drinking 65

Figure 1.6: Effects of 1-day withdrawal from 3-week chronic intermittent access (IA) alcohol drinking on brain NAE abundances in male mice (n=6) 67

Figure 1.7: Anandamide (AEA) abundances after chronic 3-week chronic intermittent access (IA) alcohol drinking or long-term withdrawal in male mice 68

Figure 1.8: No FAAH changes in male mice (n=6) after acute or long-term withdrawal from 3-week chronic intermittent access (IA) alcohol drinking 69

Figure S1.1: Effects of 1-day withdrawal from 3-week chronic intermittent access (IA) alcohol drinking on brain NAE abundances.....80

Figure S1.2: No NAE changes after chronic IA or long-term withdrawal.....81

Figure 2.1: Mass spectrometry-based untargeted metabolite profiling of patient fibroblast cell lines reveals an sALS subgroup 99

Figure 2.2: Serine tracing shows increased GSH, GSSG and cystathionine synthesis in the sALS-1 subgroup 102

Figure 2.3: Glucose tracing shows increased glutamate, GSH, and GSSG synthesis in the sALS-1 subgroup 105

Figure 2.4: Glucose tracing shows increased incorporation into A) lactate and TCA cycle intermediates (B) citrate, C) oxoglutarate, D) succinate, E) fumarate and F)malate) in the sALS-1 subgroup fibroblasts 108

Figure 2.5: Glucose tracing shows increasd incorporation into amino acids (A) alanine and B) aspartate) and acylcarnitines (C) acetylcarnitine and D) Butyrylcarnitine) in the sALS-1 subgroup fibroblasts 110

Figure 2.6: An integrated analysis of transcriptomic (microarray) and metabolomics data reveal common pathways that are perturbed in sALS-1 fibroblasts 111

Figure 2.7: Untargeted metabolite profiling of plasma reveals differentially abundant metabolites for sALS-1 subgroup vs. control that support the multiomic results from the same patients' fibroblast cell lines (abnormal trans-sulfuration related metabolites) 114

Figure S2.1: Heatmap of metabolites found to be differentially abundant in both the original and repeat fibroblast metabolite profiling experiments 120

Figure S2.2: Gene Set Enrichment analysis (GSA) identified significantly different transcript levels of genes involved in the regulation of catabolic processes, apoptosis and epigenetic regulations in sALS-1 compared to sALS-2 and control groups 121

Figure S2.3: The sALS-1 subgroup showed increased mitochondria membrane potential and mitochondria membrane potential to mass ratio compared to control and non-stratified (sALS-2) patients 122

Figure 3.1: Reproduced from [Palmieri 2013]. Phylogenetic tree of human SLC25 family mitochondrial transporters 132

Figure 3.2: Sequence alignment of knockout cell line alleles with wild-type sequence for two cell lines for *SLC25A38*, *SLC25A39* and *SLC25A40* 138

Figure 3.3: Comparison of the abundance of N-acetylglycine in *SLC25A38* KO cells compared to control cells 149

Figure 3.4: Comparison of the abundance of N-acetylglycine in *SLC25A38* KO cells compared to control cells 151

Figure 3.5: Comparison of the abundance of betaine in *SLC25A40* KO cells compared to control cells..... 155

Figure 3.6: MS/MS fragmentation spectra of m/z 118.08 at collision energies of 0 (top panel), 40 (2nd panel), 20 (3rd panel) and 10 (bottom panel)..... 156

Figure 3.7: Comparison of the MS/MS fragmentation of m/z 118.08 (pos, bottom panel) to the Metlin fragmentation database for betaine (top panel) confirms the compound's identity as betaine 158

Figure 3.8: Choline tracing reveals reduced choline-derived betaine in *SLC25A40 KO* cells..... 159

Figure 3.9: Comparison of the abundance of deuterated betaine in SLC25A40 KO cells compared to control cells 162

Figure 3.10: Comparison of the abundance of guanidinoacetic acid in *SLC25A39* KO cells compared to control cells 165

Figure 3.11: Comparison of the abundance of guanidinoacetate in *SLC25A39* KO cells compared to control cells 167

Figure 3.12: Representative confocal fluorescent image of SLC25A39-EGFP expressing cells..... 168

Figure 3.13: Comparison of the abundance of guanidinoacetate in *SLC25A39* KO cells compared to A39-EGFP reinsertion clonal lines....169

Figure 3.14: Next-Generation Amplicon Sequencing of knockout cell line mRNA supports 'TA-cloning' genotyping results but reveals high prevalence of alternate splicing 171

Figure 3.15: Western blots showing expression of SLC25A38, SLC25A39, SLC25A40 and SLC25A44 with C-terminal FLAG tags....179

Figure 3.16: Examples of metabolites that were greatly increased in one of the protein expressing lines compared to the others....181

Figure 3.17: Protein expression of most of the SLC25A family carboxylate transporters 183

Figure 3.18: SLC25A11-FLAG OE and empty vector *Lactococcus Lactis* 30 min time course transport assay 184

Figure 3.19: SLC25A11-FLAG OE *Lactococcus Lactis* substrate concentration test....186

Figure 3.20: SLC25A11-FLAG OE and empty vector *Lactococcus Lactis* 30 min transport assay....187

Figure 3.21: Various numbers of 3 second washes were tested to determine the optimal number of washes to remove excess external substrate without internal substrate leakage 188

Figure 3.22: SLC25A11-FLAG OE and empty vector *Lactococcus Lactis* 30 min transport assay....189

Figure 3.23: SLC25A11-FLAG OE and empty vector *Lactococcus Lactis* transport assay 191

Figure 3.24: SLC25A11 OE and empty vector *Lactococcus Lactis* transport assay 192

Figure 3.25: SLC25A11 OE and empty vector *Lactococcus Lactis* transport assay 193

Figure 3.26: SLC25A21-FLAG OE and empty vector *Lactococcus Lactis* 30 min time course transport assay 195

Figure 3.27: SLC25A38 OE and empty vector *Lactococcus Lactis* transport assays 196

Figure 3.28: SLC25A40-FLAG OE and empty vector *Lactococcus Lactis* transport assays 197

Figure 3.29: SLC25A8 OE and empty vector *Lactococcus Lactis* transport assays 198

LIST OF TABLES

Page

Table 1.1: The 3-week chronic intermittent access (IA) alcohol drinking model (15% alcohol vs. water, 24 hours) every other day with selective FAAH inhibitor URB597, CB1 antagonist AM251 or their combination 39

Table 1.2: The 3-week chronic alcohol drinking-in-the-dark (DID) model (15% alcohol, 4 hours) every day with selective FAAH inhibitor URB597.....42

Table 1.3: The alcohol deprivation effect (ADE) model after 1-week withdrawal from a 3-week chronic intermittent access (IA) drinking (15% alcohol vs. water) with selective FAAH inhibitor URB597.....45

Table 1.4: LC gradient used to separate NAEs 49

Table 1.5: Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 7.5% or 30% alcohol intake and preference ratio at 4 hours in male mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking.....52

Table 1.6: Pretreatment with selective CB1R antagonist AM251 (0.1 mg/kg) blocks the effect of single, acute URB597 (URB, 0.5 mg/kg) on reducing 15% alcohol drinking in male mice after 4 hours of alcohol access.....53

Table 1.7: No effect of single, acute administration of URB 597 (URB, 0.5mg/kg) on 2% (A) and 4% (B) sucrose intake and their preference ratio in male mice after 4 hours of sucrose access.....59

Table S1.1: No effects of single, acute administration of URB 597 (URB, 0.5mg/kg) on total fluid intake in male mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking.....78

Table S1.2: Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 15% alcohol intake and preference ratio in both male (presented in Figure 1) and female (presented in Figure S1) mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking at 4 hours.....79

Table 3.1: Statistically significant, differentially abundant features of an unbiased metabolomics experiment of *SLC25A38* KO cells compared to parental control cells.....148

Table 3.2: Statistically significant, differentially abundant features of an unbiased metabolomics experiment of a second *SLC25A38* KO cell line compared to parental control cells.....152

Table 3.3: Statistically significant, differentially abundant features of an unbiased metabolomics experiment of *SLC25A40* KO cells compared to parental control cells.....154

Table 3.4: Statistically significant, differentially abundant features of two unbiased metabolomics experiment of *SLC25A39* KO cells compared to parental control cells.....164

Table 3.5: Overview of most prevalent ‘other’ sequences detected with Next-Generation Amplicon Sequencing.....175

Table 3.6: Overview of in-frame sequence read percentage detected with Next-Generation Amplicon Sequencing.....177

INTRODUCTION: MASS SPECTROMETRY-BASED METABOLOMICS

Metabolomics: small molecules, big deal

Metabolomics, or metabolite profiling, seeks to survey diverse small molecules in a biological sample for changes in relative abundances. Importantly, small molecule levels are the most direct determinant of phenotype in biological systems, distinguishing physiology from pathophysiology [Aretz et al. 2016]. For this reason, the study of the ebbs and flows of metabolite abundances is crucial for furthering our understanding of practically all normal and abnormal biological processes. However, metabolomic analyses present unique challenges - unlike DNA, RNA and proteins that are built of a defined set of repeating elements, metabolite structures and their physical properties are extremely varied, incompletely defined, and turnover can be rapid [de Koning et al. 1992; Lu et al. 2017]; the Human Metabolome Database (HMDB) includes >8000 known detected compounds and >80,000 compounds that are predicted [Wishart et al. 2018]. This makes the isolation, detection and interpretation of abundances for these important biological players a distinctly complicated task, explaining why access to this knowledge has only recently become accessible. Indeed, over the past few years, technological advances have enabled remarkable ability to quantify metabolites with ever-increasing ease, accuracy and sensitivity and have pushed metabolomics to the forefront of next-generation research [Beger et al. 2016].

Mass spectrometry is a powerful analytical tool for metabolism-centered biomedical research and clinical application that has gained considerable prominence and capabilities over the past decade (Figure I.1).

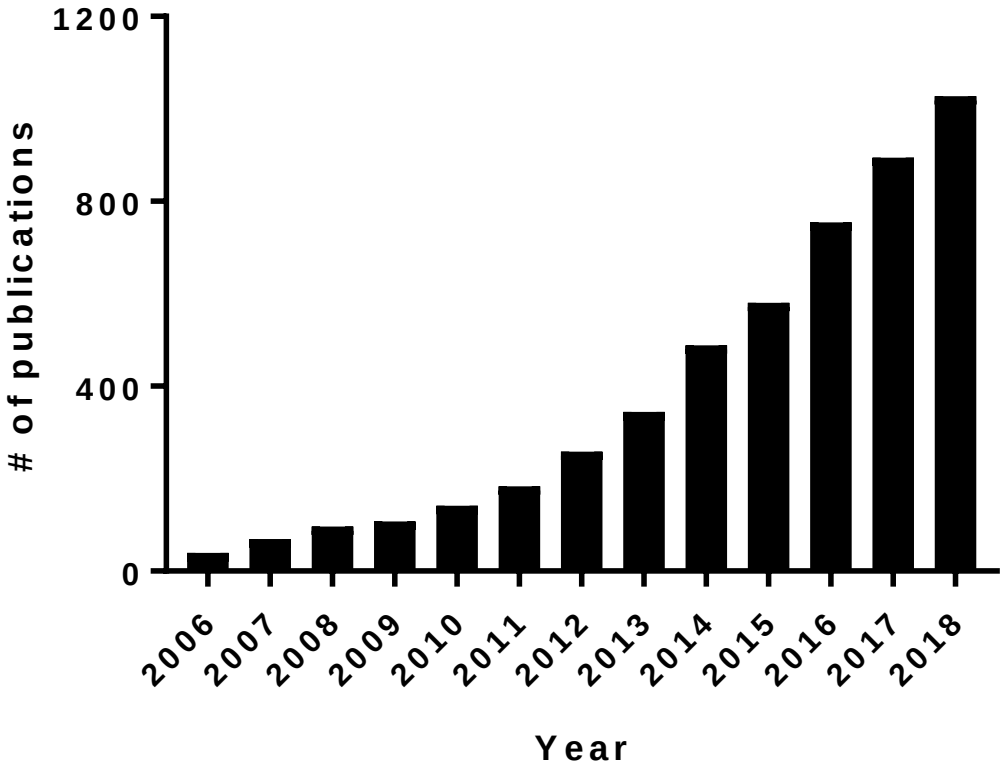


Figure I.1. *Interest in and use of mass spectrometry metabolomics is growing rapidly. Number of publications per year containing "mass spectrometry" and "metabolomics" in the title or abstract are shown (PubMed).*

Mass spectrometry detection relies on the ionization of compounds and quantification as *ion counts*, based on mass-to-charge ratio (m/z). The ability to quantify metabolite abundances using mass spectrometry in either a targeted or untargeted manner provides for both hypothesis-testing and hypothesis-generation. In this *Introduction*, I discuss the current state of mass spectrometry-based metabolite profiling (aka *metabolomics*), considering sample preparation, liquid chromatography first-dimension separations, data collection, data analysis and data interpretation. I will also discuss related topics, such as stable isotope tracing, multi-omic data integration and next-generation MS applications/methods. In the three chapters that follow this *Introduction*, I will describe research applications that include the use of both targeted and untargeted metabolite profiling, highlighting examples of the triumphs and challenges of mass spectrometry-based metabolite profiling strategies in biomedical research.

Metabolite extraction and sample preparation for metabolomics analysis

Once the overall experiment design for a metabolomic study has been established, the next step in experimental workflow is to define an appropriate sample preparation strategy (Figure I.2). Metabolite profiling is extremely versatile and can be performed on a variety of samples for a vast number of purposes, ranging from environmental toxicology to cancer biomarker discovery [Bundy et al. 2009; Armitage et al. 2014]. For biomedical research, metabolites are most often harvested for analysis from biological fluids (e.g. urine, plasma, CSF), tissue biopsies (e.g. skin, liver, kidney) and cells grown in culture.

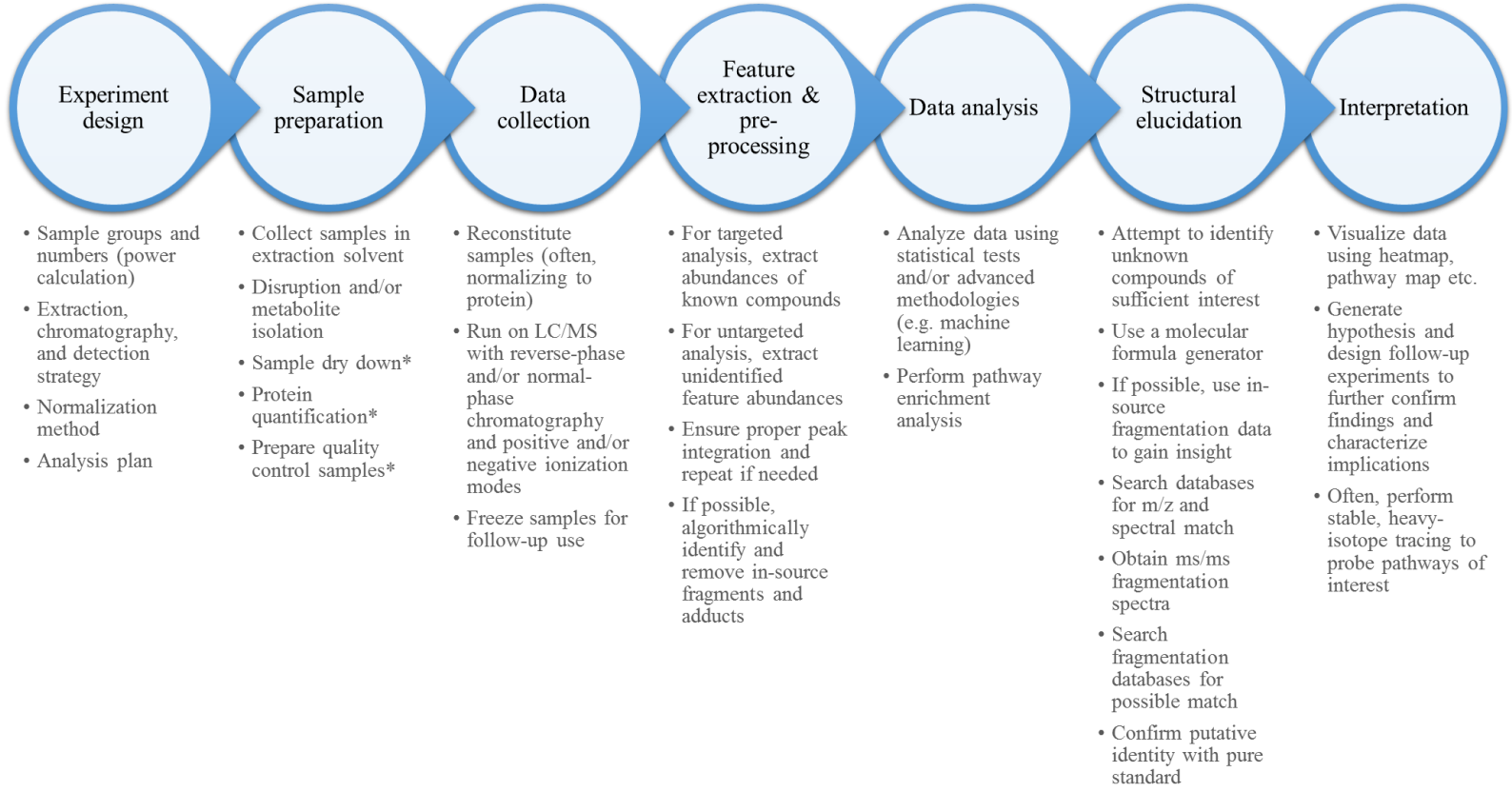


Figure I.2. *Schematic overview of a typical mass spectrometry-based metabolite profiling workflow. * denotes "if applicable"*

Metabolites of interest are commonly extracted from a relevant biological source using a suitable solvent and disruption method. For biofluids, samples can generally be directly diluted into a suitable solvent, diluting interfering salts and precipitating insoluble proteins/DNA, then centrifuged and analyzed using liquid chromatography(LC)-coupled mass spectrometry [Vuckovic et al. 2012; Vuckovic et al. 2013]. We employ this simple methodology for plasma metabolomics in Chapter 2 and show sufficient metabolite coverage for comparison with cell extract-derived metabolite findings. The relatively slow LC component of LC/MS-based profiling can sometimes be omitted and replaced by direct flow-injection of sample extracts for MS-only analysis. Alternatively, more complicated metabolite enrichment protocols, such as solid phase extraction (SPE), may be required for detection of particular metabolites of interest [Dettmer et al. 2007].

For analysis of adherent cells grown in culture, culture media is commonly removed by aspiration, and cells are then washed sequentially with a buffer (e.g. phosphate-buffered saline; PBS), to limit metabolite carry-over from culture medium, and ddH_2O, to reduce salt levels that otherwise diminish MS sensitivity. In an experiment comparing the efficacy of 1, 2 or 3 rapid water washes, I determined that 2 rapid water washes is optimal for reducing ion suppression while providing for reproducible detection of a broad array of endogenous metabolites (Figure I.3). Often, ice cold buffer and water are used to slow metabolism during the process of metabolite extraction, although the effects of cold-shock on cell metabolism warrants consideration [Lu et al. 2017; Wittmann et al. 2004; Wellerdiek et al. 2009].

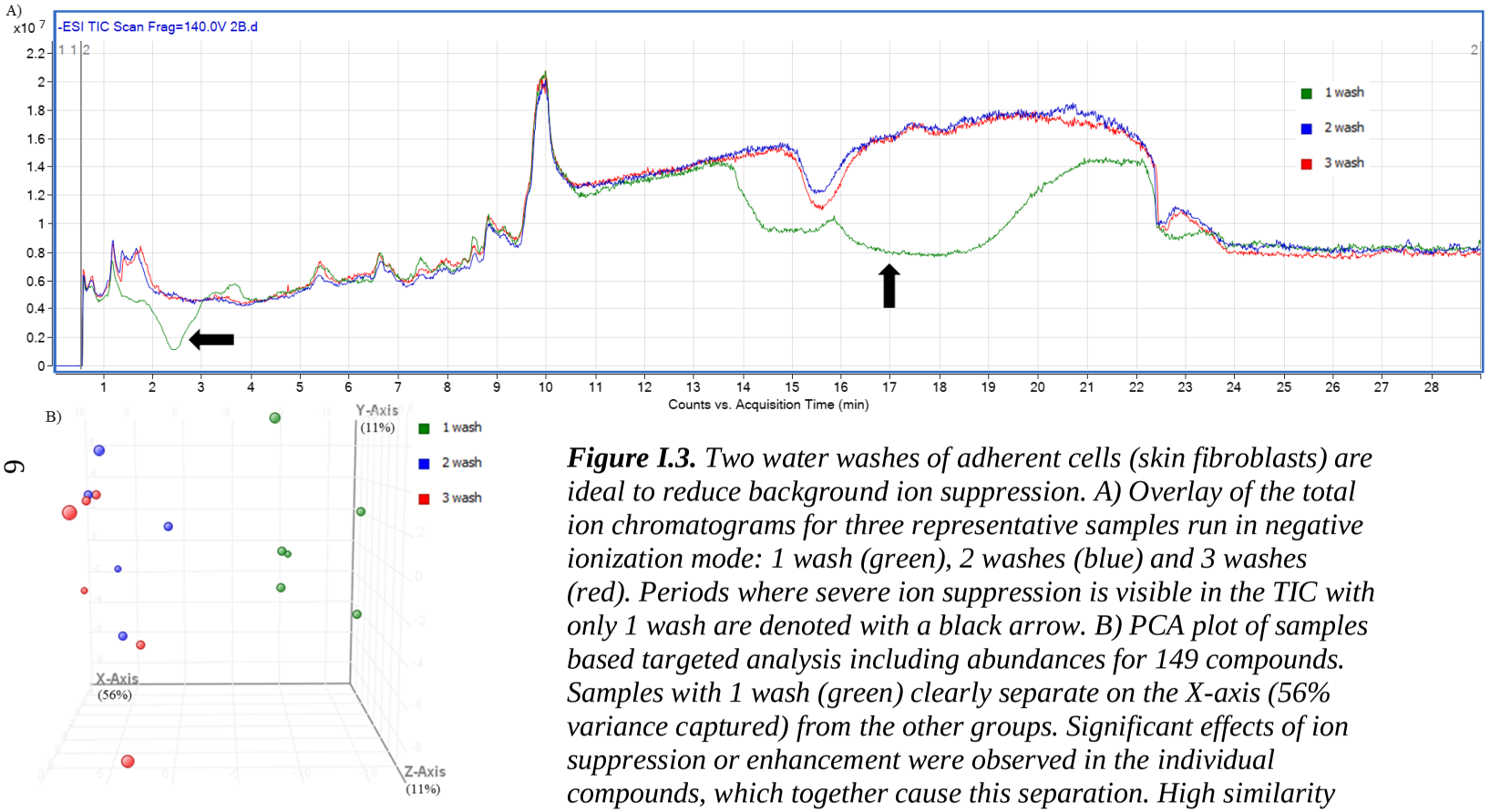


Figure I.3. *Two water washes of adherent cells (skin fibroblasts) are ideal to reduce background ion suppression. A) Overlay of the total ion chromatograms for three representative samples run in negative ionization mode: 1 wash (green), 2 washes (blue) and 3 washes (red). Periods where severe ion suppression is visible in the TIC with only 1 wash are denoted with a black arrow. B) PCA plot of samples based targeted analysis including abundances for 149 compounds. Samples with 1 wash (green) clearly separate on the X-axis (56% variance captured) from the other groups. Significant effects of ion suppression or enhancement were observed in the individual compounds, which together cause this separation. High similarity was observed between 2 and 3 water wash groups, indicating that 2 washes is an ideal.*

Regardless, cold wash steps should be as swift as possible in order to both rapidly quench metabolism and minimize the potential cell-leakage of otherwise informative metabolites.

Following all cell-wash steps, metabolites can be extracted by scrape-harvesting cells in a suitable solvent, often water with an organic solvent or a mixture of organic solvents (e.g. 80:20 methanol:water). Dettmer et al. explored the use of various extraction solvents with adherent mammalian cells and determined that ultra-cold 80:20 methanol:water maximizes overall metabolite yields [Dettmer et al. 2011]. Cold extraction solvents, to quench metabolism, or hot solvents, to denature proteins can alternatively be employed for metabolomic sample preparation [Dettmer et al. 2011; Sellick et al. 2010; Hiller et al. 2007; Winder et al. 2008]. Consistent with these findings, I employ -70°C 80% methanol as an extraction solvent for untargeted metabolomics experiments unless application-specific knowledge dictates otherwise (e.g. to assess changes in the abundance of specific lipid classes). Depending on the cell/tissue type, a disruption method may be additionally needed for efficient extraction of metabolites of interest. Cell/tissue extraction methods may include vortexing, cycles of free/thaw, sonication, and bead-beating. In any case, care must be taken to maintain a low temperature during extraction and to minimize processing time in order to limit potential metabolite degradation. Cell debris and precipitated proteins are removed from sample extracts by centrifugation.

For metabolomic analysis of tissues, tissue are dissected, minced and flash-frozen in either liquid nitrogen, or dry ice/acetone. Various homogenization techniques can be subsequently employed, but stainless steel bead-beating is commonly use (as herein) as an efficient technique that allows multiple samples to be

extracted simultaneously [Beckonert et al. 2007; Römisch-Margl et al. 2012; Wu et al. 2008; Want et al. 2013]. The method and duration of homogenization for metabolomic sample preparation is tissue-dependent. For example, brain tissue, which is relatively soft, requires far less physical force for effective disruption than kidney, a relatively connective tissue-rich tissue. After metabolite extraction, remaining debris and precipitated protein/DNA can be effectively pelleted and removed by centrifugation.

Depending on the specific metabolite profiling application, sample pre-treatment can be performed using techniques that include liquid-liquid extraction, solid phase extraction, and electromigration [Raterink et al. 2014]. These methods can serve to enrich the metabolites of interest by diminishing contaminant levels or isolating compound classes of specific interest. With SPE, for example, metabolite extracts are exposed to a resin that selectively retains certain classes of molecules (such as hydrophobic lipids on a C18 resin), allowing for the removal of salts and other impurities that may interfere with MS data acquisition. Notably, these strategies can only be employed if specific compounds with similar properties are of interest and not in a completely untargeted experiment. The benefits of increased purity and metabolite enrichment should be weighed against the required time and expense and effect on the integrity of samples for analysis. As I show in subsequent Chapters 1-3, further purification is often not necessary to quantify relative abundances of a diverse array of metabolites in biosamples.

As noted above, metabolite extracts must often be concentrated to increase the detection of compounds that are either present in relatively low abundance, or which ionize poorly. To address this potential analytical limitation, metabolite extracts are often dried using a vacuum centrifuge or lyophilizer [Lu et al. 2017]. This also allows

samples to be normalized by varying reconstitution volumes (e.g. based on protein or DNA content) to correct for variation in cell number or tissue mass. Toward this end, samples are reconstituted in a suitable solvent for injection into the LC/MS analytical system. The degradation of sample metabolites (yielding new non-biologically-derived metabolites) is a key concern at this step. If a metabolite of interest is known to be sensitive to sample dry-down, this sample preparation step will obviously need to be avoided. Regardless, samples from all treatment groups must be directly comparable and prepared in an identical manner. Ideally, following data analysis, metabolites of interest should be confirmed to not be artifacts of sample dry-down.

Chromatography

Chromatography is often employed as a first-dimension metabolite separation approach prior to electrospray ionization for MS-based analyte quantification. There are two primary reasons for a chromatography front end. First, it attenuates ionization-related issues, such as ion suppression and enhancement, thereby increasing MS detection sensitivity and reducing data variation. Metabolites compete for charge during the ionization process and, therefore, separating metabolites based on chromatographic retention time (RT) increases the extent to which metabolites enter the MS system for detection [Annesley 2003; Jessome et al. 2006; Furey et al. 2013]. Second, although mass spectrometers have achieved impressive resolution capabilities, metabolites with sufficiently similar or identical masses (in the case of structural isomers) cannot be discerned based on m/z alone. In these cases,

chromatographic retention time adds another dimension for confident compound identification [Jang et al. 2018].

There are two major types of chromatography used in combination with MS, liquid chromatography (LC) and gas chromatography (GC), which both have advantages and disadvantages. GC can be used for separation of volatile compounds, or those which can be chemically derivitized to be made volatile. It is highly reproducible, including reproducible fragmentation that leads to easier compound identification using available databases (e.g. NIST). However, derivatization to produce volatile compounds can be time consuming and can lead to confounding factors, such as incomplete conversion or generation of multiple species for a given metabolite. Also, the requirement of volatilization means that GC-MS is suitable for fewer compounds and therefore results in more limited coverage [Theodoridis et al. 2012].

LC-MS, on the other hand, is suitable for the separation and detection of far more compounds and is therefore widely used for metabolite profiling. Although sample run times can be relatively long (usually 20-40 min) and ion suppression is a technical concern, samples do not require derivatization, which simplifies sample preparation and increases coverage capabilities. In terms of LC, there are two general resin types used in the columns, hydrophobic or “reversed-phase” (RP), and hydrophilic, or “normal phase.” These chromatographic matrices are complimentary in that they retain molecule with opposite biophysical properties. RPLC columns retain relatively hydrophobic species, with low to moderate polarity, whereas normal phase columns preferably retain charged metabolites with high polarity. Unfortunately, no one extraction/chromatography method is sufficient to study all species and samples

can be run using both types of columns for greatest metabolite coverage. RPLC has a distinct advantage in that salts are not retained on these columns, avoiding salt-mediated interferences with MS detection. Depending on the compound(s) of interest, this can diminish the sample preparation requirement for sample desalting, and even eliminate the need for a water wash that can be stressful for cells [Lu et al. 2017]. However, many of the most important intermediary metabolites are highly polar compounds (e.g., nucleotide di- and tri-triphosphates, TCA cycle intermediates, amino acids, etc.), requiring hydrophilic chromatography for adequate separation prior to MS analysis [Theodoridis et al. 2012; Gika et al. 2014; Spagou 2010]. Notably, nucleotide di- and tri-triphosphates can alternatively be detected by RPLC using an ion pairing or derivatization approach, but there are drawbacks to each [Lu et al. 2010]). For this reason, the studies described herein generally use normal phase chromatography on silica hydride resin with rapid and efficient washing for hydrophilic metabolite detection and relative quantification. This single approach provides good coverage, with chromatographic separation relative quantification of key metabolic intermediates and even of many weakly hydrophilic compounds. Lastly, organic acids and bases, the metal chelator EDTA and other mobile phase additives are often used to enhance the ionization and separation of positively- and negatively-charged metabolites [Yanes et al. 2011; Gao et al. 2005].

Targeted vs. untargeted

Mass spectrometry-based metabolomics can be done in a targeted or untargeted manner. In targeted metabolomics, abundances of structurally known

compounds (defined by a specific m/z and RT database entry) are relatively quantified. If absolute quantification is required for a more limited array of metabolites, a triple quadrupole mass spectrometer can be used in multiple-reaction-monitoring mode, providing optimal sensitivity of specific compounds/transitions that are defined by the data acquisition method - this can enable the quantification of hundreds of metabolites in biologically-complex mixtures [Lu et al. 2008; Roberts et al. 2012; Zhou et al. 2012]. Notably, this method must first be developed, and LC-MS instrument parameters optimized, using pure standards for all analytes of interest. Specific methods can be produced for optimized detection of classes of compounds or metabolite members in pathways of interest. This approach was employed herein for the lipid analyses described in Chapter 1. Commercially available, ready-to-go triple-quad LC-MS methods can be purchased for detection of certain compound classes (e.g. Biocrates AbsoluteIDQ® p180 or p400 HR).

If some or all metabolites of interest in a given experiment are undefined, an untargeted MS detection approach is warranted. Towards this end, a high-resolution mass spectrometer is employed to rapidly scan and record abundances for all detectable small molecules within a defined m/z range of interest (e.g., 50-1000 Da in studies described in Chapters 2 and 3) [Lu et al. 2017; Gika et al. 2014; Zhou et al. 2012].

Data acquired in an untargeted manner can be analyzed in a targeted manner, taking advantage of a pre-established *m/z* and RT database, or analyzed in an untargeted manner, by extracting ion counts (relative abundance) for m/z values at specific retention times [Lu et al. 2008]. Advanced software (e.g. Agilent MassHunter Profinder) is available to extract this information, with user defined parameters/filters,

in an automated manner. In the latter case, structurally-undefined feature abundances are quantified as ion counts for a given m/z at a specific RT, surveying thousands of features. Differentially-expressed features of sufficient interest can then be subsequently identified structurally based on observed MS/MS fragmentation patterns, as compared with MS/MS fragmentation of pure chemical standards or predictions.

Metabolite identification

One of the most challenging aspects of untargeted metabolomic studies can be the confident specification of structural identities for observed *features*. Indeed, features of interest that are not already included in the labs existent metabolite databases will often necessitate the challenging process of *de novo* structure elucidation. As a first pass, a molecular formula generator can be used to limit possible empirical formulae for a given feature based strictly on the accurate mass observed (where highest MS resolution provides greatest confidence in formula assignments). Considering accurate mass and isotope spacing, the feature can also be searched against a number of available databases (e.g. METLIN) for possible *hits* [Sana et al. 2008]. These searches can yield tentative identities, but often lead to an excess of structural possibilities. For manual structural elucidation, MS/MS fragmentation must be performed and compared with fragmentation patterns produced from a reference standard or a theoretical, in-silico fragmentation algorithm. Here, fragmentation spectra can be obtained for a compound over a range of defined collision energies (for collision–induced dissociation). If sufficient and *clean* fragmentation data is obtained, the fragments can be interpreted by a skilled mass

spectrometrist/chemist to obtain a putative structure [Rathahao-Paris et al. 2016]. MS/MS fragmentation databases (e.g. METLIN) are now available and can be searched to at least gain helpful insight into structure, but results may require additional confirmatory evidence. Unfortunately, fragmentation can vary based on analytical conditions and collision cell geometry, so conditions and MS instrumentation used for database creation would be most optimal for potential structure assignment. In Chapter 3, I describe the successful use of *de novo* MS/MS fragmentation coupled with METLIN MS/MS database matching to confirm the identity of betaine. However, there are often no convincing MS/MS database structure match for an unknown compound of interest [Rathahao-Paris et al. 2016]. *In silico* fragmentation spectra are also available for many compounds, although these are less reliable than those established using pure reference standards. In the end, all of this information can be used together by the mass spectrometrist to attempt to confidently elucidate the structure of an unknown compound [Dunn et al. 2013]. A major limitation of de novo structure assignment is that the compound concentration and ionization efficiency need to be sufficient to obtain an informative fragmentation spectra inasmuch as MS2 fragment signals are significantly lower than that of the original metabolite. In cases of MS/MS signal insufficiency, a strategy would be to inject more concentrated samples or greater volumes of sample for analysis.

MS/MS fragmentation runs can include hundreds of compounds for simultaneous analysis. However, to obtain the best data, fragmentation of compounds with overlapping RT should ideally be performed separately – this allows scan time to be dedicated to a single compound. Also, different compounds will likely require different parameters for optimal fragmentation and target-specific optimization of

MS/MS may be required. Auto MS/MS is now available for facile untargeted fragmentation data acquisition. Here, m/z's are selected for fragmentation when the signal reaches a specific ion abundance threshold, facilitating the fragmentation of *bona fide* molecular features rather than *noise*. Auto MS/MS is a promising new instrument capability that is anticipated to become increasingly effective with the development of improved algorithms and MS instrumentation.

Advances in the process of identifying unknown features, including enhanced instrument performance, database breadth and applicability, fragmentation prediction algorithms, interpretive software and unique methods to gain structure-related information (such as retention time prediction algorithms [Cao et al. 2015] or the use of in-source MS-only fragmentation for structural elucidation assistance [Godzien et al. 2015; Seitzer et al. 2018]), are likely to be key contributors to enhancing the speed, usability and efficacy of metabolomics in the future.

Stable isotope tracing

Stable isotope tracing coupled with mass spectrometry detection is a powerful tool for understanding active metabolism of a biological system [Jang et al. 2018]. Stable heavy isotopes are atoms that contain one or more additional neutrons, which provide an increase in mass that can be readily tracked using mass spectrometry. Stable isotopes are naturally abundant on earth with a specific frequency and do not radioactively decay, as opposed to unstable radioisotopes. Stable heavy isotopes that are useful for metabolic tracing are present in nature at significantly lower abundance than a primary ('light') isotope. Importantly, compounds can be chemically

synthesized that contain heavy labeled atoms at specific positions for metabolic tracing experiments. These species (or *isotopologues*) are biologically and biophysically identical to the light isotope versions, but are readily distinguished using MS. Tracing experiments can therefore be performed, where the temporal fate of heavy-labeled atoms are mapped. For example, heavy [^{13}C]-labeled glucose can be fed to cells in culture, followed by the harvest of metabolites and analysis of ^{13}C-incorporation into glycolytic intermediates. By comparing two groups, a perturbation in glycolysis could be confirmed by detecting differential ^{13}C-glucose trafficking into downstream metabolites. Multiple time points can be used to obtain detailed information about the metabolic flux of glucose (or other metabolites of interest) as a function of time.

As opposed to simple metabolite profiling, which provides a snapshot of abundances of compounds, isotope tracing provides key information about reaction rates and metabolite origin/fate, and is integral for probing molecular mechanisms of metabolic perturbations [Zamboni et al. 2015; Johnson et al. 2016]. In Chapter 2, we effectively utilize stable isotope tracing with both [^{13}C]-glucose and [^{13}C]-glutamate to show increased glucose and transsulfuration metabolism in skin fibroblast cells grown from a newly-recognized subgroup of patients with sporadic amyotrophic lateral sclerosis (sALS). In Chapter 3, [D_U]-choline is successfully used to show decreased choline-derived betaine in a cell line with a knockout in a functionally undefined mitochondrial transporter protein (SLC25A40). Later in Chapter 3, I take advantage of the ability of MS to distinguish heavy labeled compounds from their light counterparts to detect newly transported 2-oxoglutarate, glycine and aspartate. Many isotopically labeled compounds are commercially available with a defined isotopic enrichment,

enabling metabolic tracer studies. However, uncommon or difficult-to-produce compounds may not be available and are certain to be expensive.

Data processing, normalization, analysis, interpretation and visualization

The most challenging aspect of untargeted metabolite profiling can be the data analysis and interpretation. As mentioned above, *m/z* features at specific retentions times are extracted from the acquired MS spectral data using software that filters, extracts and aligns feature peaks across samples in a recursive manner. Users must make decisions regarding optimal parameter and filter decisions for data analysis. For statistical and ease-of-interpretation reasons, it is important to avoid extracting low-abundance noise (or impossible to interpret metabolites) from the acquired MS data. Automated integration of chromatograms (i.e., assessment of ion abundance vs. time) affords a facile relative abundance readout in terms of *ion counts*, but can be inaccurate (especially for compounds with low abundance with closely neighboring peaks or RT shifts) and confident abundance data often requires time-consuming manual curation. As discussed in more detail below, many, if not most, features do not correspond to *bona fide* cellular metabolites, but instead arise as in-source fragments or adducts of a given metabolite created as an MS ionization artefact [Xu et al. 2015; Brown et al. 2009; Lu et al. 2017]. Unfortunately, determining which *m/z* features are authentic biological metabolites, rather than in-source artefacts, can be challenging and can confound confident metabolite identification [Vinaixa et al. 2016]. Methods have been produced to try and automatically identify and group features that are related (and even to attempt to use the information to assist in structural elucidation

[Seitzer et al. 2018]). I have used a simple method based on RT similarity (in-source changes, such as fragments, adducts or multimers have the same RT) and abundance correlation between samples (in-source fragments should generally have the same relative abundance between samples) to identify and remove putative fragments. In one example, by using stringent parameters of ±0.03 min RT and sample abundance Pearson correlation >0.95 for fragment identification, I was able to algorithmically remove 29% of features as likely in-source compounds (Figure I.4). This makes the statistical analysis of remaining features more accurate (i.e., multiple measurements of the same compound are removed from consideration) and provides a clearer picture of the actual number of individual statistically significant compounds for data interpretation. Notably, removed in-source generated features are not discarded, but instead utilized to determine which peak is likely the parent metabolite, as well as to gain information for structure assignment or confirmation.

As mentioned above, samples are typically normalized prior to MS data acquisition (e.g. based on protein content) to control for variations in cell number or tissue quantity in extracts. However, a second normalization is often required to control for post-reconstitution variability in metabolite detection, such as changes in MS instrument sensitivity, LC drifts in retention time, and time-dependent sample degradation. Such normalization is generally essential for large, multi-day or multi-instrument experiments. In any case, to minimize bias in data acquisition, MS samples are routinely analyzed in a staggered order (e.g. disease, control, disease, control, etc.) so that potential batch effects are minimized.

Mass	RT	Duplicate?	Mass	RT	Duplicate?
767.541	4.14	0	612.146	9.86	0
769.55	4.23	0	307.073	9.87	1
741.524	4.25	0	450.139	9.89	0
795.568	4.26	0	462.084	9.99	0
525.301	4.28	0	719.115	10.05	0
715.511	4.3	0	432.02	10.07	1
727.543	4.31	0	390.063	10.09	0
163.049	4.31	0	427.026	10.1	0
743.54	4.33	0	97.976	10.12	1
701.529	4.36	0	175.954	10.13	1
689.493	4.36	0	79.966	10.14	1
717.525	4.38	0	123.975	10.17	0
306.048	4.39	0	506.991	10.18	0
124.988	4.51	0	412.044	10.21	0
80.964	4.51	1	314.07	10.25	0
127.01	4.51	1	155.067	10.39	0
500.054	4.51	1	211.033	10.7	0
750.082	4.51	1			
625.068	4.51	1			
875.096	4.51	1			
627.065	4.51	1			
752.079	4.51	1			
877.092	4.51	1			
251.025	4.52	1			
252.024	4.52	1			
352.106	4.68	0			
230.14	4.73	0			
216.038	4.81	0			
120.041	4.81	0			

Figure I.4. *Features produced in-source can be algorithmically identified. Excerpts from a list of features (sorted by RT) identified in an untargeted analysis are shown. Highlighted are two groups of features where multiple in-source compounds (fragments, adducts and multimers, denoted with a 1) corresponding to the same parent compound (denoted with a 0) were algorithmically labeled. A 1 was assigned to a feature with a RT of ±0.03min and had a Pearson correlation (among the 72 samples) of 0.95 or greater with any other feature. In the comparison, the feature with lower average abundance was assumed to be produced in source and assigned a 1 (the feature with higher average abundance was assumed to be the parent compound and assigned a 0). In this example, this analysis identified 219 features out of 747 as in-source (~29%).*

One normalization technique that is often employed is to normalize based on a continually re-analyzed *quality control* (QC) sample, created from a mixture of the experimental samples, which is aliquoted, frozen and repeatedly analyzed at regular intervals (i.e. every 6-10 runs). Abundances for each individual metabolite can then be normalized based on the abundance in flanking quality control sample LC-MS runs [Kamleh et al. 2012; Chen et al. 2014]. In Chapter 2, we demonstrate the use of this technique to compare samples analyzed on the mass spectrometer over the course of one month.

A common normalization method for targeted MS analyses is to normalize to a heavy isotope-labeled internal standard that is added to each sample. However, heavy labeled compounds are not always available (or can be expensive). An array of internal standards with various retention times have been used to correct for compounds across an LC-MS run [Sysi-Aho et al. 2007]. Unfortunately, compounds are so diverse that this approach is not robust for the detection of many compounds, where it can suppress biological variation, and certainly cannot be utilized for untargeted experiments [De Livera et al. 2012].

Other methods of normalization are data-driven methods, which generally normalize based on the array of metabolite abundances within each sample and can range from median fold change or quantile normalization to complex machine learning algorithm-based normalization [Ejigu et al. 2013; van den Berg et al. 2006; Shen et al. 2016; Wu et al. 2016; De Livera et al. 2015; Karpievitch et al. 2014]. For any experiment with samples being run at different times or on different machines a robust data normalization strategy must be selected during the experiment design phase to ensure the ability to effectively compare all the samples.

After the feature abundances have been obtained and normalized, the data is analyzed based on the goals of the experiment either using data science tools (e.g. R, Matlab) or metabolomics-specific applications (e.g. Agilent Mass Profiler Professional, MetaboAnalyst). These can include basic tests of statistical significance or advanced machine learning applications [Hendriks et al. 2011]. Interpretation and visualization of complex metabolomics results can be challenging. Often, pathway enrichment analyses (e.g. metabolite set enrichment analysis and pathway topology analysis) and are employed to connect individual metabolic difference and highlight perturbations in a more global way (these pathways become good candidates for stable isotope tracing experiments) [Johnson et al. 2016; Xia et al. 2010; Xia et al. 2011; Xia et al. 2016; Brown et al. 2016; Chen et al. 2015]. Multiple metabolite perturbations are generally visualized in heatmaps or pathway/metabolic maps [Sugimoto et al. 2012], as we do in Chapter 2 for transsulfuration-related compounds. In the end, an orthogonal approach to confirm metabolomics results should be employed, if possible.

New applications and methods

Mass spectrometry-based metabolomics is a rapidly growing field and new applications and methodologies are continuously being developed. Unique approaches to solving many of the challenges (e.g. data normalization or improving chromatography for greater metabolite coverage) are being explored in individual labs. Many times, the application of the newest approaches requires specialized instrumentation or in-depth understanding of advanced statistical/computational concepts. However, those new approaches that are most promising eventually gain

mainstream interest and are broadly incorporated into the research of a number of metabolomics laboratories. Three relatively recent and advanced uses of metabolomics which further highlight its powerful, developing capabilities are multi-omics, mass spectrometry imaging of small molecules and single cell metabolomics.

As the name implies, a multi-omic analysis integrates data from at least two different ‘omics (metabolomics, transcriptomics, proteomics, genomics) experiments to provide an even more insightful and encompassing understanding of the topic under study. Combining the different ‘omic datasets (e.g. making connections between metabolites and RNA) and performing a shared pathway interpretation is challenging, but tools to assist with integration and interpretation of these data are continually being developed [Chong et al. 2018]. In Chapter 2 we use metabolomics and transcriptomics in a multi-omic analysis to help characterize a subgroup of sALS patient fibroblasts. In the coming years, the continued work into the integration of different ‘omics data is likely to provide far deeper and multi-layer insight in diverse research endeavors.

Another metabolomics application that has recently become feasible due to enhanced instrumentation is mass spectrometry imaging. In MS imaging, metabolite abundances can be sequentially obtained across a sample surface (such as an unfixed tissue cryosection). Matrix-assisted laser desorption/ionization (MALDI), where a laser is used to ionize metabolites from the sample provides the most common source for metabolite-on-tissue ionization. The spatial resolution of this ionization technique and the detection sensitivity of mass spectrometers is now sufficiently high to obtain single cell resolution metabolomics data [Hansen et al. 2018; Wang et al. 2017; Kompauer et al. 2016; Balluff et al. 2018]. This can be collected across a sample

surface (e.g. a brain slice) to obtain information about the relative abundance and distribution of compounds across the tissue section. This developing field is likely to add significantly to our understanding of endogenous-location-specific metabolism in a myriad of biological samples types.

An additional cutting-edge application in metabolomics is single cell metabolomics. Again, methodological and instrumentation-related improvements have begun to make the detection of metabolites from single cells possible. Generally, single cells from a population are isolated using various techniques, such as micropipettes or microfluidic devices, and metabolites from individual cells are sequentially ionized by various methods, such as by MALDI or by directly trapping and spraying a live cell using an electrospray ionization tip [Emara et al. 2017]. This developing technology will continue the expansion of research related to the study of single cells from primarily RNA transcript level to the level of cellular small molecule metabolism.

Conclusion

In summary, mass spectrometry-based global metabolite profiling is a powerful tool that serves to both generate and test novel hypotheses in biomedical research. Advances in MS instrumentation, chromatographic separations, data acquisition and statistical and computational/bioinformatic analyses are continuing to expand the use of MS-based metabolomics as a powerful tool in the armamentarium of biomedical researchers. Judicious selection of metabolite extraction techniques, chromatography and MS detection platforms now enable the detection of thousands of

metabolite with greatly disparate structures, biophysical properties and biological functions.

Researchers must be careful, however, when designing and implementing experiments in order to ensure the collection of informative data and avoid possible pitfalls. Since there are no universally-accepted methodologies as yet, sample preparation protocols must be established in each laboratory to ensure reproducibility and to minimize common confounding factors, such as ion suppression.

Due to the extremely large number of detectable features and undefined inventory of resident biomolecules, unique issues arise. These include confident feature identification as a bottleneck and the potential to over fit data (i.e. the curse of dimensionality), requiring both diligence and expertise on the part of experimenters. However, these challenges are clearly surmountable and metabolomics analysis, as utilized in the subsequent chapters, offers a powerful tool for small molecule-centered research.

In the next three chapters I discuss a number of applications of mass spectrometry-based metabolomics. Many of the topics considered in this *Introduction* are explicitly or implicitly explored and applied. In Chapter 1, I discuss the endocannabinoid system in the context of acute withdrawal from binge alcohol drinking and use targeted RPLC-MS methodology to detect multiple N-acyl ethanolamine family endocannabinoids in mouse brain samples. In Chapter 2, I provide an example of the use of aqueous normal phase LC-MS unbiased metabolite profiling to generate and test a hypothesis regarding sALS. Specifically, we identify a previously unrecognized subgroup of sALS patient fibroblasts that are defined by a

unique metabotype. We use stable isotope tracing and a multi-omics analysis to confirm the hypothesis and further characterize the metabolic perturbations of this novel sALS subgroup. Finally, in Chapter 3, I seek to utilize MS-based targeted and untargeted small molecule detection in research focused on characterizing the function of members of the SLC25 family of mitochondrial transporter proteins.

REFERENCES

Annesley TM. Ion Suppression in Mass Spectrometry. Clin Chem. 2003;49(7):1041-1044.

Aretz I, Meierhofer D. Advantages and pitfalls of mass spectrometry based metabolome profiling in systems biology. Int J Mol Sci. 2016;17(5).

Armitage EG, Barbas C. Metabolomics in cancer biomarker discovery: Current trends and future perspectives. J Pharm Biomed Anal. 2014;87:1-11.

Balluff B, McDonnell LA. Mass Spectrometry Imaging of Metabolites. In: Clinical Metabolomics.; 2018:345-357.

Beckonert O, Keun HC, Ebbels TMD, et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc. 2007;2(11):2692-2703.

Beger RD, Dunn W, Schmidt MA, et al. Metabolomics enables precision medicine: "A White Paper, Community Perspective." Metabolomics. 2016;12(10).

Brown DG, Rao S, Weir TL, et al. Metabolomics and metabolic pathway networks from human colorectal cancers, adjacent mucosa, and stool. Cancer Metab. 2016;4(1):1-12.

Brown M, Dunn WB, Dobson P, et al. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. Analyst. 2009;134(7):1322-1332.

Bundy JG, Davey MP, Viant MR. Environmental metabolomics: A critical review and future perspectives. Metabolomics. 2009;5(1):3-21.

Cao M, Fraser K, Huege J, Featonby T, Rasmussen S, Jones C. Predicting retention time in hydrophilic interaction liquid chromatography mass spectrometry and

its use for peak annotation in metabolomics. Metabolomics. 2015;11(3):696-706.

Chen HH, Tseng YJ, Wang SY, et al. The metabolome profiling and pathway analysis in metabolic healthy and abnormal obesity. Int J Obes. 2015;39(8):1241-1248.

Chen M, Rao RSP, Zhang Y, Zhong CX, Thelen JJ. A modified data normalization method for GC-MS-based metabolomics to minimize batch variation. Springerplus. 2014;3(1):1-7.

Chong J, Soufan O, Li C, et al. MetaboAnalyst 4.0: Towards more transparent and integrative metabolomics analysis. Nucleic Acids Res. 2018;46(W1):W486-W494.

de Koning W, van Dam K. A Method for the determination of Changes of Glycolytic Metabolites in Yeast on a Subsecond Time Scale Using Extraction at Neutral pH. Anal Biochem. 1992;204:118-123.

De Livera AM, Dias DA, De Souza D, et al. Normalizing and integrating metabolomics data. Anal Chem. 2012;84(24):10768-10776.

De Livera AM, Sysi-Aho M, Jacob L, et al. Statistical Methods for Handling Unwanted Variation in Metabolomics Data. Anal Chem. 2015;87(7):3606-3615.

Dettmer K, Aronov PA, Hammock BD. Mass Spectrometry-Based Metabolomics. Mass Spectrom Rev. 2007;26:51-78.

Dettmer K, Nürnberger N, Kaspar H, Gruber MA, Almstetter MF, Oefner PJ. Metabolite extraction from adherently growing mammalian cells for metabolomics studies: Optimization of harvesting and extraction protocols. Anal Bioanal Chem. 2011;399(3):1127-1139.

Dunn WB, Erban A, Weber RJM, et al. Mass appeal: Metabolite identification in mass spectrometry-focused untargeted metabolomics. Metabolomics. 2013;9(SUPPL.1):44-66.

Ejigu BA, Valkenborg D, Baggerman G, et al. Evaluation of Normalization Methods to Pave the Way Towards Large-Scale LC-MS-Based Metabolomics Profiling Experiments. Omi A J Integr Biol. 2013;17(9):473-485.

Emara S, Amer S, Ali A, Abouleila Y, Oga A, Masujima T. Single-Cell Metabolomics. In: Sussulini A, ed. Metabolomics: From Fundamentals to Clinical Applications.; 2017:353.

Furey A, Moriarty M, Bane V, Kinsella B, Lehane M. Ion suppression; A critical review on causes, evaluation, prevention and applications. Talanta. 2013;115:104-122.

Gao S, Zhang ZP, Karnes HT. Sensitivity enhancement in liquid chromatography/atmospheric pressure ionization mass spectrometry using derivatization and mobile phase additives. J Chromatogr B Anal Technol Biomed Life Sci. 2005;825(2):98-110.

Gika HG, Theodoridis GA, Plumb RS, Wilson ID. Current practice of liquid chromatography-mass spectrometry in metabolomics and metabonomics. J Pharm Biomed Anal. 2014;87:12-25.

Godzien J, Armitage EG, Angulo S, et al. In-source fragmentation and correlation analysis as tools for metabolite identification exemplified with CE-TOF untargeted metabolomics. Electrophoresis. 2015;36(18):2188-2195.

Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306-313.

Hansen RL, Lee YJ. High-Spatial Resolution Mass Spectrometry Imaging: Toward Single Cell Metabolomics in Plant Tissues. Chem Rec. 2018;18(1):65-77.

Hendriks MMWB, Eeuwijk FA va., Jellema RH, et al. Data-processing strategies for metabolomics studies. TrAC - Trends Anal Chem. 2011;30(10):1685-1698.

Hiller J, Franco-Lara E, Weuster-Botz D. Metabolic profiling of Escherichia coli cultivations: Evaluation of extraction and metabolite analysis procedures. Biotechnol Lett. 2007;29(8):1169-1178.

Jang C, Chen L, Rabinowitz JD. Metabolomics and Isotope Tracing. Cell. 2018;173(4):822-837.

Jessome LL, Volmer DA. Ion Suppression: A Major Concern in Mass Spectrometry. LCGC North Am. 2006;24(5):498-510.

Johnson CH, Ivanisevic J, Siuzdak G. Metabolomics: Beyond biomarkers and towards mechanisms. Nat Rev Mol Cell Biol. 2016;17(7):451-459.

Kamleh MA, Ebbels TMD, Spagou K, Masson P, Want EJ. Optimizing the use of quality control samples for signal drift correction in large-scale urine metabolic profiling studies. Anal Chem. 2012;84(6):2670-2677.

Karpievitch Y V., Nikolic SB, Wilson R, Sharman JE, Edwards LM. Metabolomics data normalization with EigenMS. PLoS One. 2014;9(12):1-10.

Kompauer M, Heiles S, Spengler B. Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4-μm lateral resolution. Nat Methods. 2016;14(1):90-96.

Lu W, Bennett BD, Rabinowitz JD. Analytical strategies for LC-MS-based targeted metabolomics. J Chromatogr B Anal Technol Biomed Life Sci. 2008;871(2):236-242.

Lu W, Clasquin MF, Melamud E, Amador-Noguez D, Caudy AA, Rabinowitz JD. Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. Anal Chem. 2010;82(8):3212-3221.

Lu W, Su X, Klein MS, Lewis IA, Fiehn O, Rabinowitz JD. Metabolite Measurement: Pitfalls to Avoid and Practices to Follow. Annu Rev Biochem. 2017;86(1):277-304.

Raterink RJ, Lindenburg PW, Vreeken RJ, Ramautar R, Hankemeier T. Recent developments in sample-pretreatment techniques for mass spectrometry-based metabolomics. TrAC - Trends Anal Chem. 2014;61:157-167.

Rathahao-Paris E, Alves S, Junot C, Tabet JC. High resolution mass spectrometry for structural identification of metabolites in metabolomics. Metabolomics. 2016;12(1):1-15.

Roberts LD, Souza AL, Gerszten RE, Clish CB. Targeted metabolomics. Curr Protoc Mol Biol. 2012;1(SUPPL.98):1-24.

Römisch-Margl W, Prehn C, Bogumil R, Röhring C, Suhre K, Adamski J. Procedure for tissue sample preparation and metabolite extraction for high-throughput targeted metabolomics. Metabolomics. 2012;8(1):133-142.

Sana TR, Roark JC, Li X, Waddell K, Fischer SM. Molecular formula and METLIN personal metabolite database matching applied to the identification of compounds generated by LC/TOF-MS. J Biomol Tech. 2008;19(4):258-266.

Seitzer PM, Searle BC. Incorporating In-Source Fragment Information Improves Metabolite Identification Accuracy in Untargeted LC–MS Data Sets. J Proteome Res. 2018:acs.jproteome.8b00601.

Sellick CA, Knight D, Croxford AS, et al. Evaluation of extraction processes for intracellular metabolite profiling of mammalian cells: Matching extraction approaches to cell type and metabolite targets. Metabolomics. 2010;6(3):427-438.

Shen X, Gong X, Cai Y, et al. Normalization and integration of large-scale metabolomics data using support vector regression. Metabolomics. 2016;12(5):1-12.

Spagou K, Tsoukali H, Raikos N, Gika H, Wilson ID, Theodoridis G. Hydrophilic interaction chromatography coupled to MS for metabonomic/metabolomic studies. J Sep Sci. 2010;33(6-7):716-727.

Sugimoto M, Kawakami M, Robert M, Soga T, Tomita M. Bioinformatics Tools for Mass Spectroscopy-Based Metabolomic Data Processing and Analysis. Curr Bioinform. 2012;7(1):96-108.

Sysi-Aho M, Katajamaa M, Yetukuri L, Orešič M. Normalization method for metabolomics data using optimal selection of multiple internal standards. BMC Bioinformatics. 2007;8:1-17.

Theodoridis GA, Gika HG, Want EJ, Wilson ID. Liquid chromatography-mass spectrometry based global metabolite profiling: A review. Anal Chim Acta. 2012;711:7-16.

van den Berg R a, Hoefsloot HCJ, Westerhuis J a, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. BMC Genomics. 2006;7:142.

Vinaixa M, Schymanski EL, Neumann S, Navarro M, Salek RM, Yanes O. Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects. TrAC - Trends Anal Chem. 2016;78:23-35.

Vuckovic D. Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. Anal Bioanal Chem. 2012;403(6):1523-1548.

Vuckovic D. Sample Preparation in Global Metabolomics of Biological Fluids and Tissues. Elsevier; 2013.

Wang X, Han J, Hardie DB, Yang J, Pan J, Borchers CH. Metabolomic profiling of prostate cancer by matrix assisted laser desorption/ionization-Fourier transform ion cyclotron resonance mass spectrometry imaging using Matrix Coating Assisted by an Electric Field (MCAEF). Biochim Biophys Acta - Proteins Proteomics. 2017;1865(7):755-767.

Want EJ, Masson P, Michopoulos F, et al. Global metabolic profiling of animal and human tissues via UPLC-MS. Nat Protoc. 2013;8(1):17-32.

Wellerdiek M, Winterhoff D, Reule W, Brandner J, Oldiges M. Metabolic quenching of corynebacterium glutamicum: Efficiency of methods and impact of cold shock. Bioprocess Biosyst Eng. 2009;32(5):581-592.

Winder CL, Dunn WB, Schuler S, et al. Global metabolic profiling of Escherichia coli cultures: An evaluation of methods for quenching and extraction of intracellular metabolites. Anal Chem. 2008;80(8):2939-2948.

Wishart DS, Feunang YD, Marcu A, et al. HMDB 4.0: The human metabolome database for 2018. Nucleic Acids Res. 2018;46(D1):D608-D617.

Wittmann C, Krömer JO, Kiefer P, Binz T, Heinzle E. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. Anal Biochem. 2004;327(1):135-139.

Wu H, Southam AD, Hines A, Viant MR. High-throughput tissue extraction protocol for NMR- and MS-based metabolomics. Anal Biochem. 2008;372(2):204-212.

Wu Y, Li L. Sample normalization methods in quantitative metabolomics. J Chromatogr A. 2015;1430:80-95.

Xia J, Wishart DS. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. Nat Protoc. 2011;6(6):743-760.

Xia J, Wishart DS. Using metaboanalyst 3.0 for comprehensive metabolomics data analysis. Curr Protoc Bioinforma. 2016;2016(September):14.10.1-14.10.91.

Xia J, Wishart DS, Valencia A. MetPA: A web-based metabolomics tool for pathway analysis and visualization. Bioinformatics. 2011;27(13):2342-2344.

Xu YF, Lu W, Rabinowitz JD. Avoiding misannotation of in-source fragmentation products as cellular metabolites in liquid chromatography-mass spectrometry-based metabolomics. Anal Chem. 2015;87(4):2273-2281.

Yanes O, Tautenhahn R, Patti GJ, Siuzdak G. Expanding coverage of the metabolome for global metabolite profiling. Anal Chem. 2011;83(6):2152-2161.

Zamboni N, Saghatelian A, Patti GJ. Defining the Metabolome: Size, Flux, and Regulation. Mol Cell. 2015;58(4):699-706.

Zhou B, Xiao JF, Tuli L, Ressom HW. LC-MS-based metabolomics. Mol Biosyst. 2012;8(2):470-481.

CHAPTER ONE*

BLOCKADE OF ALCOHOL ESCALATION AND "RELAPSE" DRINKING BY PHARMACOLOGICAL FAAH INHIBITION IN MALE AND FEMALE C57BL/6J MICE

INTRODUCTION

The endocannabinoid (eCB) system consists of cannabinoid receptors (CB1 and CB2) and endogenous cannabinoids (including anandamide [AEA] and 2-arachidonoyl glycerol). Pharmacological blockade of CB1 receptors reduces alcohol drinking in mice [Arnone et al. 1997], decreases alcohol intake and the motivation to consume alcohol [Colombo et al. 1998; Gallate and McGregor 1999] and blocks alcohol seeking in rats [McGregor et al. 2005], suggesting that the eCB/CB1 system is involved in the positive reinforcing properties and consumption of alcohol. In addition, mice lacking CB1 have reduced alcohol drinking or preference [Hungund et al. 2003; Wang et al. 2003; Naassila et al. 2004], and alcohol-induced reward [Houchi et al. 2005]. Together, the increased eCB/CB1 activity is likely to promote alcohol intake during early stages of alcohol drinking by potentiating acute alcohol reward [Manzanares et al. 1999]. In apparent contrast, eCB/CB1 signaling deficiency occurs after chronic alcohol exposure and protracted withdrawal, which may promote alcohol

intake through a negative reinforcement mechanism. Several findings support this notion: [1] down-regulation of CB1 expression and function is observed during protracted alcohol withdrawal [Mitrirattanakul et al. 2007; Varodayan et al. 2016]; [2] in humans, *in vivo* imaging studies find decreased CB1 availability in heavy-drinking alcoholics that persists into abstinence [Hirvonen et al. 2013; Ceccarini et al. 2014]; and [3] alcohol dependent patients have lowered AEA levels in plasma during recent abstinence [Mangieri et al, 2009].

The AEA-dependent signaling has been shown to be regulated by the enzyme involved in its catabolism, fatty acid amide hydrolase (FAAH) [Cravatt et al. 1996, 2001]. Several lines of evidence have demonstrated that AEA is involved in the behavioral effects of alcohol: inhibition or genetic deletion of FAAH, with resultant increases in AEA levels [Cravatt et al. 2001], increases alcohol consumption and preference [Basavarajappa et al. 2006; Blednov et al. 2007]. The FAAH C385A polymorphism (with increased eCB activity due to impaired FAAH function) has been reported to be associated with increased alcohol abuse and dependency [e.g., Sipe et al. 2002; Sloan et al. 2017], and consistently, we found that the knock-in mice with the human FAAH C385A polymorphism had increased alcohol consumption [Zhou et al. 2016], with decreased anxiety-like behavior [Dincheva et al. 2015].

Withdrawal from alcohol is associated with increased stress responsivity and persistent negative affective symptoms, such as anxiety and depression, the severity of which are closely associated with alcohol relapse susceptibility [Koob & Kreek 2007]. eCBs exert prominent modulatory influence on the extended amygdala and corticostriatal circuits, and stress exposure disrupts eCB-enriched regions that participate in emotional control [Serrano et al. 2012; Dincheva et al. 2015; Morena et

al. 2016]. Specifically, the relatively deficient eCB function is associated with increased anxiety and depression, and as such, impaired eCB activity may contribute to the negative affective states and increased stress responsivity that underlie negative reinforcement mechanisms driving alcohol drinking by dependent individuals and that contribute to alcohol relapse following periods of abstinence [Parsons & Hurd, 2015]. Pharmacologic and genetic manipulations (knockout or knock-in) of FAAH have been implicated in anxiolytic and anti-depressive behaviors [Bortolato et al. 2007; Gunduz-Cinar et al. 2013; Kathuria et al. 2003; Moreira et al. 2008; Carnevali et al. 2015]. However, there is no study on FAAH inhibitors in rodent models with alcohol withdrawal from high intake, though FAAH inhibition decreases anxiety-like behaviors that are present during alcohol withdrawal [Cippitelli et al. 2008].

Our hypothesis is that enhanced eCB signaling by FAAH inhibition opposes the effect of alcohol withdrawal driving alcohol escalation and "relapse" drinking. In this study, therefore, we first investigated whether a selective FAAH inhibitor URB597 alters voluntary alcohol drinking in mice during acute (1 day) withdrawal from chronic (3-week) intermittent access (IA) alcohol drinking, to explore its potential for development as a therapeutic agent for alcoholism. In the IA model, the mice had access to alcohol drinking for 3 weeks in a two-bottle choice paradigm with alcohol drinking every other day, and the mice developed high alcohol consumption (15-25 g/kg/day). Thus, chronic IA mice constitute an appropriate animal model for studying voluntary alcohol drinking, as they display rapid escalation of high alcohol intakes [Hwa et al. 2011] and develop the alcohol deprivation effect (ADE, modeling the relapse episodes that occur in human alcoholics) [Zhou et al. 2017]. Consequently, we determined the pharmacological effects of URB597 in an ADE model, which

allows for access to alcohol after 1 week of withdrawal. For comparison, we compared the effect of URB597 in the drinking-in-the-dark (DID) model, which allows for limited access (4 h/day) with low intake (<5-6 g/kg/day) [Rhodes et al. 2005], and models "binge" drinking to the point of intoxication. Finally, we assessed the eCB system at a biochemical level in the IA model, by measuring relative abundances of AEA, other N-acyl ethanolamide (NAE) lipids and of FAAH protein after chronic IA and acute (1 day) or long-term (1 and 2 weeks) withdrawal, providing a dynamic profile of the eCB system in response to chronic alcohol IA consumption and subsequent withdrawal.

MATERIALS AND METHODS

1. Animals. Male adult C57BL/6J (B6) mice (8 weeks of age) were obtained from The Jackson Laboratory (Bar Harbor, ME, USA) and housed in a temperature-controlled room (21 ºC). Mice were placed on a 12-hour reverse light-dark cycle (lights off at 7:00 am) upon arrival, and acclimated for a week prior to testing. During the first day of testing (i.e., session 1), all mice were 9 weeks of age. Mice were individually housed in ventilated cages fitted with steel lids and filter tops and given *ad libitum* access to food and water. Animal care and experimental procedures were conducted according to the *Guide for the Care and Use of Laboratory Animals* (Institute of Laboratory Animal Resources Commission on Life Sciences 1996), and were approved by the Institutional Animal Care and Use Committee of the Rockefeller University.

2. Materials. FAAH inhibitor URB597 (Cyclohexylcarbamic acid 3'-[aminocarbonyl]-[1,1'-biphenyl]-3-yl ester) and selective CB1 antagonist AM251 [Tocris Inc., Bristol, UK] were dissolved in 1% DMSO plus physiological saline and in 5% Cremophor in physiological saline, respectively. Alcohol solutions (7.5%, 15% and 30% v/v) were prepared from 190 proof absolute ethyl alcohol (Pharmco-AAPER, Brookfield, CT, USA) and dissolved in tap water. Sucrose and saccharin [Sigma-Aldrich Inc., St. Louis, MO, USA) were diluted in tap water. Optima liquid chromatography-mass spectrometry (LC/MS) grade acetonitrile [Fisher Scientific, Pittsburgh, PA, USA] was used for metabolite extractions and liquid chromatography. Deuterated anandamide standard was purchased from Cayman Chemical [Ann Arbor, Michigan, USA].

3. Procedures.

3.1 Chronic Intermittent Access (IA). This model of alcohol drinking in C57BL/6J mice has been widely used for several years by many laboratories [e.g., Hwa et al. 2011; Zhou et al. 2017] (**Table 1.1**).

3.1.1. The 3-week IA model. Mice had access to alcohol drinking in their home cages for 3 weeks. Food and water were available at all times in this two-bottle choice paradigm with chronic alcohol exposure every other day. This IA model was similar to earlier protocols [e.g., Hwa et al. 2011], with our modifications [Zhou et al. 2017]: At the time when the mice started individual housing (1 week before the experiments), the original water bottles were exchanged for those with sipper tubes to acclimate the mice to the sipper tubes (without ball bearings). Starting at 3 hours after lights off (10:00 am), both the water and alcohol (7.5%, 15% or 30%) solution sipper tubes were placed on their home cages.

Table 1.1. *The 3-week chronic intermittent access (IA) alcohol drinking model (15% alcohol vs. water, 24 hours) every other day with selective FAAH inhibitor URB597, CB1 antagonist AM251 or their combination.*

Week 1	Week 2	Week 3	1-day withdrawal (day 22)
IA 24-h alcohol vs water	IA 24-h alcohol vs water	IA 24-h alcohol vs water	**Single URB597 with AM251** 24-h alcohol vs water

These sipper tubes were 10-ml pipettes that were cut at both ends, sealed with a rubber stopper, and fitted with a stainless steel straight sipper tubing. The sipper tubing contained a ball bearing at the end to prevent alcohol leakage. The left/right position of the tubes was randomly set to avoid the development of side preference. The alcohol pipettes were refilled with fresh alcohol solution, and left for 24 hours before being replaced with the water tubes. Alcohol solutions were prepared fresh every 48 hours by mixing alcohol with tap water to reach the appropriate (v/v) alcohol concentration. Body weights were recorded every 2 days. Alcohol and water intake values were recorded after 4, 8 and 24 hours of alcohol access in the drinking days (to the nearest 0.1 ml). These data were used to calculate consumed alcohol intake (i.e., g/kg) and relative preference for alcohol (i.e., alcohol intake/total fluid intake).

Mice exposed to the 2-bottle "alcohol *vs.* water" choice regimen every other day rapidly acquired alcohol drinking behaviour, and subsequently daily alcohol intake rose to levels averaging approximately 15-20 g/kg/day after 3 weeks with 15% alcohol, with a high preference ratio above 0.8-0.9. This 3-week voluntary IA alcohol exposure was found to give rise to blood alcohol concentrations (BAC) around 0.50 - 0.63 mg/ml, which were within the range required to produce specific psycho-pharmacological effects [Hwa et al. 2011; Zhou et al. 2017].

The mice randomized as the vehicle-treated and drug-treated groups had matched body weight and similar alcohol intake 24 hours before the test day. The drug dissolved in vehicle was administered by an experimenter, blinded to the treatments given to the experimental groups.

3.1.2. Single, acute administration in the 3-week IA model. After 1 day of withdrawal, alcohol (7.5%, 15% or 30% concentrations) was presented 30 min after a single injection of URB597 (0.125, 0.25, 0.5, or 1 mg/kg in 1% DMSO, i.p.) or vehicle, and then alcohol and water intake values were recorded (**Table 1.1**). The URB597 doses chosen (0.125-1 mg/kg) were based on the literature [Blednov et al. 2007; Manwell et al. 2009].

In this experiment, we also tested the effect of URB597 at 0.5 mg/kg after 1 day of withdrawal in female mice exposed to 3 weeks of IA. We did not observe any sex differences in the effects of URB597 after either single or repeated administration in the following experiments, suggesting that the estrous cycle and associated hormones might not be important factors in the response to these treatments in females.

Lastly, we used selective CB1R antagonist AM251 to pharmacologically block the CB1R to confirm that the URB597 effects were mediated via CB1R. Mice were pretreated with AM251 (0.1 mg/kg) in 5% Cremophor plus saline (i.p.) 30 min before one URB597 (0.5 mg/kg) or vehicle (1% DMSO) injection 30 min before the drinking test. The AM251 dose chosen (0.1 mg/kg) was based on our recent study [Zhou et al. 2016].

3.2 Chronic (3-week) drinking-in-the-dark (DID) (**Table 1.2**). This model of alcohol drinking in C57BL/6J mice has been widely in use by many laboratories [e.g., Rhodes et al. 2005]. In this one-bottle paradigm, alcohol exposure was every day, and alcohol drinking was measured during 4 hours of alcohol access.

Table 1.2. *The 3-week chronic alcohol drinking-in-the-dark (DID) model (15% alcohol, 4 hours) every day with selective FAAH inhibitor URB597.*

Week 1	Week 2	Week 3	1-day withdrawal (day 22)
4-h alcohol	4-h alcohol	4-h alcohol	**Single URB597** 4-h alcohol

This chronic DID model was similar to earlier protocols [Rhodes et al. 2005], with our modifications [Zhou et al. 2017]; The procedures were identical to the above IA with the following exceptions: Starting at 3 hours after lights off (10:00 am), the water bottle was replaced with one 10-ml alcohol (15%) pipette, and left for 4 hours before being replaced with the water bottle. Alcohol intake values were recorded after 4 hours of alcohol access every day (to the nearest 0.1 ml). These data were used to calculate alcohol intake (i.e., g／kg). After 3 weeks of DID, the mice assigned to the vehicle-treated and URB597-treated groups with matched body weight had similar 15% alcohol intake 1 day before the test day. The URB597 doses (0.25 and 0.5 mg/kg) were chosen based on the results of the above IA alcohol study. On the test day (20 hours of withdrawal), alcohol (15% concentration) was presented 30 min after a single injection of URB597 or vehicle. Control groups: mice experienced DID for 3 weeks, and received one vehicle injection (1% DMSO, i.p.) before the 4-h drinking test. Test groups: mice experienced DID for 3 weeks, and received one URB597 injection (0.25 or 0.5 mg/kg in 1% DMSO, i.p.) before the test.

3.3. Sucrose (caloric reinforcer) and saccharin (non-caloric reinforcer) drinking. With 0.5 mg/kg URB597 (the most effective dose tested for reducing alcohol intake), we further tested the specificity of the action of URB597 on alcohol intake using sucrose or saccharin drinking behavior after single, acute administration in alcohol naïve mice or following the IA. The URB597 dose was chosen also based on the literature for sucrose drinking tests (0.1-1 mg/kg) [Bortolato et al. 2007; Rademacher & Hillard, 2007]. The sucrose or saccharin preference test in mice is sensitive to the function of brain rewarding systems and is widely used to measure the expression of anhedonia after chronic stress [Lim et al. 2012] and during alcohol

abstinence [Pang et al. 2013]. In the following experiments, 15% alcohol IA exposure was identical to those in the above experiment as described in section 3.1.1. After 3 weeks of IA, the alcohol tube was switched to sucrose for 3 sessions with stable intakes. The mice assigned to the vehicle-treated or URB597-treated groups had similar sucrose intake 24 hours before the test day. On the test day (after 8 days of alcohol withdrawal), sucrose (2% or 4%) and water intake values were recorded after 4, 8 and 24 hours of sucrose access. In parallel separate experiments, saccharin drinking (0.1% or 0.2%) was tested after 3 weeks of IA with an identical procedure.

3.3.1. Single, acute administration of URB597 on sucrose or saccharin drinking after 3-week IA. A single i.p. injection of URB597 (0.5 mg/kg) in 1% DMSO or vehicle was given 30 min before the sucrose and saccharin were presented. Mice were assigned to one of two treatment groups: vehicle as control and URB597.

3.3.2. Single, acute administration of URB597 on saccharin drinking in alcohol-naïve mice. The procedures were identical to the above, except the mice were exposed to saccharin or sucrose only (after 3 training sessions).

3.4. The alcohol deprivation effect (ADE) after 1-week withdrawal from 3-week IA (**Table 1.3**). Similar to the above, mice were given access to 15% alcohol and water for 24 h, and then received 24-h access to alcohol on alternating days for 3 weeks.

In the baseline session, 15% alcohol and water intake values were recorded at 4, 8 and 24 hours after 3 weeks of IA. The alcohol bottles were then removed from the cages leaving the mice with access to food and water *ad libitum*.

Table 1.3. *The alcohol deprivation effect (ADE) model after 1-week withdrawal from a 3-week chronic intermittent access (IA) drinking (15% alcohol vs. water) with selective FAAH inhibitor URB597.*

Week 1	Week 2	Week 3	Week 4	Day 30
IA 24-h alcohol vs water	IA 24-h alcohol vs water	IA 24-h alcohol vs water Baseline (BL)	1-week withdrawal **5 repeated URB597**	**Single URB597** ADE

After 7 days of withdrawal, the alcohol (30%) bottles were presented to the mice 3 hours after the dark cycle and the alcohol and water intake values were recorded at 4, 8 and 24 hours in the ADE session. In the following two experiments, we also tested the effect of URB597 at 0.5 mg/kg in female mice with the ADE model.

3.4.1. Single, acute administration in the ADE model. The mice assigned to the vehicle-treated or URB597-treated groups had similar alcohol intake in the baseline session. Control groups: mice received one vehicle injection (1% DMSO, i.p.) before the ADE test after 1-week withdrawal; and Test groups: mice received one URB597 injection (0.5 mg/kg in 1% DMSO, i.p.) before the ADE test after 1-week withdrawal. Then, alcohol was presented 30 min after URB597 or vehicle injection, and then alcohol and water intake values were recorded after 4, 8 and 24 hours of alcohol access. The URB597 dose chosen (0.5 mg/kg) was based on the above IA experiments.

3.4.2. Repeated administration in the ADE model. In the following experiments, alcohol exposure procedure was identical to the above experiment in section 3.4.1, with the following exceptions: mice received 5 consecutive administrations of vehicle or URB597 (0.5 mg/kg) during the 1-week withdrawal. On the ADE test day, alcohol (15%) was presented 1 day after the last URB597 injection, and then alcohol and water intake values were recorded after 4, 8 and 24 hours of alcohol access.

3.5. N-Acyl Ethanolamide (NAE) extraction for mass spectrometry (MS) quantification.

3.5.1. Brain tissue collection and extraction. Alcohol exposed and control (water) male mice from the 3-week chronic IA model were sacrificed by decapitation

in the following groups: No alcohol withdrawal (immediately following the last day of alcohol exposure), 1-day withdrawal, 1-week withdrawal, and 2-week withdrawal. Brain regions (nucleus accumbens [NAc], prefrontal cortex [PFC], basolateral amygdala [BLA] and cerebellum) were dissected and frozen on dry ice within 1 minute of sacrifice. Brain regions from two mice were combined for each sample to ensure sufficient sample size.

NAEs were extracted from brain tissue using a protocol adapted from a previously described method [Patel et al. 2003]: Stainless steel bead-beating beads were added to the Eppendorf tubes containing brain tissue. 400µL of acetonitrile (ACN) was added to the samples. 1µL of 1ng/µL deuterated anandamide ([$^{2}H_{8}$]AEA) was added as an internal standard (IS). Bead beating homogenization was performed using a TissueLyser II bead beater (Qiagen) for 1 minute at 30/s frequency. Samples were maintained on dry ice until homogenization, after which they were kept on ice. Homogenates were centrifuged in a table top centrifuge cooled to 4°C at maximum speed for 10 min to pellet brain tissue. The supernatant was removed to a new Eppendorf and brain tissue was homogenized twice more in 400µL ACN to ensure complete extraction of NAEs. Supernatants were combined and stored overnight at -20°C to precipitate particulates. Remaining brain tissue protein pellets were stored at -80°C until protein quantification. Brain extracts were centrifuged at max speed for 10 min and supernatants were removed to a new Eppendorf. Extracts were dried under vacuum in a vacufuge, reconstituted in 500 µL ACN to remove metabolites from the side of the tube and dried again. Dried extracts were stored at -80°C until reconstitution for mass spectrometry.

3.5.2. Protein quantification. Remaining protein pellets from brain tissue homogenization were dried under vacuum in a vacufuge and dissolved in 0.2N NaOH at 95°C and protein content was determined using standard Bio-Rad Bradford protein quantification protocol.

3.5.3. Liquid chromatography-mass spectrometry (LC-MS) NAE quantification. Dried brain extracts were reconstituted to the equivalent of 1µL 70% ACN in water per 100µg protein to correct for differences in brain tissue size. The relative abundances of NAEs were determined by liquid chromatography-electrospray ionization mass spectrometry using a 1290 Infinity LC system (Agilent Technologies) coupled to a 6460 Triple Quadrupole Mass Spectrometer (Agilent Technologies). Metabolites were separated on a ZORBAX SB-Aq Rapid Resolution reversed-phase column (2.1 X 100mm, 1.8 micron) using mobile phase A (deionized water, 1% ACN, 1mM ammonium formate, 0.1% formic acid) and mobile phase B (ACN, 1% deionized water, 1mM ammonium formate, 0.1% formic acid). Samples were eluted using the following gradient (**Table 1.4)**:

NAEs were detected in positive ion mode. Standard multiple reaction monitoring methodology was employed for detection of anandamide (*m/z* 348.3 → 62.1; retention time = 4.8 min), deuterated anandamide (*m/z* 356.3 → 62.1; retention time = 4.8 min), palmitoyl ethanolamide (*m/z* 300.3 → 62.1; retention time = 5.5 min) and oleoyl ethanolamide (*m/z* 326.3 → 62.1; retention time = 6.2 min).

3.5.4. MS data analysis. MS data was extracted and integrated using Agilent Technologies Masshunter Quantitative Analysis software.

Table 1.4. *LC gradient used to separate NAEs. A: deionized water, 1% ACN, 1mM ammonium formate, 0.1% formic acid, B: ACN, 1% deionized water, 1mM ammonium formate, 0.1% formic acid.*

Time (min)	A [%]	B [%]	Flow [mL/min]
0	65	35	0.4
0.5	65	35	0.4
0.6	45	55	0.4
10	45	55	0.4
10.1	0	100	1
13.1	0	100	1
13.2	65	35	0.4
16.2	65	35	0.4

Raw ion count abundances were then normalized as follows: internal standard ([$^2H_8$]AEA) abundances were normalized by multiplying by $\frac{sample\ protein\ abundance}{average\ protein\ abundance}$ to correct for different reconstitution volumes. These normalized IS abundances were then used to normalize the raw abundances by multiplying by $\frac{average\ [^2H_8]\text{AEA abundance}}{sample\ [^2H_8]\text{AEA abundance}}$ to correct for handling error. [$^2H_8$]AEA was used as the IS for all NAEs with the assumption that physical properties of the compounds are similar.

3.6. Western Blot. Alcohol exposed and control (water) male mice from the 3-week chronic IA model were sacrificed by decapitation in the following groups: water control, 1-day withdrawal, 1-week withdrawal, and 2-week withdrawal. Prefrontal cortex was dissected and frozen on dry ice within 1 minute of sacrifice, and the tissue from two mice were combined for each sample to ensure sufficient sample size. Tissue lysates were prepared in lysis buffer containing 1% NP-40, 1% Triton, 10% Glycerol, 1% protease and phosphatase inhibitors (Calbiochem, San Diego, USA) in 1x buffered saline on ice. Samples were triturated and spun down at 12,000 rpm for 10 min at 4°C. Supernatants were collected and protein concentration measured by Bradford (Bio-Rad, Hercules, USA). Proteins were resolved by SDS–PAGE electrophoresis using ready NuPAGE 10% Bis-Tris gel (Invitrogen, Carlsbad, USA) and probed as previously described (Dincheva et al. 2015) using anti-FAAH (1:500). Secondary HRP-conjugated antibodies were used at 1:3000 (Santa Cruz, Dallas, USA).

4. Data analysis. We performed power analyses to determine the number of mice required to provide statistically significant results, based on the levels of differences seen previously [Zhou et al. 2016, 2017], and predicted that these studies require 6-8 animals per group. In the experiments with URB597, alcohol intake, water intake,

total fluid and preference ratio differences between the different groups were analyzed using two-way ANOVA with repeat measures for treatment (vehicle vs drug) and for time (0-4h, 5-8h, vs 9-24h interval) in each sex (**Figure 1.1**). For dose response analysis on URB597, group differences for alcohol intake and preference ratios at the 4-hour recording time were analyzed using one-way ANOVA for treatments with different doses (**Figure 1.2**). In the experiments with URB597 with 7.5 and 30% alcohol, alcohol intake and preference ratio differences across the different groups were analyzed using two-way ANOVA for treatment (vehicle vs drug) and for concentration (7.5% vs 30%) (**Table 1.5**). In the experiments with URB597 combination with AM251, alcohol intake and preference ratio differences across the different groups were analyzed using one-way ANOVA for different treatments (**Table 1.6**).

In the experiments with URB597 with 0.2 and 0.4% saccharin, saccharin intake and preference ratio differences across the different groups were analyzed using two-way ANOVA for treatment (vehicle vs drug) and for concentration (0.2% vs 0.4%) (**Figure 1.3**). In the ADE experiments, alcohol intake differences across the different groups were analyzed using two-way ANOVA for treatment (vehicle vs drug) and for session (baseline vs ADE) in each sex, with testing our *a priori* hypothesis that there was an ADE based on the published findings [e.g., Vengeliene et al. 2014; Zhou et al. 2017] (**Figures 1.4, 1.5**). For NAE quantification, abundance differences between water and alcohol groups were analyzed using a two-tailed Student's t-test followed by Bonferroni *post-hoc* tests (**Figure 1.6**) or one-way ANOVA with different time-points (**Figure 1.7**). All the ANOVA analysis was followed by Newman-Keuls *post-hoc* tests.

Table 1.5. *Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 7.5% or 30% alcohol intake and preference ratio at 4 hours in male mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking. * p<0.05 or ** p<0.01 vs. vehicle control at the same alcohol concentration, and ++ p<0.01 vs. vehicle group at 7.5% alcohol.*

	7.5% Alcohol		30% alcohol	
	Vehicle (n = 6)	0.5 mg/kg URB (n = 6)	Vehicle (n = 6)	0.5 mg/kg URB (n = 6)
Intake, g/kg	3.6± 0.45	2.7 ± 0.40	11 ± 0.24 ++	4.7 ± 0.81 **
Preference ratio	0.75 ± 0.05	0.76 ± 0.08	0.70 ± 0.08	0.40 ± 0.06 **

Table 1.6. *Pretreatment with selective CB1R antagonist AM251 (0.1 mg/kg) blocks the effect of single, acute URB597 (URB, 0.5 mg/kg) on reducing 15% alcohol drinking in male mice after 4 hours of alcohol access. * p<0.05 vs. vehicle + vehicle control; + p<0.05 or ++ p<0.01 vs. vehicle + URB.*

	Vehicle + vehicle (n = 8)	Vehicle + URB (n = 6)	AM251 + vehicle (n = 7)	AM251 + URB (n = 6)
Intake, g/kg	5.6 ± 0.65	3.6 ± 0.41 *	5.1 ± 0.25	4.8 ± 0.41 +
Preference	0.81 ± 0.07	0.55 ± 0.04 *	0.82 ± 0.05	0.74 ± 0.04 ++

The accepted level of significance for all tests was $p < 0.05$. All statistical analyses were performed using *Statistica* (version 5.5, StatSoft Inc, Tulsa, OK).

RESULTS

1. Single, acute administration of URB597 reduced alcohol (but not sucrose) intake and preference after 1 day of withdrawal from chronic IA.

1.1. Single, acute URB597 at 0.5 mg/kg reduced 15% alcohol intake and preference in male and female mice. In male mice, for intake, two-way ANOVA with repeat measure revealed a significant effect of URB597 treatment [$F(1, 13) = 4.6$, $p<0.05$], and *post hoc* analysis showed that the URB597-treated males had less intake than the vehicle-treated ones at 4 hours [$p<0.01$] (**Figure 1.1A**). In comparison to the vehicle controls, 0.5 mg/kg URB597 reduced mean alcohol intake by 40-60% after 4 hours, associated with a compensatory increase in water intake (resulting in virtually unchanged total fluid intake [**Table S1.1**]). For preference ratio at this dose, two-way ANOVA with repeat measure revealed a significant interaction between URB597 treatment and time interval [$F(1, 26) = 6.2$, $p<0.01$], and *post hoc* analysis showed that the URB597-treated males had less preference than the vehicle-treated ones at 4 hours [$p<0.05$] (**Figure 1.1B**).

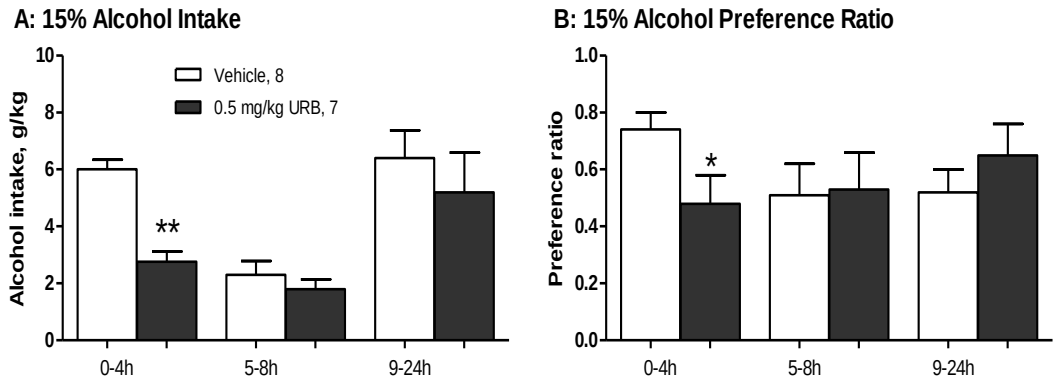
In this experiment, we also tested the effect of URB597 at 0.5 mg/kg in female mice after 3 weeks of IA, and URB597 had a similar reducing effect on alcohol intake and preference. As no significant sex difference in the effects of URB597 was found (**Table S1.2**), the female data are presented separately (**Figures 1.1C, 1.1D**). For intake, two-way ANOVA with repeat measure revealed a significant effect of

URB597 treatment [F (1, 14) = 4.8, $p<0.05$], and *post hoc* analysis showed that the URB597-treated females had less intake than the vehicle-treated ones at 4 hours [$p<0.05$] (**Figure 1.1C**). For preference ratio, two-way ANOVA with repeat measure revealed a significant effect of URB597 treatment [F (1, 14) = 8.6, $p<0.05$], and the URB597-treated females had less preference than the vehicle-treated ones at both 0-4 and 5-8 hour intervals [$p<0.01$ for both, *post hoc* analysis] (**Figure 1.1D**).

1.2. Dose-response of URB597 in male mice. At the 4-hour time point, the full-dose response of single, acute URB597 administration (0, 0.125, 0.25, 0.5 and 1 mg/kg) in terms of 15% alcohol intake and preference is presented in **Figure 1.2**. For intake (**Figure 1.2A**), one-way ANOVA revealed a significant effect of URB597 [F (4, 28) = 9.6, $p<0.001$], and *post hoc* analysis showed that: (1) in comparison with the vehicle group, the URB597-treated males had less intake than the vehicle-treated ones at 0.25, 0.5 and 1.0 mg/kg doses [$p<0.05$, $p<0.01$ and $p<0.05$, respectively]; and (2) the reductions at 0.5 mg/kg was greater than the one at 0.125 mg/kg [$p<0.05$]. For preference ratio (**Figure 1.2B**), one-way ANOVA revealed a significant effect of URB597 [F (4, 28) = 5.4, $p<0.005$], and *post hoc* analysis showed that (1) in comparison with the vehicle group, the URB597-treated males had less preference than the vehicle-treated ones at both 0.5 and 1.0 mg/kg doses [$p<0.01$ and $p<0.05$, respectively]; and (2) the reduction at 0.5 mg/kg was greater than the one at 0.125 mg/kg and 0.25 mg/kg [$p<0.05$ for both].

1.3. Single, acute URB597 (0.5 mg/kg) reduced 30% (but not 7.5%) alcohol intake and preference in male mice.

A and B, Male



C and D, Female

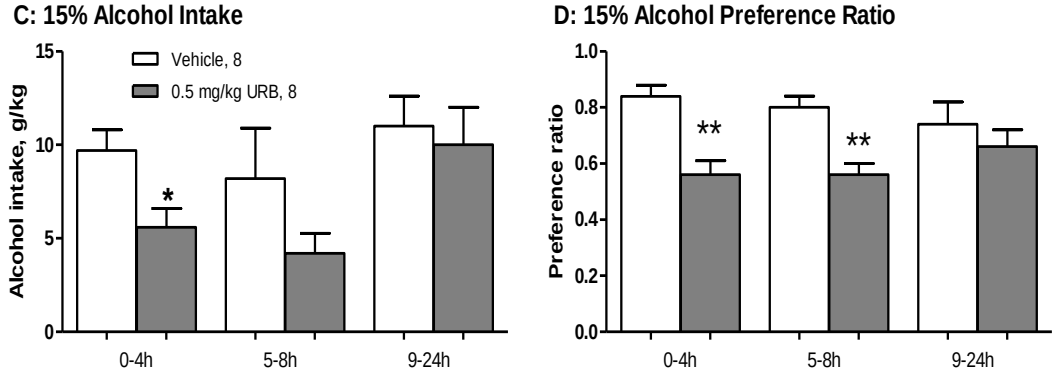


Figure 1.1. *Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 15% alcohol intake and preference ratio in male (A and B, n=7-8) and female (C and D, n=8) mice after 1 day of withdrawal from 3-week chronic intermittent access alcohol drinking. Alcohol and water intake values were recorded after 4, 8 and 24 hours of alcohol access. * $p<0.05$ and ** $p<0.01$ vs. vehicle control at the same time point.*

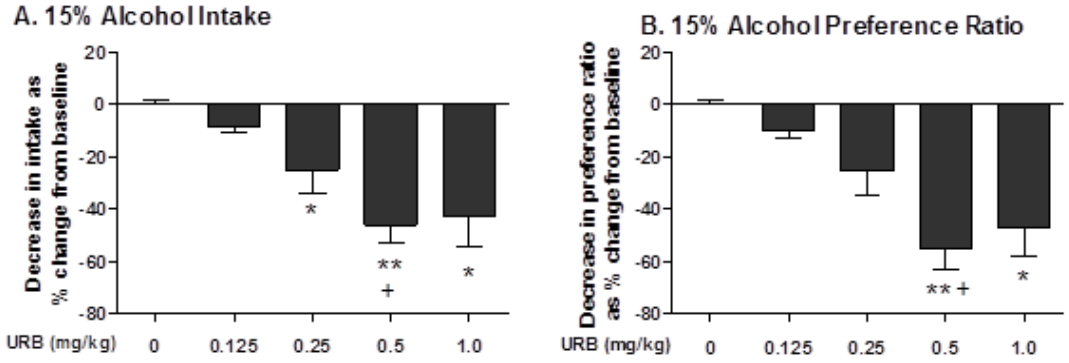


Figure 1.2. *Dose responses of single, acute administration of URB597 (URB, 0, 0.125-1 mg/kg) on reducing 15% alcohol intake (A) and preference ratio (B) in male mice (n=6-7) after 1 day of withdrawal from 3-week chronic intermittent access alcohol drinking. Data were collected at the 4-hour time point on the testing day and are expressed as a percentage of alcohol intake in vehicle-treated mice. * $p<0.05$ or ** $p<0.01$ vs. control (URB597 at 0 mg/kg), and + $p<0.05$ vs. 0.125 or 0.25 dose.*

For alcohol intake after 4 hours (**Table 1.5**), two-way ANOVA revealed a significant effect of URB597 treatment [$F(1, 20) = 7.5$, $p<0.05$], alcohol concentration [$F(1, 20) = 14.6$, $p<0.005$], and a significant interaction between URB597 and alcohol concentration [$F(1, 20) = 4.7$, $p<0.05$]. *Post hoc* analysis showed that: (1) the URB597-treated males had less intake than the vehicle-treated ones at 30% concentration [$p<0.01$]; and (2) without URB597 treatment, mice had more intake at 30% concentration than at 7.5% [$p<0.01$]. For preference ratio, two-way ANOVA revealed a significant effect of URB597 treatment [$F(1, 20) = 4.9$, $p<0.05$], concentration [$F(1, 20) = 8.2$, $p<0.01$], and a significant interaction between URB597 and concentration [$F(1, 20) = 5.5$, $p<0.05$]. *Post hoc* analysis showed that the URB597-treated mice had less preference than the control ones at 30% concentration only [$p<0.01$].

1.4. Pretreatment with selective CB1R antagonist AM251 (0.1 mg/kg) blocked the effect of acute URB597 (0.5 mg/kg) on reducing 15% alcohol drinking in male mice. One-way ANOVA revealed a significant effect of URB597 treatment [$F(3, 23) = 5.1$, $p<0.01$], and *post hoc* analysis showed that the URB597-treated males had less intake than the vehicle-treated ones [$p<0.05$]. Although pretreatment with AM251 at 0.1 mg/kg alone had no effect on alcohol drinking per se, it blunted the effect of URB597 at 0.5 mg/kg on alcohol drinking during 4 hours [$p<0.05$] (**Table 1.6**). For preference ratio, one-way ANOVA revealed a significant effect of URB597 treatment [$F(3, 23) = 5.2$, $p<0.01$], and *post hoc* analysis showed that (1) the URB597-treated males had less preference than the vehicle-treated ones [$p<0.05$]; and (2) although pretreatment with AM251 at 0.1 mg/kg alone had no effect on preference ratio per se, it blunted the reducing effect of URB597 [$p<0.01$].

Table 1.7. *No effect of single, acute administration of URB 597 (URB, 0.5mg/kg) on 2% (A) and 4% (B) sucrose intake and their preference ratio in male mice after 4 hours of sucrose access.*

A

2% sucrose	Vehicle (n = 6)	0.5 mg/kg URB (n = 6)
Intake, g/kg	1.69 ± 0.14	2.10 ± 0.26
Preference	0.89 ± 0.04	0.94 ± 0.02

B

4% sucrose	Vehicle (n = 6)	0.5 mg/kg URB (n = 6)
Intake, g/kg	7.10 ± 0.98	7.70 ± 0.70
Preference	0.94 ± 0.02	0.93 ± 0.02

1.5. No effects of single, acute URB597 (0.5 mg/kg) on sucrose intake and preference in male mice. As alcohol is a caloric reinforcer, the specificity of the action of URB597 on alcohol intake was tested after single, acute administration of 0.5 mg/kg URB597 on sucrose (caloric reinforcer) drinking in mice after chronic IA. In these experiments, the chronic 15% alcohol exposure procedures were identical to those in the above experiments. After 3 weeks of chronic IA, the alcohol tube was switched to sucrose for 3 sessions over 6 days, with stable intake observed. The male mice assigned to the vehicle- or URB597-treated groups had similar sucrose intake 24 hours before the test day. URB597 had no significant effect on either intake or preference at 2% (**Table 1.7A**) or 4% (**Table 1.7B**) sucrose concentration. Further, in alcohol-naïve male mice (n = 6), there was no effect of 0.5 mg/kg URB597 on 4% sucrose intake (control mice: 7.5 ± 0.77 g/kg; URB597 mice: 7.7 ± 1.1 g/kg) or preference ratio (control: 0.93 ± 0.05; URB597: 0.96 ± 0.03).

1.6. Single, acute URB597 (0.5 mg/kg) increased saccharin preference ratio, but not intake in male mice. The specificity of the action of URB597 on alcohol intake was further tested on saccharin (non-caloric reinforcer) intake in mice. Like the sucrose experiments, after 3 weeks of chronic IA, the alcohol tube was switched to saccharin for 3 sessions over 6 days, and after stable intake observed, the male mice assigned to the vehicle-treated or URB597-treated groups had similar saccharin intake 24 hours before the test day. For saccharin intake after 4 hours (**Figure 1.3A**), two-way ANOVA revealed a significant effect of saccharin concentration [$F(1, 20) = 86$, $p<0.00005$] only. *Post hoc* analysis showed that with or without URB597 treatment, mice had more intake at 0.2% concentration than at 0.1% [$p<0.01$].

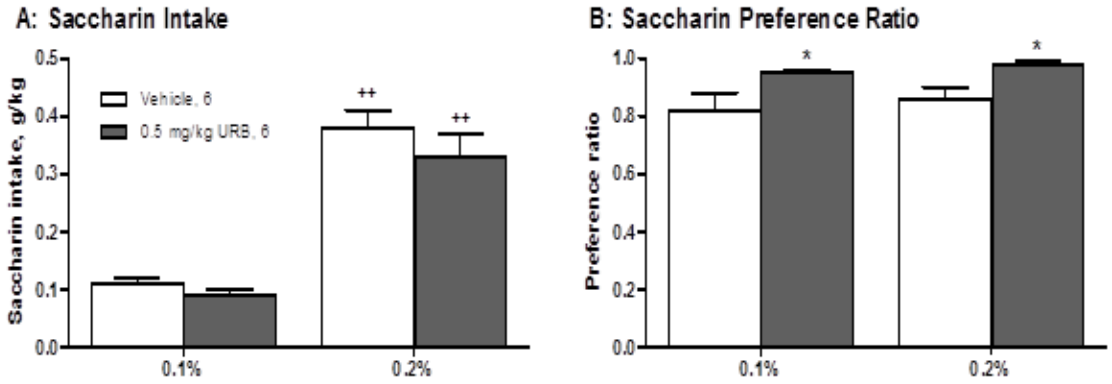


Figure 1.3. *Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 0.1% or 0.2% saccharin intake (**A**) and preference ratio (**B**) in male mice (n=6) after 3-week chronic intermittent access alcohol drinking. Saccharin and water intake values were recorded after 4 hours of saccharin access. *p<0.05 vs. vehicle control at the same saccharin concentration, and ++ p<0.01 vs. 0.1% saccharin.*

For preference ratio (**Figure 1.3B**), two-way ANOVA revealed a significant effect of URB597 treatment only [$F(1, 20) = 13$, $p<0.005$]. *Post hoc* analysis showed that the URB597-treated mice had more preference than the control mice at both 0.1% and 0.2% concentrations [$p<0.05$ for both]. Furthermore, in alcohol-naïve male mice (n=6), there was no effect of 0.5 mg/kg URB597 on 0.1% saccharin intake (control mice: 0.15 ± 0.02 g/kg; URB597 mice: 0.13 ± 0.03 g/kg) or preference ratio (control: 0.95 ± 0.02; URB597: 0.97 ± 0.02).

2. Effect of single, acute or repeated administration of URB597 on ADE in both male and female mice.

2.1. Single, acute URB597. In the single dose URB597 experiment, we tested the effect of URB597 at 0.5 mg/kg in both male and female mice, and found similar reducing effects on the ADE. As there were no significant sex differences in the effects of URB597, the data of each sex are presented separately (**Figure 1.4**).

In males at 4 hours (**Figure 1.4A**), two-way ANOVA revealed a significant effect of URB597 treatment [$F(1, 24) = 4.6$, $p<0.05$], and a significant interaction between Session and Treatment [$F(1, 24) = 4.5$, $p<0.05$], and *post hoc* analysis showed that the URB597-treated males had less intake than the vehicle-treated ones in the ADE session [$p<0.05$]. To test our *a priori* hypothesis that there was an ADE, we included the *post-hoc* result showing that in the vehicle-treated mice, the ADE was significant after 4 hours [$p<0.05$], though 2-way ANOVA did not reveal any significant effect of Session after 4 hours (**Figure 1.4A**). There were no significant effects of ADE or acute URB597 after 8 or 24 hours (data not shown).

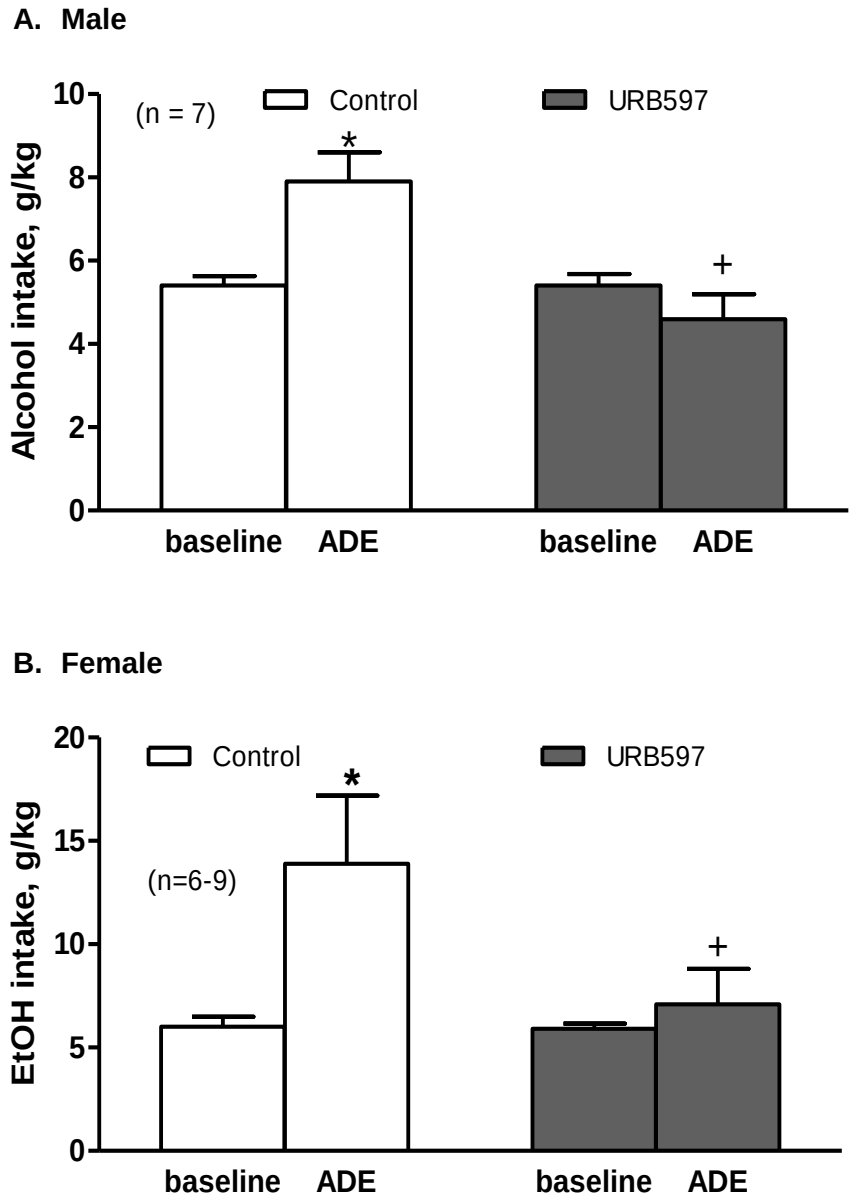


Figure 1.4. *Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on alcohol intake in an alcohol deprivation effect (ADE) model at 4 hours in male (A, n=7) and female (B, n=6-9) mice after 1 week of withdrawal from 3-week chronic intermittent access alcohol drinking. * $p<0.05$ vs. control baseline, and + $p<0.05$ vs. ADE.*

In females at 4 hours (**Figure 1.4B**), two-way ANOVA revealed a significant effect of Session [$F(1, 26) = 4.3$, $p<0.05$], and *post hoc* analysis showed that the females had more intake in the ADE session than the baseline [$p<0.05$]. Though overall Treatment difference was not significant, *post hoc* analysis showed that the URB597-treated females had less intake than the vehicle-treated ones in the ADE session [$p<0.05$] (**Figure 1.4B**). Similar to the male mice, there were no significant effects of ADE or acute URB597 treatment after 8 or 24 hours in females (data not shown).

2.2. Repeated URB597. In the repeated URB597 experiment, we tested the effect of URB597 at 0.5 mg/kg in both male and female mice, and found similar reducing effects on the ADE. As there were no significant sex differences in the effects of URB597, the data of each sex are presented separately (**Figure 1.5**).

In males at 4 hours (**Figure 1.5A**), two-way ANOVA revealed a significant effect of URB597 treatment [$F(1, 20) = 5.8$, $p<0.05$], and a significant interaction between Session and Treatment [$F(1, 20) = 5.9$, $p<0.05$], and *post hoc* analysis showed that the URB597-treated males had less intake than the vehicle-treated ones in the ADE session [$p<0.05$]. To test our *a priori* hypothesis that there was an ADE, we included the *post-hoc* result showing that in the vehicle-treated mice, the ADE was significant after 4 hours [$p<0.05$] (**Figure 1.5A**), though 2-way ANOVA did not reveal any significant effect of Session. There were no significant effects of ADE or repeated URB597 after 8 or 24 hours (data not shown).

In females at 4 hours (**Figure 1.5B**), two-way ANOVA revealed a significant effect of Session [$F(1, 26) = 15.1$, $p<0.001$] and *post hoc* analysis showed that the females had more intake in the ADE session than the baseline [$p<0.01$].

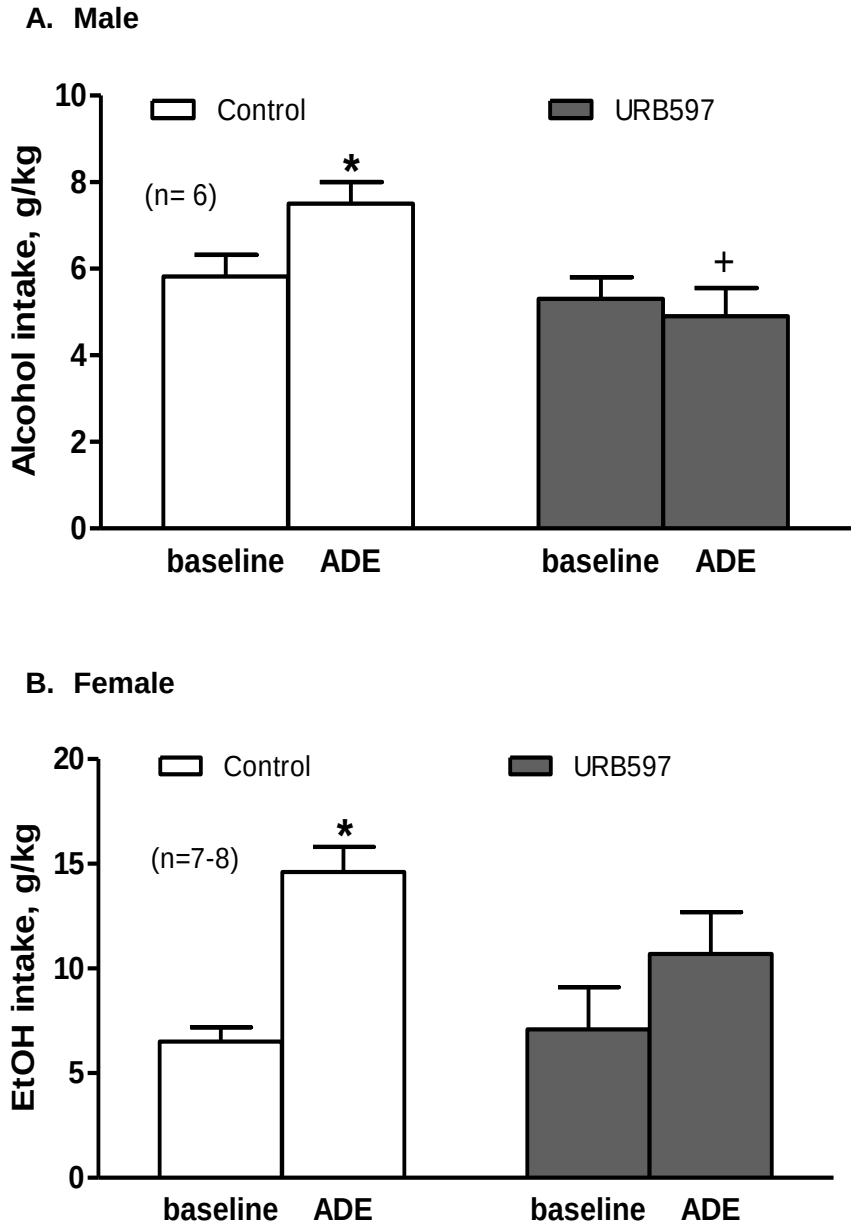


Figure 1.5. *Effects of five repeated administration of URB597 (URB, 0.5 mg/kg) on alcohol intake in an alcohol deprivation effect (ADE) model at 4 hours in male (A, n=6) and female (B, n=7-8) mice after 1 week of withdrawal from 3-week chronic intermittent access drinking. During the withdrawal, mice received five daily URB597 or vehicle injections with the last one 1 day before alcohol bottles were presented again. * $p<0.05$ vs. control baseline, and + $p<0.05$ vs. ADE.*

However, there was no significant effect of the repeated URB597 treatment after 8 or 24 hours (data not shown).

3. Acute (1 day) alcohol withdrawal transiently caused global brain eCB system dysregulation in male mice.

3.1. Acute (1 day) withdrawal. Relative NAE abundances were detected in four mouse brain regions from male mice subjected to the chronic IA model following 1 day alcohol withdrawal as well as from control (water) male mice. All four brain regions showed a significant increase in the NAEs (**Figure 1.6, S1.1**). AEA levels were increased the greatest, with a 3.1, 2.7, 2.1 and 1.9 -fold increase in the NAc, BLA, cerebellum and PFC, respectively (**Figure 1.6A**). Palmitoyl ethanolamide (PEA) and oleoyl ethanolamide (OEA) were similarly increased (**Figures 1.6B, 1.6C, S1.1**), as were other NAEs (data not shown).

3.2. Chronic IA and long-term withdrawal. Relative NAE abundances were further detected in the BLA of mice subjected to the chronic IA model as well as from that of control (water) mice in three groups: immediately after 3 week IA and long-term (1- and 2-week) withdrawal. This was done in order to determine: (A) whether the increased NAEs after 1 day withdrawal were due to chronic alcohol drinking or due to acute withdrawal, and (B) if NAE elevation after acute withdrawal was a short-term response to withdrawal or whether it persisted. Neither of these groups showed changed NAE levels, indicating that NAE elevation was a transient response to acute withdrawal (**Figures 1.7A, S1.2**).

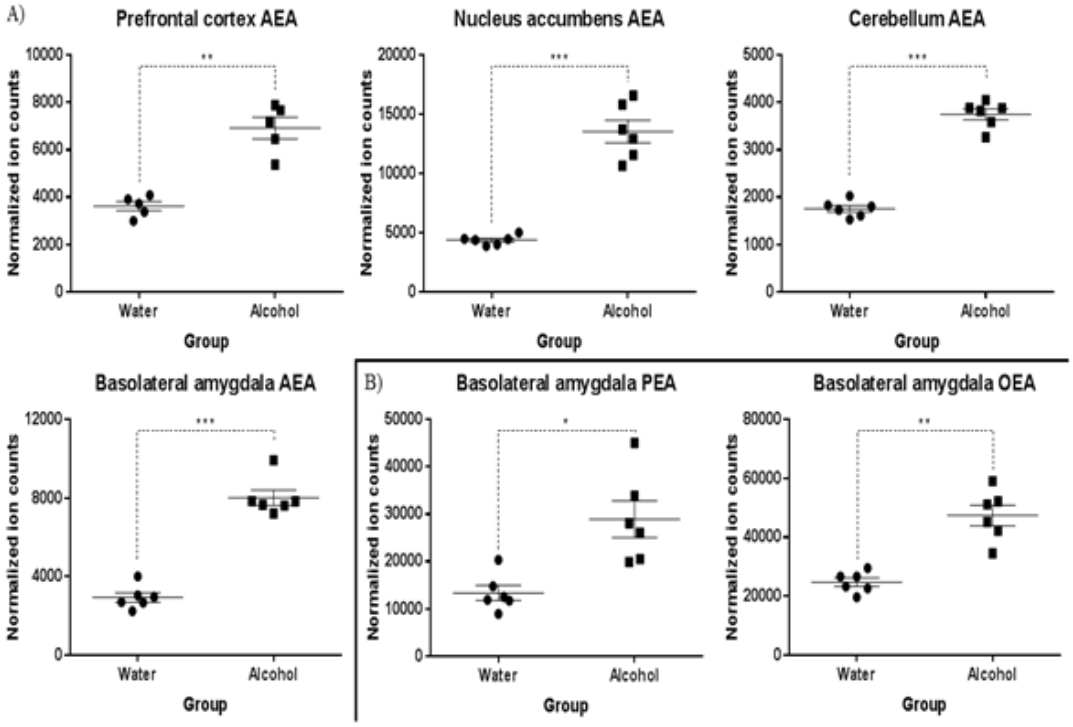


Figure 1.6. *Effects of 1-day withdrawal from 3-week chronic intermittent access (IA) alcohol drinking on brain NAE abundances in male mice (n=6). Test mice are labelled Alcohol and control mice are labelled Water. (A) Anandamide (AEA) abundances in four brain regions after 1 day alcohol withdrawal. (B) Palmitoyl ethanolamide (PEA) and oleoyl ethanolamide (OEA) abundances in the basolateral amygdala after 1 day alcohol withdrawal. Normalized, relative abundances are shown in all graphs. *p<0.05 **p<0.001 and ***p<0.0001 vs. control, error bars indicate standard error of the mean.*

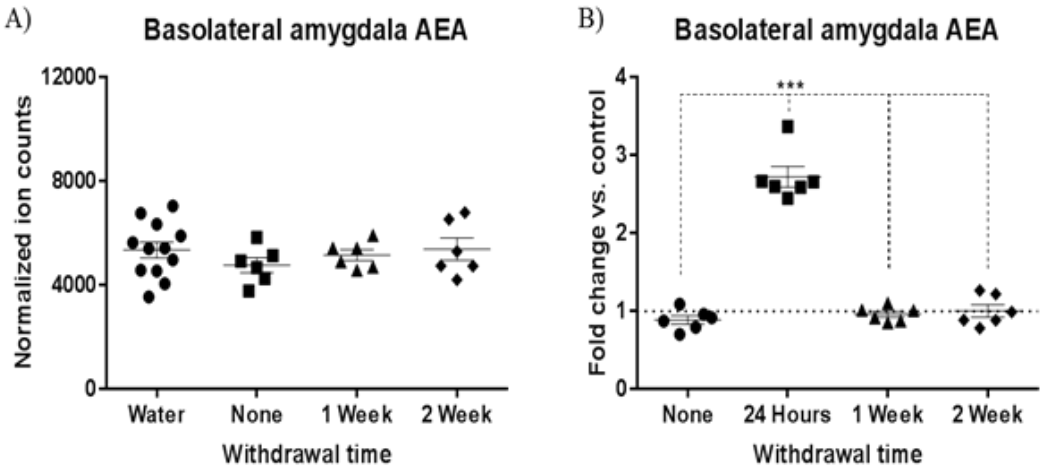


Figure 1.7. *Anandamide (AEA) abundances after chronic 3-week chronic intermittent access (IA) alcohol drinking or long-term withdrawal in male mice. (A) AEA abundances after chronic IA (labelled "none"), 1-week or 2-week withdrawal; Alcohol naïve control mice labelled Water; (B) AEA abundances presented as fold change vs. control for each tested time point. Normalized, relative abundances are shown in all graphs. ***p<0.0001 vs. other time points, error bars indicate standard error of the mean.*

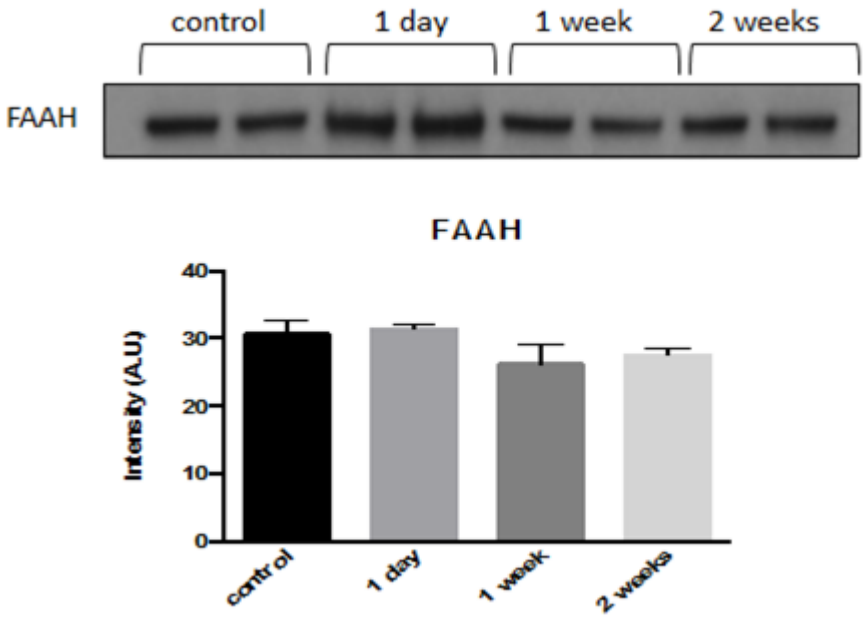


Figure 1.8. *No FAAH changes in male mice (n=6) after acute or long-term withdrawal from 3-week chronic intermittent access (IA) alcohol drinking. A Western blot of FAAH at each time point and the corresponding intensity are shown. Control mice were given only water. Test mice were exposed to chronic IA alcohol (15%) drinking for 3 weeks and then 1 day, 1 week or 2 week of withdrawal.*

As presented as fold change of AEA (**Figure 1.7B**), one-way ANOVA revealed a significant effect of time [$F(3, 20) = 109$, $p<0.000001$], and *post hoc* analysis showed that the 1-day (24 hours) withdrawal had higher ADE levels than other time points examined [$p<0.0001$].

4. FAAH levels in the prefrontal cortex after withdrawal from chronic IA in male mice. Protein lysates prepared from the PFC of male mice were analyzed by Western Blotting. No difference in FAAH levels was detected after acute (1 day) or long-term (1 and 2 weeks) withdrawal, as compared to the water control (**Figure 1.8**).

DISCUSSION

Our first objective in the present study was to investigate the potential of an FAAH inhibitor, URB597, in reducing alcohol consumption in mice after acute (1 day) withdrawal from chronic IA alcohol drinking. Single acute administration of URB597 (a long-lasting inhibitor) significantly reduced 15% alcohol intake in mice in a dose-dependent manner (0.125-1 mg/kg) (**Figure 1.2A**). The URB597-induced reduction of alcohol intake was coupled with a full compensatory increase in water intake, resulting in no change in total fluid intake (**Table S1.1**). This indicates the absence of nonspecific sedative effects by this FAAH inhibitor. Consistently, URB597 (0.125-1 mg/kg) produced a dose-dependent reduction in alcohol preference (**Figure 1.2B**). In the present study, we further confirmed that URB597 reduced alcohol drinking in a CB1-dependent manner, as selective CB1 antagonist AM251 blocked the effects of URB597 (**Table 1.6**). It is unlikely that the effect of URB597 in reducing alcohol intake was secondary to a general suppression of appetitive (anhedonic effect)

or consumption behaviors, since URB597 at 0.5 mg/kg did not decrease sucrose (caloric reinforcer) (**Table 1.7**) or saccharin (non-caloric reinforcer) (**Figure 1.3**) intake. The present study also found a relatively long duration (4 hours) of the effect of single acute URB597 administration on alcohol drinking, which is likely due to its long-lasting bioactivity *in vivo* [Fegley et al. 2005; Basavarajappa et al. 2014]. Together, these results clearly demonstrate that the URB597-induced inhibition of FAAH plays a role in modulating alcohol drinking in mice. Therefore, the development of FAAH inhibitors with improved pharmacokinetics has the potential to yield useful compounds for the treatment of alcoholism and other drug addictions [Panlilio et al. 2013; Zhou & Kreek 2014].

Though URB597 does not produce any reward [Gobbi et al. 2005], repeated administration of URB597 has been found to correct the reduction in sucrose intake after chronic stress in both mice [Rademacher & Hillard, 2007] and rats [Bortolato et al. 2007]. Our results showed that single acute URB597 administration at 0.5 mg/kg (the most effective dose tested for reducing alcohol intake) had no effect on sucrose or saccharin preference in alcohol naïve mice, but increased saccharin preference in mice after 1-week alcohol withdrawal (**Figure 1.3**). Consistently, previous work has found that URB597 does *not* affect saccharin or sucrose drinking in stress-naïve animals (not rewarding) [Basavarajappa et al. 2006; Blednov et al. 2007; Manwell et al. 2009], but increases sucrose preference and/or intake in stress exposed animals due to "anti-depression" properties [Rademacher & Hillard, 2007; Bortolato et al. 2007]. Together, our data suggest that the enhanced eCB activity resulting from the FAAH inhibitor can modulate sensitivity to saccharin reward during alcohol withdrawal. Alternatively, URB597 as an indirect eCB agonist may increase preference for more palatable

reinforcers, such as saccharin (**Figure 1.3**) and sucrose (slight, but not significant, increases at 2% [**Table 1.7A**]), with decreased preferences for less palatable reinforcers (15% and 30 % alcohol), though URB597 itself is not rewarding [Gobbi et al. 2005].

With low alcohol intake models, previous studies tested the FAAH inhibitor URB597 on alcohol-related behaviors: [1] on reinstatements of alcohol-seeking behavior, no effect was found after acute URB597 administration in rats with alcohol intake < 2 g/kg during 1 hour of alcohol access [Cippitelli et al. 2008]; and [2] URB597 was reported to increase alcohol drinking in mice with alcohol intake < 4 g/kg/day during 24 hours of alcohol access [Blednov et al. 2007]. In contrast to mice in the 4-hour limited-access DID paradigm, which had low alcohol intake (<5-6 g/kg/day) [Rhodes et al. 2007], mice in the chronic IA paradigm showed "heavy" drinking with 15-20 g/kg/day in males and 20-25 g/kg/day in females [Hwa et al. 2011; Zhou et al. 2017]. For these reasons, we purposely compared the effects of URB597 on chronic IA with DID, and found that single acute URB597 administration at 0.5 mg/kg reduced alcohol drinking after 1 day withdrawal from chronic IA (**Figure 1.2**), with no effect after 1 day withdrawal from chronic DID (**SI section, Results S1.1**). When tested in other alcohol concentrations, we observed that single acute administration of URB597 at 0.5 mg/kg reduced high levels of alcohol intake with 30% alcohol, but had no effect on low levels of alcohol intake with 7.5% alcohol (**Table 1.5**), consistent with the above notion. Together, our results demonstrate that URB597 modulates alcohol drinking in mice with long-access and chronic exposure to high levels of alcohol intake, but not with either limited access (like DID) or low levels of alcohol intake at 7.5% concentration. In accord with previous research, our

data further indicate that the protective role of the eCBs becomes effective as the withdrawal from "heavy" alcohol drinking is pronounced [Morena et al. 2016]. Drug escalation to high level intake is widely considered a hallmark of the transition from alcohol abuse to drug dependence in humans. Thus, our preclinical study may provide new information about the medical potential of URB597 in the treatment of alcoholism.

Our second main objective was to investigate the potential of URB597 in reducing relapse-like drinking in mice after long-term withdrawal. The phenomenon of a transient increase in alcohol consumption after a period of imposed abstinence has been termed the alcohol deprivation effect (ADE). It can be regarded as an animal model of relapse behavior and craving with good predictive validity [Vengeliene et al. 2014]. In the present study, mice of both sexes displayed the ADE with increased alcohol intake after 1 week of withdrawal (**Figures 1.4, 1.5**). In contrast, the mice pretreated with single acute URB597 administration at 0.5 mg/kg had a blunted ADE in both sexes (**Figure 1.4**), suggesting the involvement of FAAH activity in the relapse-like drinking behavior. This finding with alcohol would constitute an interesting extension to the anti-relapse properties of URB597 observed in cocaine [Chauvet et al. 2014], nicotine [Scherma et al. 2008] or opiate seeking behaviors [Manwell et al. 2009]. One concern about the use of URB597 is that its FAAH inhibitory activity after repeated administration could result in the development of tolerance. Therefore, we tested the effectiveness of the multiple-dosing regimen (5 daily injections before the ADE during the 1-week withdrawal) to mimic the multiple-dosing treatment in the clinic. Like the effect of single acute URB597 administration at 0.5 mg/kg on reducing the ADE in both sexes (**Figure 1.4**), the repeated

pretreatments of URB597 at the same dose during the withdrawal week efficiently prevented the ADE drinking in mice (**Figure 1.5**), showing no tolerance development after this multiple-dosing regimen. It is worth mentioning that the prevention of the ADE by this multiple-dosing regimen occurred with the last URB597 administration 1 day before alcohol re-exposure ADE. Thus, our findings show promising *in vivo* data indicating that the long-lasting FAAH inhibitor URB597 may shows therapeutic potential in alcoholism treatment models.

Previous lines of experimental evidence have suggested chronic alcohol-exposed rats and mice as an animal model of alcohol withdrawal-induced "anxiety" and "depression". These emotional states, and the search for the anxiolytic and antidepressant-like effects of alcohol, might in turn contribute to promoting the high levels of alcohol intake and preference that characterize the escalation drinking observed in the present study in mice [Colombo et al. 1995; Koob and Volkow 2010]. The stress-response eCB system is critically involved in "anxiety"-related behaviors [Dincheva et al. 2015; Parsons & Hurd 2015; Pardini et al et al. 2012; Spagnolo et al. 2016], and FAAH inhibition reverses enhanced anxiety-like behavior present during alcohol withdrawal [Cippitelli et al. 2008]. Our mouse escalation model suggests a role for FAAH activity in the alcohol withdrawal-induced tendency of mice towards "anxiety"- and "depression"- like states. Thus, the increased AEA/CB1 activity after URB597 administration, by inhibiting FAAH activity [Basavarajappa et al. 2014], could be partially responsible for the anxiolytic action of the FAAH inhibitor and subsequent reduction in alcohol intake after acute withdrawal in both male and female mice (**Figure 1.1**). In support of this concept, CB1 knockout mice with higher

“anxiety”-like behavior [Bowers & Ressler 2016] display higher alcohol consumption than the wild-type littermates during the first days [Racz et al. 2003].

Therefore, we sought to determine the effects of chronic high alcohol intake and subsequent alcohol withdrawal on the eCB system at a molecular level to better understand the dynamic changes in different brain regions. We used a targeted mass spectrometry method to quantify the relative abundance of NAE family of eCBs in the NAc, BLA, PFC and cerebellum of mice. Interestingly, after acute 1-day withdrawal from 3-week chronic IA alcohol drinking, all four brain regions showed significant increases in the AEA, OEA and PEA, with the NAc and BLA displaying the largest changes (**Figure 1.6**). As previous studies showed alterations of AEA levels in the amygdala and other brain regions of rats after chronic alcohol or withdrawal [Rubio et al. 2008; Robinson et al. 2016; Henricks et al. 2017], we further determined dynamic profiles of relative abundances of NAEs in the BLA following chronic IA drinking (no withdrawal) and after one and two weeks of alcohol withdrawal, all of which showed no change compared to control (**Figure 1.7A, S1.2**). Together, this indicates that global NAE dysregulation occurs as a response to acute alcohol withdrawal only. This transient increase may be due to decreased FAAH activity, as found in the rats after 1-day alcohol withdrawal [Serrano et al. 2012], but not the FAAH expression level as observed in the present study (**Figure 1.8**). To our knowledge, this is the first description of global NAE dysregulation, both in terms of multi-localization and number of endogenous substrates (AEA, OEA, and PEA), in response to acute alcohol withdrawal, and indicates that a strong eCB response may act as an important molecular coping mechanism in acute alcohol withdrawal, thereby allowing the mice to better cope with alcohol withdrawal stress. As a possible allostatic process, the

AEA surges may be resultant from the absence of an AEA release during 1-day acute withdrawal. Indeed, a recent study has found that the AEA release in the rat nucleus accumbens is observed only after acute alcohol, but not chronic alcohol or chronic withdrawal [Ceccarini et al. 2013], though other reports showed that acute alcohol decreases [Ferrer et al. 2007] or did not change AEA release in this region [Alvarez-Jaimes et al. 2009]. AEA and other NAE elevations were detected at 1-day withdrawal, whereas behavioral responses to the FAAH inhibitor URB597 (inhibition of alcohol drinking) were apparent at both 1-day and 1-week withdrawal time points. With the available data, the best explanation of this phenomenon is that naturally or pharmacologically elevated NAEs are protective from alcohol withdrawal stress. Naturally, this response (observed in tissue content, but not release) occurs only transiently as an acute response to withdrawal and does not persist after a week. However, after one week of alcohol withdrawal, pharmacological elevation of NAEs with an FAAH inhibitor still provides the same anti-stress protection, leading to reduce "relapse" drinking. As the role of eCBs is complicated by their involvement in modulation of alcohol escalation and "relapse" drinking, two other possible explanations should be considered. The AEA changes in other brain regions rather than the four areas tested could be correlated to alcohol withdrawal. Alternatively, our data suggest that the CB1 binding and/or function, and the downstream effects of CB1 receptor signaling can be more involved in the behavioral effects of the FAAH inhibition. Further study to elucidate these findings is warranted. In summary: In line with previous studies on cocaine, nicotine and opiate seeking behaviour [Panlilio et al. 2013], our findings show initial, promising *in vivo* data indicating that

pharmacological FAAH inhibition decreases alcohol escalation and “relapse” drinking in mice.

SUPPLEMENTARY INFORMATION

Results S1.1. No effect of single, acute administration of URB597 (0.25 or 0.5 mg/kg) on 15% alcohol intake after 3-week drinking-in-dark (DID) model. To assess the effect of URB597 on the short-access "binge" alcohol drinking, the mice (n = 6/per group) received URB597 at 0, 0.25 or 0.5 mg/kg after 3 weeks of DID. There was no effect on alcohol intake after 4 hours (control mice: 4.3 ± 0.27 g; 0.25 mg/kg URB597 mice: 3.7 ± 0.39 g; 0.5 mg/kg URB597 mice: 3.9 ± 0.44 g).

Table S1.1. *No effects of single, acute administration of URB 597 (URB, 0.5mg/kg) on total fluid intake in male mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking.*

Total fluid intake, ml	Vehicle (n = 7)	0.5 mg/kg URB (n = 6)
4h	1.7 ± 0.2	1.6 ± 0.1
8h	1.1 ± 0.3	1.3 ± 0.5
24h	2.9 ± 0.3	2.8 ± 0.3

Table S1.2. *Effects of single, acute administration of URB597 (URB, 0.5 mg/kg) on 15% alcohol intake and preference ratio in both male (presented in Figure 1) and female (presented in Figure S1) mice after 1 day of withdrawal from 3-week chronic intermittent access (IA) alcohol drinking at 4 hours. Statistical analysis was run for both male and female mice together. For intake, two-way ANOVA revealed a significant effect of URB597 treatment [F (1, 27) = 24.5, p<0.00005] and a significant effect of sex [F (1, 27) = 9.4, p<0.01], with no significant interaction between URB597 treatment and sex. Post hoc analysis showed that: (1) the vehicle-treated females had more intake than the vehicle-treated males (p<0.05); and (2) both the URB597-treated males and females had less intake than the vehicle-treated ones at 4 hours [p<0.05 and p<0.01, respectively]. For preference ratio, two two-way ANOVA revealed a significant effect of URB597 treatment [F (1, 27) = 18, p<0.0005], with no significant interaction between URB597 treatment and sex. Post hoc analysis showed that both the URB597-treated males and females had less preference than the vehicle-treated ones at 4 hours [p<0.05 for both]. * p<0.05 and **p<0.01 vs. vehicle control in the same sex; + p<0.05 vs. male.*

	Male		female	
	Vehicle (n = 8)	0.5 mg/kg URB (n = 7)	Vehicle (n = 6)	0.5 mg/kg URB (n = 6)
Intake, g/kg	6.0± 0.35	2.8 ± 0.40 *	9.7 ± 1.1 +	5.6 ± 1.0 **
Preference ratio	0.74 ± 0.06	0.48 ± 0.10 *	0.84 ± 0.04	0.56 ± 0.05 *

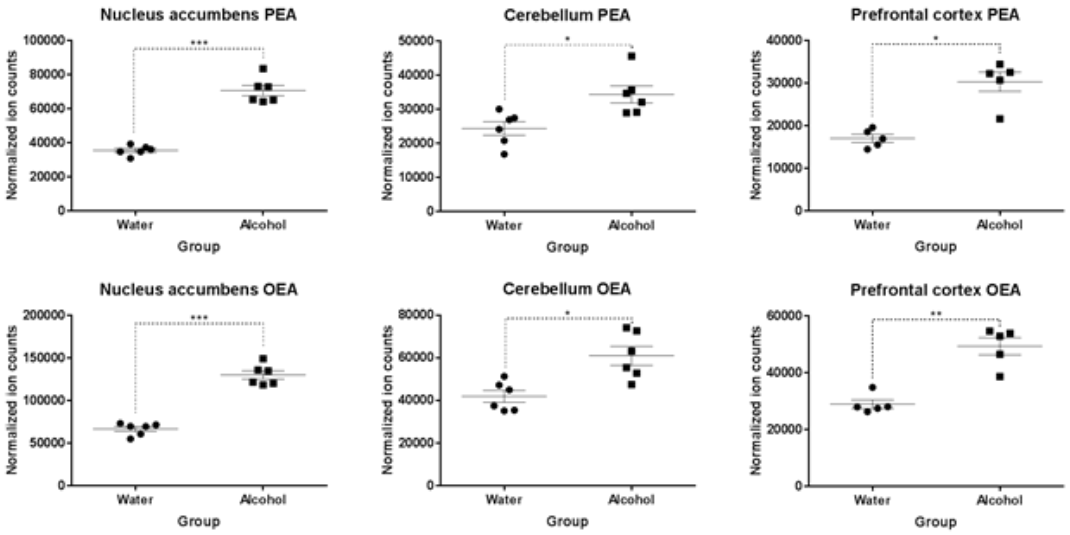


Figure S1.1. *Effects of 1-day withdrawal from 3-week chronic intermittent access (IA) alcohol drinking on brain NAE abundances. Test male mice (labelled Alcohol, n = 6, each n is combined brain regions from 2 mice) were exposed to chronic IA alcohol (15%) drinking for 3 weeks and then 1 day of withdrawal. Control mice (labelled Water, n = 6, each n is combined brain regions from 2 mice) were given only water. NAEs were then extracted and quantified from four brain regions on an LC-MS system. The family of NAEs is generally elevated globally after 1 day alcohol withdrawal, as shown by AEA (shown in Figure 6), palmitoyl ethanolamide (PEA) and oleoyl ethanolamide (OEA) abundances in the basolateral amygdala (shown in Figure 6), nucleus accumbens, cerebellum, and prefrontal cortex. Normalized, relative abundances are shown in all graphs. *p<0.05 **p<0.001 and ***p<0.0001 vs. control, error bars indicate standard error of the mean.*

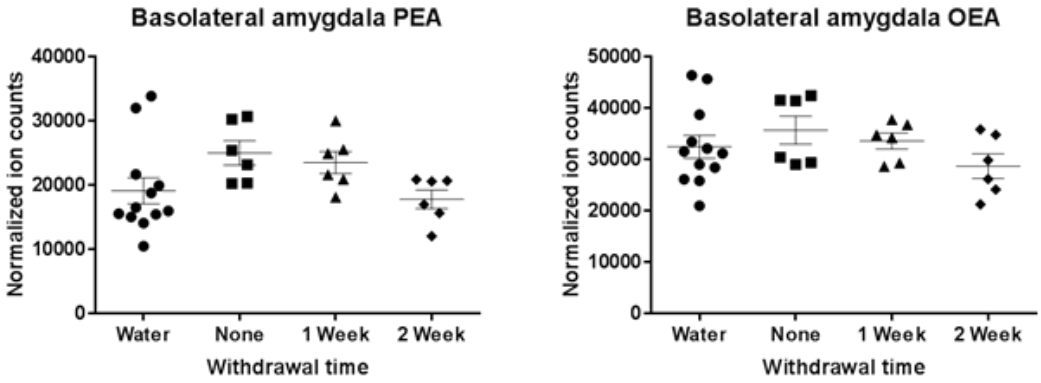


Figure S1.2. *No NAE changes after chronic IA or long-term withdrawal. Test male mice (n = 6, each n is combined brain regions from 2 mice) were exposed to chronic IA alcohol (15%) drinking for 3 weeks and then 1 week or 2 weeks of alcohol withdrawal. Control mice (labelled Water, n = 12, each n is combined brain regions from 2 mice) were given only water. NAEs were then extracted from the basolateral amygdala and quantified on an LC-MS system. Palmitoyl ethanolamide (PEA) and oleoyl ethanolamide (OEA) abundances were unchanged after chronic IA, 1-week or 2-week withdrawal. Normalized, relative abundances are shown in all graphs. Error bars indicate standard error of the mean.*

REFERENCES

Alvarez-Jaimes L, Stouffer DG, Parsons LH (2009) Chronic ethanol treatment potentiates ethanol-induced increases in interstitial nucleus accumbens endocannabinoid levels in rats. J Neurochem 111: 37-48.

Arnone M, Maruani J, Chaperon F, Thiebot MH, Poncelet M, Soubrie P, Le Fur G (1997) Selective inhibition of sucrose and ethanol intake by SR 141716, an antagonist of central cannabinoid (CB1) receptors. Psychopharmacology (Berlin) 132: 104-106.

Basavarajappa BS, Yalamanchili R, Cravatt BF, Cooper TB, Hungund BL (2006) Increased ethanol consumption and preference and decreased ethanol sensitivity in female FAAH knockout mice. Neuropharmacology 50: 834-844.

Basavarajappa BS, Nagre NN, Xie S, Subbanna S (2014) Elevation of endogenous anandamide impairs LTP, learning, and memory through CB1 receptor signaling in mice. Hippocampus 24: 808-818.

Blednov YA, Cravatt BF, Boehm SL 2nd, Walker D, Harris RA (2007) Role of endocannabinoids in alcohol consumption and intoxication: studies of mice lacking fatty acid amide hydrolase. Neuropsychopharmacology 32: 1570-1582.

Bortolato M, Mangieri RA, Fu J, Kim JH, Arguello O, Duranti A, Tontini A, Mor M, Tarzia G, Piomelli D (2007) Antidepressant-like activity of the fatty acid amide hydrolase inhibitor URB597 in a rat model of chronic mild stress. Biol Psychiatry 62:1103-1110.

Bowers ME, Ressler KJ (2016) Sex-dependence of anxiety-like behavior in cannabinoid receptor 1 (Cnr1) knockout mice. Behav Brain Res 300:65-69.

Carnevali L, Vacondio F, Rossi S, Callegari S, Macchi E, Spadoni G, Bedini A, Rivara S, Mor M, Sgoifo A (2015) Antidepressant-like activity and cardioprotective effects of fatty acid amide hydrolase inhibitor URB694 in socially stressed Wistar Kyoto rats. Eur Neuropsychopharmacol 25: 2157-2169.

Ceccarini J, Casteels C, Koole M, Bormans G, Van Laere K (2013) Transient changes in the endocannabinoid system after acute and chronic ethanol exposure and abstinence in the rat: a combined PET and microdialysis study. Eur J Nucl Med Mol Imaging 40: 1582-1594.

Ceccarini J, Hompes T, Verhaeghen A, Casteels C, Peuskens H, Bormans G, Claes S, Van Laere K (2014) Changes in cerebral CB1 receptor availability after acute and chronic alcohol abuse and monitored abstinence. J Neurosci 34: 2822–2831.

Chauvet C, Nicolas C, Thiriet N, Lardeux M, Duranti A, Solinas M (2014) Chronic stimulation of the tone of endogenous anandamide reduces cue- and stress-induced relapse in rats. Int J Neuropsychopharmacol 18: 1-11.

Cippitelli A, Cannella N, Braconi S, Duranti A, Tontini A, Bilbao A, Defonseca FR, Piomelli D, Ciccocioppo R (2008) Increase of brain endocannabinoid anandamide levels by FAAH inhibition and alcohol abuse behaviors in the rat. Psychopharmacology 198: 449-460.

Colombo G, Agabio R, Fà M, Guano L, Lobina C, Loche A, Reali R, Gessa GL (1998) Reduction of voluntary ethanol intake in ethanol-preferring sP rats by the cannabinoid antagonist SR-141716. Alcohol Alcohol 33: 126-130.

Colombo G, Agabio R, Lobina C, Reali R, Zocchi A, Fadda F, Gessa GL (1995) Sardinian alcohol-preferring rats: A genetic animal model of anxiety. Physiol Behav. 57: 1181-1185.

Cravatt BF, Giang DK, Mayfield SP, Boger DL, Lerner RA, Gilula NB (1996) Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. Nature 384: 83-87.

Cravatt BF, Demarest K, Patricelli MP, Bracey MH, Giang DK, Martin BR, Lichtman AH (2001) Supersensitivity to anandamide and enhanced endogenous cannabinoid signaling in mice lacking fatty acid amide hydrolase. Proc Natl Acad Sci U S A 98: 9371-9376.

Dincheva I, Drysdale AT, Hartley CA, Johnson DC, Jing D, King EC, Ra S, Gray JM, Yang R, DeGruccio AM, Huang C, Cravatt BF, Glatt CE, Hill MN, Casey BJ, Lee FS (2015) FAAH genetic variation enhances fronto-amygdala function in mouse and human. Nat Commun 6: 6395-6404.

Fegley D, Gaetani S, Duranti A, Tontini A, Mor M, Tarzia G, Piomelli D (2005) Characterization of the fatty acid amide hydrolase inhibitor cyclohexyl carbamic acid 30-carbamoyl-biphenyl-3-yl ester (URB597): effects on anandamide and oleoylethanolamide deactivation. J Pharmacol Exp Ther 313: 352-358.

Ferrer B, Bermúdez-Silva FJ, Bilbao A, Alvarez-Jaimes L, Sanchez-Vera I, Giuffrida A, Serrano A, Baixeras E, Khaturia S, Navarro M, Parsons LH, Piomelli D, Rodríguez de Fonseca F (2007) Regulation of brain anandamide by acute administration of ethanol. Biochem J 404: 97-104.

Gallate JE, McGregor IS (1999) The motivation for beer in rats: effects of ritanserin, naloxone and SR 141716. Psychopharmacology 142: 302-308.

Gobbi G, Bambico FR, Mangieri R, Bortolato M, Campolongo P, Solinas M, Cassano T, Morgese MG, Debonnel G, Duranti A, Tontini A, Tarzia G, Mor M, Trezza V, Goldberg SR, Cuomo V, Piomelli D (2005) Antidepressant-like activity and modulation of brain monoaminergic transmission by blockade of anandamide hydrolysis. Proc Natl Acad Sci U S A 102:18620-5.

Gunduz-Cinar O, MacPherson KP, Cinar R, Gamble-George J, Sugden K, Williams B, Godlewski G, Ramikie TS, Gorka AX, Alapafuja SO, Nikas SP, Makriyannis A, Poulton R, Patel S, Hariri AR, Caspi A, Moffitt TE, Kunos G, Holmes A (2013) Convergent translational evidence of a role for anandamide in amygdala-mediated fear extinction, threat processing and stress-reactivity. Mol Psychiatry 18, 813–823.

Henricks AM, Berger AL, Lugo JM, Baxter-Potter LN, Bieniasz KV, Petrie G, Sticht MA, Hill MN, McLaughlin RJ. (2017) Sex- and hormone-dependent alterations in alcohol withdrawal-induced anxiety and corticolimbic endocannabinoid signaling. Neuropharmacology (in press).

Hirvonen J, Zanotti-Fregonara P, Umhau JC, George DT, Rallis-Frutos D, Lyoo CH, Li CT, Hines CS, Sun H, Terry GE, Morse C, Zoghbi SS, Pike VW, Innis RB, Heilig M (2013) Reduced cannabinoid CB1 receptor binding in alcohol dependence measured with positron emission tomography. Mol. Psychiatry 18: 916–921.

Houchi H, Babovic D, Pierrefiche O, Ledent C, Daoust M, Naassila M (2005) CB1 receptor knockout mice display reduced ethanol-induced conditioned place preference and increased striatal dopamine D2 receptors. Neuropsychopharmacology 30: 339-349.

Hungund BL, Szakall I, Adam A, Basavarajappa BS, Vadasz C (2003) Cannabinoid CB1 receptor knockout mice exhibit markedly reduced voluntary alcohol consumption and lack alcohol-induced dopamine release in the nucleus accumbens. J Neurochem 84: 698-704.

Hwa LS, Chu A, Levinson SA, Kayyali TM, DeBold JF, Miczek KA (2011) Persistent escalation of alcohol drinking in C57BL/6J mice with intermittent access to 20% alcohol. Alcohol Clin Exp Res 35: 1938-1947.

Kathuria S, Gaetani S, Fegley D, Valiño F, Duranti A, Tontini A, Mor M, Tarzia G, La Rana G, Calignano A, Giustino A, Tattoli M, Palmery M, Cuomo V, Piomelli D (2003) Modulation of anxiety through blockade of anandamide hydrolysis. Nat Med 9: 76-81.

Koob GF, Kreek MJ (2007) Stress, dysregulation of drug reward pathways, and the transition to drug dependence. Am J Psychiatry 164: 1149-1159.

Koob GF, Volkow ND (2010) Neurocircuitry of addiction. Neuropsychopharmacology 35: 217-238.

Lim BK, Huang KW, Grueter BA, Rothwell PE, Malenka RC (2012) Anhedonia requires MC4R-mediated synaptic adaptations in nucleus accumbens. Nature 487: 183-189.

Mangieri RA, Hong KI, Piomelli D, Sinha R (2009) An endocannabinoid signal associated with desire for alcohol is suppressed in recently abstinent alcoholics. Psychopharmacology 205, 63-72.

Manwell LA, Satvat E, Lang ST, Allen CP, Leri F, Parker LA (2009) FAAH inhibitor URB-597 promotes extinction and CB1 antagonist SR141716 inhibits extinction of conditioned aversion produced by naloxone-precipitated morphine withdrawal, but not extinction of conditioned preference produced by morphine in rats. Pharmacol Biochem Behav 94: 154-162.

Manzanares J, Corchero J, Romero J, Fernandez-Ruiz JJ, Ramos JA, Fuentes JA (1999) Pharmacological and biochemical interactions between opioids and cannabinoids. Trends Pharmacol Sci 20: 287-294.

McGregor IS, Dam KD, Mallet PE, Gallate JE (2005) Delta9-THC reinstates beer- and sucrose-seeking behaviour in abstinent rats: comparison with midazolam, food deprivation and predator odor. Alcohol Alcohol 40: 35-45.

Mitrirattanakul S, López-Valdés HE, Liang J, Matsuka Y, Mackie K, Faull KF, Spigelman I (2007) Bidirectional alterations of hippocampal cannabinoid 1 receptors and their endogenous ligands in a rat model of alcohol withdrawal and dependence. Alcohol Clin Exp Res 31: 855-867.

Moreira FA, Kaiser N, Monory K, Lutz B (2008) Reduced anxiety-like behaviour induced by genetic and pharmacological inhibition of the endocannabinoid-degrading enzyme fatty acid amide hydrolase (FAAH) is mediated by CB1 receptors. Neuropharmacology 54: 141-150.

Morena M, Patel S, Bains JS, Hill MN (2016) Neurobiological interactions between stress and the endocannabinoid system. Neuropsychopharmacology 41: 80-102.

Naassila M, Pierrefiche O, Ledent C, Daoust M (2004) Decreased alcohol self-administration and increased alcohol sensitivity and withdrawal in CB1 receptor knockout mice. Neuropharmacology 46: 243-253.

Pang TY, Du X, Catchlove WA, Renoir T, Lawrence AJ, Hannan AJ (2013) Positive environmental modification of depressive phenotype and abnormal hypothalamic-pituitary-adrenal axis activity in female C57BL/6J mice during abstinence from chronic ethanol consumption. Front Pharmacol. 4:93. eCollection 2013.

Panlilio LV, Justinova Z, Goldberg SR (2013) Inhibition of FAAH and activation of PPAR: new approaches to the treatment of cognitive dysfunction and drug addiction. Pharmacol Ther 138: 84-102.

Pardini M, Krueger F, Koenigs M, Raymont V, Hodgkinson C, Zoubak S, Goldman D, Grafman J (2012) Fatty-acid amide hydrolase polymorphisms and post-traumatic stress disorder after penetrating brain injury. Transl Psychiatry. 2:e75.

Parsons LH, Hurd YL (2015) Endocannabinoid signalling in reward and addiction. Nat Rev Neurosci 16: 579-594.

Patel S, Rademacher DJ, Hillard CJ (2003). Differential regulation of the endocannabinoids anandamide and 2-arachidonylglycerol within the limbic forebrain by dopamine receptor activity. J Pharmacol 306: 880-888.

Racz I, Bilkei-Gorzo A, Toth ZE, Michel K, Palkovits M, Zimmer A (2003) A critical role for the cannabinoid CB1 receptors in alcohol dependence and stress-stimulated ethanol drinking. J Neurosci 23: 2453-2458.

Rademacher DJ, Hillard CJ (2007) Interactions between endocannabinoids and stress-induced decreased sensitivity to natural reward. Prog Neuropsychopharmacol Biol Psychiatry 31: 633-641.

Rhodes JS, Best K, Belknap JK, Finn DA, Crabbe JC (2005) Evaluation of a simple model of ethanol drinking to intoxication in C57BL/6J mice. Physiol Behav 84: 53-63.

Robinson SL, Alexander NJ, Bluett RJ, Patel S, McCool BA (2016) Acute and chronic ethanol exposure differentially regulate CB1 receptor function at glutamatergic synapses in the rat basolateral amygdala. Neuropharmacology 108: 474-484.

Rubio M, Fernández-Ruiz J, de Miguel R, Maestro B, Michael Walker J, Ramos JA (2008) CB1 receptor blockade reduces the anxiogenic-like response and ameliorates the neurochemical imbalances associated with alcohol withdrawal in rats. Neuropharmacology 54: 976-988.

Scherma M, Panlilio LV, Fadda P, Fattore L, Gamaleddin I, Le Foll B, Justinová Z, Mikics E, Haller J, Medalie J, Stroik J, Barnes C, Yasar S, Tanda G, Piomelli D, Fratta W, Goldberg SR (2008) Inhibition of anandamide hydrolysis by cyclohexyl carbamic acid 3'-carbamoyl-3-yl ester (URB597) reverses abuse-related behavioral and neurochemical effects of nicotine in rats. J Pharmacol Exp Ther. 327: 482-490.

Serrano A, Rivera P, Pavon FJ, Decara J, Suárez J, Rodriguez de Fonseca F, Parsons LH (2012) Differential effects of single versus repeated alcohol withdrawal on the expression of endocannabinoid system-related genes in the rat amygdala. Alcohol Clin Exp Res 36: 984-994.

Sipe JC, Chiang K, Gerber AL, Beutler E, Cravatt BF (2002) A missense mutation in human fatty acid amide hydrolase associated with problem drug use. Proc Natl Acad Sci USA 99: 8394-8399.

Sloan ME, Gowin JL, Yan J, Schwandt ML, Spagnolo PA, Sun H, Hodgkinson CA, Goldman D, Ramchandani VA (2017) Severity of alcohol dependence is associated with the fatty acid amide hydrolase Pro129Thr missense variant. Addict Biol (in press).

Spagnolo PA, Ramchandani VA, Schwandt ML, Kwako LE, George DT, Mayo LM, Hillard CJ, Heilig M (2016) FAAH gene variation moderates stress response and symptom severity in patients with posttraumatic stress disorder and comorbid alcohol dependence. Alcohol Clin Exp Res 40: 2426-2434.

Varodayan FP, Soni N, Bajo M, Luu G, Madamba SG, Schweitzer P, Parsons LH, Roberto M (2016) Chronic ethanol exposure decreases CB1 receptor function at GABAergic synapses in the rat central amygdala. Addict Biol. 21: 788-801.

Vengeliene V, Bilbao A, Spanagel R (2014) The alcohol deprivation effect model for studying relapse behavior: a comparison between rats and mice. Alcohol 48: 313-320.

Wang L, Liu J, Harvey-White J, Zimmer A, Kunos G (2003) Endocannabinoid signaling via cannabinoid receptor 1 is involved in ethanol preference and its age-dependent decline in mice. Proc Natl Acad Sci USA 100: 1393-1398.

Zhou Y, Kreek MJ (2014) Alcohol: a stimulant activating brain stress responsive systems with persistent neuroadaptation. Neuropharmacology 87: 51-58.

Zhou Y, Huang T, Lee F, Kreek MJ (2016) Involvement of endocannabinoids in alcohol "binge" drinking: Studies of mice with human fatty acid amide hydrolase genetic variation and after CB1 receptor antagonists. Alcohol Clin Exp Res 40: 467-473.

Zhou Y, Rubinstein M, Low MJ, Kreek MJ (2017) Hypothalamic-specific proopiomelanocortin-deficiency reduces alcohol drinking in male and female mice. Genes, Brain and Behavior 16: 449-461.

CHAPTER TWO[†]

IDENTIFICATION OF A DISTINCT METABOLOMIC SUBTYPE OF SPORADIC ALS PATIENTS

INTRODUCTION

Sporadic amyotrophic lateral sclerosis (sALS) is a debilitating neurodegenerative disease for which effective therapeutic options are essentially non-existent [Kiernan et al. 2011]. With a clinical presentation involving gradual failure of motor neurons all around the body, average survival from diagnosis is 2-5 years [Wijesekera et al. 2009]. Our understanding of sALS at a molecular level is severely lacking. To date, no clinically useful biomarkers of the disease exist and the molecular basis for disease pathogenesis is yet to be discovered [Turner et al. 2013].

A number of studies looking at metabolic abnormalities in blood plasma and cerebral spinal fluid of ALS patients suggest that ALS is closely linked with dysregulated metabolism [Lawton et al. 2012; Wuolikainen et al. 2012; Wuolikainen et al. 2011; Wuolikainen et al. 2016; Blasco et al. 2010; Blasco et al. 2013; Kumar et al. 2010]. However, these prior metabolomics studies have focused on familial ALS and consideration of these abnormalities in sALS are hindered by the fact that these studies have been limited to post-mortem tissues and biological fluids, rather than living cells. The *ALS Stratification Prize - Using the Power of Big Data and Crowdsourcing for Catalyzing Breakthroughs in ALS* aims to work towards

[†] Adapted from Chen et al. Identification of a Distinct Metabolomic Subtype of Sporadic ALS Patients. *bioRxiv preprint.* 2018.

understanding heterogeneity in ALS. Several clinical parameters have been identified, such as plasma uric acid, creatinine and blood pressure [Kuffner et al. 2015]. However, disease mechanism-relevant metabolic abnormalities would be more useful in understanding and interpreting patient heterogeneity. In support of cellular metabolism being integral to sALS, our collaborator recently showed that sALS patient-derived skin fibroblasts display bioenergetics abnormalities, including increased mitochondrial membrane potential (MMP), increased oxygen consumption rate (OCR), and increased extracellular acidification rate (ECAR) [Kirk et al. 2014]. In sum, it seems that the study of small metabolites offers a new vantage point to possibly uncover pivotal molecular mechanisms that are relevant to sALS disease.

Identification of molecular biomarkers would potentially allow for stratification of sALS patients into subgroups, providing an opportunity for attempts at more personalized therapeutic approaches. A knowledge of quantifiable sALS disease biomarkers, as opposed to only long-term clinical outcomes, and a recognition of distinct mechanistic subtypes of sALS for patient stratification, would predictably allow for faster and more targeted clinical trials (i.e. using drugs efficacious in specific subgroups), leading to accelerated clinical advances. Finally, an understanding of molecular mechanisms and metabolic pathways involved in disease initiation and/or progression would provide new therapeutic targets for development of successful pharmaceutical interventions.

For highly relevant results, the metabolism-centered study of sALS should be done using live, metabolically active patient-derived cells. For this purpose, skin fibroblasts can be obtained with a minimally invasive punch biopsy and cultured until metabolite extraction and analysis [Vangipuram et al. 2013]. Because these cells share

the genetic make-up of motor neurons, if sALS has a genetic or epigenetic basis, they offer readily accessible disease-relevant insights. Notably, skin abnormalities in ALS, such as reduced elasticity and a delayed return phenomenon, indicate that fibroblasts are also affected by the disease [Ono et al. 1988]. ALS patient skin samples have also been found to display changes in ALS-associated proteins and metabolites, such as FUS [Oketa et al. 2013], TDP-43 [Suzuki et al. 2010], VCP [Ishikawa et al. 2012], VEGF [Suzuki et al. 2009], ubiquitin [Watanabe et al. 2010], MMP-2, MMP-9 [Fang et al. 2009; Fang et al. 2010], TNF-α [Fukazawa et al. 2013], IL-6 [Ono et al. 2001], angiogenin [Higashida et al. 2013], hyaluronic acid [Ono et al. 2000], and collagen [Ono et al. 2000]. Finally, astrocytes generated from ALS fibroblasts have been shown to be toxic to motor neurons, similar to regular ALS astrocytes [Meyer et al. 2014]. In all, these facts make patient fibroblasts a suitable model for live-cell metabolomics studies of sALS.

The following study was based on the hypothesis that subgroups of patient derived sALS skin fibroblasts would have a distinct pattern of metabolic abnormalities that is shared with motor neurons and that is a common hallmark of sALS. The innovation of this study is two-pronged. First, it focuses on live, human-derived cells, a complex system which has not yet been used in the study of sALS. This system allows for unprecedented disease relevance and coverage of small metabolites. Second, this project utilizes mass spectrometry based untargeted global metabolite profiling, an extremely powerful analytical technique, whose full capabilities have not been exercised in pursuing information about sALS.

We first did a number of experiments aimed at the characterization of metabolic profiles of sALS vs. control fibroblasts in order to identify subgroups of

sALS patients based on a shared ‘metabotype.’ We identified an sALS subgroup defined by increased abundance of homocysteine and other compounds related to the trans-sulfuration pathway in the largest such study to date (77 patient-derived dermal fibroblast lines and 45 age- and sex-matched controls). We next worked to validate these results and further characterize this subgroup using stable isotope tracing, untargeted metabolite profiling of plasma from these patients, and a multi-omics analysis including mRNAs and microRNAs from the patient fibroblasts. In all, this study highlights the possibility of stratification of sALS patients based on small molecule metabolites, which could lead to a better understanding of the disease and contribute to the development of personalized therapies.

MATERIALS AND METHODS

Reagents. LC-MS grade acetonitrile (ACN), isopropanol (IPA) and methanol (MeOH) were purchased from Fischer Scientific. High purity deionized water (ddH2O) was filtered from Millipore (18 OMG). OmniTrace glacial acetic acid and ammonium hydroxide were obtained from EMD Chemicals. [U-13C] glucose, [2,3,3-^{2}H]-serine were purchased from Cambridge Isotope Laboratory. Ammonium acetate and all other chemicals and standards were obtained from Sigma Aldrich in the best available grade.

Cell culture: A total of 77 sALS and 45 control fibroblast cell lines were obtained from cases and propogated as previously described [Konrad et al. 2017]. 75,000 cells/well for each line were plated in duplicate 6-well plates and cultured in DMEM medium containing 5 mM glucose and 4 mM glutamine, 10% FBS, 1% of 100X

antibiotic/antimycotic (Ab+F; which contains sterile-filtered 10,000 units penicillin, 10 mg streptomycin and 25 µg amphotericin B per mL, and 2.5 µg/ml Plasmocin). Cells were harvested at 80% confluency and metabolites extracted for LC/MS metabolomic analysis.

Metabolite extraction: Each cell line was cultured and initially extracted as two biological replicates, for independent LC/MS metabolomic data acquisition. Cells were washed rapidly twice with ice-cold PBS and twice with ice-cold water, followed by metabolite extraction using -70ºC 80:20 methanol:water (LC-MS grade methanol, Fisher Scientific). The tissue–methanol mixture was subjected to multiple episodes of vortexing or were bead-beaten for 45 sec using a Tissuelyser cell disrupter (Qiagen). Extracts were centrifuged for 5 min at 5,000 rpm to pellet insoluble material and supernatants were transferred to clean tubes. The extraction procedure was repeated two additional times and all three supernatants were pooled, dried in a Vacufuge (Eppendorf) and stored at -80ºC until analysis. The methanol-insoluble protein pellet was solubilized in 0.2 M NaOH at 95ºC for 20 min and protein was quantified using a BioRad DC assay. On the day of metabolite analysis, dried cell extracts were reconstituted in 70% acetonitrile at a relative protein concentration of 1 µg/ml, and 4 µl of this reconstituted extract was injected for LC/MS-based untargeted metabolite profiling.

These 244 patient-derived skin fibroblast extracts (122 samples with 2 biological replicates) were analyzed by LC-QTOF metabolomics profiling in random sequence. To adjust for day-to-day and batch-to-batch LC/MS instrument drift, metabolite stability, and other experimental factors that may contribute to systematic error, metabolite measurements were normalized to one another using flanking quality

control (QC) samples run at intervals of every 6 injections. These QC samples were prepared from a pool of all samples and this normalization procedure enabled the comparative analysis of data acquired over a 1-month collection period.

Plasma metabolites were extracted by 1:20 dilution of plasma in 70% acetonitrile in ddH2O (vol:vol). The mixture was briefly vortexed and then centrifuged for 5 min at 16,000 × g to pellet precipitated proteins. An aliquot of the resulting extract (3 μl) was subjected to untargeted metabolite profiling using, applying both positive and negative ion monitoring MS.

Untargeted metabolite profiling by LC/MS: Cell extracts were analyzed by LC/MS as described previously [Chen et al. 2012; Ismailoglu et al. 2014], using a platform comprised of an Agilent Model 1290 Infinity II liquid chromatography system coupled to an Agilent 6550 iFunnel time-of-flight MS analyzer. Chromatography of metabolites utilized aqueous normal phase (ANP) chromatography on a Diamond Hydride column (Microsolv). Mobile phases consisted of: (A) 50% isopropanol, containing 0.025% acetic acid, and (B) 90% acetonitrile containing 5 mM ammonium acetate. To eliminate the interference of metal ions on chromatographic peak integrity and electrospray ionization, EDTA was added to the mobile phase at a final concentration of 6 μM. The following gradient was applied: 0-1.0 min, 99% B; 1.0-15.0 min, to 20% B; 15.0 to 29.0, 0% B; 29.1 to 37min, 99% B. Raw data were analyzed using MassHunter Profinder 8.0 and MassProfiler Professional (MPP) 14.9.1 software (Agilent technologies). Mann Whitney t-tests ($p<0.05$) were performed to identify significant differences between groups.

Metabolite Structure Specification To ascertain the identities of differentially expressed metabolites ($P<0.05$), LC/MS data was searched against an in-house

annotated metabolite database created using MassHunter PCDL manager 7.0 (Agilent Technologies), based on monoisotopic neutral masses (<5 ppm mass accuracy) and chromatographic retention times. A molecular formula generator (MFG) algorithm in MPP was used to generate and score empirical molecular formulae, based on a weighted consideration of monoisotopic mass accuracy, isotope abundance ratios, and spacing between isotope peaks. A tentative compound ID was assigned when PCDL database and MFG scores concurred for a given candidate molecule. Tentatively assigned molecules were verified based on a match of LC retention times and/or MS/MS fragmentation spectra for pure molecule standards.

Stable isotope tracing of [U-^{13}C] glucose and [2,3,3-^{2}H] serine: An in-house untargeted stable isotope tracing (USIT) workflow was employed using Agilent metabolite profiling software MassHunter Qualitative Analysis 7.0, MassProfinder 8.0 and MPP 14.9.1. Labelled metabolites were identified on the basis of differential abundance in cells cultured with supplemental heavy isotope-labelled metabolite vs. natural (light) isotope, as described previously [Leung et al. 2017; Chen et al. 2018]. Selected metabolites were identified on the basis of previously curated isotopologues. Notably, the USIT workflow calculates and corrects for the natural abundance of ^{13}C and ^{2}H isotope in samples.

Microarray analysis: Total RNA (combined mRNA and miRNA) from skin fibroblasts was extracted from 27 sALS patients and 27 controls using Trizol agent (Invitrogen). Total RNA was further extracted using the Agilent Absolutely RNA miRNA Kit. The total RNA concentration and RNA sample integrity were verified by a Bio-Analyzer 2100 (Agilent Technologies, Waldbronn, Germany). The quality of isolated RNA was determined using an Agilent 2200 TapeStation system and

Bioanalyzer with mRNA labeling and microarray processing were performed according to the manufacturer's recommendations. miRNA labeling was done using an Agilent miRNA Complete Labeling and Hyb Kit with gene expression and miRNA data extracted using Agilent Feature Extraction Software.

Extracted mRNA and miRNA data were analyzed using the respective workflows in GeneSpring GX 14.9.1 (Agilent Technologies, CA). Signal intensities for each mRNA probe was normalized to 75th percentile values with baseline transformation for gene expression analysis. Separately, miRNA data was normalized to 90th percentile and baseline transformed for miRNA analysis. The differential miRNA list ($P<0.05$, $FC>1.2$) was used to identify the gene targets based on a target prediction database incorporated in GeneSpring GX (e.g. TargetScan, PicTar, microRNA.org). The differentially expression mRNA was combined with validated miRNA targets and metabolites for integrated multi-Omics pathway analysis using databases linked to KEGG, Wiki pathways and Biocyc.

Statistics: All values are averages of at least three independent measurements. Error bars indicate standard deviation (S.D.) or standard error of the mean (S.E.M.). Statistically significant differences between two groups were estimated by unpaired two-tailed Student's test with significance set at $p<0.05$.

RESULTS

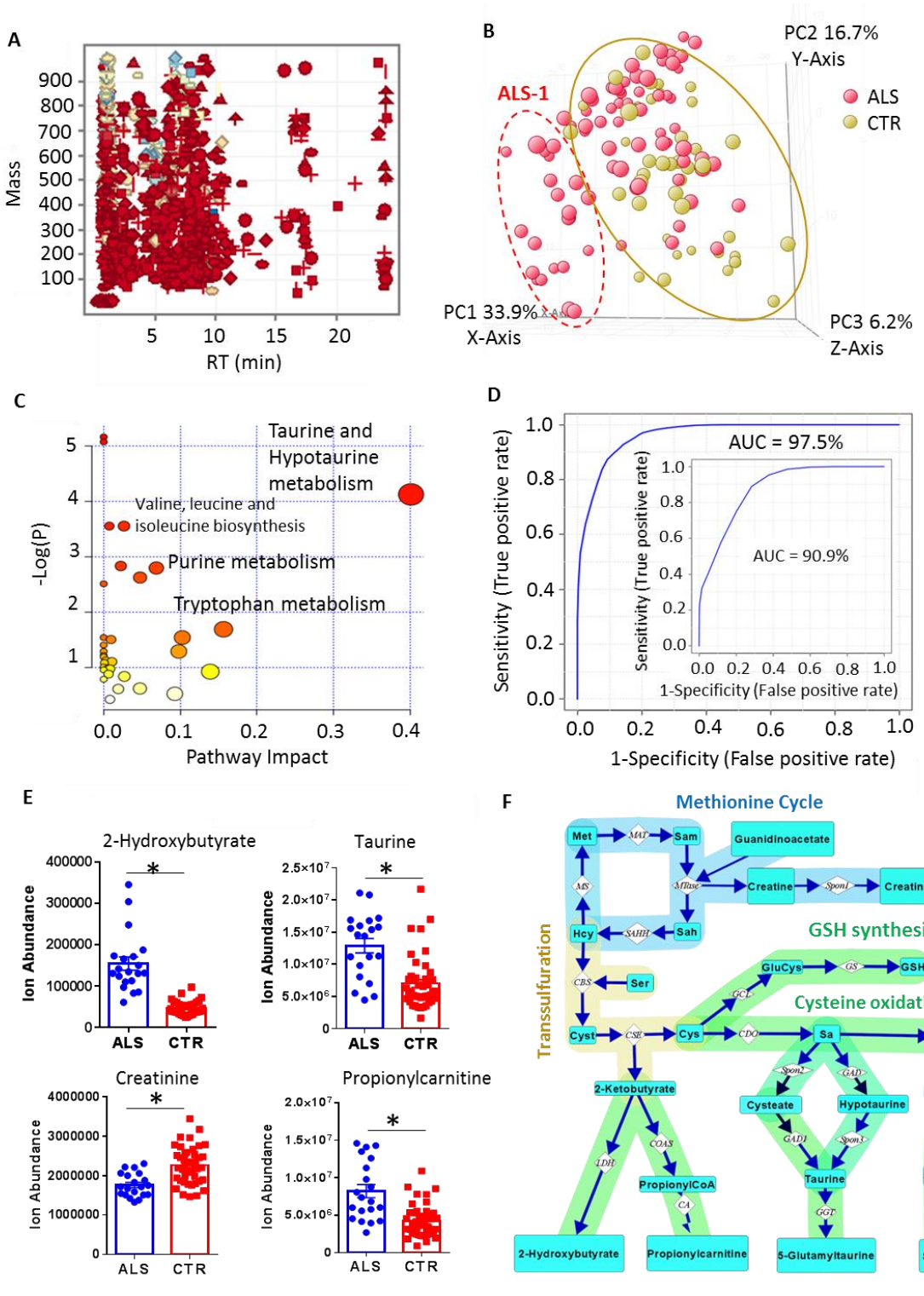
Untargeted metabolite profiling of human skin fibroblasts identifies an sALS subgroup. Metabolite profiling was performed on cultured skin fibroblasts obtained from 77 sALS patients and 45 healthy controls (age and sex matched). 507 of 1172

detected features were differentially abundant in sALS vs control (non-parametric Mann-Whitney test ($P<0.05$)) (Figure 2.1A). Principal component analysis (PCA) using these metabolite abundances highlighted a distinct subgroup of 18 sALS ('sALS-1') cell lines that clustered along PC1 (33.9% variance), revealing a subgroup with a distinct 'metabotype' (Figure 2.1B).

We repeated metabolic profiling on these 18 sALS cell lines (cells were re-thawed, grown and metabolites extracted) and a randomly selected group of 18 control cell lines. 41 metabolites were differentially abundant in both experiments when the sALS-1 was compared to all control lines (Figure S2.1), including significant increases for 2-hydroxybutyrate, taurine, and C3-C5 carnitines and significant decreases in creatinine, ATP and several amino acid (Figure 2.1E). The taurine and hypotaurine pathway was identified as the most significantly impacted in a pathway enrichment and topology analysis (Figure 2.1C).

We generated a 4-metabolite linear support vector machine (SVM) model (using 2-hydroxybutyrate, taurine, propionylcarnitine and creatinine) that predicted the sALS-1 group with an accuracy of 97% in the original group (with 1/45 false-positives and 1/18 false-negatives; Figure 2.1D) and 90.9% accuracy in the repeat analysis. These metabolites are related through the methionine cycle, methylation and trans-sulfuration pathways (Figure 2.1F) and these metabolite profiling results indicate abnormalities in these pathways in the sALS-1 subgroup. Homocysteine can enter the trans-sulfuration pathway as a precursor for cystathionine. Cystathionine cleavage produces 2-ketobutyrate, which can be reduced to 2-hydroxybutyrate or added to propionylcarnitine for β-oxidation in mitochondria.

Figure 2.1. *Mass spectrometry-based untargeted metabolite profiling of patient fibroblast cell lines reveals an sALS subgroup. A) Mass and retention time (RT) scatterplot of 1165 features detected in skin fibroblasts. Metabolites with >70% frequency in one group were included. B) Principal component analysis plot of 77 sALS (red) and 45 control (yellow circles) cell lines based on metabolite abundances. A subgroup of >18 sALS patients (dashed red oval) clearly separates form other cell lines (yellow oval). C) An enrichment and pathway topology analysis of metabolites of interest. Taurine and hypotaurine are identified as the most significantly affected pathways in sALS-1. D) An ROC curve from the support vector machine model based on 4 metabolites. The AUC for sALS-1 prediction (i.e. the overall accuracy, specificity and sensitivity) is 97.5%. The inset is the ROC curve and AUC for the repeat experiment. E) Detected ion abundance of 4 metabolites of interest. *, P<0.05 (2-tailed Student t-test). F) Schematic representation of the methionine cycle, methylation and trans-sulfuration pathways. Met, methionine; Hcy: homocystieine; Sam: S-adenosylmethionine; Sah: S-adenosylhomocysteine; Cyst: cystathionine; Sa: sulfoalanine; 3-SP: 3-sulfopyruvate. The enzyme name abbreviations were adapted from KEGG pathway. MAT: methionine adenosyltransferase**;** MS: methionine synthase; SAHH: S-adenosylhomocysteine hydrolase; MTase: methyltransferase; CBS: cstathionine beta synthase; CSE: cystathionine gamma lyase; GCL: glutamate cysteine ligase; GS: GSH synthase; CDO: cysteine dioxygenase; LDH: lactate dehydrogenase; COAS: propionylCoA synthetase; CA: carnitine acyltransferase; GAD1: glutamate decarboxylase 1; GGT: Gamma-glutamyltransferase; GOT1: glutamic-oxaloacetic transaminase 1; SO: sulfite oxidase; Spon1, Spon2 and Spon3 refer to non-enzymatic spontaneous degradation process.*

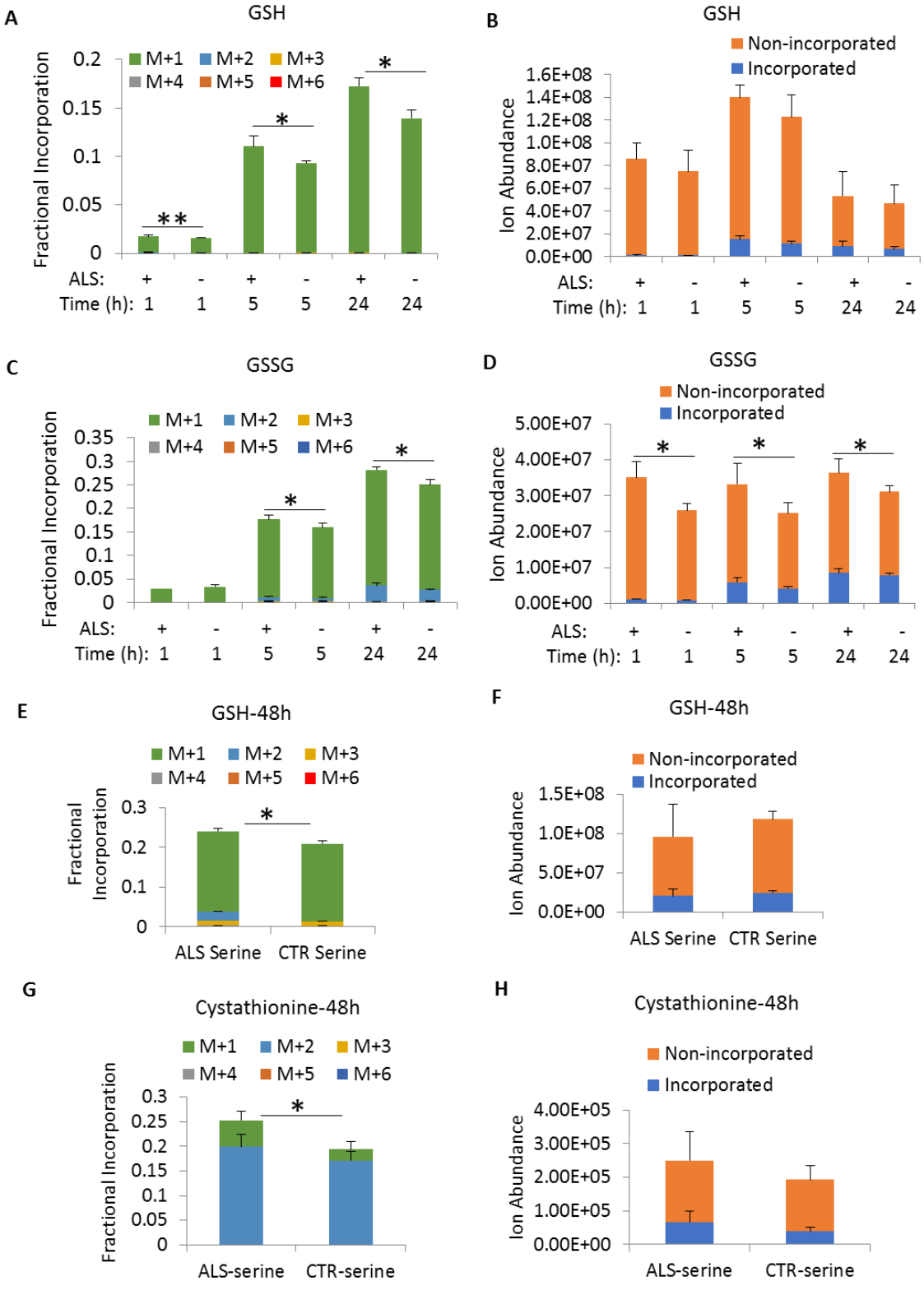

A
Mass
RT (min)
B
PC2 16.7%
Y-Axis
ALS-1
ALS
CTR
PC1 33.9%
X-Axis
PC3 6.2%
Z-Axis
C
-Log(P)
Pathway Impact
Taurine and
Hypotaurine
metabolism
Valine, leucine and
isoleucine biosynthesis
Purine metabolism
Tryptophan metabolism
D
Sensitivity (True positive rate)
1-Specificity (False positive rate)
AUC = 97.5%
AUC = 90.9%
E
2-Hydroxybutyrate
Taurine
Creatinine
Propionylcarnitine
Ion Abundance
ALS
CTR
F
Methionine Cycle
Met
Sam
Guanidinoacetate
Creatine
Creatinine
Hcy
Sah
GSH synthesis
GluCys
GSH
Ser
Cysteine oxidation
Cyst
Cys
Sa
Transsulfuration
2-Ketobutyrate
Cysteate
Hypotaurine
3-SP
PropionylCoA
Taurine
Sulfite
2-Hydroxybutyrate
Propionylcarnitine
5-Glutamyltaurine
Sulfate

Creatine, which is synthesized by methylation of guanidinoacetate [Wyss et al. 2000], and can spontaneously degrade to creatinine. A lower abundance of creatinine indicates reduced S-adenosylmethionine (SAM) mediated methylation.

The sALS-1 subgroup has increased serine incorporation into glutathione. In order to further characterize the trans-sulfuration related abnormality in sALS-1 fibroblasts, we performed stable isotope tracing with deuterated serine. The goal was to assess if these fibroblasts display increased glutathione production from serine-derived glycine. sALS-1 (3 cell lines with high intracellular 2-hydroxyglutarate) and control fibroblasts (3 randomly selected) were grown in 2 mM [2,3,3-^{2}H]-serine supplemented medium. sALS-1 fibroblasts indeed displayed increased serine flux to into GSH and GSSG (Figure 2.2A, B, C, D).

Surprisingly, we did not detect any serine incorporation into cystathionine, indicating that trans-sulfuration from homocysteine to *de novo* synthesized cysteine was insignificant. Serine, together with homocysteine, is a precursor for cystathionine (Figure 2.1F). Deuterium incorporated into cystathionine can therefore indicate the activity of the trans-sulfuration pathway**.** It is possible that cystine in the culture medium was at a high enough concentration to support GSH production. To test this, we cultured cells with deuterated serine for a longer time (48h), when cystine is significantly decreased in the media. Increased serine-derived GSH (Figure 2.2E, F) and cystathionine (Figure 2.2G-H) was indeed observed in the sALS-1 group vs. control at this time point. The cystathionine gamma-lyase/hydrogen sulfide system was previously shown to be essential for maintaining cellular glutathione status [Lee et al. 2014].

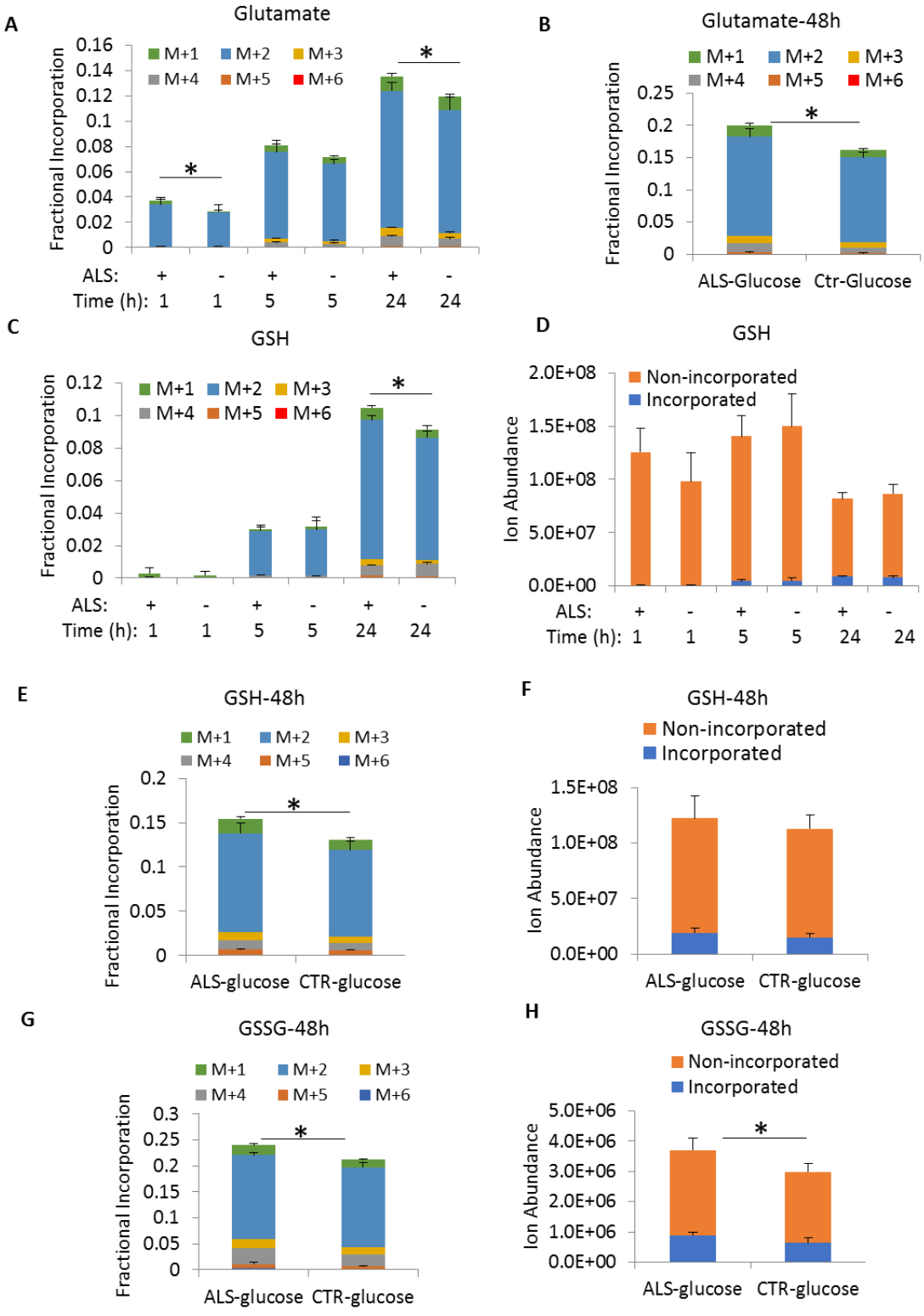
Figure 2.2. *Serine tracing shows increased GSH, GSSG and cystathionine synthesis in the sALS-1 subgroup. A) isotope enrichment and B) total ion abundance of GSH in sALS-1 and control fibroblasts after growth with [2,3,3-^{2}H]-serine for indicated durations. C) isotope enrichment and D) total ion abundance of GSSG in sALS-1 and control fibroblasts after growth with [2,3,3-^{2}H]-serine for indicated durations. E) isotope enrichment and F) total ion abundance of GSH in sALS-1 and control fibroblasts after growth with [2,3,3-^{2}H]-serine for 48h. G) isotope enrichment and H) total ion abundance of cystathionine in sALS-1 and control fibroblasts after growth with [2,3,3-^{2}H]-serine for 48h. 3 sALS-1 and 3 control cell lines were used. Error bars are mean ± S.D. *, P<0.05 (2-tailed Student t-test).*


A
GSH
Fractional Incorporation
M+1
M+2
M+3
M+4
M+5
M+6
ALS:
Time (h):
B
GSH
Non-incorporated
Incorporated
Ion Abundance
C
GSSG
D
GSSG
E
GSH-48h
ALS Serine
CTR Serine
F
GSH-48h
G
Cystathionine-48h
ALS-serine
CTR-serine
H
Cystathionine-48h

The minor incorporation into GSH observed here suggests that the cysteine is primarily used toward oxidation to taurine and reduction to hydrogen sulfide to maintain current GSH supply, as a possible mechanism for protection against oxidative stress.

The sALS-1 subgroup has increased glucose incorporation into glutathione and TCA cycle intermediates. Next, we studied the flux of glucose carbon atoms to GSH, which provides information about the contribution of glutamate to GSH production. Through TCA cycle-derived alpha ketoglutarate, glucose carbons are incorporated into glutamate. We traced glucose carbons in 3 sALS-1 and 3 control fibroblast cell lines by adding 5 mM [U-^{13}C]-glucose to cell culture media and metabolites were extracted after 1h, 5h, 24h and 48h time points. Incorporation of glucose into glutamate was observed at all time points (Figure 2.3A-B). sALS-1 fibroblasts had significantly increased GSH incorporation from glutamate (M+2) at 24h and 48h time points (Figure 2.3C, E). Increased glutamate-derived GSH but unchanged total GSH abundance (Figure 2.3D, F) in the sALS-1 subgroup indicates increased GSH consumption, which is supported by increased ^{13}C labeled GSSG and total GSSG (Figure 2.3G, H). This may be explained as accelerated GSH oxidation to GSSG to provide more antioxidant defense in the sALS-1 group. We also analyzed these samples to determine if the sALS-1 subgroup displayed an overall increased metabolic rate, by looking at the glucose-derived metabolite abundances of intermediates in glycolysis, the TCA cycle, and other metabolic pathways.

Figure 2.3. *Glucose tracing shows increased glutamate, GSH, and GSSG synthesis in the sALS-1 subgroup. A) and B) isotope enrichment of glutamate in sALS-1 and control fibroblasts after growth with [U-^{13}C]-glucose for indicated durations. C) and E) isotope enrichment and D) and F) total ion abundance of GSH in sALS-1 and control fibroblasts after growth with [U-^{13}C]-glucose for indicated durations. G) isotope enrichment and H) total ion abundance of GSSG in sALS-1 and control fibroblasts after growth with [U-^{13}C]-glucose for 48h. 3 sALS-1 and 3 control cell lines were used. Error bars are mean ± S.D. *, P<0.05 (2-tailed Student t-test).*

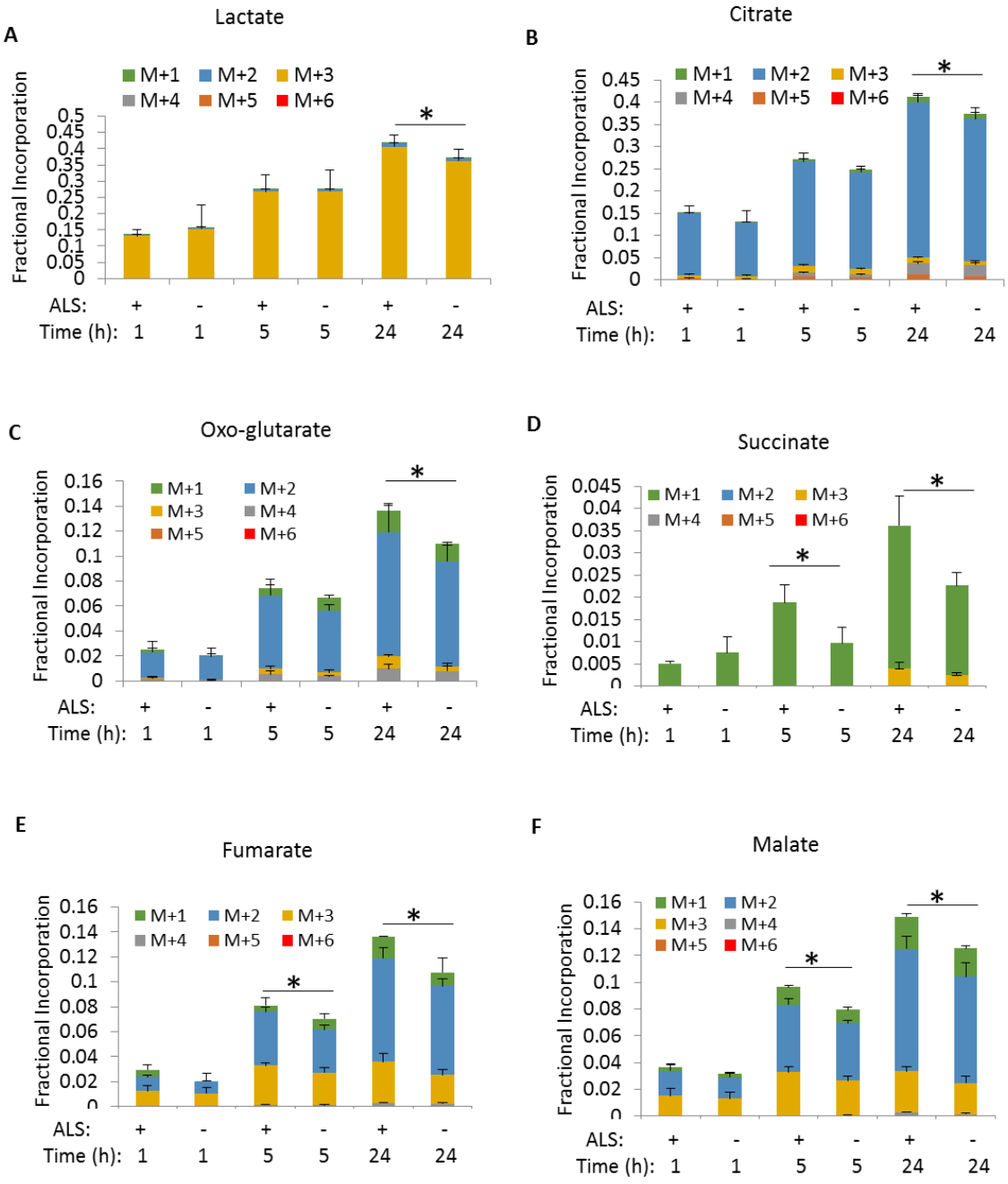

A
Glutamate
Fractional Incorporation
M+1
M+2
M+3
M+4
M+5
M+6
ALS: + - + - + -
Time (h): 1 1 5 5 24 24
B
Glutamate-48h
ALS-Glucose
Ctr-Glucose
C
GSH
D
GSH
Ion Abundance
Non-incorporated
Incorporated
E
GSH-48h
ALS-glucose
CTR-glucose
F
GSH-48h
G
GSSG-48h
H
GSSG-48h

sALS-1 fibroblasts had increased incorporation of glucose into lactate, TCA cycle intermediates, amino acids (aspartate and alanine), and C2 and C4 acyl-carnitines (Figure 2.4A-F and Figure 2.5A-D). These differences indicate an accelerated TCA cycle flux in the sALS-1 subgroup, supportive of increased energy metabolism as has been previously described in sALS fibroblasts [Konrad et al. 2017; Kirk et al. 2014].

sALS-1 subgroup transcriptomic changes support the metabolomic findings. mRNA and microRNA microarray analysis was performed on 27 sALS (11 sALS-1 subgroup & 16 other sALS (sALS-2)) and 27 control cell lines to determine whether the metabolic pathway peculiarities described above in the sALS-1 subgroup were manifest in transcriptional changes. We curated the searchable database to include potential candidate gene sets based on previously published sALS studies, then identified genes based on differentially-expressed transcripts and microRNA targets in sALS subgroups. Using the curated database, Gene Set Enrichment analysis (GSA) identified significantly increased expression of genes involved in the regulation of catabolic processes, apoptosis, and epigenetic regulation by H3K4Me2 in sALS-1 vs. sALS-2 fibroblasts (Figure S2.2). Epigenetic regulation of senescence was suppressed in both sALS groups compared to control.

For a multiomic interpretation, we integrated transcriptome, microRNAome and metabolome data into metabolic pathways queried by KEGG, Wiki-pathways and BioCyc. The super-trans-sulfuration pathway was among the top enriched in sALS-1 ($P<2.3e-4$ by Fischer's exact test, Figure 2.6A).

Figure 2.4. *Glucose tracing shows increased incorporation into A) lactate and TCA cycle intermediates (B) citrate, C) oxoglutarate, D) succinate, E) fumarate and F)malate) in the sALS-1 subgroup fibroblasts. Fibroblasts were cultured for the designated duration with [U-^{13}C]-glucose. Time points where significantly increased incorporation was observed are designated for each metabolite. 3 sALS-1 and 3 control cell lines were used. Error bars are mean ± S.D. *, P<0.05 (2-tailed Student t-test). Note: Due to the overlapping of succinate M+2 with the reference ion (m/z 119.03) used for mass accuracy correction, the calculated M+2 incorporation in succinate is much smaller than the actual incorporated value.*


A
Lactate
B
Citrate
C
Oxo-glutarate
D
Succinate
E
Fumarate
F
Malate
M+1
M+2
M+3
M+4
M+5
M+6
Fractional Incorporation
ALS:
Time (h):
*

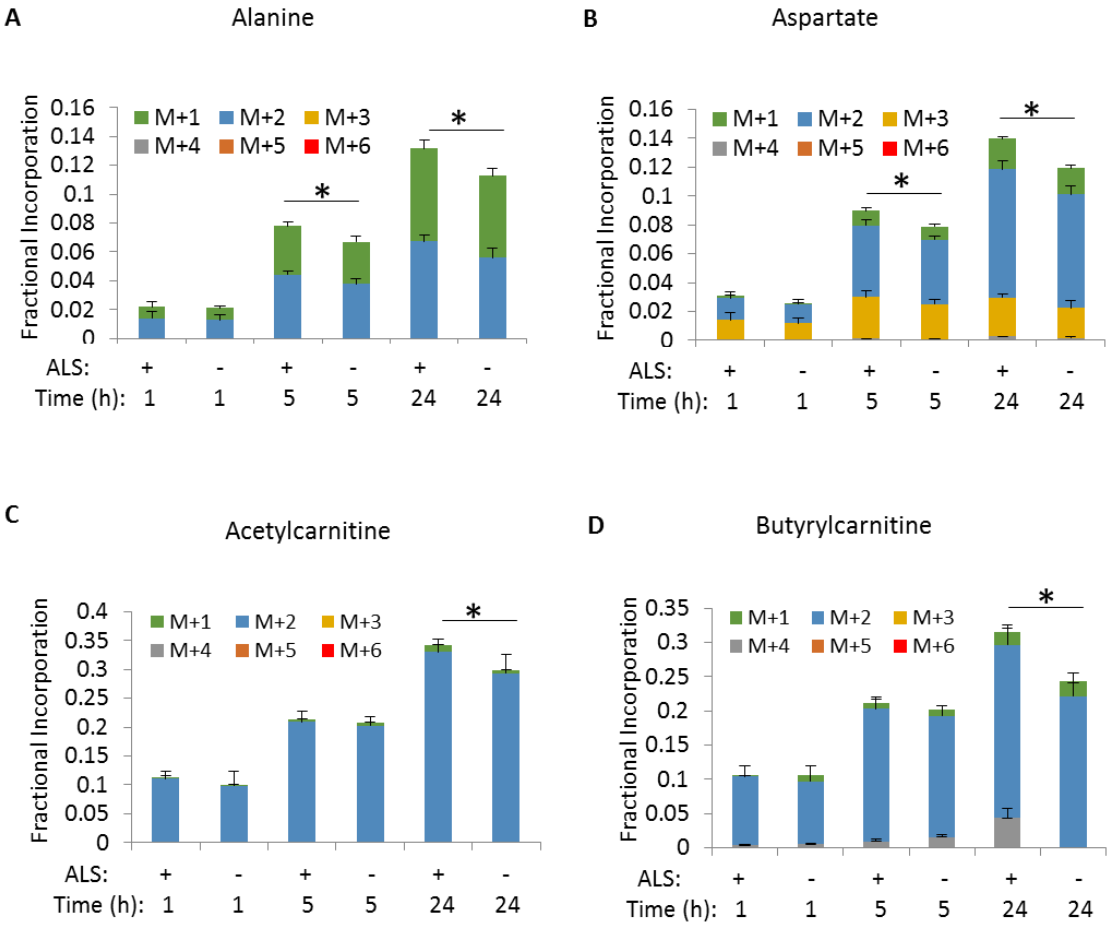


Figure 2.5. *Glucose tracing shows increased incorporation into amino acids (A) alanine and B) aspartate) and acylcarnitines (C) acetylcarnitine and D) Butyrylcarnitine) in the sALS-1 subgroup fibroblasts. Fibroblasts were cultured for the designated duration with [U-^{13}C]-glucose. Time points where significantly increased incorporation was observed are designated for each metabolite. 3 sALS-1 and 3 control cell lines were used. Error bars are mean ± S.D. *, P<0.05 (2-tailed Student t-test).*

Figure 2.6. *An integrated analysis of transcriptomic (microarray) and metabolomics data reveal common pathways that are perturbed in sALS-1 fibroblasts. A) Enrichment analysis of multiomic data identifies the super-transsulfuration pathway as the most significantly different in sALS-1 compared to control group (sALS-2 similarly shows this pathway as significant, but less so than sALS-1) ($p<0.05$, Fisher's exact test). Analysis is based on differentially-abundant (metabolites or transcripts) targets from each 'omic experiement. The metabolic pathways used for mapping were obtained from KEGG, Wiki-pathways and Biocyc and imported into Agilent GeneSpring 14.9.1 for pathway analysis. B) Schematic representation of the affected pathways with differentially expressed mRNA (cyan with purple boxes), miRNA (cyan without purple boxes) and metabolites (yellow) indicated. Charts next to entities represent normalized abundance in control (left bar) and sALS-1 (right bar). Transcript abundances of C) SLC7A11, the cystine-glutamate antiporter D) SLC1A3 , the glutamate aspartate transporter and E) glutathione peroxidase 6 (GPX6) show differential expression in sALS-1 fibroblasts vs. control fibroblasts and highlight the manifestation of the perturbed transsulfuration pathway in sALS-1 subgroup at the transcript level. *, $P<0.05$ (2-tailed Student t-test).*

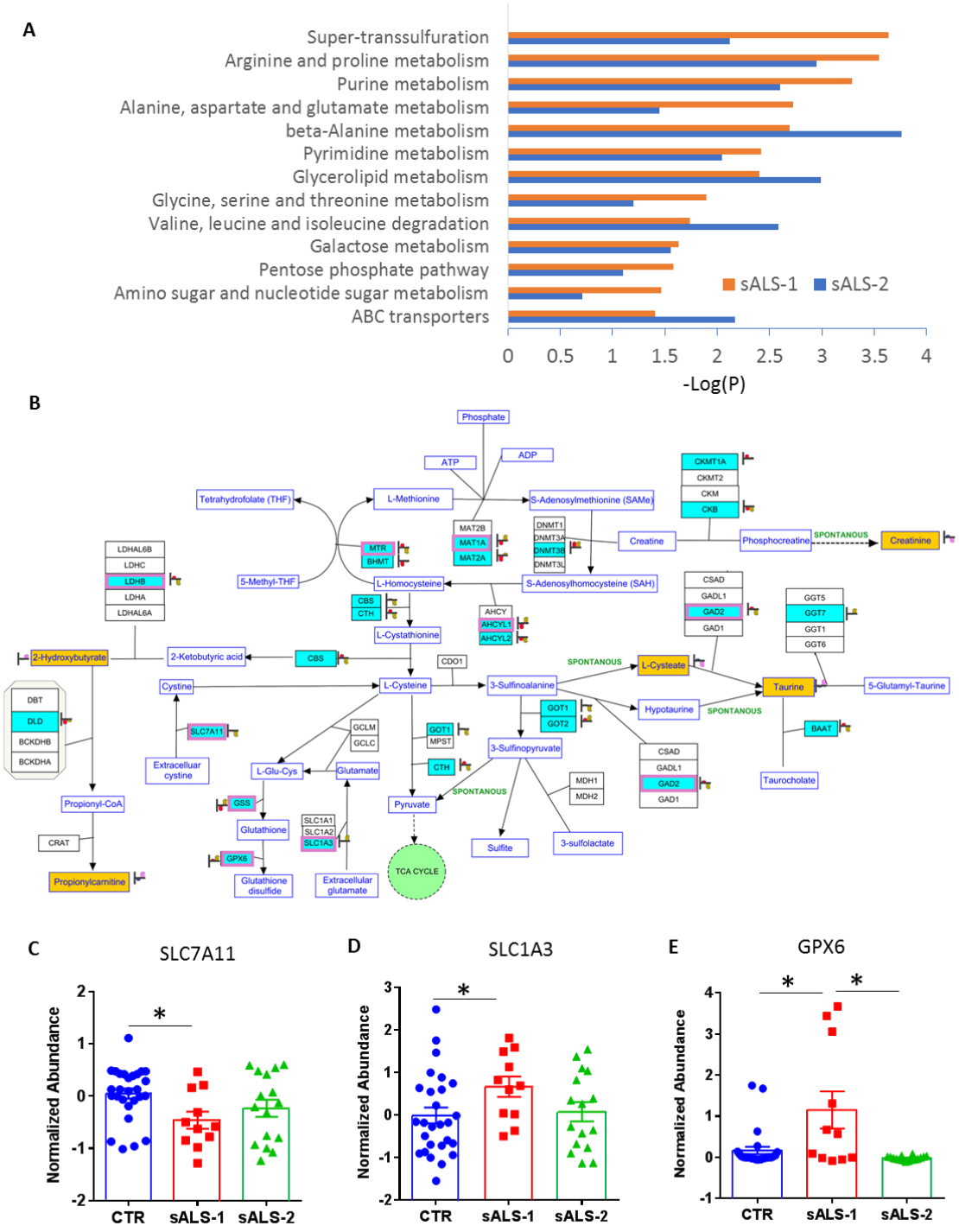
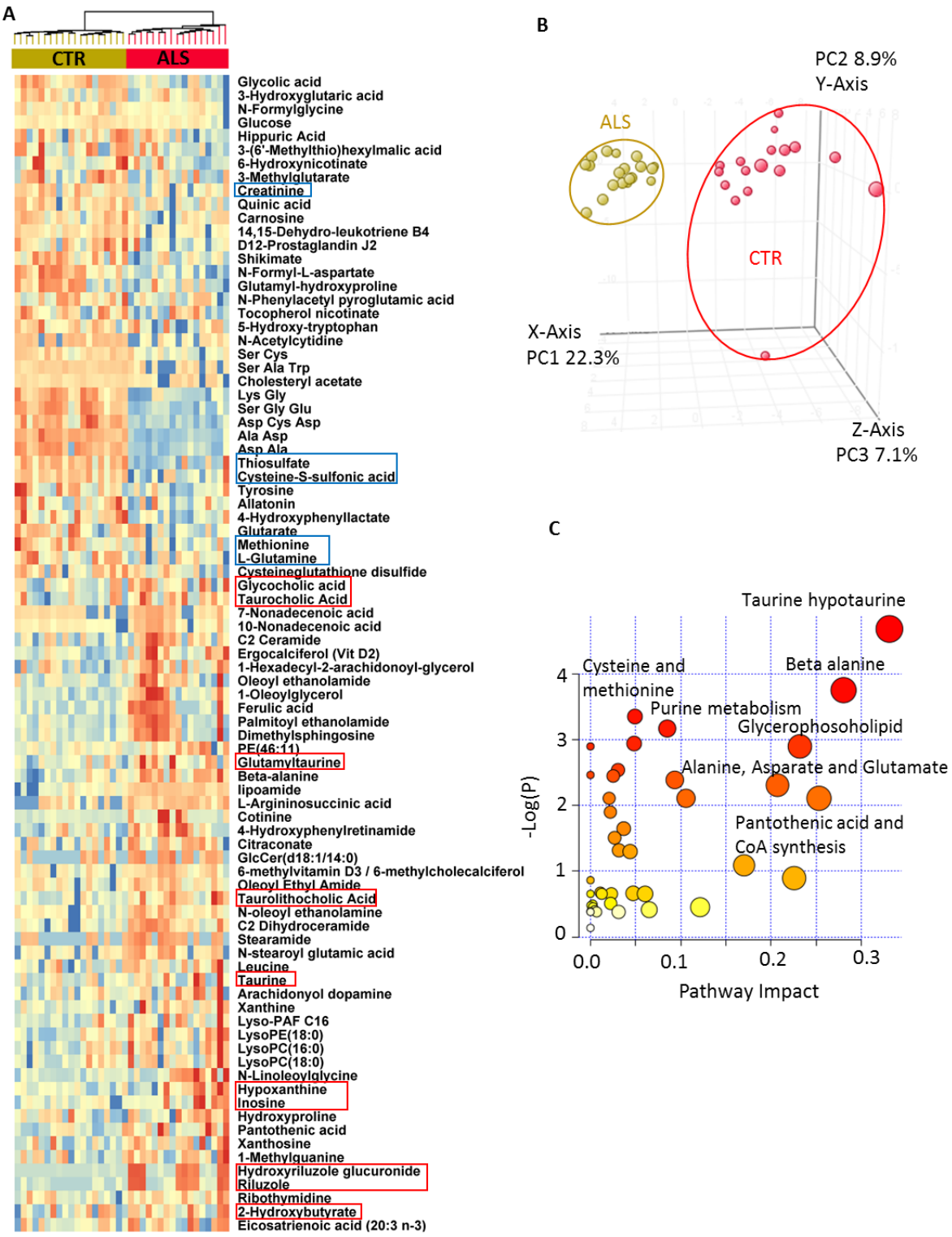

A
Super-transsulfuration
Arginine and proline metabolism
Purine metabolism
Alanine, aspartate and glutamate metabolism
beta-Alanine metabolism
Pyrimidine metabolism
Glycerolipid metabolism
Glycine, serine and threonine metabolism
Valine, leucine and isoleucine degradation
Galactose metabolism
Pentose phosphate pathway
Amino sugar and nucleotide sugar metabolism
ABC transporters
sALS-1
sALS-2
0 0.5 1 1.5 2 2.5 3 3.5 4
-Log(P)
B
Phosphate
ATP
ADP
Tetrahydrofolate (THF)
L-Methionine
S-Adenosylmethionine (SAMe)
Creatine
Phosphocreatine
SPONTANOUS
Creatinine
5-Methyl-THF
L-Homocysteine
S-Adenosylhomocysteine (SAH)
L-Cystathionine
2-Hydroxybutyrate
2-Ketobutyric acid
Cystine
L-Cysteine
3-Sulfinoalanine
L-Cysteate
Taurine
5-Glutamyl-Taurine
Hypotaurine
BAAT
Taurocholate
Extracellular cystine
L-Glu-Cys
Glutamate
Glutathione
Glutathione disulfide
Extracellular glutamate
Pyruvate
TCA CYCLE
3-Sulfinopyruvate
Sulfite
3-sulfolactate
Propionyl-CoA
CRAT
Propionylcarnitine
C
SLC7A11
Normalized Abundance
CTR
sALS-1
sALS-2
D
SLC1A3
E
GPX6

Figure 2.6B is a visualization of both the metabolomic and transcriptomic changes related to the trans-sulfuration pathway and highlights the overlap/support of the multiomic findings (sALS-1 fibroblasts vs. control). Differentially-expressed mRNA transcripts and microRNA targets were related to the methionine cycle, homocysteine to cysteine trans-sulfuration, and glutathione and cysteine metabolism.

The sALS-1 subgroup showed significant transcriptional changes related to glutathione synthesis and oxidation, such as a decrease in SLC25A11 and an increase in SLC1A3 and GPX6 transcripts (Figure 2.6C-E). SLC7A11 is the cystine/glutamate antiporter (system xc-) and is of great interest in ALS, as it is linked to both oxidative stress and excitotoxicity. Cystine transport into the cell is critical for GSH generation, whereas glutamate release can potentially lead to excitotoxicity. SLC7A11 has been shown to be dysregulated in both ALS patients and animal models [Mesci et al. 2015; Albano et al. 2013; Figuera-Losada et al. 2017]. SLC1A3 is the glutamate/aspartate antiporter and transports glutamate into cells or mitochondria. Upregulation possibly indicates increased malate-aspartate shuttle, which is consistent with the increased TCA cycle flux found observed above. GPX6 is an isoform of glutathione peroxidase, which is involved in the oxidation of GSH to GSSG and the upregulation in this transcript is consistent with the accumulation of trans-sulfuration pathway metabolites.

sALS-1 subgroup plasma displays altered trans-sulfuration pathway metabolites. Metabolite profiling of plasma from the same 18 sALS-1 subgroup patients and 20 control subjects was performed. >1000 metabolite features were detected, with 85 differentially-abundant in sALS-1 vs. control (Figure 2.7A, B).

Figure 2.7. *Untargeted metabolite profiling of plasma reveals differentially abundant metabolites for sALS-1 subgroup vs. control that support the multiomic results from the same patients' fibroblast cell lines (abnormal trans-sulfuration related metabolites). A) A heatmap of metabolites found to be significantly different in sALS-1 subgroup plasma compared to control plasma. Increased (red) and decreased (blue) metabolites related to altered taurine and trans-sulfuration pathway, amino acids, lipid metabolism and purine nucleosides are indicated. sALS drug Riluzole and its drug metabolite hydroxyriluzole glucuronide are also highlighted. B) A principal component analysis plot using 88 differentially abundant plasma metabolites shows a clear separation of the sALS-1 subgroup (yellow) from control (red). C) Metabolite pathway enrichment analysis (KEGG pathways, MetaboAnalyst 3.0) highlights the Taurine/Hypotaurine pathway metabolites as the most significantly different in sALS-1 plasma.*


A
CTR
ALS
Glycolic acid
3-Hydroxyglutaric acid
N-Formylglycine
Glucose
Hippuric Acid
3-(6'-Methylthio)hexylmalic acid
6-Hydroxynicotinate
3-Methylglutarate
Creatinine
Quinic acid
Carnosine
14,15-Dehydro-leukotriene B4
D12-Prostaglandin J2
Shikimate
N-Formyl-L-aspartate
Glutamyl-hydroxyproline
N-Phenylacetyl pyroglutamic acid
Tocopherol nicotinate
5-Hydroxy-tryptophan
N-Acetylcytidine
Ser Cys
Ser Ala Trp
Cholesteryl acetate
Lys Gly
Ser Gly Glu
Asp Cys Asp
Ala Asp
Asp Ala
Thiosulfate
Cysteine-S-sulfonic acid
Tyrosine
Allatonin
4-Hydroxyphenyllactate
Glutarate
Methionine
L-Glutamine
Cysteineglutathione disulfide
Glycocholic acid
Taurocholic Acid
7-Nonadecenoic acid
10-Nonadecenoic acid
C2 Ceramide
Ergocalciferol (Vit D2)
1-Hexadecyl-2-arachidonoyl-glycerol
Oleoyl ethanolamide
1-Oleoylglycerol
Ferulic acid
Palmitoyl ethanolamide
Dimethylsphingosine
PE(46:11)
Glutamyltaurine
Beta-alanine
lipoamide
L-Argininosuccinic acid
Cotinine
4-Hydroxyphenylretinamide
Citraconate
GlcCer(d18:1/14:0)
6-methylvitamin D3 / 6-methylcholecalciferol
Oleoyl Ethyl Amide
Taurolithocholic Acid
N-oleoyl ethanolamine
C2 Dihydroceramide
Stearamide
N-stearoyl glutamic acid
Leucine
Taurine
Arachidonyol dopamine
Xanthine
Lyso-PAF C16
LysoPE(18:0)
LysoPC(16:0)
LysoPC(18:0)
N-Linoleoylglycine
Hypoxanthine
Inosine
Hydroxyproline
Pantothenic acid
Xanthosine
1-Methylguanine
Hydroxyriluzole glucuronide
Riluzole
Ribothymidine
2-Hydroxybutyrate
Eicosatrienoic acid (20:3 n-3)
B
PC2 8.9%
Y-Axis
ALS
CTR
X-Axis
PC1 22.3%
Z-Axis
PC3 7.1%
C
Taurine hypotaurine
Cysteine and
methionine
Beta alanine
Purine metabolism
Glycerophosoholipid
Alanine, Asparate and Glutamate
Pantothenic acid and
CoA synthesis
-Log(P)
0
1
2
3
4
0.0
0.1
0.2
0.3
Pathway Impact

Decreased abundance of glutamine, glutamate, methionine, tyrosine and creatinine were observed in sALS-1 samples. Increased metabolites in these samples included taurine, glutamyl-taurine, bile acids, 2-hydroxybutyrate, purine metabolites, phospholipids, ceramides and fatty acid amides. The taurine/hypotaurine pathway was identified as the most significantly different (Figure 2.7C). These results are consistent with the metabolic perturbations identified in fibroblasts and suggest potential for the use of plasma biomarkers for clinical recognition of patients belonging to the sALS-1 subgroup.

DISCUSSION

Our goal was to utilize mass spectrometry-based untargeted metabolite profiling to identify metabolic abnormalities in sALS patient-derived fibroblasts and to attempt to identify sALS subgroups based on these biomarkers. We successfully identified a subgroup of sALS fibroblasts characterized by increased trans-sulfuration pathway activity, displaying increased GSH synthesis and accelerated glucose metabolism. In a multiomic approach, we also found evidence of abnormal trans-sulfuration pathway in fibroblast mRNAs and microRNA. This metabotype was supported by plasma metabolites from the same patients, indicating the possibility of usable plasma (easy to obtain) biomarkers for clinical subgrouping.

Hypermetabolism in ALS has been described before. In fact, we (the Manfredi lab) previously showed that these sALS skin fibroblasts display abnormal mitochondrial bioenergetics, with increased mitochondrial membrane potential (MMP) and MMP to mitochondrial mass (MMP:MM ratio) [Konrad et al. 2017].

Additionally, some ALS patient brains have 10% elevated resting energy expenditure [Cistaro et al. 2012; Weijs et al. 2011; Desport et al. 2001]. On average, the presence of hypermetabolism in ALS results in a worse prognosis [Vucic et al. 2017; Jesus et al. 2017]. Clinically, ^{18}F-FDG-PET imaging has been used to assess differences in glucose uptake in ALS patients, with increased uptake being associated with worse cognitive deficiency [Cistaro et al. 2012; Canosa et al. 2016; Turner et al. 2000]. Notably, an analysis of our previous bioenergetics data from these fibroblasts revealed that the sALS-1 subgroup fibroblasts showed higher MMP and lower MM compared to other ALS patients and controls (Figure S2.3). Our findings suggest that the increased antioxidants in the sALS-1 subgroup may be a protective response to the consequences of hypermetabolism associated with the disease.

Using untargeted mass spectrometry methods, we showed that sALS patients can be subgrouped based on fibroblast metabotype. The identified subgroup displayed increased GSH synthesis, glucose metabolism, and 2-hydroxybutyrate and taurine abundances. We hypothesize that, in about 25% of sALS patients (18/77), increased glucose metabolism leads to greater production of ROS spurring antioxidant GSH synthesis as a defense. Recently, van Weehaeghe et al. demonstrated using PET imaging that stages of ALS disease correlate with regional brain glucose metabolism, disease duration and forced vital capacity [van Weehaeghe et al. 2017]. This indicates that the sALS-1 subgroup may also fall along a spectrum of glucose hypermetabolism with a common metabotype resulting from a cysteine-dependent antioxidant response.

Our findings are supported by the fact that changes in plasma levels 2-hydroxybutyrate, taurine, amino acids and creatinine have been previously reported in ALS patients, with oxidative stress being credited for many of these changes

[Wuolikainen et al. 2016; Lawton et al. 2014; Lawton et al. 2012; Patel et al. 2013]. Decreased plasma creatinine has previously been suggested as an ALS risk factor [van Eijk et al. 2017; Chio et al. 2014]. The fact that metabolites of interest are observed in both fibroblasts and plasma suggests the possibility of plasma metabolomics for clinical stratification of ALS patients.

In the current study, bile acids and bile acid conjugates, known to play a protective role in neurodegenerative disorders [Ackerman et al. 2016], were also increased in the sALS-1 subgroup (Figure 2.7A). There are two primary bile acids produced by the liver: cholic acid and chenodeoxycholic acid, which undergo conjugation with glycine (glycocholic acid and glycochenodeoxycholic acid) or taurine (taurocholic acid and taurochenodeoxycholic acid) prior to secretion in the bile. Increased plasma abundances of taurine and taurocholic acid may also be part of an oxidative stress defense mechanism.

In summary, we described the first application of mass spectrometry-based untargeted metabolite profiling and stable isotope tracing as a means to stratify, predict and characterize a subgroup of sALS patients using metabolically active skin fibroblasts. Our work also highlights the utility of a multiomic approach to support initial findings and more comprehensively characterize results. We identified a distinct sALS metabotype displaying an increase in trans-sulfuration and glucose hypermetabolism. The presence of these changes in patient plasma indicates the possibility of easy clinical stratification of sALS patients using plasma metabolite profiling. As an example, the sASL-1 subtype patients could be identified for yet-undiscovered, targeted anti-oxidant therapies. We hypothesize that the upregulation of the trans-sulfuration pathway arises as an adaptation to limiting cysteine and hence

GSH, potentially triggered by hypermetabolism that accelerates GSH oxidation. The knowledge obtained from metabotypic stratification sALS patients may provide a much-needed first step toward developing metabotype-personalized therapy.

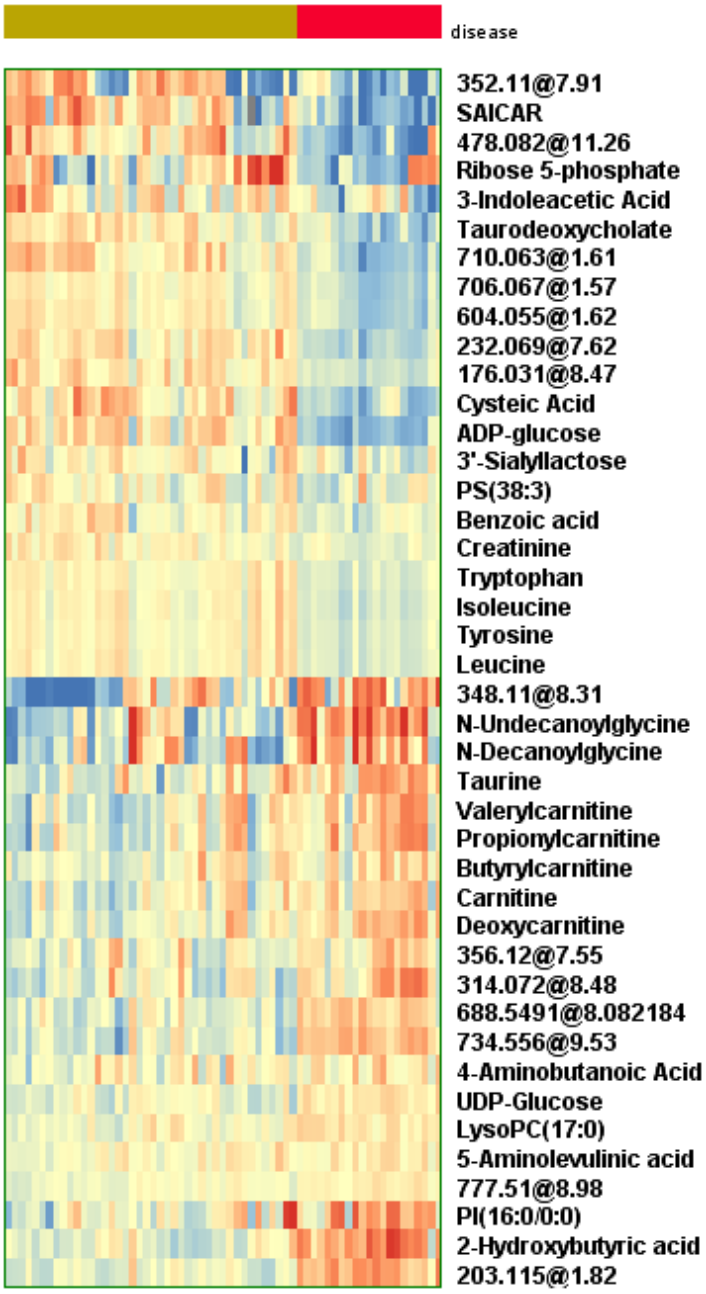


Figure S2.1. *Heatmap of metabolites found to be differentially abundant in both the original and repeat fibroblast metabolite profiling experiments. Unknown metabolites were expressed as accurate mass@retention time. Blue denotes downregulation and red denotes upregulation.*

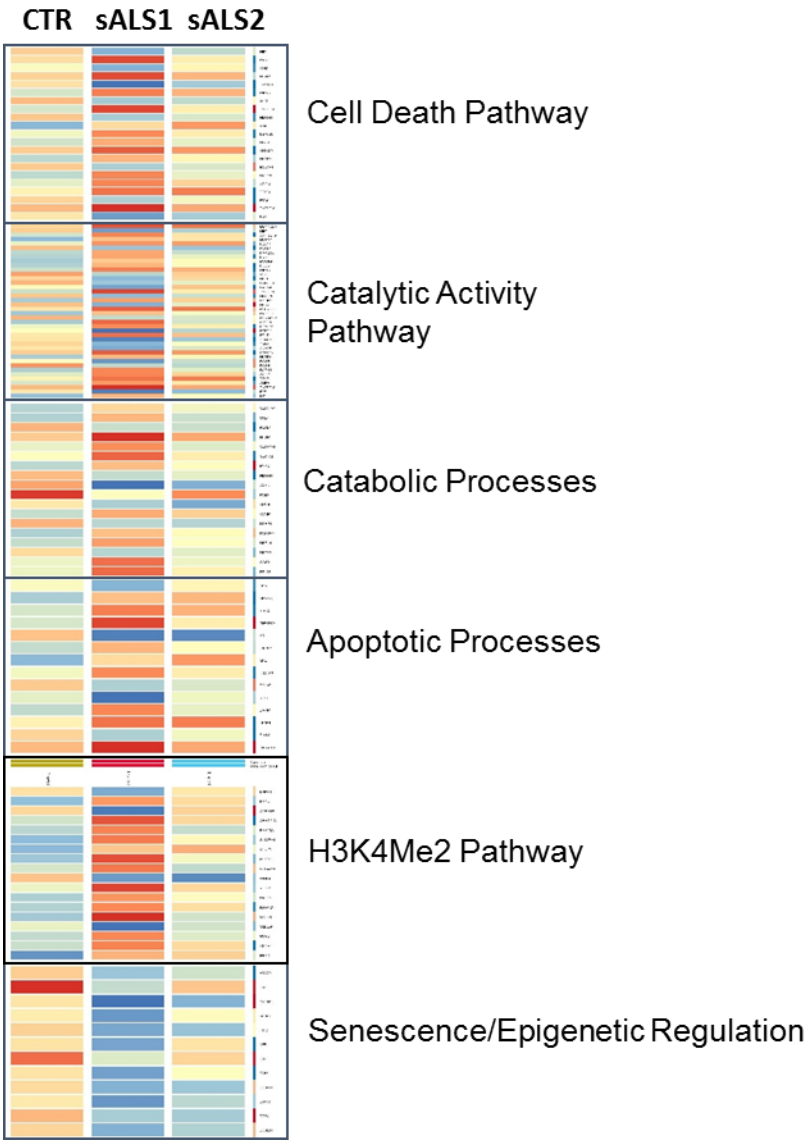


Figure S2.2. *Gene Set Enrichment analysis (GSA) identified significantly different transcript levels of genes involved in the regulation of catabolic processes, apoptosis and epigenetic regulations in sALS-1 compared to sALS-2 and control groups.*

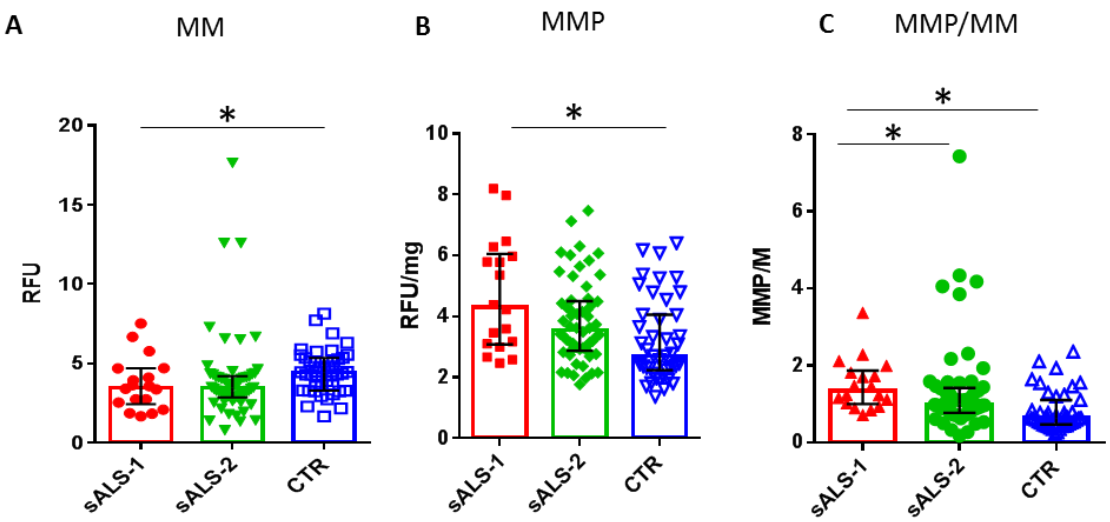


Figure S2.3. *The sALS-1 subgroup showed increased mitochondria membrane potential and mitochondria membrane potential to mass ratio compared to control and non-stratified (sALS-2) patients. The Y axis for A) and B) are relative fluorescence units for the respective dyes used to measure mitochondrial mass (MM) and mitochondrial membrane potential (MMP). Data was obtained as previously described [Konrad et al. 2017]. The numbers of sALS-1, sALS-2 and control cell lines are 18, 58 and 43, respectively. *, P<0.05 (2-tailed Student t-test).*

REFERENCES

Ackerman HD, and Gerhard GS. Bile Acids in Neurodegenerative Disorders. *Front Aging Neurosci.* 2016;8(263.

Albano R, Liu X, and Lobner D. Regulation of system x(c)- in the SOD1-G93A mouse model of ALS. *Exp Neurol.* 2013;250(69-73.

Blasco H, Corcia P, Moreau C, et al. 1H-NMR-based metabolomic profiling of CSF in early amyotrophic lateral sclerosis. *PLoS One*. 2010;5(10):e13223.

Blasco H, Corcia P, Pradat P-F, et al. Metabolomics in cerebrospinal fluid of patients with amyotrophic lateral sclerosis: an untargeted approach via high-resolution mass spectrometry. *J Proteome Res*. 2013;12(8):3746-3754.

Canosa A, Pagani M, Cistaro A, Montuschi A, et al. 18F-FDG-PET correlates of cognitive impairment in ALS. *Neurology.* 2016;86(1):44-9.

Chen Q, Park HC, Goligorsky MS, Chander P, et al. Untargeted plasma metabolite profiling reveals the broad systemic consequences of xanthine oxidoreductase inactivation in mice. *PLoS One.* 2012;7(6):e37149.

Chen Q, Kirk K, Shurubor YI, Zhao D, et al. Rewiring of Glutamine Metabolism Is a Bioenergetic Adaptation of Human Cells with Mitochondrial DNA Mutations. *Cell Metab.* 2018.

Chio A, Calvo A, Bovio G, Canosa A, et al. Amyotrophic lateral sclerosis outcome measures and the role of albumin and creatinine: a population-based study. *JAMA Neurol.* 2014;71(9):1134-42.

Cistaro A, Valentini MC, Chio A, Nobili F, et al. Brain hypermetabolism in amyotrophic lateral sclerosis: a FDG PET study in ALS of spinal and bulbar onset. *Eur J Nucl Med Mol Imaging.* 2012;39(2):251-9.

Desport JC, Preux PM, Magy L, Boirie Y, et al. Factors correlated with hypermetabolism in patients with amyotrophic lateral sclerosis. *Am J Clin Nutr.* 2001;74(3):328-34.

Fang L, Huber-Abel F, Teuchert M, et al. Linking neuron and skin: Matrix metalloproteinases in amyotrophic lateral sclerosis (ALS). *J Neurol Sci*. 2009;285(1-2):62-66.

Fang L, Teuchert M, Huber-Abel F, et al. MMP-2 and MMP-9 are elevated in spinal cord and skin in a mouse model of ALS. *J Neurol Sci*. 2010;294(1-2):51-56.

Figuera-Losada M, Thomas AG, Stathis M, Stockwell BR, et al. Development of a primary microglia screening assay and its use to characterize inhibition of system xc(-) by erastin and its analogs. *Biochem Biophys Rep.* 2017;9(266-72.

Fukazawa H, Tsukie T, Higashida K, Fujikura M, Ono S. An immunohistochemical study of increased tumor necrosis factor-α in the skin of patients with amyotrophic lateral sclerosis. *J Clin Neurosci*. 2013;20(10):1371-1376.

Higashida K, Tsukie T, Fukazawa H, Fujikura M, Ono S. Immunohistochemical studies of angiogenin in the skin of patients with amyotrophic lateral sclerosis. *J Neurol Sci*. 2013;326(1-2):18-23.

Ishikawa H, Yasui K, Oketa Y, Suzuki M, Ono S. Increased expression of valosin-containing protein in the skin of patients with amyotrophic lateral sclerosis. *J Clin Neurosci*. 2012;19(4):522-526.

Ismailoglu I, Chen Q, Popowski M, Yang L, et al. Huntingtin protein is essential for mitochondrial metabolism, bioenergetics and structure in murine embryonic stem cells. *Dev Biol.* 2014;391(2):230-40.

Jesus P, Fayemendy P, Nicol M, Lautrette G, et al. Hypermetabolism is a deleterious prognostic factor in patients with amyotrophic lateral sclerosis. *Eur J Neurol.* 2017.

Kiernan MC, Vucic S, Cheah BC, Turner MR, et al. Amyotrophic lateral sclerosis. *Lancet.* 2011;377(9769):942-55.

Kirk K, Gennings C, Hupf JC, Tadesse S, et al. Bioenergetic markers in skin fibroblasts of sporadic amyotrophic lateral sclerosis and progressive lateral sclerosis patients. *Ann Neurol.* 2014;76(4):620-4.

Konrad C, Kawamata H, Bredvik KG, Arreguin AJ, et al. Fibroblast bioenergetics to classify amyotrophic lateral sclerosis patients. *Mol Neurodegener.* 2017;12(1):76.

Kuffner R, Zach N, Norel R, Hawe J, et al. Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression. *Nat Biotechnol.* 2015;33(1):51-7.

Kumar A, Bala L, Kalita J, et al. Metabolomic analysis of serum by (1) H NMR spectroscopy in amyotrophic lateral sclerosis. *Clin Chim Acta*. 2010;411(7-8):563-567.

Lawton KA, Cudkowicz ME, Brown MV, Alexander D, et al. Biochemical alterations associated with ALS. *Amyotroph Lateral Scler.* 2012;13(1):110-8.

Lawton KA, Brown MV, Alexander D, Li Z, et al. Plasma metabolomic biomarker panel to distinguish patients with amyotrophic lateral sclerosis from disease mimics. *Amyotroph Lateral Scler Frontotemporal Degener.* 2014;15(5-6):362-70.

Lee ZW, Low YL, Huang S, Wang T, et al. The cystathionine gamma-lyase/hydrogen sulfide system maintains cellular glutathione status. *Biochem J.* 2014;460(3):425-35.

Leung KY, Pai YJ, Chen Q, Santos C, et al. Partitioning of One-Carbon Units in Folate and Methionine Metabolism Is Essential for Neural Tube Closure. *Cell Rep.* 2017;21(7):1795-808.

Liu X, Rush T, Zapata J, and Lobner D. beta-N-methylamino-l-alanine induces oxidative stress and glutamate release through action on system Xc(-). *Exp Neurol.* 2009;217(2):429-33.

Mesci P, Zaidi S, Lobsiger CS, et al. System xC- is a mediator of microglial function and its deletion slows symptoms in amyotrophic lateral sclerosis mice. *Brain.* 2015;138(Pt 1):53-68.

Meyer K, Ferraiuolo L, Miranda CJ, et al. Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. *Proc Natl Acad Sci U S A.* 2014;111(2):829-832.

Oketa Y, Higashida K, Fukasawa H, Tsukie T, Ono S. Abundant FUS-immunoreactive pathology in the skin of sporadic amyotrophic lateral sclerosis. *Acta Neurol Scand*. 2013;128(4):257-264.

Ono S, Imai T, Shimizu N, Nakayama M, Yamano T, Tsumura M. Serum markers of type I collagen synthesis and degradation in amyotrophic lateral sclerosis. *Eur Neurol*. 2000;44(1):49-56.

Ono S, Imai T, Tsumura M, et al. Increased serum hyaluronic acid in amyotrophic lateral sclerosis: relation to its skin content. *Amyotroph Lateral Scler Other Motor Neuron Disord*. 2000;1(3):213-218.

Ono S, Toyokura Y, Mannen T, Ishibashi Y. "Delayed return phenomenon" in amyotrophic lateral sclerosis. *Acta Neurol Scand*. 1988;77(2):102-107.

Ono S, Hu J, Shimizu N, Imai T, Nakagawa H. Increased interleukin-6 of skin and serum in amyotrophic lateral sclerosis. *J Neurol Sci*. 2001;187(1-2):27-34.

Patel SS, Molnar MZ, Tayek JA, Ix JH, et al. Serum creatinine as a marker of muscle mass in chronic kidney disease: results of a cross-sectional study and review of literature. *J Cachexia Sarcopenia Muscle.* 2013;4(1):19-29.

Suzuki M, Mikami H, Watanabe T, et al. Increased expression of TDP-43 in the skin of amyotrophic lateral sclerosis. *Acta Neurol Scand*. 2010;122(5):367-372.

Suzuki M, Watanabe T, Mikami H, et al. Immunohistochemical studies of vascular endothelial growth factor in skin of patients with amyotrophic lateral sclerosis. *J Neurol Sci*. 2009;285(1-2):125-129.

Turner MR, Bowser R, Bruijn L, et al. Mechanisms, models and biomarkers in amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener*. 2013;14 Suppl 1:19-32.

Turner MR, and Leigh PN. Positron emission tomography (PET)--its potential to provide surrogate markers in ALS. *Amyotroph Lateral Scler Other Motor Neuron Disord.* 2000;1 Suppl 2(S17-22.

van Eijk RPA, Eijkemans MJC, Ferguson TA, Nikolakopoulos S, et al. Monitoring disease progression with plasma creatinine in amyotrophic lateral sclerosis clinical trials. *J Neurol Neurosurg Psychiatry.* 2017.

van Weehaeghe D, Ceccarini J, Willekens SM, de Vocht J,et al. Is there a glucose metabolic signature of spreading TDP-43 pathology in Amyotrophic Lateral Sclerosis? *Q J Nucl Med Mol Imaging.* 2017.

Vangipuram M, Ting D, Kim S, Diaz R, Schüle B. Skin punch biopsy explant culture for derivation of primary human fibroblasts. *J Vis Exp*. 2013;(77):e3779.

Vucic S. Hypermetabolism appears to be an adverse prognostic biomarker in amyotrophic lateral sclerosis: a potential for therapeutic intervention? *Eur J Neurol.* 2017.

Watanabe T, Okeda Y, Yamano T, Ono S. An immunohistochemical study of ubiquitin in the skin of sporadic amyotrophic lateral sclerosis. *J Neurol Sci*. 2010;298(1-2):52-56.

Weijs PJ. Hypermetabolism, is it real? The example of amyotrophic lateral sclerosis. *J Am Diet Assoc.* 2011;111(11):1670-3.

Wijesekera LC, Leigh PN. Amyotrophic lateral sclerosis. *Orphanet J Rare Dis*. 2009;4:3

Wuolikainen A, Andersen PM, Moritz T, Marklund SL, Antti H. ALS patients with mutations in the SOD1 gene have an unique metabolomic profile in the cerebrospinal fluid compared with ALS patients without mutations. *Mol Genet Metab*. 2012;105(3):472-478.

Wuolikainen A, Jonsson P, Ahnlund M, Antti H, et al. Multi-platform mass spectrometry analysis of the CSF and plasma metabolomes of rigorously matched amyotrophic lateral sclerosis, Parkinson's disease and control subjects. *Mol Biosyst.* 2016;12(4):1287-98.

Wuolikainen A, Moritz T, Marklund SL, Antti H, Andersen PM. Disease-related changes in the cerebrospinal fluid metabolome in amyotrophic lateral sclerosis detected by GC/TOFMS. *PLoS One*. 2011;6(4):e17947.

Wyss M, and Kaddurah-Daouk R. Creatine and creatinine metabolism. *Physiol Rev.* 2000;80(3):1107-213.

CHAPTER THREE

TOWARDS CHARACTERIZING SLC25 FAMILY MITOCHONDRIAL TRANSPORTERS WITH MASS SPECTROMETRY-BASED METABOLOMICS

INTRODUCTION

The solute carrier SLC25 family of proteins is a group of sequence-homologous proteins that localize to the inner mitochondrial membrane and serve to transport a diverse array of small molecule metabolites in and out of the mitochondria. In humans, there are 53 *SLC25* genes, which are considered to be integral for physiological metabolism [Palmieri et al. 2016]. Mutation or interruption of normal function for many of these proteins have been associated with disease [Palmieri 2014]. For example, SLC25A8, an uncoupling protein and carboxylate transporter, has been linked to obesity and type 2 diabetes [Zhang et al. 2001; Zhang et al. 2006], and mutations of SLC25A38, the glycine transporter, have been shown to compromise transport function and give rise to congenital sideroblastic anemia [Guernsey et al. 2009]. Reciprocally, upregulation of SLC25A38 was observed to occur in acute lymphoblastic leukemia cases [Chen et al. 2014]. SLC25 family proteins alternatively employ antiport (most common), symport or uniport as a fundamental modes of transport. Transport is energy-independent, driven by the trans-membrane concentration gradients of cognate substrates, and is sometimes impacted by the chemical or electrical potentials across the mitochondrial membrane [Palmieri 2013].

Understanding the function of these transporters is integral to our understanding of cellular metabolism in health and disease. However, many of these transporters remain either uncharacterized, with no known substrates, or incompletely characterized, with a few identified substrates but the potential to transport additional substrates [Palmieri 2013]. Similarly, we know of many compounds with integral metabolic or pathologic roles which must traverse the mitochondrial membrane with transporter assistance (charged/polar compounds), but for which no transporter has yet been discovered, such as formate, 2-hydroxyglutarate and, until recently, serine.

In this study, we were specifically interested in identifying the mitochondrial 2-hydroxyglutarate and serine transporters. While during the course of this investigation a mitochondrial serine transporter was identified, it is notable that this did not prove to be an SLC25 family member [Kory et al. 2018]. Notwithstanding, we hypothesize that there may also be an SLC transporter family member that contributes to serine transport.

D-2-hydroxyglutarate is an oncometabolite that is produced at high levels in gliomas and acute myeloid leukemia from the precursor 2-oxoglutarate by mutations causing hyperactivity of mitocondrial isocitrate dehydrogenase (IDH2) [Xu et al. 2011]. IDH-1 is the cytosolic form of this enzyme that, when mutated, can also provide an oncogenic source [Achouri et al. 2004; Dang et al. 2017]. Additionally, L-2-hydroxyglutarate is an oncometabolite that can be produced under some circumstances by malate dehydrogenase (MDH1 in the cytosol or MDH2 in the mitochondria) or the cytosolic enzyme lactate dehydrogenase A (LDHA) [Rzem et al. 2006; Ye et al. 2018]. Both the L and D forms of 2-hydroxyglutarate are metabolized by mitochondria-localized enzymes, L-2-hydroxyglutarate dehydrogenase or D-2-

hydroxyglutarate dehydrogenase, respectively [Dang et al. 2017; Ye et al. 2018]. Notably, although it is appreciated that cytosol produced 2-hydroxyglutarate (from IDH1 or LDHA) must be transported from cytosol to mitochondria for metabolic consumption/detoxification, the relevant mitochondrial transporter has eluded discovery [Engqvist et al. 2014].

In the case of serine, multiple important physiological roles are recognized, including function as a precursor for glycine and methylene tetrahydrofolate (THF) formation, ultimately providing the one carbon donor for synthesis of purines, thymidine, phospholipids, and methylation reactions involving DNA, proteins and small molecules [Locasale 2013]. At the start of this work, no mitochondrial serine transporter was recognized, although a novel transporter was recently found [Kory et al. 2018].

Whereas the specific cargo of current "orphan" SLC25 family transporters is unknown, general class functions can be inferred based on homology to known *SLC25* family members (such as 'amino acid' or 'carboxylate' transporters (see Figure 3.1, from [Palmieri 2013])).

Studies were performed focusing on functional characterization of SLC25A38, a presumed glycine transporter, which has not been fully characterized in terms of its cargo. Importantly, SLC25A38 has been presumed, without experimental proof, to additionally transport (or antiport) serine [Guernsey et al. 2009; Fernández-Murray et al. 2016]. We also selected the homologous amino acid transporters SLC25A39, SLC25A40 and SLC25A44 as alternative mitochondrial serine transporter candidates.

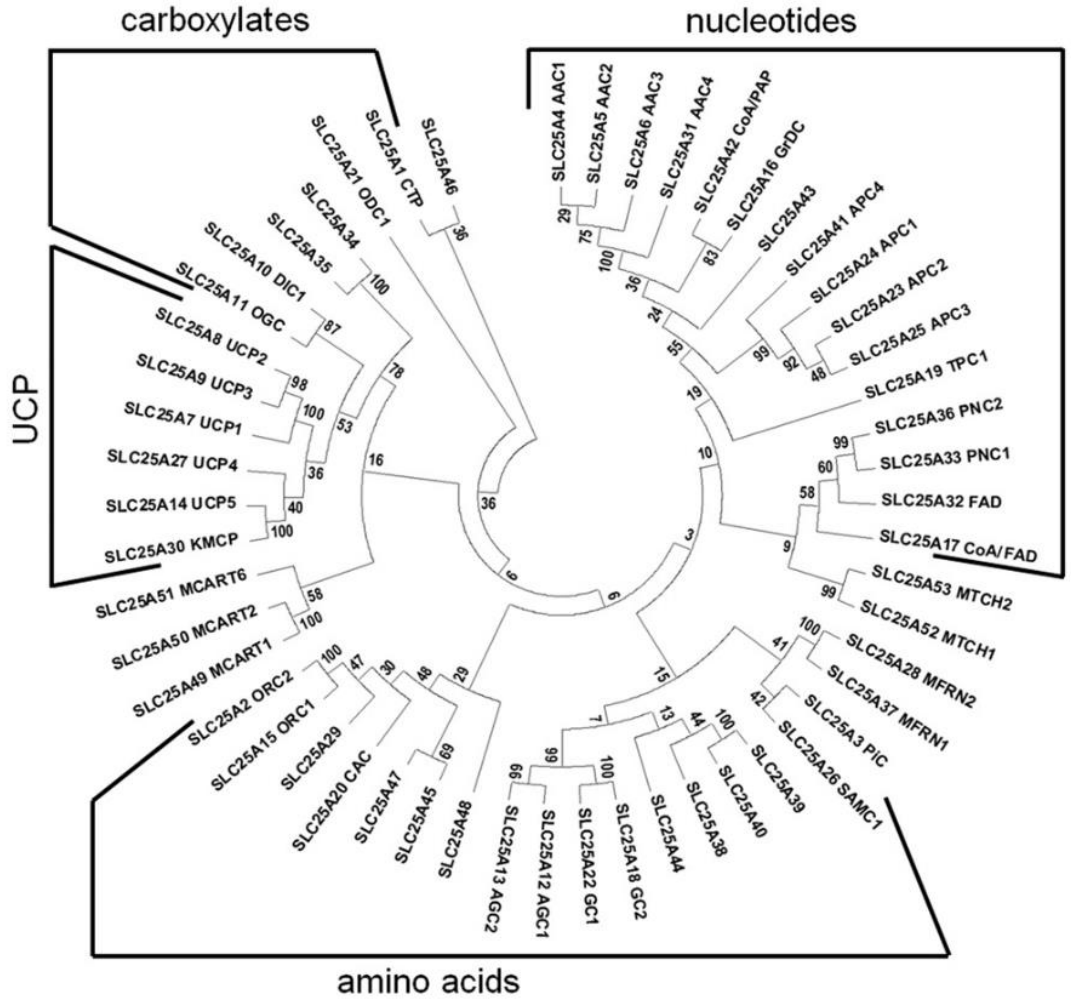


Figure 3.1. *Reproduced from [Palmieri 2013]. Phylogenetic tree of human SLC25 family mitochondrial transporters. Homologous transporters can be broken down into subgroups based on substrates (likely to be) transported, such as carboxylates and amino acids. 'Carboxylate' and 'UCP' transporters are relatively homologous and transporters from both groups have been shown to transport carboxylates.*

It should be noted that although we had a specific interest in identifying the mitochondrial serine transporter as a critical source of support for folate-mediated one-carbon reactions, we further sought to characterize these transporters as potential carriers for other amino acids. We also selected 12 SLC family transporters with the closest structural homology to known carboxylate and uncoupling protein (UCP) carriers as candidates for the 2-hydroxyglutarate transporter. Although most of these proteins are not fully characterized, some of the uncoupling protein transporters have been shown to carry carboxylate species, including aspartate, malate and 2-oxoglutarate [Vozza et al. 2014; Monné et al. 2018]. SLC25A11 and SLC25A21 were of particular interest as they were shown to transport 2-oxoglutarate, the direct metabolic precursor to 2-hydroxyglutarate [Indiveri et al. 1987; Monné et al. 2013; Palmieri et al. 2001; Fiermonte et al. 2001]. In fact, 2-hydroxyglutarate was shown to competitively inhibit 2-oxoglutarate transport into rat liver mitochondria even before the isolation and characterization of SLC25A11 or other 2-oxoglutarate transporters, evidence of interaction with the transporter's active sight [Palmieri et al. 1972]. Similarly, 2-hydroxyglutarate was found to inhibit 2-oxoglutarate transport by yeast SLC25A21 in reconstituted proteoliposomes [Palmieri et al. 2001].

We initially set out to assess function of SLC25 transporters in the MDA-MB-468 human breast cancer cell line using CRISPR-Cas9 knockouts. The rationale for this approach was that loss of transporter function could cause significant metabolic abnormalities in the knockout cell lines compared to control, which could be detected in an unbiased manner by untargeted metabolite profiling. Profound changes in metabolite abundances could be used to generate a hypothesis as to the substrate of the knocked-out transporter, which could then be confirmed by various methods, such as

stable isotope tracing and studying isolated mitochondria or purified proteins. The underlying assumption for studies of *SLC25* family CRISPR KO cell lines was that these transporters are important enough for their absence to cause metabolic perturbations, but not essential so knockout cell lines would be viable. The disruption of important transporter function may directly impact metabolite abundances (such as elevated levels of a compound that is metabolized in the mitochondria) or lead to detectable metabolic compensation mechanisms that indirectly inform on cargos for a given transporter. However, due to findings considered below, we later shifted to a simpler and more targeted discovery approach that involved study of heterologous protein expression in *Lactococcus Lactis* bacteria.

Lactococcus Lactis is a gram-positive bacteria that has been used for heterologous protein expression of functional membrane proteins. There are a number of features that make this bacterium attractive for membrane protein expression. First, membrane proteins, even eukaryotic mitochondrial transporters, have been found to fold properly and localize to the bacterial membrane, with little or no formation of inclusion bodies. The expressed proteins can be inducibly expressed by taking advantage of a nisin promoter that resides in the pNZ8048 expression plasmid. Nisin is an anti-bacterial peptide that can be administered at very low concentrations to drive protein expression, allowing for controlled levels of over-expression (this may be important for toxic proteins) [Kunji et al. 2003; Seigneurin-Berny et al. 2016]. In the past, this system has been used to express membrane proteins that were subsequently reconstituted into proteoliposomes to enable radioligand transport studies. However, there have also been reports of studies demonstrating functional membrane transport activity in intact *L. lactis* bacteria (also using radioisotopes) [Janvilisri et al. 2003;

Kunji et al. 2003; Burgess et al. 2006; Frelet-Barrand et al. 2010; Linares et al. 2010; Herzig et al. 2012; Mifsud et al. 2012; Seigneurin-Berny et al. 2016].

We reasoned that the study of SLC25 protein function in intact bacteria coupled with mass spectrometric detection of substrates would save significant time compared to studying the reconstituted proteins in liposomes and would open the possibility for more rapid study or even screening experiments. The use of heavy stable isotope-labelled substrates for transport assays and LC/MS-based detection affords a number of advantages over studies involving radioisotopes. First, with stable isotope studies radioactive contamination and safety is not a concern, which significantly reduces the hassle of working with the samples and equipment. Second, unlike radioisotopes, stable isotopes can enable reliable metabolic fate-tracing. Finally, mass spectrometry can easily distinguish between different labeled compounds and potentially enable the screening of transporters for activity with multiple compounds at one time, facilitating the comprehensive study of uncharacterized transporters. We therefore set out to establish a whole cell *Lactococcus Lactis* protocol for studying small molecule membrane protein transport using mass spectrometry, and to assess the feasibility of using *Lactococcus Lactis* to study the human SLC25 family of transporters.

MATERIALS AND METHODS

1. Reagents. LC-MS grade acetonitrile (ACN), isopropanol (IPA) and methanol (MeOH) were purchased from Fischer Scientific. High purity deionized water (ddH_2O) was filtered from Millipore (18 OMG). OmniTrace glacial acetic acid and

ammonium hydroxide were obtained from EMD Chemicals. ^{9}D-Choline, ^{6}D-2-oxoglutarate, ^{13}C-glycine, and ^{15}N-aspartate were purchased from Cambridge Isotope Laboratory.

2. CRISPR knockout cell line production.

2.1 Plasmid construction for CRISPR genome editing (eSpCas9). Double-stranded oligomers were synthesized commercially and cloned into BbsI restricted eSpCas9(1.1) vector (Addgene, Cat.# 71814). The vector was transformed into NEB Stable chemically competent *E. coli* (NEB) and positive clones were identified and DNA sequenced using standard methods. The eSpCas9 constructs used for this study were produce with the following gRNA sequences, targeting the genomic DNA near endogenous restriction enzyme cut sites:

GGTAGAGCAGGTCCCACTGA (*SLC25A38* KO1),

ACTCTTGAAGGTGGTTCGCA (*SLC25A38* KO2),

GGACACCATTGCAATACAGG (*SLC25A39*),

GCAGGTAAAATAAATAACTG (*SLC25A40*),

GAGTAACACAGTCAAATCAC (*SLC25A44*).

2.2 Generation of knockout cell lines. eSpCas9 constructs were transiently transfected into MDA-MB-468 triple-negative breast cancer cells using lipofectamine3000 (Invitrogen, CA) and selected for cell uptake based on viability in puromycin media for 3 days. Viable cells were seeded at low density (~750 cells per 100 mm dish), then individual colonies (clones) were selected and transferred to 24-well plates for clonal expansion.

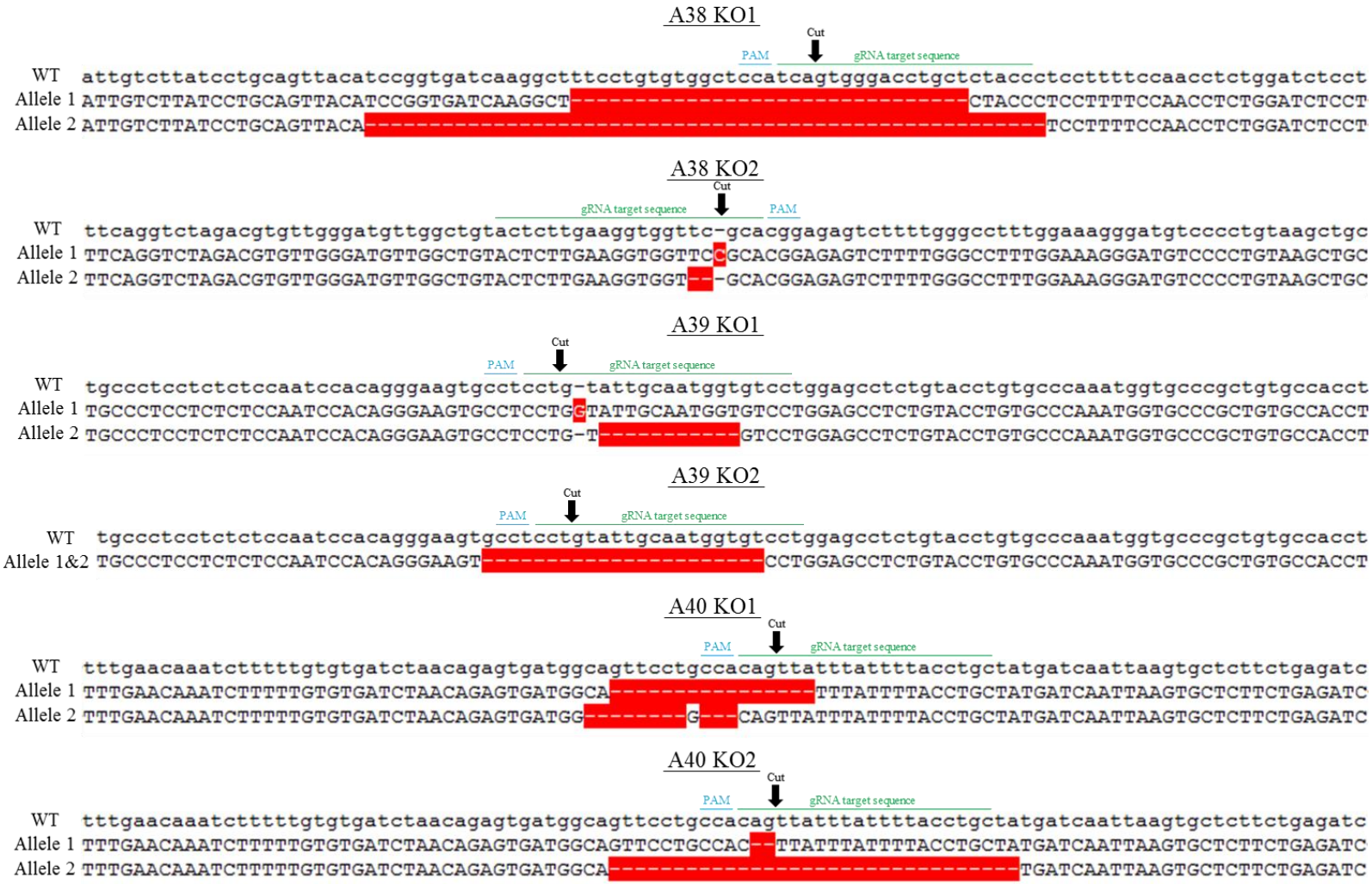
2.3 Cut-site screening. Genomic DNA was harvested from each clonal line for genotyping. Colonies were screened by PCR amplification of the target region and endonuclease restriction digestion to identify edits in the restriction site near the target CRISPR cut site (restriction resistant PCR products indicate edited genomic DNA). The PCR bands with modified restriction-sites were sequenced by TA-cloning, as follows. Purified PCR products were cloned into the pGEM-T Easy vector (Promega, WI), and transformed into NEB Stable *E.coli*. Plasmid DNA was isolated from individual bacterial colonies and sequenced using the T7(+) primer. Sufficient colonies were screened to obtain sequence information for both alleles (Figure 3.2).

2. Protein reinsertion cell line production.

2.1 Plasmid construction for gene reinsertion. Gene sequences with either N-terminal FLAG or C-terminal EGFP and the appropriate restriction sites were PCR amplified. Purified PCR products were cloned into pcDNA3.1 vector (Invitrogen) and transformed into bacteria. Positive clones were identified using standard bacterial screening and confirmed by DNA sequencing.

2.2 Generation of reinsertion cell lines. We transiently introduced the constructs into the appropriate MDA-MB-468 triple-negative breast cancer knockout cell line using lipofectamine3000 (Invitrogen, CA) and selected for neomycin resistance of cells during a 7 day treatment. We then seeded the cells at low density (~750 cells per 100 mm dish) to obtain clonal lines. Individual colonies were selected under a microscope and transferred to 24-well plates for clonal expansion. EGFP positive cell lines were confirmed under a fluorescent microscope and later imaged by confocal microscopy.

Figure 3.2. *Sequence alignment of knockout cell line alleles with wild-type sequence for two cell lines for SLC25A38, SLC25A39 and SLC25A40. Top sequences of each alignment indicate the wild-type sequence (lower case). Bottom sequences indicate sequence obtained from genotyping alleles by TA cloning (two sequences) or by directly sequencing the PCR product (one sequence). Each edited allele has an insertion or deletion that causes a frame-shift. The gRNA target sequence (green line), Protospacer adjacent motif (PAM), and cut site (black arrow) are indicated for each CRISPR design.*


A38 KO1
Cut
PAM
gRNA target sequence
WT attgtcttatcctgcagttacatccggtgatcaaggctttcctgtgtggctccatcagtgggacctgctctaccctccttttccaacctctggatctcct
Allele 1 ATTGTCTTATCCTGCAGTTACATCCGGTGATCAAGGCT----------CTACCCTCCTTTTCCAACCTCTGGATCTCCT
Allele 2 ATTGTCTTATCCTGCAGTTACA----------TCCTTTTCCAACCTCTGGATCTCCT
A38 KO2
Cut
gRNA target sequence
PAM
WT ttcaggtctagacgtgttgggatgttggctgtactcttgaaggtggttc-gcacggagagtcttttgggcctttggaaagggatgtcccctgtaagctgc
Allele 1 TTCAGGTCTAGACGTGTTGGGATGTTGGCTGTACTCTTGAAGGTGGTTCCGCACGGAGAGTCTTTTGGGCCTTTGGAAAGGGATGTCCCCTGTAAGCTGC
Allele 2 TTCAGGTCTAGACGTGTTGGGATGTTGGCTGTACTCTTGAAGGTGGT--GCACGGAGAGTCTTTTGGGCCTTTGGAAAGGGATGTCCCCTGTAAGCTGC
A39 KO1
Cut
PAM
gRNA target sequence
WT tgccctcctctctccaatccacagggaagtgcctcctg-tattgcaatggtgtcctggagcctctgtacctgtgcccaaatggtgcccgctgtgccacct
Allele 1 TGCCCTCCTCTCTCCAATCCACAGGGAAGTGCCTCCTGGTATTGCAATGGTGTCCTGGAGCCTCTGTACCTGTGCCCAAATGGTGCCCGCTGTGCCACCT
Allele 2 TGCCCTCCTCTCTCCAATCCACAGGGAAGTGCCTCCTG-T----------GTCCTGGAGCCTCTGTACCTGTGCCCAAATGGTGCCCGCTGTGCCACCT
A39 KO2
Cut
PAM
gRNA target sequence
WT tgccctcctctctccaatccacagggaagtgcctcctgtattgcaatggtgtcctggagcctctgtacctgtgcccaaatggtgcccgctgtgccacct
Allele 1&2 TGCCCTCCTCTCTCCAATCCACAGGGAAGT----------CCTGGAGCCTCTGTACCTGTGCCCAAATGGTGCCCGCTGTGCCACCT
A40 KO1
Cut
PAM
gRNA target sequence
WT tttgaacaaatctttttgtgtgatctaacagagtgatggcagttcctgccacagttatttatttacctgctatgatcaattaagtgctcttctgagatc
Allele 1 TTTGAACAAATCTTTTTGTGTGATCTAACAGAGTGATGGCA----------TTTATTTTACCTGCTATGATCAATTAAGTGCTCTTCTGAGATC
Allele 2 TTTGAACAAATCTTTTTGTGTGATCTAACAGAGTGATGG--------G---CAGTTATTTATTTTACCTGCTATGATCAATTAAGTGCTCTTCTGAGATC
A40 KO2
Cut
PAM
gRNA target sequence
WT tttgaacaaatctttttgtgtgatctaacagagtgatggcagttcctgccacagttatttatttacctgctatgatcaattaagtgctcttctgagatc
Allele 1 TTTGAACAAATCTTTTTGTGTGATCTAACAGAGTGATGGCAGTTCCTGCCAC--TTATTTATTTTACCTGCTATGATCAATTAAGTGCTCTTCTGAGATC
Allele 2 TTTGAACAAATCTTTTTGTGTGATCTAACAGAGTGATGGCA----------TGATCAATTAAGTGCTCTTCTGAGATC

3. MDA-MB-468 derived cell line metabolite profiling.

3.1 Cell culture. 150,000 cells/well were plated in 6-well plates and cultured in DMEM medium containing 25 mM glucose, 10% FBS, and 1% HEPES for 48 hours. Cells were harvested at 80% confluency and extracted for LC/MS metabolomic analysis.

3.2 Metabolite extraction. One plate was extracted at a time. Cells in each individual well (samples) were harvested one-by-one, rapidly washed twice with ice-cold PBS, then twice with ice-cold water, followed by addition of -70ºC 80:20 methanol:water (LC-MS grade methanol, Fisher Scientific). Cells were scraped and transferred to an Eppendorf tube. For quantitative recovery of metabolite extracts, an additional 400uL of 80% methanol was added to cell culture wells 2 additional times and pooled with the initial extract. Cells suspensions/extracts were vortexed for 15 seconds before being placed on dry ice and stored overnight at -80 ºC. On the following day, extracts were thawed on ice, vortexed, and then centrifuged at 16,000 RPM for 10 min to pellet insoluble material (protein/nucleotides). Supernatants were transferred to clean tubes, dried in a Vacufuge (Eppendorf), and stored at -80ºC until metabolomic analysis by LC/MS. The methanol-insoluble protein pellet was solubilized in 0.2 M NaOH at 95ºC for 20 min and protein content was quantified using a BioRad DC assay and findings were used to normalize the volume of extracts for metabolite profiling. On the day of metabolite analysis, dried cell extracts were reconstituted in 70% acetonitrile at a relative protein concentration of 3 µg/ml, and 4 µl of this reconstituted extract was injected for LC/MS-based untargeted metabolite profiling. Samples were analyzed on an LC-QTOF system in an alternating sequence (i.e, test, control, test, control etc.).

3.3 Untargeted metabolite profiling by LC/MS. Cell extracts were analyzed by LC/MS as described previously [Chen et al. 2012; Ismailoglu et al. 2014], using a platform comprised of an Agilent Model 1290 Infinity II liquid chromatography system coupled to an Agilent 6550 iFunnel time-of-flight MS analyzer (Agilent Technologies). Chromatography of metabolites utilized aqueous normal phase (ANP) chromatography on a Diamond Hydride column (Microsolv). Mobile phases consisted of: (A) 50% isopropanol, containing 0.025% acetic acid, and (B) 90% acetonitrile containing 5 mM ammonium acetate. To eliminate the interference of metal ions on chromatographic peak integrity and electrospray ionization, EDTA was added to the mobile phase at a final concentration of 6 µM. The following gradient was applied: 0-1.0 min, 99% B; 1.0-15.0 min, to 20% B; 15.0 to 29.0, 0% B; 29.1 to 37min, 99% B. Raw data were analyzed using MassHunter Profinder 8.0 and Mass Profiler Professional (MPP) 14.9.1 software (Agilent Technologies).

3.4 Metabolite structure specification. To ascertain the identities of unknown compounds, LC/MS data was searched against an in-house annotated metabolite database created using MassHunter PCDL manager 7.0 (Agilent Technologies), based on monoisotopic neutral masses (<5 ppm mass accuracy) and chromatographic retention times. A molecular formula generator (MFG) algorithm in MPP was used to generate and score empirical molecular formulae, based on a weighted consideration of monoisotopic mass accuracy, isotope abundance ratios, and spacing between isotope peaks. A tentative compound ID was assigned when PCDL database and MFG scores concurred for a given candidate molecule. Tentatively assigned molecules were verified based on a match of LC retention times and/or MS/MS fragmentation spectra for pure molecule standards.

3.5 Stable isotope tracing. Deuterated choline (6D) was added to choline free DMEM and added to cells at specific times prior to harvest (1, 5 and 24 hours). Metabolites were harvested and LC-MS metabolomics was performed as described above. A database of predicted m/z values for deuterium-labeled choline-derived metabolites was produced based on retention times from our in-house database and used in MassProfinder 8.0 to quantify metabolite abundance data.

4. Next-Generation Amplicon Sequencing of *SLC25* knockout cell line mRNAs. mRNA was reverse transcribed using random hexamer primers to produce cDNA from two knockout cell lines per protein (SLC25A38, SLC25A39, SLC25A40). cDNA were PCR amplified using primers spanning the CRISPR target site (~200bp on either side) for the respective gene. Primers were designed based on exon sequences (processed mRNA) and spanned known intron splice sites so that amplified DNA at the predicted size (~400bp) could not be due to genomic DNA contamination. PCR products were sequenced by Next-Generation Amplicon Sequencing (GENEWIZ). Alignment to the wild-type reference sequence for each protein (based on the most common NCBI transcript sequence) was performed using the CRISPRESSO 2.0 web tool.

5. Transporter overexpression in *Lactococcus Lactis*.

5.1 Electrocompetent *Lactococcus Lactis* production. Overnight cultures (~8 mL) of *Lactococcus Lactis* grown in M17 were diluted into 200 mL SGGM17 (M17 broth containing 0.5M sucrose, 0.5% glucose and 2.5%) and grown to an O.D. of 0.3-0.4 at 600 nm. Cells were pelleted by centrifugation at 6,000 x g at 4°C for 8 min.

Cells were resuspended in 100 mL of 0.5M sucrose/10% glycerol solution and pelleted again. Cells were then resuspended in 100mL 0.5M sucrose/10% glycerol with 50mM EDTA pH 8.0 and placed on ice for 15 minutes and pelleted at 6,500 x g for 12 min, then twice resuspended/pelleted in 125mL 0.5M sucrose/10% glycerol solution. Finally, cells pellets were resuspended in 1 mL of 0.5M sucrose/10% glycerol, flash frozen in 40uL aliquots, and stored at -80°C until analysis.

5.2 Plasmid construction for gene reinsertion. Gene sequences were PCR amplified from endogenous homo sapiens sequences, for *SLC25A4*0 and *SLC25A44*, or from *Lactococcus Lactis* 'codon context' optimized sequences for *SLC25A38*, *SLC25A39*, carboxylate transporters (optimized using Codon Optimization On-Line (COOL) tool, gene fragments from GenScript)) with the appropriate restriction sites and with or without a C-terminus FLAG tag sequence. Generally, BbsI/KpnI endonuclease sites were used for cloning into the NcoI/KpnI restricted pnZ8048 vector backbone. Vectors were precipitated by addition of sodium acetate pH 5.2 and 100% EtOH, followed by incubation at -20°C for at least 1 hour. DNA was pelleted by centrifugation, washed 3 times with 70% EtOH, dried (not to completion) and dissolved in 5 µL water. A 1 µl volume of the resulting vector solution was electroporated into electrocompetent *Lactococcus Lactis* using a 0.2 cm cuvette (2,000 V, 25 µF, 200 Ω), incubated in recovery media (M17 broth containing 0.5 M sucrose 0.5% glucose, 20 mM $MgCl_2$ and 2 mM $CaCl_2$) at 30°C for 1.5 hours, plated on M17 agar plates containing 0.5% glucose and 10 ug/mL chloramphenicol and incubated at 30°C for 2 days. Positive clones were identified using standard bacterial screening (cells were incubated in a 'resuspension buffer' (10 mM Tris pH 8.0, 50 mM NaCl, 10 mM EDTA, 20% sucrose, 10 mg/mL lysozyme (added fresh)) for 20 minutes at 55°C,

before continuing at the lysis buffer step using a DNA Miniprep Kit (Sigma). Construct sequences were confirmed by DNA sequencing.

5.3 Transporter overexpression. Overnight cultures of *Lactococcus Lactis* that express pnZ8048 vector with an incorporated a gene of interest, were diluted 10-fold in M17 broth containing 0.5% glucose and 5 ug/mL chloramphenicol. Cells were grown at 30°C to an OD of 0.6 at 600nm (~1 hour). Protein expression was induced by addition of 1 ng/µL nisin (Sigma) and bacteria were allowed to grow for 3 h until use.

5.4 Western blot and dot-blot for FLAG tagged transporters. 2 mL of M17 broth was inoculated with *Lactococcus Lactis* and transporter expression was induced as described above. Bacteria were harvested by centrifugation, resuspended in a 50 mM Tris pH 8.0, 200 mM NaCl, 5 mg/mL lysozyme (added fresh) solution and incubated at 55°C for 20 min. Cells were pelleted and the supernatant removed. The pellet was suspended in an 8 M urea, 50 mM Tris-HCl pH 6.8, 10 mM EDTA, 50 mM DTT solution and incubated at room temperature for 10 min. The debris was pelleted and 10 µL of the resulting supernatant was added to 10 µL 10% glycerol/bromophenol blue loading dye. Samples were loaded on an 8-16% tris-glycine gel and run at 120V for 1.5 hr. Protein was transferred onto an activated PVDF membrane, blocked with 2.5% milk in TBS-T, then treated with monoclonal anti-flag-M2 antibody (1:1000 dilution; Sigma) overnight. For the dot blots, PVDF membranes were dried, blocked and then treated the same as the western blots. Membranes were washed in TBS-T, treated for 1 hour with secondary anti-mouse antibody, washed and then imaged by chemilumiescence after treatment with the SuperSignal ‘Pico’ or ‘Femto’ chemiluminescent reagent (ThermoFisher).

6. Metabolite profiling of *Lactococcus Lactis*.

6.1 Initial metabolite extraction protocol. *Lactococcus Lactis* cell lines were cultured and transporter expression induced with nisin as described above (2 mL M17 broth inoculated with 200 uL overnight culture). Cells were isolated by centrifugation and at 12,000 x g for 1 min and resuspended in phosphate buffered saline (PBS) pH 7.4. Cell were pelleted again at 18,000 x g. Wash steps were repeated 2 times in iced PBS and then ddH_2O. After the final wash, 1 mL -70°C 80% MeOH was added to extract metabolites and samples were frozen on dry ice. Samples were subsequently subjected to bead-beating using a TissueLyser II bead beater (Qiagen) and stainless steel beads for 2.5 minutes, then centrifuged at 16,000 x g for 5 minutes and placed on dry ice. This process was continued until the samples were bead-beaten 4 times. Samples were then placed at -80°C overnight. After thawing the following day, samples were spun a final time and supernatants were transferred to new Eppendorf tubes followed by drying in a vacufuge and storage at -80°C until analysis. Remaining protein in cell pellets (after metabolite extraction) was dried under vacuum in a vacufuge, solubilized in 0.2 N NaOH at 95°C, then protein was quantified using a Bio-Rad Bradford protein assay. On the day of analysis, metabolites were reconstituted to the equivalent of 1 µL 70% ACN per 3 µg protein to normalize for sample differences in bacteria quantity.

6.2 Improved metabolite extraction protocol. *Lactococcus Lactis* cell lines were cultured and transporter expression was induced as described above (in 35 mL M17 broth/chloramphenicol inoculated with 3.5 mL overnight culture, 2 times per cell line for 12 samples, 4 times for 24 samples). Cells were collected at 6,000 x g for 10 min at 4°C and washed with ice-cold 50 mM potassium phosphate buffer pH 7.0.

Pellets were respun and washed a second time. After a final spin (20 min), pellets were resuspended in 7 mL potassium phosphate buffer (14 mL for 24 samples) and stored on ice. Aliquots of bacterial suspensions (500 µL) were added to 5 mL culture tubes. The test transport substrate of interest was added to the incubation solution and the bacteria were maintained in a 30°C water bath for the desired time interval. Samples were simultaneously harvested on glass fiber filter paper (GF/B, Brandel) and rapidly washed with water 4 times (3 seconds/wash) using a Brandel M-24T cell harvester. Filter discs were punched out of the glass fiber sheet by hand, and added to 2 mL round bottom Eppendorf tubes on dry ice. A 1.8 ml volume of methanol:ddH_2O (80:20; -70°C) was added to each sample. Samples were bead-beaten for 2.5 minutes, centrifuged at 16,000 x g for 10 min at 4°C and then refrozen on dry ice. This extraction procedure was repeated until each sample was beaten 4 times in total (with a final 5 min extraction). Samples were stored overnight at -80°C. On the following day, samples were bead-beaten (5 min, then 10 min), centrifuged and frozen two additional times. Finally, samples were centrifuged at 16000 x g for 30 min at 15°C. Small squares of no lint cotton (Agilent Technologies) were pushed into p200 pipette tips to create a filter. Supernatant (750 µL) was pushed through these filters into a clean Eppendorf tube using a p1000 pipette in order to remove glass fibers that were too fine to spin down. Samples were dried in a vacuum centrifuge and stored at -80°C overnight until analysis. On the day of analysis, potential transported substrates were suspended in 70% ACN (at specified volumes; 40 µl, 80 µl or 300 µl, depending on compound of interest).

RESULTS

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A38* knockout cell line reveals modest metabolite changes compared to parental control. Untargeted metabolite profiling was performed on an MDA-MB-468 *SLC25A38* knockout clonal cell line, compared to the parental control cell line (n=12). Cells were grown for 48h, followed by metabolite extraction and untargeted profiling as described above. Unbiased data was acquired and analysis was performed in a targeted manner (with reference to an in-house retention time and mass-annotated database) and also considering structurally-unidentified compounds. Notably, no unidentified compounds were found to exhibit sufficient differential expression and significance to warrant the effort of structure elucidation. However, there were multiple statistically significant identified compounds, although also showing only modest metabolite changes in KO cells compared to control (Table 3.1). It should be noted that these identities are tentative and confident assignment of some would require MS/MS fragmentation for definitive assignment by comparison with reference standards. A feature matching the retention time and m/z of an N-acetylglycine acetic acid adduct (acetic acid is present in the mobile phase) was the most significant change, with a nearly 2-fold decrease in *SLC25A38* KO cells (Figure 3.3). Due to the fact that SLC25A38 was reported to transport glycine and increases in other acetylated amino acids were concomitantly observed, N-acetylglycine was a focus of follow-up experiments.

Table 3.1. *Statistically significant, differentially abundant features of an unbiased metabolomics experiment of SLC25A38 KO cells compared to parental control cells. Tentative molecular identities were assigned with reference to an in-house database. N-acetylglycine, the most significant compound of interest is highlighted. P-values are Benjamini-Hochberg false discovery rate corrected. Features from each ionization mode are corrected independently).*

Compound	p value	Fold change	Mode
N-Acetylglycine	3.36E-09	-1.95	Neg
3-Phosphoglyceroinositol	3.10E-07	-1.40	Pos
Valerylcarnitine	3.10E-07	1.54	Pos
D-Sorbitol	8.63E-07	1.93	Neg
Galacturonic Acid	8.91E-06	2.10	Pos
Acetylcysteine	1.02E-05	1.50	Neg
Glucaric acid	1.94E-04	1.49	Neg
Shikimic Acid	5.99E-04	1.43	Neg
LysoPC(18:0)	0.0024	-1.61	Pos
N-Acetylleucine	0.0027	1.57	Neg
N-Acetylphenylalanine	0.0030	2.67	Pos
O-Acetylserine	0.0040	1.90	Pos
Gluconic Lactone	0.0333	1.50	Neg
2-Acetamido-2-Deoxy-beta-glucosylamine	0.0366	-1.48	Neg
Pterin	0.0385	1.76	Neg
Shikimic Acid	0.0388	1.47	Pos

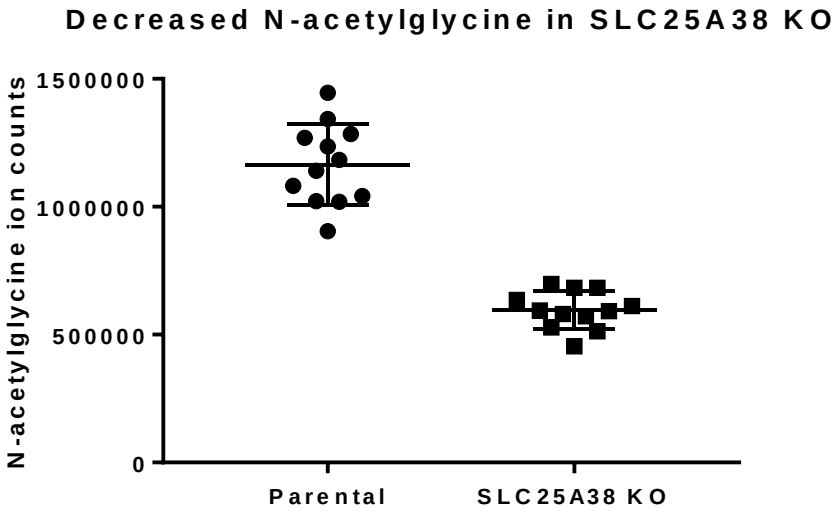


Figure 3.3. *Comparison of the abundance of N-acetylglycine in SLC25A38 KO cells compared to control cells. A38 KO cells have significantly reduce abundance, possibly due to reduced glycine transport into the mitochondria.*

Metabolite profiling validation using a second *SLC25A38* knockout cell line shows a less significant decrease in N-acetylglycine. Untargeted metabolite profiling was performed on a second MDA-MB-468 *SLC25A38* knockout clonal cell line produced with different guide RNA (in order to control for off target effects), using the parental cell line as a control (n=12). Cells were grown for 48 h, followed by metabolite extraction and untargeted profiling as described above. Unbiased data analysis was performed in a targeted manner, using an in-house retention time and mass-annotated database as well as in an untargeted mode. As shown in Figure 3.4, abundance of the N-acetylglycine feature was found to be only modestly reduced. Additionally, other features of statistical significance showed no overlap with those from the first knockout cell line, suggesting that observed differences arose primarily as clonal peculiarities, off target CRISPR effects or were cell culture condition dependent (Table 3.2). No shared features were promising enough to follow up on.

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A40* knockout cell line reveals modest metabolite changes compared to parental control. Untargeted metabolite profiling was performed on an MDA-MB-468 *SLC25A40* knockout clonal cell line, compared to the parental control cell line (n=12). Cells were grown for 48h, followed by metabolite extraction and untargeted profiling as described above. Unbiased data was acquired and analysis was performed in a targeted manner (with reference to an in-house retention time and mass-annotated database) and also considering structurally-unidentified compounds. Notably, no unidentified compounds were found to exhibit sufficient differential expression and significance to warrant the effort of structure elucidation.

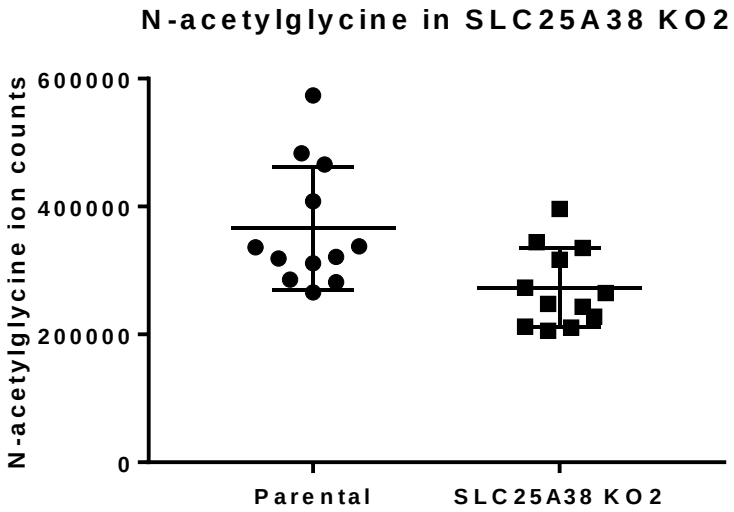


Figure 3.4. *Comparison of the abundance of N-acetylglycine in SLC25A38 KO cells compared to control cells. A38 KO cells have mildly reduced abundance of unconvincing importance to the study of the knockout protein.*

Table 3.2. *Statistically significant, differentially abundant features of an unbiased metabolomics experiment of a second SLC25A38 KO cell line compared to parental control cells. Tentative identities were assigned based on an in-house database. No interesting compounds shared between the two knockout cell lines were identified. P values are Benjamini-Hochberg false discovery rate corrected (features from each ionization mode corrected separately).*

Compound	p value	Fold change	Mode
5'-Deoxyadenosine	7.67E-07	-1.54	Pos
Homocystine	3.26E-05	-1.76	Neg
LysoPC(18:0)	1.66E-04	1.55	Neg
Succinoadenosine	3.69E-04	-1.54	Neg
Gluconic Lactone	3.69E-04	-1.46	Neg
Phenylacetic Acid	3.69E-04	-1.47	Neg
LysoPC(16:0)	5.48E-04	1.49	Neg
3-Aminoisobutyric acid	6.39E-04	-1.68	Neg
LysoPC(18:2)	0.0027	1.46	Neg
Triiodothyronine	0.0038	1.93	Neg
4-Aminobutanoic Acid (GABA)	0.0038	-1.42	Pos
S-Adenosylhomocysteine	0.0053	-2.03	Neg
N-Acetylglucosamine-1-phosphate	0.0062	-1.43	Neg
Succinoadenosine	0.0070	-1.60	Pos
Phosphatidylethanolamine(32:0)	0.0079	1.49	Neg
Urea	0.0086	1.41	Neg
GMP	0.0242	-1.42	Neg
Acetyl-CoA	0.0374	-1.55	Neg
Uric acid	0.0389	-1.51	Neg
S-Adenosylmethionine (SAM)	0.0418	-1.69	Pos
Pterin	0.0419	-1.83	Pos

However, there were multiple statistically significant identified compounds, although also showing only modest metabolite changes in KO cells compared to control (Table 3.3). It should be noted that these identities are tentative and confident assignment of some would require MS/MS fragmentation for definitive assignment by comparison with reference standards. A feature matching the retention time and m/z of betaine was one of the most significant differentially abundant compounds, with a 1.73 fold decrease in *SLC25A40* KO cells (Figure 3.5). As a compound produced exclusively in the mitochondria from a choline precursor, this generated the hypothesis that SLC25A40 could be a mitochondrial choline transporter.

MS/MS fragmentation confirms betaine identity. MS/MS fragmentation was performed on the 118.08 m/z (positive mode) feature, the putative betaine compound, using collision energies of 0, 10, 20 and 40 (Figure 3.6). Comparison to the Metlin fragmentation database confirmed the compounds identity as betaine (Figure 3.7).

Heavy isotope labeled (9 deuterated) choline tracing confirms decreased choline-derived betaine in the *SLC25A40* knockout cell line. Mass spectrometry based stable isotope tracing was performed using deuterated choline (9 deuterium) in *SLC25A40* KO cells and parental control cells.

Cell culture media was exchanged at 24h, 5h or 1h time points before metabolite extraction for media containing 9D choline (n=6). Metabolite abundances for betaine, choline and other choline-derived compounds were analyzed (Figure 3.8A-F). Reduced choline-derived betaine (deuterated) was observed in *SLC25A40* KO cells (Figure 3.8B), further indicating that choline mitochondrial transport is reduced in these cells.

Table 3.3. *Statistically significant, differentially abundant features of an unbiased metabolomics experiment of SLC25A40 KO cells compared to parental control cells. Tentative identities were assigned based on an in-house database. Betaine, the most significant and primary compound of interest is highlighted. P values are Benjamini-Hochberg false discovery rate corrected (features from each ionization mode corrected separately).*

Compound	p value	Fold change	Mode
Kynurenine	1.00E-07	-1.49	Neg
Betaine	5.28E-08	-1.73	Pos
D-Sorbitol	1.07E-05	1.71	Neg
Acetylputrescine	3.39E-04	1.44	Pos
Valerylcarnitine	3.39E-04	-1.45	Pos
Biotin	9.19E-04	-1.43	Neg
Cotinine	0.009782	1.54	Pos
1-Oleoylglycerol	0.01623	-1.42	Pos

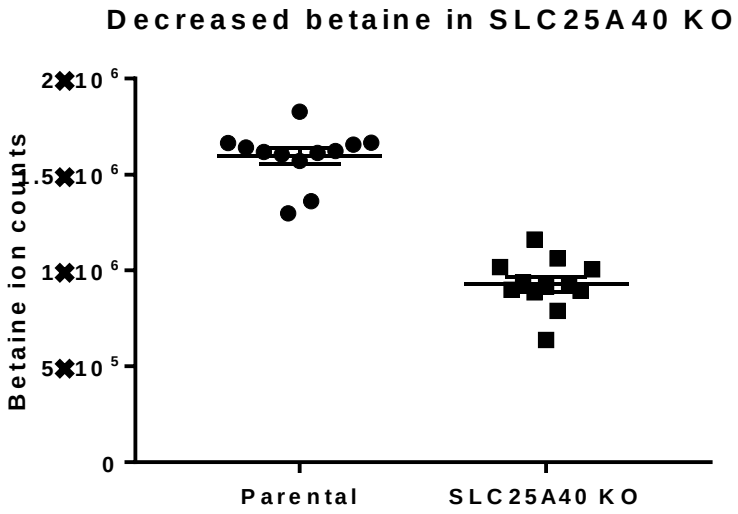
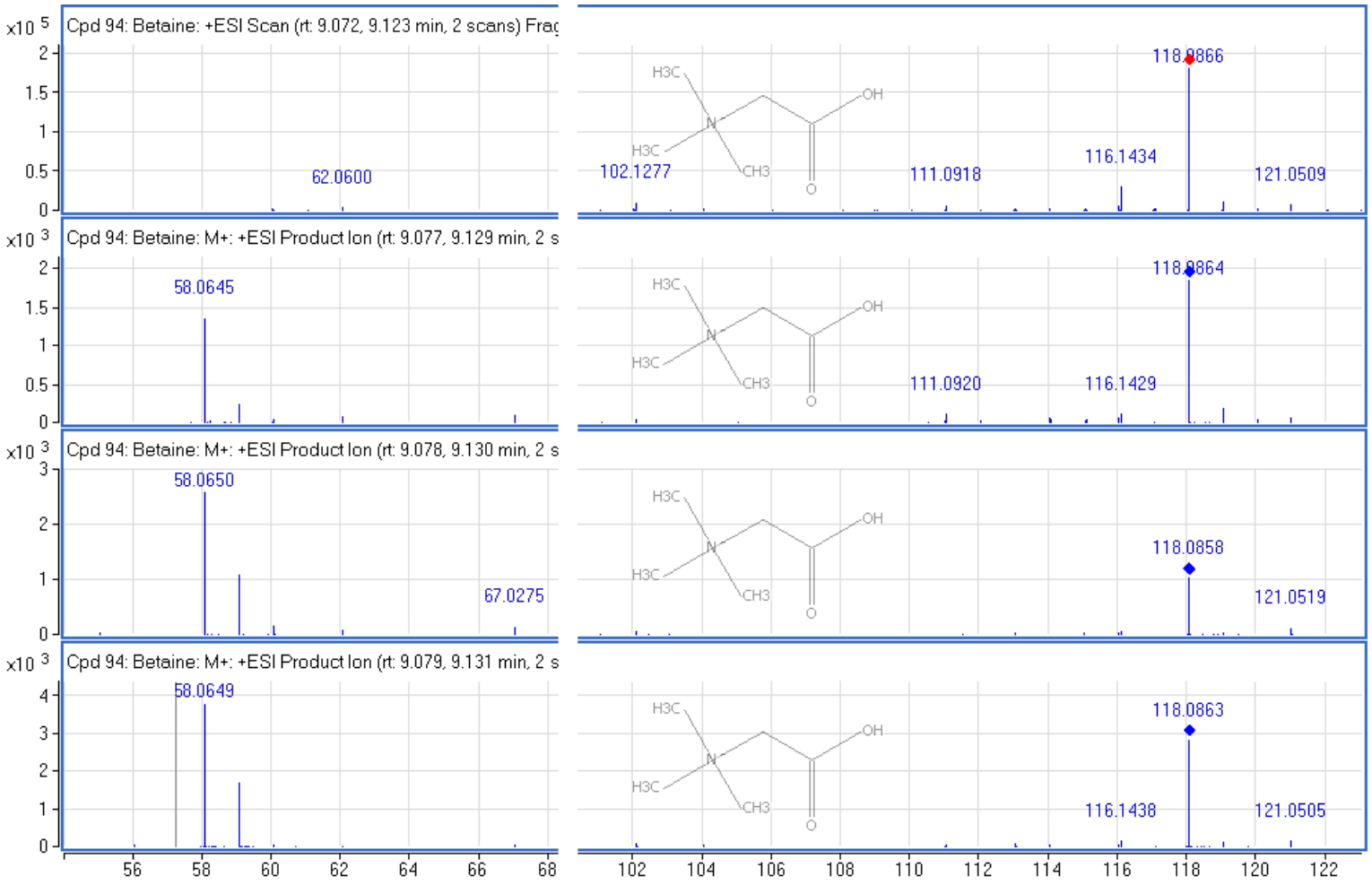


Figure 3.5. *Comparison of the abundance of betaine in SLC25A40 KO cells compared to control cells. A40 KO cells have significantly reduce abundance, possibly due to reduced choline transport into the mitochondria.*

Figure 3.6. *MS/MS fragmentation spectra of m/z 118.08 at collision energies of 0 (top panel), 40 (2nd panel), 20 (3rd panel) and 10 (bottom panel). Images is cropped in the middle (where no significant peaks are present) for visibility.*


Cpd 94: Betaine: +ESI Scan (rt: 9.072, 9.123 min, 2 scans) Frag
x10 5
62.0600
102.1277
111.0918
116.1434
118.0866
121.0509
Cpd 94: Betaine: M+: +ESI Product Ion (rt: 9.077, 9.129 min, 2 s
x10 3
58.0645
111.0920
116.1429
118.0864
Cpd 94: Betaine: M+: +ESI Product Ion (rt: 9.078, 9.130 min, 2 s
x10 3
58.0650
67.0275
118.0858
121.0519
Cpd 94: Betaine: M+: +ESI Product Ion (rt: 9.079, 9.131 min, 2 s
x10 3
58.0649
116.1438
118.0863
121.0505
H3C
N
OH
CH3
O

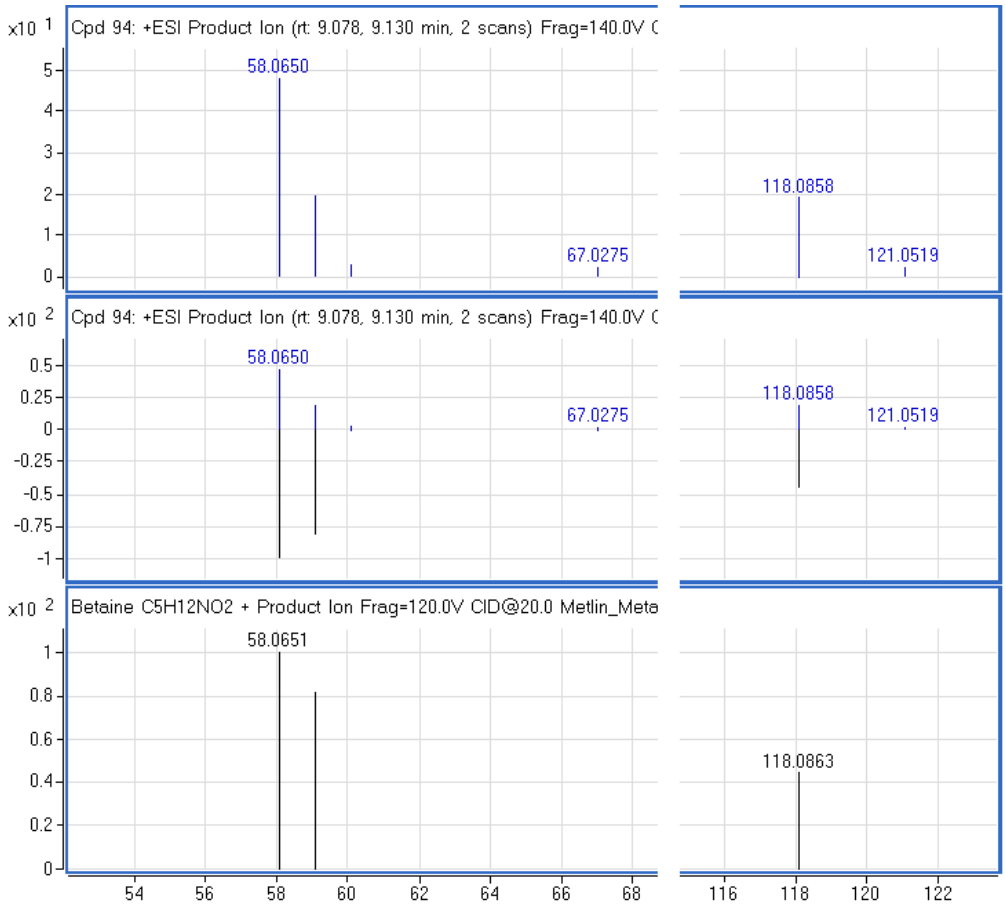
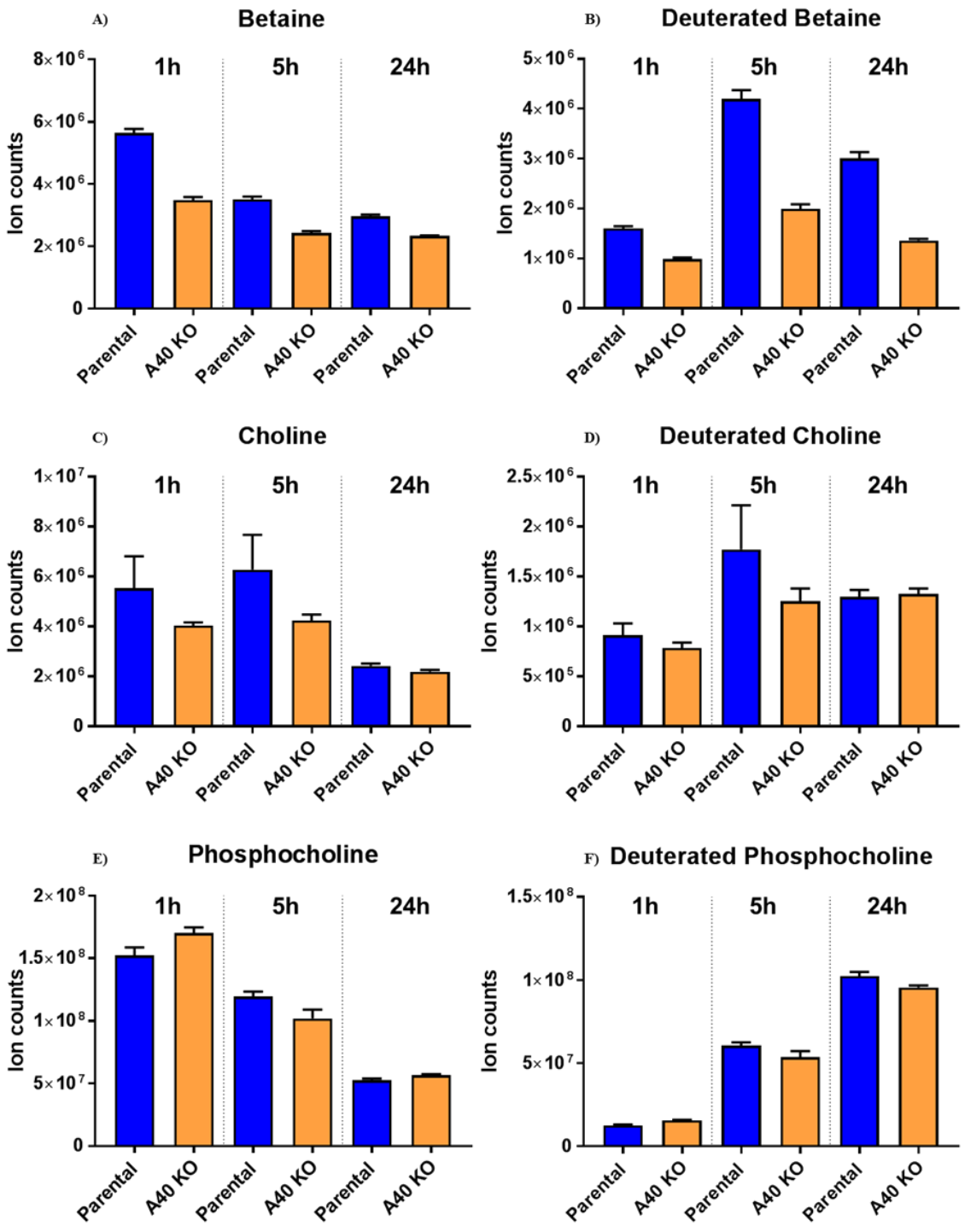


Figure 3.7. *Comparison of the MS/MS fragmentation of m/z 118.08 (pos, bottom panel) to the Metlin fragmentation database for betaine (top panel) confirms the compound's identity as betaine. The middle panel shows the direct comparison of both spectra.*

Figure 3.8. *Choline tracing reveals reduced choline-derived betaine in SLC25A40 KO cells. A) Unlabeled betaine abundance is initially different between the cell lines. At later time points the abundances even out, as betaine is no longer being derived from unlabeled choline. B) Choline-derived betaine is significantly reduced in A40 KO cells at all time points. C), E) Choline and phosphocholine abundances are the same at all time point. D) Total cellular deuterated choline uptake and F) extramitochondrial produced metabolites, such as phosphocline, are unchanged, implicating a mitochondria related defect in the reduced betaine production (enhanced betaine degeration could aslo be possible).*


A)
Betaine
B)
Deuterated Betaine
C)
Choline
D)
Deuterated Choline
E)
Phosphocholine
F)
Deuterated Phosphocholine
1h
5h
24h
Ion counts
Parental
A40 KO

Importantly, overall cellular choline uptake and choline-derived compounds produced outside the mitochondria, such as phosphocholine were unchanged (Figure 3.8D, F).

Choline tracing of additional *SLC25A40* knockout cell lines shows unchanged choline-derived betaine abundance. 24h choline tracing was performed on a second, third and fourth MDA-MB-468 *SLC25A40* knockout clonal cell line produced with the same guide RNA (this approach does not control for CRISPR off target effects, but as a first pass should arguably be more likely to be similar to the initial knockout cell line) and on the parental cell line as a control (n=6). Unfortunately, the original A40 KO cell line did not grow properly when plating for this experiment and is therefore not included in the comparison. The abundance of choline-derived betaine was not found to be affected in any of the other cell lines (Figure 3.9). Additionally, other features of statistical significance showed no overlap with those from the first knockout cell line, indicating again that differences were merely clonal peculiarities, cell culture condition dependent, or off-target CRISPR effects.

Untargeted metabolite profiling of an MDA-MB-468 *SLC25A39* knockout cell line reveals highly reproducible metabolite changes compared to parental control. Untargeted metabolite profiling was performed on an MDA-MB-468 *SLC25A39* knockout clonal cell line, compared to the parental control cell line (n=12). Cells were grown for 48h, followed by metabolite extraction and untargeted profiling as described above. Unbiased data was acquired and analysis was performed in a targeted manner (with reference to an in-house retention time and mass-annotated database) and also considering structurally-unidentified compounds.

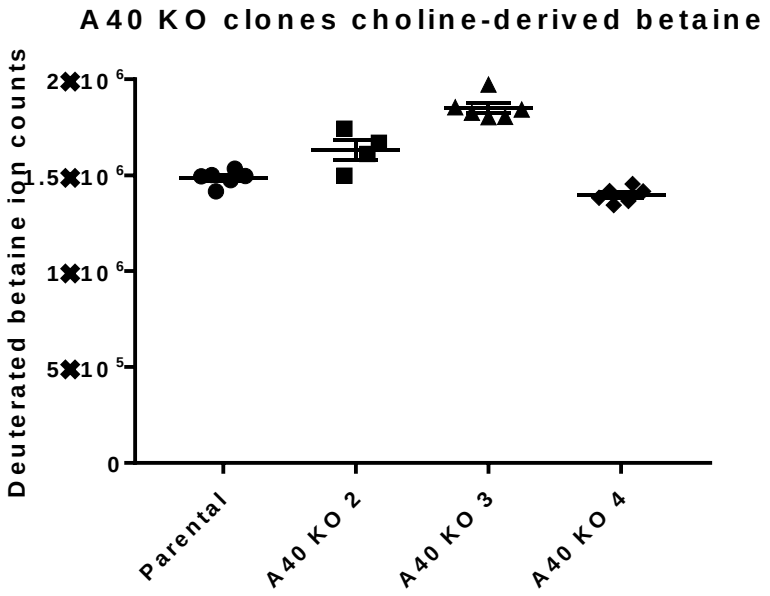


Figure 3.9. *Comparison of the abundance of deuterated betaine in SLC25A40 KO cells compared to control cells. The three other A40 KO clonal cell lines have unchanged betaine, indicating that the above observations were not SLC25A40 related.*

A number of unidentified compounds were differentially abundant, but tentative IDs could not be assigned and, given the clonal nature of changes observed in other experiments, structural elucidation was not pursued. There were a number of statistically significant identified compounds showing >2 fold change in KO cells compared to control (Table 3.4). It should be noted that these identities are tentative and confident assignment of some would require MS/MS fragmentation for definitive assignment by comparison with reference standards. A feature matching the retention time and m/z of guanidinoacetate was one of the most significant changes, with a 2.7-fold decrease in *SLC25A39* KO cells (Figure 3.10A, B). Guanidinoacetate is a compound produced in the mitochondria from the precursors glycine and arginine and this suggested that A39 may serve to transport either of these compounds. Due to the fact that A39 is homologous A38, a presumed glycine transporter, this compound was the focus of follow-up experiments.

This experiment was repeated (n=6) with highly similar results (Table 3.4, Figure 3.10A, B), confirming these metabolite differences in this cell line and highlighting the robustness of the metabolite extraction protocol.

Table 3.4. *Statistically significant, differentially abundant features of two unbiased metabolomics experiment of SLC25A39 KO cells compared to parental control cells. Tentative identities were assigned based on an in-house database. guanidinoacetate, one of the most significant and primary compound of interest is highlighted. P values are Benjamini-Hochberg false discovery rate corrected. Only positive mode features are shown for clarity. On the left are the results of the original experiment. On the right, the repeat experiment shows strikingly similar results for these compounds.*

Original:			**Repeat:**		
Compound	p value	**Fold Change**	Compound	p value	**Fold Change**
2,3-Diaminopropionic acid	3.20E-05	**-1.34**	2,3-Diaminopropionic acid	0.0106	**-1.38**
2-Oxoadipic Acid	4.89E-07	**-1.61**	2-Oxoadipic Acid	0.0220	**-1.50**
3-Aminoisobutyric acid	1.16E-05	**-1.38**	3-Aminoisobutyric acid	0.0204	**-1.56**
4-Guanidinobutyric Acid	4.09E-05	**-1.41**	4-Guanidinobutyric Acid	0.0021	**-1.80**
Ascorbic acid	2.77E-08	**1.81**	Ascorbic acid	2.00E-04	**1.91**
Aspartic Acid	3.18E-04	**-1.38**	Aspartic Acid	0.0220	**-1.21**
Butyrylcarnitine	8.57E-05	**-1.35**	Butyrylcarnitine	3.73E-05	**-1.91**
Carnitine	9.81E-10	**-1.51**	Carnitine	0.013523	**-1.61**
Cytidine 5'-diphosphocholine (CDP-choline)	6.63E-05	**1.97**	Cytidine 5'-diphosphocholine (CDP-choline)	3.73E-05	**6.63**
Fumaric Acid	3.18E-04	**-1.37**	Fumaric Acid	0.025114	**-1.20**
Galacturonic Acid	8.40E-06	**1.65**	Galacturonic Acid	2.00E-04	**1.55**
Glucuronic acid	4.56E-07	**1.72**	Glucuronic acid	0.003453	**1.57**
Guanidinoacetic Acid	1.69E-11	**-2.71**	Guanidinoacetic Acid	1.86E-05	**-3.43**
Taurine	1.45E-07	**-1.47**	Taurine	5.12E-05	**-1.95**

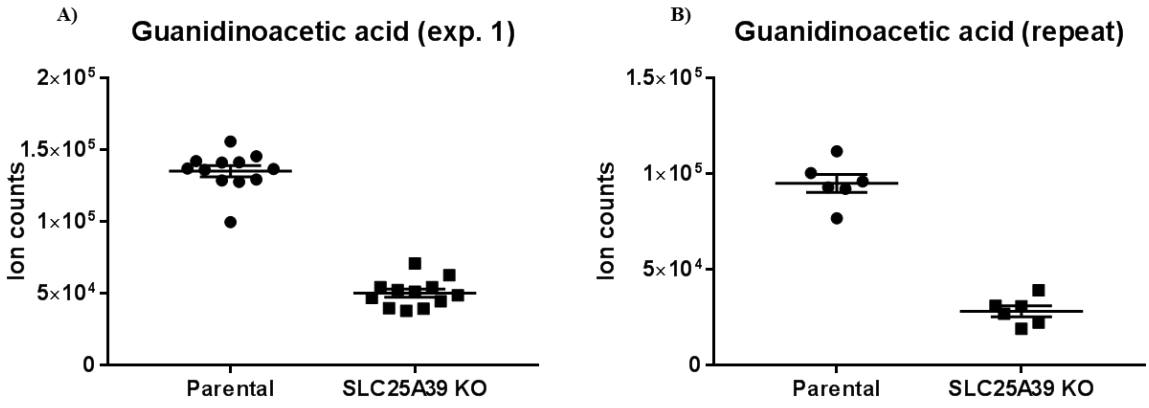


Figure 3.10. *Comparison of the abundance of guanidinoacetic acid in SLC25A39 KO cells compared to control cells. A) A39 KO cells have significantly reduce abundance, possibly due to diminished glycine or arginine transport into the mitochondria. B) Similar results were obtained in a repeat experiment, confirming this difference as a feature of this cell line and highlighting the reliability of the experimental protocol.*

Metabolite profiling of additional *SLC25A39* knockout cell lines show no profound shared metabolite changes and unchanged guanidinoacetate abundance. Metabolite profiling of the original knockout cell line alongside a second and third MDA-MB-468 *SLC25A39* knockout clonal cell line produced with the same guide RNA (does not control for CRISPR off target effects, but as a first pass should arguably be more likely to be similar to the initial knockout cell line) and on the parental cell line as a control (n=6). The abundance of the putative guanidinoacetate compound was not found to be affected in any of the other cell lines (Figure 3.11). Additionally, other features of statistical significance showed no overlap with those from the first knockout cell line, indicating again that differences were clonal peculiarities, cell culture condition dependent, or off-target CRISPR effects. No shared features were promising enough to follow-up on.

Metabolite profiling of *SLC25A39* EGFP tagged reinsertion cell lines do not rescue the observed metabolic perturbations. Metabolite profiling of the original knockout cell line alongside two clonal cell lines in which the *SLC25A39*-EGFP tagged gene was reinserted (constitutively overexpressed, Figure 3.12) and on the parental cell line as a control (n=6). The abundance of guanidinoacetate was not found to be rescued with in the reinsertion lines in any of the other cell lines (Figure 3.13). No shared changes within the reinsertion cell lines and compared to the knockout cell line were promising enough to follow up on.

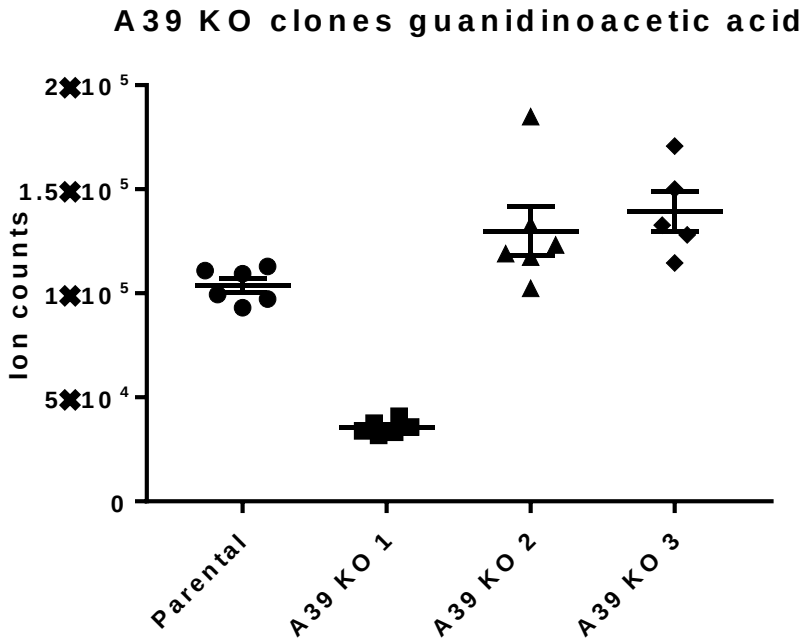


Figure 3.11. *Comparison of the abundance of guanidinoacetate in SLC25A39 KO cells compared to control cells. Although the original cell line once again shows a profound depletion of guanidinoacetate, the two other A39 KO clonal cell lines have unchanged abundance, suggesting that the above observations were not SLC25A39 related.*

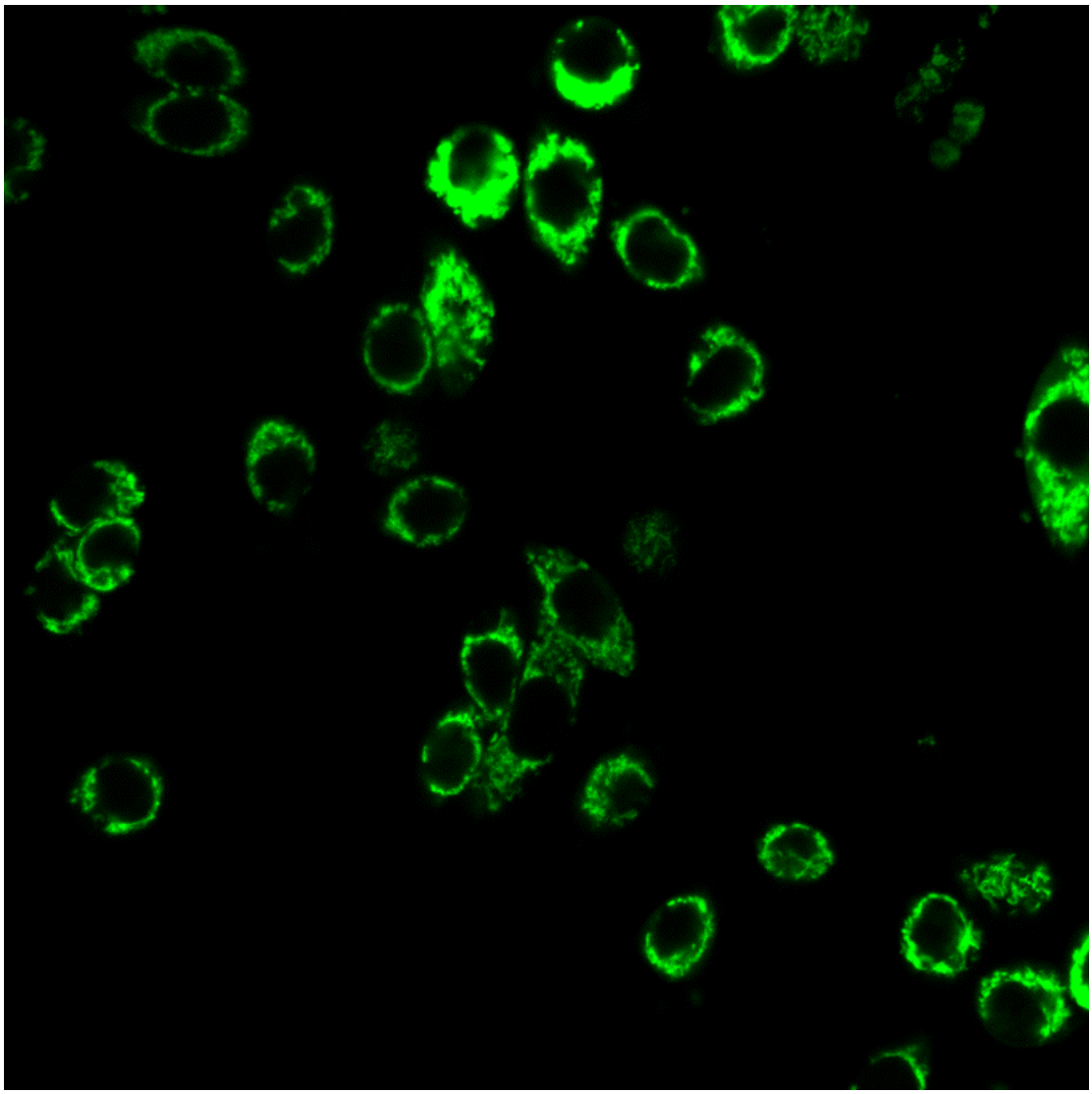


Figure 3.12. *Representative confocal fluorescent image of SLC25A39-EGFP expressing cells. Perinuclear, punctate staining consistent with mitochondrial localization of the reinserted protein.*

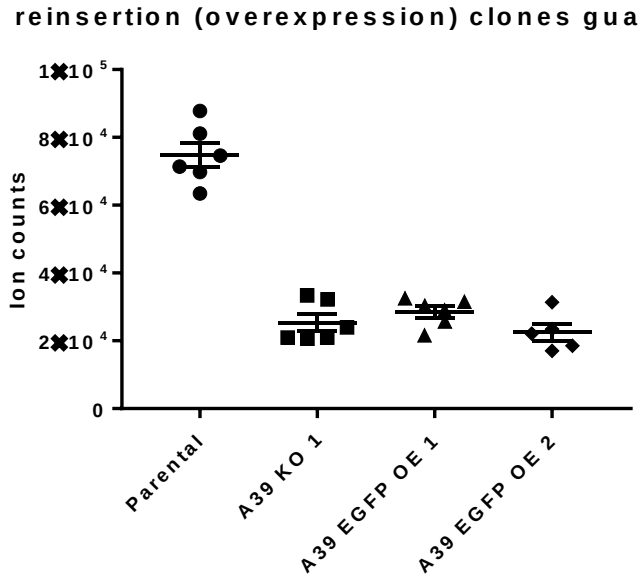
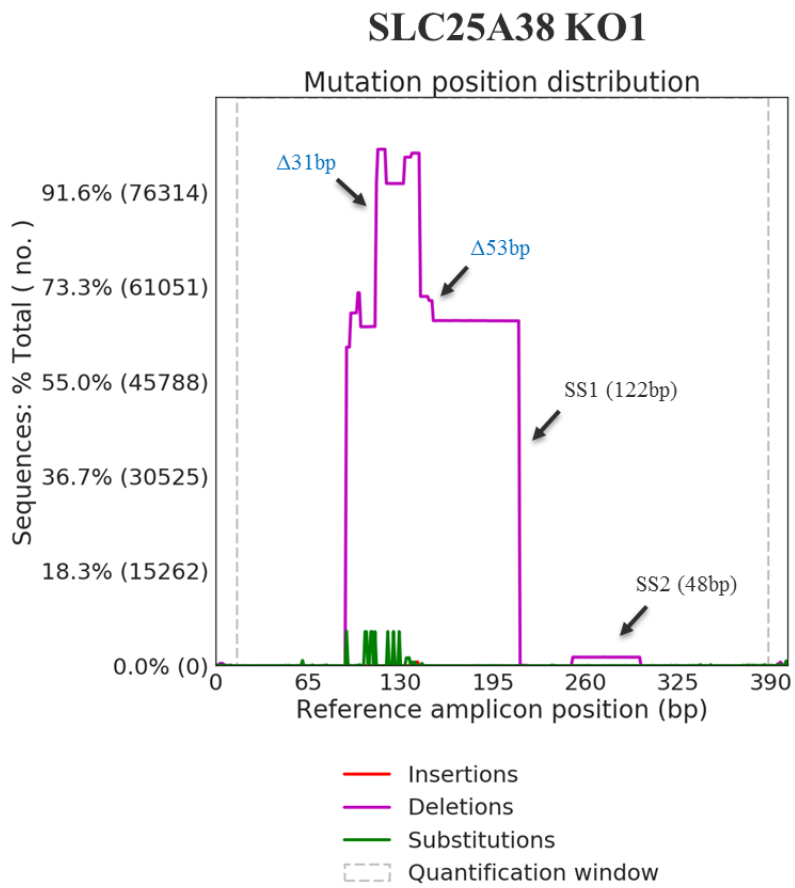


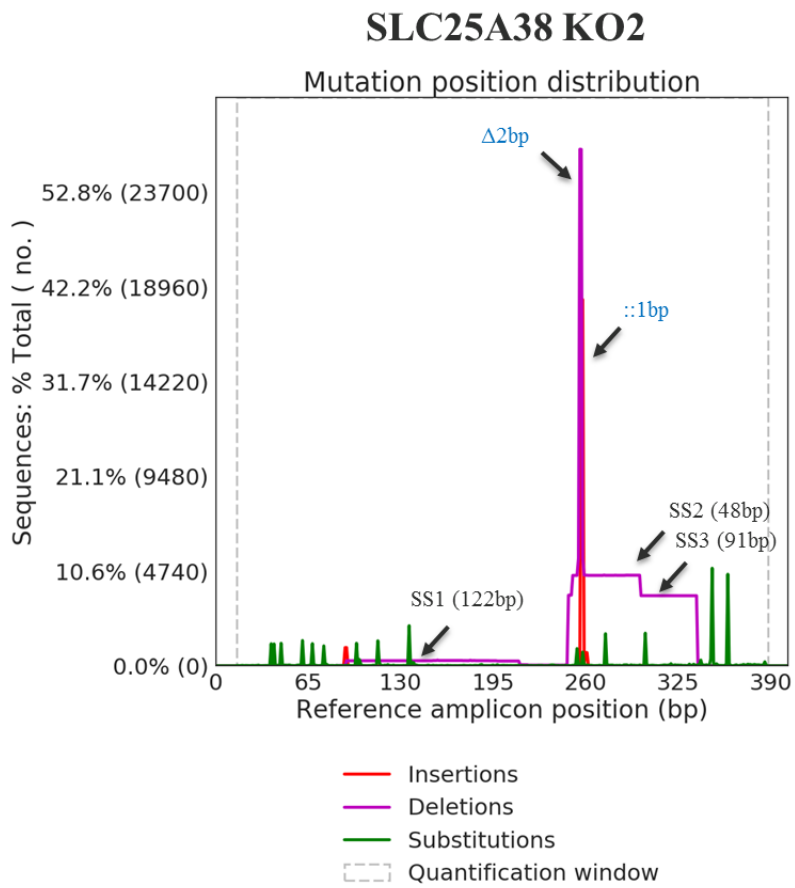
Figure 3.13. *Comparison of the abundance of guanidinoacetate in SLC25A39 KO cells compared to A39-EGFP reinsertion clonal lines. The reinsertion cell lines do not show rescue of guanidinoactate abundance, further indicating that the above observations were not SLC25A39 related.*

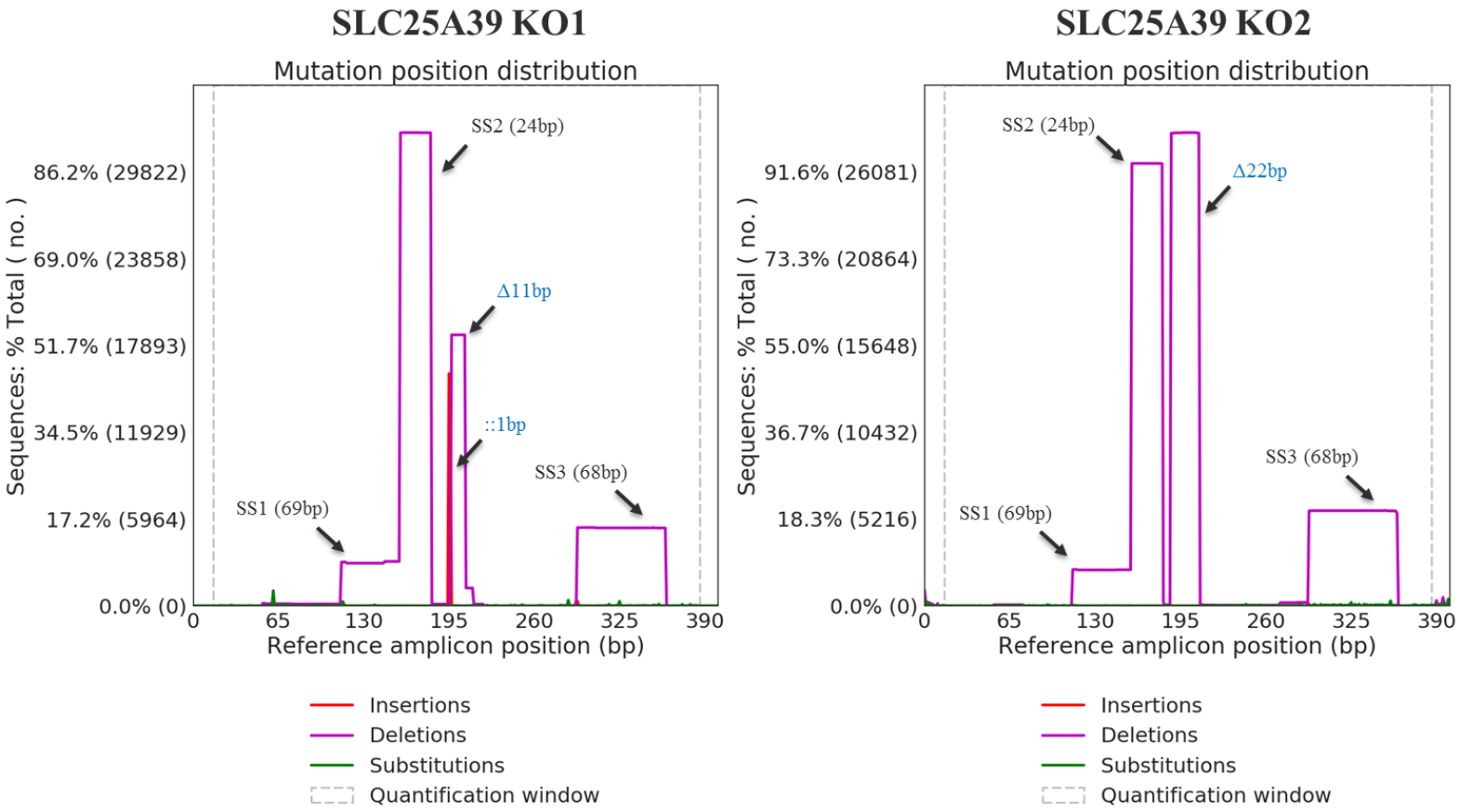
Untargeted metabolite profiling of an MDA-MB-468 *SLC25A44* knockout cell line reveals highly reproducible unidentified metabolite changes compared to parental control. Untargeted metabolite profiling was performed on an MDA-MB-468 *SLC25A44* knockout clonal cell line, compared to the parental control cell line (n=12). Cells were grown for 48h, followed by metabolite extraction and untargeted profiling as described above. Unbiased data was acquired and analysis was performed in a targeted manner (with reference to an in-house retention time and mass-annotated database) and also considering structurally-unidentified compounds. A number of unidentified compounds were differentially abundant, but tentative IDs were not confidently assigned (some matched the predicted m/z of polypeptides). Molecular structure elucidation based on MS/MS fragmentation was attempted, but abundances were insufficient to obtain informative spectra, data not shown.

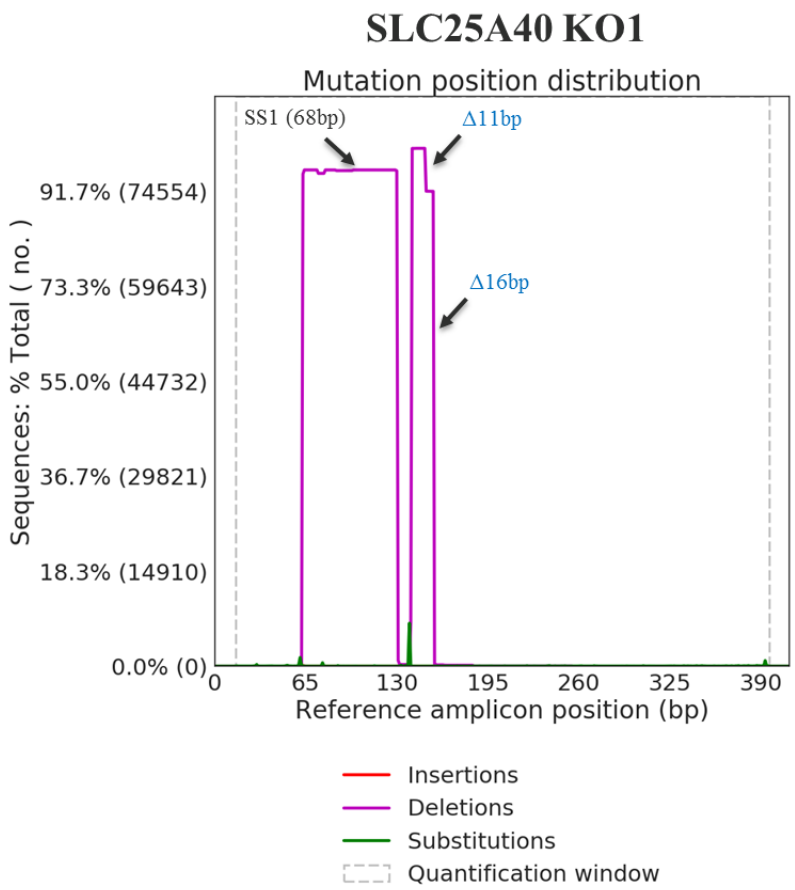
Next-generation sequencing of *SLC25* knockout cell line mRNAs shows no indication of expression of additional alleles but reveals presence of alternate transcript splice variants. mRNA was reverse transcribed to produce cDNA from two knockout cell lines per protein (SLC25A38, SLC25A39, SLC25A40). cDNA were PCR amplified using primers spanning the CRISPR target site (~200bp on either side). PCR products were sequenced using Next-Generation Amplicon Sequencing, which resulted in ~30,000 – 90,000 reads per sample. Alignment to the wild-type reference sequence for each protein (based on the most common NCBI transcript sequence) confirmed the two allele edits identified using *TA-cloning* (shown in Figure 3.2) and showed no indication of additional wild-type (Figure 3.14). Evidence of an additional edited allele sequence was absent or minimal (Table 3.5).

Figure 3.14. *Next-Generation Amplicon Sequencing of knockout cell line mRNA supports 'TA-cloning' genotyping results but reveals high prevalence of alternate splicing. Displayed are graphical overviews of amplicon sequence alignment to wild-type sequence (most common isoform on based on NCBI AceView) indicating the percent of total sequencing reads for which each position on the amplicon is deleted, substituted or has an insertion. Visible CRISPR-mediated insertions and deletions (blue text) and prominent splice sites ('SS', black text) are indicated. Splice sites detected in the SLC25A38 and SLC25A40 genes have not been previously reported.*


SLC25A38 KO1
Mutation position distribution
Δ31bp
Δ53bp
SS1 (122bp)
SS2 (48bp)
91.6% (76314)
73.3% (61051)
55.0% (45788)
36.7% (30525)
18.3% (15262)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window


SLC25A38 KO2
Mutation position distribution
Δ2bp
::1bp
SS1 (122bp)
SS2 (48bp)
SS3 (91bp)
52.8% (23700)
42.2% (18960)
31.7% (14220)
21.1% (9480)
10.6% (4740)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window


SLC25A39 KO1
Mutation position distribution
SS2 (24bp)
Δ11bp
::1bp
SS3 (68bp)
SS1 (69bp)
86.2% (29822)
69.0% (23858)
51.7% (17893)
34.5% (11929)
17.2% (5964)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window
SLC25A39 KO2
Mutation position distribution
SS2 (24bp)
Δ22bp
SS3 (68bp)
SS1 (69bp)
91.6% (26081)
73.3% (20864)
55.0% (15648)
36.7% (10432)
18.3% (5216)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window


SLC25A40 KO1
Mutation position distribution
SS1 (68bp)
Δ11bp
Δ16bp
91.7% (74554)
73.3% (59643)
55.0% (44732)
36.7% (29821)
18.3% (14910)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window

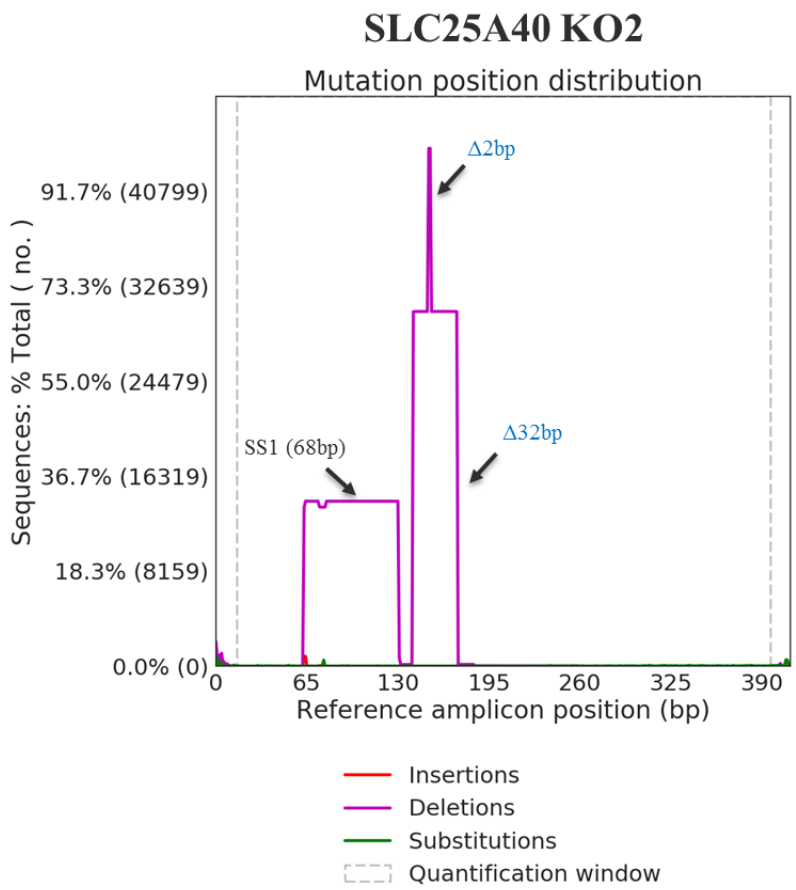

SLC25A40 KO2
Mutation position distribution
Δ2bp
SS1 (68bp)
Δ32bp
91.7% (40799)
73.3% (32639)
55.0% (24479)
36.7% (16319)
18.3% (8159)
0.0% (0)
Sequences: % Total (no.)
0
65
130
195
260
325
390
Reference amplicon position (bp)
Insertions
Deletions
Substitutions
Quantification window

Table 3.5. *Overview of most prevalent 'other' sequences detected with Next-Generation Amplicon Sequencing. This analysis was done to address the question of if there are additional alleles present. 'Other' sequences are defined as those which have a sequences varying form the wild-type sequence at the CRISPR target site. * indicates sequence deviation that is most likely due to splicing, as a region is removed that spans exactly to an exon start site. A38KO2 and A39KO1 are unlikely to represent an additional allele, given the low percent of representation and the absence of detection in 'TA-cloning' genotyping. More likely explanations for these sequences are described in the discussion.*

	1st 'other' sequence % of reads
A38 KO1	NA
A38 KO2	0.47%
A39 KO1	1.85%
A39 KO2	NA
A40 KO1	0.09%*
A40 KO2	0.18%*

There are a number of explanations for these ‘other’ sequences of minor representation in the A38 KO2 and A40 KO1, which are discussed below. However, sequence alignments revealed the presence of transcript variants for all genes. For A39, the 3 primary splice sites detected have been reported previously (NCBI). For A38 and A40 KO cell lines transcript variants that were detected have not been previously described or predicted (based on NCBI, AceView, Ensembl, and ASPicDB databases). These variants include complete exon skipping and removal of portions of predicted exons.

For SLC25A38 KO1 the targeted exon can be skipped entirely but still results in a frame-shift (122bp). >99% of transcript reads indicate a frame-shift (overview of frame-shift analysis for all cell lines shown in Table 3.6). For SLC25A38 KO2 the targeted region can be skipped in two separate variants (skipping 91bp or 38bp). ~97% of detected sequences maintain a frame-shift, with a splice variant representing ~1.5% of reads resulting in an in-frame transcript (with 48bp being skipped).

For SLC25A39, 3 splice variants are observed which do not overlap with the CRISPR target site. For KO1 ~97.5% of detected sequences maintain a frame-shift, with another sequence representing ~1.85% of reads that, combine with the dominant splice variant, results in an in-frame transcript. For KO2 the combination of splice variant 2, CRISPR mediated deletion and splice variant 3 cause the transcript to be in-frame. This represents ~14% of transcript reads.

Table 3.6. *Overview of in-frame sequence read percentage detected with Next-Generation Amplicon Sequencing.*

	% In frame
A38 KO1	<1%
A38 KO2	~3%
A39 KO1	~2.5%
A39 KO2	~15%
A40 KO1	~80%
A40 KO2	~1.5%

For SLC25A40, one splice variant is observed which does not overlap with the CRISPR target site. For KO1 the combination of the splice variant and one allele edit cause the transcript to be in-frame. This transcript represents ~80% of the reads. For KO2 all significantly represented reads maintain a frame-shift, with only ~1.5% of reads in-frame (likely PCR or sequencing artifact).

Human SLC25A38, SLC25A39, SLC25A40 and SLC25A44 are successfully expressed in *Lactococcus Lactis* bacteria. SLC25A38, SLC25A39, SLC25A40 and SLC25A44 with C-terminal FLAG tags were inducibly expressed in *Lactococcus Lactis* bacteria, as shown by western blot (Figures 3.15A, B). Nisin was used to induce protein expression from the pnZ8048 plasmid containing the gene sequences.

Untargeted metabolite profiling of SLC25A38, SLC25A39, SLC25A40 and SLC25A44 (both FLAG and non-FLAG tagged) expressing *Lactococcus Lactis* shows metabolites with strongly increased abundance in specific bacterial lines, likely protein related peptides. Untargeted metabolite profiling was performed on SLC25A38, SLC25A39, SLC25A40 and SLC25A44 clonal knockout cell line and on the parental cell line as a control (n=8).

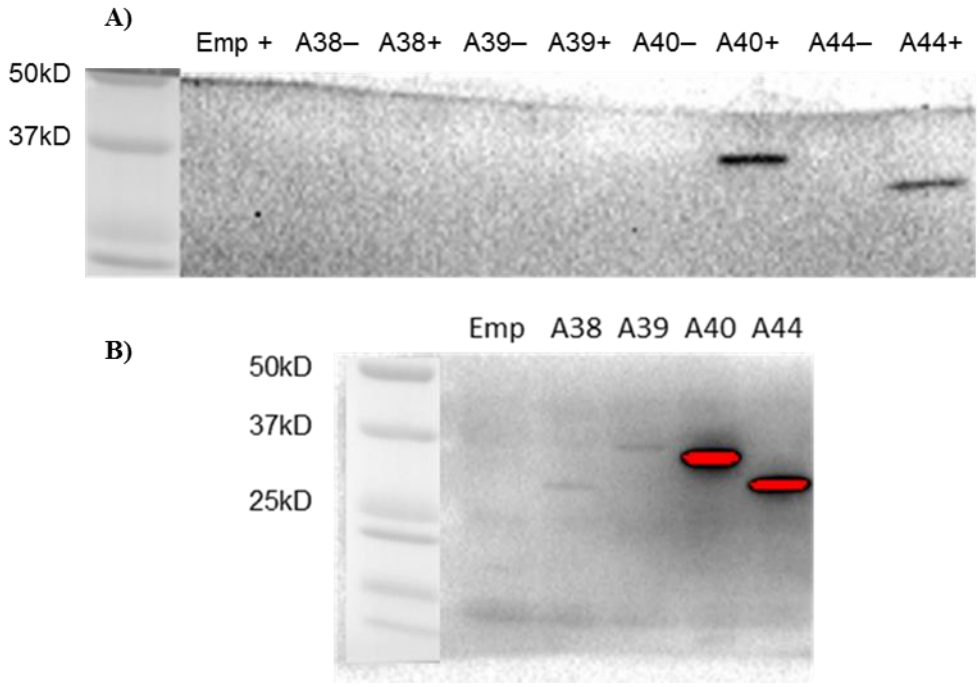


Figure 3.15. *Western blots showing expression of SLC25A38, SLC25A39, SLC25A40 and SLC25A44 with C-terminal FLAG tags. A) Inducible expression of A40 and A44. + or – indicate presence or absence of the inducer nisin, respectively. B) A38 and A39 expression are also detectable, albeit at relatively lower levels.*

Bacteria were grown from dilution of an overnight culture until induction at O.D. 600nm of 0.6, followed by an additional 3 hours of proliferation and protein expression. For this experiment, cells were washed with a series of PBS and water washes with intermittent centrifugation to remove the wash media (initial extraction protocol). Untargeted profiling was done as described above. Unbiased data analysis was performed in an untargeted (unknown compound ID, data not shown) and targeted manner (using an in-house retention time and mass annotated database). Due to massively different metabotypes which would lead to hundreds of differentially abundant compounds if compared to control (which would be uninterpretable), groups were compared to all the others in order to highlight only the most significant differences that were unique to bacteria with a specific SLC25 transporter expression. This analysis highlighted a number of untargeted compounds that were greatly more abundant with expression of a given protein (these were shown to not be clonal, as the same results were found in both FLAG and non-FLAG tag protein expressing clones) (Figure 3.16). Many of these unknowns matched the predicted m/z of polypeptides and may be degradant peptides of the overexpressed protein. Most of these compounds were not detectable in the bacteria culture media, indicating that it was not being transported into the bacteria (i.e. not the expressed protein's substrate). Coupled with these possibilities, concerns related to the extraction protocol (metabolite leakage due to extensive time spent in wash steps) led us to forgo follow-up on these compounds and to rather develop a superior extraction methodology and a more targeted analytical approach.

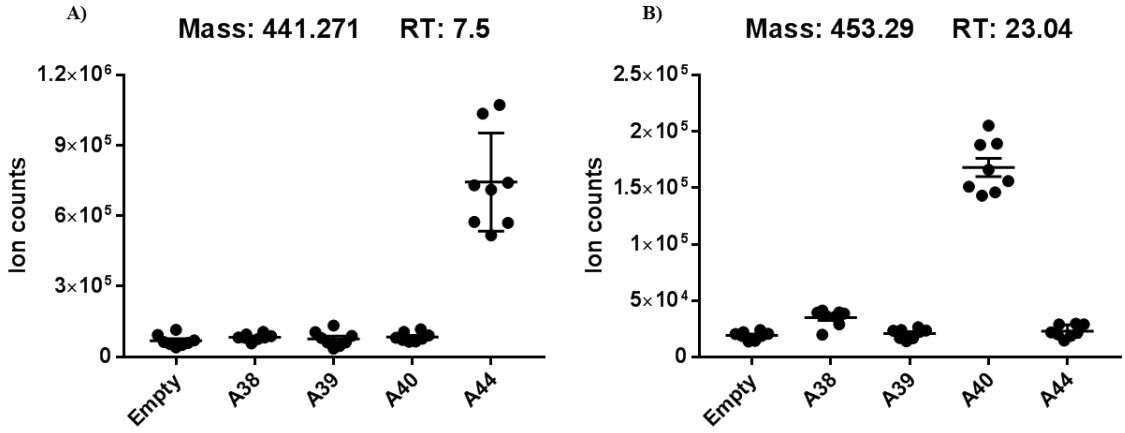


Figure 3.16. *Examples of metabolites that were greatly increased in one of the protein expressing lines compared to the others. A) A coumpound with mass 441.27 at retention time 7.5 min was very abundant in SLC25A44 expressing Lactococcus Lactis compared to all the other cell lines. B) A coumpound with mass 453.29 at retention time 23.04 min was very abundant in SLC25A40 expressing Lactococcus Lactis compared to all the other cell lines.*

Most human SLC25A family carboxylate transporters are successfully expressed in *Lactococcus Lactis* bacteria. SLC25A7, 8, 9, 10, 11, 14, 21, 27, 30, 34, and 35 with C-terminal FLAG tags were successfully inducibly expressed in *Lactococcus Lactis* bacteria, as shown by western blot (Figures 3.17 A, B). Nisin was used to induce protein expression from the pnZ8048 plasmid containing the gene sequences. Proteins were expressed with highly varying abundances. SLC25A10 showed a faint band on the western blot but a dot blot of the same samples indicates that the low detection may be due to a technical issue (e.g. inefficient transfer to the membrane) (Figure 3.17C). SLC25A35 shows significant degradation and is likely not a good candidate for study/expression in these bacteria. SLC25A1 shows no conspicuous expression, although the dot blot suggets very low expression. As the homologs of this protein from other species have been expressed before in this system, it may be possible with varying nisin concentration. In all, 10 out of 12 proteins were found to be clearly expressed without significant degradation.

Transport assays with SLC25A11 expressing bacteria (with and without FLAG tag) fail to show increased 2-oxoglutarate transport with protein expression. Initial transport assays were attempted with unlabeled 2-oxoglutarate in order to get a sense for the proper conditions, such as concentration and time, and in order to confirm that the metabolite extraction protocol was effective. First, 2-oxoglutaric acid (200 μM) was assessed for transport during a 30 min interval using the assay described above. Both SLC25A11 overexpressing bacteria and empty vector control did not reveal significant transport over this time course (Figure 3.18A, B).

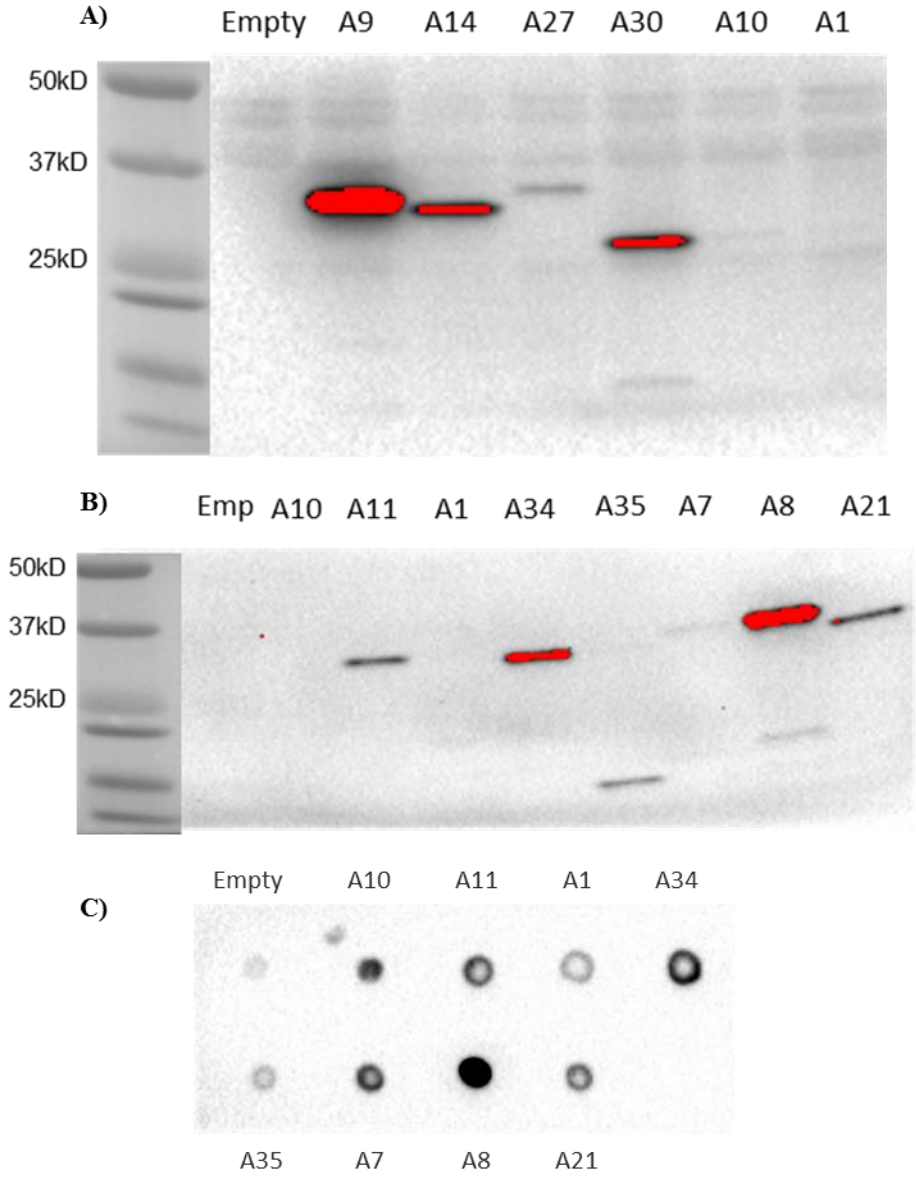


Figure 3.17. *Protein expression of most of the SLC25A family carboxylate transporters. A) Western blot showing expression of SLC25A9, 14, 27, 30, 10. A1 may have a very faint band or may not be expressed. B) Western blot showing expression of SLC25A11, 34, 35, 7, 8 and 21. A35 shows significant degradation. C) Dot blot of the samples shown in B, indicating possible expression of A10 and A1.*

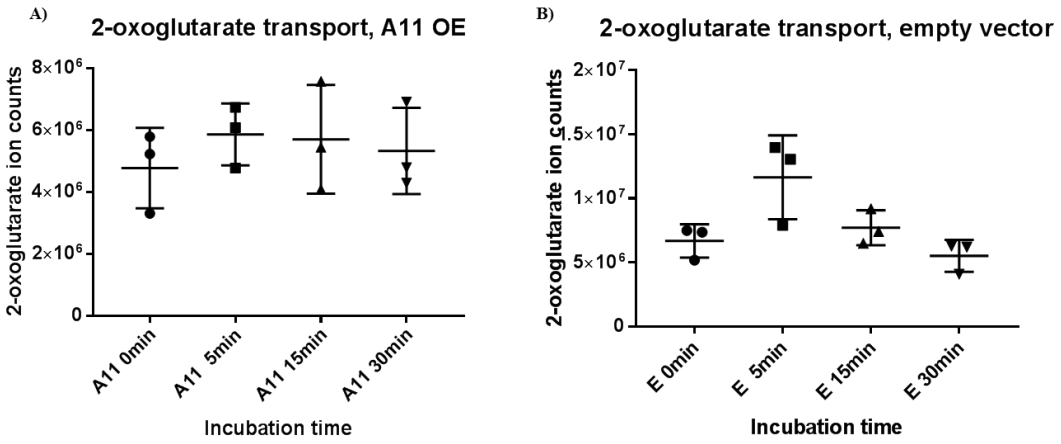


Figure 3.18. *SLC25A11-FLAG OE and empty vector Lactococcus Lactis 30 min time course transport assay. 2-oxoglutarate was used in a transportation assay at a concentration of 200 µM with A) SLC25A11 overexpressing Lactococcus Lactis and B) empty vector Lactococcus Lactis. No indication of transport was observed.*

This may have been caused by an ineffective choice of substrate concentration, so an additional experiment was performed to test the potential impact of substrate concentration on uptake. Malate transport was also attempted, as SLC25A11 was reported to exchange malate for 2-oxoglutarate (and possibly other carboxylates). Although no concentration of either substrate showed clear evidence of SLC25A11-selective transport, 2mM 2-oxoglutrate was markedly increased in 2/3 samples (Figure 3.19A, B). Although this is now believed to have been protocol or compound related variation (error), this concentration was used for all future experiments.

This concentration/time point was repeated, this time with an empty vector control to check for increased transport with SLC25A11 overexpression. A difference between the substrate treated and untreated groups was observed, but the data variation within groups was excessive (Figure 3.20). We reasoned that this may be due to substrate carry-over due to insufficient washing of the bacteria. To evaluate this possibility, an experiment was performed to assess the effect of repeated 3-sec wash steps on substrate detection levels. Toward this end, the bacteria were gently vortexed to break up clumps and aspirated along with the 1st wash (as opposed to before) to facilitate uniform distribution over the filter membrane. As shown in Figure 3.21, a single additional wash step was sufficient to remove unbound substrate and at least 4 more wash steps did not lead to bacteria-associated substrate leakage. This protocol (including careful vortex and aspiration steps) was used to repeat the SLC25A11 overexpression vs. empty vector control experiment. 2-oxoglutarate was increased in subtrate treated vs. untreated groups, suggesting uptake into the bacteria, but increased uptake with A11 expression was not clear (comparison of A11 2-OG group vs. empty vector 2-OG group) (Figure 3.22).

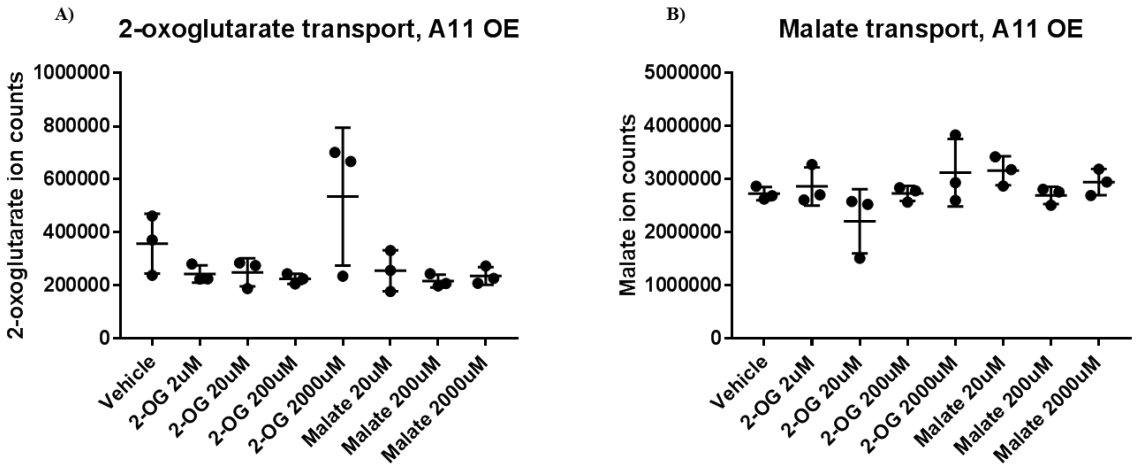


Figure 3.19. *SLC25A11-FLAG OE Lactococcus Lactis substrate concentration test. Various concntrations of A) 2-oxoglutarate and B) malate were tested for transport in SLC25A11 overexpressing Lactococcus Lactis during a 30 min time interval. 2mM 2-oxoglutarate showed the most promising results.*

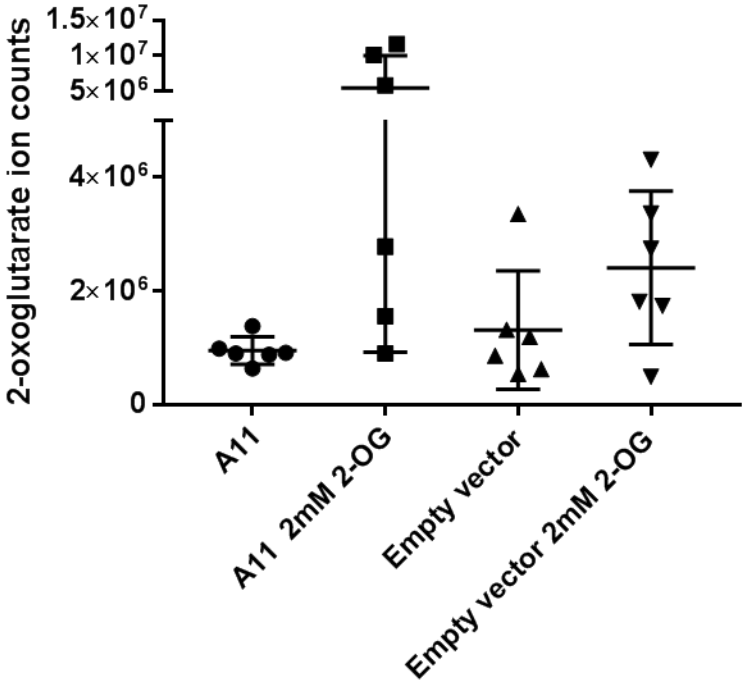


Figure 3.20. *SLC25A11-FLAG OE and empty vector Lactococcus Lactis 30 min transport assay. 2-oxoglutarate was used in a transportation assay at a concentration of 2mM with SLC25A11 overexpressing Lactococcus Lactis and an empty vector control. Indication of transport with substrate vs. without was observed. However, significant data variation was also present.*

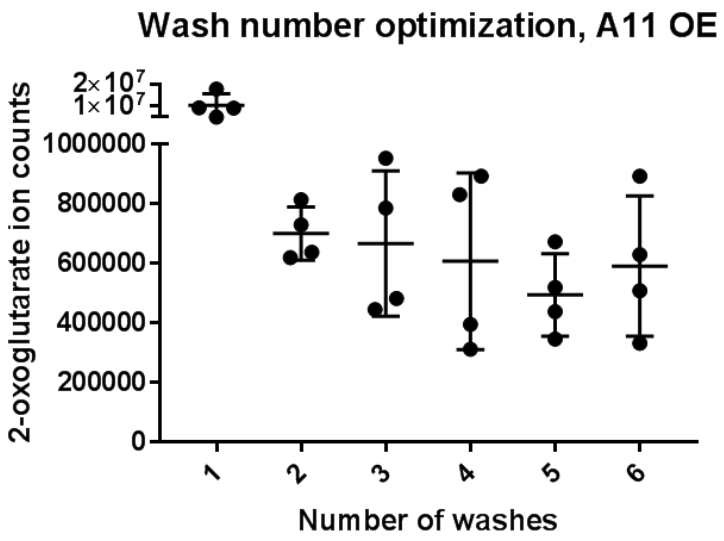


Figure 3.21. *Various numbers of 3 second washes were tested to determine the optimal number of washes to remove excess external substrate without internal substrate leakage. Any number of washes 2 or greater appear sufficient. No significant leakage was observed with up to 6 washes.*

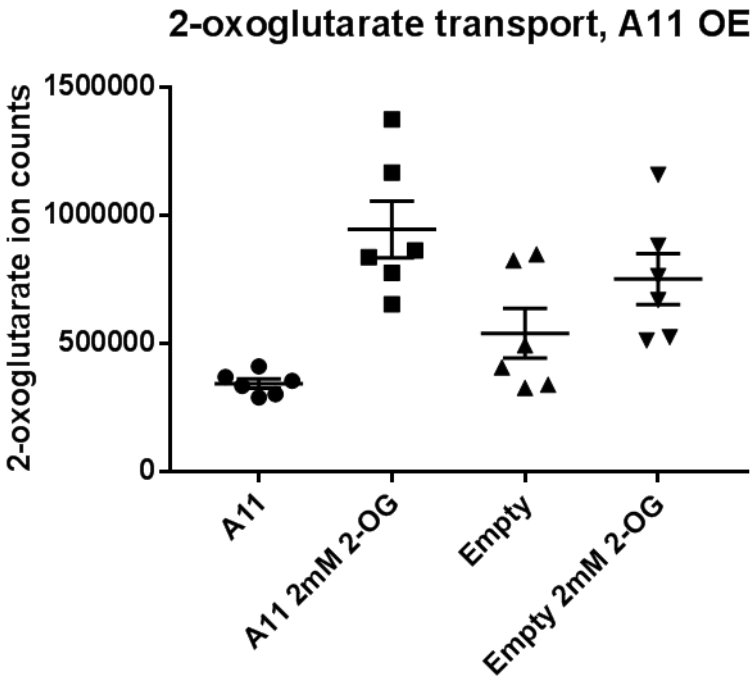


Figure 3.22. *SLC25A11-FLAG OE and empty vector Lactococcus Lactis 30 min transport assay. 2-oxoglutarate was used in a transportat assay at a concentration of 2 mM with SLC25A11 overexpressing Lactococcus Lactis and an empty vector control. Indication of transport with substrate vs. without was observed, but SLC25A11 mediated transport was not clear (comparison of A11 2-OG group vs. empty vector 2-OG group).*

Now that the protocol was established and transport (at least endogenous) of 2-oxoglutarate was indicated, experiments with heavy isotope labeled substrate, which has no background signal, were performed. The time frame was reduced, as possible transport saturation could cause the difference between overexpression and control to be masked. [6D]-2-oxoglutarate was transported at 5 min and 14 min time points (Figure 3.23). Although substrate was observed to be associated with bacteria (consistent with transport), no difference was seen between the groups. Surprisingly, no time-dependent difference was observed, possibly indicating rapid saturation. Because the addition of a FLAG-tag could potentially perturb transporter function, a follow-up experiment was performed using non-tagged SLC25A11 overexpressing bacteria and very brief time points (2min/5min) (Figure 3.24). At 2 min, SLC25A11 overexpression trended towards increased [6D]-2-oxoglutarate abundance, but sample number was too low to be conclusive. Again, however, no increase was observed from 2 min to 5 min. Therefore, the experiment was repeated with a 1 min time point and more samples per group. This experiment was also inconclusive, as data variation was once again too high (Figure 3.25). This level of variation seen in some experiements was only ever observed with 2-oxoglutarate and may indicate that this compound is, for unknown reasons, not a good candidate for this transport assay setup.

Transport assays with FLAG tagged SLC25A21 expressing bacteria fail to show increased 2-oxoglutarate transport. SLC25A21 is an oxoadipate/oxoglutarate exchanger and was tested (with a FLAG tag) for its ability to increase 2-oxoglutarate transport compared to an empty vector control. As shown in Fig 23, SLC25A21 expression was not associated with increased transport.

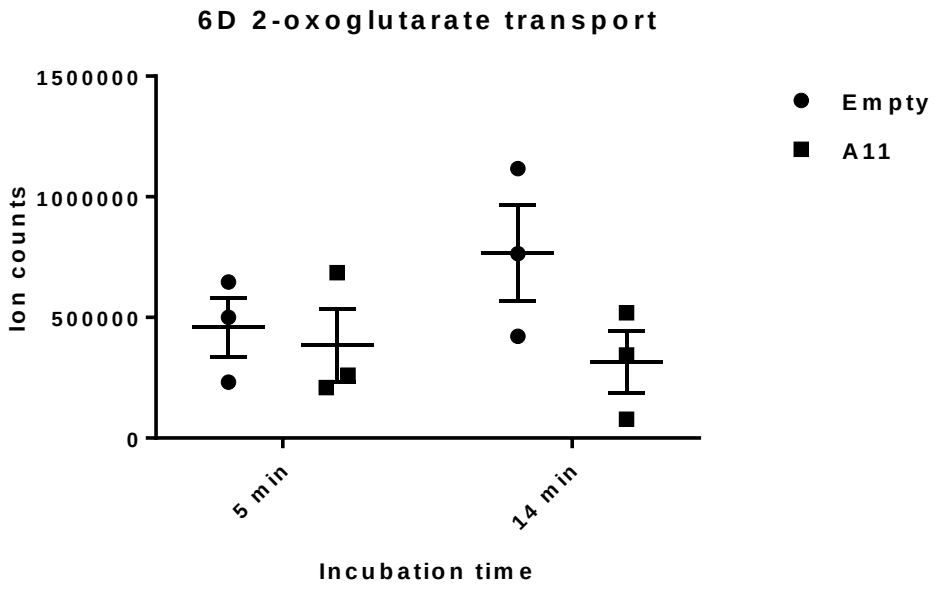


Figure 3.23. *SLC25A11-FLAG OE and empty vector Lactococcus Lactis transport assay. [6D]-2-oxoglutarate was studied at a concentration of 2mM with SLC25A11 overexpressing Lactococcus Lactis and an empty vector control at 5 min and 14 min time points. [6D]-2-oxoglutarate is clearly entering the cells, but no difference between groups and no increase from 5 min to 14 min were observed.*

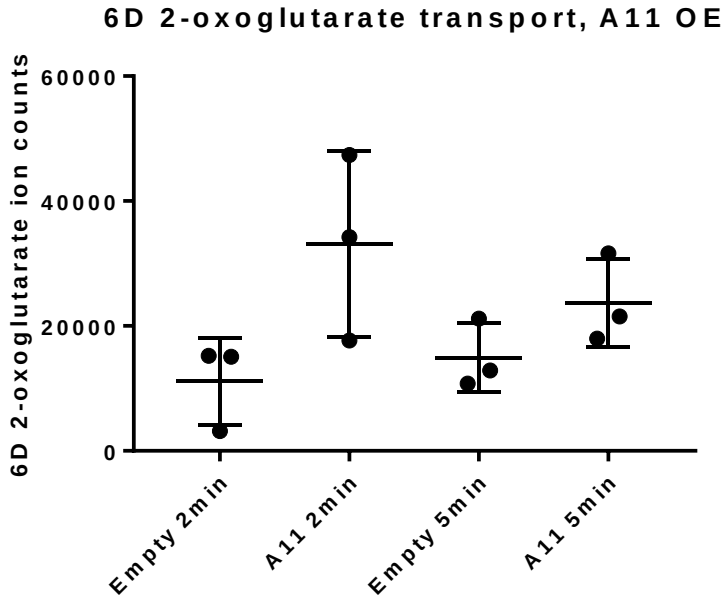


Figure 3.24. *SLC25A11 OE and empty vector Lactococcus Lactis transport assay. 6D 2-oxoglutarate was used in a transportation assay at a concentration of 2mM with SLC25A11 (no tag) overexpressing Lactococcus Lactis and an empty vector control at 2min and 5min time points. [6D]-2-oxoglutarate is clearly entering the cells, but no clear difference between groups was observed (although a possible increase with protein expression at 2 min). No increase from 2 min to 5 min was observed.*

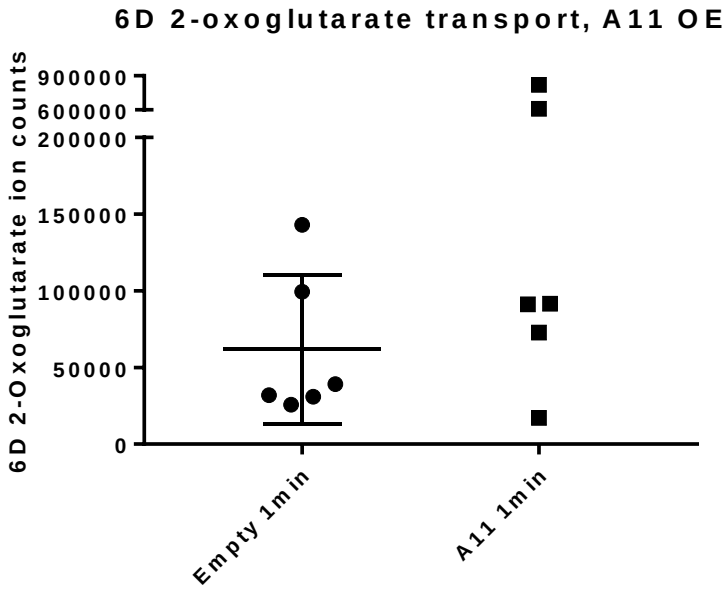


Figure 3.25. *SLC25A11 OE and empty vector Lactococcus Lactis transport assay. 6D 2-oxoglutarate was used in a transportation assay at a concentration of 2mM with SLC25A11 (no tag) overexpressing Lactococcus Lactis and an empty vector control at a 1min time point. [6D]-2-oxoglutarate is clearly entering the cells, but no clear difference between groups was observed. Again, significant data variation led to inconclusive results.*

Transport assays with SLC25A38 expressing bacteria (with and without FLAG tag) fail to show increased glycine transport. SLC25A38 has been recognized as a glycine transporter and was also used in an attempt to show protein mediated transport in *Lactococcus Lactis*. Heavy isotope (^{13}C labeled) glycine was used in the transport assay described above with both FLAG tagged (Figure 3.27A) and non-tagged (Figure 3.27B) SLC25A38 overexpressing bacteria. Although endogenous transport was evident and replicate data were tight, no effect of SLC25A38 on glycine transport was observed.

Transport assays with SLC25A40 expressing bacteria (with FLAG tag) fail to show increased glycine transport with protein expression. SLC25A40 is homologous to the SLC25A38 glycine transporter and showed robust protein expression in *Lactococcus Lactis*. Therefore, it was tested whether glycine may be a substrate for transport by SLC25A40 in *Lactococcus Lactis*. For this purpose, [^{13}C]-glycine was studied for uptake into SLC25A40-FLAG overexpressing bacteria (Figure 3.28). Although endogenous transport was evident and replicate data very tight, no effect of transporter expression on relative uptake was observed.

Transport assays with SLC25A8 expressing bacteria (without FLAG tag) fail to show increased aspartate transport with protein expression. SLC25A8 is recognized as an aspartate transporter and was also used in an attempt to demonstrate protein mediated transport in *Lactococcus Lactis* – notably, robust protein expression was observed (Figure 3.17B). Heavy isotope labeled aspartate (^{15}N) was used in the transport assay described above with SLC25A8 (no tag) overexpressing bacteria. Although endogenous transport was clear and replicate data very tight, no effect of protein expression was observed (Figure 3.29).

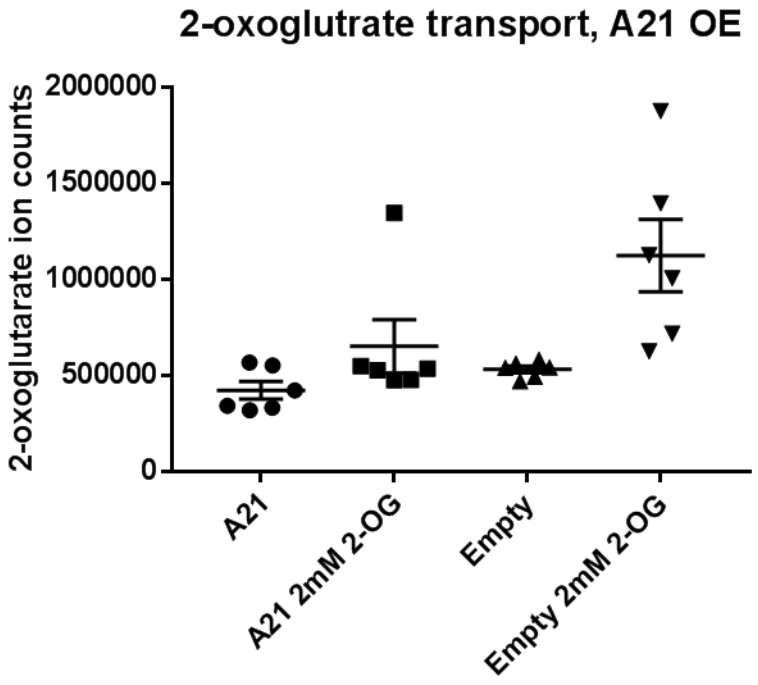


Figure 3.26. *SLC25A21-FLAG OE and empty vector Lactococcus Lactis 30 min time course transport assay. 2-oxoglutarate was used in a transportation assay at a concentration of 200 µM with SLC25A21 overexpressing Lactococcus Lactis and empty vector Lactococcus Lactis. Although endogenous transport was indicated, no indication of SLC25A21 transport was observed.*

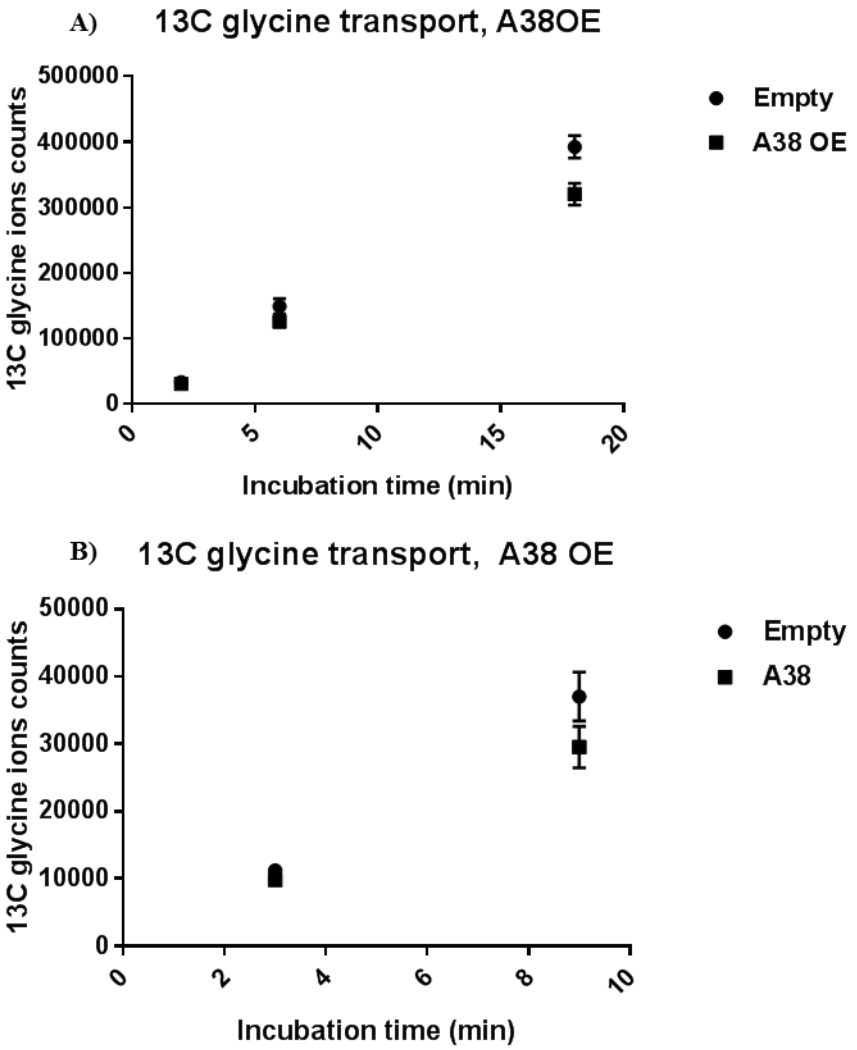


Figure 3.27. *SLC25A38 OE and empty vector Lactococcus Lactis transport assays. 13C glycine was used in a transport assay at a concentration of 2mM with A) SLC25A38-FLAG overexpressing Lactococcus Lactis and an empty vector control at 2min, 6min and 18min time points and B) SLC25A38 (no tag) overexpressing Lactococcus Lactis and an empty vector control at 3min and 9min time points. Linear endogenous transport and highly reproducible data is observed, but there is no indication of SLC25A38 mediated transport.*

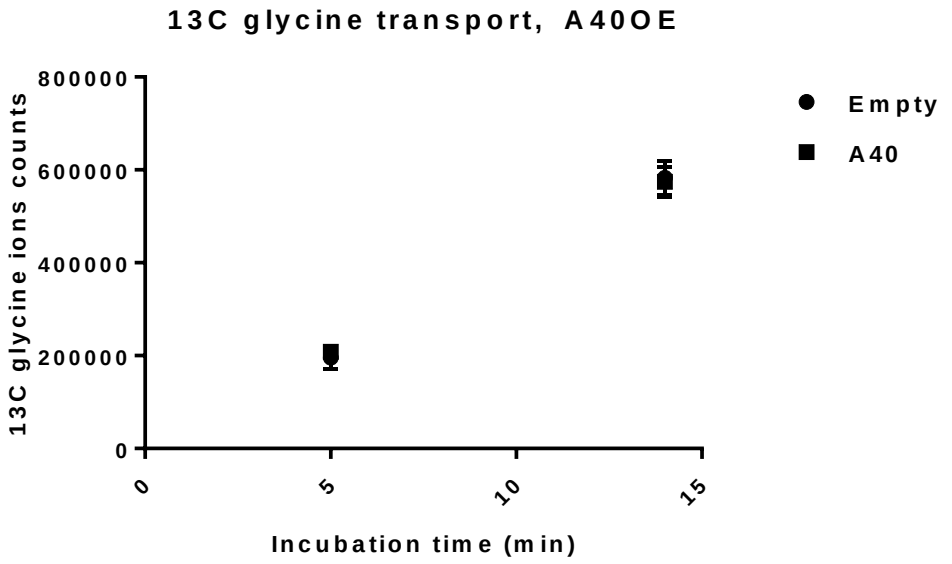


Figure 3.28. *SLC25A40-FLAG OE and empty vector Lactococcus Lactis transport assays. 13C glycine was used in a transportation assay at a concentration of 2mM with SLC25A40-FLAG overexpressing Lactococcus Lactis and an empty vector control at 5min and 14min time points. Linear endogenous transport and highly reproducible data is observed, but there is no indication of SLC25A40 mediated transport.*

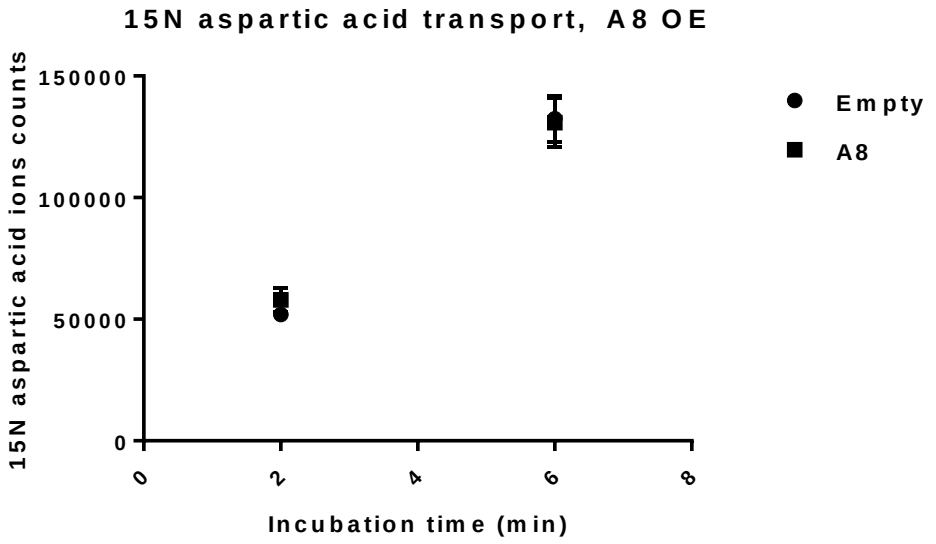


Figure 3.29. *SLC25A8 OE and empty vector Lactococcus Lactis transport assays. 15N aspartate was used in a transportation assay at a concentration of 2mM with SLC25A8 (no tag) overexpressing Lactococcus Lactis and an empty vector control at 2min and 6min time points. Linear endogenous transport and highly reproducible data is observed, but there is no indication of SLC25A8 mediated transport.*

DISCUSSION

Initially, our primary goal was to identify the substrates transported by orphan mitochondrial transporters from the SLC25 family of proteins. Specifically, we focus on SLC25A39, SLC25A40 and SLC25A44, which have high homology to SLC25A38, a putative glycine transporter (Figure 3.1, from [Palmieri 2013]). Our hypothesis was that these transporters may transport glycine, glycine-like compounds, such as serine (SLC25A38 has been hypothesized to be a glycine/serine antiporter [Guernsey et al. 2009; Fernández-Murray et al. 2016]), and/or other amino acids. We included SLC25A38 in our study, which has not yet been fully characterized.

We produced monoclonal homozygous knockout cell lines for each individual protein of interest in MDA-MB-468 triple negative breast cancer cells with the idea that the cells could then be probed for profound abnormalities in metabolism that could be traced back to the lost function of an SLC25-type transporter protein. Due to the lack of characterized western blot antibodies for these SLC25 transporter proteins, homozygous knockout was confirmed using a PCR-based screen for genome edits, followed by allele sequencing by *TA cloning*. Cell lines with a confirmed frameshift-causing edit in both alleles were used for the experiments.

In a first experiment, untargeted metabolite profiling was performed with one *SLC25A38* KO cell line, one *SLC25A40* KO cell line and the parental MDA-MB-468 as the control cell line for each. Metabolic changes compared to control were found to be overall modest (fold change < 2) for both cell lines (Tables 3.1 and 3.3).

SLC25A38

For the *SLC25A38* KO cell line features possibly corresponding to acetylglycine (an acetic acid adduct), acetylleucine and acetylserine (based on mass and retention time matches to an in-house reference database) were among the most potentially relevant changes (Table 3.1). The relatively diminished level of acetylglycine (Figure 3.3) may be explained by reduced glycine import into the mitochondria, where it can be acetylated by the mitochondrial localized protein glycine-N-acyltransferase [Nandi et al. 1979]. The reduction in acetylglycine production could lead to diversion of acyl-CoA towards other products, including acetylleucine and acetylserine, explaining the increased abundances of these compounds.

In attempt to reproduce these results, an experiment was performed using a second *SLC25A38* KO cell line made with a different guide RNA (controlling for potential off-target CRISPR effects or clonal peculiarities). Acetylglycine abundance was less severely decreased in this follow-up experiment (1.3 fold; Figure 3.4). Taken together, these experiments suggest further validation of SLC25A38 as the glycine transporter. The most profound changes from both targeted and untargeted analyses of both experiments were compared, but the levels of no other compounds of interest were significantly changed in both cell lines (Tables 3.1 and 3.2).

Worthy of note, sorbitol was increased significantly in the knockout cell lines for many of the mitochondrial proteins. This change may be a response to CRIPSR gene editing and we therefore did not further pursue it. However, it could also be a true metabolic perturbation due to the loss of the function of these transporters, which

would possibly suggest overlapping metabolism among these transporters, and could be worthy of further scrutiny.

SLC25A40

With regard to SLC25A40, as mentioned above, metabolic changes in the knockout compared to control were modest (fold change < 2; Table 3.3). One of the most significant changes was a reduction in betaine in the knockout cell line (1.73 fold change; Figure 3.5). Betaine's identity was initially based on comparison to our in-house database and then confirmed by MS/MS fragmentation followed by matching to the METLIN fragmentation database (Figure 3.6 and Figure 3.7). Betaine is produced exclusively in the mitochondria via a two-step enzymatic process from its precursor, choline. Importantly, betaine can be used in one carbon trafficking and thereby contribute to purine salvage [Ueland et al. 2015; Obeid 2013] as well as other one-carbon donor fates. Therefore, we hypothesized that SLC25A40 could function as a choline transporter, with KO resulting in diminished choline import into the mitochondria, explaining the observed decrease in cell betaine. Notably, no mitochondria-specific choline transporter has yet been identified, although the plasma membrane choline transporter has been shown to also localize to the mitochondria [Michel, et al. 2009].

Surprisingly, in a follow-up experiment the same cell line showed only a mild decrease in betaine (data not shown). However, confidence in the decreased betaine in the original *SLC25A40* knockout cell line was restored in a choline tracing experiment.

Indeed, we traced deuterated choline ([9D]-choline) supplied to the cells in the cell culture media to assess whether *SLC25A40* KO cells do indeed exhibit a decrease

in choline-derived betaine (Figure 3.8B). Diminished [9D]-choline-derived betaine was confirmed at 1, 5 and 24 hrs. Importantly, total deuterated choline in cells was not different between the *SLC25A40* KO and control cell lines, indicating that overall choline uptake from the media by both cell lines was indistinguishable (Figure 3.8D). Also, the abundance of extramitochondrial-produced deuterated choline-derived compounds, such as phosphocholine and phosphatidylcholines, were not different between the two cell lines, further indicating that the perturbation observed with regard to betaine could be due to reduced choline import into the mitochondria (Figure 3.8F).

In attempt to validate this conclusion, deuterated betaine uptake was quantified in four other knockout cell lines produced from the same guide RNA (two of these had only a single sequencing read, suggesting an identical edit on both alleles and a high likelihood of being a homozygous knockout). This was a first pass at validation before endeavoring to engineer knockout cell lines using an alternative guide RNA. Cell lines produced with the same guide RNA would be expected to show a similar phenotype if the observed metabolic changes observed were indeed due to the function of the knocked out transporter. Surprisingly, none of the other knockout cell lines had decreased betaine production compared to the control cell line, suggesting that the reduced betaine in the original cell line may be growth state-dependent, or a clone-specific metabolic peculiarity that is unrelated to SLC25A40 function (Figure 3.9).

SLC25A39

We performed a similar experiment for SLC25A39, comparing metabolite abundances for an *SLC25A39* knockout cell line to the control parental line. Here,

there were a few profound metabolite changes (>2 fold) in the knockout compared to control, such as compounds with tentative identifications of guanidinoacetic acid and CDP-choline, as well as unidentified features from untargeted analyses (Table 3.4 left side, Figure 3.10A). These changes were highly reproducible in this cell line in a repeat experiment (Table 3.4, Figure 3.10B), validating the reliability of the cell growth conditions and metabolite extraction protocol. In fact, guanidinoacetate was found to be severely depleted in this cell line in every follow-up experiment. Guanidinoacetate is produced from the precursors glycine and arginine by the mitochondrial protein arginine:glycine amidinotransferase [Humm et al. 1997], which led to the hypothesis that SLC25A39 could be an arginine or glycine transporter.

However, similar to findings with SLC25A40, this metabolic perturbation was not observed in other *SLC25A39* KO clonal cell lines produced with the same guide RNA, indicating again that reduced guanidinoacetate may be cell growth-state dependent or arise as a clonal specific peculiarity unrelated to the transporter (Figure 3.11). In support of this interpretation, reinsertion of EGFP tagged SLC25A39 into the knockout cell line was unable to rescue this phenotype (Figure 3.12; Figure 3.13).

SLC25A44

We performed a similar experiment for SLC25A44, comparing metabolite abundances for an *SLC25A44* knockout cell line to the control parental line. Here, there were also a few profoundly differentially abundant features (>2 fold) in the knockout compared to control in an untargeted analysis. These changes were highly reproducible in this cell line in a repeat experiment. However, the identity of these features could not be established by MS/MS analysis due to relatively low species

abundance and poor fragmentation spectra obtained. Considering the prevalence of what may be clonal peculiarities, based on our analysis of other *SLC25* transporter knockout cell lines, the molecular identities of these species and relevance of their changing levels was not pursued further (data not shown).

Suitability of knockout cell lines as a system for untargeted mitochondrial transporter studies

In all, these experiments highlight a number of challenges and considerations for studying mitochondrial transporters in knockout eukaryotic cell lines. Apparent clone-specific metabolite abundances and inability to confirm findings were a major pitfall of this study. These could be caused by off-target CRISPR edits, genetic variation in the cancer cell line, epigenetic variation, or growth state-dependent metabolism. In the first possibility, the occurrence of off-target CRISPR-mediated genome edits in some knockout clones and not others could lead to metabolic variation that would explain the observed metabotypic difference. Although the possibility of CRISPR off target effects were known at the beginning of the study [Zhang et al. 2015], the likelihood/prevalence of this affecting metabolism may not have been adequately appreciated.

In terms of genetics or epigenetics, it is possible that there is a wide array of metabolite variation within a multi-clonal population of parental cell lines. This variation in metabolite abundances among clones would likely fall along a normal distribution for each metabolite. Conceivably, certain metabolites would have larger variation in abundance than others. Selecting one clone from among a population of clones and then performing metabolite profiling, surveying thousands of compounds,

would have a high likelihood of finding one or more metabolites for which that clone has an abundance that deviates markedly from average.

In support of this possibility, Zong et al. showed genetic variation in a single cell selected from SW480 colon cancer cell line compared to the population average [Zong et al. 2012]. Specifically, they detected distinct copy number variation (CNV) and single nucleotide variation (SNV) in single cell genome sequencing. CNVs were detected in single cells selected from the population of SW480 cells obtained from American Type Culture Collection (ATCC). Importantly, for SNV they first expanded a single cell into a colony in order to ensure a single genetic origin and grew for 20 generations before selecting 1 cell for comparison to the bulk. They estimate that 49 mutation occurred during the 20 generations, indicating a mutation rate of 2.5 nucleotides per generation (in addition to detectable copy number variation). This highlight the possibility of genetic variation among cultured cells, which could conceivably cause metabotypic variation. It also exemplifies that heterogeneity is present in human tumor cell lines, even when obtained from repositories. The authors estimate that SW480 cells have been cultured for at least 150 generations; the MDA-MB-468 cell line used in the present study was established around the same time (1977) as the SW480 cell line and has likely undergone comparable expansion. Notably, these cell lines are established from a tumor bulk that is generally assumed to be syngeneic. However, clonal genome, genetic and phenotypic variation is well known within solid tumors and derived cell lines [Enger et al. 1986; Nistér et al. 1986; Meyer et al. 2015; Greaves et al. 2012; Marusyk et al. 2010; McGranahan et al. 2017; Hastings et al. 1983]. Other studies have also shown variation among individual cells from a cultured population, such as in the growth rate of T lymphoblastoid Molt3

clones selected from a population obtained from ATCC [Tomelleri at al. 2008] and in various phenotypes of MCF-7 breast cancer line clones [Baguley et al. 2011]. Famously, HeLa cell lines (from a single origin) have been shown to display marked heterogeneity across different labs even when cultured side-by-side [Frattini et al. 2015; Liu et al. 2018]. Indeed, Stockholm et al. described the emergence of phenotypic heterogeneity in cultured, clonal mammalian cells as "a rule rather than exception" [Stockholm et al. 2007].

Furthermore, Xu et al. showed significant chromosomal variation in single cell clones selected from a parental MDA-MB-468 population, the same cell line used for this study [Xu et al. 2008]. In the karyotyping of 27 cells, they found a range 45-65 chromosomes, including significant chromosomal aberrations. Importantly, these numbers varied from previous reports, which they interpreted as an indication of genomic instability in this cell line. Notably, the ATCC listing for MDA-MB-468 recognizes the presence of karyotype variation among the population. Abdallah et al. show that cell lines exhibiting genomic instability can vary in a short time frame (1-2 months) and emphasize that in these cases considering an average of a heterogeneous population can be misleading [Abdallah et al. 2013]. Conceivably, genomic instability and the common occurrence of chromosomal variation in clones from MDA-MB-468 could lead to cell-specific metabotypes.

The fact that MDA-MB-468 is aneuploid provides another possible explanation for clonal variation. Initially, we assume that there would be two copies of each gene of interest and therefore performed *TA cloning* until 2 allele sequences were confirmed. However, should the gene be present additional times in certain clones of this cell line, there could be remaining yet-undetected wild-type or edited but in-frame

alleles of the protein in some cell lines. Should one cell line be a true knockout and the comparison lines contain a functional allele, metabolic differences would be expected.

We therefore explored this possibility using Next-Generation Amplicon Sequencing of a PCR amplification of cDNA that is reverse transcribed from cell line-derived RNA (to obtain mRNA sequences). This sequencing revealed that for all cell lines there were no or minimal sequencing reads that could be considered a third allele (Table 3.5). The very minor representation of sequence deviations in SLC25A40 KO cell lines are most likely due to splicing, as in each case a region is removed that spans exactly to an exon start site. The other sequences in SLC25A38 KO2 and SLC25A39 KO1 are unlikely to be due to a third allele, considering the very low representation in mRNA sequencing reads and the lack of detection by *TA cloning*. There are a number of other explanations for these findings, such as RNA splicing, PCR or sequencing artifacts, minor contamination with another cell line (clonal selection error) and continued CRISPR-mediated cutting during clonal colony growth. The latter possibility is supported by the fact that these two cell lines are the only two where a frame-shift edit is a 1bp insertion. This could leave sufficient homology to the original gRNA recognition site for a second editing event if Cas9 was still expressed during growth of the clonal colonies. This would cause a minor subpopulation of cells in that colony to have the re-edited allele.

More importantly, this mRNA sequencing revealed prevalent mRNA splicing as an unexpected consideration. For each gene, alternative splice sites were identified (Figure 3.14). With the exception of those for SLC25A39, these splice sites have not been previously reported. In the case of SLC25A38, both cut sites sit on top of possible splice regions. However, in this case, these splice variants overwhelmingly

result in out-of-frame sequences, which should not affect the validity of the knockout. However, in two cell lines (SLC25A39 KO2 and SLC25A40 KO1), the combination of the splice variants and the CRISPR edit results in in-frame sequences that are significantly represented in the mRNA (Table 3.6). As these could possibly lead to functional (albeit edited) transporter, the validity of these two knockout cell lines are questionable.

In the case of SLC25A39, this prospect of functional protein in KO2 is a possible explanation for the irreproducibility of results seen in SLC25A39 KO1 (mRNA from KO3 could potentially be sequenced to further assess this possibility).

For SLC25A40, if it is assumed that KO1 has functional protein, the metabolic variation from the parental population is a concern (as the observed decrease in betaine would not be due to protein knockout). On the other hand, if this KO clone does not have functional protein, the irreproducibility in KO2 remains an issue. Additionally, the metabolic variation observed in the case of the SLC25A38 KO cell lines is not explained by the sequencing results, as sequence data indicates both cell lines to be knockouts. Therefore, in the end, the issue of clonal metabolic variation unrelated to the function of the protein of interest persists.

A possible solution to the issue of comparing a single cell-derived cell line to the original population average would be to select a single clone to expand before producing knockout cell lines. This strategy would allow for a comparison of a knockout clone to a syngeneic control rather than to a population average of clonal variants. This would ensure that the knockout and control cell line would be as similar

as possible (fewer passages between splitting the lineages), providing for a 'cleaner' comparison.

Another possible explanation for the differences detected between clonal cell lines stems from differences in compensation for the loss of transporter function. Conceivably, knockout of very important proteins would force cells to compensate metabolically for the loss in order to live, likely by altering the expression of other proteins [El-Brolosy et al. 2017]. Should there be multiple compensatory mechanisms that enable survival of a given *SLC25* knockout cell line, one would anticipate the selection of multiple clonal variants with marked differences in metabolite abundances. Regardless of the cause, this study highlights that two independently produced CRISPR knockout cell lines should always be analyzed initially to avoid being misled by knockout-independent phenotypic differences.

The goal of studying knockout cell lines was to observe profound metabolite changes due to the knockout that could be traced back to aberrations in mitochondrial transporter function. Part of the problem in achieving this goal was that, for the most part, metabolite changes in the knockout cell lines were small. This leads to low confidence in the metabolite changes being due to loss of the transporter protein and difficulty in selecting high-confidence species to focus on for follow-up experiments. Conceivably, the SLC25 transporter proteins studied may be non-essential due to functional redundancy, compensation, or cell type-specific metabolism that make a given knockout cell lines a poor system for de-orphanizing mitochondrial transporters. Notably, the selection of study cell type could be a key issue, as certain transporters can be critically important in some cell types, while completely dispensable in others. As an example, betaine is an important osmoregulator kidney cells [Grunewald et al.

1995], but may have little role in the breast cancer epithelial cells, such as the MDA-MB-468 line studied herein. Therefore, a choline mitochondrial transporter knockout would predictably have more profound metabolic effects in kidney cells than breast cancer cells. Unfortunately, when studying transporters of unknown function it is impossible to know which cell type the transporter is most functionally relevant to. An educated guess should be made based on information known about homologous transporters or the suspected class of the substrate.

Another challenge highlighted by this study is that of investigating the function of mitochondrial proteins within the context of whole cell metabolism. This challenge is compounded by the generic problem of studying a protein with unknown function in an untargeted manner. Although untargeted metabolite profiling can be a powerful tool, the combination of so many unknowns in a complex system (i.e., global cellular metabolism) will predictably lead to difficult-to-interpret results. The issue with studying mitochondrial transporters in a whole cell context is that even if profound metabolic changes are observed, they could be significantly downstream metabolic effects that would be very difficult to trace to a specific substrate for a given SLC25 transporter. Instead, studying short-term transport in isolated mitochondria from knockout cells could be considered as an alternative approach that affords more readily interpretable results. However, this strategy would engender new methodological challenges for rapid isolation of large numbers of functional mitochondria or mitochondria that retain their metabolite profiles despite extensive purification procedures.

As previously mentioned, during the course of this study SFXN1 was reported to serve as a mitochondrial serine transporter [Kory et al. 2018]. Importantly, this

report recognized the possibility of functional redundancy and there may indeed be an SLC25 family transporter involved in serine transport. In this paper, the investigators took a more tailored approach towards identifying the serine transporter. Assuming that the serine transporter would display synthetic lethality in the absence of serine hydroxymethyl transfrase-1 (SHMT1), the authors performed a CRISPR knockout screen targeting 3000 proteins in SHMT1 knockout cells to identify candidate proteins. Additionally, in order to reduce possible confounding effects of redundant mechanisms, they depleted serine from the media, thereby sensitizing cells to the knockout of a potential serine transporter. This strategic approach led to successful identification of a serine transporter and reveals the benefit of a more targeted knowledge-driven discovery approach for identifying transport proteins with a specific function.

Lactococcus Lactis transporter expression studies

Due to the challenges with eukaryotic cell studies, we sought a simpler and more confident model system for establishment of SLC25-type transporter functions. In this regard, we became aware that the bacteria *Lactococcus Lactis* has previously been found to be a suitable host for the inducible expression of functional membrane proteins. In fact, functional mitochondrial transporters have been expressed a number of times in these bacteria, with very little degradation and no detectable inclusion body formation. Monné et al. showed expression of 11 functional yeast SLC25 family mitochondrial transporters [Monné et el. 2005]. Included in this study were the homologs of SLC25A1, SLC25A10 and SLC2511, three of the carboxylate transporters, which were demonstrated to be functional in proteoliposome-based transport studies. Human SLC25A10 and SLC25A11 have also been previously

expressed in *Lactoccoccus Lactis* and were functional in proteoliposomes [Booty et al. 2015]. Due to the fact that bacteria do not have mitochondria, these transporters are localized to the bacterial outer cell membrane. This membrane protein localization obviates the challenge of studying mitochondrial metabolism within the context of an intact eukaryotic cell.

Although transport assays using *Lactococcus Lactis*-expressed membrane proteins have previously been done with isolated proteins resuspended in artificial liposomes, a number of reports have successfully shown transport into intact bacteria cells. Most notably, the mouse mitochondrial pyruvate carrier (MPC) complex (co-expression of MPC1 and MPC2 proteins) [Herzig et al. 2012] and the human mitochondrial ADP/ATP carrier (SLC25A4) were characterized in intact *L. Lactis* transport studies [Mifsud et al. 2013]. Similarly, transport assays with intact bacteria were also successfully used to study the riboflavin transport protein (RibU) [Burgess et al. 2006], the Arabidopsis ATP/ADP translocator (AtAATP1) [Frelet-Barrand], the *L. Lactis* di- and tripeptide transport protein (DtpT) [Hagting et al. 1994], sterol transport by the Human Breast Cancer Protein (BRCP) [Janvilisri et al. 2003], a yeast mitochondrial ATP/ADP carrier (AAC3), the *Neocallimastix* hydrogenosomal ADP/ATP carrier (AAC_{hyd}) [Kunji et al. 2003], the *Streptococcus Thermophilus* lactose transporter (LacS- ΔIIA) and *L. Lactis* branched-chain amino acid permease (BcaP) [Linares et al. 2010]. Due to the simplicity of this system, and the conspicuous advantages it affords for study of mitochondrial transporters, we set out to use this model system in combination with mass spectrometry as a novel platform to assess function for the SLC25A family of mitochondrial transporters.

Initially, *Lactococcus Lactis* were engineered to express a pNZ8048 plasmid that encodes *SLC25A38*, *SLC25A39*, *SLC25A40* or *SLC25A44*. Later, bacterial lines containing FLAG-tagged version of these proteins were also produced in to facilitate the detection of nisin-induced transporter expression. Expression of all four proteins were detected by anti-FLAG western blot analysis. Whereas, SLC25A40 and SLC25A44 showed robust expression, levels of SLC25A38 and SLC25A39 were relatively lower (Figure 3.15). An untargeted metabolite profiling experiment was performed on bacterial metabolite extracts collected after 3 hours of SLC25 protein expression of each of the above proteins (and an empty vector control). These cell lines displayed profoundly different metabotypes and generally appeared to fall into two groups. Surprisingly, this indicated that general metabolism for each cell line was distinct, possibly due to cell density (induction of protein expression can increase or decrease bacterial growth rate, leading to different densities/metabolic states). This meant that in comparison to the control cell line there could be hundreds of differentially-abundant metabolites for each protein overexpression line, challenging the interpretation of SLC25 cargo function. Therefore, in order to identify the most profound changes that were unique to each cell line, and therefore more likely due to transporter function, the metabolite abundances of each cell line were compared to the aggregate of all the others as reference "control" group. This indeed identified a number of untargeted features with extremely increased abundance in certain transporter-expressing cell lines (Figure 3.16). In experiments performed using both FLAG-tagged and non-tagged SLC25 transporter cell lines, identical features were differentially abundant, indicating that observed changes were indeed related to transporter protein expression. These features could indicate compounds that are

transported by the overexpressed protein, but may alternatively arise as peptide degradation products of the overexpressed protein. In support of the latter possibility, many of the m/z values for differentially-expressed features were found to match values predicted for polypeptides. Also, metabolite profiling of culture media in which these transporter expressing lines were grown indicated no change for these features (for the few that were detectable in the media) and many were not detectable in the media. This supports the possibility that the differentially-expressed features were not metabolites being transported, but rather other species that arose from SLC25 overexpression (such as peptides).

Moving forward, in order to avoid these confounding issues involved with studying uncharacterized transporters with untargeted profiling, we decided to take a more targeted approach. Here, in addition to the above transporters, we set our sights on the transporter of 2-hydroxyglutarate, a recognized cancer-driving metabolite (oncometabolite) produced from 2-oxoglutarate by several alternative enzymes [Dang et al. 2017; Ye et al. 2018]. Notably, 2-hydroxyglutarate is often synthesized in the cytosol and must be transported into mitochondria for metabolism, but the cognate mitochondrial transporter has not been identified [Engqvist et al. 2014]. Based on homology, there are 12 SLC25 family transporters that have been shown to or are likely to transport carboxylic acids (Figure 3.1, from [Palmieri 2013]). We hypothesized that one or more of these 12 transporters conducts 2-hydroxyglutarate.

Gene fragments for each of these SLC25 proteins were codon optimized for *Lactococcus Lactis* expression and bacterial cell lines were produced for FLAG-tagged versions of these transporter. Western blots showed successful protein expression for at least ten out of the twelve transporters (Figure 3.17).

Due to the fact that the previously employed extraction protocol required extensive wash time in phosphate-buffered saline (PBS) and water, an experimental approach was developed for more rapid extraction followed by targeted LC/MS-based analysis.

In terms of experiment design, cells lines were grown in larger quantities for protein induction. Bacteria were then normalized to culture density, collected by centrifugation, washed, resuspended in HEPES buffer and split into multiple samples. The metabolite of interest would then be added to replicate bacterial aliquots and incubated for the desired time at 30°C, followed by cell harvesting and metabolite extraction. After the desired incubation time, bacteria were simultaneously harvested on glass fiber filter paper and rapidly washed using a Brandel M-24T cell harvester. This harvesting methodology has traditionally been used for radiolabel-based studies. The goal for this approach was to avoid metabolite leakage during the bacterial wash steps, followed by bead-beating of bacteria on filter paper for metabolite extraction and analysis. Should this experiment design and metabolite extract protocol suffice for transport assays, this would provide a novel approach to study small molecule transport in a bacterial model system. As opposed to radioactivity based assays, LC/MS-based approach used would allow the possibility of simultaneously screening a multitude of compounds for transport.

To establish this system for studying metabolite transport in *Lactococcus Lactis*, positive control experiments were attempted using transporters with known cargo, including: SLC25A11 (2-oxoglutarate transporter), SLC25A21 (2-oxoglutarate transporter), SLC25A38 (glycine transporter) and SLC25A8 (aspartate transporter).

SLC25A11 is an 2-oxoglutarate/malate antiporter, and due to the structural similarity between 2-oxoglutarate and 2-hydroxyglutarate, it was a candidate for 2-hydroxyglutarate transport. In a first experiment, the time course of 2-oxoglutarate uptake was assessed using 200 µM 2-oxoglutarate added to the resuspended bacterial aliquots – an empty vector time course was performed concurrently as a negative control. Under these conditions, no evidence of 2-oxoglutarate transport was detected (Figure 3.18). To determine if there is an optimal concentration to detect transport, we did a follow up experiment at a 30 min time point where we screened a range of 2-oxoglutarate concentrations for an indication of transport (2 µM, 20 µM, 200 µM, and 2 mM). Various concentrations of malate were also tested for potential transport (20 µM, 200 µM, and 2 mM). Although the results were not conclusive, there was an indication of 2-oxoglutarate transport when studied at its highest concentration (i.e., 2 mM; Figure 3.19). Although this was later assumed to be due to washing inefficiency, 2 mM substrate was used as concentration for all following experiments.

2mM 2-oxoglutarate was tested again at a 30min time point with more replicates and a control cell line. The samples with added 2-oxoglutarate once again showed large variation, suggesting a potential issue with washing away the external 2-oxoglutarate (Figure 3.20). Therefore, we tested using the effects of up to six 3-second washes on levels of quantified 2-oxoglutarate, potentially informing on either under-washing (i.e., unincorporated metabolite) or over-washing (i.e., metabolite leakage). The findings suggested that 2 washes were sufficient, and that extensive washing did not cause significant additional substrate loss (Figure 3.21). Four washes were used in subsequent experiments.

The previous experiment with 2-oxoglutarate wash repeated with more careful attention to vortexing bacteria during wash steps (to break up clumps), vortexing after metabolite addition (to ensure adequate mixing), allowing cell harvester probes to only minimally submerge during washing (to avoid leakage from the probes during washing). Also, ice cold water was used for washes. Results indicated that 2-oxoglutarate was indeed being transported into the bacteria, despite a lack of clear evidence that SLC25A11-FLAG overexpressing bacteria had increased transport (Figure 3.22). The same experiment done with SLC25A21-FLAG overexpressing bacteria showed no increased transport with protein expression (Figure 3.26).

In these preliminary experiments, unlabeled 2-oxoglutarate was used. Ideally, heavy-isotope labeled substrate could be used to better differentiate transported 2-oxoglutarate from endogenous sources of 2-oxoglutarate. Therefore, transport assays were done using deuterated [6D]-2-oxoglutarate. As a standard, this compound is primarily detected as [4D]-2-oxoglutarate due to the rapid exchange of two deuterium atoms. However, during the extraction and drying process, additional deuterium atoms can apparently be exchanged resulting in detection of both [2D] and [3D] forms of 2-oxoglutarate. In general, [4D] and [3D] were the most abundant forms and their abundance generally follows the same pattern between the samples. Although transport of labeled 2-oxoglutarate was detectably present in the bacteria, at 5min and 14min time points no differences was detected between SLC25AA11-FLAG overexpression bacteria and empty vector bacteria (Figure 3.23). Interestingly, no difference was observed between the time points, possibly indicating that transport reached steady-state saturation.

Similar experiments were attempted using SLC25A38-FLAG and SLC25A40-FLAG overexpressing bacteria, studying the potential for uptake of [2-^{13}C]-glycine. Note that SLC25A40 was tested for glycine transport because it is homologous to SLC25A38 and additionally exhibits far more robust expression in *Lactococcus Lactis*. These studies revealed no overexpression mediated transport by either SLC25 carrier protein (Figure 3.27A, Figure 3.28). However, tight data and a linear increase in [2-^{13}C]-glycine abundance validated the robustness of the extraction methodology and indicated that appropriate time points were being used (i.e. transport was not saturated), allowing detection of potential differences in glycine transport rate.

Until this point, C-terminus FLAG tagged proteins were being expressed to enable convenient detection of transporter protein expression. However, the addition of a tag sequence could affect protein function or localization, contributing to the observed lack of expressed protein function. Therefore, we additionally engineered *Lactococcus Lactis* bacterial lines to express non-tagged forms of SLC25A11, SLC25A38 and SLC25A8, with the assumption that the removal of the 8 amino acid tag on the C-terminus would not detrimentally affect protein expression. SLC25A8 was also included as it displayed robust protein expression by western blot and was shown to transport aspartate [Vozza et al. 2014], which is readily available as a heavy isotope-labelled species.

In a first experiment with the SLC25A11 overexpressing bacterial line, there was a trend of increased [6D]-2-oxoglutarate transport at 2 min, although there were not enough samples per group to be convincing (Figure 3.24). Once again, no increase in 2-oxoglutarate was observed between 2 min and 5 min, possibly indicating rapid saturation of transport. Therefore, the experiment was repeated with a 1 min time point

and more replicates. However, this experiment was inconclusive, as there was seemingly samples that did not efficiently wash and the other samples showed wide variation (Figure 3.25).

Transport assays with SLC25A38 and SLC25A8 overexpressing bacteria and glycine and aspartate, respectively, were performed. Once again, no protein function was detected although linear import and highly reliable replicates were observed in both cases (Figure 3.27B; Figure 3.29).

In the end, *Lactococcus Lactis* seems to be a suitable system for the expression of many mitochondrial transporters, although it may not be suitable for intact bacteria functional assays. Expression of 11/12 carboxylate or UCP and 4/4 glycine-related SLC25 family transporters were detectable, albeit with varying levels of expression. Protein expression could likely be increased for some of the transporters by varying the concentration of the inducer, nisin. Also, low protein abundance may not be an issue for studying the proteins, depending on the assay sensitivity (protein transportation rate compared to noise).

A transport assay protocol, including simultaneous, rapid, filter-based wash, followed by metabolite extraction for mass spectrometry was established. This assay was found to be reliable (despite the need to quickly normalize based on cell culture density, to remove salts and impurities for mass spectrometry and to extract metabolites from bacteria imbedded in a glass-fiber filter), as exemplified by the glycine and aspartate transport experiments, where replicates were tight, even across different cell lines, and linearly time-dependent. Of course, a high level of endogenous transport may mean that these compounds are not good candidates for study in these

bacteria. However, sufficient exogenous protein activity could possibly overcome this issue.

2-oxoglutarate at times displayed a puzzling combination of compound-specific highly variant data without evidence of time-dependent transport. Although this was initially taken to mean rapid saturation of transport, it may also indicate that 2-oxoglutarate transport into the cells was not being observed at all. In this case, detected deuterated 2-oxglutarate would be a background signal. It unclear where this would come from, as the water washes seem to be sufficient and there is no background in samples not treated with deuterated 2-oxoglutarate. However, support for the view that no transport was being observed comes from Monné et el., where no/low background for 2-oxoglutarate was observed in proteoliposomes prepared from membrane isolates of control Lactoccous Lactis (which should retain endogenous transporters) [Monné et el. 2005].

In any case, expressed protein-mediated transport was, unfortunately, not detected, which may indicate that these transporters are not suitable for study in intact *Lactococcus Lactis*. There could be a number of explanations, such as insufficient abundance of internal substrates for antiport (seemingly unlikely due to detection of endogenous transport), improper transport conditions (e.g. pH, substrate concentration, membrane potential) insufficient protein abundance (e.g. SLC25A38), incorrect protein localization, or an unsuitable membrane lipid composition. This suggests that functional assessment of SLC25 transporters in reconstituted proteoliposomes, rather than intact bacteria, may be unavoidable. Although this would add significantly to assay complexity and diminish throughput, it could still lead to significant discovery and would be a worthwhile next step. In this regard, human SLC25A10 and

SLC25A11 have already been shown to be functional in proteoliposome studies. If the isolation and wash protocol described above can be suitably adapted to proteoliposomes, mass spectrometry centered studies may be more effectively used to characterize SLC25 mitochondrial transporters expressed in *Lactococcus Lactis*.

REFERENCES

Abdallah BY, Horne SD, Stevens JB, et al. Single cell heterogeneity: Why unstable genomes are incompatible with average profiles. Cell Cycle. 2013;12(23):3640-3649.

Achouri Y, Noel G, Vertommen D, Rider MH, Veiga-Da-Cunha M, van Schaftingen E. Identification of a dehydrogenase acting on D-2-hydroxyglutarate. Biochem J. 2004;381(1):35-42.

Baguley BC, Leung E. Heterogeneity of Phenotype in Breast Cancer Cell Lines. In: Breast Cancer - Carcinogenesis, Cell Growth and Signalling Pathways.Vol 5. InTech; 2011:89.

Booty LM, King MS, Thangaratnarajah C, et al. The mitochondrial dicarboxylate and 2-oxoglutarate carriers do not transport glutathione. *FEBS Lett*. 2015;589(5):621-628.

Burgess CM, Slotboom DJ, Geertsma ER, Duurkens RH, Poolman B, Van Sinderen D. The riboflavin transporter RibU in *Lactococcus lactis*: Molecular characterization of gene expression and the transport mechanism. *J Bacteriol*. 2006;188(8):2752-2760.

Chen H, Lu Q, Zhang Y, Zhang C, Zhang H, Xu H. Overexpression of SLC25A38 protein on acute lymphoblastic leukemia cells. *Oncol Lett*. 2014;7(5):1422-1426.

Chen Q, Park HC, Goligorsky MS, Chander P, et al. Untargeted plasma metabolite profiling reveals the broad systemic consequences of xanthine oxidoreductase inactivation in mice. *PLoS One.* 2012;7(6):e37149.

Dang L, Su S-SM. Isocitrate Dehydrogenase Mutation and (R)-2-Hydroxyglutarate: From Basic Discovery to Therapeutics Development. Annu Rev Biochem. 2017;86(1):305-331.

El-Brolosy MA, Stainier DYR. Genetic compensation: A phenomenon in search of mechanisms. *PLoS Genet*. 2017;13(7):1-17.

Enger MD, Tesmer JG, Travis GL, et al. Clonal variation of cadmium response in human tumor cell lines Clonal variation of cadmium response in human tumor cell lines. Am J Physiol. 1986;(15):256-263.

Engqvist MKM, Eßer C, Maier A, Lercher MJ, Maurino VG. Mitochondrial 2-hydroxyglutarate metabolism. *Mitochondrion*. 2014;19(PB):275-281.

Fernández-Murray JP, Prykhozhij S V, Dufay JN, et al. Glycine and Folate Ameliorate Models of Congenital Sideroblastic Anemia. *PLoS Genet*. 2016;12(1).

Fiermonte G, Dolce V, Palmieri L, et al. Identification of the Human Mitochondrial Oxodicarboxylate Carrier. *J Biol Chem*. 2001;276(11):8225-8230.

Frattini A, Fabbri M, Valli R, et al. High variability of genomic instability and gene expression profiling in different HeLa clones. Sci Rep. 2015;5:1-9.

Frelet-Barrand A, Boutigny S, Moyet L, et al. *Lactococcus lactis*, an alternative system for functional expression of peripheral and intrinsic arabidopsis membrane proteins. *PLoS One*. 2010;5(1).

Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012;481(7381):306-313.

Grunewald RW, Eckstein A. Osmotic regulation of the betaine metabolism in immortalized renal cells. *Kidney Int*. 1995;48(6):1714-1720.

Guernsey DL, Jiang H, Campagna DR, et al. Mutations in mitochondrial carrier family gene SLC25A38 cause nonsyndromic autosomal recessive congenital sideroblastic anemia. *Nat Genet*. 2009;41(6):651-653.

Hagting A, Kunjis ERS, Leenhoutsg KJ, Poolmanm B. The Di- and Tripeptide Transport Protein of *Lactococcus lactis*. J Biol Chem. 1994;269(15):11391-11399.

Hastings RJ, Franks LM. Cellular heterogeneity in a tissue culture cell line derived from a human bladder carcinoma. Br J Cancer. 1983;47(2):233-244.

Herzig S, Raemy E, Montessuit S, et al. Identification and functional expression of the mitochondrial pyruvate carrier. *Science (80-)*. 2012;336(6090):93-96.

Humm A, Fritsche E, Steinbacher S, Huber R. Crystal structure and mechanism of human L-arginine:glycine amidinotransferase: A mitochondrial enzyme involved in creatine biosynthesis. *EMBO J*. 1997;16(12):3373-3385.

Indiveri C, Palmieri F, Bisaccia F, Krämer R. Kinetics of the reconstituted 2-oxoglutarate carrier from bovine heart mitochondria. *BBA - Bioenerg*. 1987;890(3):310-318.

Ismailoglu I, Chen Q, Popowski M, Yang L, et al. Huntingtin protein is essential for mitochondrial metabolism, bioenergetics and structure in murine embryonic stem cells. *Dev Biol.* 2014;391(2):230-40.

Janvilisri T, Venter H, Shahi S, Reuter G, Balakrishnan L, Van Veen HW. Sterol transport by the human breast cancer resistance protein (ABCG2) expressed in *Lactococcus lactis*. *J Biol Chem*. 2003;278(23):20645-20651.

Kory N, Wyant GA, Prakash G, et al. SFXN1 is a mitochondrial serine transporter required for one-carbon metabolism. *Science (80-)*. 2018;362(6416):eaat9528.

Kunji ERS, Slotboom DJ, Poolman B. *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochim Biophys Acta - Biomembr*. 2003;1610(1):97-108.

Linares DM, Geertsma ER, Poolman B. Evolved *Lactococcus lactis* strains for enhanced expression of recombinant membrane proteins. *J Mol Biol*. 2010;401(1):45-55.

Liu Y, Mi Y, Mueller T, et al. Genomic, Proteomic and Phenotypic Heterogeneity in HeLa Cells across Laboratories: Implications for Reproducibility of Research Results. bioRxiv. 2018:307421.

Locasale JW. Serine, glycine and one-carbon units: Cancer metabolism in full circle. *Nat Rev Cancer*. 2013;13(8):572-583.

Lucas V, Webster T. Benzoyl-Coenzyme Coenzyme A : Glycine Mitochondria A : Glycine from and Phenylacetyl- Bovine Liver. *J Biol Chem*. 1979;254(15):7230-7237.

Marusyk A, Polyak K. Tumor heterogeneity: Causes and consequences. Biochim Biophys Acta - Rev Cancer. 2010;1805(1):105-117.

McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. Cell. 2017;168(4):613-628.

Meyer M, Reimand J, Lan X, et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. Proc Natl Acad Sci. 2015;112(3):851-856.

Michel V, Bakovic M. The solute carrier 44A1 is a mitochondrial protein and mediates choline transport. *FASEB J*. 2009;23(8):2749-2758.

Mifsud J, Ravaud S, Krammer EM, et al. The substrate specificity of the human ADP/ATP carrier AAC1. *Mol Membr Biol*. 2013;30(2):160-168.

Monné M, Chan KW, Slotboom D, Kunji ERS. Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. *Protein Sci*. 2005;14(12):3048-3056.

Monné M, Daddabbo L, Gagneul D, et al. Uncoupling proteins 1 and 2 (UCP1 and UCP2) fromArabidopsis thaliana are mitochondrial transporters of aspartate, glutamate, and dicarboxylates. *J Biol Chem*. 2018;293(11):4213-4227.

Monné M, Miniero DV, Bisaccia F, Fiermonte G. The mitochondrial oxoglutarate carrier: From identification to mechanism. *J Bioenerg Biomembr*. 2013;45(1-2):1-13.

Nistér M, Heldin CH. Clonal Variation in the Production of a Platelet-derived Growth Factor-like Protein and Expression of Corresponding Receptors in a Human Malignant Glioma. Cancer Res. 1986;46(1):332-340.

Obeid R. The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients*. 2013;5(9):3481-3495.

Palmieri F. The mitochondrial transporter family SLC25: Identification, properties and physiopathology. *Mol Aspects Med*. 2013;34(2-3):465-484.

Palmieri F. Mitochondrial transporters of the SLC25 family and associated diseases: A review. *J Inherit Metab Dis*. 2014;37(4):565-575.

Palmieri F, Monne M. Discoveries, metabolic roles and diseases of mitochondrial carriers: A review. *Biochim Biophys Acta - Mol Cell Res*. 2016;1863(10):2362-2378.

Palmieri F, Quagliariello E, Klingenberg M. Kinetics and Specificity of the Oxoglutarate Carrier in Rat-Liver Mitochondria. *Eur J Biochem*. 1972;29(3):408-416.

Palmieri L, Agrimi G, Runswick MJ, Fearnley IM, Palmieri F, Walker JE. Identification in Saccharomyces cerevisiae of two isoforms of a novel mitochondrial transporter for 2-oxoadipate and 2-oxoglutarate. *J Biol Chem*. 2001;276(3):1916-1922.

Rzem R, Van Schaftingen E, Veiga-Da-Cunha M. The gene mutated in L-2-hydroxyglutaric aciduria encodes L-2-hydroxyglutarate dehydrogenase. Biochimie. 2006;88(1):113-116.

Seigneurin-Berny D, King MS, Sautron E, et al. *Heterologous Expression of Membrane Proteins*.; 2010.

Seigneurin-Berny D, King MS, Sautron E, et al. Membrane Protein Production in *Lactococcus lactis* for Functional Studies. In: Mus-Veteau I, ed. *Heterologous*

Expression of Membrane Proteins: Methods and Protocols. New York, NY: Springer New York; 2016:79-101.

Stockholm D, Benchaouir R, Picot J, et al. The origin of phenotypic heterogeneity in a clonal cell population in vitro. PLoS One. 2007;2(4).

Tomelleri C, Milotti E, Dalla Pellegrina C, et al. A quantitative study of growth variability of tumour cell clones in vitro. Cell Prolif. 2008;41(1):177-191.

Ueland PM, Holm PI, Hustad S. Betaine: A key modulator of one-carbon metabolism and homocysteine status. *Clin Chem Lab Med*. 2005;43(10):1069-1075.

Vozza A, Parisi G, De Leonardis F, et al. UCP2 transports C4 metabolites out of mitochondria, regulating glucose and glutamine oxidation. *Proc Natl Acad Sci*. 2014;111(3):960-965.

Xu J, Chambers AF, Tuck AB, Rodenhiser DI. Molecular cytogenetic characterization of human breast cancer cell line MDA-MB-468 and its variant 468LN, which displays aggressive lymphatic metastasis. Cancer Genet Cytogenet. 2008;181(1):1-7.

Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer Cell*. 2011;19(1):17-30.

Ye D, Guan KL, Xiong Y. Metabolism, Activity, and Targeting of D- and L-2-Hydroxyglutarates. Trends in Cancer. 2018;4(2):151-165.

Zhang CY, Baffy G, Perret P, et al. Uncoupling protein-2 negatively regulates insulin secretion and is a major link between obesity, β cell dysfunction, and type 2 diabetes. *Cell*. 2001;105(6):745-755.

Zhang CY, Parton LE, Ye CP, et al. Genipin inhibits UCP2-mediated proton leak and acutely reverses obesity- and high glucose-induced β cell dysfunction in isolated pancreatic islets. *Cell Metab*. 2006;3(6):417-427.

Zhang XH, Tee LY, Wang XG, Huang QS, Yang SH. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol Ther - Nucleic Acids*. 2015;4(11):e264.

Zong C, Lu S, Chapman AR, Xie XS. Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. Science (80-). 2012;338(December):1622-1627.

NEXT STEPS

Chapter 1

The endocannabinoid system is an integral signaling system in mammals and contributes to a wide array of physiological and pathological topics, such as learning, memory, metabolism and hunger, depression, anxiety and drug addiction. Further research into the characterization of this important group of receptors and related ligands will shed light on the various roles of the endocannabinoid system and may allow the biomedical research community to take advantage of it for therapeutic gain. In our research, we showed that the endocannabinoid system is important in the acute metabolic response to alcohol withdrawal in mice. Future experiments to confirm results in human studies are warranted and further mouse-centered studies could focus on endocannabinoid fluctuations over the entire addiction time course (before, during and after alcohol addiction), as well as the mechanisms of potential alcohol-abuse therapeutics in relation to the endocannabinoid system. Further metabolic characterization of the endocannabinoid system in alcohol dependence and withdrawal using mass spectrometry-based metabolomics could provide mechanistic insights into addiction and withdrawal, and may be useful in developing novel therapies. Our research could also be expanded to survey endocannabinoid levels in response to use of and withdrawal from other drugs of addiction. Additionally, the endocannabinoid system is known to have an important actions in the periphery, in addition to the brain (e.g. immune cell function). Mass spectrometry-based studies of other cannabinoid-regulated tissues would provide a more comprehensive understanding of the dynamics of cannabinoid receptor signaling in genetic mouse models and diseases. Other areas

of endocannabinoid research would benefit from the application of metabolite profiling to assist interpretation of behavioral data at a more granular, metabolic level, such as studies focused on the common human C385A fatty acid amide hydrolase polymorphism known to be related to stress and alcohol abuse tendency, the effects of cannabis on the endocannabinoid system or the dynamics of the endocannabinoid system during development. These topics could all be studied using the metabolomics methods discussed in our work.

Chapter 2

Amyotrophic lateral sclerosis remains a poorly understood and devastating disease. Further metabolism-centered studies are warranted in attempt to gain insight into the molecular mechanisms of this disease and its apparent variants/subtypes. We showed that metabolic subgroups can be identified among sALS patients, specifically revealing the sALS-1 subgroup. Future research should focus on the continued characterization of sALS-1 patients, seeking to further validate our results in a separate cohort, using metabolomics to probe the mechanisms of disease in patient-derived fibroblasts. These experiments will likely include surveying the cell lines in stress conditions for differential response in the sALS-1 subgroup, which could help in clarifying the source(s) of the metabolic perturbations. Metabolite profiling and stable isotope tracing would shed light on the metabolic effects causing the observed phenotype. Also, motor neurons could be produced from the patient fibroblasts using induced pluripotent stem cell technology and would provide the most disease-relevant system for metabolic characterization of the subgroup. This technology could also be used to determine if the causes of the metabotype are genetic or epigenetic by

reprogramming the motor neurons back into fibroblasts, as epigenetic markers are wiped during the conversion process.

Generally, our work indicates that larger metabolite profiling studies are warranted to pursue other sALS subgroups. Importantly, machine learning-based subgrouping endeavors gain significant power with increased sample size. A larger, collaborative metabolic characterization of thousands of sALS patient-derived samples would likely reveal additional, less common subgroups exhibiting a shared metabotype.

Chapter 3

Further work is needed to establish a system/methodology for the reliable characterization of SLC25 family mitochondrial transporters. In regard to the existing MDA-MB-468 human breast cancer knockout cell lines, cell culture scenarios could be screened to identify conditions in which the missing transporter becomes more important for cell function, predictably exacerbating metabolic abnormalities associated with the knockout metabotype. These conditions may be selected based on hypothesized transport cargo, such as testing for auxotrophic growth-dependence on the presence of serine and/or glycine in culture medium when trying to identify a suspected serine transporter. Ideally, a growth, cell morphology or other noticeable phenotype would be observed in knockout cells, compared to control, which would allow for rapid identification of conditions and pathways worth studying with a more focused metabolic profiling effort. This is an example of an approach to screen for protein knockout-related phenotypes or pathways of interest that could then be followed up with using metabolomics. These sorts of preliminary screen-type

experiments could help focus the study and avoid the confusion and misdirection that can come from untargeted metabolomics when used as a first approach. Another more focused discovery approach may include a CRISPR knockout screen specifically designed to identify transporter candidates for a specific compound, as was done in studies that identified SFXN1 as a serine transporter. In any case, any further studies using knockout cell lines should include a second knockout cell line produced using a different guide RNA and a compensated knockout (reinsertion) cell line from the beginning.

The study of mitochondria isolated from these knockout cell lines could also be considered to remove the issue of studying mitochondrial metabolism within the background of cytosolic metabolites. However, this introduces significant consideration and would require careful protocol establishment and validation. There are two likely approaches to studying isolated mitochondria. The first would be to endeavor to study the endogenous metabolites present in the mitochondria. Here, mitochondria would have to isolated and washes rapidly to avoid leakage of metabolites of interest. The applicability of this approach would likely be metabolite dependent, as incubation times of at least 10 minutes (most protocols without specialized cell lines take more than 1 hour) in wash buffer are currently unavoidable, which could lead to metabolite leakage or continued metabolism. The second approach would be to isolate functional mitochondria and study metabolism in isolation. These could include either simple substrate transport assays or studying the small molecule content of mitochondria after a given period of time of active metabolism in a complex milieu of metabolic substrates. Again, a rapid wash protocol would have to be established (the one developed for *L. Lactis* studies could be

considered). Regardless of the approach, isolated mitochondria experiments are probably more amenable to targeted or semi-targeted experiments, as replicates will generally limited and the sensitive, specialized process will likely lead to more replicate data variation than cell metabolite extract replicates.

The production of new knockout cell lines in a different parental cell line should be considered. As discussed above, the MDA-MB-468 cell line may be genetically unstable and may exhibit clonal metabolic variation, which can be a confounding factor in this knockout cell line-based approach and may make it a poor choice of experimental system. Selection of a more stable cell type could mitigate some of these concerns. The detected prevalence of unreported splice variants suggests that a parental cell line's mRNA should be sequenced before pursuing a knockout to check for cell line-specific splice variants that could confound knockout attempts. I would also recommend that clonal selection from the parental population before attempting knockout to establish the most homogenous and comparable control cell line possible. Other human cell line-based approaches could also be considered, such as the use of inducible hairpin knockdown cell lines or studying CRISPR knockout in a polyclonal population-based approach.

With regard to *Lactococcus Lactis* bacteria protein expression protocol, further work should focus on determining if there are conditions where intact bacteria could be used to generally characterize SLC25 family transporters. Numerous changes can be considered, such as attempting to increase protein expression (e.g. by increasing nisin concentration), varying pH, "preloading" bacteria with antiported substrate and de-energizing (with 2-deoxyglucose) and energizing bacteria (with glucose). Other carboxylate transport proteins could be tested for successful positive control transport,

such as the mitochondrial citrate transporter SLC25A1. Use of a transporter that has previously been shown to be functional in intact bacteria transport experiments, such as the human mitochondrial ADP/ATP carrier (SLC25A4), could be useful in establishing the validity of the protocol before troubleshooting proteins which have not previously been shown to be functional in whole *L. Lactis*.

Should intact bacteria experiments continue to fail to show overexpressed protein-mediated transport, the protocol should be adapted to reconstituted proteoliposome-based studies. Once again, these studies should first use transporters that have previously been shown to be active in identical experiments (albeit with radiolabeled substrates), such as human SLC25A10 or SLC25A11, as positive controls to establish the validity of the protocol.