

8-2015

Metabolomics of transformed MDA-MB-231 cell lines expressing different levels of human arylamine N-acetyltransferase 1 (NAT1).

Samantha Marie Carlisle
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>

Part of the [Pharmacy and Pharmaceutical Sciences Commons](#)

Recommended Citation

Carlisle, Samantha Marie, "Metabolomics of transformed MDA-MB-231 cell lines expressing different levels of human arylamine N-acetyltransferase 1 (NAT1)." (2015). *Electronic Theses and Dissertations*. Paper 2218.
<https://doi.org/10.18297/etd/2218>

This Master's Thesis is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

METABOLOMICS OF TRANSFORMED MDA-MB-231 CELL LINES EXPRESSING DIFFERENT
LEVELS OF HUMAN ARYLAMINE *N*-ACETYLTRANSFERASE 1 (NAT1)

By

Samantha Marie Carlisle
B.S., University of Louisville, 2012

A Thesis
Submitted to the Faculty of the
School of Medicine of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science

In Pharmacology and Toxicology

Department of Pharmacology and Toxicology
University of Louisville
Louisville, Kentucky

August 2015

Copyright 2015 by Samantha Marie Carlisle

All rights reserved

METABOLOMICS OF TRANSFORMED MDA-MB-231 CELL LINES EXPRESSING DIFFERENT
LEVELS OF HUMAN ARYLAMINE *N*-ACETYLTRANSFERASE 1 (NAT1)

By

Samantha Marie Carlisle
B.S., University of Louisville, 2012

A Thesis Approved on

July 9, 2015

by the following Thesis Committee:

David W. Hein, Ph.D.

J. Christopher States, Ph.D.

Xiang Zhang, Ph.D.

Carolyn M. Klinge, Ph.D.

Shesh N. Rai, Ph.D.

ACKNOWLEDGEMENTS

I would first and foremost like to thank my mentor, Dr. David Hein, for all of his advice and direction over the past three years. I would also like to thank my committee members, Dr. J. Christopher States, Dr. Xiang Zhang, Dr. Carolyn Klinge, and Dr. Shesh Rai, for all of their feedback and helpful suggestions throughout this process. Lastly, I would like to acknowledge and thank all of my friends and family who have provided unwavering encouragement and support in all of my endeavors. I am tremendously grateful for all of you.

ABSTRACT

METABOLOMICS OF TRANSFORMED MDA-MB-231 CELL LINES EXPRESSING DIFFERENT LEVELS OF HUMAN ARYLAMINE *N*-ACETYLTRANSFERASE 1 (NAT1)

Samantha Marie Carlisle

July 9, 2015

Human arylamine *N*-acetyltransferase 1 (NAT1) is a phase II xenobiotic metabolizing enzyme that is found in almost all tissues of the body. Expression of NAT1 is commonly elevated in several cancers including breast cancer. However, the exact mechanism by which NAT1 expression affects cancer risk and progression remains unclear. Three MDA-MB-231 breast adenocarcinoma cell lines that stably express wild-type, increased, and decreased levels of human NAT1 have been constructed and characterized for NAT1 mRNA, NAT1 acetylation activity, cell doubling rates, and endogenous acetyl-CoA levels. Differences in polar metabolic profile between these three cell lines were investigated using a comprehensive GC×GC-TOF MS system. Increased levels of human NAT1 in the transformed cell lines resulted in a statistically significant decreased abundance of the metabolite palmitoleic acid with one-way ANOVA $p = 0.0004$, when compared to normal and decreased levels of human NAT1. We hypothesize that increased NAT1 activity leads to increased hydrolysis of acetyl-CoA thus leaving less acetyl-CoA available to participate in other reactions. The fatty acid synthesis pathway utilizes acetyl-CoA in the first two reactions of the pathway and eventually leads to the synthesis of palmitoleic acid (16:1). These data suggest a link between increased levels of NAT1 activity and decreased flux of acetyl-CoA through this portion of the fatty acid synthesis pathway.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS.....	iii
ABSTRACT.....	iv
LIST OF FIGURES.....	vi
INTRODUCTION.....	1
MATERIALS AND METHODS.....	9
RESULTS.....	17
Characterization of Samples.....	17
Metabolomics.....	20
Multivariate/Multivariable Analyses.....	34
DISCUSSION.....	39
SUMMARY AND CONCLUSIONS.....	45
Strengths of this Work.....	45
Caveats and Weaknesses.....	45
Future Directions.....	47
REFERENCES.....	49
APPENDICES.....	52
CURRICULUM VITAE.....	53

LIST OF FIGURES

FIGURE	PAGE
1. Construction of MDA-MB-231 Breast Cancer Cells That Express Varying Levels of NAT1 Activity.....	11
2. Metabolomics Experimental Design Diagram.....	14
3. Metabolomics Data Analyses Workflow.....	15
4. NAT1 mRNA of the Three Transformed Cell Lines.....	18
5. NAT1 Activity of the Three Transformed Cell Lines.....	19
6. Doubling Rate of the Three Transformed Cell Lines.....	21
7. Endogenous Acetyl-CoA levels of the Three Transformed Cell Lines.....	22
8. Scrambled versus Down Volcano Plot.....	28
9. Scrambled versus Up Volcano Plot.....	29
10. Down versus Up Volcano Plot.....	30
11. Box Plots Comparing the Abundance Distributions of L-leucine, L-aspartic acid, and L-valine.....	31
12. Box Plot Showing the Abundance Distribution of Palmitoleic Acid.....	32
13. Box Plot of Amino Acids that are Broken Down to Form Acetyl-CoA.....	33
14. Principal Component Analysis Scores Plot.....	35
15. Partial Least Squares Discriminate Analysis Scores Plot	36
16. Orthogonal Partial Least Squares Discriminate Analysis Scores Plot	37
17. OPLS-DA Jack-Knife Loadings Plot.....	38
18. Acetyl-CoA is a Central Molecule Involved in a Plethora of Reactions.....	40
19. Fatty Acid Synthesis Pathway (abbreviated).....	43

INTRODUCTION

It is estimated that in 2015 breast cancer will account for 29% of all new cancer cases diagnosed in women and 15% of all cancer related deaths in women [1]. From 2002-2003 there was a sharp decline in breast cancer incidence in non-Hispanic Caucasian women correlating with the decreased use of hormone replacement therapy, but that trend did not continue into the subsequent years [2]. From 2004 to 2010, breast cancer incidence rates have not significantly changed, thus emphasizing the necessity of further investigation into what factors are affecting breast cancer risk and progression. As more is known about breast cancer, preventative and treatment options can be improved.

Current breast cancer treatment options include surgery, sentinel lymph node biopsy followed by surgery, radiation therapy, hormone therapy, targeted therapy, and chemotherapy [3]. Also, for people with a history of familial breast cancer, genetic screenings for mutations in the tumor suppressor genes BRCA1 and BRCA2 serve as a way to identify breast cancer risk. Potential onset of the disease can be prevented if mastectomies are performed on those who test positive for mutations in those genes. Although there are many treatment options for breast cancer, all come with complicated physical and emotional side effects that decrease the quality of life of patients. Unfortunately, these treatment and preventative options are not always 100% effective at eliminating all the cancerous cells, leaving the risk of relapse in addition to the harmful side effects. Also some cancers, during the course of treatment, become resistant to radiation therapy, hormone therapy, targeted therapy, and chemotherapy meaning additional measures must be taken. Due to the harmful side effects, restricted effectiveness, and risk for relapse associated with current therapies it is important that new treatment options and targets are explored and elucidated.

Early detection of breast cancer is extremely important. When breast cancer is detected in the localized stage, the five-year relative survival rate is an incredible 99%, but when detected in the distant stage, the five-year relative survival rate is a disappointing 24% [1]. In the earliest stages of breast cancer, the cancerous cells are localized at the point of origin and can be removed by surgery with no further intervention required. As breast cancer progresses, more mutations are incurred and this frequently leads to the breast cancer becoming more aggressive, invasive, and harder to treat. Once breast cancer has undergone metastasis, harsher and more aggressive treatments are required. A better understanding of what influences disease risk and drives disease progression is needed to increase detection of breast cancer in the earliest stage when it is the most treatable.

Breast cancer is a heterogeneous disease that can be subclassified according to multiple classifiers including pathology, tumor grade and stage, gene and protein expression, and immunohistochemical characteristics. One of the most commonly used methods to classify breast cancer is immunohistochemically according to estrogen receptor (ER), progesterone receptor (PR), and HER2/neu status. Using this classification system, a triple negative breast cancer would not express any of these receptors while a triple positive breast cancer would express all three receptors. Recently, classification based on gene and protein expression has become of more interest because it may provide clues about what treatment options will be the most effective for each specific cancer case. Sorlie et al. conducted a study in which gene expression data of 85 cDNA microarray experiments representing 78 cancers, three fibroadenomas, and four normal breast tissues were evaluated [4]. Based on analysis by hierarchical clustering, six subtypes were defined. These breast cancer subtypes are 1. Luminal Subtype A, 2. Luminal Subtype B, 3. Luminal Subtype C, 4. Basal-like, 5. ERBB2+, and 6. Normal Breast-like; these subtypes are associated with different clinical outcomes in addition to the differences in gene and protein expression. Other studies have since been conducted looking at the prevalence of each of these subtypes and have investigated further differences between them [5].

Human arylamine *N*-acetyltransferase 1 and 2 (NAT1 and NAT2) are cytosolic phase II xenobiotic metabolizing isozymes that acetylate a wide range of aromatic and heterocyclic

amines via a ping-pong bi-bi reaction mechanism [6],[7]. NAT1 and NAT2 genes are both located on chromosome 8p22 [8]. Although NAT1 and NAT2 share approximately 87% coding sequence homology [9] functional studies have demonstrated significant differences between the two isozymes in terms of substrate specificity and tissue localization. NAT1 shows substrate specificity for p-aminosalicylate and p-aminobenzoic acid (PABA) while NAT2 shows substrate specificity for isoniazid, sulfamethazine, and procainamide [10]. These distinct but overlapping substrate specificities have been shown to be due to differences in the substrate binding site of the enzymes [11]. NAT1 and NAT2 also vary in terms of tissue expression with NAT1 being expressed in most body tissues but NAT2 expression is limited primarily to the liver and GI tract tissues. Although NAT1 has been shown to be expressed in most body tissues, the only known endogenous substrate of NAT1, to date, is p-aminobenzoyl-glutamate (PABA-glu), a folate catabolite [12]. We suspect NAT1 has other endogenous substrates that have yet to be elucidated, given the extensive tissue distribution and presence in almost all species.

Humans are frequently exposed to NAT substrates. Many therapeutic drugs, including arylamine drugs such as procainamide, dapson, and aminoglutethimide, hydrazine drugs such as isoniazid, hydralazine and phenelzine, and secondary arylamine and hydrazine metabolites such as sulfasalazine, clonazepam, acebutolol, and caffeine are metabolized by NAT [13]. Aromatic amines, another class of NAT substrates, are found as the byproducts of the manufacturing of compounds such as polyurethane foams, dyes, pesticides, pharmaceuticals, and semi-conductors [14]. Aromatic amines are also found in environmental pollution such as diesel exhaust, combustion of wood chips and rubber, tobacco smoke and well-done grilled meats [15]. Several of these NAT substrates are environmental carcinogens and their acetylation leads to either deactivation or bioactivation. One of the most well known and studied NAT substrates is 4-aminobiphenyl (ABP), a carcinogenic aromatic amine that is a major constituent of cigarette smoke [16]. The NAT's play a major role in the deactivation and bioactivation of carcinogenic aromatic amines and arylamines. Carcinogenic NAT substrates, such as ABP, can either be *N*-acetylated (usually deactivation) by NAT, or if first acted upon by a cytochrome P450 enzyme, they can be *O*-acetylated (usually bioactivation) by NAT. Once bioactivated, these

compounds can spontaneously form arylnitrenium ions leading to DNA adducts. If not repaired by nucleotide excision repair (NER), these DNA adducts can lead to mutagenesis and possible cancer initiation.

Recently it has been shown that human NAT1 can also catalyze the direct hydrolysis of acetyl-CoA in the absence of a NAT1 arylamine substrate using folate as a cofactor [17]. In this reaction, the folate is not acetylated, instead the acetate released from acetyl-CoA is excreted into the surrounding bulk solvent. Interestingly, the same reaction was not detected in the presence of human NAT2 instead of human NAT1. The discovery of this additional function of NAT1 provides support for the hypothesis that NAT1 has an undetermined endogenous role in addition to the role NAT1 serves as a phase II xenobiotic metabolizing enzyme.

Historically, NAT2 has been of more interest because of the role single nucleotide polymorphisms (SNPs) in NAT2 play in the metabolism of isoniazid, an anti-tuberculosis drug. Decreased NAT2 activity due to the presence of SNP's in "slow" acetylators increases the risk of developing hepatotoxicity during treatment because isoniazid is eliminated slower when compared to "rapid" acetylators and accumulates in the liver [18-20]. NAT1 has recently become of more interest because of the possible link between NAT1 expression and cancer risk and progression. Examination of publicly available microarray data by others, has shown that changes in the levels of NAT1 mRNA are associated with specific cancers and cancer subtypes [21]. Of all cancers, the association between NAT1 and breast cancer has been the most extensively investigated. A microarray study of 66 breast cancer samples in 2009 by Casey et al., revealed NAT1 expression was higher in invasive carcinoma compared with normal tissue [21, 22]. Also, Yuan et al. showed that NAT1 is commonly amplified in breast carcinomas using computational approaches and a previously published microarray dataset containing 89 breast cancer samples [23, 24]. Adam et al. have shown by comparative proteomic analysis of 8 normal breast samples and 25 breast cancer derived tissues that NAT1 is elevated in invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC) when compared to the normal tissue [25]. In the same study it was also shown that overexpression of NAT1 in HB4a, conditionally immortalized human mammary luminal epithelial derived cells, enhanced cell growth and resistance to etoposide, a

cytotoxic anticancer drug, relative to control cells [25]. Smid et al showed in a study of 107 breast cancer samples that high NAT1 levels in primary tumors were significantly associated with increased metastasis to the bone [26].

Initially, NAT1 was thought to be monomorphic in terms of acetylation activity but additional studies have since revealed NAT1 is subject to separate polymorphisms from NAT2 [27-29]. NAT1 contains SNP's located in the protein region and in the 3'-untranslated regions of the gene that affect the level of enzymatic activity. Although SNP's in NAT1 are less common than those found in NAT2 they still confer distinct differences in NAT1 activity in the human population. To date, there have been 28 NAT1 alleles identified. Due to the infrequency at which they occur, roughly half of the NAT1 alleles have yet to be phenotypically defined as conferring normal, elevated, or decreased NAT1 activity. Data regarding human NAT1 alleles have been collected and presented at http://nat.mbg.duth.gr/Human%20NAT1%20alleles_2013.htm [30]. The *NAT1*4* allele is defined as the reference allele. *NAT1*14A*, *NAT1*14B*, *NAT1*17*, and *NAT1*22* have all been defined as having activity lower than *NAT1*4*. *NAT1*15*, *NAT1*19A*, and *NAT1*19B* have been defined as coding for a truncated protein that has no activity. *NAT1*20*, *NAT1*21*, *NAT1*23*, *NAT1*24*, *NAT1*25*, and *NAT1*27* have all been defined as having activity equivalent to the reference allele, *NAT1*4*. There have been conflicting reports about the activity of the *NAT1*10* allele, some report increased NAT1 activity when compared to the reference allele but others have seen no difference [31-33]. Interestingly, in a study by Wakefield et al. that reported the NAT1 activity and genotypes of multiple breast cancer cell lines, the cell lines ZR-75-1 and MDA-MB-231 had the same NAT1 genotype (*NAT1*4/NAT1*4*) but varied in activity by greater than 1,000 fold [34]. In the ZR-75-1 cell line, elevated levels of NAT1 were not due to alterations in gene copy number. Further analysis has indicated that the NAT1 transcripts in the ZR-75-1 cell line are produced from the distal promoter NATa while the NAT1 transcripts produced in all the other cell lines tested were derived from the NATb promoter [35]. There have also been conflicting data regarding the extent to which these polymorphisms in NAT1 modify risk of developing breast, colorectal, urinary bladder, head and neck, lung, and prostate cancers. Adding to the confusion, NAT1 activity is also modulated by a plethora of post-translational and

environmental factors, including direct chemical modification of the active-site cysteine by reactive chemical species [36, 37], factors that induce oxidative stress and alter GSH levels [21], cytokines [38], and through inhibition by small molecule inhibitors [39-41], heavy metals [42], plant extracts [37], nanoparticles [43], and therapeutic agents [44]. The NAT1 gene copy number can also be increased or decreased when compared to the standard of two, thus causing differences in NAT1 activity.

Several studies have demonstrated that expression of NAT1 is significantly linked to estrogen receptor status in breast cancer samples [45-47]. An evaluation of previously published microarray data that compared gene expression between ER-positive and ER-negative tissues by Wakefield et al. showed that NAT1 is in the top three most overexpressed genes in ER-positive tissue in over 50% of the studies [34]. ER-positive breast cancer can further be subclassified as luminal A (classically defined as ER^{high} Ki67^{low}) and luminal B (classically defined as ER^{low} Ki67^{high}). In a recent study investigating differences in the luminal A and luminal B subtypes of ER-positive breast cancer published by Endo et al., it was shown that the microRNA (miRNA), miR-1290 is downregulated in luminal A tumors when compared to luminal B tumors [48]. The same group has shown in a later publication that miR-1290 binds to a site in the *NAT1* 3'-UTR and NAT1 is directly downregulated in a dose-dependent manner by miR-1290 [49]. These data suggests that downregulation of miR-1290 in luminal A breast cancer leads to overexpression of NAT1. Interestingly, overexpression of NAT1 in ER-positive breast cancer has been shown to be correlated with a decreased 5-year relapse rate [50] and NAT1 expression has been shown to be negatively correlated with tumor size [48].

There have been multiple studies investigating the effects the inhibition of NAT1 has on immortalized cancer cells. In one study by Tiang et al., in which NAT1 activity in HT-29 cells, immortalized human colorectal adenocarcinoma cells, was decreased through the use of shRNA, significant changes in cell growth and morphology were observed when cells were at high density as opposed to sub-confluent density [51]. In the same study it was also shown that NAT1 knockdown led to increased E-cadherin protein expression and impeded anchorage-independent growth [51]. Studies have also been conducted in which NAT1 was inhibited by small molecules.

In one such study Rhod-o-hp, a rhodanine derivative, was used to decrease NAT1 activity in MDA-MB-231 cells, breast adenocarcinoma cells. This study revealed that Rhod-o-hp inhibits anchorage independent growth and invasiveness of MDA-MD-231 cells, which was comparable to the effects seen with knockdown of NAT1 via shRNA, suggesting Rhod-o-hp's effect on MDA-MB-231 cells are caused by inhibition of NAT1 [40]. We have seen similar results in our lab with the NAT1 specific small molecule inhibitor 9,10-dihydro-9,10-dioxo-1,2-anthracenediyl diethyl ester (DDADE) [41]. DDADE was identified through an *in silico* screening of over 700,000 small molecules in the ZINC (a free database of commercially-available compounds for virtual screening) library using the crystal structure of human NAT1 to predict binding affinity. Studies in our lab have shown that DDADE decreases NAT1 catalyzed *N*-acetylation and *O*-acetylation of arylamine substrates in human breast adenocarcinoma cells. DDADE has also been shown to decrease cell adhesion and cell invasion in human breast cancer cells.

Cancer cells characteristically exhibit rapid, uncontrolled growth therefore requiring increased levels of energy and cellular components when compared to normal cells. Reprogramming of cellular metabolism to meet these needs is one of the next generation hallmarks of cancer proposed by Hanahan and Weinberg [52]. Although, the exact mechanism by which cellular metabolism is reprogrammed remains unclear, it is likely specific to each type of cancer and is also influenced by a combination of factors. Metabolomics studies are able to generate insights into which pathways have been modified, by measuring the abundance of metabolites and then comparing the abundances between groups with defined differences.

Historically, studies investigating the role NAT1 plays in cancer risk and progression have focused on NAT1's role as a xenobiotic metabolizing enzyme that can transform carcinogens, but we propose NAT1 affects cancer risk and progression through regulation of acetyl-CoA levels in the human body. It is our hypothesis that changing the level of NAT1 activity will lead to alterations in the levels of acetyl-CoA. Increased NAT1 activity would lead to increased hydrolysis of acetyl-CoA therefore less acetyl-CoA would be available. Conversely, decreased NAT1 activity would lead to decreased hydrolysis of acetyl-CoA therefore more free acetyl-CoA would be available. Preceding this study, metabolomics had not been utilized to investigate how differences

in NAT1 activity alter the abundance distribution of metabolites in breast cancer cells. It is expected that if NAT1 activity is involved in cancer risk or progression a change in metabolite abundance distribution could be observed between breast cancer cell lines representing varying levels of NAT1 activity. These potential changes are important to elucidate because understanding how varying levels of NAT1 affect metabolite abundances can give us a better understanding of which cellular pathways are altered as a result of varying levels of NAT1 activity. An improved understanding of cellular reprogramming as a result of NAT1 activity could lead to earlier cancer detection and more effective treatment options.

MATERIALS AND METHODS

CONSTRUCTION OF THE THREE STABLY TRANSFECTED MDA-MB-231 CELL LINES

The breast adenocarcinoma cell line, MDA-MB-231, was purchased from ATCC (Manassas, VA) and cultured in high-glucose Dulbecco's Modified Eagle Medium (DMEM), with 10% fetal bovine serum, 5% glutamine, and 5% penicillin/streptomycin added. The Flp-In™ System (Life Technologies, Grand Island, NY) was used previously in our lab (by Mark Doll) to create a single Flp Recombination Target (FRT) site in the parent MDA-MB-231 cell line according to manufacturer's instructions. SureSilencing™ NAT1 specific shRNA plasmids and a corresponding scrambled shRNA plasmid were purchased from SA Biosciences (Valencia, CA). It has previously been shown that the SureSilencing™ NAT1 shRNA Clone 3 is the most effective of the four provided by SA Biosciences at silencing NAT1 activity in the MDA-MB-231 cell line, therefore Clone 3 was chosen as the shRNA plasmid that would be used when generating a stable NAT1 knockdown MDA-MB-231 cell line. Next, the shRNA plasmids containing the scrambled and NAT1 clone 3 shRNA sequences were integrated into the pcDNA5/FRT vector from the pGeneclip™ vector, independently. For the construction of the MDA-MB-231 cell line overexpressing NAT1, the NATb/NAT1*4 plasmid in the pcDNA5/FRT vector previously constructed in our lab [53] was chosen. The construction of a MDA-MB-231 cell line containing a single FRT site in the transcriptionally active region allowed us to transfect each of the plasmids into the same genetic location in different cells, thus creating three new MDA-MB-231 cell lines whose only difference is the level of NAT1 activity. The pcDNA5/FRT plasmids containing the scrambled shRNA, NAT1 clone 3 shRNA plasmid or NATb/NAT1*4 overexpression plasmid, were co-transfected into the MDA-MB-231 cell line containing the single FRT site with pOG44 (Life Technologies), a Flp recombinase expression plasmid, using Amaxa Cell Line Nucleofector Kit V (Lonza, Basel, Switzerland) following manufacturer's instructions. Resulting cells were then

cultured in complete DMEM media containing 500 µg/mL hygromycin to select for cells that had successfully undergone transfection and stably integrated the pcDNA5/FRT plasmid into their genomic DNA since the pcDNA5/FRT vector contains a hygromycin resistance cassette. Hygromycin resistant colonies were selected approximately two weeks after transfection and isolated using cloning cylinders. The resulting transformed cell lines, “Scrambled”, “Down”, and “Up”, express parental, decreased, and increased NAT1 activity, respectively (Figure 1).

CHARACTERIZATION OF SAMPLES

Total RNA was isolated from each transformed cell line using the RNeasy Mini Kit (Qiagen, Germantown, MD) per manufacturer’s instructions. The High Capacity Reverse Transcriptase kit (Life Technologies, Grand Island, NY) was then used to synthesize cDNA according to manufacturer’s instructions. NAT1 mRNA was quantitated for each transformed cell line in triplicate using quantitative real-time polymerase chain reaction (qRT-PCR) and NAT1 specific primers and probes (Life Technologies, Grand Island, NY). The PCR conditions suggested by the manufacturer were used.

NAT1 *N*-acetylation activity was measured for each transformed cell line in triplicate via high pressure liquid chromatography (HPLC) in the presence of the NAT1 specific substrate *p*-aminobenzoic acid (PABA) by measuring the appearance of acetylated product. Cells from each transformed cell line were collected using a cell-scraper to detach the cells from the plate and then the cells were rinsed with 1 X PBS. Cells were then lysed in 20 mM sodium phosphate, 1 mM EDTA, 0.2% Triton X-100 (pH 7.4) with the addition of 1 mM DTT, 100 µM PMSF, 1 µg/ml Aprotinin, and 2 µM pepstatin A. Cell lysate was centrifuged at 15,000 x g for 10 min and the supernatant saved for subsequent NAT1 activity assays. The resulting cell lysates were incubated with 10 µM *para*-aminobenzoic acid (PABA) and 100 µM acetyl-CoA at 37°C for 10 minutes. The reaction was stopped by the addition of 1 M acetic acid at a volume equal to 1/10 of the total reaction volume. Samples were then centrifuged at 10,000 x g for 10 minutes and the resulting supernatant was collected for analysis of *N*-acetyl-PABA levels via HPLC.

To determine a cell doubling rate each transformed cell line was plated at an initial density of 25,000 cells per well in 12 well plates and then counted on a cell counter daily for 5

Figure 1

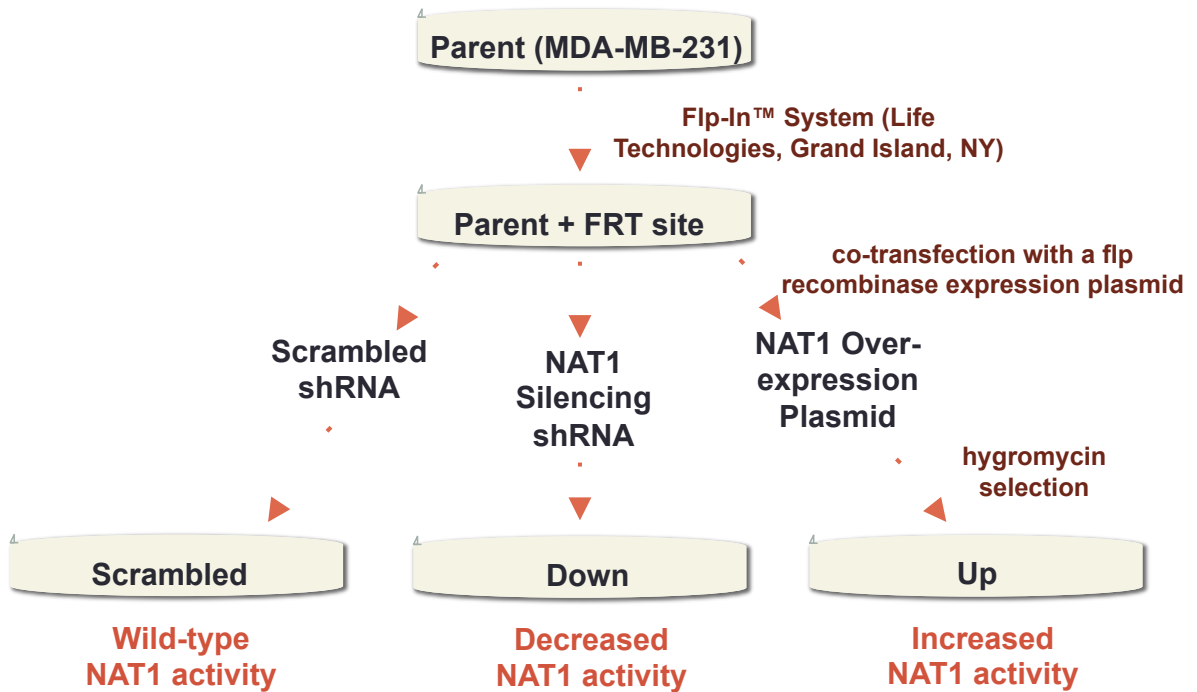


Figure 1: Construction of MDA-MB-231 Breast Cancer Cells That Express Varying Levels of NAT1 Activity

The Flp-In™ System (Life Technologies, Grand Island, NY) was used to insert a single FRT site into the parent MDA-MB-231 cell line. The MDA-MB-231 cell line containing the FRT site was then stably transfected with either with a plasmid containing a nonspecific shRNA plasmid into the FRT site (Scrambled), a plasmid containing NAT1 specific shRNA into the FRT site (Down), or a plasmid containing the NATb/NAT1*4 vector into the FRT site (Up). Each plasmid contained a hygromycin resistant cassette therefore cells that had successfully been stably transfected were then selected for using hygromycin

days. Cells were plated on day 1 and allowed to adjust for 24 hours before counting on day 2. Data were plotted to ensure cells were in the logarithmic growth phase. A cell doubling rate was calculated for each day using Equation 1 (shown below) and then rates were averaged over the 5 day time period to get a final cell doubling rate.

$$\text{Equation 1: Cell Doubling Rate} = 1/(((\log(b) - \log(a)) \times 3.32)/(t_b - t_a))$$

In the equation, b = cell number on day n , a = cell number on day $n-1$, t_b = hours passed since plating when cells were counted on day n , and t_a = hours passed since plating when cells were counted on day $n-1$.

Endogenous acetyl-CoA levels of the three transformed MDA-MB-231 cell lines were measured in triplicate via HPLC-UV. Cells were cultured in high-glucose Dulbecco's Modified Eagle Medium (DMEM), with 10% fetal bovine serum, 5% glutamine, and 5% penicillin/streptomycin added in 35 mm plates to approximately 80% confluence. Medium was then aspirated from each plate followed by a rinse with 10 mL of 1 X PBS. After removal of the PBS, cells were released from the plate with the addition of 0.5 mL trypsin per plate and then resuspended in 5 mL of 20% medium and 80% PBS. Cells were then counted via a Coulter automated cell counter. Cell suspensions were then centrifuged in 15 mL tubes for 2.5 minutes at 400 X g at 4°C. Supernatant was removed and cell pellets were resuspended in 1 mL of cold PBS. Cell suspensions were centrifuged again in 15 mL tubes for 2.5 minutes at 400 X g at 4°C. Cells were transferred to cooled 1.5 mL eppendorf tubes by resuspending them in 1 mL of cold PBS. Resulting cell suspensions were then centrifuged in the cold room at 300 X g for approximately 5 minutes. Supernatant was removed and cells were lysed in 100 μ L of 5% 5-sulfosalicylic acid by pipetting up and down. The resulting suspension was centrifuged in the cold room at 15,000 X g for approximately 10 minutes. Supernatant (100 μ L) was collected and injected onto a C18 reverse-phase HPLC column (250 mm \times 4 mm; 5 μ M pore size). Reactants and products were separated and quantitated using a Beckman System Gold high performance liquid chromatography (HPLC) system. HPLC separation of CoA and acetyl-CoA was achieved using a linear gradient of 100:0 sodium phosphate (NaH_2PO_4) pH 4: methanol to 0:100 sodium phosphate pH 4: methanol over 20 minutes and was quantitated by absorbance at 260 nm.

Since the cells were lysed in 5% 5-sulfosalicylic acid, this buffer solution was used as a reference to subtract any background acetyl-CoA coming from the lysis buffer.

METABOLOMICS

Each sample of cells for metabolomic analyses contained approximately five million cells with a wet weight of 20 mg. The cells were lysed by homogenization in 200 μ L of deionized water. Next, samples were mixed using a vortex mixer for two minutes after the addition of 800 μ L of ice cold methanol. Samples were then stored on ice for ten minutes followed by three more minutes of mixing on the vortex mixer. Next, samples were centrifuged at 12,000 X g for ten minutes. Supernatant (800 μ L) was then transferred to a 1.5 mL eppendorf tube and dried via speed vacuum until dry (roughly three hours). Cell pellets used for metabolomics samples were frozen at -80°C until day of lysis and extraction.

Polar metabolites were extracted from cell pellet samples of each transformed cell line in octuplicate using a mixture of 80% methanol and 20% water. Samples were then derivitized with N-tert-butyltrimethylsilyl-N-methyltrifluoroacetamide (MTBSTFA) followed by GC \times GC-TOF MS to identify metabolites. The GC \times GC-TOF MS data were then processed using LECO's instrument control software ChromaTOF for peak picking and tentative metabolite identification, followed by retention index matching, peak merging and peak list alignment (Figure 2). Based on the results of our analyses (described in Statistical Analysis), a list of metabolites were subject to technical validation i.e., analysis of standards for the given metabolites on GC \times GC-TOF MS and comparison of those results to the experimental data to provide further evidence of correct metabolite identification.

STATISTICAL ANALYSIS

One-way ANOVA was conducted on the data collected during characterization of the transformed cell lines; Post-hoc *t*-tests were conducted when the one-way ANOVA statistic was significant. The analyses were performed using Prism® software (GraphPad software, La Jolla, CA). Two approaches were used to analyze the metabolomics data obtained, 1. univariate and 2. multivariate/multivariable analyses (Figure 3). Imputation, normalization, and statistical significance tests, including student's *t*-tests and one-way ANOVA were performed using R

Figure 2

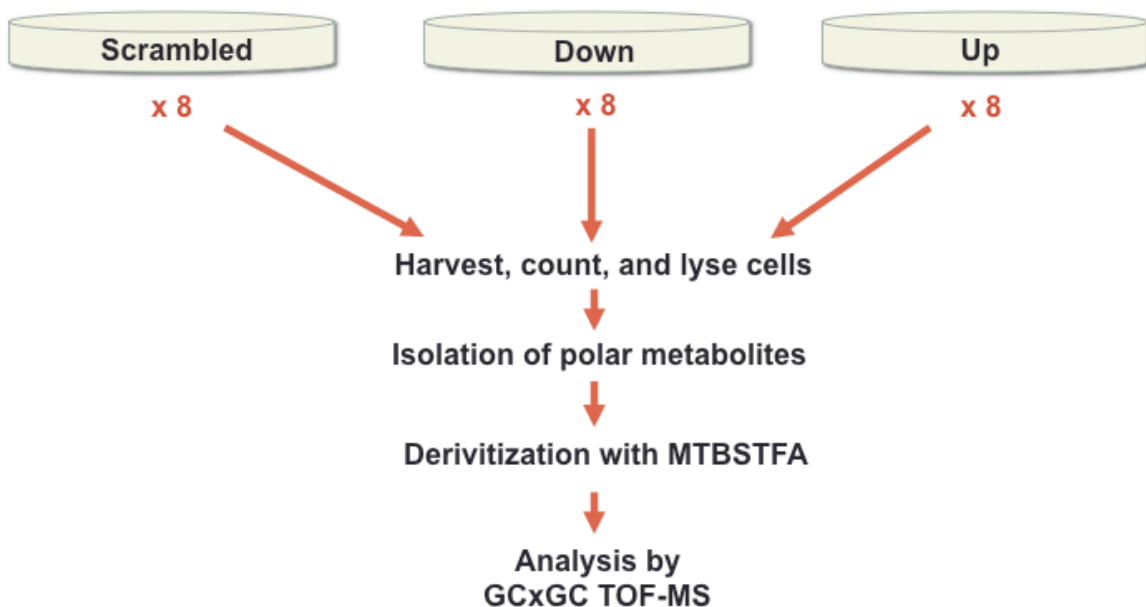


Figure 2: Metabolomics Experimental Design

In this figure, Scrambled refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing a nonspecific shRNA plasmid into the FRT site, Down refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing NAT1 specific shRNA into the FRT site, and Up refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing the NATb/NAT1*4 vector into the FRT site. Cells from eight biological replicates of each cell line were first harvest, counted, and lysed. Next polar metabolites were isolated, followed by derivitization with MTBSTFA. The resulting samples were then analyzed by two-dimensional gas chromatography time-of-flight mass spectrometry.

Figure 3

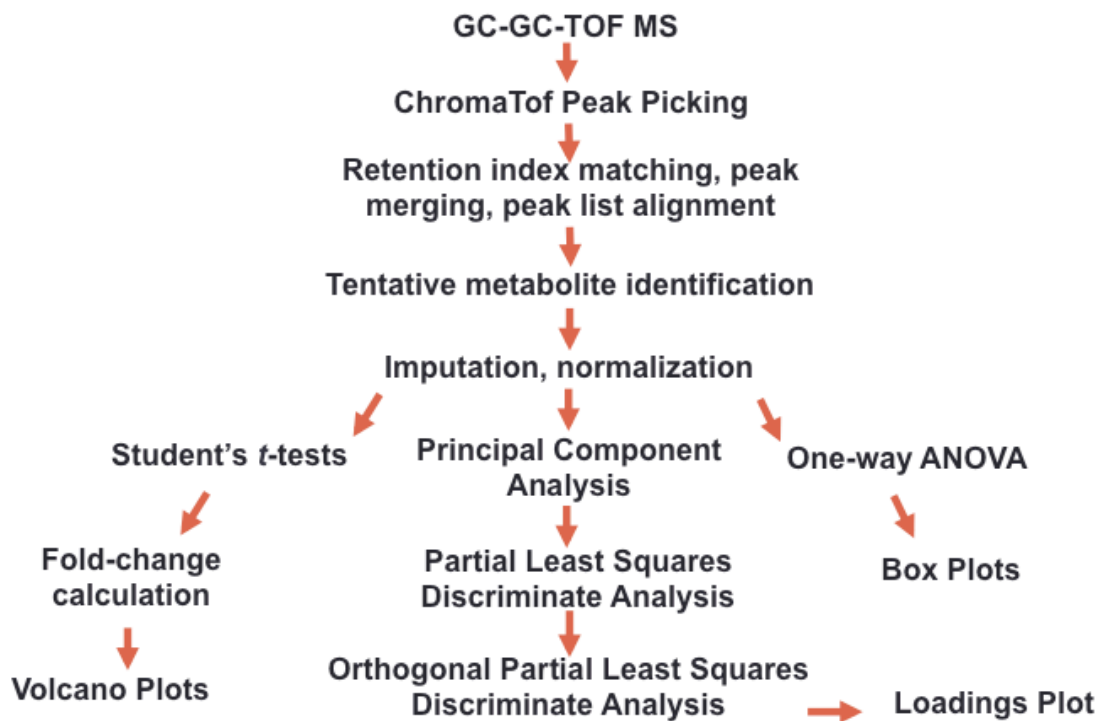


Figure 3: Metabolomics Data Analyses Workflow

In this figure, the workflow for the analyses of the data collected from the two-dimensional gas chromatography time-of-flight mass spectrometer is shown.

software (R Core Team (2015), Vienna, Austria). Q-values for the one-way ANOVA and *t*-test comparisons were computed from the raw p-values, respectively, using the R package qvalue (John Storey, R package version 2.0.0.). To better visualize the data, volcano plots and box plots were created using R software. Metabolite abundance fold-change between the Scrambled & Up, Scrambled & Down, and Down & Up groups were calculated using Equation 2 (shown below).

$$\text{Equation 2: } FC = 2^{\left| \log_2 \frac{\bar{B}}{\bar{A}} \right|}$$

In the equation, *FC* = fold-change, \bar{A} = average metabolite abundance in the reference group, \bar{B} = average metabolite abundance in the comparison group. A negative sign was added to the fold-change when metabolite abundance was lower in the comparison group compared to the reference group. The volcano plots plot the log base 2 of the fold change between the two specific groups versus the negative log base 10 of the unadjusted student's *t*-test p-value between the two specific groups. The box plots plot the median normalized abundance levels for each metabolite. Then, Principal Component Analysis (PCA), Partial Least Squares Discriminate Analysis (PLS-DA), and Orthogonal Partial Least Squares Discriminate Analysis (OPLS-DA) were conducted to give a more complete picture of metabolite interactions. The loadings of Component 1 from the OPLS-DA results were plotted to visualize which metabolites were contributing the most to the separation of the groups.

PATHWAY ANALYSIS

Pathway analysis was conducted on only those metabolites that were verified and found to be significant following student's *t*-tests and/or one-way ANOVA. We began the pathway analysis portion of this study by first locating acetyl-CoA on the Roche Biochemical Pathways Chart (F. Hoffmann-La Roche AG, Basel, Switzerland) and then following all possible pathways to and from acetyl-CoA to locate the metabolites that were verified and significant. Some metabolites were not located in a pathway involving acetyl-CoA therefore they have not been considered further at this time. For those that were located in a pathway involving acetyl-CoA, we then developed hypotheses that could explain how each metabolite's abundance distribution across the three groups related to the amount of free acetyl-CoA available.

RESULTS

CHARACTERIZATION OF SAMPLES

Three stably transfected MDA-MB-231 cell lines, Scrambled, Up, and Down, expressing parental, increased, and decreased levels of NAT1 activity, respectively, have been constructed. The three cell lines were characterized to confirm we had successfully constructed stable cell lines that express parental, increased, and decreased NAT1 activity relative to the parent MDA-MB-231 cell line. First, the levels of NAT1 mRNA were measured via quantitative real-time PCR with NAT1 specific primers and probes. Differences in NAT1 mRNA expression of the Parent, Scrambled, Down, and Up cell lines were tested by one-way ANOVA ($p < 0.0001$) and found to be statistically significant (Figure 4). The NAT1 mRNA of the parent MDA-MB-231 cell line that contains the FRT site but was not transfected with shRNA plasmids was measured and found to express 90% of the NAT1 mRNA that the Scrambled cell line does. This difference in NAT1 mRNA between the Parent and Scrambled cell lines was not statistically significant ($p > 0.05$) when tested further by a Tukey's post-hoc test. Relative to the Scrambled cell line, the Down cell line expressed 50% of the NAT1 mRNA and the Up cell line expressed 3100% of the NAT1 mRNA. Following a Tukey's post-hoc test NAT1 mRNA expression was significantly different between the Scrambled & Up ($p \leq 0.001$) and Down & Up ($p \leq 0.001$) cell lines but not between the Scrambled & Down cell lines ($p > 0.05$).

Next, the levels of NAT1 activity were measured via detection of acetylated product by HPLC using the NAT1 specific substrate PABA. Differences in NAT1 PABA acetylation activity of the Parent, Scrambled, Down, and Up cell lines were tested by one-way ANOVA ($p < 0.0001$) and found to be statistically significant (Figure 5). The parent MDA-MB-231 cell line, that contains only the FRT site, was found to have a NAT1 activity of 19.2 ± 1.2 nmol PABA acetylated/min/mg protein. The Scrambled cell line, that contains the FRT site and was transfected with the NAT1

Figure 4

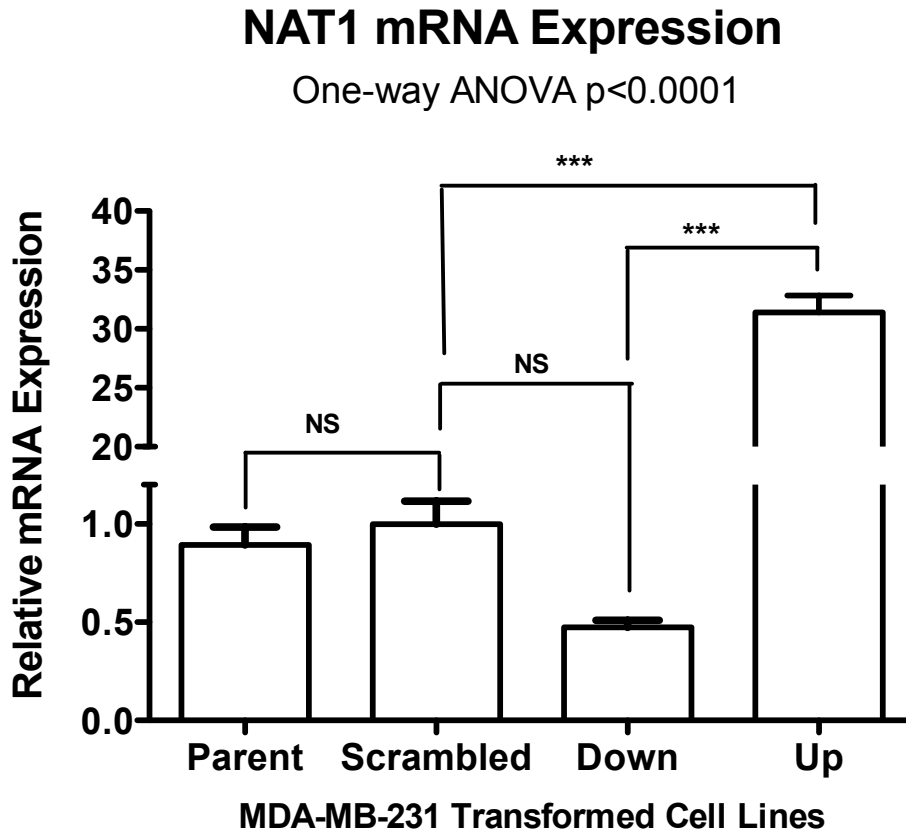


Figure 4: NAT1 mRNA expression of MDA-MB-231 transformed cell lines.

In this figure, Scrambled refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing a nonspecific shRNA plasmid into the FRT site, Down refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing NAT1 specific shRNA into the FRT site, and Up refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing the NATb/NAT1*4 vector into the FRT site. Relative NAT1 mRNA levels in cell lines normalized to scrambled shRNA cell line. Bars represent results from three separate determinations with mean \pm SD. Statistical significance was determined using one-way ANOVA with $p < 0.0001$ followed by a Tukey's post-hoc test. NS = not significant, *** = $p \leq 0.001$.

Figure 5

NAT1 Acetylation Activity

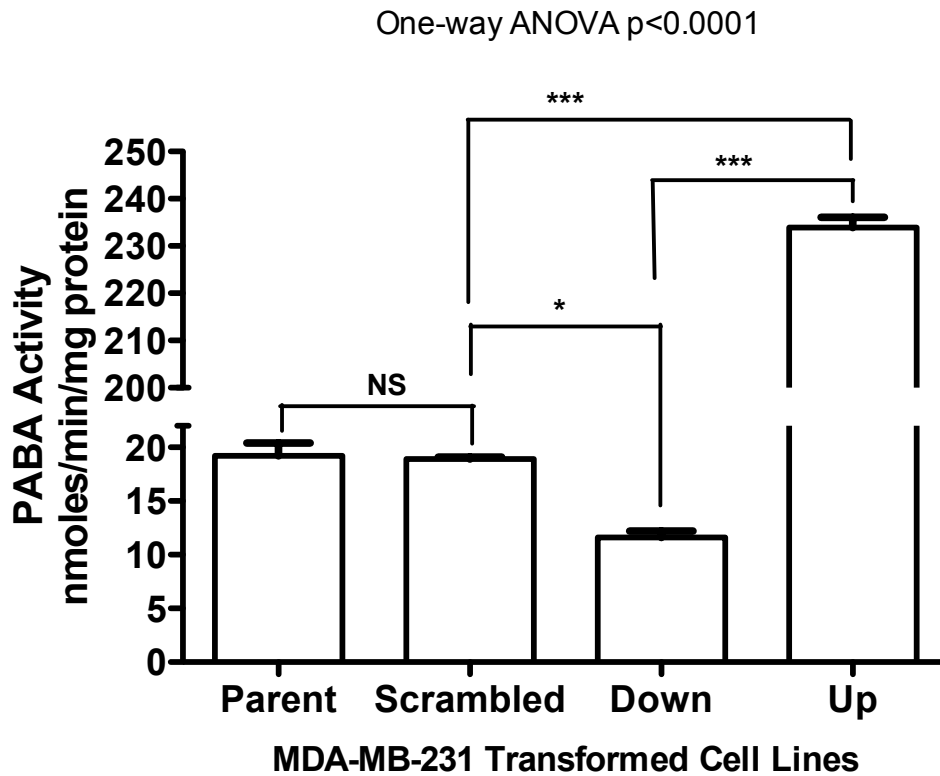


Figure 5: NAT1 PABA acetylation activity of MDA-MB-231 transformed cell lines.

PABA N-acetyltransferase activity of transformed cell line lysates were measured via HPLC. In this figure, Scrambled refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing a nonspecific shRNA plasmid into the FRT site, Down refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing NAT1 specific shRNA into the FRT site, and Up refers to the MDA-MB-231 cell line that was stably transfected with a plasmid containing the NATb/NAT1*4 vector into the FRT site. Bars represent results from three separate determinations with mean \pm SD. Statistical significance was determined using one-way ANOVA ($p < 0.0001$) followed by a Tukey's post-hoc test. NS = not significant, * = ≤ 0.05 , *** = $p \leq 0.001$.

scrambled shRNA, was found to have a NAT1 activity of 18.9 ± 0.2 nmol PABA acetylated/min/mg protein. The Up cell line, that contains the FRT site and was transfected with the NATb/NAT1*4 overexpression plasmid, was found to have a NAT1 activity of 233.9 ± 2.2 nmol PABA acetylated/min/mg protein. Finally, the Down cell line, that does contain the FRT site and was transfected with the NAT1 clone 3 shRNA, was found to have a NAT1 activity of 11.6 ± 0.6 nmol PABA acetylated/min/mg protein. When compared to the Scrambled cell line, the Down cell line had 61% of the NAT1 activity and the Up cell line had 1240% of the NAT1 activity. Following a Tukey's post-hoc test the NAT1 activity of the parent MDA-MB-231 cell line that contains the FRT site but was not transfected with shRNA plasmids did not significantly vary ($p > 0.05$) in activity from the NAT1 activity of the Scrambled cell line. Further analysis does not include the parent MDA-MB-231 cell line since there was no statistically significant difference between its levels of NAT1 mRNA & NAT1 activity and that of the Scrambled cell line.

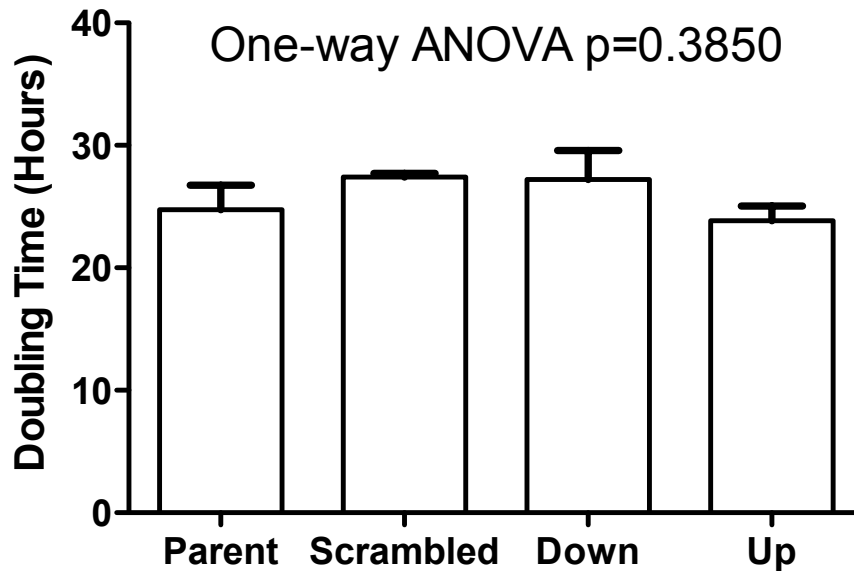
Next, the doubling times of the three cell lines were measured by counting the cells over a period of 6 days. The doubling rate of the Scrambled, Up, and Down cell lines were calculated to be 27.4 ± 0.3 hours, 23.9 ± 1.2 hours, and 27.2 ± 2.4 hours, respectively. Following a one-way ANOVA ($p = 0.3850$) the doubling rates of the three transformed cell lines were not found to be significantly different (Figure 6). Finally, endogenous acetyl-CoA levels were measured via HPLC-UV. Differences in endogenous acetyl-CoA levels were tested by one-way ANOVA ($p = 0.0218$) and found to be statistically significant (Figure 7). The acetyl-CoA levels of the Scrambled, Up, and Down transformed cell lines were found to be 0.046 ± 0.0025 nmoles acetyl-CoA/1,000,000 cells, 0.031 ± 0.0026 nmoles acetyl-CoA/1,000,000 cells, and 0.049 ± 0.0072 nmoles acetyl-CoA/1,000,000 cells, respectively. Tukey's post-hoc test showed that only the difference in endogenous acetyl-CoA levels between the Down and Up group was statistically significant ($p < 0.05$).

METABOLOMICS

All metabolites were given a metabolite number (arbitrarily) to provide a unique identification since some metabolites appeared multiple times and others were unnamed. Of all

Figure 6

Doubling Time of Transformed MDA-MB-231 Cell Lines



Transformed MDA-MB-231 Cell Lines

Figure 6: Doubling Time of MDA-MB-231 transformed cell lines.

Determination of doubling time in the transformed cell lines. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. Bars represent results from three separate determinations with mean \pm SD. Statistical significance was tested by one-way ANOVA ($p=0.3850$).

Figure 7

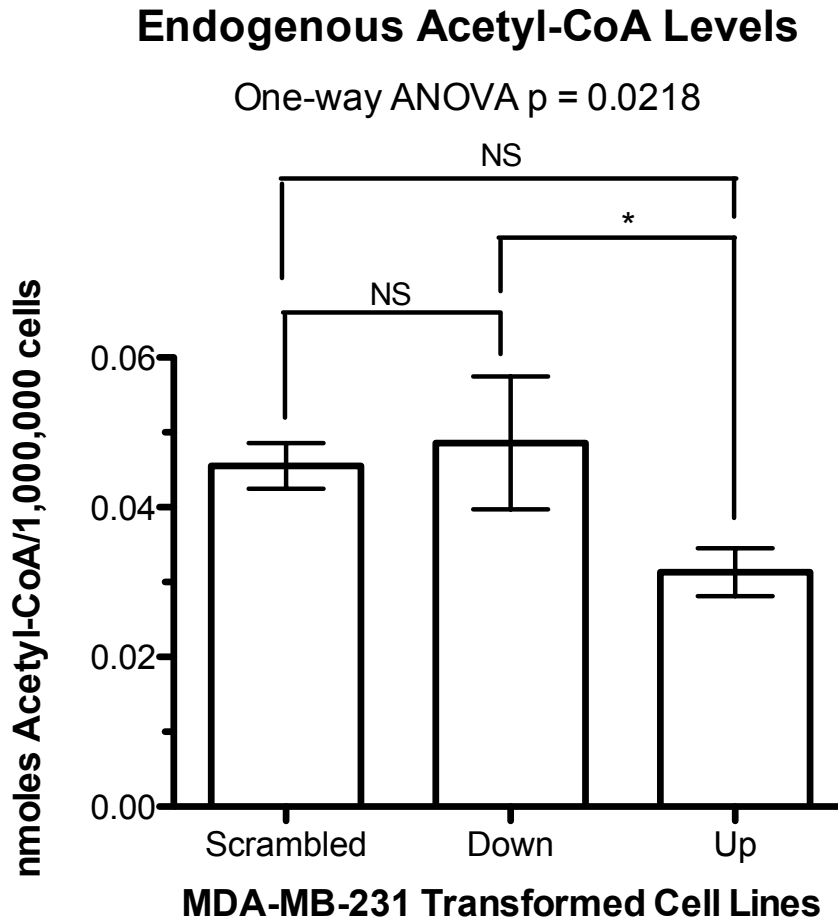


Figure 7: Endogenous Levels of Acetyl-CoA in the Three Transformed MDA-MB-231 Cell Lines.

Endogenous levels of acetyl-CoA were measured in via HPLC-UV in the three transformed MDA-MB-231 cell lines. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. The bars represent results from a single determination in triplicate in nanomoles of acetyl-CoA per one million cells \pm standard deviation. Bars represent results from a single determination. Statistical significance was tested by one-way ANOVA ($p=0.0218$) followed by a Tukey's post-hoc test. NS = not significant, * = ≤ 0.05

metabolites detected by GC×GC-TOF MS, 52.3% were tentatively identified based on mass spectrum and retention index matching. Following student's *t*-tests, 8.3% of metabolites detected by GC×GC-TOF MS had a *p*-value ≤ 0.05 and were verified with standards (Table 1). Most of these metabolites are amino acids. Two metabolites, L-asparagine and palmitoleic acid, were found to be significant across two student's *t*-test comparisons. Following a one-way ANOVA, 29.0% of tentatively identified metabolites had a *p* ≤ 0.05 . After technical validation, only 7.6% of metabolites had a one-way ANOVA *p* ≤ 0.05 and were verified with standards (Table 2). Of the metabolites we submitted for technical validation, 52.3% were either misidentified or had no standard available for comparison (Table 3).

After stratification of the metabolites by student's *t*-test *p*-value and fold-change via volcano plots, it was observed that L-asparagine and trans-9-octadeconic acid had both a *p*-value ≤ 0.05 and a fold-change of at least ± 1.5 between the Scrambled and Down groups (Figure 8); between the Scrambled and Up groups, L-threonine, palmitoleic acid, and L-asparagine had both a *p*-value ≤ 0.05 and a fold-change of at least ± 1.5 (Figure 9). From the volcano plot comparing the Down and Up groups, L-threonine, palmitoleic acid, and L-ornithine had both a *p*-value ≤ 0.05 and a fold-change of at least ± 1.5 (Figure 10). In the volcano plots, only those metabolites that had a both a *p*-value ≤ 0.05 and a fold-change of at least ± 1.5 were labeled.

Box plots were created for all metabolites that were identified to be significant and had a fold change of ± 1.5 between at least 2 of the groups. The metabolites L-leucine, L-aspartic acid, and L-valine display the same abundance distribution pattern between the Scrambled, Down, and Up groups (Figure 11). The Down group has the lowest abundance, the Scrambled group has a mid-level abundance, and the Up group has the highest abundance. The box plot for palmitoleic acid shows that the Up group has the lowest abundance while the Scrambled group has the highest abundance (Figure 12). Of the seven amino acids that are degraded to acetyl-CoA, only leucine, tyrosine, and threonine were verified with standards and showed a significant difference in abundance distribution between the three groups (Figure 13).

Table 1

	Metabolite #	Metabolite	t-test p-value	t-test q-value	Fold Change	¹ t _R (s)	² t _R (s)	CAS no.
Up vs Down*	M73	L-threonine	0.0006	0.01	-1.64	1744.578695	0.991514109	72-19-5
	M96	L-aspartic acid	0.0008	0.01	1.44	2243.252686	0.991516707	56-84-8
	M102	palmitoleic acid	0.006	0.04	-1.40	2375.717124	1.043582822	373-49-9
	M66	L-valine	<0.0001	0.004	1.36	1640	0.9275	72-18-4
	M69	L-leucine	0.0003	0.008	1.27	1693.142084	0.91843099	61-90-5
	M103	L-ornithine	0.03	0.1	1.73	2376.053736	0.962442604	70-26-8
Down vs Scrambled [#]	M106	L-asparagine	0.002	0.02	-1.54	2395.754445	1.064343458	70-47-3
	M113	trans-9-octadecenoic acid	0.01	0.07	-1.90	2573.274268	1.091159053	112-79-8
Up vs Scrambled [§]	M73	L-threonine	0.0009	0.01	-1.62	1744.578695	0.991514109	72-19-5
	M102	palmitoleic acid	0.001	0.02	-1.57	2375.717124	1.043582822	373-49-9
	M106	L-asparagine	0.002	0.02	-1.53	2395.754445	1.064343458	70-47-3

Table 1: Metabolites Altered by an Increase or Decrease in NAT1 Activity in Breast Cancer Cells

Metabolites differed significantly between designated groups following student's *t*-test, and were technically validated. *Down is the reference group, therefore a positive fold change indicates that metabolite was higher in the Up group whereas a negative value indicates the metabolite was lower in the Up group. [#]Scrambled is the reference group, therefore a positive fold change indicates that metabolite was higher in the Down group whereas a negative value indicates the metabolite was lower in the Down group. [§]Scrambled is the reference group, therefore a positive fold change indicates that metabolite was higher in the Up group whereas a negative value indicates the metabolite was lower in the Up group. ANOVA q-values were calculated to adjust p-values for the multiple comparisons we are testing. ¹t_R is the first dimension retention time in seconds and ²t_R is the second dimension retention time. The CAS number is a unique numerical identifier assigned to chemical compounds by the Chemical Abstracts Service (CAS).

Table 2

Metabolite #	Metabolite	ANOVA p-value	ANOVA q-value	¹ t _R (s)	² t _R (s)	CAS no.
M73	L-threonine	0.0002	0.0009	1744.578695	0.991514109	72-19-5
M106	L-asparagine	0.002	0.003	2395.754445	1.064343458	70-47-3
M66	L-valine	<0.0001	0.0004	1640	0.9275	72-18-4
M69	L-leucine	<0.0001	0.0004	1693.142084	0.91843099	61-90-5
M96	L-aspartic acid	0.0008	0.002	2243.252686	0.991516707	56-84-8
M84	L-methionine	0.05	0.03	2034.28795	1.052523863	63-68-3
M102	palmitoleic acid	0.0001	0.0006	2375.717124	1.0435882822	373-49-9
M65	beta alanine	0.03	0.02	1626	0.96	107-95-9
M77	L- proline	0.04	0.03	1789.30765	1.01797484	147-85-3
M80	glycerol	0.02	0.02	1921.391	0.87520052	56-81-5

Table 2: Metabolites Altered by an Increase and Decrease in NAT1 Activity in Breast**Cancer Cells**

Ten metabolites were found to differ significantly following a one-way ANOVA and were also technically validated with standards. These metabolites were included in subsequent pathway analysis. ANOVA q-values were calculated to adjust p-values for the multiple comparisons we are testing. ¹t_R is the first dimension retention time in seconds and ²t_R is the second dimension retention time. The CAS number is a unique numerical identifier assigned to chemical compounds by the Chemical Abstracts Service (CAS).

Table 3

Verification Status	Metabolite	Metabolite #	¹ t _R (s)	² t _R (s)
V	L-glutamine	M112	2519.1796	1.0830347
V	L-threonine	M73	1744.578695	0.991514109
V		M86	2097.140471	0.934089146
V	L-asparagine	M106	2395.754445	1.064343458
V	trans-9-octadecenoic acid	M113	2573.274268	1.091159053
V	L-valine	M66	1640	0.9275
V	L-leucine	M69	1693.142084	0.91843099
V	L-aspartic acid	M96	2243.252686	0.991516707
V	L-tyrosine	M119	2748	1.28
V	L-methionine	M84	2034.28795	1.052523863
V	palmitoleic acid	M102	2375.717124	1.043582822
V	malic acid	M94	2195.57733	0.9831648
V	fumaric acid	M78	1827.812518	0.97596729
V	beta-alanine	M65	1626	0.96
V	succinic acid	M75	1783.060069	1.017216049
V	citric acid	M117	2712	1.15
V	L-alanine	M57	1483.14905	0.925
V	pidolic acid	M83	2016.031208	1.15
V	L-proline	M77	1789..30765	1.01797484
V	glycerol	M80	1921.391	0.87520052
V	L-ornithine	M103	2376.053736	0.962442604

*continued on next page

Table 3-Continued

U	phosphoric acid	M87	2109.995458	0.914484667
U		M88	2069.295505	0.938482329
U		M93	2098.260648	0.924620606
U		M99	2313.63	1.039414761
U		M105	2420.345586	0.795902017
U		M107	2428.017212	0.791466107
U		M115	2664.376829	1.095429665
U		M128	1996.460852	0.989310526
U		M129	2063.537718	0.943364826
U		4-bromobutan-1-ol	M41	1194
U	cholesterol	M131	3135.889469	1.344001898
U		M132	3141.608204	1.362317152
U	adenine	M108	2415.504303	1.335384303
U	8-chlorooctan-1-ol	M123	1056.667071	0.961182658
M	L-glutamine	M116	2696.95982	1.052917405
M	oxalic acid	M92	2186.059441	1.025892253
M		M47	1277.936778	0.848586571
M	malic acid	M120	2769.82079	1.149911156
M		M109	2436.8759	0.93113281
M	beta-alanine	M79	1891.871716	1.073551252
M	pidolic acid	M90	2242.519873	0.954952753
M		M100	2353.592825	0.888429472
M	L-lysine	M110	2478.422412	0.977623334

Table 3: Tentatively Identified Metabolites Altered by Increased and/or Decreased NAT1**Activity Subjected to Technical Validation**

V = verified with standards

U = standard unavailable

M = retention time does not match standard therefore misidentified

Figure 8

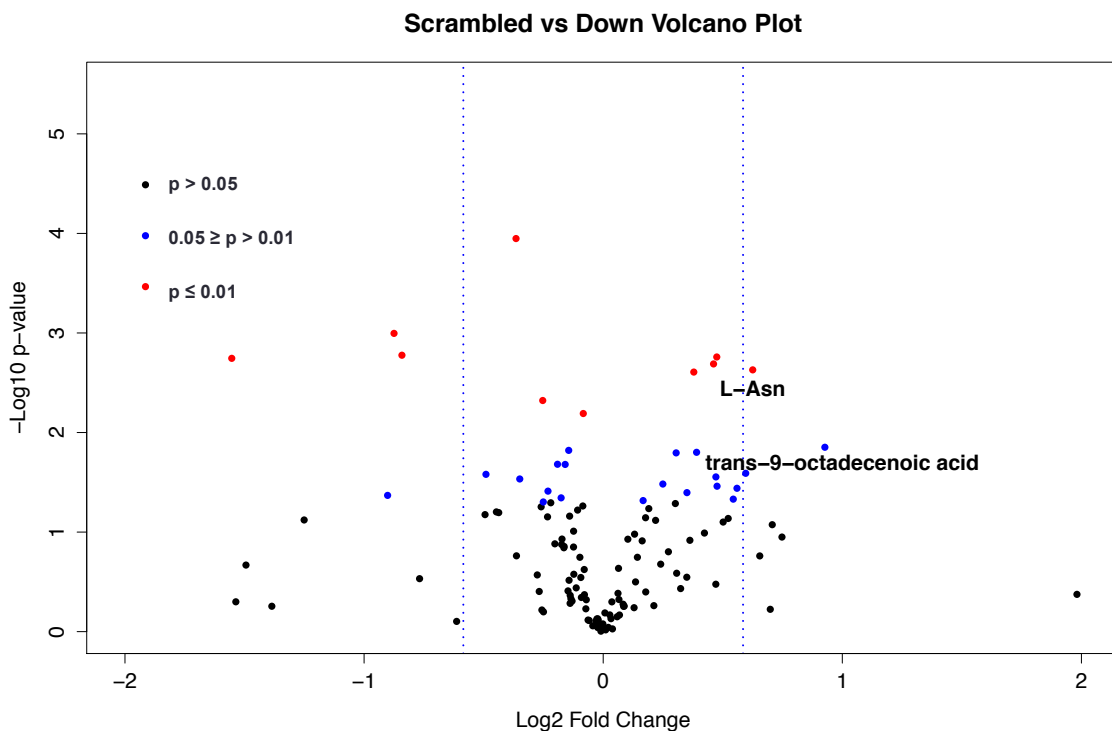


Figure 8: Volcano plot comparing metabolite fold-change and student t-test unadjusted p-value between the Scrambled and Down transformed cell lines.

Fold-change of metabolite abundance was calculated and plotted versus unadjusted student's *t*-test p-values to compare the transformed cell lines. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. Metabolites located at the top left and top right of the plot indicate that metabolite has both a substantial fold-change and a significant p-value. The dotted vertical blue lines indicate a fold-change of ± 1.5 .

Figure 9

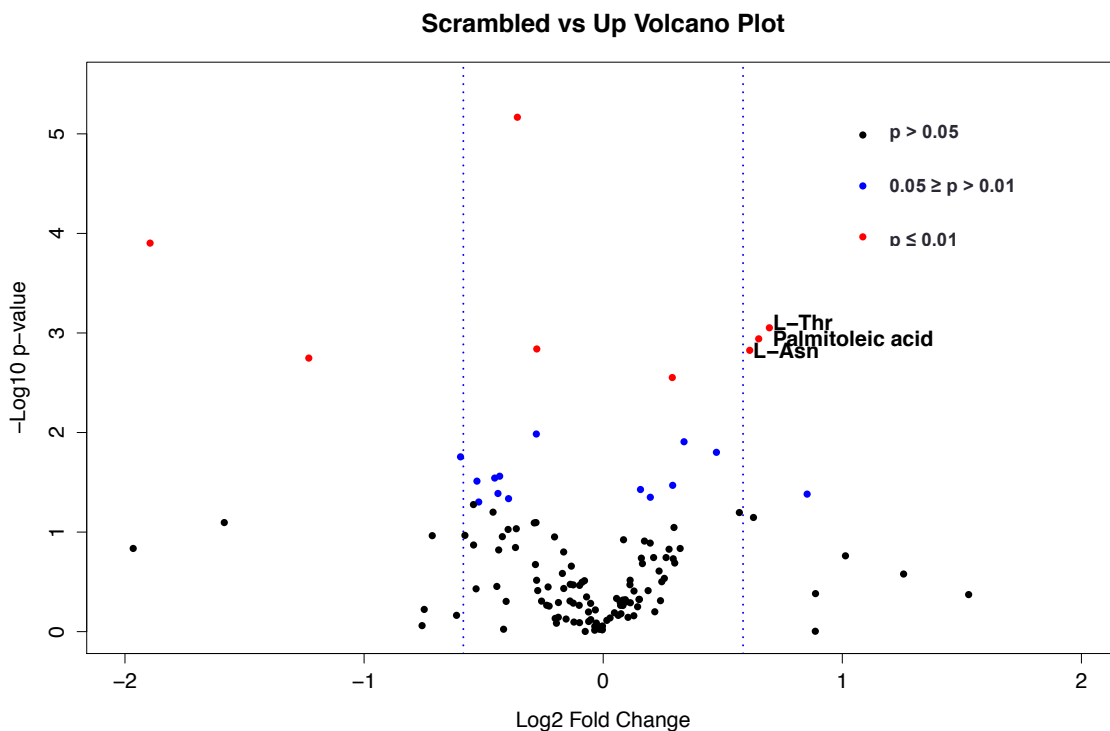


Figure 9: Volcano plot comparing metabolite fold-change and student t-test unadjusted p-value between the Scrambled and Up transformed cell lines.

Fold-change of metabolite abundance was calculated and plotted versus unadjusted student's *t*-test p-values to compare the transformed cell lines. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. Metabolites located at the top left and top right of the plot indicate that metabolite has both a substantial fold-change and a significant p-value. The dotted vertical blue lines indicate a fold-change of ± 1.5 .

Figure 10

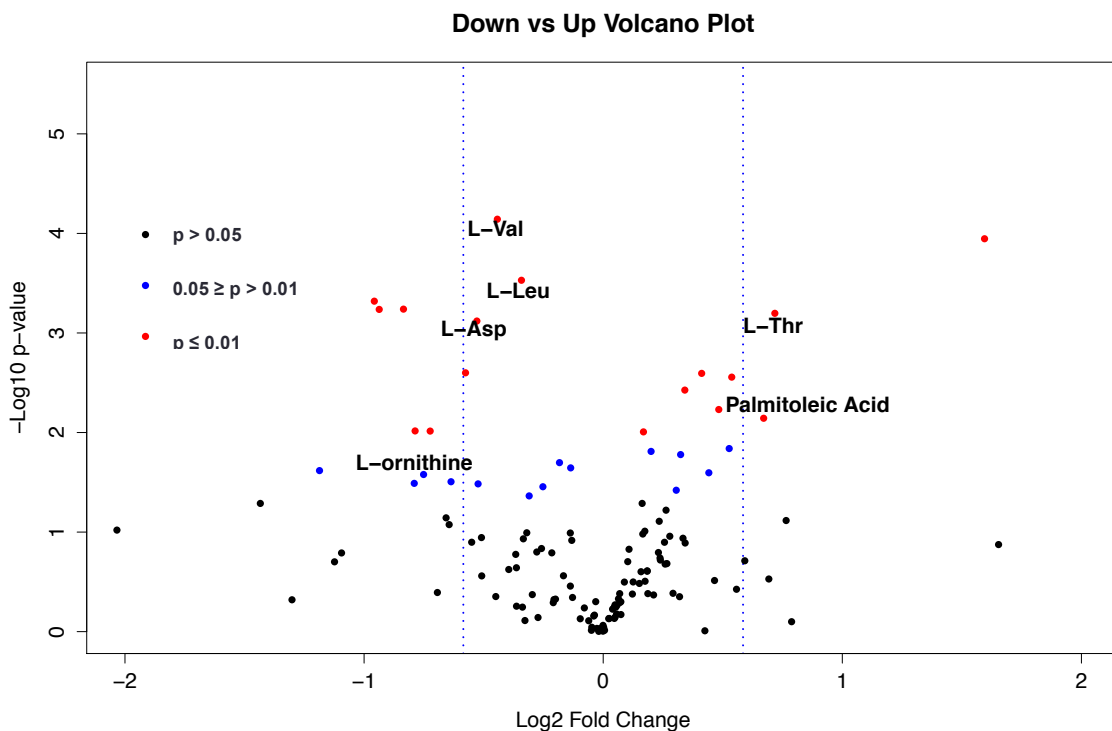


Figure 10: Volcano plot comparing metabolite fold-change and student t-test unadjusted p-value between the Down and Up transformed cell lines.

Fold-change of metabolite abundance was calculated and plotted versus unadjusted student's *t*-test p-values to compare the transformed cell lines. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. Metabolites located at the top left and top right of the plot indicate that metabolite has both a substantial fold-change and a significant p-value. The dotted vertical blue lines indicate a fold-change of ± 1.5 .

Figure 11

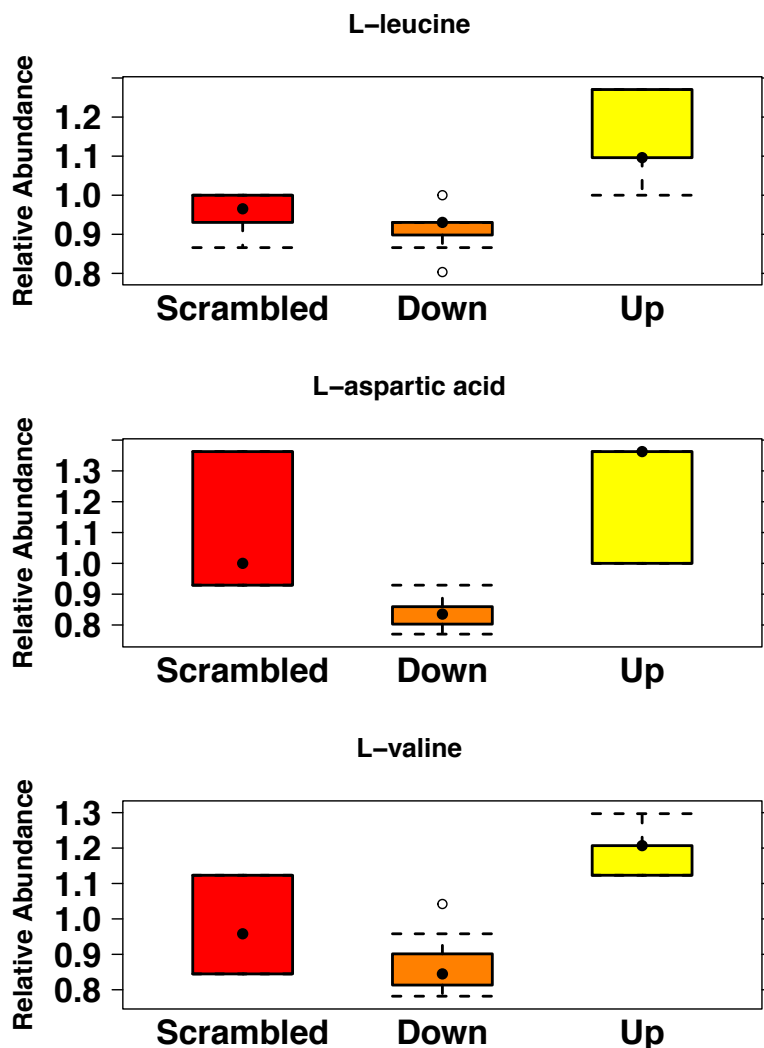


Figure 11: Box Plots Comparing the Abundance of Specific Metabolites Between the Transformed Cell Lines.

Box plots showing the abundance distribution of the metabolites L-leucine, L-aspartic acid, and L-valine in the three groups. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. Abundances were median scaled for each metabolite. The solid black dots represent the mean and the open dots represent outliers. These three metabolites all follow the same pattern of abundance distribution.

Figure 12

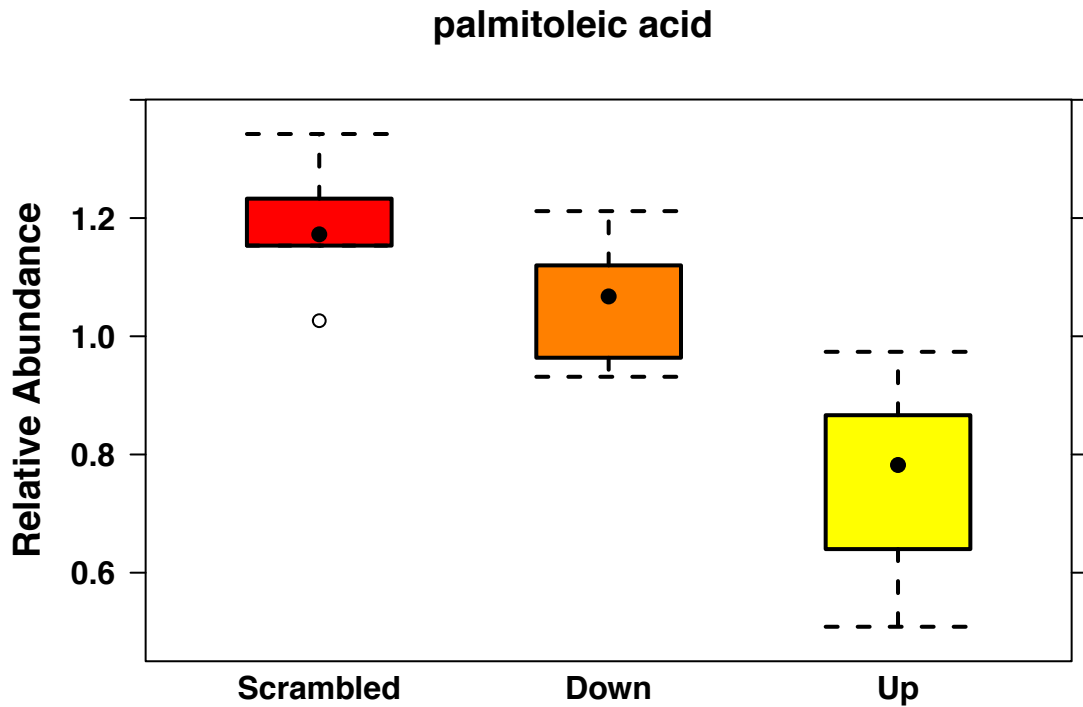


Figure 12: Box Plot Comparing the Abundance of the Metabolite Palmitoleic Acid Between the Transformed Cell Lines.

Box plot showing the abundance distribution of the metabolite palmitoleic acid in the three groups. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. The solid black dots represent the mean and the open dots represent outliers. Abundances were median scaled.

Figure 13

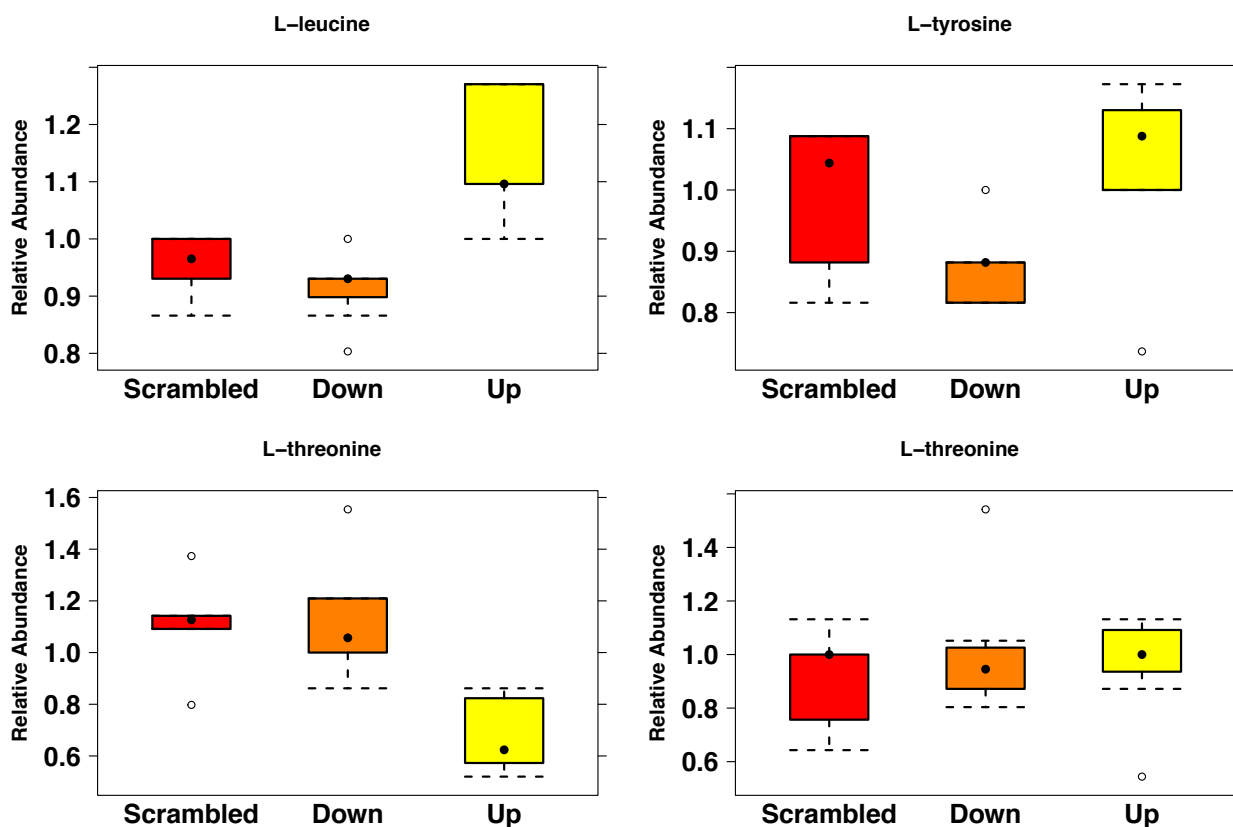


Figure 13: Box Plot of Amino Acids that are Broken Down to Form Acetyl-CoA.

Box plot showing the abundance distribution of amino acids that are degraded to form acetyl-CoA. In this figure, Scrambled refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, Down refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and Up refers to the MDA-MB-231 cell line that has increased NAT1 activity. The solid black dots represent the mean and the open dots represent outliers. Abundances were median scaled for each metabolite. L-leucine and L-tyrosine exhibit the same distribution pattern with metabolite levels increased in the Up group when compared to the Scrambled and Down groups. The two L-threonine box plots represent different derivitizations of that amino acid.

MULTIVARIATE/MULTIVARIABLE ANALYSIS

The goal of using multivariate/multivariable analyses methods is to identify groups of metabolites that, together, are driving between group variation versus within-group variation. Principal Component Analysis (PCA) was the first multivariate approach used. Principal component 1 (PC1) of PCA separated the Down group from the Scrambled and Up groups but two component PCA did not discriminate between the Scrambled and Up groups (Figure 14). Therefore, Partial Least Squares- Discriminate Analysis (PLS-DA) methods were used in an attempt to achieve separation between all three groups with a single principle component. Two component PLS-DA led to complete group separation but within group variation and between group variation were both contributing to each component (Figure 15). We then proceeded to the multivariate analysis method, Orthogonal Partial Least Squares- Discriminate Analysis (OPLS-DA). Two component OPLS-DA showed complete group separation with component 1 explaining group predictive variation and component 2 explaining variation uncorrelated with group membership (Figure 16). Since component 1 of the OPLS-DA showed complete group separation, the loadings of that component were plotted to show which metabolites were contributing the most to the separation (Figure 17). The loadings of component 1 from OPLS-DA show that L-threonine, L-glutamine, and oxalic acid are driving the separation between the groups and are higher in the Down group and conversely lower in the Up group. The loadings of component 1 from OPLS-DA also show L-valine, L-leucine, and L-aspartic acid are driving the separation between the groups but are higher in the Up group and conversely lower in the Down group.

Figure 14

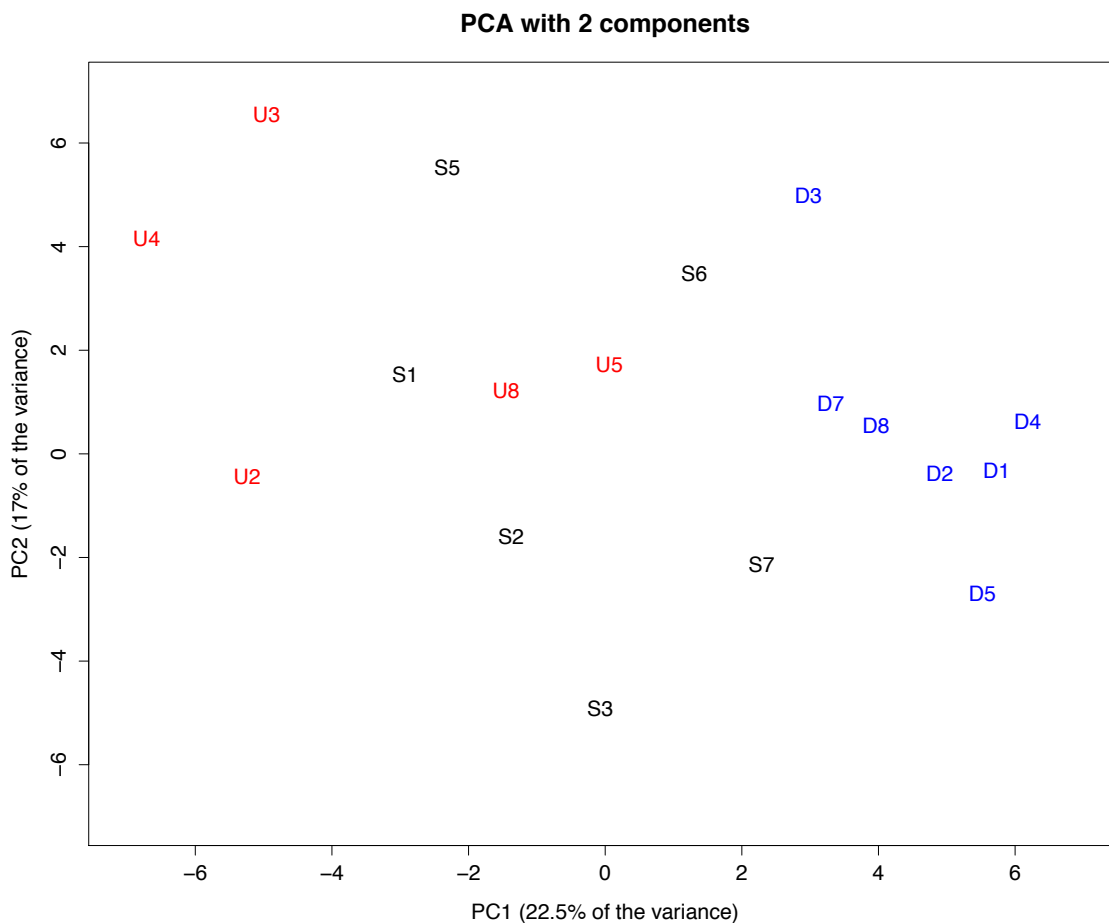


Figure 14: Principal Component Analysis (PCA) Scores Plot

Two component PCA explains 40% of the total variance. In this figure, S (in black) refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, D (in blue) refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and U (in red) refers to the MDA-MB-231 cell line that has increased NAT1 activity. The number following each letter indicates the individual sample number. Each component on the axes represents a linear combination of all metabolites.

Figure 15

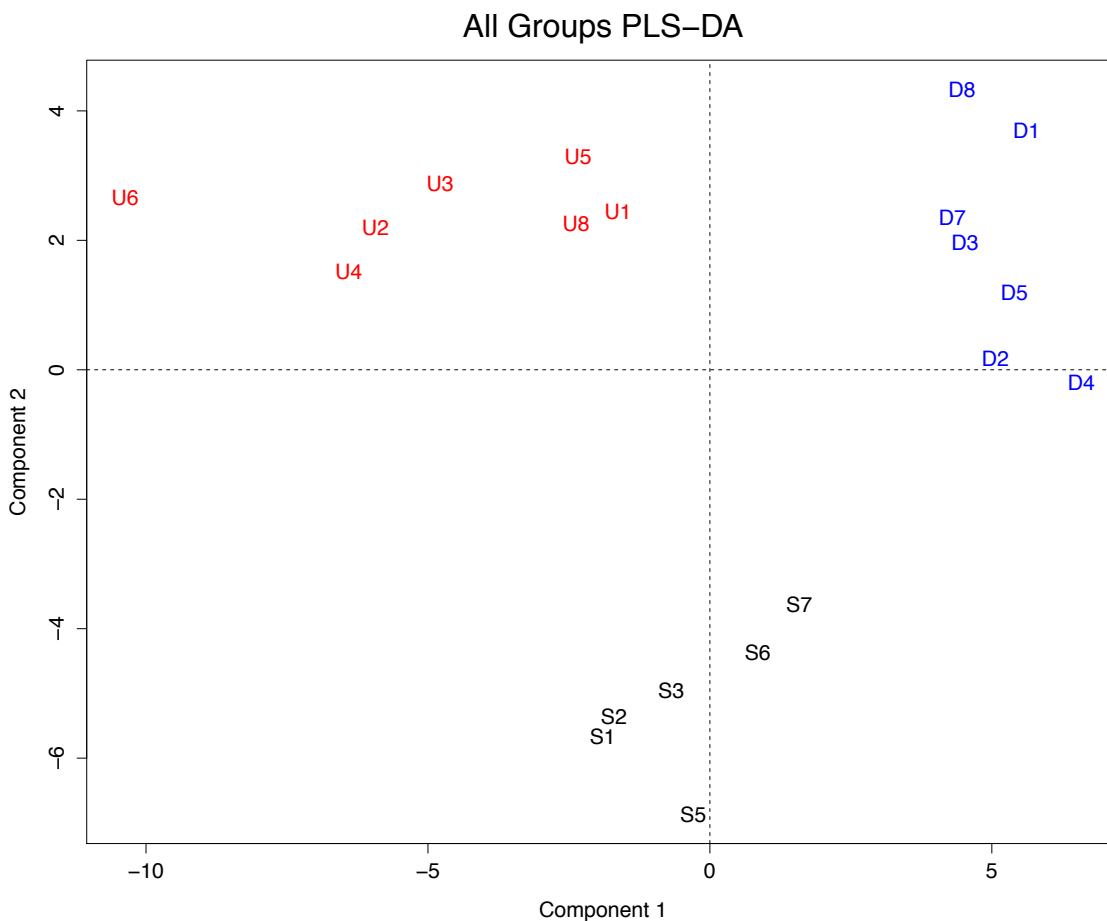


Figure 15: Partial Least Squares-Discriminate Analysis (PLS-DA) Scores Plot

Two component PLS-DA showed separation between the groups. In this figure, S (in black) refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, D (in blue) refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and U (in red) refers to the MDA-MB-231 cell line that has increased NAT1 activity. The number following each letter indicates the individual sample number. Each component represents a linear combination of all metabolites.

Figure 16:

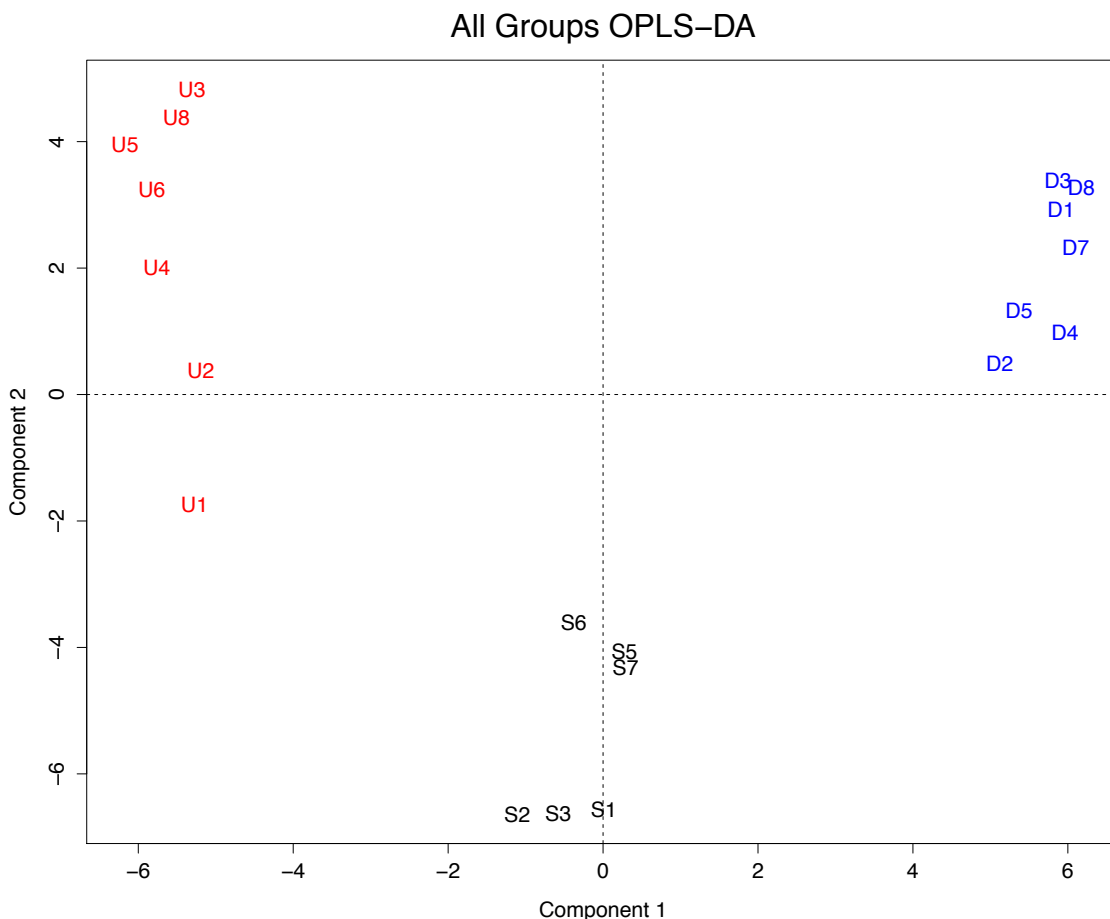


Figure 16: Orthogonal Partial Least Squares Discriminate Analysis (OPLS-DA) Scores Plot

OPLS-DA achieved complete separation between all groups. In this figure, S (in black) refers to the MDA-MB-231 cell line that has parental levels of NAT1 activity, D (in blue) refers to the MDA-MB-231 cell line that has decreased NAT1 activity, and U (in red) refers to the MDA-MB-231 cell line that has increased NAT1 activity. The number following each letter indicates the individual sample number. Each component represents a linear combination of all metabolites.

Figure 17

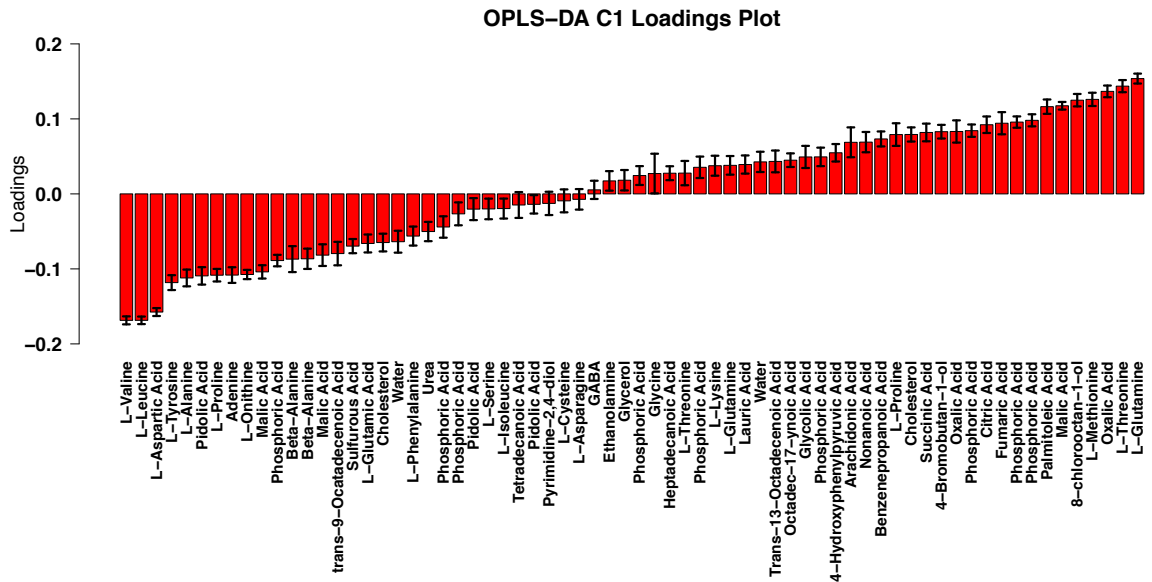


Figure 17: Jack-knife loadings plot for component 1 of the OPLS-DA.

The metabolites at the extremes of the plot are the largest contributors to the separation of the groups. Error bars represent mean \pm SD.

DISCUSSION

The purpose of this study was to identify metabolic changes induced by varying levels of human NAT1 in MDA-MB-231 breast cancer cells. It is our hypothesis that changing the level of NAT1 activity will lead to alterations in the levels of acetyl-coenzyme A (acetyl-CoA). Increased NAT1 activity would lead to increased hydrolysis of acetyl-CoA therefore less free acetyl-CoA would be available. Conversely, decreased NAT1 activity would lead to decreased hydrolysis of acetyl-CoA therefore more free acetyl-CoA would be available. It is our hypothesis that due to fluctuations in the amount of free acetyl-CoA available, cellular pathways will be altered. Acetyl-CoA is a central mediator of many cellular pathways (Figure 18).

Many of the breast cancer cell lines used in research have innately different levels of human NAT1 activity. Wakefield et al. have reported that the NAT1 activity in breast cancer cell lines range from 0.1 ± 0.02 nmol PABA acetylated/min/mg protein in Cal51 cells to 202.2 ± 28 nmol PABA acetylated/min/mg protein in ZR-75-1 cells [34]. With our transformed MDA-MB-231 cell lines, not only do they express varying levels of NAT1 activity, ranging from 11.6 ± 0.6 nmol PABA acetylated/min/mg protein in the Down cell line to 233.9 ± 2.2 nmol PABA acetylated/min/mg protein in the Up cell line, but they also have the same genetic background. This removes any confounding factors due to genetic differences that would arise if we performed this same study using the cell lines that innately have different NAT1 activity. Ideally, we wanted the Down cell line to have little to no NAT1 activity but when utilizing shRNA that targets NAT1 the most we have been able to knock-down NAT1 activity has been by 50% when compared to the parent cell line. We have recently constructed two MDA-MB-231 complete NAT1 knockout cell lines in our lab using the CRISPR® system that will be used in further studies (see future directions for more details).

Figure 18

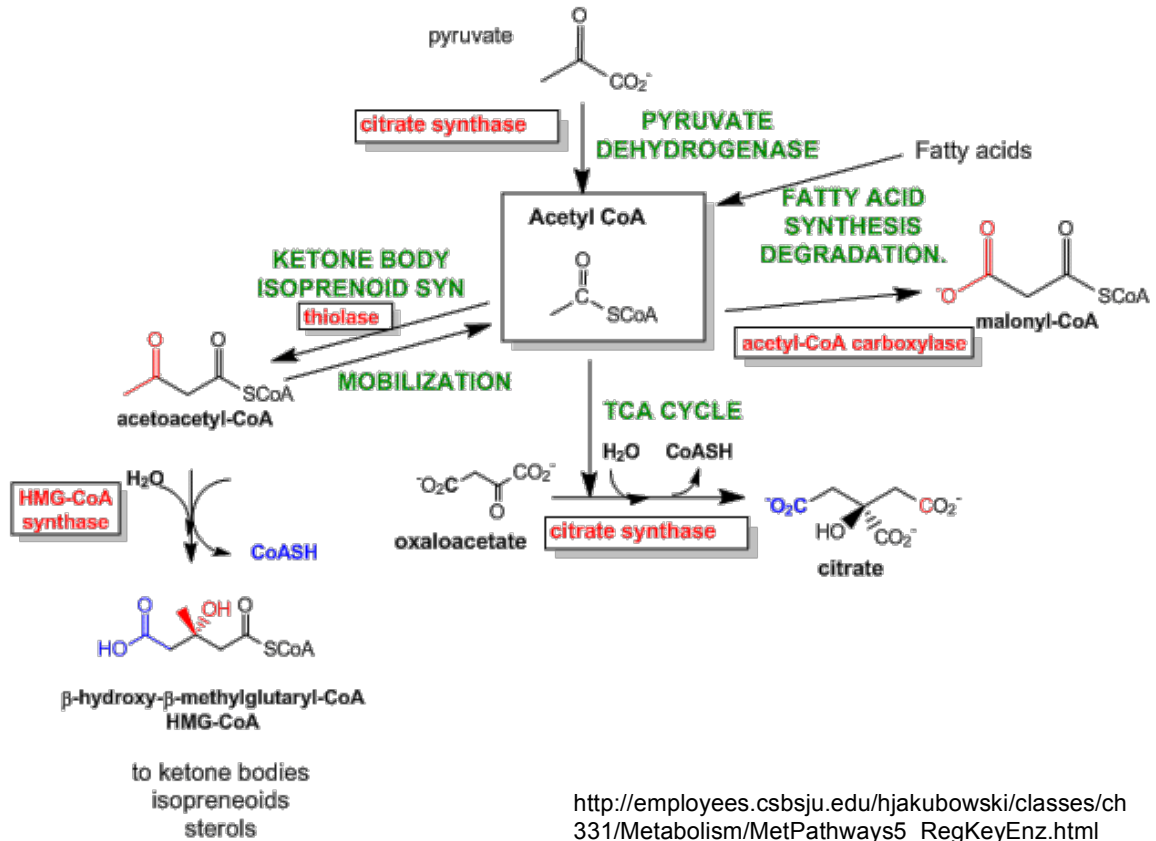


Figure 18- Acetyl-CoA is a Central Molecule Involved in a Plethora of Reactions.

Acetyl-CoA is a central molecule involved in many pathways including fatty acid synthesis and degradation, the citric acid cycle (TCA Cycle), and ketone body and isoprenoid synthesis.

In our characterization of the transformed cell lines, we found that relative to the Scrambled cell line, the Down cell line expressed 50% of the NAT1 mRNA and the Up cell line expressed 31 times the NAT1 mRNA. We also found that the Down cell line had 61% of the NAT1 activity and the Up cell line had 1240% the NAT1 activity of the Scrambled cell line. The difference in NAT1 mRNA expression between the Scrambled and Down groups was not statistically significant by a Tukey's post-hoc test but, the difference in NAT1 activity between the Scrambled and Down groups was statistically significant by a Tukey's post-hoc test. We postulate that this is due to the difference in standard error of the mean (SEM) between the two measurements. The SEM is much larger for the NAT1 mRNA measurement in the Scrambled group when compared to the SEM for the NAT1 activity measurement in the same group.

When we began this study it was our hypothesis that changing the level of NAT1 activity would change the level of free acetyl-CoA available and this change in acetyl-CoA levels would thus induce alterations in cellular pathways. In this study we observed that the levels of endogenous acetyl-CoA are decreased in the transformed MDA-MB-231 cell line expressing increased NAT1 acetylation activity when compared to the transformed MDA-MB-231 cell line expressing parental levels of NAT1 acetylation activity. This is the first study to compare endogenous levels of acetyl-CoA in samples that express varying levels of NAT1. This finding, combined with our original hypothesis, led us to begin the pathway analysis portion of this study by first following pathways that involved acetyl-CoA and identifying where the metabolites that were classified as significant and were verified with standards were located in those pathways. We then applied our hypothesis to see if the change in abundance of metabolite were biologically plausible. In the metabolomics portion of this study, the abundance of palmitoleic acid was shown to be decreased in the Up group when compared to the Scrambled group with a student's t-test p-value < 0.005 and a fold change of 1.57. Our refined hypothesis is that in the Up group, acetyl-CoA hydrolysis by NAT1 is increased when compared to the Scrambled group, leaving less free acetyl-CoA available for other reactions.

We expected to see an increase in the flux of acetyl-CoA through the fatty acid synthesis pathway in the cell line that had increased NAT1 activity since cancer cells require increased

amounts of cellular components and energy due to their rapidly dividing nature and overexpression of NAT1 in cell lines has been shown to increase cell growth and invasion. The fatty acid synthesis pathway (Figure 19) begins with the synthesis of malonyl-CoA by acetyl-CoA carboxylase (ACC) from acetyl-CoA and bicarbonate. Acetyl-CoA carboxylase (ACC) is a biotin-dependent enzyme that catalyzes the irreversible carboxylation of acetyl-CoA to produce malonyl-CoA through its two catalytic activities. In the next step of fatty acid synthesis, malonyl-CoA and acetyl-CoA are combined by fatty acid synthase I (FAS1) to form palmitate (16:0). Palmitate is a branch point in the fatty acid synthesis pathway. Either stearate (18:0) or palmitoleic acid (16:1 (Δ^9)) can be synthesized from palmitate. Palmitoleic acid is synthesized from palmitate by fatty acyl-CoA desaturase. Palmitate negatively feeds back on acetyl-CoA carboxylase (ACC) thus preventing further palmitate generation. In this study we saw a decrease in the abundance of palmitoleic acid (16:1 (Δ^9)) when NAT1 activity was increased. We now propose that the abundance of palmitoleic acid is decreased in the cell line that has increased NAT1 activity because there is less free acetyl-CoA available therefore palmitate is selectively shunted to the other branch of the fatty acid synthesis pathway, which synthesizes stearate. To test this we need to measure the levels of stearate to see how they compare between the three cell lines.

There are seven amino acids that are degraded to acetyl-CoA; these amino acids are 1. tryptophan, 2. lysine, 3. phenylalanine, 4. tyrosine, 5. leucine, 6. isoleucine, and 7. threonine. If our hypothesis that different levels of human NAT1 in the MDA-MB-231 cells lead to altered levels of acetyl-CoA and these altered levels of acetyl-CoA lead to alterations in cellular pathways is correct, it is plausible that there could be increased or decreased degradation of one or more of these seven amino acids to compensate for the altered levels of acetyl-CoA. In our data, leucine, tyrosine, and threonine were verified with standards and showed a significant difference in abundance distribution between the three groups. Of those three amino acids, leucine and

Figure 19

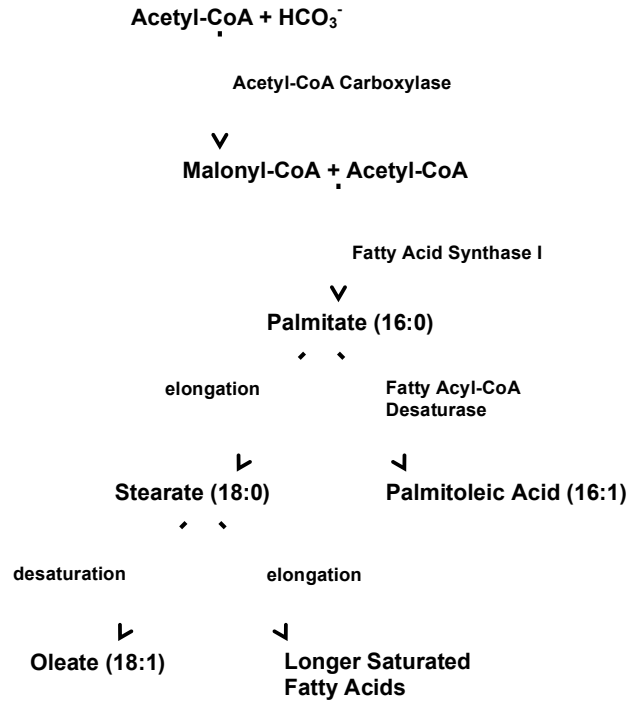


Figure 19: Abbreviated Fatty Acid Synthesis Pathway.

This pathway shows the utilization of Acetyl-CoA in the fatty acid synthesis pathway. Palmitoleic acid is synthesized from palmitate by the enzyme fatty acyl-CoA desaturase. Palmitate can also be elongated to form stearate which can then be further elongated to form longer saturated fatty acids. The numbers in parentheses next to the names of the fatty acids are lipid numbers which have the form $C:D$, where C is the number of carbon atoms and D is the number of double bonds in the fatty acid.

tyrosine had the same relative abundance distribution suggesting there may be a shared mechanism by which these metabolites are being regulated.

As with any study involving a high number of comparisons, there is the chance that some of the observed changes in abundance of metabolites are due to random chance. We have decreased the likelihood of pursuing metabolites where this may be the case by choosing metabolites for further investigation that are involved in pathways that also involve acetyl-CoA. By doing this we can develop a rationale for why the abundance of that metabolite would be changed between cells expressing varying levels of NAT1 activity.

In the metabolomics data there were several misidentified metabolites that were both statistically significant and had a fold change of 1.5 or greater. One such metabolite, M116, with a one-way ANOVA $p < 0.0001$ was initially identified as L-glutamine but was later shown to be misidentified when compared to the L-glutamine standard. Unfortunately, there is currently not a way for us to retroactively identify these metabolites; we can only see that they are misidentified. As a result we may be missing important metabolites that could provide a key link between varying levels of NAT1 and altered cellular metabolism. As the field of metabolomics progresses, libraries containing spectral data of metabolites will become more complete and lead to a larger number of identified metabolites in future studies.

If our proposed hypothesis holds true, we would expect the abundance distribution of any metabolite that is changed, as a result of the level of NAT1 activity, to show the Scrambled group in the middle with the Up and Down groups above and below or below and above. This is not necessarily true though as there may be compensatory mechanisms that are regulating the amount of acetyl-CoA. Also, there may be a threshold in the amount of acetyl-CoA that must be altered in order for a change in metabolite abundance to be observed.

Another concern in metabolomics studies is the threshold of detection. Some metabolites are present in cells at levels that are too small to be accurately measured or even detected but that does not mean the abundance of those metabolites are not changing. A small change in these metabolites could have a greater impact than metabolites that are present in the cells at higher abundances and can be easily measured.

SUMMARY AND CONCLUSIONS

STRENGTHS OF THIS WORK

The construction of three stably transfected cell lines provided stable breast cancer cell line models whose only genetic difference is the level of human NAT1 activity. Before the construction of these cell lines, comparison between cell lines that expressed different levels of NAT1 activity was limited to evaluating cell lines whose endogenous NAT1 activity varied. In the cell lines that endogenously express varying levels of NAT1 activity there are also other differences that have to be taken into account, including mutation differences such as estrogen receptor (ER), progesterone receptor (PR), and Her2/neu status, among many others, when comparing the results of those studies. With these new cell lines, the effect NAT1 activity has on cancer progression can be more thoroughly evaluated with less confounding factors. We have utilized this advantageous new resource to show for the first time that endogenous acetyl-CoA levels vary between samples whose only difference is the level of NAT1 activity. Acetyl-CoA is a central molecule involved in a plethora of reactions, including fatty acid synthesis and the TCA cycle.

Also, this study is the first to investigate the effect of varying human NAT1 activity on the metabolome of a breast cancer cell line. Metabolomics has developed into an extremely beneficial and information-rich area of research as of late, but has been under-utilized in regards to research on the NAT isozymes. Analysis of metabolomics data can help reveal what pathways are altered as a result of varying levels of NAT1 activity thus providing insights into the role NAT1 has in breast cancer disease and progression. A greater understanding of the role NAT1 has in breast cancer could lead to better detection and treatment methods.

CAVEATS AND WEAKNESSES

MDA-MB-231 triple-negative breast cancer cells were chosen for manipulation of NAT1 activity and subsequent analyses in this study. MDA-MB-231 cells are known to have mutations in 5 genes, including BRAF, CDKN2A, KRAS, NF2, and TP53. These mutations may complicate evaluation of the effect NAT1 has on cancer risk. Since these cells are already cancerous in nature, evaluating the risk NAT1 levels play in cancer risk is muddled at best. A fundamental complementation to this study would include the manipulation of human NAT1 activity in other breast cancer cell lines and cells from normal breast tissue.

Comparison of NAT1 activity between the three transformed cell lines revealed that while the Up cell line has approximately thirteen times the NAT1 activity of the Scrambled cell line, the NAT1 activity in the Down cell line is only decreased by approximately 50%. The difference in NAT1 activity between the Scrambled and Up cell lines is so drastic that it may overshadow differences between the Scrambled and Down cell lines. Comparison of endogenous acetyl-CoA levels also only showed a statistically significant difference between the Scrambled and Up cell lines. We were hoping to completely knockdown the NAT1 activity in our down cell line however an approximate 50% decrease in activity was the most that could be achieved when utilizing our methods. Additional methods that knockout rather than knockdown the NAT1 gene have helped us achieve a cell line with no NAT1 activity, that will be used in future studies.

An additional caveat to this study is the use of cell lines as opposed to patient derived tumor samples. Tumors are not made up of a single cell-type but multiple cell types that each play a unique role while cell lines are a homogenous collection of cells. Also, these multiple cell types found in tumors exist in a tumor micro-environment (TME) that plays a big role in cancer progression. It has been shown that cancer cells secrete paracrine growth factors that recruit stromal cells then the stromal cells secrete cytokines which accelerate the aggressiveness of cancer cells [54]. It has also been shown that the TME allows complex metabolic interactions between the tumor and surrounding cell types. Unfortunately, since we are using cell lines as our samples in this study, we may be missing out on important metabolic changes due to the TME.

The samples prepared for metabolomics were not incubated with any NAT1 substrates. Humans are exposed to NAT1 substrates frequently, therefore the condition under which the cells were collected may not represent the most biologically relevant conditions. However, this does allow us to investigate a possible endogenous role for NAT1. Fellow researchers who work with NAT1 have postulated that NAT1 has an undiscovered endogenous role given that it is distributed throughout most tissues of the body and found in almost every species.

In cell lines, metabolites can be found both within the cells themselves and also within the medium the cells have been cultured in since some metabolites are excreted from the cells. In this study we chose to collect only the cell pellets for metabolomics sample preparation and discarded the media they were cultured in. The metabolites excreted into the culture medium may also provide valuable insights into cellular pathway alternations that were missed by analyzing metabolites from cell pellets only. We may have missed important differences in the metabolites that were excreted into the media.

In this study we only extracted and examined polar metabolites from the cell pellet samples. Examples of polar metabolites include sugars and most amino acids while examples of non-polar metabolites include membrane lipids and long-chain fatty acids. We expect to observe a greater difference in non-polar metabolites than polar metabolites since acetyl-CoA is a central mediator of fatty acid synthesis and we hypothesize that varying levels of NAT1 activity influence cellular metabolism through regulation of acetyl-CoA. By increasing the number of platforms used to detect metabolites in our samples, a wider range of metabolites can be detected thus giving a more complete picture of any metabolic flux or alterations that are occurring.

FUTURE DIRECTIONS

In the future we will be performing further metabolomics studies that cover a wider range of metabolites, including both polar and non-polar metabolites. This will be accomplished by applying a comprehensive metabolomics approach that will involve measurement of metabolites across six different platforms including 1. GCXGC-TOF MS, 2. LC-MS, 3. UHPLC-MS/MS (+ESI), 4. UHPLC-MS/MS (-ESI), 5. UHPLC-MS/MS (HILIC), and 6. GC-MS (+EI). Once metabolic differences between the cell lines are identified, we will then use biochemical approaches such as

high pressure liquid chromatography-ultraviolet detection (HPLC-UV) and spectrophotometric enzyme-coupled assays to verify the metabolites' identities and then hypothesize a mechanism by which NAT1 contributed to these differences. The biochemical approach chosen will depend on the metabolite.

Our lab has recently completed the creation of heterozygous and homozygous NAT1 knockout MDA-MB-231 cells using the CRISPR™ system. The CRISPR™ system utilizes guide RNA's (gRNA's) that specifically target one's gene of interest, which in our case is human NAT1, and induce DNA damage in that gene. The DNA damage and subsequent repair mechanisms lead to mutations in the gene which result in the knockout of the gene. These two additional cell lines with altered NAT1 activity are an incredibly valuable addition to the three transformed cell lines used for this project. Further metabolomics studies will include two CRISPR™ homozygous NAT1 knockout MDA-MB-231 cell lines that were constructed using two different gRNA's. By using MDA-MB-231 NAT1 knockouts that were generated using two different gRNA's we will be able to differentiate changes in metabolites that are due to NAT1 levels as opposed to possible off target effects of the gRNA's. We are also working on creating NAT1 knockout MCF-7 and ZR-75-1 cells using the CRISPR™ system.

Also, our lab is currently in the process of using the FLP-In System™ to construct MCF-7, MCF-10a, and ZR-75-1 transfected breast cancer cell lines that stably overexpress and underexpress human NAT1. By repeating our study with these additional cell lines, we will be able to show that the observed changes from our initial study are repeatable and not merely a consequence of using MDA-MB-231 cells as the parental cell line.

REFERENCES

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2015*. CA Cancer J Clin, 2015. **65**(1): p. 5-29.
2. DeSantis, C., et al., *Breast cancer incidence rates in U.S. women are no longer declining*. Cancer Epidemiol Biomarkers Prev, 2011. **20**(5): p. 733-9.
3. National Cancer Institute. *PDQ® Breast Cancer Treatment*. 10/24/2014 [cited 2014 11/18]; Available from: <http://cancer.gov/cancertopics/pdq/treatment/breast/Patient>.
4. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications*. Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
5. Munirah, M.A., et al., *Identification of different subtypes of breast cancer using tissue microarray*. Rom J Morphol Embryol, 2011. **52**(2): p. 669-77.
6. Hein, D.W., et al., *Molecular genetics and epidemiology of the NAT1 and NAT2 acetylation polymorphisms*. Cancer Epidemiol Biomarkers Prev, 2000. **9**(1): p. 29-42.
7. Hein, D.W., *Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis*. Mutat Res, 2002. **506-507**: p. 65-77.
8. Matas, N., et al., *Mapping AAC1, AAC2 and AACP, the genes for arylamine N-acetyltransferases, carcinogen metabolising enzymes on human chromosome 8p22, a region frequently deleted in tumours*. Cytogenet Cell Genet, 1997. **77**(3-4): p. 290-5.
9. Blum, M., et al., *Human arylamine N-acetyltransferase genes: isolation, chromosomal localization, and functional expression*. DNA Cell Biol, 1990. **9**(3): p. 193-203.
10. Sim, E., et al., *Arylamine N-acetyltransferase*. Biochem Soc Trans, 1992. **20**(2): p. 304-9.
11. Liu, L., et al., *Arylamine N-acetyltransferases: characterization of the substrate specificities and molecular interactions of environmental arylamines with human NAT1 and NAT2*. Chem Res Toxicol, 2007. **20**(9): p. 1300-8.
12. Minchin, R.F., *Acetylation of p-aminobenzoylglutamate, a folic acid catabolite, by recombinant human arylamine N-acetyltransferase and U937 cells*. Biochem J, 1995. **307** (Pt 1): p. 1-3.
13. Weber, W.W. and D.W. Hein, *N-acetylation pharmacogenetics*. Pharmacol Rev, 1985. **37**(1): p. 25-79.
14. DeBruin, L.S., J.B. Pawliszyn, and P.D. Josephy, *Detection of monocyclic aromatic amines, possible mammary carcinogens, in human milk*. Chem Res Toxicol, 1999. **12**(1): p. 78-82.
15. Possanzini, M. and V. Dipalo, *Improved Hplc Determination of Aliphatic-Amines in Air by Diffusion and Derivatization Techniques*. Chromatographia, 1990. **29**(3-4): p. 151-154.
16. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans., World Health Organization., and International Agency for Research on Cancer., *Some aromatic amines, organic dyes, and related exposures*. IARC monographs on the evaluation of carcinogenic risks to humans., 2010, Lyon: IARC Press. viii, 692 p.
17. Laurieri, N., et al., *From arylamine N-acetyltransferase to folate-dependent acetyl CoA hydrolase: impact of folic acid on the activity of (HUMAN)NAT1 and its homologue (MOUSE)NAT2*. PLoS One, 2014. **9**(5): p. e96370.
18. Clark, D.W., *Genetically determined variability in acetylation and oxidation. Therapeutic implications*. Drugs, 1985. **29**(4): p. 342-75.
19. Eichelbaum, M., H.K. Kroemer, and G. Mikus, *Genetically-Determined Differences in Drug-Metabolism as a Risk Factor in Drug Toxicity*. Toxicology Letters, 1992. **64-5**: p. 115-122.

20. Ohno, M., et al., *Slow N-acetyltransferase 2 genotype affects the incidence of isoniazid and rifampicin-induced hepatotoxicity*. Int J Tuberc Lung Dis, 2000. **4**(3): p. 256-61.
21. Butcher, N.J. and R.F. Minchin, *Arylamine N-acetyltransferase 1: a novel drug target in cancer development*. Pharmacol Rev, 2012. **64**(1): p. 147-65.
22. Casey, T., et al., *Molecular signatures suggest a major role for stromal cells in development of invasive breast cancer*. Breast Cancer Res Treat, 2009. **114**(1): p. 47-62.
23. Chin, K., et al., *Genomic and transcriptional aberrations linked to breast cancer pathophysiology*. Cancer Cell, 2006. **10**(6): p. 529-41.
24. Yuan, Y., et al., *A sparse regulatory network of copy-number driven gene expression reveals putative breast cancer oncogenes*. IEEE/ACM Trans Comput Biol Bioinform, 2012. **9**(4): p. 947-54.
25. Adam, P.J., et al., *Arylamine N-acetyltransferase-1 is highly expressed in breast cancers and conveys enhanced growth and resistance to etoposide in vitro*. Mol Cancer Res, 2003. **1**(11): p. 826-35.
26. Smid, M., et al., *Genes associated with breast cancer metastatic to bone*. J Clin Oncol, 2006. **24**(15): p. 2261-7.
27. Vatsis, K.P. and W.W. Weber, *Structural heterogeneity of Caucasian N-acetyltransferase at the NAT1 gene locus*. Arch Biochem Biophys, 1993. **301**(1): p. 71-6.
28. Weber, W.W. and K.P. Vatsis, *Individual variability in p-aminobenzoic acid N-acetylation by human N-acetyltransferase (NAT1) of peripheral blood*. Pharmacogenetics, 1993. **3**(4): p. 209-12.
29. Grant, D.M., et al., *Human acetyltransferase polymorphisms*. Mutat Res, 1997. **376**(1-2): p. 61-70.
30. N-acetyltransferase Gene Nomenclature Committee. *Human NAT1 Alleles (Haplotypes)*. 1-11-13 [cited 2015 Jan 30]; Available from: http://nat.mbg.duth.gr/Human NAT1 alleles_2013.htm.
31. Badawi, A.F., et al., *Role of aromatic amine acetyltransferases, NAT1 and NAT2, in carcinogen-DNA adduct formation in the human urinary bladder*. Cancer Res, 1995. **55**(22): p. 5230-7.
32. Bell, D.A., et al., *Polyadenylation polymorphism in the acetyltransferase 1 gene (NAT1) increases risk of colorectal cancer*. Cancer Res, 1995. **55**(16): p. 3537-42.
33. de Leon, J.H., K.P. Vatsis, and W.W. Weber, *Characterization of naturally occurring and recombinant human N-acetyltransferase variants encoded by NAT1*. Mol Pharmacol, 2000. **58**(2): p. 288-99.
34. Wakefield, L., et al., *Arylamine N-acetyltransferase 1 expression in breast cancer cell lines: a potential marker in estrogen receptor-positive tumors*. Genes Chromosomes Cancer, 2008. **47**(2): p. 118-26.
35. Chatterjee, Malay, and Khosrow Kashfi. "Human Arylamine N-acetyltransferase 1: From Drug Metabolism to Drug Target." *Cell Signaling & Molecular Targets in Cancer*. New York: Springer, 2012. 30-31.
36. Dupret, J.M., et al., *Inactivation of human arylamine N-acetyltransferase 1 by hydrogen peroxide and peroxyxynitrite*. Methods Enzymol, 2005. **400**: p. 215-29.
37. Rodrigues-Lima, F., J. Dairou, and J.M. Dupret, *Effect of environmental substances on the activity of arylamine N-acetyltransferases*. Curr Drug Metab, 2008. **9**(6): p. 505-9.
38. Buranrat, B., et al., *Inflammatory cytokines suppress arylamine N-acetyltransferase 1 in cholangiocarcinoma cells*. World J Gastroenterol, 2007. **13**(46): p. 6219-25.
39. Russell, A.J., et al., *Selective small molecule inhibitors of the potential breast cancer marker, human arylamine N-acetyltransferase 1, and its murine homologue, mouse arylamine N-acetyltransferase 2*. Bioorg Med Chem, 2009. **17**(2): p. 905-18.
40. Tiang, J.M., N.J. Butcher, and R.F. Minchin, *Small molecule inhibition of arylamine N-acetyltransferase Type I inhibits proliferation and invasiveness of MDA-MB-231 breast cancer cells*. Biochem Biophys Res Commun, 2010. **393**(1): p. 95-100.
41. Leggett, C., *Role of Human Arylamine N-Acetyltransferase in Carcinogen Metabolism and Human Breast Cancer Progression*. Diss. University of Louisville, 2012.

42. Rangunathan, N., et al., *Cadmium alters the biotransformation of carcinogenic aromatic amines by arylamine N-acetyltransferase xenobiotic-metabolizing enzymes: molecular, cellular, and in vivo studies*. Environ Health Perspect, 2010. **118**(12): p. 1685-91.
43. Sanfins, E., et al., *Carbon black nanoparticles impair acetylation of aromatic amine carcinogens through inactivation of arylamine N-acetyltransferase enzymes*. ACS Nano, 2011. **5**(6): p. 4504-11.
44. Paterson, S., et al., *Histone deacetylase inhibitors increase human arylamine N-acetyltransferase-1 expression in human tumor cells*. Drug Metab Dispos, 2011. **39**(1): p. 77-82.
45. Perou, C.M., et al., *Distinctive gene expression patterns in human mammary epithelial cells and breast cancers*. Proc Natl Acad Sci U S A, 1999. **96**(16): p. 9212-7.
46. Zhao, H., et al., *Different gene expression patterns in invasive lobular and ductal carcinomas of the breast*. Mol Biol Cell, 2004. **15**(6): p. 2523-36.
47. Wang, Y., et al., *Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer*. Lancet, 2005. **365**(9460): p. 671-9.
48. Endo, Y., et al., *miR-1290 and its potential targets are associated with characteristics of estrogen receptor alpha-positive breast cancer*. Endocr Relat Cancer, 2013. **20**(1): p. 91-102.
49. Endo, Y., et al., *Immunohistochemical determination of the miR-1290 target arylamine N-acetyltransferase 1 (NAT1) as a prognostic biomarker in breast cancer*. BMC Cancer, 2014. **14**: p. 990.
50. Bieche, I., et al., *Relationship between intratumoral expression of genes coding for xenobiotic-metabolizing enzymes and benefit from adjuvant tamoxifen in estrogen receptor alpha-positive postmenopausal breast carcinoma*. Breast Cancer Res, 2004. **6**(3): p. R252-63.
51. Tiang, J.M., et al., *RNAi-mediated knock-down of arylamine N-acetyltransferase-1 expression induces E-cadherin up-regulation and cell-cell contact growth inhibition*. PLoS One, 2011. **6**(2): p. e17031.
52. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
53. Millner, L.M., et al., *NATb/NAT1*4 promotes greater arylamine N-acetyltransferase 1 mediated DNA adducts and mutations than NATa/NAT1*4 following exposure to 4-aminobiphenyl*. Mol Carcinog, 2012. **51**(8): p. 636-46.
54. Romero, I.L., et al., *Molecular pathways: trafficking of metabolic resources in the tumor microenvironment*. Clin Cancer Res, 2015. **21**(4): p. 680-6.

ABBREVIATIONS

Acetyl-CoA	acetyl-coenzyme a
ANOVA	analysis of variance
BRCA1	breast cancer 1, early onset
BRCA2	breast cancer 2, early onset
cDNA	complementary deoxyribonucleic acid
FRT	flippase recognition target
GC×GC-TOF MS	two dimensional gas chromatography time-of-flight mass spectrometry
gRNA	guide ribonucleic acid
HPLC	high pressure liquid chromatography
mRNA	messenger ribonucleic acid
MTBSTFA	N-tert-butyltrimethylsilyl-N-methyltrifluoroacetamide
OPLS-DA	orthogonal partial least squares- discriminant analysis
NAT1	arylamine <i>N</i> -acetyltransferase 1
NAT2	arylamine <i>N</i> -acetyltransferase 2
PABA	para-aminobenzoic acid
PCA	principle component analysis
PLS-DA	partial least squares- discriminant analysis
qRT-PCR	quantitative real time polymerase chain reaction
SEM	standard error of the mean
shRNA	short hairpin ribonucleic acid
TME	tumor microenvironment

CURRICULUM VITAE

Samantha Marie Carlisle
University of Louisville School of Medicine
Department of Pharmacology and Toxicology
Louisville, KY 40202
smcarl06@louisville.edu
502-852-6284

EDUCATION

2008-2012 University of Louisville
B.S. Chemistry conc. Biochemistry
GPA 3.5

PROFESSIONAL EXPERIENCE

2012 NCI R25 Cancer Education Program
2011 James Graham Brown Cancer Center Summer Intern

HONORS

2015 NIH Travel Fellowship to Attend the 3rd Annual Workshop on
Metabolomics at UAB
2014 Best Poster Presentation by a M.S. Graduate Student, Ohio Valley
Society of Toxicology Annual Meeting
2012-2014 Graduate Research Fellowship, Integrated Programs in Biomedical
Science
2011 Kroger Scholars Scholarship
2008-2012 University of Louisville Trustee's Scholarship
2008-2012 Kentucky Educational Excellence Scholarship

LEADERSHIP

2014-present Vice President, Pharmacology-Toxicology Graduate Student
Organization
2012-present Class Representative, University of Louisville

PROFESSIONAL SOCIETIES

2013-present Society of Toxicology
2013-present American Society for Pharmacology and Experimental Therapeutics

PUBLICATIONS

Abstracts - Regional

1. Carlisle SM, Trainor PJ, Zhang X, Yin X, Doll MA, States JC, Hein DW (2015)
Investigating pathway changes associated with varying levels of human arylamine *N*-

- acetyltransferase 1 (NAT1) Activity. OVSOT Summer Student/Postdoc Meeting University of Cincinnati, Cincinnati, OH
2. Carlisle SM, Trainor PJ, Zhang X, Yin X, Doll MA, States JC, Hein DW (2014) Metabolomics of transformed MDA-MB-231 cell lines expressing different levels of human arylamine *N*-acetyltransferase 1 (NAT1). Brown Cancer Center Retreat Louisville, KY
 3. Carlisle SM, Trainor PJ, Zhang X, Yin X, Doll MA, States JC, Hein DW (2014) Metabolomics of transformed MDA-MB-231 cell lines expressing different levels of human arylamine *N*-acetyltransferase 1 (NAT1). OVSOT Dayton, OH
 4. Carlisle SM, Trainor PJ, Zhang X, Yin X, Doll MA, States JC, Hein DW (2014) Metabolomics of transformed MDA-MB-231 cell lines expressing different levels of human arylamine *N*-acetyltransferase 1 (NAT1). Research!Louisville Louisville, KY
 5. Carlisle SM, Doll MA, Stepp MW, States JC, Hein DW (2014) DDADE is an effective inhibitor of arylamine *N*-acetylation but not folate-dependent direct hydrolysis of acetyl-coenzyme A by human arylamine *N*-acetyltransferase 1 (NAT1). Great Lakes Drug Metabolism and Disposition Group Meeting Indianapolis, IN
 6. Carlisle SM, Doll MA, States JC, Hein DW (2013) The effect of a human arylamine *N*-acetyltransferase 1 specific inhibitor and curcumin or resveratrol on the proliferation of breast cancer cell lines. Research!Louisville Louisville, KY
 7. Carlisle SM, Doll MA, States JC, Hein DW (2013) The effect of a human arylamine *N*-acetyltransferase 1 specific inhibitor and curcumin or resveratrol on the proliferation of breast cancer cell lines. OVSOT Louisville, KY
 8. Carlisle SM, Doll MA, States JC, Hein DW (2013) The effect of a human arylamine *N*-acetyltransferase 1 specific inhibitor and curcumin or resveratrol on the proliferation of breast cancer cell lines. Brown Cancer Center Retreat Louisville, KY
 9. Carlisle SM, Leggett CS, Trent JO, Doll MA, States JC, Hein DW (2012) *In silico* screening for novel human arylamine *N*-acetyltransferase 1 inhibitors. Research!Louisville Louisville, KY

Abstracts - National/International

1. Carlisle SM, Doll MA, States JC, Hein DW (2013) The effect of a human arylamine *N*-acetyltransferase 1 specific inhibitor and curcumin or resveratrol on the proliferation of breast cancer cell lines. Sixth International Workshop on Arylamine *N*-Acetyltransferases Toronto, Ontario, Canada

PRESENTATIONS

1. Research Seminar, 11/13, The effect of a human arylamine *N*-acetyltransferase 1 specific inhibitor and curcumin or resveratrol on the proliferation of breast cancer cell lines. University of Louisville, Pharmacology and Toxicology Student Group Presentations, Louisville, KY.
2. Research Seminar, 03/13, Human arylamine *N*-acetyltransferase 1 inhibition by compound 10 in combination with curcumin and/or resveratrol. University of Louisville, Seminar in Pharmacology and Toxicology, Louisville, KY.