

**PHYSIOLOGICAL AND MOLECULAR OUTCOMES OF CHRONIC POST-WEANING
FOLATE DEFICIENCY**

A Dissertation

Presented to the Faculty of the Graduate School of Liberty University

In partial fulfillment of the Requirement for the Degree of
Doctor of Philosophy

by

Abigail Joan Lawton

May 2020



**PHYSIOLOGICAL AND MOLECULAR OUTCOMES OF CHRONIC POST-WEANING
FOLATE DEFICIENCY**

By Abigail Joan Lawton

A Dissertation Approved for the Department of Biology and Chemistry

Approved By

Dr. Gary D. Isaacs (Chairperson)

Dr. Gregory M. Raner (Committee Member)

Dr. Michael S. Price (Committee Member)

© Copyright by Abigail Joan Lawton 2020

All rights reserved.

ABSTRACT

Folate is an essential B vitamin that serves as a primary one-carbon carrier in a number of essential cellular reactions; among these reactions are synthesis of purines, thymidylate, and methionine. As a result of these synthesis reactions, folate plays a key role in regulating nucleic acid metabolism, amino acid metabolism, maintenance of DNA stability, and production of S-adenosylmethionine for methylation of nucleic acids, neurotransmitters, phospholipids, histones, and other proteins. As such, folate deficiency is associated with a number of adverse outcomes including various cancers, cardiovascular disease, and cognitive defects. Despite mandatory folate fortification in grains in many developed countries, a number of people groups remain at risk for folate deficiency due to various disease conditions and lifestyle choices such as smoking and alcoholism, many of which arise in early adulthood and persist for the remainder of life.

Using a mouse model of folate deficiency beginning post-weaning and persisting through the duration of the adult life, we examine the outcome of chronic folate deficiency in two key tissues: the hippocampus and the liver. The hippocampus was chosen as a tissue of focus due to its central role in cognition. The liver was chosen as the second tissue of focus because of its central role in folate storage and metabolism.

Folate deficient mice exhibited memory deficits beginning in early adulthood and persisting through late adulthood. These deficits were linked to differential expression of a number of hippocampal genes in a tissue-specific manner, and enrichment of several transcription factor binding sites were found to be associated with the differentially expressed genes. These genes may represent targets for understanding and treating diseases associated with cognitive decline, such as Alzheimer's disease.

Further, significant site-specific changes in methylation of promoter regions of hepatic genes were observed; however, no significant changes in gene expression were determined, both by global expression analysis using microarrays and by specific expression analysis of a number of tumor suppressor genes by quantitative PCR. Since promoter methylation is often inversely related to gene expression, the lack of expression changes was surprising. This may indicate that many of the differentially methylated promoter regions do not significantly impact gene expression. *Gad1*, a tumor suppressor gene associated with liver cancer, exhibited significant promoter hypermethylation in response to folate deficiency. Bisulfite sequencing of portions of the *Gad1* promoter region showed differential methylation patterns that may indicate that methylation downstream of the transcription start site of this gene has a greater impact on gene expression than methylation in upstream promoter regions. Together, these results show a relative resiliency of the liver in resisting adverse outcomes of chronic folate deficiency, such as differential expression of tumor suppressor genes, despite significant changes in methylation. The mechanisms by which the liver so propitiously resists negative outcomes of this methyl donor deficiency may serve as a template for treating this deficiency in less resilient tissues by comparing liver-specific pathways to similar pathways in other tissues.

To my parents and husband, for your unwavering support in both my studies and my life
beyond academics.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Gary D. Isaacs, for his continual encouragement and guidance. He served as an exceptional role model in research, teaching, and life outside of academics, and I am certain that I will continue to reap the benefits of lessons learned through his efforts for many years. I could not have asked for a better mentor. I would also like to thank my committee members, Dr. Greg Raner and Dr. Michael Price, for their willingness not only to advise me but also to embark on this unique journey of pioneering a new program.

I cannot imagine completing this degree without the support of my husband Kurt. He has believed in me at times when I did not believe in myself. He has helped me to harness my tenacity and stubbornness and maintain perspective. He has sacrificed so much to help me pursue my goals, even as he was in the midst of pursuing his own doctorate. I love you, and I am so thankful for you!

To my parents Ken and Carol Lenz for the many ways that they have shown their love and support throughout my life; I am so grateful. They have set examples of strong work ethics and have taught me to maintain balance in life. They have always displayed deep and genuine interest in everything I pursue, no matter how strange some of my research methods and teaching assignments may seem at times. To my brother and sister, Karl and Susanna, for always “letting” me win; your support has helped me to succeed at many things. To Grandpa Lenz, your constant prayers have impacted me greatly; I would not be the person I am today without your love and example. To my in-laws, Mike and Janet Lawton and all my new siblings and nieces and nephews, thank you for always being such a welcoming and refreshing reprieve from my academic toils.

I am grateful for each member of the faculty and staff in the Biology and Chemistry Department. So many of you have impacted me in so many ways, but the thing that stands out the most is the time you take to demonstrate your genuine care and support for me. I could not have chosen a better place to spend the past decade.

I am also grateful for my research team, Britton Upchurch and Feifan Xu, for all of your help and your company. We may have been a small team, but we accomplished a lot together. You often brightened my day and made my time in lab more enjoyable.

Finally, I would like to thank my fellow trailblazers, Crystal Passburg and Lauren Caulfield, for all of the interesting times we faced together. We are all very different and bring unique perspectives to every situation, but from the very beginning we have been in this together supporting one another through everything we face. We have pushed each other and learned from each other, and I am grateful for the growth I have experienced as a result of your friendship.

As I end the final few months of my journey as a student in such unprecedented circumstances of quarantine and physical isolation, I am reminded all the more that this journey could not have been done alone. Though I lack the space to adequately express my gratitude to the many people who have impacted me, I know that I would not have made it this far without you. Therefore I will boast all the more gladly of my weaknesses, so that the power of Christ may rest upon me.

TABLE OF CONTENTS

| | |
|---|-----|
| ABSTRACT..... | iv |
| DEDICATION..... | vi |
| ACKNOWLEDGEMENTS..... | vii |
| TABLE OF CONTENTS..... | ix |
| LIST OF FIGURES..... | xii |
| LIST OF TABLES..... | xiv |
| LIST OF ABBREVIATIONS..... | xv |
| | |
| CHAPTER 1: Overview of Folate Structure, Absorption, Function, and Deficiency | |
| Overview of folate..... | 2 |
| Structure of folates..... | 2 |
| Intestinal absorption of folates..... | 4 |
| Circulation and cellular uptake of folates..... | 6 |
| Folate-dependent one-carbon metabolism..... | 9 |
| Serine and glycine interconversion..... | 9 |
| Thymidylate synthesis..... | 12 |
| Purine synthesis..... | 13 |
| Methionine synthesis..... | 14 |
| Folate deficiency..... | 16 |
| Causes of folate deficiency..... | 16 |
| Outcomes of folate deficiency..... | 18 |
| Rationale and goals of project..... | 21 |

| | |
|-----------------|----|
| References..... | 23 |
|-----------------|----|

CHAPTER 2: Folate-dependent cognitive impairment associated with specific gene networks in the adult mouse hippocampus

| | |
|-------------------------------|----|
| Abstract..... | 34 |
| Introduction..... | 35 |
| Methods..... | 37 |
| Results..... | 43 |
| Discussion..... | 54 |
| Acknowledgements..... | 59 |
| References..... | 60 |
| Supplemental information..... | 65 |

CHAPTER 3: Chronic folate deficiency induces changes in hepatic site-specific methylation but not gene expression in the aging mouse

| | |
|-------------------------------|-----|
| Abstract..... | 73 |
| Introduction..... | 74 |
| Methods..... | 76 |
| Results..... | 90 |
| Discussion..... | 105 |
| References..... | 113 |
| Supplemental information..... | 118 |

Chapter 4: Initial characterization of Zfp410

| | |
|-------------------------------|-----|
| Abstract..... | 137 |
| Introduction..... | 137 |
| Methods..... | 138 |
| Results..... | 142 |
| Discussion..... | 150 |
| References..... | 154 |
| Supplemental information..... | 156 |

APPENDIX

| | |
|--|-----|
| Appendix A. Hippocampus microarray analysis R code..... | 157 |
| Appendix B. Boxplot for weight comparisons of mice R code..... | 161 |
| Appendix C. Volcano plots of mRNA and lncRNA R code..... | 163 |
| Appendix D. HELP assay microarray analysis R code..... | 164 |

LIST OF FIGURES

| | |
|---|-----|
| Figure 1.1 Structures of folates..... | 3 |
| Figure 1.2 Folate absorption in the small intestine..... | 5 |
| Figure 1.3 Hepatic handling of folates..... | 8 |
| Figure 1.4 Folate cycle..... | 10 |
| | |
| Figure 2.1 Study design..... | 44 |
| Figure 2.2 Novel object test for object memory..... | 46 |
| Figure 2.3 Microarray analysis of mice with and without folic acid..... | 47 |
| Figure 2.4 qPCR confirmations of specific genes shown to be differentially expressed in microarray data..... | 48 |
| Figure 2.5 Folate-dependent genes are not ubiquitously expressed..... | 51 |
| Figure 2.6 Sequence motifs enriched in folate-regulated genes..... | 55 |
| Supplemental Figure 2.1 Expression of hippocampus-regulated genes in other tissues..... | 67 |
| | |
| Figure 3.1 HELP assay overview..... | 80 |
| Figure 3.2 Bisulfite sequencing workflow..... | 86 |
| Figure 3.3 Global DNA methylation levels remain constant in folic acid deficiency..... | 92 |
| Figure 3.4 Folate deficiency induces significant methylation changes at specific loci..... | 94 |
| Figure 3.5 Relative gene locations of loci with significant methylation changes..... | 97 |
| Figure 3.6 Folate deficiency does not induce significant gene expression changes..... | 99 |
| Figure 3.7 qPCR confirms tumor suppressor genes to not be differentially expressed..... | 101 |

| | |
|---|-----|
| Figure 3.8 Folate-induced differential methylation in Gad1 promoter regions..... | 102 |
| Supplemental Figure 3.1 Verification of differential methylation patterns in Microarray probes for tumor suppressor genes..... | 130 |
| Supplemental Figure 3.2 Confirmation of bisulfite PCR..... | 133 |
| Supplemental Figure 3.3 Sequencing alignment for Gad1..... | 134 |
| | |
| Figure 4.1 Zfp410 study design for identifying binding regions..... | 140 |
| Figure 4.2 Distance between TSS and Zfp410 binding motifs..... | 146 |
| Figure 4.3 Zfp410 binding motifs are associated with cell division processes..... | 148 |
| Figure 4.4 Isolation of Zfp410 protein..... | 149 |
| Figure 4.5 Perinuclear localization of Zfp410 antibody in the human hippocampus..... | 151 |

LIST OF TABLES

| | |
|---|-----|
| Table 2.1 Gene ontology of folate-regulates genes..... | 53 |
| Supplemental Table 2.1 Teklad Envigo Custom folate, B6, B12 dietary chow..... | 65 |
| Supplemental Table 2.2 qPCR primers for gene expression..... | 66 |
| Supplemental Table 2.3 Genes whose expression was altered at both 4 and 18 mo..... | 68 |
| Supplemental Table 2.4 Genes whose expression was altered at 18 mo in both the hippocampus and the liver..... | 69 |
| Supplemental Table 2.5 Gene ontology with all genes listed..... | 70 |
| Supplemental Table 2.6 Gene ontology of genes whose expression was altered at 6 mo and 18 mo of folic acid deficiency..... | 71 |
| Supplemental Table 3.1 qPCR primers for expression analysis..... | 118 |
| Supplemental Table 3.2 Bisulfite sequencing primers..... | 119 |
| Supplemental Table 3.3 Genes associated with hypermethylated probes..... | 120 |
| Table 4.1 Genes associated with multiple Zfp410 binding sites..... | 143 |
| Table 4.2 Zfp410 regulated genes associated with differentially expressed genes in cognitive decline..... | 147 |
| Supplemental Table 4.1 Gene ontology categories for genes associated with Zfp410 binding motifs..... | 156 |

LIST OF ABBREVIATIONS

| | |
|-------------|---|
| 5,10-MTHF | 5,10-methyleneTHF |
| 5,10-mTHF | 5,10-methenylTHF |
| 5MTHF | 5-methyltetrahydrofolate |
| 10FTHF | 10-formylTHF |
| 5MTHFR | 5-methylenetetrahydrofolate reductase |
| AICAR Tfase | aminoimidazolecarboxamide ribonucleotide transformylase |
| CpG | cytosine-(phosphate)-guanine dinucleotide |
| DHF | dihydrofolate reductase |
| DHFR | dihydrofolate reductase |
| dTMP | deoxy-thymidine-5-monophosphate |
| dUMP | 2-deoxy-uridine-5-monophosphate |
| FA | folic acid |
| FDR | false discovery rate |
| FPGS | folylpoly-gamma-glutamate synthetase |
| GAR Tfase | glycinamide ribonucleotide transformylase |
| GREAT | Genomic Regions Enrichment of Annotations Tool |
| HELP | HpaII tiny fragment Enriched by Ligation-Mediated PCR |
| IPTG | isopropyl β -d-1-thiogalactopyranoside |
| LM-PCR | ligation mediated-PCR |
| LB | Luria Bertani |
| MRP | multidrug-resistance protein |
| MS | methionine synthase |
| MTHFD | methylenetetrahydrofolate dehydrogenase |
| NADH | nicotinamide adenine dinucleotide phosphate |
| p-ABA | para-aminobenzoate moiety |
| PCFT | proton-coupled folate transporter |
| PCIAA | phenol-cholorform-isoamyl alcohol |

| | |
|-------|--|
| qPCR | quantitative PCR |
| RFT | reduced folate transporter |
| SAH | S-adenosylhomocysteine |
| SAM | S-adenosylmethionine |
| SHMT | serine hydroxymethyltransferase |
| SUMO | small ubiquitin-like modifier |
| THF | tetrahydrofolate |
| TSS | transcription start site |
| TYMS | thymidylate synthase |
| X-gal | 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside |

CHAPTER 1

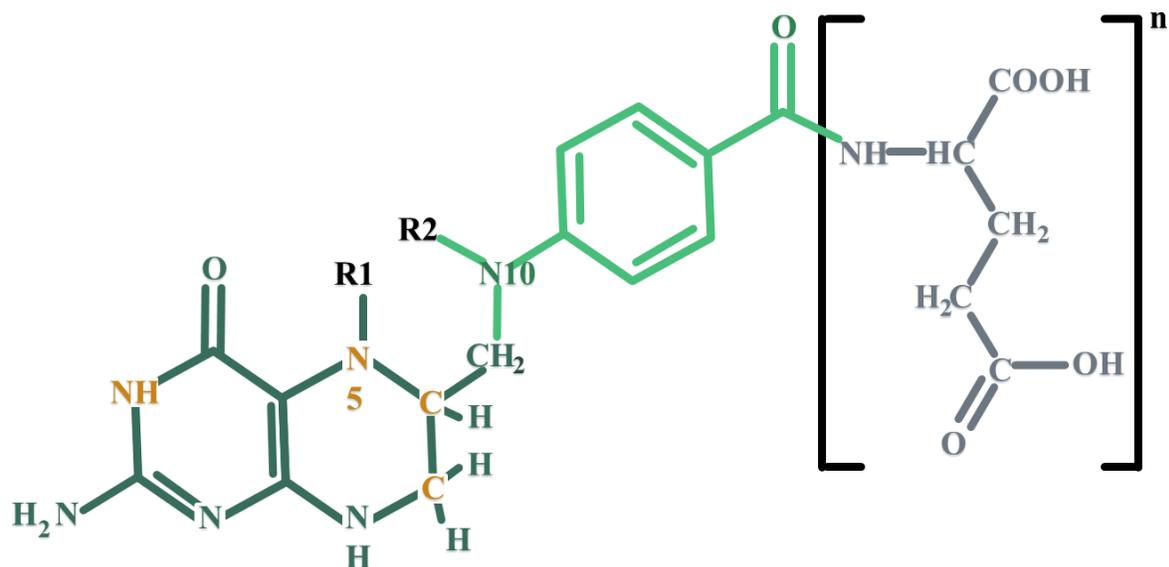
Overview of Folate Structure, Absorption, Function, and Deficiency

OVERVIEW OF FOLATE

Folate is a general term used to describe an essential water soluble B vitamin (B₉) that can be found in a variety of forms; folic acid is the primary synthetic form used in supplements and food fortification while 5-methyltetrahydrofolate (5MTHF) is the biologically active form found naturally in foods (Allen, 2008; Patanwala et al., 2014). Since mammalian cells are unable to produce folate, folate needs must be met through dietary sources such as leafy green vegetables and legumes, although some intestinal bacteria such as *Bifidobacterium* and *Lactobacillus* are able to give minor contributions to folate production (Qin & Wade, 2018; Zhao, Matherly, & Goldman, 2009). Since the late 1990s the mandatory folate fortification of grains in Western cultures represents another primary source of folate for many people.

Structure of Folates

The general structure of folates consists of a 2-amino-4-hydroxy-pteridine ring linked by a methylene group to a para-aminobenzoate moiety (p-ABA), which is linked through an amide bond to the α -amino group of a tail containing varying numbers (ranging from 1 to 8) of L-glutamates (Blancquaert, De Steur, Gellynck, & Van Der Straeten, 2014; Strobbe & Van Der Straeten, 2017). This general structure can exist in a number of different oxidation states, with tetrahydrofolate (THF) being the most reduced form and folic acid (FA) being the most oxidized form. Additionally, folates can vary in carbon units attached to the pteridine (N5) moiety and the p-ABA (N10) moiety; FA is fully oxidized and contains no varied carbon units, while all reduced folates contain at least one carbon unit (Figure 1.1). FA is an oxidized monoglutamate that is more stable and more readily absorbed in the intestine than 5MTHF; 5MTHF is a reduced



| R1 | R2 | Folate Name |
|-------|-----|-------------------|
| H | H | THF |
| CH3 | H | 5-methylTHF |
| CHO | H | 5-formylTHF |
| HC=NH | H | 5-formiminoTHF |
| H | CHO | 10-formylTHF |
| CH2 | CH2 | 5,10-methyleneTHF |
| CH | CH | 5,10-methenylTHF |

Figure 1.1. Structures of folates. Folates are composed of three distinct moieties: a pteridine ring (dark green), a para-aminobenzoate (p-ABA) moiety (light green), and a glutamate tail (grey). A variety of forms of folates exist, which can be broadly categorized as synthetic and natural folate. The primary synthetic folate is folic acid, which is fully oxidized on the pteridine ring, contains a hydrogen (H) on R2 of the para-aminobenzoate moiety, and contains a single glutamate. Although this figure depicts folate in a reduced form, the locations of folic acid oxidation are noted in orange. Natural folates are not fully oxidized and thus contain an R group on the pteridine and p-ABA moieties (R1 and R2, respectively), each of which is listed in the table along with the associated name of the folate compound based on R groups. Additionally, natural folates contain polyglutamate tails, ranging in size from 1-8.

polyglutamate that must be unconjugated to form a monoglutamate for intestinal absorption (Patanwala et al., 2014). Although 5MTHF is the most common natural form of folate, a number of other forms of natural folates occur and are characterized by the carbon groups attached to the pteridine and p-ABA moieties.

Intestinal absorption of folates

Following ingestion, various forms of folate are handled differently by the intestine. As mentioned previously, polyglutamates must be hydrolyzed to monoglutamates prior to absorption. This process is carried out by an enzyme known as glutamate carboxypeptidase II (commonly referred to as pteroylpolyglutamate hydrolase when referring specifically to enzymatic activity on polyglutamic folates), which is found in the apical membrane of jejunal brush border enterocytes (Figure 1.2) (Chandler, Wang, & Halsted, 1986). Monoglutamic folates are transported across the apical border of enterocytes primarily via either the proton-coupled folate transporter (PCFT), which has a high affinity for all folate and functions optimally at low pH, or via the reduced folate transporter (RFT), which has a low affinity for FA and functions optimally at higher pH (Visentin, Diop-Bove, Zhao, & Goldman, 2014; Zhao, Diop-Bove, Visentin, & Goldman, 2011). Although FA is in the preferred form for absorption into enterocytes, 5MTHF is the preferred form once inside enterocytes. 5MTHF is the primary form of folate used by the body, so folate enzymes within enterocytes attempt to metabolize all folates to 5MTHF before exportation (using 5-methylenetetrahydrofolate reductase [5MTHFR] for most natural folates and dihydrofolate reductase [DHF] for most synthetic folate). However, DHF availability limits the extent of this conversion for FA (Patanwala et al., 2014). Both metabolized and unmetabolized FA are transported through the basolateral membrane of enterocytes via multidrug resistance proteins (MRPs), particularly MRP3s (Zhao et al., 2009). However, MRP3

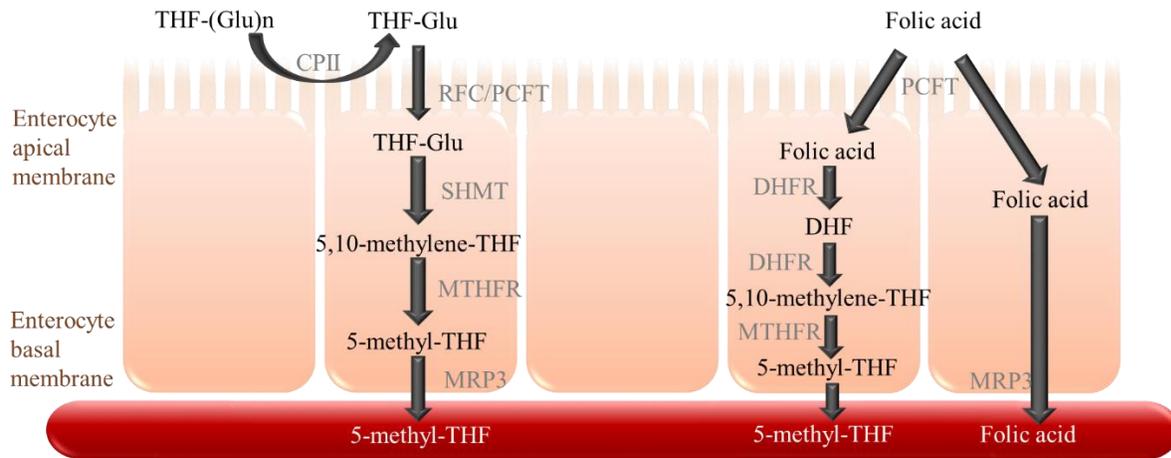


Figure 1.2. Folate absorption in the small intestine. Absorption of folates in enterocytes of the small intestine occurs by slightly differing mechanisms depending on the folate source. Natural folates, represented in the left pathway by tetrahydrofolate with a polyglutamate tail (THF-(Glu)_n), must first be hydrolyzed to monoglutamates, an action carried out by the enzyme glutamate carboxypeptidase II (CPII). These monoglutamates can then be taken up by the cell through the action of either the reduced folate carrier (RFC) or the proton-coupled folate transporter (PCFT). Once inside the enterocyte, THF-Glu is converted to 5,10-methylene-THF by the action of serine hydroxymethyltransferase (SHMT), followed by reduction to 5-methyl-THF by methylenetetrahydrofolate (MTHFR). 5-methyl-THF is transported across the basal membrane of the enterocyte and into circulation by multidrug-resistance protein 3 (MRP3). The absorption of synthetic folates follows a slightly different pathway, as represented by folic acid in the right schematic. Folic acid can be absorbed directly by the cell without need for conversion since it already exists in the monoglutamate form. However, the RFC has a very low affinity for this oxidized form of folate, making PCFT fully responsible for folic acid uptake. Like natural folates, folic acid is converted to 5-methyl-THF within the enterocyte. This process is initiated by the reduction of folic acid to dihydrofolate (DHF) then 5,10-methylene-THF through the action of dihydrofolate reductase (DHFR). 5,10-methylene-THF is then converted to 5-methyl-THF by MTHFR then transported into the bloodstream by MRP3. However, DHFR activity in enterocytes is limited, so not all folic acid can be converted to the preferred 5-methyl-THF form in enterocytes. Thus, some is released into the bloodstream in its fully oxidized form (far right). Folic acid cannot be taken up by all cell types and thus must be sent to the liver for further processing to a bioactive form.

is more efficient at transporting reduced THF than FA (Zeng, Chen, Belinsky, Rea, & Kruh, 2001). Thus, although synthetic folate is absorbed more readily into enterocytes, natural folates are more readily transported into circulation, thus allowing them to become more bioavailable following absorption.

Circulation and cellular uptake of folates

Upon transport out of enterocytes, folates enter hepatic portal circulation, where they are either taken up by hepatocytes for further processing and storage via RFT, PCFT, or receptor-mediated endocytosis of the folate receptors present on plasma membrane surfaces (Ebara, 2017). Once inside a hepatocyte, folate is either converted to a polyglutamate form for storage or to methyltetrahydrofolate, which is either secreted in the bile or hydrolyzed to a monoglutamate and delivered into systemic circulation for use by metabolically active cells (Steinberg, Campbell, & Hillman, 1982; Zhao et al., 2009).

Folate receptors are highly specific for monoglutamate folate, the form that is most commonly found in the blood (Ratanasthien, Blair, Leeming, Cooke, & Melikian, 1974). However, cells favor polyglutamate folates both for folate retention and for use by folate-dependent enzymes and their substrates; therefore once folate has entered a cell, it is typically converted to a polyglutamate form using the enzyme folylpoly-gamma-glutamate synthetase (FPGS) (Matthews, Ghose, Green, Matthews, & Dunlap, 1987; Barry Shane, 2009). While DHF and THF are excellent substrates for FPGS, FA and 5MTHF (the primary circulating form of folate) are poor substrates for FPGS (Cichowicz, Hynes, & Shane, 1988). Thus, conversion of folates to THF by methionine synthase (MS) and DHF, if necessary, is thought to be the first step in folate metabolism in the cell, followed polyglutamation by FPGS (Cook, Cichowicz, George,

Lawler, & Shane, 1987). FPGS is expressed in most tissue in two distinct isoforms, one of which targets polyglutamate folates to the mitochondria, and the other of which targets polyglutamate folates to the cytoplasm (Lawrence et al., 2014). Thus, two parallel one carbon metabolic pathways exist for folate in two different cellular compartments; these pathways are connected by small metabolites that are able to cross the mitochondrial membrane, such as serine, glycine, and formate (Zheng & Cantley, 2019). The majority of cellular folate metabolism takes place in one of these two cellular locations.

Should a cell have an excess of folate, it must be converted to a monoglutamate form for excretion, a process which is carried out by the enzyme γ -glutamyl-hydrolase in lysosomes (Yao, Rhee, & Galivan, 1995). The monoglutamate can then re-enter circulation, where it is taken up by hepatocytes. Hepatocytes have a high concentration of FPGS, making them efficient cells for folate storage (Visentin et al., 2014). However, hepatocytes also play an essential role in folate excretion. Two organic anion transporters, OATP1B1 and OATP1B3, are present on the basolateral membrane of hepatocytes surrounding the central vein, mediating folate uptake into hepatocytes from sinusoidal blood (Abe et al., 2001; Konig, Cui, Nies, & Keppler, 2000). Once inside a hepatocyte, folate can exit either on the basolateral membrane from which it entered or on the apical membrane of the cell. On the basolateral membrane, MRP3 and MRP4 transport folate into hepatic sinusoids, from which they re-enter circulation for delivery to tissues in need of folate. On the canalicular (apical) membrane of the hepatocytes, MRP2 and BCRP (breast cancer resistance protein) act to transport folate into bile for either intestinal re-absorption or for secretion in feces (Visentin et al., 2014). Thus, the liver plays a central role in storing, excreting, and mobilizing folates to maintain folate homeostasis (Figure 1.3).

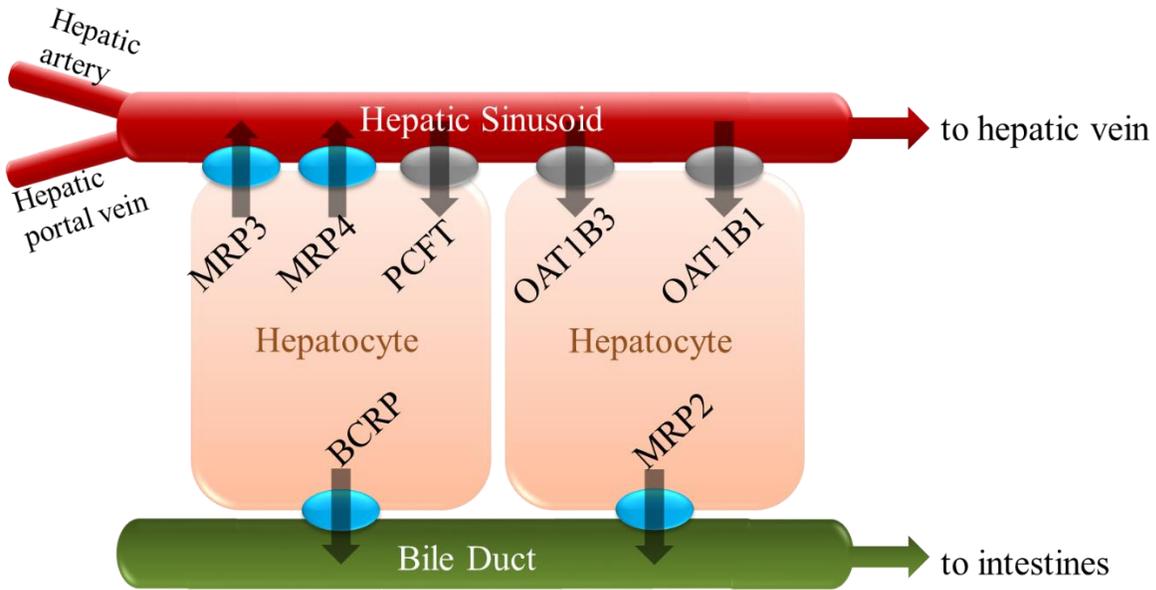


Figure 1.3. Hepatic handling of folates. Hepatocytes are able to receive folates from circulation, either through the hepatic portal vein following intestinal absorption or from the hepatic artery from systemic circulation. In both scenarios, folate will enter hepatic sinusoids, from which they can be taken up through the basolateral hepatocyte membrane by the proton-coupled folate transporter (PCFT), organic anion transporting polypeptide (OAT)1B3, or OAT1B1. Inside a hepatocyte, folate is typically polyglutamated for storage or for immediate use. If folates are converted back to the monoglutamate form in which they entered, they can be transported either out through the canicular or the basolateral membrane. Folate is transported through the canicular membrane using breast cancer resistance protein (BCRP) or multidrug resistance protein 2 (MRP2); it then enters the bile duct for transportation to the intestines where it is either re-absorbed or excreted as waste. Folate can also be transported through the basolateral membrane through MRP3 or MRP4 for transportation from hepatic sinusoids to the hepatic vein, from which it can enter systemic circulation for delivery to other tissues.

FOLATE-DEPENDENT ONE-CARBON METABOLISM

Polyglutamate folates are essential one-carbon carriers for a number of cellular reactions, contributing to DNA and RNA synthesis, methionine synthesis, and interconversion of serine and glycine (Pietrzik, Bailey, & Shane, 2010). Each of the various roles of folate are interconnected through a cycle referred to in this review as the folate cycle (Figure 1.4). As mentioned previously, folates must be converted to a polyglutamate form to be retained within a cell and used in cellular reactions; specifically, FPGS must produce a glutamate chain length of greater than three to ensure optimal folate usage and retention (B. Shane, 1989).

Although reactions have been noted in a number of subcellular compartments, the vast majority occur in the cytoplasm and mitochondria (P. J. Stover & Field, 2011). Folates become polyglutamated by FPGS primarily in one of these two cellular compartments; once they are polyglutamated they largely remain in that compartment to participate in cellular reactions. However, cells are able to address fluctuations in metabolic needs throughout various portions of the cell cycle by regulating expression and activity of folate enzymes, some of which may be capable of transporting polyglutamate folates between cellular compartments (Lan, Field, & Stover, 2018).

Serine and glycine interconversion

Serine is a non-essential amino acid that can either be acquired exogenously through diet or produced *de novo* by an intermediate metabolite in glycolysis, 3-phosphoglycerate, or by the triose phosphate pool generated from carbon intermediates from the citric acid cycle (Kalhan & Hanson, 2012; Locasale, 2013). Serine acts as a major carbon donor in folate metabolism by donating its β -carbon to THF using the enzyme serine hydroxymethyltransferase (SHMT), a

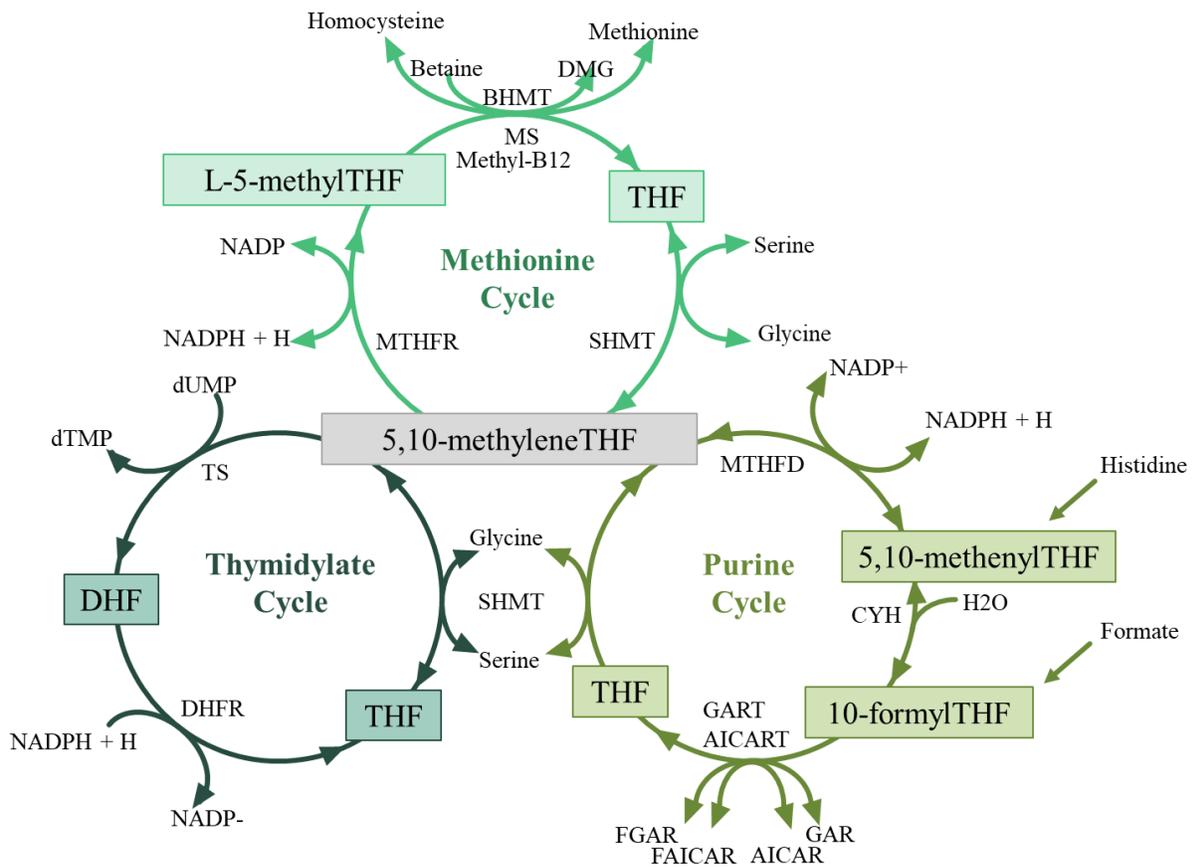


Figure 1.4. Folate cycle. The folate cycle can be divided into three different sub-cycles: the methionine cycle, the thymidylate cycle, and the purine cycle. All of these sub-cycles share one common central form of folate, 5,10-methylene-tetrahydrofolate (THF). AICAR = aminomidazole carboxamide ribinucleotide; AICART = AICAR formyltransferase; B12 = vitamin B12 (cobalamin); BHMT = betaine-homocysteine methyltransferase; CYH – methenyl-THF-cyclohydrolase; DHF = dihydrofolate; DHFR = DHF reductase; DMG = dimethylglycine; dTMP = deoxythymidine monophosphate; dUMP = deoxyuridine monophosphate; FAICAR = formyl-AICAR; FGAR = formyl-GAR; GAR = glycinamide ribonucleotide; GART = GAR formyltransferase; MS = methionine synthase; MTHFD = methylene-THF-dehydrogenase; MTHFR = 5,10-methylene-THF-reductase; NADP = nicotinamide adenine dinucleotide phosphate; NADPH – reduced NADP; SHMT = serine hydroxymethyltransferase; THF = tetrahydrofolate; TS = thymidylate synthase.

vitamin B₆-dependent enzyme (Appaji Rao, Ambili, Jala, Subramanya, & Savithri, 2003; Gregory et al., 2000). SHMT exists in both a cytosolic form (SHMT1) and a mitochondrial form (SHMT2), both of which catalyze the same reaction in different cellular compartments (Labuschagne, van den Broek, Mackay, Vousden, & Maddocks, 2014). In this reversible reaction serine is converted to glycine, and THF is converted to 5,10-methyleneTHF (5,10-MTHF). This reaction serves to provide the majority of one-carbon units for cellular use (P. Stover & Schirch, 1990). The 5,10-MTHF form of folate resulting from serine's carbon donation is a central folate which interconnects the cycles involved in folate's major functions of nucleic acid and methionine synthesis (Figure 1.4).

Not only is serine able to donate a carbon to the folate cycle, but the byproduct of this reaction, glycine, is also able to donate a carbon to the folate cycle. Like serine, glycine is a non-essential amino acid; it can be obtained from the diet or from biosynthetic pathways, including serine degradation and the glycine cleavage system. The glycine cleavage system involves a reversible reaction in which glycine is decarboxylated, forming ammonia, carbon dioxide and a carbon unit; similar to the serine pathways, this carbon unit can be added to THF to produce 5,10-MTHF (Kikuchi, 1973). Demethylation reactions involving choline, betaine, dimethylglycine and sarcosine are also able to produce glycine, thus providing a carbon unit for the folate cycle (Locasale, 2013).

While serine and glycine are both able to donate a carbon to THF forming 5,10-MTHF, this reaction is reversible; 5,10-MTHF can donate its carbon to form serine or glycine (Kikuchi, Motokawa, Yoshida, & Hiraga, 2008; Miyo et al., 2017). In turn, these amino acids can be used in synthesis reactions for a large number of proteins, or they can act as neurotransmitters

(glycine) and neuromodulators (serine) (Lopez-Corcuera, Geerlings, & Aragon, 2001; Wolosker, 2006). Thus, serine and glycine represent major carbon donors and receptors in the folate cycle.

Thymidylate synthesis

Once 5,10-MTHF has been produced by a carbon donation from serine or glycine to THF, this methylated folate form is able to act as a cofactor in a number of cellular reactions, including the formation of nucleic acids. Folate participates in pyrimidine synthesis by acting as a cofactor in reaction producing thymidylate. In thymidylate synthesis reactions, 5,10-MTHF donates its newly acquired carbon unit to 2-deoxy-uridine-5-monophosphate (dUMP), forming deoxy-thymidine-5-monophosphate (dTMP) in a reaction catalyzed by thymidylate synthase (TYMS); in the process 5,10-MTHF is oxidized to DHF (Hardy et al., 1987). This reaction represents the only *de novo* mechanism of dTMP synthesis, making it critical for DNA replication and repair (Hori, Ayusawa, Shimizu, Koyama, & Seno, 1984). DHF can be reduced through the transfer of a hydride from nicotinamide adenine dinucleotide phosphate (NADH) by the enzyme dihydrofolate reductase (DHFR) to produce THF (Schnell, Dyson, & Wright, 2004). THF is then able to receive a carbon unit from serine or glycine, reforming 5,10-MTHF (Figure 1.4). The dTMP synthesis pathway is distinct from that of other nucleotides in that it is localized to the two primary locations of DNA replication, the nucleus and the mitochondria (Anderson, Quintero, & Stover, 2011; Anderson, Woeller, & Stover, 2007; Woeller, Anderson, Szebenyi, & Stover, 2007). Since the folate cycle primarily functions in the cytosol and mitochondria, the nuclear thymidylate pathway requires small ubiquitin-like modifier (SUMO)-dependent nuclear translocation of each of TYMS, DHFR, and SHMT to occur (Anderson et al., 2007). The mechanism by which folates arrive in the nucleus remains to be elucidated, with theories

including transport bound to thymidylate pathway enzymes or non-specific sequestration following nuclear membrane breakdown with each cell division (P. J. Stover & Field, 2011). This pathway is present in the cytoplasm in the G₁ phase of the cell cycle, and upon initiation of S phase SUMO translocates the enzymes to the nucleus (MacFarlane et al., 2011). Mitochondria encode each of these enzymes and polyglutamate folate, allowing the pathway to function without the need for translocation. Failure of this pathway to function properly may result in uracil misincorporation into DNA or DNA strand breaks (Blount et al, 1997).

Purine synthesis

With the exception of the thymidylate synthesis pathway, all other *de novo* nucleotide synthesis occurs in the cytoplasm. However, the thymidylate synthesis pathway is not unique in its dependence on folate cofactors to occur; *de novo* purine synthesis also requires folates. The form of folate used in this pathways is 10-formylTHF (10FTHF), which donates carbons 2 and 8 on the purine ring using the enzymes glycinamide ribonucleotide transformylase (GAR Tfase) and aminoimidazolecarboxamide ribonucleotide transformylase (AICAR Tfase) (Figure 1.4) (P. J. Stover & Field, 2011).

As with thymidylate biosynthesis, this purine biosynthesis pathway can be traced back to the central branch point of folate metabolism, 5,10-MTHF. To begin purine synthesis, 5,10-MTHF is converted to 5,10-methenylTHF (5,10-mTHF) using the NADP⁺ dependent enzyme methylenetetrahydrofolate dehydrogenase (MTHFD) (Yeh & Greenberg, 1965). 5,10-mTHF is then reversibly converted to 10FTHF by the enzyme methenyltetrahydrofolate cyclohydrolase (Tabor & Wyngarden, 1959). Two molecules of 10FTHF are then able to donate formyl groups to phosphoribosyl pyrophosphate using GAR Tfase and AICAR Tfase, resulting in the formation

of inosine monophosphate, a precursor to adenosine monophosphate and guanosine monophosphate, and THF (Yamaoka et al., 1997). THF can then be methylated by SHMT, forming 5,10-MTHF and bringing folate back to its central form for use in purine, thymidylate, or methionine synthesis.

Methionine synthesis

In addition to acting as a key carbon donor in the nucleic acid synthesis pathways described above, 5,10-MTHF acts as a key carbon donor in methionine synthesis. To begin the methionine cycle, 5,10-MTHF is converted to L-5-MTHF in a reaction catalyzed by MTHFR, the rate-limiting enzyme of the methionine cycle (Goyette et al., 1994). The methyl group of L-5-MTHF can then be donated to homocysteine by MS, which uses vitamin B₁₂ as an intermediate and produces methionine and THF as end product (Banerjee & Matthews, 1990). SHMT is then able to methylate THF, thereby recreating the cellular pool of 5,10-MTHF to function in continued methionine synthesis or in nucleic acid synthesis.

Methionine is an essential amino acid in mammals as they are unable to synthesize either this amino acid or its homocysteine precursor. Because dietary methionine intake is very limited, methionine salvage is important to maintain adequate levels in the body (B. Shane, 2008). As such, a methionine cycle exists to salvage the homocysteine backbone to refill the methionine pool. In this cycle, methionine and ATP are converted to S-adenosylmethionine (SAM) by the enzyme methionine adenosyltransferase (Kotb & Geller, 1993). SAM donates its methyl group in a number of methylation reactions catalyzed by methyltransferases, resulting in the formation of S-adenosylhomocysteine (SAH). SAH is an inhibitor of many methyltransferases, thus enacting a feedback mechanism in the methionine cycle (B. Shane, 2008). SAH hydrolase then cleaves

SAH to homocysteine and adenosine; homocysteine can be methylated to form methionine using a carbon donation either from 5-MTHF via methionine synthase in all cells or from betaine via betaine homocysteine methyltransferase in the liver using betaine arising from choline oxidation in hepatocyte mitochondria (Murin, Vidomanova, Kowtharapu, Hatok, & Dobrota, 2017; Titus & Moran, 2000). The methionine cycle is primarily dependent on 5-MTHF rather than betaine as a carbon donor, with betaine likely representing about 30% of homocysteine methylation reactions in hepatocytes (Mudd & Poole, 1975). Thus, when folate intake is inadequate, the body suffers from a large decrease in the ability to re-methylate homocysteine (Shane, 2008).

SAM is the primary methyl donor in the vast majority of methylation reactions (Cantoni, 1953). These methylation reactions are important to a wide variety of biological molecules in mammals including DNA, RNA, proteins, and lipids (Bauerle, Schwalm, & Booker, 2015). DNA methylation epigenetically regulates gene expression, often in an inverse manner with decreased methylation resulting in increased transcription, and *vice versa* (Veland, 2017). RNA methylation acts as a post-transcriptional mechanism by which to further regulate gene expression by altering the way RNA interacts with cellular components (Mongan, Emes, & Archer, 2019). Protein methylation is able to regulate protein-protein interactions, stability, localization, and enzyme activity (Staudacher, 2012). Perhaps the most extensively studied are histones since their methylation can act as an epigenetic regulation mechanism for gene expression (Grewal & Rice, 2004). Phosphatidylcholine, a major lipid membrane component, is synthesized from choline through SAM adenylation (Locasale, 2013). Thus, the methionine cycle acts to regulate a wide variety of essential cellular processes; because of this cycle's large dependency on folate, inadequate folate intake can severely impact all of the aforementioned cellular processes.

FOLATE DEFICIENCY

In 1998 the Food and Nutrition Board of the Institute of Medicine of the United State established dietary recommendations for folate, other B vitamins, and choline; these recommendations remain in place today (Choumenkovitch et al., 2002). The recommendations are given in the form of dietary folate equivalents, which serve to equalize the different bioavailability of different folate sources. In this system 1 μg dietary folate equivalent (DFE) is equal to 1 μg of food folate, 0.6 μg of synthetic folate consumed with food, and 0.5 μg of synthetic folate consumed without food. The recommended DFE at birth is 65 μg and increases throughout childhood, reaching its maximum of 400 μg at 18 years of age; this maximum value remains constant through all of adulthood apart from an increase to 600 μg during pregnancy and 500 μg during lactation. In the same year that these recommendations were established, the Food and Drug Administration began requiring fortification of grain products with 140 μg FA/100 g. This mandatory fortification resulted in a mean folate increase of 190 $\mu\text{g}/\text{day}$ and represented a primary folate source for many Americans (Choumenkovitch et al., 2002). Since the advent of mandatory FA supplementation, most people in the United States and other higher income countries obtain adequate folate intake (Rogers et al., 2018). However, despite efforts to maintain adequate dietary levels of this essential vitamin, several people groups remain at risk for folate deficiency.

Causes of folate deficiency

A variety of conditions and circumstances may result in chronic post-natal folate deficiency. Some lifestyle choices or socioeconomic status-related conditions such as poor diet, chronic alcoholism, and chronic smoking may result in chronic folate deficiency (Allen, 2008;

Hoffbrand & Weir, 2001; Ong, Moreno, & Ross, 2011). Alcohol impairs intestinal and hepatic folate uptake, accelerates folate breakdown, and increases renal folate excretion (Bailey et al., 2015). A number of compounds in cigarette smoke including organic nitrites, nitrous oxide, cyanates, and isocyanates, have been shown to transform folates into inactive compounds (Abu Khaled, Watkins, & Krumdieck, 1986). Further cigarette smoking results in oxidative stress, perhaps altering a cell's ability to metabolize and store folate (Relton, Pearce, & Parker, 2005). In addition to the negative effects of smoking itself, people who smoke tend to exhibit significantly decreased folate intake compared to non-smokers, thus compounding the negative effects of smoking (Ortega et al., 2004).

A number of diseases may also contribute to inadequate folate status including Crohn's disease, celiac disease, other chronic inflammatory bowel diseases (IBD), obesity, liver disease, and gene polymorphisms on folate cycle enzymes (Bermejo et al., 2013; Davis et al., 2005; Devlin et al., 2000; Dickey et al., 2008; Elsborg & Larsen, 1979; Gloria et al., 1997; Halsted, Villanueva, & Devlin, 2002; Jansen et al., 2004; Kaidar-Person, Person, Szomstein, & Rosenthal, 2008; Kremer, Galivan, Streckfuss, & Kamen, 1986; Longstreth & Green, 1983; Mojtabai, 2004; Molloy et al., 1997; Morrell, 2002; Salojin et al., 2011; Wierdsma, van Bokhorst-de van der Schueren, Berkenpas, Mulder, & van Bodegraven, 2013). IBDs result in decreased absorption of a number of nutrients through morphological changes in intestinal mucosa in both inflamed and non-inflamed regions as well as changes in activity of luminal enzymes; additionally, nutrient intake may be decreased with IBD due to dietary restrictions (Pan et al., 2017; Peuhkuri, Vapaatalo, & Korpela, 2010). Similarly, obesity leads to chronic inflammation and oxidative stress, which both alters intestinal absorption and alters cellular metabolism and storage of micronutrients (Ellulu, Patimah, Khaza'ai, Rahmat, & Abed, 2017). To make matters worse,

some drugs that are used to treat the aforementioned diseases also result in decreased folate. For example, methotrexate, which is used to treat Crohn's disease in addition to other autoimmune diseases and cancers, targets DHR and is thus designed to be an antifolate (Kremer et al., 1986; Zeng et al., 2001). Similarly, sulfasalazine, which is used to treat Crohn's disease and other inflammatory bowel diseases as well as arthritis, decreases folate levels by potentially inhibiting RFT (Jansen et al., 2004). Additionally, many drugs that treat epilepsy, such as phenytoin, carbamazepine, and valproate negatively impact folate levels (Morrell, 2002).

These individuals who are at risk of folate deficiency due to lifestyle, disease, and medications represent a significant portion of the American population. According to the 2018 National Survey on Drug Use and Health by the NIH, 14.4 million Americans had alcohol use disorders ("National Survey on Drug Use and Health (NSDUH)," 2018). According to the CDC, close to 34 million Americans smoke (Creamer et al., 2019). Roughly 3 million Americans suffer from IBDs (Dahlhamer, Zammitti, Ward, Wheaton, & Croft, 2016). Most astoundingly, 35% of men (about 56 million) and 40% of women (about 66 million) in America are obese (Flegal, Kruszon-Moran, Carroll, Fryar, & Ogden, 2016). Thus, a large portion of the United States is at risk for folate deficiency.

Outcomes of folate deficiency

Perhaps the most influential discovery in the history of folate was the association between neural tube defects and folate deficiency as early as 1965 (Hibbard, Hibbard, & Jeffcoate, 1965). Over the course of the next few decades numerous studies further developed this connection, and in the early 1990s, the United States began officially recommending FA supplementation to all women of child-bearing age (Honein, Paulozzi, Mathews, Erickson, &

Wong, 2001). Since the establishment of the direct link between folate deficiency and neural tube defects, folate research has expanded beyond fetal development to post-natal populations. Post-natal folate deficiency is associated with a number of unfavorable outcomes including cardiovascular disease, various cancers, and cognitive defects.

One of the earliest noted outcomes of post-natal folate deficiency is the presence of megaloblastic anemia due to limited cell division from decreased DNA synthesis and repair (Green & Miller, 1999). Since this outcome presents quickly and with high prevalence, it has historically been the first indicator of folate deficiency (Snow, 1999). Another major cardiovascular risk factor associated with folate deficiency is increased risk of stroke, which is thought to be caused by increased homocysteine since its conversion to methionine is highly inefficient without folate; elevated homocysteine results in accumulation of reactive oxygen species and posttranslational protein modifications, often affecting enzymes involved in redox balance (Casas, Bautista, Smeeth, Sharma, & Hingorani, 2005; Lehotsky et al., 2016). To a slightly lesser extent, folate deficiency also results in increased risk of coronary artery disease (Klerk et al., 2002). The mechanism linking coronary artery disease and folate deficiency may not be a direct link, as indicated by the inability of folate supplementation to reduce coronary artery disease (Huang et al., 2012). However, this same supplementation can effectively prevent strokes, indicating a more direct association between strokes and folate deficiency.

The role of folate deficiency in cancer appears to be similarly murky; some trials report folate deficiency to increase cancer risk while other report folate supplementation to increase cancer risk (Kim, 2018). It is likely that these discrepancies arise from the timing of folate deficiency or supplementation. Since folate is crucial to DNA synthesis and thus cell divisions, folate administration to individuals who already have cancer or pre-cancerous growths can aid

cancer growth by supplying nutrients for cancer cell division (Mason, 2011). Conversely, folate deficiency may aid in cancer development due to aberrant DNA synthesis, repair, and methylation patterns. Indeed, folate status is inversely related to a number of cancers including colorectal, lung, pancreatic, esophageal, stomach, cervical, ovarian, breast, bladder (Bailey et al., 2015; Collin et al., 2010; Kim, 1999; Xu & Chen, 2009).

A somewhat more definitive link between folate and disease is the association between low folate levels and increased risk of cognitive impairment and dementia (Hooshmand et al., 2012; Ravaglia et al., 2005; Smith & Refsum, 2016). However, cognitive impairments induced by folate deficiency appear to be less quickly established and less readily reversed than megaloblastic anemia induced by folate deficiency, likely because red blood cells are constantly being replaced while neurons are more stagnant in their establishment. One study found that folate supplementation in mild-to-moderate Alzheimer's disease did not slow the rate of cognitive decline (Aisen et al., 2008). However, another study that included baseline dietary folate intake found that individuals with low dietary folate exhibited a decreased rate of cognitive decline when given folate supplements, while those with normal dietary folate received no cognitive benefit from folate supplementation (Kang et al., 2008). Other studies corroborate these findings: folate supplementation does not alter cognitive decline in individuals with normal folate levels, but supplementation to those with low folate increases cognitive function (Durga et al., 2007; Smith & Refsum, 2016; Smith et al., 2018). Thus, folate is not a miracle treatment to reverse all cognitive decline, but folate deficiency may cause some cognitive impairments in a partially reversible manner.

RATIONALE AND GOALS OF PROJECT

Post-natal folate deficiency is known to be associated with a number of poor outcomes including increased risk of cardiovascular disease, various cancers, and decline in cognitive function. However, the vast majority of studies on this topic focus on short term deficiency or deficiency in early life, providing little understanding of the impacts of chronic deficiency throughout adult life (Berrocal-Zaragoza et al., 2014; Miller & Sweatt, 2007; Jadavji, Deng, Malysheva, Caudill, & Rozen, 2015). Literature for long-term folate deficiency rarely studies deficiency beyond 12 weeks in rodents and at maximum extends to 20-30 weeks, representing about half of the rodent's adult lifespan (Crott, Choi, Ordovas, Ditelberg, & Mason, 2004; Esfandiari, Villanueva, Wong, French, & Halsted, 2005; McMartin et al., 1989; Riggs, Spiro, Tucker, & Rush, 1996; Varela-Moreiras & Selhub, 1992). Thus, although many habits and diseases may lead to chronic folate deficiency for the duration of adult life, little is known, regarding the effects of this deficiency in the later stages of adulthood. Further, most studies on folate deficiency induce the deficiency early in life, either *in utero* or during weaning (Berrocal-Zaragoza et al., 2014; Mejos, Kim, Lim, & Chang, 2013). Although valuable, these studies provide no understanding of the many populations who receive adequate folate early in life but become deficient during adulthood (such as alcoholics, chronic smokers, individuals with bowel diseases that present during teenage or early adult years, individuals taking cancer and arthritis medications, and dietary choices that a person may not have freedom to make until they are no longer dependent on parents for food). As such, it has yet to be determined to what extent adequate folate intake during early life can protect against folate deficiency later in life, despite the fact that this scenario likely affects many people.

This project seeks to determine some of the late life outcomes of chronic post-weaning folate deficiency in two of the tissues which are strongly linked to folate. Late life is defined as 18 mo in this study, which correlates to a time when senescent biomarkers are widely present in mice; a later time point was not used since survivorship of mice becomes increasingly problematic with age (Flurkey K, Curren JM, Harrison DE. 2007). First, we hypothesize that chronic post-weaning folate deficiency will result in cognitive deficits, which can be linked to tissue-specific changes in gene expression (Chapter II). Second, we hypothesize that chronic post-weaning folate deficiency will result in altered methylation patterns in hepatic DNA, perhaps shedding light on mechanisms by which folate deficiency in this major storage organ may lead to downstream changes in other body tissues (Chapter III).

REFERENCES

- Abe, T., Unno, M., Onogawa, T., Tokui, T., Kondo, T. N., Nakagomi, R., . . . Matsuno, S. (2001). LST-2, a human liver-specific organic anion transporter, determines methotrexate sensitivity in gastrointestinal cancers. *Gastroenterology*, *120*(7), 1689-1699. doi: 10.1053/gast.2001.24804
- Abu Khaled, M., Watkins, C. L., & Krumdieck, C. L. (1986). Inactivation of B12 and folate coenzymes by butyl nitrite as observed by NMR: implications on one-carbon transfer mechanism. *Biochem Biophys Res Commun*, *135*(1), 201-207. doi: 10.1016/0006-291x(86)90963-0
- Aisen, P. S., Schneider, L. S., Sano, M., Diaz-Arrastia, R., van Dyck, C. H., Weiner, M. F., . . . Alzheimer Disease Cooperative, Study. (2008). High-dose B vitamin supplementation and cognitive decline in Alzheimer disease: a randomized controlled trial. *JAMA*, *300*(15), 1774-1783. doi: 10.1001/jama.300.15.1774
- Allen, L. H. (2008). Causes of vitamin B12 and folate deficiency. *Food Nutr Bull*, *29*(2 Suppl), S20-34; discussion S35-27. doi: 10.1177/15648265080292S105
- Anderson, D. D., Quintero, C. M., & Stover, P. J. (2011). Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *Proc Natl Acad Sci U S A*, *108*(37), 15163-15168. doi: 10.1073/pnas.1103623108
- Anderson, D. D., Woeller, C. F., & Stover, P. J. (2007). Small ubiquitin-like modifier-1 (SUMO-1) modification of thymidylate synthase and dihydrofolate reductase. *Clin Chem Lab Med*, *45*(12), 1760-1763. doi: 10.1515/CCLM.2007.355
- Appaji Rao, N., Ambili, M., Jala, V. R., Subramanya, H. S., & Savithri, H. S. (2003). Structure-function relationship in serine hydroxymethyltransferase. *Biochim Biophys Acta*, *1647*(1-2), 24-29. doi: 10.1016/s1570-9639(03)00043-8
- Bailey, L. B., Stover, P. J., McNulty, H., Fenech, M. F., Gregory, J. F., 3rd, Mills, J. L., . . . Raiten, D. J. (2015). Biomarkers of Nutrition for Development-Folate Review. *J Nutr*, *145*(7), 1636S-1680S. doi: 10.3945/jn.114.206599
- Banerjee, R. V., & Matthews, R. G. (1990). Cobalamin-dependent methionine synthase. *FASEB J*, *4*(5), 1450-1459. doi: 10.1096/fasebj.4.5.2407589
- Bauerle, M. R., Schwalm, E. L., & Booker, S. J. (2015). Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J Biol Chem*, *290*(7), 3995-4002. doi: 10.1074/jbc.R114.607044
- Bermejo, F., Algaba, A., Guerra, I., Chaparro, M., De-La-Poza, G., Valer, P., . . . Gisbert, J. P. (2013). Should we monitor vitamin B12 and folate levels in Crohn's disease patients? *Scand J Gastroenterol*, *48*(11), 1272-1277. doi: 10.3109/00365521.2013.836752
- Berrocal-Zaragoza, M. I., Sequeira, J. M., Murphy, M. M., Fernandez-Ballart, J. D., Abdel Baki, S. G., Bergold, P. J., & Quadros, E. V. (2014). Folate deficiency in rat pups during weaning causes learning and memory deficits. *Br J Nutr*, *112*(8), 1323-1332. doi: 10.1017/S0007114514002116

- Blancquaert, D., De Steur, H., Gellynck, X., & Van Der Straeten, D. (2014). Present and future of folate biofortification of crop plants. *J Exp Bot*, *65*(4), 895-906. doi: 10.1093/jxb/ert483
- Blount, B. C., Mack, M. M., Wehr, C. M., MacGregor, J. T., Hiatt, R. A., Wang, G., . . . Ames, B. N. (1997). Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *Proc Natl Acad Sci U S A*, *94*(7), 3290-3295. doi: 10.1073/pnas.94.7.3290
- Cantoni, G. L. (1953). S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate. *J Biol Chem*, *204*(1), 403-416.
- Casas, J. P., Bautista, L. E., Smeeth, L., Sharma, P., & Hingorani, A. D. (2005). Homocysteine and stroke: evidence on a causal link from mendelian randomisation. *Lancet*, *365*(9455), 224-232. doi: 10.1016/S0140-6736(05)17742-3
- Chandler, C. J., Wang, T. T., & Halsted, C. H. (1986). Pteroylpolyglutamate hydrolase from human jejunal brush borders. Purification and characterization. *J Biol Chem*, *261*(2), 928-933.
- Choumenkovitch, S. F., Selhub, J., Wilson, P. W., Rader, J. I., Rosenberg, I. H., & Jacques, P. F. (2002). Folic acid intake from fortification in United States exceeds predictions. *J Nutr*, *132*(9), 2792-2798. doi: 10.1093/jn/132.9.2792
- Cichowicz, D. J., Hynes, J. B., & Shane, B. (1988). Substrate specificity of mammalian folylpoly-gamma-glutamate synthetase for 5,8-dideazafolates and 5,8-dideaza analogues of aminopterin. *Biochim Biophys Acta*, *957*(3), 363-369. doi: 10.1016/0167-4838(88)90227-0
- Collin, S. M., Metcalfe, C., Refsum, H., Lewis, S. J., Zuccolo, L., Smith, G. D., . . . Martin, R. M. (2010). Circulating folate, vitamin B12, homocysteine, vitamin B12 transport proteins, and risk of prostate cancer: a case-control study, systematic review, and meta-analysis. *Cancer Epidemiol Biomarkers Prev*, *19*(6), 1632-1642. doi: 10.1158/1055-9965.EPI-10-0180
- Cook, J. D., Cichowicz, D. J., George, S., Lawler, A., & Shane, B. (1987). Mammalian folylpoly-gamma-glutamate synthetase. 4. In vitro and in vivo metabolism of folates and analogues and regulation of folate homeostasis. *Biochemistry*, *26*(2), 530-539. doi: 10.1021/bi00376a027
- Creamer, M. R., Wang, T. W., Babb, S., Cullen, K. A., Day, H., Willis, G., . . . Neff, L. (2019). Tobacco Product Use and Cessation Indicators Among Adults - United States, 2018. *MMWR Morb Mortal Wkly Rep*, *68*(45), 1013-1019. doi: 10.15585/mmwr.mm6845a2
- Crott, J. W., Choi, S. W., Ordovas, J. M., Ditelberg, J. S., & Mason, J. B. (2004). Effects of dietary folate and aging on gene expression in the colonic mucosa of rats: implications for carcinogenesis. *Carcinogenesis*, *25*(1), 69-76. doi: 10.1093/carcin/bgg150
- Dahlhamer, J. M., Zammitti, E. P., Ward, B. W., Wheaton, A. G., & Croft, J. B. (2016). Prevalence of Inflammatory Bowel Disease Among Adults Aged ≥ 18 Years - United

- States, 2015. *MMWR Morb Mortal Wkly Rep*, 65(42), 1166-1169. doi: 10.15585/mmwr.mm6542a3
- Davis, S. R., Quinlivan, E. P., Shelnutt, K. P., Maneval, D. R., Ghandour, H., Capdevila, A., . . . Gregory, J. F., 3rd. (2005). The methylenetetrahydrofolate reductase 677C->T polymorphism and dietary folate restriction affect plasma one-carbon metabolites and red blood cell folate concentrations and distribution in women. *J Nutr*, 135(5), 1040-1044. doi: 10.1093/jn/135.5.1040
- Devlin, A. M., Ling, E. H., Peerson, J. M., Fernando, S., Clarke, R., Smith, A. D., & Halsted, C. H. (2000). Glutamate carboxypeptidase II: a polymorphism associated with lower levels of serum folate and hyperhomocysteinemia. *Hum Mol Genet*, 9(19), 2837-2844.
- Dickey, W., Ward, M., Whittle, C. R., Kelly, M. T., Pentieva, K., Horigan, G., . . . McNulty, H. (2008). Homocysteine and related B-vitamin status in coeliac disease: Effects of gluten exclusion and histological recovery. *Scand J Gastroenterol*, 43(6), 682-688. doi: 10.1080/00365520701881118
- Durga, J., van Boxtel, M. P., Schouten, E. G., Kok, F. J., Jolles, J., Katan, M. B., & Verhoef, P. (2007). Effect of 3-year folic acid supplementation on cognitive function in older adults in the FACIT trial: a randomised, double blind, controlled trial. *Lancet*, 369(9557), 208-216. doi: 10.1016/S0140-6736(07)60109-3
- Ebara, S. (2017). Nutritional role of folate. *Congenit Anom (Kyoto)*, 57(5), 138-141. doi: 10.1111/cga.12233
- Ellulu, M. S., Patimah, I., Khaza'ai, H., Rahmat, A., & Abed, Y. (2017). Obesity and inflammation: the linking mechanism and the complications. *Arch Med Sci*, 13(4), 851-863. doi: 10.5114/aoms.2016.58928
- Elsborg, L., & Larsen, L. (1979). Folate deficiency in chronic inflammatory bowel diseases. *Scand J Gastroenterol*, 14(8), 1019-1024.
- Esfandiari, F., Villanueva, J. A., Wong, D. H., French, S. W., & Halsted, C. H. (2005). Chronic ethanol feeding and folate deficiency activate hepatic endoplasmic reticulum stress pathway in micropigs. *Am J Physiol Gastrointest Liver Physiol*, 289(1), G54-63. doi: 10.1152/ajpgi.00542.2004
- Flegal, K. M., Kruszon-Moran, D., Carroll, M. D., Fryar, C. D., & Ogden, C. L. (2016). Trends in Obesity Among Adults in the United States, 2005 to 2014. *JAMA*, 315(21), 2284-2291. doi: 10.1001/jama.2016.6458
- Flurkey K, Currer JM, Harrison DE. 2007. The Mouse in Aging Research. In *The Mouse in Biomedical Research 2nd Edition*. Fox JG, et al, editors. American College Laboratory Animal Medicine (Elsevier), Burlington, MA. pp. 637-672.
- Gloria, L., Cravo, M., Camilo, M. E., Resende, M., Cardoso, J. N., Oliveira, A. G., . . . Mira, F. C. (1997). Nutritional deficiencies in chronic alcoholics: relation to dietary intake and alcohol consumption. *Am J Gastroenterol*, 92(3), 485-489.

- Goyette, P., Sumner, J. S., Milos, R., Duncan, A. M., Rosenblatt, D. S., Matthews, R. G., & Rozen, R. (1994). Human methylenetetrahydrofolate reductase: isolation of cDNA mapping and mutation identification. *Nat Genet*, *7*(4), 551.
- Green, R., & Miller, J. W. (1999). Folate deficiency beyond megaloblastic anemia: hyperhomocysteinemia and other manifestations of dysfunctional folate status. *Semin Hematol*, *36*(1), 47-64.
- Gregory, J. F., 3rd, Cuskelly, G. J., Shane, B., Toth, J. P., Baumgartner, T. G., & Stacpoole, P. W. (2000). Primed, constant infusion with [2H3]serine allows in vivo kinetic measurement of serine turnover, homocysteine remethylation, and transsulfuration processes in human one-carbon metabolism. *Am J Clin Nutr*, *72*(6), 1535-1541. doi: 10.1093/ajcn/72.6.1535
- Grewal, S. I., & Rice, J. C. (2004). Regulation of heterochromatin by histone methylation and small RNAs. *Curr Opin Cell Biol*, *16*(3), 230-238. doi: 10.1016/j.ceb.2004.04.002
- Halsted, C. H., Villanueva, J. A., & Devlin, A. M. (2002). Folate deficiency, methionine metabolism, and alcoholic liver disease. *Alcohol*, *27*(3), 169-172.
- Hardy, L. W., Finer-Moore, J. S., Montfort, W. R., Jones, M. O., Santi, D. V., & Stroud, R. M. (1987). Atomic structure of thymidylate synthase: target for rational drug design. *Science*, *235*(4787), 448-455. doi: 10.1126/science.3099389
- Hibbard, B. M., Hibbard, E. D., & Jeffcoate, T. N. (1965). Folic acid and reproduction. *Acta Obstet Gynecol Scand*, *44*(3), 375-400. doi: 10.3109/00016346509155874
- Hoffbrand, A. V., & Weir, D. G. (2001). The history of folic acid. *Br J Haematol*, *113*(3), 579-589.
- Honein, M. A., Paulozzi, L. J., Mathews, T. J., Erickson, J. D., & Wong, L. Y. (2001). Impact of folic acid fortification of the US food supply on the occurrence of neural tube defects. *JAMA*, *285*(23), 2981-2986. doi: 10.1001/jama.285.23.2981
- Hooshmand, B., Solomon, A., Kareholt, I., Rusanen, M., Hanninen, T., Leiviska, J., . . . Kivipelto, M. (2012). Associations between serum homocysteine, holotranscobalamin, folate and cognition in the elderly: a longitudinal study. *J Intern Med*, *271*(2), 204-212. doi: 10.1111/j.1365-2796.2011.02484.x
- Hori, T., Ayusawa, D., Shimizu, K., Koyama, H., & Seno, T. (1984). Chromosome breakage induced by thymidylate stress in thymidylate synthase-negative mutants of mouse FM3A cells. *Cancer Res*, *44*(2), 703-709.
- Huang, T., Chen, Y., Yang, B., Yang, J., Wahlqvist, M. L., & Li, D. (2012). Meta-analysis of B vitamin supplementation on plasma homocysteine, cardiovascular and all-cause mortality. *Clin Nutr*, *31*(4), 448-454. doi: 10.1016/j.clnu.2011.01.003
- Jansen, G., van der Heijden, J., Oerlemans, R., Lems, W. F., Ifergan, I., Scheper, R. J., . . . Dijkmans, B. A. (2004). Sulfasalazine is a potent inhibitor of the reduced folate carrier: implications for combination therapies with methotrexate in rheumatoid arthritis. *Arthritis Rheum*, *50*(7), 2130-2139. doi: 10.1002/art.20375

- Kaidar-Person, O., Person, B., Szomstein, S., & Rosenthal, R. J. (2008). Nutritional deficiencies in morbidly obese patients: a new form of malnutrition? Part A: vitamins. *Obes Surg*, *18*(7), 870-876. doi: 10.1007/s11695-007-9349-y
- Kalhan, S. C., & Hanson, R. W. (2012). Resurgence of serine: an often neglected but indispensable amino Acid. *J Biol Chem*, *287*(24), 19786-19791. doi: 10.1074/jbc.R112.357194
- Kang, J. H., Cook, N., Manson, J., Buring, J. E., Albert, C. M., & Grodstein, F. (2008). A trial of B vitamins and cognitive function among women at high risk of cardiovascular disease. *Am J Clin Nutr*, *88*(6), 1602-1610. doi: 10.3945/ajcn.2008.26404
- Kikuchi, G. (1973). The glycine cleavage system: composition, reaction mechanism, and physiological significance. *Mol Cell Biochem*, *1*(2), 169-187. doi: 10.1007/bf01659328
- Kikuchi, G., Motokawa, Y., Yoshida, T., & Hiraga, K. (2008). Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia. *Proc Jpn Acad Ser B Phys Biol Sci*, *84*(7), 246-263. doi: 10.2183/pjab.84.246
- Kim, Y. I. (1999). Folate and carcinogenesis: evidence, mechanisms, and implications. *J Nutr Biochem*, *10*(2), 66-88. doi: 10.1016/s0955-2863(98)00074-6
- Kim, Y. I. (2018). Folate and cancer: a tale of Dr. Jekyll and Mr. Hyde? *Am J Clin Nutr*, *107*(2), 139-142. doi: 10.1093/ajcn/nqx076
- Klerk, M., Verhoef, P., Clarke, R., Blom, H. J., Kok, F. J., Schouten, E. G., & Group, Mthfr Studies Collaboration. (2002). MTHFR 677C-->T polymorphism and risk of coronary heart disease: a meta-analysis. *JAMA*, *288*(16), 2023-2031. doi: 10.1001/jama.288.16.2023
- Konig, J., Cui, Y., Nies, A. T., & Keppler, D. (2000). A novel human organic anion transporting polypeptide localized to the basolateral hepatocyte membrane. *Am J Physiol Gastrointest Liver Physiol*, *278*(1), G156-164. doi: 10.1152/ajpgi.2000.278.1.G156
- Kotb, M., & Geller, A. M. (1993). Methionine adenosyltransferase: structure and function. *Pharmacol Ther*, *59*(2), 125-143. doi: 10.1016/0163-7258(93)90042-c
- Kremer, J. M., Galivan, J., Streckfuss, A., & Kamen, B. (1986). Methotrexate metabolism analysis in blood and liver of rheumatoid arthritis patients. Association with hepatic folate deficiency and formation of polyglutamates. *Arthritis Rheum*, *29*(7), 832-835.
- Labuschagne, C. F., van den Broek, N. J., Mackay, G. M., Vousden, K. H., & Maddocks, O. D. (2014). Serine, but not glycine, supports one-carbon metabolism and proliferation of cancer cells. *Cell Rep*, *7*(4), 1248-1258. doi: 10.1016/j.celrep.2014.04.045
- Lan, X., Field, M. S., & Stover, P. J. (2018). Cell cycle regulation of folate-mediated one-carbon metabolism. *Wiley Interdiscip Rev Syst Biol Med*, *10*(6), e1426. doi: 10.1002/wsbm.1426
- Lawrence, S. A., Titus, S. A., Ferguson, J., Heineman, A. L., Taylor, S. M., & Moran, R. G. (2014). Mammalian mitochondrial and cytosolic folylpolyglutamate synthetase maintain the subcellular compartmentalization of folates. *J Biol Chem*, *289*(42), 29386-29396. doi: 10.1074/jbc.M114.593244

- Lehotsky, J., Tothova, B., Kovalska, M., Dobrota, D., Benova, A., Kalenska, D., & Kaplan, P. (2016). Role of Homocysteine in the Ischemic Stroke and Development of Ischemic Tolerance. *Front Neurosci*, *10*, 538. doi: 10.3389/fnins.2016.00538
- Locasale, J. W. (2013). Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nat Rev Cancer*, *13*(8), 572-583. doi: 10.1038/nrc3557
- Longstreth, G. F., & Green, R. (1983). Folate status in patients receiving maintenance doses of sulfasalazine. *Arch Intern Med*, *143*(5), 902-904.
- Lopez-Corcuera, B., Geerlings, A., & Aragon, C. (2001). Glycine neurotransmitter transporters: an update. *Mol Membr Biol*, *18*(1), 13-20.
- MacFarlane, A. J., Anderson, D. D., Flodby, P., Perry, C. A., Allen, R. H., Stabler, S. P., & Stover, P. J. (2011). Nuclear localization of de novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA. *J Biol Chem*, *286*(51), 44015-44022. doi: 10.1074/jbc.M111.307629
- Mason, J. B. (2011). Unraveling the complex relationship between folate and cancer risk. *Biofactors*, *37*(4), 253-260. doi: 10.1002/biof.174
- Matthews, R. G., Ghose, C., Green, J. M., Matthews, K. D., & Dunlap, R. B. (1987). Folylpolylglutamates as substrates and inhibitors of folate-dependent enzymes. *Adv Enzyme Regul*, *26*, 157-171. doi: 10.1016/0065-2571(87)90012-4
- McMartin, K. E., Collins, T. D., Eisenga, B. H., Fortney, T., Bates, W. R., & Bairnsfather, L. (1989). Effects of chronic ethanol and diet treatment on urinary folate excretion and development of folate deficiency in the rat. *J Nutr*, *119*(10), 1490-1497. doi: 10.1093/jn/119.10.1490
- Mejos, K. K., Kim, H. W., Lim, E. M., & Chang, N. (2013). Effects of parental folate deficiency on the folate content, global DNA methylation, and expressions of FRalpha, IGF-2 and IGF-1R in the postnatal rat liver. *Nutr Res Pract*, *7*(4), 281-286. doi: 10.4162/nrp.2013.7.4.281
- Miyo, M., Konno, M., Colvin, H., Nishida, N., Koseki, J., Kawamoto, K., . . . Ishii, H. (2017). The importance of mitochondrial folate enzymes in human colorectal cancer. *Oncol Rep*, *37*(1), 417-425. doi: 10.3892/or.2016.5264
- Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron*, *53*(6), 857-869. doi: 10.1016/j.neuron.2007.02.022
- Mojtabai, R. (2004). Body mass index and serum folate in childbearing age women. *Eur J Epidemiol*, *19*(11), 1029-1036.
- Molloy, A. M., Daly, S., Mills, J. L., Kirke, P. N., Whitehead, A. S., Ramsbottom, D., . . . Scott, J. M. (1997). Thermolabile variant of 5,10-methylenetetrahydrofolate reductase associated with low red-cell folates: implications for folate intake recommendations. *Lancet*, *349*(9065), 1591-1593. doi: 10.1016/S0140-6736(96)12049-3
- Mongan, N. P., Emes, R. D., & Archer, N. (2019). Detection and analysis of RNA methylation. *F1000Res*, *8*. doi: 10.12688/f1000research.17956.1

- Morrell, M. J. (2002). Folic Acid and Epilepsy. *Epilepsy Curr*, 2(2), 31-34. doi: 10.1046/j.1535-7597.2002.00017.x
- Mudd, S. H., & Poole, J. R. (1975). Labile methyl balances for normal humans on various dietary regimens. *Metabolism*, 24(6), 721-735. doi: 10.1016/0026-0495(75)90040-2
- Murin, R., Vidomanova, E., Kowtharapu, B. S., Hatok, J., & Dobrota, D. (2017). Role of S-adenosylmethionine cycle in carcinogenesis. *Gen Physiol Biophys*, 36(5), 513-520. doi: 10.4149/gpb_2017031
- . National Survey on Drug Use and Health (NSDUH). (2018): Substance Abuse and Mental Health Services Administration.
- Nicolas Veland, Taiping Chen. (2017). Mechanisms of DNA Methylation and Demethylation During Mammalian Development. In T. O. Tollefsbol (Ed.), *Handbook of Epigenetics* (Second ed., pp. 11-24).
- Ong, T. P., Moreno, F. S., & Ross, S. A. (2011). Targeting the epigenome with bioactive food components for cancer prevention. *J Nutrigenet Nutrigenomics*, 4(5), 275-292. doi: 10.1159/000334585
- Ortega, R. M., Requejo, A. M., Lopez-Sobaler, A. M., Navia, B., Mena, M. C., Basabe, B., & Andres, P. (2004). Smoking and passive smoking as conditioners of folate status in young women. *J Am Coll Nutr*, 23(4), 365-371. doi: 10.1080/07315724.2004.10719380
- Pan, Y., Liu, Y., Guo, H., Jabir, M. S., Liu, X., Cui, W., & Li, D. (2017). Associations between Folate and Vitamin B12 Levels and Inflammatory Bowel Disease: A Meta-Analysis. *Nutrients*, 9(4). doi: 10.3390/nu9040382
- Patanwala, I., King, M. J., Barrett, D. A., Rose, J., Jackson, R., Hudson, M., . . . Jones, D. E. (2014). Folic acid handling by the human gut: implications for food fortification and supplementation. *Am J Clin Nutr*, 100(2), 593-599. doi: 10.3945/ajcn.113.080507
- Peuhkuri, K., Vapaatalo, H., & Korpela, R. (2010). Even low-grade inflammation impacts on small intestinal function. *World J Gastroenterol*, 16(9), 1057-1062. doi: 10.3748/wjg.v16.i9.1057
- Pietrzik, K., Bailey, L., & Shane, B. (2010). Folic acid and L-5-methyltetrahydrofolate: comparison of clinical pharmacokinetics and pharmacodynamics. *Clin Pharmacokinet*, 49(8), 535-548. doi: 10.2165/11532990-000000000-00000
- Qin, Y., & Wade, P. A. (2018). Crosstalk between the microbiome and epigenome: messages from bugs. *J Biochem*, 163(2), 105-112. doi: 10.1093/jb/mvx080
- Ratanasthien, K., Blair, J. A., Leeming, R. J., Cooke, W. T., & Melikian, V. (1974). Folates in human serum. *J Clin Pathol*, 27(11), 875-879. doi: 10.1136/jcp.27.11.875
- Ravaglia, G., Forti, P., Maioli, F., Martelli, M., Servadei, L., Brunetti, N., . . . Licastro, F. (2005). Homocysteine and folate as risk factors for dementia and Alzheimer disease. *Am J Clin Nutr*, 82(3), 636-643. doi: 10.1093/ajcn.82.3.636

- Relton, C. L., Pearce, M. S., & Parker, L. (2005). The influence of erythrocyte folate and serum vitamin B12 status on birth weight. *Br J Nutr*, *93*(5), 593-599. doi: 10.1079/bjn20041395
- Riggs, K. M., Spiro, A., 3rd, Tucker, K., & Rush, D. (1996). Relations of vitamin B-12, vitamin B-6, folate, and homocysteine to cognitive performance in the Normative Aging Study. *Am J Clin Nutr*, *63*(3), 306-314. doi: 10.1093/ajcn/63.3.306
- Rogers, L. M., Cordero, A. M., Pfeiffer, C. M., Hausman, D. B., Tsang, B. L., De-Regil, L. M., . . . Bailey, L. B. (2018). Global folate status in women of reproductive age: a systematic review with emphasis on methodological issues. *Ann N Y Acad Sci*, *1431*(1), 35-57. doi: 10.1111/nyas.13963
- Salojin, K. V., Cabrera, R. M., Sun, W., Chang, W. C., Lin, C., Duncan, L., . . . Oravec, T. (2011). A mouse model of hereditary folate malabsorption: deletion of the PCFT gene leads to systemic folate deficiency. *Blood*, *117*(18), 4895-4904. doi: 10.1182/blood-2010-04-279653
- Schnell, J. R., Dyson, H. J., & Wright, P. E. (2004). Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct*, *33*, 119-140. doi: 10.1146/annurev.biophys.33.110502.133613
- Shane, B. (1989). Folylpolyglutamate synthesis and role in the regulation of one-carbon metabolism. *Vitam Horm*, *45*, 263-335. doi: 10.1016/s0083-6729(08)60397-0
- Shane, B. (2008). Folate and vitamin B12 metabolism: overview and interaction with riboflavin, vitamin B6, and polymorphisms. *Food Nutr Bull*, *29*(2 Suppl), S5-16; discussion S17-19. doi: 10.1177/15648265080292S103
- Shane, Barry. (2009). *Folate in Health and Disease* (L. B. Bailey Ed. Second ed.). Florida, USA: CRC Press Boca Raton.
- Smith, A. D., & Refsum, H. (2016). Homocysteine, B Vitamins, and Cognitive Impairment. *Annu Rev Nutr*, *36*, 211-239. doi: 10.1146/annurev-nutr-071715-050947
- Smith, A. D., Refsum, H., Bottiglieri, T., Fenech, M., Hooshmand, B., McCaddon, A., . . . Obeid, R. (2018). Homocysteine and Dementia: An International Consensus Statement. *J Alzheimers Dis*, *62*(2), 561-570. doi: 10.3233/JAD-171042
- Snow, C. F. (1999). Laboratory diagnosis of vitamin B12 and folate deficiency: a guide for the primary care physician. *Arch Intern Med*, *159*(12), 1289-1298. doi: 10.1001/archinte.159.12.1289
- Staudacher, E. (2012). Methylation--an uncommon modification of glycans. *Biol Chem*, *393*(8), 675-685. doi: 10.1515/hsz-2012-0132
- Steinberg, S. E., Campbell, C. L., & Hillman, R. S. (1982). The role of the enterohepatic cycle in folate supply to tumour in rats. *Br J Haematol*, *50*(2), 309-316. doi: 10.1111/j.1365-2141.1982.tb01921.x
- Stover, P. J., & Field, M. S. (2011). Trafficking of intracellular folates. *Adv Nutr*, *2*(4), 325-331. doi: 10.3945/an.111.000596

- Stover, P., & Schirch, V. (1990). Serine hydroxymethyltransferase catalyzes the hydrolysis of 5,10-methenyltetrahydrofolate to 5-formyltetrahydrofolate. *J Biol Chem*, 265(24), 14227-14233.
- Strobbe, S., & Van Der Straeten, D. (2017). Folate biofortification in food crops. *Curr Opin Biotechnol*, 44, 202-211. doi: 10.1016/j.copbio.2016.12.003
- Tabor, H., & Wyngarden, L. (1959). The enzymatic formation of formiminotetrahydrofolic acid, 5,10-methenyltetrahydrofolic acid, and 10-formyltetrahydrofolic acid in the metabolism of formiminoglutamic acid. *J Biol Chem*, 234(7), 1830-1846.
- Titus, S. A., & Moran, R. G. (2000). Retrovirally mediated complementation of the glyB phenotype. Cloning of a human gene encoding the carrier for entry of folates into mitochondria. *J Biol Chem*, 275(47), 36811-36817. doi: 10.1074/jbc.M005163200
- Varela-Moreiras, G., & Selhub, J. (1992). Long-term folate deficiency alters folate content and distribution differentially in rat tissues. *J Nutr*, 122(4), 986-991. doi: 10.1093/jn/122.4.986
- Visentin, M., Diop-Bove, N., Zhao, R., & Goldman, I. D. (2014). The intestinal absorption of folates. *Annu Rev Physiol*, 76, 251-274. doi: 10.1146/annurev-physiol-020911-153251
- Wierdsma, N. J., van Bokhorst-de van der Schueren, M. A., Berkenpas, M., Mulder, C. J., & van Bodegraven, A. A. (2013). Vitamin and mineral deficiencies are highly prevalent in newly diagnosed celiac disease patients. *Nutrients*, 5(10), 3975-3992. doi: 10.3390/nu5103975
- Woeller, C. F., Anderson, D. D., Szebenyi, D. M., & Stover, P. J. (2007). Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *J Biol Chem*, 282(24), 17623-17631. doi: 10.1074/jbc.M702526200
- Wolosker, H. (2006). D-serine regulation of NMDA receptor activity. *Sci STKE*, 2006(356), pe41. doi: 10.1126/stke.3562006pe41
- Xu, X., & Chen, J. (2009). One-carbon metabolism and breast cancer: an epidemiological perspective. *J Genet Genomics*, 36(4), 203-214. doi: 10.1016/S1673-8527(08)60108-3
- Yamaoka, T., Kondo, M., Honda, S., Iwahana, H., Moritani, M., Ii, S., . . . Itakura, M. (1997). Amidophosphoribosyltransferase limits the rate of cell growth-linked de novo purine biosynthesis in the presence of constant capacity of salvage purine biosynthesis. *J Biol Chem*, 272(28), 17719-17725. doi: 10.1074/jbc.272.28.17719
- Yao, R., Rhee, M. S., & Galivan, J. (1995). Effects of gamma-glutamyl hydrolase on folyl and antifolylpolyglutamates in cultured H35 hepatoma cells. *Mol Pharmacol*, 48(3), 505-511.
- Yeh, Y. C., & Greenberg, D. M. (1965). Purification and properties of N5, N10-Methylenetetrahydrofolate dehydrogenase of calf thymus. *Biochim Biophys Acta*, 105(2), 279-291. doi: 10.1016/s0926-6593(65)80152-7
- Zeng, H., Chen, Z. S., Belinsky, M. G., Rea, P. A., & Kruh, G. D. (2001). Transport of methotrexate (MTX) and folates by multidrug resistance protein (MRP) 3 and MRP1: effect of polyglutamylation on MTX transport. *Cancer Res*, 61(19), 7225-7232.

- Zhao, R., Diop-Bove, N., Visentin, M., & Goldman, I. D. (2011). Mechanisms of membrane transport of folates into cells and across epithelia. *Annu Rev Nutr*, *31*, 177-201. doi: 10.1146/annurev-nutr-072610-145133
- Zhao, R., Matherly, L. H., & Goldman, I. D. (2009). Membrane transporters and folate homeostasis: intestinal absorption and transport into systemic compartments and tissues. *Expert Rev Mol Med*, *11*, e4. doi: 10.1017/S1462399409000969
- Zheng, Y., & Cantley, L. C. (2019). Toward a better understanding of folate metabolism in health and disease. *J Exp Med*, *216*(2), 253-266. doi: 10.1084/jem.20181965

CHAPTER 2

Folate-dependent cognitive impairment associated with specific gene networks in the adult mouse hippocampus*

* This chapter was completed with the following contributions: Caroline R. Morgan, Caleb R. Schreiner, and Chris G. Schreiner contributed to animal care and cognition testing; Britton Upchurch and Feifan Xu contributed to gene ontology; Jacqueline Baumann and Michael S. Price contributed to bioinformatics.

ABSTRACT

Short-term folate deficiency has been linked to cognitive defects. Given folate's role in regulating nucleotide synthesis and DNA and histone methylation, these changes are often linked to altered gene expression and might be controlled by specific regulatory networks. In our study we examined the effects of folic acid deficient or replete diets in mice, containing either no source of folate or normal folic acid intake, beginning post-weaning and persisting through the end of adult life at 18 mo. Our goal was to assess levels of cognition in these mice using the novel object test and then connect the cognitive results to genetic changes. Folic acid deficient mice showed significant memory impairment ($P < 0.05$) compared to control counterparts beginning at 5 mo and persisting through 17 mo, as determined by the novel object test. These deficits were associated with 363 significantly downregulated and 101 significantly upregulated genes ($P < 0.05$, fold change > 2) in the deficient condition compared to the control condition in microarray analysis of hippocampal tissue. Many of these gene expression changes were determined to be specific to the hippocampus. Significant ontological categories for differential genes included nucleotide regulation, ion channel activity, and MAPK signaling; while some of these categories contain genes previously mapped to cognitive decline, other genes have not previously been associated with cognition. To determine proteins possibly involved in regulation of these genes, we performed bioinformatics analysis and found enriched motifs for MafB and Zfp410 binding sites. These genes and enriched motifs may represent targets for treatment or investigation of memory-related diseases.

INTRODUCTION

Vitamin B9 can be found in a number of forms, referred to together as folates. The two most common folates are 5-methyltetrahydrofolate (5MTHF), the biologically active form found naturally in food, and folic acid (FA), the synthetic form used in supplements and food fortification (Patanwala et al., 2014). Folates are important one-carbon carriers in a number a cellular reactions including nucleic acid metabolism, amino acid metabolism, maintenance of DNA stability, and production of S-adenosylmethionine (SAM) for methylation of nucleic acids, neurotransmitters, phospholipids, histones, and other proteins (L. B. Bailey & Gregory, 1999; L. B. Bailey et al., 2015; Bottiglieri, Hyland, & Reynolds, 1994; Paik, Paik, & Kim, 2007). Gene methylation and expression often have an inverse relationship with decreased methylation resulting in increased expression (Garcia, Luka, Loukachevitch, Bhanu, & Wagner, 2016; Nicolas Veland, 2017). Therefore, maintenance of proper folate levels helps avoid aberrant DNA methylation patterns, thus ensuring healthy transcriptional regulation. Similarly, decreased folate levels can lead to altered nucleic acid metabolism which can be expected to result in altered gene functions. As such, folate deficiency has been associated with a number of diseases including various cancers, cardiovascular disease, and cognitive defects (Stover, 2004). Although many foods in Western cultures contain mandatory folic acid fortification, a variety of circumstances can lead to chronic folate deficiency including poor diet, intestinal diseases, medications, and gene polymorphisms (Allen, 2008; Elsborg & Larsen, 1979; Finnell, Shaw, Lammer, & Rosenquist, 2008; Kremer, Galivan, Streckfuss, & Kamen, 1986).

Evidence has accumulated in recent years linking folate deficiency at many stages of life to neurological deficits, with the largest investigative emphasis on early life since the discovery of the critical role of folate in neural tube development (Pogribny et al., 2008; Tomizawa et al.,

2015; Tsankova, Renthal, Kumar, & Nestler, 2007). Folate deficiency in rat pups during weaning has been shown to induce cognitive impairments that cannot be fully rescued by post-weaning supplementation (Berrocal-Zaragoza et al., 2014). Such deficits have been linked to expression changes in genes associated with DNA methyltransferases (DMNTs) in the rat hippocampus (Miller & Sweatt, 2007). Additionally, folate and choline deficiencies in a rat pup's mother during gestation results in increased apoptosis and altered choline metabolism in the pup's hippocampus (Jadavji, Deng, Malysheva, Caudill, & Rozen, 2015).

This observed alteration in hippocampus structure due to folate deficiency holds true in the adult hippocampus as well; dietary folate deficiency during early adult life in mice reduces the number of proliferating cells in the hippocampus (Kruman, Mouton, Emokpae, Cutler, & Mattson, 2005). Further, studies showing associations between low folate levels and altered brain status in adults extend beyond the rodent population to clinical observations that many dementia patients presenting with low plasma folate levels (Fioravanti et al., 1998). While these low folate levels may be due in part to inadequate ingestion, it is likely that inadequate absorption and utilization in late life are likely also important factors. A study of elderly Swedish patients receiving no folate or vitamin B12 supplementation determined that low serum levels of folate or B12 are indicative of a nearly doubled the risk of developing dementia compared to patients with normal folate or B12 levels (H. X. Wang et al., 2001). However, while clinical observations appear to implicate folate as an important factor in maintaining hippocampal health, many of the molecular mechanisms underlying this relationship have yet to be elucidated.

While it is well established that folate deficiency during early life results in cognitive impairments, it is not clear whether adequate folate during major developmental milestones followed by inadequate folate in later life will exhibit the same effect. Because of the essential

role folates play not only in methylation but also in nucleic acid metabolism, we hypothesize that chronic post-weaning folate deficiency will result in expression changes in many additional hippocampal genes beyond DNMTs and that those genes may be linked to cognitive deficits. Because of the chronic nature of many conditions leading to folate deficiency in adult life, we sought to study the effects of deficiency in both of the primary folate sources (5MTHF and FA) beginning post-weaning and extending through the duration of the adult mouse's life, perhaps shedding light on additional therapeutic avenues for treating or diagnosing cognitive decline.

METHODS

Mice.

Twelve female outbred CD-1 mice (Charles River Laboratories, Wilmington, MA) were crossed with males of the same strain to produce pup for this study. This strain was chosen for the high genetic diversity, healthy offspring, and large litter size associated with outbred strains. Two weeks prior to breeding half of the mice were placed on a custom chow supplemented with FA. Parental mice were approximately 42 days old at the time of breeding. Female pups remained on the parental diet with their mothers until they had been weaned. Once the pups were weaned, half of each litter continued on the parental FA diet while the other half began a FA-deficient diet. Each dietary condition contained three independent litters with multiple pups. At least one mouse from each litter at 6 mo, 12 mo, and 18 mo was euthanized by carbon dioxide asphyxiation, and all efforts were made to minimize suffering. Experimental protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at Liberty University (protocol 3.160309). Mice were kept on a 12 hr light/dark cycle in a temperature and moisture controlled room for the duration of the study.

Diets.

Parental mice were given a custom vitamin B-deficient chow (Envigo-Teklad Diets, Madison, WI; Supplemental Table 2.1) containing 1% succinylsulfathiazole to inhibit microbial folate production in the gut. Vitamins B₆ and B₁₂ were supplemented (7 mg/kg and 25 mg/kg, respectively, Solgar, Leonia, NJ) via water for all mice since the custom chow lacked these two B Vitamins and our study wanted to determine the effects of folate (B₉) only. FA supplemented mice were given water also fortified with FA (2 mg/kg dissolved in 5 mL of water), following normal dietary recommendations (Reeves, Nielsen, & Fahey, 1993). Initial calculations (data not shown) of average daily water consumption was 5mL per day and determined the supplemental vitamin concentrations in the water. Fortified water was made fresh and changed every 3 days.

Behavioral Tests.

Novel object tests were performed on at least three mice representing independent litters for each of the dietary groups at 5 and 17 mo. Each test was composed of a 5-min interaction with two identical objects followed by a second 5-min interaction 24 hr later with one of the objects replaced with a new object. Two stopwatches recorded the amount of time spent with the two objects during this second trial. Data was recorded as the percentage of time spent investigating the new object relative to the total time investigating both objects. Familiar and novel object sets were different for each time point tested to preserve the integrity of the test and provide multiple object pairs testing reliability. At 5 mo the familiar objects were two T25 cell culture flasks filled with blue-colored water and the novel object was a green Lego object of similar shape and size. At 17 mo the familiar objects were two green turtles made of rubber and the novel object was a blue dolphin of similar size and material. Results were analyzed using a

one-tailed homoscedastic student t-test at the 0.05 level comparing deficient to control diets at each time point.

Tissue extraction and RNA Isolation.

After the completion of behavioral studies, mice were sacrificed to obtain tissue samples. After isolation of whole brain samples, a sagittal cut divided the hemispheres, and the samples were placed in RNALater (Applied Biosystems, Foster City, CA) for 10 min to preserve RNA and dehydrate the tissue. The hippocampal dissection procedure followed a previously described account (<https://www.youtube.com/watch?v=tdEvicXkMck>). Isolated hippocampi were soaked in RNALater for 24 hr at 4°C before storage at -80°C.

RNA was extracted from hippocampal tissue using Trizol (Invitrogen) according to the manufacturer's instructions. Briefly, tissue was homogenized in 0.1x Trizol volume and incubated at room temperature for 5 min. Chloroform was added up to 0.2x Trizol volume, and samples were incubated an additional 2 min at room temperature before being centrifuged at 12,000 x g at 4°C for 15 min. The aqueous phase was separated, mixed with 0.5x volume of isopropanol, incubated for 10 min at room temperature, then centrifuged at 12,000 x g at 4°C for 10 minutes. The pellet was resuspended in 75% ethanol, vortexed, and centrifuged at 12,000 x g at 4°C for 5 min. Finally, the sample was dried and resuspended in 30 µl of RNase-free water and incubated at 55°C for 10 min. The sample was quantified and purity was assessed using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

Microarray Analysis.

A 60K expression microarray was obtained from ArrayStar for each dietary condition (N=3) at each time point. This V3.0 Mouse LncRNA Array had a total of 35,923 lncRNAs and

collected from all major public databases and repositories, such as Refseq, UCSC Known Genes, Ensembl, Fantom, RNAdb, NRED, and scientific publications, and a total of 24,881 protein coding mRNAs. A total of 5000 ng of RNA was provided at a concentration of 200 ng/ μ l. RNA integrity and quantity were verified using standard denaturing agarose gel electrophoresis and the NanoDrop ND-1000. For microarray analysis, Agilent Array platform was employed, and results were normalized by Lowess normalization using R version 3.1.0 (R Foundation for Statistical Computing, Vienna, Austria) and the package “ggplot2” for images. The sample preparation and microarray hybridization were performed based on the manufacturer’s standard protocols (Shi & Shang, 2016).

Quantitative PCR.

Gene-specific quantitative PCR (qPCR) confirmations were used to validate the microarray data. A High-Capacity RNA-to-cDNA Kit (Applied BioSystems, Foster City, CA) was used to convert RNA to cDNA according to manufacturer’s instructions. Briefly, 2 μ g of RNA was incubated for 60 min in a 30 μ l reaction containing 15 μ l of 2x RT Buffer mix and 1.5 μ l of 20x RT enzyme mix brought up with nuclease-free water. The reaction was stopped by heating to 95°C for 5 min. The cDNA was diluted 1:100 with nuclease-free water for use in qPCR reactions. Primers for qPCR were designed using the UCSC BLAT tool on ArrayStar probes for significant differentially expressed genes observed in the microarray data (<https://genome.ucsc.edu/cgi-bin/hgBlat>). This tool allowed us to map all primers to the same transcript as the one targeted by ArrayStar probes by designing primers to contain the probes region when possible or to be within the same exon as the probe if the exact probe region was not possible. Primer3 was then used to design primers based on the genetic regions determined

in BLAT (<http://bioinfo.ut.ee/primer3-0.4.0/primer3/>). Parameters for Primer3 software were as follows: primer T_m was $59\pm 2^\circ\text{C}$, primer size was $20\pm 2\text{bp}$, and GC clamps were used when possible; all other criteria were left as the default software settings. Primer sequences are given in Supplemental Table 2.2. The $25\mu\text{l}$ qPCR reactions contained $0.02\mu\text{g}$ of cDNA, $12.5\mu\text{l}$ of 2x SybrGreen PCR Supermix, and forward and reverse primers with a final concentration of $0.625\mu\text{M}$ each. All reactions were performed in duplicate for each cDNA pool (N=3 independent biological replicates for each condition) using *Gapdh* as the control gene; reactions lacking cDNA template were included as controls for primer self-annealing and amplification. Amplification was performed using a BioRad MJ Mini Personal Thermal Cycler. The qPCR cycling parameters were as follows: 95°C for 3 min (1 cycle), 95°C for 10 sec followed by 59°C for 1 min (40 cycles), finished with a melt curve analysis. The amplification graphs were generated using BioRad CFX manager 2.0. The quantification cycle (C_q) values were obtained for all samples and used for quantification with the $2\Delta C_q$ method (Livak & Schmittgen, 2001).

Gene Ontology.

Gene ontology was performed on significantly differentially expressed genes from microarray analysis using GeneCodis, a gene annotation website (Carmona-Saez, Chagoyen, Tirado, Carazo, & Pascual-Montano, 2007; Nogales-Cadenas et al., 2009; Tabas-Madrid, Nogales-Cadenas, & Pascual-Montano, 2012). Two random gene lists were also analyzed using GeneCodis and compared to the list of differentially expressed genes. The following criteria were used to determine significant single enrichment gene ontologies: GeneCodis level 7 (most stringent), minimum of 28 genes associated with the ontology, Chi-square value greater than Chi-square from any random list. Additionally, the following criteria were used to determine

significant pathway enrichment gene ontologies: GeneCodis level 7, at least twice as many genes as any random list, Chi-square value greater than Chi-square from any random list. The enrichment value was calculated using the following formula:

$$\frac{\frac{\text{Number of annotated genes in input list}}{\text{Number of genes in reference list}}}{\frac{\text{Total number of genes in input list}}{\text{Total number of genes in reference list}}}$$

where the input list consisted of all significantly expressed genes from microarray analysis (with separate input lists for upregulated and downregulated genes) and the reference list consisted of all expressed genes on the microarray.

Motif Analysis.

A publicly available motif discovery tool called HOMER (Hypergeometric Optimization of Motif EnRichment) was used to identify transcription factor motifs within the target genes regulated by FA (Heinz et al., 2010). In this analysis, two files were run separately: one containing the up regulated genes and one containing the down regulated genes from the RNA microarray results of the FA diets at 5 mo. The HOMER script findMotifs.pl was used for the genome of “mouse.” Parameters included a starting point of 2000 bases downstream of the TSS and 500 upstream were specified with a background list of all the genes from the microarray test. Results consisted of *de novo* and known motif sequences generated by position weight matrixes with their p-value, percent of the target and background, and the best match of the protein that HOMER found based on their database from UCSC.

RESULTS

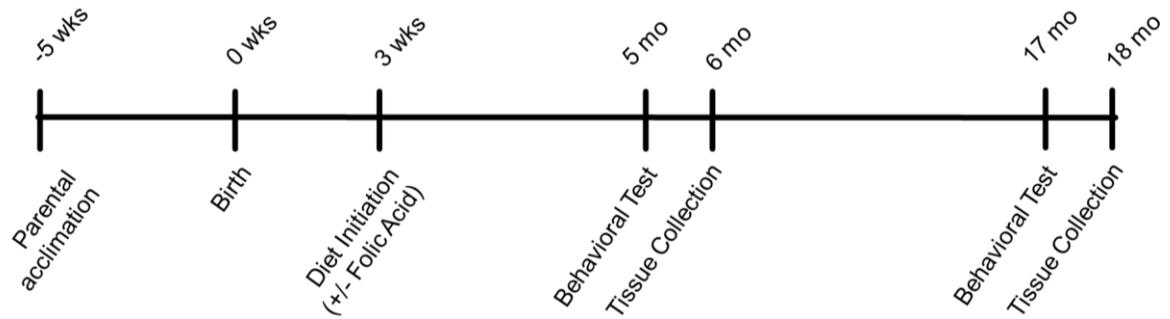
Post-weaning effects of folate on growth.

Since little is known about long-term effects of folate deficiency, we designed a study to follow folate deficient mice from infancy through late adulthood (Figure 2.1a). Cohorts of 3 FA deficient and 3 FA control mice were followed for each folate source. Since mice remained on their designated diet for the duration of their life, we weighted them at 3 stages of their adult life to ensure the diet did not affect their overall health (6 mo, 12 mo, and 18 mo – representing young adult, middle-aged adult, and senior adult). Although all mice in each diet gained weight as they aged, at any given time point there was no significant difference between the weights of mice on a folate diet compared to mice on a folate deficient diet (Figure 2.1b, $P \geq 0.6$ in all conditions).

Effects of folate restriction on cognition.

The novel object test for memory was administered to mice on all four diets to assess cognitive abilities associated with each diet. Although the weights of the mice were not affected by dietary condition, a significant difference was noted in cognitive abilities beginning at 5 mo (young adult) and persisting through 17 mo (old adult). Mice on the FA deficient diet spent significantly less time with the novel object (60% and 69% of time with novel object in control mice compared to 47% and 53% of time in deficient mice at 5 and 17 mo, respectively; $P < 0.05$ for all comparisons) compared to mice on the FA control diet, demonstrating decreased memory capabilities as assayed over a 24-hr period (Figure 2.2). Even though cognitive deficit persisted through 17 mo, the initial observance at 5 mo suggests that the effects of folate deficiency maybe

A



B

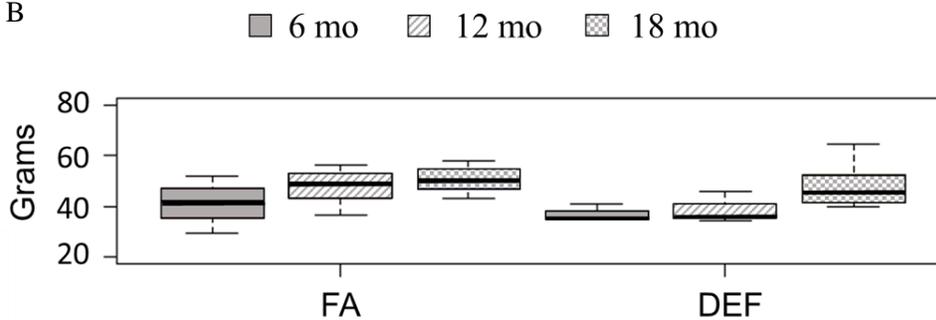


Figure 2.1. Study design.

(A) Outbred 7 week old CD1 mice were acclimated for this study with the introduction of a custom chow containing folic acid (FA, 2 mg/kg) for 5 weeks prior to the birth of their pups. Mothers and pups were kept on this diet during weaning. After weaning (3 week post birth), the pups of half of the litters remained on the parental diet, and the pups of the other half of the litters were put on a folate-deficient diet. Behavioral tests were administered 5 mo and 17 mo after birth, and tissues were collected following the completion of behavioral tests (about one month after their initiation). Results from littermates were averaged and considered to be N =1. (B) Mice were weighed throughout the course of the study to determine if diet had any effect on overall size of the mouse. Although all mice gained weight with age, there was no statistical difference between the various dietary conditions at any single time point (one-tailed homoscedastic student t-test: 6 mo FA vs DEF P = 0.29, 12 mo FA vs DEF = P 0.14, 18 mo FA vs DEF P = 0.36).

be established at this early time. Any differences that may be observed between earlier and later times cannot be determined to result directly from the folate deficiency or from secondary compensatory mechanisms.

Microarray analysis to determine folate-regulated genes.

In order to characterize molecular mechanisms associated with the cognitive decline observed in mice who were chronically FA deficient, hippocampal gene expression in the FA deficient mice was compared to expression in FA control mice using RNA microarrays. These microarrays were obtained for the 6 mo and 18 mo time points; however, the 6 mo time point was the primary focus for the remainder of the study in order to assess primary genetic changes associated with the diet-regulated cognitive decline. These 6 mo arrays found a high correlation of expressed transcripts among all RNA preps both within and between dietary groups reflecting the large similarity of the transcriptomes from these mice (Figure 2.3a, $r = 0.99$). We identified 363 transcripts to be significantly downregulated and 101 transcripts to be significantly upregulated ($P < 0.05$, fold change > 2) in the deficient condition compared to the FA condition (Figure 2.3b). A subset of 16 of these transcripts were confirmed by qPCR to be differentially regulated due to FA supplementation (Figure 2.4a-c, $P \leq 0.05$).

Further, significant genes from the 6 and 18 mo microarrays were compared to determine persistent changes in gene expression. A total of 114 genes were determined to exhibit differential expression beginning at 6 mo and persisting through 18 mo (Supplemental Table 2.3). Additionally, a microarray was obtained for liver tissue at 18 mo to determine if any endpoint genes were altered both in the tissue exhibiting significant physiological changes

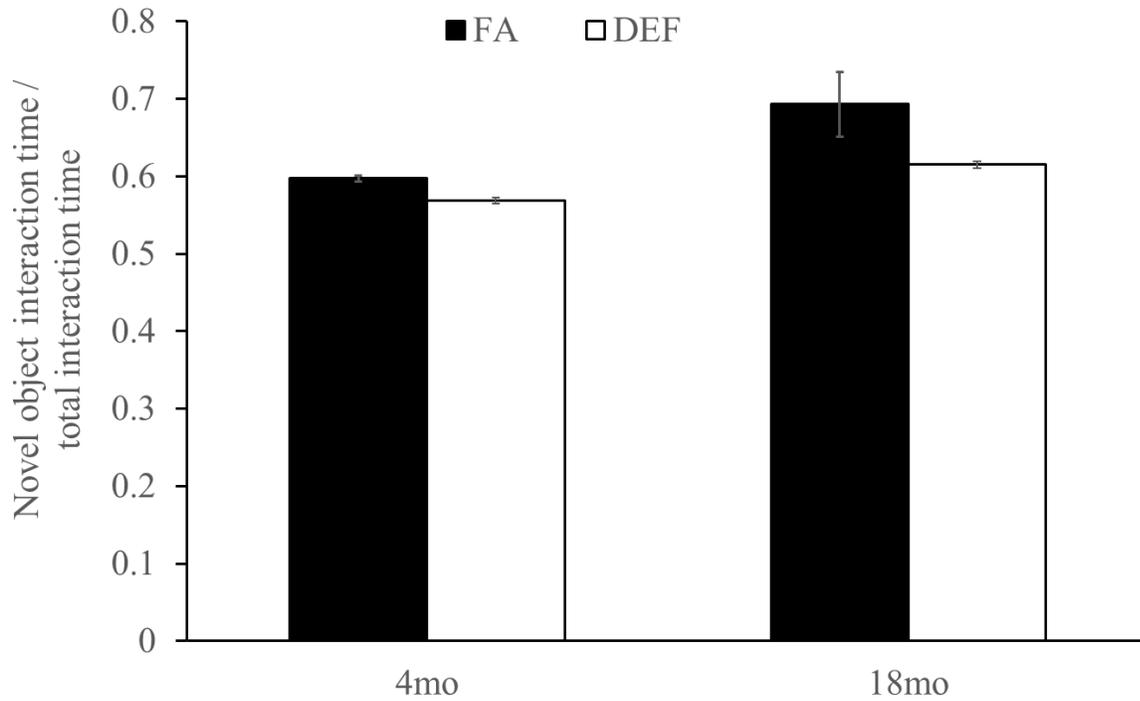


Figure 2.2. Novel object test for object memory.

A novel object test was used to assess memory in each dietary condition by comparing percent time the mouse spends with a new object versus an old object. Three litters of mice were used per condition with multiple litter mates averaged to represent one data point. Five month old folic acid (FA) litters contained 4 mice and FA-deficient (DEF) litters contained 3 mice. Each 17 month old litter contained 3 mice. All comparisons represent $P \leq 0.03$. Error bars represent SEM values calculated from the independent biological replicates across all litters.

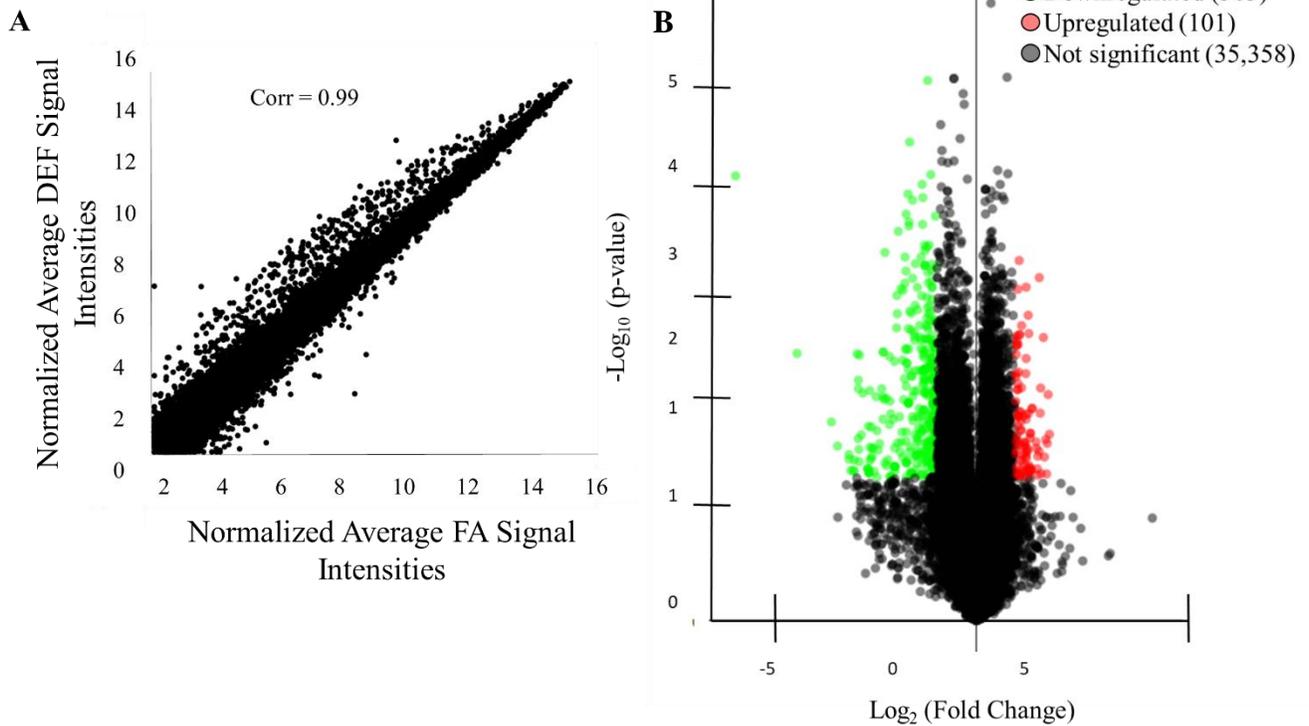
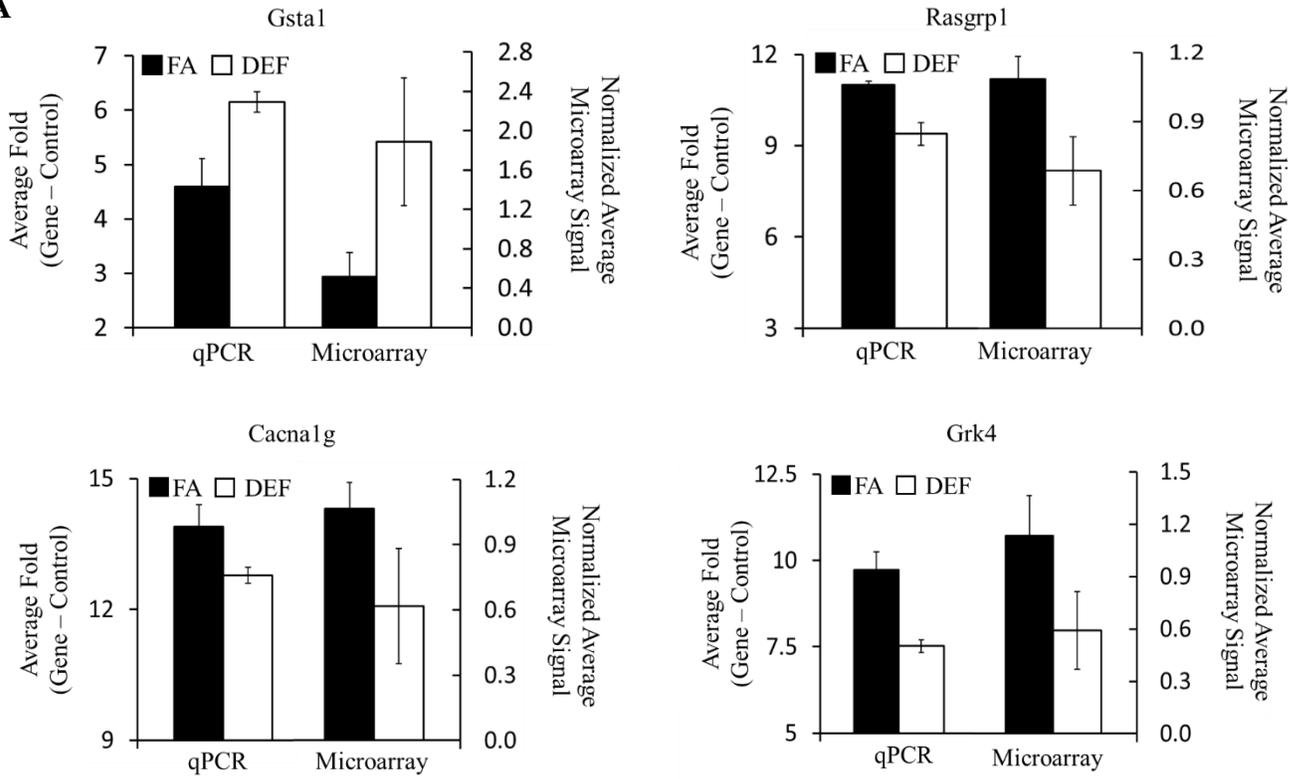
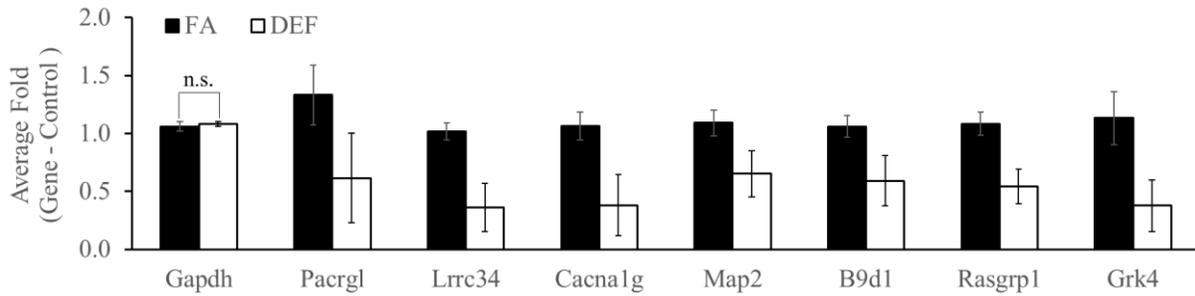
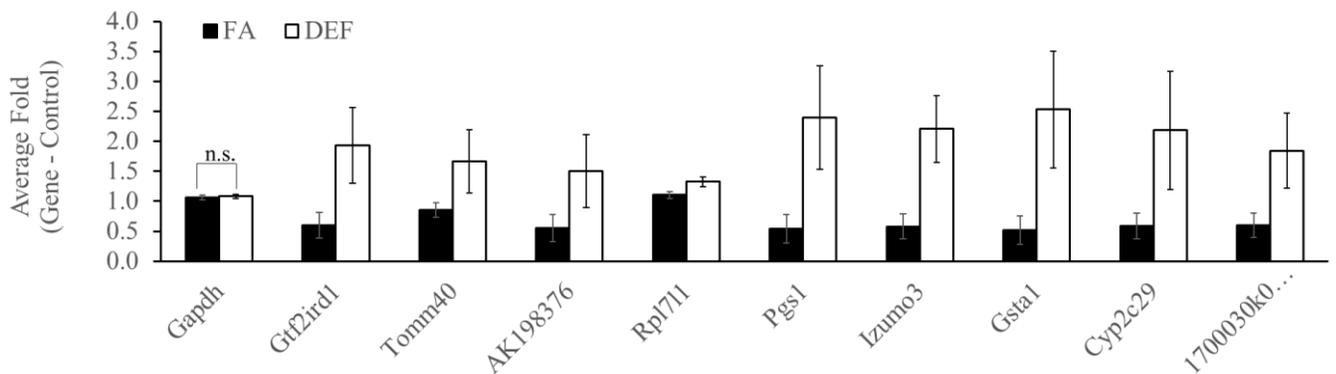


Figure 2.3. Microarray analysis of mice with and without folic acid.

(A) A scatter plot is shown comparing the microarray data of the average normalized signal intensities for the three folic acid (FA) replicates versus the average normalized signal intensities for the three deficient (DEF) replicates. The correlation coefficient was calculated using R version 3.1.0. Correlations between individual replicates within the folic acid group (0.99, 0.98, 0.98) and within the deficient group (0.98, 0.99, 0.98) were also calculated to show the similarities within each condition. (B) A total of 363 mRNA or lncRNA regions were downregulated (green) in the deficient condition compared to the folic acid condition, and a total of 101 mRNA and lncRNA regions were upregulated (red) in the same comparison as indicated by the volcano plot. An additional 35,358 mRNA and lncRNA regions were not differentially regulated (gray). Genes were considered differentially regulated if they had fold change >2 and $P < 0.05$ by microarray analysis.

Figure 2.4. qPCR confirmations of specific genes shown to be differentially expressed in microarray data.

A subset of genes considered to be significant (fold change > 2 and $P < 0.05$) by microarray analysis of the hippocampus were confirmed to be significant by qPCR ($P < 0.05$). Genes for qPCR confirmations were selected from microarray data using one of the following criteria: fold change > 2 and $P < 0.05$ in folic acid (FA) versus deficient (DEF) comparisons (N=3) or fold change > 2 and $P < 0.05$ in the folate (FA and 5-methyltetrahydrofolate) versus deficient comparisons (N=6). Primers were designed around the same region or within the same exon as the microarray probe when possible to ensure qPCR was targeting the same transcript as the microarray. Each gene was tested using at least N=2 for independent biological samples for each dietary condition. (A) Four examples of microarray versus qPCR comparisons of confirmed genes are shown. A number of genes were confirmed to be downregulated (B) or upregulated (C) in the deficient condition compared to the folic acid condition. Each experimental gene was normalized to the average *Gapdh* expression for its respective dietary category. *Gapdh* expression was not significantly altered (n.s.) in any deficient versus folate comparisons. Plotted *Gapdh* levels are relative to the expression of one FA *Gapdh* replicate to demonstrate the lack of variability between independent replicates and to demonstrate the lack of folate dependence. Error bars represent SEM values calculated from the 3 normalized array values or 3 fold values prepared from three independent biological replicates.

A**B****C**

(hippocampus) and the tissue regulating storage, modification, and distribution of FA (liver). A total of 7 genes were differentially regulated in both the hippocampus and the liver at 18 mo (Supplemental Table 2.4). Of these 7 genes, 6 were regulated in the same direction in both tissues; one gene, *Dclk1* was upregulated in the FA deficient liver and downregulated in the FA deficient hippocampus. *Dclk1* is a tumor stem cell marker that regulates pluripotency and angiogenic factors in both colorectal and pancreatic cancer, both of which are associated with folate deficiency (J. M. Bailey et al., 2014; Chandrakesan et al., 2017; Chittiboyina, Chen, Chiorean, Kamendulis, & Hocevar, 2018; Kim, 2003; Sureban et al., 2013; Vedeld, Skotheim, Lothe, & Lind, 2014).

Folate-dependent genes specific to hippocampus.

Since folate is known to have tissue-specific effects, we further investigated if any of our 16 confirmed differentially expressed genes were indeed specific for hippocampus (Anderson, Sant, & Dolinoy, 2012; Kopp, Morisset, & Rychlik, 2017). We analyzed the same 16 transcripts using cDNA pools generated from liver and heart tissue from the same mice. Interestingly, none of these genes were differentially expressed in the heart, and only one (*Pacrgl*, $P = 0.04$) was differentially expressed in the liver (Figure 2.5a, Supplemental Figure 2.1). Further, we observed that this *Pacrgl* gene was significantly downregulated in the deficient hippocampus and significantly upregulated in the deficient liver (Figure 2.5b), an observation that is consistent with previous studies that noted the hippocampus and liver exhibit opposite genetic responses to short-term folate deficiency (Alonso-Aperte & Varela-Moreiras, 1996; Kim et al., 1997; Kotsopoulos, Sohn, & Kim, 2008; Najjar et al., 2018). Thus, it is likely that many of the differentially regulated genes found on the microarray are specific to the hippocampus and may

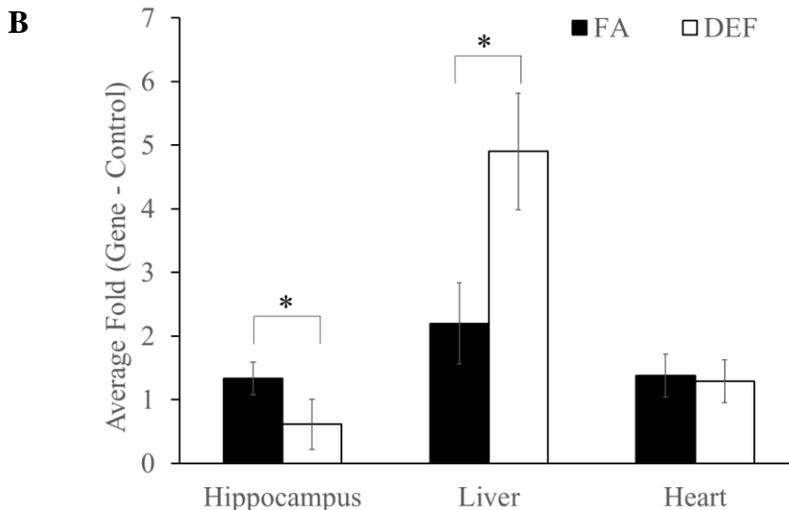
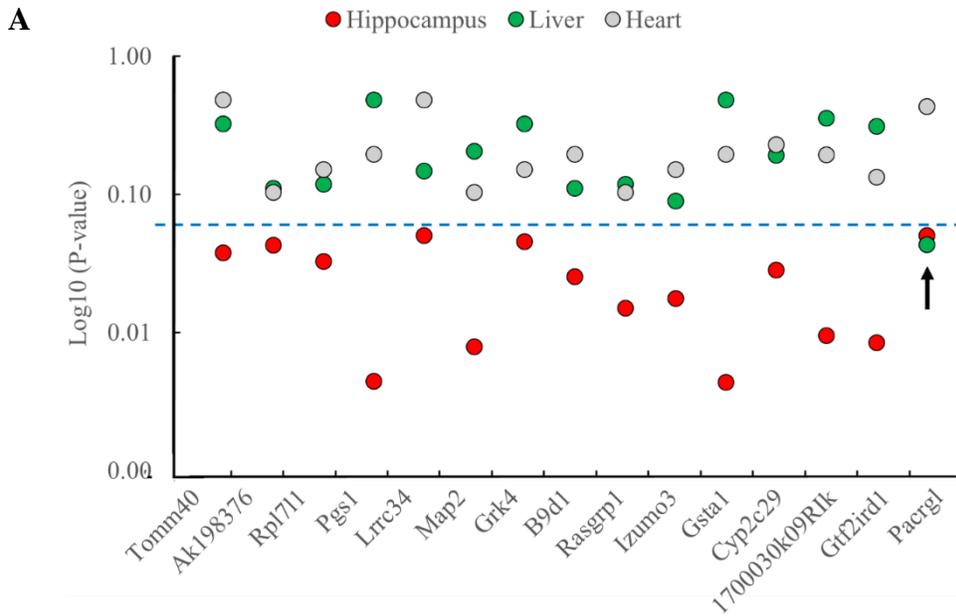


Figure 2.5. Folate-dependent genes are not ubiquitously expressed.

Each of the genes confirmed to be differentially expressed in hippocampus was also tested in liver and heart tissue of the same mice to determine if differential expression was systemic or tissue specific. Each gene was tested on 3 independent biological replicates in the folic acid (FA) and deficient (DEF) conditions. (A) Of the 15 genes tested, only one was found to be differentially expressed in the liver (*Pacrgl*), and no genes were found to be differentially expressed in the heart. The blue dotted line represents a significant p-value (≤ 0.05). (B) Interestingly, only one of the differentially expressed hippocampus genes was also differentially expressed in the liver (*Pacrgl*, hippocampus P = 0.05, liver P = 0.04). Further, *Pacrgl* was significantly downregulated in the hippocampus and significantly upregulated in the liver in the deficient versus folate comparison. This same gene was not differentially expressed in the heart (P = 0.43). Error bars represent SEM values calculated from the 3 fold values prepared from three independent biological replicates.

be involved in the observed memory deficits found in these animals.

Gene Ontology of folate-regulated genes.

To shed further light on the molecular mechanisms involved in our observed memory deficits precipitated by chronic FA deficiency, we performed gene ontology on the differentially regulated genes using GeneCodis (Table 2.1 and Supplemental Table 2.3). The largest ontological category was the 47 downregulated genes associated with the regulation of nucleobase-containing compounds, which is not surprising considering the crucial role the folate cycle plays in nucleotide synthesis. Indeed, many of the other downregulated ontological categories involved DNA and RNA associated mechanisms. Perhaps a more interesting ontological category is the 11 downregulated genes associated with MAPK signaling pathway, a number of which (*Arrb2*, *Cacna1g*, *Fgfr4*, *Taok1*, *Cacna1h*) have been implicated in Alzheimer's disease (Boshnjaku et al., 2012; Espuny-Camacho et al., 2017; Giacomini et al., 2018; Jiang et al., 2014; Rice, Berchtold, Cotman, & Green, 2014). Additionally, two downregulated ontological categories involve ion channel activity, which has also recently been implicated in Alzheimer's disease (Thei, Imm, Kaisis, Dallas, & Kerrigan, 2018).

Gene ontology was also performed on a list of 114 genes whose regulation significantly changed in the FA deficient group at 6 mo and remained differentially expressed at 18 mo, as determined by microarray analysis (Supplemental Tables 2.3 and 2.5). While not all of the 6 mo genes continued to show significant expression changes at 18 mo, a number of ontological similarities were observed in genes whose expression change persisted through the later time point. Among these persistent categories were those related to transcriptional regulation, gated ion channels, and MAPK signaling.

| Gene Set | Gene Ontology | Genes | <i>p</i> -Value | Enrichment |
|---------------------------------------|---|----------|-----------------------|------------|
| Genes Downregulated (317 genes) | Regulation of Nucleobase- Containing Compounds | 47 Genes | 0.0489 | 1.44 |
| | Transcription from RNA Polymerase II Promotor | 22 Genes | 0.0226 | 1.85 |
| | Intracellular Protein Kinase Cascade | 18 Genes | 0.00276 | 2.34 |
| | Detection of Stimulus in Sensory Perception | 7 Genes | 2.6×10^{-5} | 5.72 |
| | Ion Channel Activity | 9 Genes | 0.0397 | 2.35 |
| | Gated Channel Activity | 8 Genes | 0.0202 | 2.72 |
| | Double Stranded DNA Binding | 5 Genes | 0.0312 | 3.22 |
| | MAPK Signaling Pathway | 11 Genes | 0.000968 | 3.58 |
| Genes Upregulated (89 Genes) | Heparin Binding | 3 Genes | 2.41×10^{-6} | 10.6 |

Table 2.1. Gene ontology of folate-regulated genes.

All genes that were differentially expressed in the deficient condition compared to the folic acid condition according to microarray analysis (fold change > 2 and $P < 0.05$) were analyzed with GeneCodis using all of the microarray probes as the background list. Two random lists of genes of the same size as the upregulated and two of the same size as the downregulated experimental lists were also run as controls for false positives in gene ontologies. Annotations are presented in this table if (a) the gene ontology represented a greater number of genes and a lower *p*-value than those resulting from a corresponding random list and (b) the annotations appeared to be relevant to this study. *P*-values were calculated by GeneCodis using Chi-square tests.

Bioinformatics.

Since altered expression implies altered gene regulation, we then examined the promoter sequences of folate-regulated genes in order to determine if any transcription factor binding sites were enriched in a diet-dependent manner. We used the motif discovery tool HOMER to determine significantly enriched motifs associated with differentially regulated transcripts from microarray analysis in the deficient mice compared to the folic acid mice. To do this we compared promoter sequences (-2000bp to +500bp) of the differentially regulated genes searching for both known and *de novo* motifs. HOMER determined 2 significantly enriched motifs ($P \leq 0.0001$) as *de novo* motifs that match previously identified binding sites for MafB and Zfp410 (Figure 2.6a). This finding is consistent with experiments demonstrating a potential neuroplasticity role for *Zfp410* in rat hippocampal dendrites (Francis et al., 2014). We confirmed that *MafB* and *Zfp410* are indeed expressed in the hippocampal transcriptome using our microarray data (Figure 2.6b).

DISCUSSION

We demonstrated that chronic post-weaning folate deficiency in mice produced memory deficits arising during early adulthood and persisting through old age. Although the mice received adequate folate during major developmental milestones, this early life folate was not sufficient to protect against detrimental outcomes of late life folate deficiency. Since the hippocampus is largely responsible for memory consolidation, we mapped differences in hippocampal gene expression between FA control and FA deficient mice and found a large number of genes to be differentially expressed in the FA deficient group. We believe that many of these genes may have implications in memory functions (Table 1, S3 Table).

A

| Motif | Position Weight Matrix | Type | p-Value | % Target | % BG | Fold | Match / Detail | Expressed |
|-------|---|----------------|---------|----------|------|------|--------------------------|-----------|
| 1. |  | <i>de novo</i> | 1e-12 | 5.54 | 0.31 | 17.9 | MafB Homer (0.591) | Yes |
| 2. |  | <i>de novo</i> | 1e-12 | 6.27 | 0.49 | 12.8 | Zfp410 Jaspar (0.643) | Yes |

B

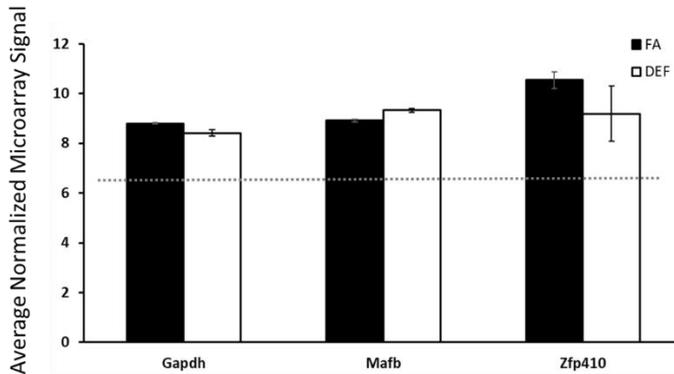


Figure 2.6. Sequence motifs enriched in folate-regulated genes.

(A) Differentially expressed RNA was analyzed with HOMER to determine significant transcription factors associated with differentially regulated genes. A target list of 224 sequences was compared to 19,048 background sequences. Transcription factor binding sites were mapped to be -2000 or +500 bp relative to the TSS of differentially expressed microarray genes. This table shows a list of motifs and their respective transcription factor that were enriched in the deficient condition compared to the folic acid condition and associated with each motif. (B) HOMER analysis showed 2 significantly enriched motifs that match the binding sequence for *Mafb* and *Zfp410*, respectively. The average normalized microarray levels (log2 values) for these genes are shown. *Mafb* and *Gapdh* are represented by one probe each while *Zfp410* (which has three probes on the array) is represented by the probe with the highest expression level. The expression of *Gapdh* is a reference along with the median normalized value of all transcripts on the array (red dotted line). Error bars represent SEM values calculated from the 3 normalized array values prepared from three independent biological replicates.

Knowing that folate is relevant in regulating nucleotide synthesis and methylation patterns in all tissue throughout the body, we compared expression of a subset of our significant hippocampal genes to expression in two other body tissues: heart and liver. None of the genes we tested were significantly different in the heart and only one was differentially expressed in the liver, demonstrating tissue-specific effects of folate deficiency. Further, the one gene (*Pacrgl*) that was differential expressed in the liver showed an opposing expression pattern between the two tissues, being downregulated in the deficient hippocampus and upregulated in the deficient liver. This is consistent with short-term folate deficiency studies that demonstrate opposite outcomes of folate deficiency in the liver versus the hippocampus (Alonso-Apperte & Varela-Moreiras, 1996; Jhaveri, Wagner, & Trepel, 2001; Kotsopoulos et al., 2008). This opposing expression pattern together with a large lack of long-term studies examining gene-specific outcomes of folate deficiency, indicate the need for future studies investigating the effects of chronic folate deficiency on other tissues. This is particularly true for both the liver due to its central role in folate metabolism and the heart due to the documented association between folate deficiency and cardiovascular disease (Blom & Smulders, 2011; Wang et al., 2007). Additionally, our tissue-specific findings demonstrate the likelihood that many of the genes we found to be statistically significant in the folate deficient hippocampus are likely associated with our observed cognitive outcomes and may serve as targets for treating memory-related diseases such as Alzheimer's and dementia.

The MAPK signaling pathway findings from our gene ontology of downregulated genes present an interesting category of potential genes to target for memory-related diseases, especially given that half of our genes in this category are already implicated in Alzheimer's disease (Supplemental Table 2.3) (Boshnjaku et al., 2012; Espuny-Camacho et al., 2017;

Giacomini et al., 2018; Jiang et al., 2014; Rice et al., 2014). Specifically, *Arrb2* polymorphisms are indicated as risk factors for late onset Alzheimer's disease; *Taok1* has been shown to actively phosphorylate tau in Alzheimer's brains; folate receptor α , which binds to promotor regions of *Fgfr4*, showed increased expression in Alzheimer's disease fibroblasts; *Cacna1g* is downregulated in Alzheimer's disease brains; and *Cacna1h* is downregulated in Alzheimer's disease human neurons. MAPK signaling in general has been implicated in Alzheimer's disease, so the additional genes noted in our ontology list may present novel specific targets for Alzheimer's investigations (Munoz & Ammit, 2010). Similarly, ion channels have been generally implicated in Alzheimer's disease, so our ion-channel ontological category genes may present specific targets for nutrition-related cognitive decline (Thei et al., 2018).

All of our ontological categories containing genes with known cognitive implications represent genes downregulated in the folate deficient condition; indeed, we found that of all of our differentially expressed genes, over three times more genes were downregulated than upregulated. This is counterintuitive since folate is an important methyl donor, and decreased DNA methylation typically leads to increased gene expression (Garcia et al., 2016; Nicolas Veland, 2017). Indeed, this inverse relationship between methylation and expression has been shown to hold true in a folate-deficient liver, with overall global hypomethylation observed in mice on folate deficient diets (Kotsopoulos et al., 2008). However, the rat liver and brain have been shown to react to folate deficiency in opposing manners with brains exhibiting hypermethylation while livers exhibit hypomethylation in response to folate deficiency (Pogribny et al., 2008). Our bioinformatics analysis of differentially expressed genes provides one possible explanation for this seeming disparity. We observed enriched motifs of binding sites for two proteins, MafB and Zfp410 (Figure 2.6). Although expression levels of these two genes

are not significantly different between dietary conditions by microarray analysis, they may play an important role in regulating a number of genes whose expression was observed to be significant in the FA deficient condition. Perhaps folate's role in altering the hippocampal expression landscape can be partially explained by post-translational modification of MafB and Zfp410 or by altered methylation status in their binding regions. This is consistent not only with folate's pivotal role in DNA methylation, but also with the implications of folate deficiency on post-translational modification of various proteins including histones, septins regulating neural tube closure, and endothelial nitric oxide synthase in cardiovascular disease (Serefidou, Venkatasubramani, & Imhof, 2019; Taylor, Dixon, Yoganayagam, Price, & Lang, 2013; Toriyama, Toriyama, Wallingford, & Finnell, 2017).

It is important to note that although these genes changed expression in response to inadequate folate intake, the exact mechanisms by which these genes changed were not determined in this study. It is tempting to assume that since folate plays a major role in DNA methylation, which in turn plays a major role in gene expression, the genes may exhibit altered expression patterns due to altered methylation status. However, methylation status was not determined in this study, nor were levels of folate or SAM in the blood and hippocampus; therefore, we cannot rule out alternative mechanisms leading to the observed behavioral deficits and associated alterations in gene expression. The role of folate in neurotransmitter and membrane phospholipid formation may be of particular importance since disruption in either of these molecules may cause dysfunctional hippocampal neurons.

In summary, we showed that folate deficiency beginning post-weaning and persisting through late adulthood induced memory deficits which can be linked to expression changes in over 400 genes, most of which were downregulated in the FA deficient condition. While many of

these genes are expectedly related to regulation of nucleotide-related functions, a significant number are also associated with MAPK signaling and ion channel activity. Further, binding sites for two proteins MafB and Zfp410 were enriched in the FA deficient condition, perhaps representing regulatory pathways altered by folate deficiency. These genes and motifs may present novel targets for treatment of memory-related diseases such as Alzheimer's disease.

ACKNOWLEDGEMENTS

The author's contributions were as follows: GI – project conception, development of overall research plan, study oversight, final decision for publication; AL – extraction of liver and heart RNA, microarray analysis, qPCR confirmations and comparisons, paper writing; CRS, CGS, and CR – animal care, data collection for behavioral tests, extraction of hippocampal RNA, production of hippocampal cDNA for microarray; JB and MP – bioinformatics; BU and FX – gene ontology

REFERENCES

- Allen, L. H. (2008). Causes of vitamin B12 and folate deficiency. *Food Nutr Bull*, 29(2 Suppl), S20-34; discussion S35-27. doi: 10.1177/15648265080292S105
- Alonso-Aperte, E., & Varela-Moreiras, G. (1996). Brain folates and DNA methylation in rats fed a choline deficient diet or treated with low doses of methotrexate. *Int J Vitam Nutr Res*, 66(3), 232-236.
- Anderson, O. S., Sant, K. E., & Dolinoy, D. C. (2012). Nutrition and epigenetics: an interplay of dietary methyl donors, one-carbon metabolism and DNA methylation. *J Nutr Biochem*, 23(8), 853-859. doi: 10.1016/j.jnutbio.2012.03.003
- Bailey, J. M., Alsina, J., Rasheed, Z. A., McAllister, F. M., Fu, Y. Y., Plentz, R., . . . Leach, S. D. (2014). DCLK1 marks a morphologically distinct subpopulation of cells with stem cell properties in preinvasive pancreatic cancer. *Gastroenterology*, 146(1), 245-256. doi: 10.1053/j.gastro.2013.09.050
- Bailey, L. B., & Gregory, J. F., 3rd. (1999). Folate metabolism and requirements. *J Nutr*, 129(4), 779-782. doi: 10.1093/jn/129.4.779
- Bailey, L. B., Stover, P. J., McNulty, H., Fenech, M. F., Gregory, J. F., 3rd, Mills, J. L., . . . Raiten, D. J. (2015). Biomarkers of Nutrition for Development-Folate Review. *J Nutr*, 145(7), 1636S-1680S. doi: 10.3945/jn.114.206599
- Berrocal-Zaragoza, M. I., Sequeira, J. M., Murphy, M. M., Fernandez-Ballart, J. D., Abdel Baki, S. G., Bergold, P. J., & Quadros, E. V. (2014). Folate deficiency in rat pups during weaning causes learning and memory deficits. *Br J Nutr*, 112(8), 1323-1332. doi: 10.1017/S0007114514002116
- Blom, H. J., & Smulders, Y. (2011). Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *J Inherit Metab Dis*, 34(1), 75-81. doi: 10.1007/s10545-010-9177-4
- Boshnjaku, V., Shim, K. W., Tsurubuchi, T., Ichi, S., Szany, E. V., Xi, G., . . . Mayanil, C. S. (2012). Nuclear localization of folate receptor alpha: a new role as a transcription factor. *Sci Rep*, 2, 980. doi: 10.1038/srep00980
- Bottiglieri, T., Hyland, K., & Reynolds, E. H. (1994). The clinical potential of ademetionine (S-adenosylmethionine) in neurological disorders. *Drugs*, 48(2), 137-152. doi: 10.2165/00003495-199448020-00002
- Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J. M., & Pascual-Montano, A. (2007). GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. *Genome Biol*, 8(1), R3. doi: 10.1186/gb-2007-8-1-r3
- Chandrakesan, P., Yao, J., Qu, D., May, R., Weygant, N., Ge, Y., . . . Houchen, C. W. (2017). Dclk1, a tumor stem cell marker, regulates pro-survival signaling and self-renewal of intestinal tumor cells. *Mol Cancer*, 16(1), 30. doi: 10.1186/s12943-017-0594-y

- Chittiboyina, S., Chen, Z., Chiorean, E. G., Kamendulis, L. M., & Hocevar, B. A. (2018). The role of the folate pathway in pancreatic cancer risk. *PLoS One*, *13*(2), e0193298. doi: 10.1371/journal.pone.0193298
- Elsborg, L., & Larsen, L. (1979). Folate deficiency in chronic inflammatory bowel diseases. *Scand J Gastroenterol*, *14*(8), 1019-1024.
- Espuny-Camacho, I., Arranz, A. M., Fiers, M., Snellinx, A., Ando, K., Munck, S., . . . De Strooper, B. (2017). Hallmarks of Alzheimer's Disease in Stem-Cell-Derived Human Neurons Transplanted into Mouse Brain. *Neuron*, *93*(5), 1066-1081 e1068. doi: 10.1016/j.neuron.2017.02.001
- Finnell, R. H., Shaw, G. M., Lammer, E. J., & Rosenquist, T. H. (2008). Gene-nutrient interactions: importance of folic acid and vitamin B12 during early embryogenesis. *Food Nutr Bull*, *29*(2 Suppl), S86-98; discussion S99-100. doi: 10.1177/15648265080292S112
- Fioravanti, M., Ferrario, E., Massaia, M., Cappa, G., Rivolta, G., Grossi, E., & Buckley, A. E. (1998). Low folate levels in the cognitive decline of elderly patients and the efficacy of folate as a treatment for improving memory deficits. *Arch Gerontol Geriatr*, *26*(1), 1-13. doi: 10.1016/s0167-4943(97)00028-9
- Francis, C., Natarajan, S., Lee, M. T., Khaladkar, M., Buckley, P. T., Sul, J. Y., . . . Kim, J. (2014). Divergence of RNA localization between rat and mouse neurons reveals the potential for rapid brain evolution. *BMC Genomics*, *15*, 883. doi: 10.1186/1471-2164-15-883
- Garcia, B. A., Luka, Z., Loukachevitch, L. V., Bhanu, N. V., & Wagner, C. (2016). Folate deficiency affects histone methylation. *Med Hypotheses*, *88*, 63-67. doi: 10.1016/j.mehy.2015.12.027
- Giacomini, C., Koo, C. Y., Yankova, N., Tavares, I. A., Wray, S., Noble, W., . . . Morris, J. D. H. (2018). A new TAO kinase inhibitor reduces tau phosphorylation at sites associated with neurodegeneration in human tauopathies. *Acta Neuropathol Commun*, *6*(1), 37. doi: 10.1186/s40478-018-0539-8
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*, *38*(4), 576-589. doi: 10.1016/j.molcel.2010.05.004
- Jadavji, N. M., Deng, L., Malysheva, O., Caudill, M. A., & Rozen, R. (2015). MTHFR deficiency or reduced intake of folate or choline in pregnant mice results in impaired short-term memory and increased apoptosis in the hippocampus of wild-type offspring. *Neuroscience*, *300*, 1-9. doi: 10.1016/j.neuroscience.2015.04.067
- Jhaveri, M. S., Wagner, C., & Trepel, J. B. (2001). Impact of extracellular folate levels on global gene expression. *Mol Pharmacol*, *60*(6), 1288-1295. doi: 10.1124/mol.60.6.1288
- Jiang, T., Yu, J. T., Wang, Y. L., Wang, H. F., Zhang, W., Hu, N., . . . Tan, L. (2014). The genetic variation of ARRB2 is associated with late-onset Alzheimer's disease in Han Chinese. *Curr Alzheimer Res*, *11*(4), 408-412.

- Kim, Y. I. (2003). Role of folate in colon cancer development and progression. *J Nutr*, 133(11 Suppl 1), 3731S-3739S. doi: 10.1093/jn/133.11.3731S
- Kim, Y. I., Pogribny, I. P., Basnakian, A. G., Miller, J. W., Selhub, J., James, S. J., & Mason, J. B. (1997). Folate deficiency in rats induces DNA strand breaks and hypomethylation within the p53 tumor suppressor gene. *Am J Clin Nutr*, 65(1), 46-52. doi: 10.1093/ajcn/65.1.46
- Kopp, M., Morisset, R., & Rychlik, M. (2017). Characterization and Interrelations of One-Carbon Metabolites in Tissues, Erythrocytes, and Plasma in Mice with Dietary Induced Folate Deficiency. *Nutrients*, 9(5). doi: 10.3390/nu9050462
- Kotsopoulos, J., Sohn, K. J., & Kim, Y. I. (2008). Postweaning dietary folate deficiency provided through childhood to puberty permanently increases genomic DNA methylation in adult rat liver. *J Nutr*, 138(4), 703-709. doi: 10.1093/jn/138.4.703
- Kremer, J. M., Galivan, J., Streckfuss, A., & Kamen, B. (1986). Methotrexate metabolism analysis in blood and liver of rheumatoid arthritis patients. Association with hepatic folate deficiency and formation of polyglutamates. *Arthritis Rheum*, 29(7), 832-835.
- Kruman, II, Mouton, P. R., Emokpae, R., Jr., Cutler, R. G., & Mattson, M. P. (2005). Folate deficiency inhibits proliferation of adult hippocampal progenitors. *Neuroreport*, 16(10), 1055-1059. doi: 10.1097/00001756-200507130-00005
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25(4), 402-408. doi: 10.1006/meth.2001.1262
- Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron*, 53(6), 857-869. doi: 10.1016/j.neuron.2007.02.022
- Munoz, L., & Ammit, A. J. (2010). Targeting p38 MAPK pathway for the treatment of Alzheimer's disease. *Neuropharmacology*, 58(3), 561-568. doi: 10.1016/j.neuropharm.2009.11.010
- Najar, R. A., Wani, N. A., Bhat, J. A., Dar, N. J., Rahat, B., Gupta, A. P., . . . Hamid, A. (2018). Modulation of dietary folate with age confers selective hepatocellular epigenetic imprints through DNA methylation. *J Nutr Biochem*, 53, 121-132. doi: 10.1016/j.jnutbio.2017.10.007
- Nicolas Veland, Taiping Chen. (2017). Mechanisms of DNA Methylation and Demethylation During Mammalian Development. In T. O. Tollefsbol (Ed.), *Handbook of Epigenetics* (Second ed., pp. 11-24).
- Nogales-Cadenas, R., Carmona-Saez, P., Vazquez, M., Vicente, C., Yang, X., Tirado, F., . . . Pascual-Montano, A. (2009). GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. *Nucleic Acids Res*, 37(Web Server issue), W317-322. doi: 10.1093/nar/gkp416
- Paik, W. K., Paik, D. C., & Kim, S. (2007). Historical review: the field of protein methylation. *Trends Biochem Sci*, 32(3), 146-152. doi: 10.1016/j.tibs.2007.01.006

- Patanwala, I., King, M. J., Barrett, D. A., Rose, J., Jackson, R., Hudson, M., . . . Jones, D. E. (2014). Folic acid handling by the human gut: implications for food fortification and supplementation. *Am J Clin Nutr*, *100*(2), 593-599. doi: 10.3945/ajcn.113.080507
- Pogribny, I. P., Karpf, A. R., James, S. R., Melnyk, S., Han, T., & Tryndyak, V. P. (2008). Epigenetic alterations in the brains of Fisher 344 rats induced by long-term administration of folate/methyl-deficient diet. *Brain Res*, *1237*, 25-34. doi: 10.1016/j.brainres.2008.07.077
- Reeves, P. G., Nielsen, F. H., & Fahey, G. C., Jr. (1993). AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr*, *123*(11), 1939-1951. doi: 10.1093/jn/123.11.1939
- Rice, R. A., Berchtold, N. C., Cotman, C. W., & Green, K. N. (2014). Age-related downregulation of the CaV3.1 T-type calcium channel as a mediator of amyloid beta production. *Neurobiol Aging*, *35*(5), 1002-1011. doi: 10.1016/j.neurobiolaging.2013.10.090
- Serefidou, M., Venkatasubramani, A. V., & Imhof, A. (2019). The Impact of One Carbon Metabolism on Histone Methylation. *Front Genet*, *10*, 764. doi: 10.3389/fgene.2019.00764
- Shi, Y., & Shang, J. (2016). Long Noncoding RNA Expression Profiling Using Arraystar LncRNA Microarrays. *Methods Mol Biol*, *1402*, 43-61. doi: 10.1007/978-1-4939-3378-5_6
- Stover, P. J. (2004). Physiology of folate and vitamin B12 in health and disease. *Nutr Rev*, *62*(6 Pt 2), S3-12; discussion S13. doi: 10.1111/j.1753-4887.2004.tb00070.x
- Sureban, S. M., May, R., Qu, D., Weygant, N., Chandrakesan, P., Ali, N., . . . Houchen, C. W. (2013). DCLK1 regulates pluripotency and angiogenic factors via microRNA-dependent mechanisms in pancreatic cancer. *PLoS One*, *8*(9), e73940. doi: 10.1371/journal.pone.0073940
- Tabas-Madrid, D., Nogales-Cadenas, R., & Pascual-Montano, A. (2012). GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res*, *40*(Web Server issue), W478-483. doi: 10.1093/nar/gks402
- Taylor, S. Y., Dixon, H. M., Yoganayagam, S., Price, N., & Lang, D. (2013). Folic acid modulates eNOS activity via effects on posttranslational modifications and protein-protein interactions. *Eur J Pharmacol*, *714*(1-3), 193-201. doi: 10.1016/j.ejphar.2013.05.026
- Thei, L., Imm, J., Kaisis, E., Dallas, M. L., & Kerrigan, T. L. (2018). Microglia in Alzheimer's Disease: A Role for Ion Channels. *Front Neurosci*, *12*, 676. doi: 10.3389/fnins.2018.00676
- Tomizawa, H., Matsuzawa, D., Ishii, D., Matsuda, S., Kawai, K., Mashimo, Y., . . . Shimizu, E. (2015). Methyl-donor deficiency in adolescence affects memory and epigenetic status in the mouse hippocampus. *Genes Brain Behav*, *14*(3), 301-309. doi: 10.1111/gbb.12207

- Toriyama, M., Toriyama, M., Wallingford, J. B., & Finnell, R. H. (2017). Folate-dependent methylation of septins governs ciliogenesis during neural tube closure. *FASEB J*, *31*(8), 3622-3635. doi: 10.1096/fj.201700092R
- Tsankova, N., Renthal, W., Kumar, A., & Nestler, E. J. (2007). Epigenetic regulation in psychiatric disorders. *Nat Rev Neurosci*, *8*(5), 355-367. doi: 10.1038/nrn2132
- Vedeld, H. M., Skotheim, R. I., Lothe, R. A., & Lind, G. E. (2014). The recently suggested intestinal cancer stem cell marker DCLK1 is an epigenetic biomarker for colorectal cancer. *Epigenetics*, *9*(3), 346-350. doi: 10.4161/epi.27582
- Wang, H. X., Wahlin, A., Basun, H., Fastbom, J., Winblad, B., & Fratiglioni, L. (2001). Vitamin B(12) and folate in relation to the development of Alzheimer's disease. *Neurology*, *56*(9), 1188-1194. doi: 10.1212/wnl.56.9.1188
- Wang, X., Qin, X., Demirtas, H., Li, J., Mao, G., Huo, Y., . . . Xu, X. (2007). Efficacy of folic acid supplementation in stroke prevention: a meta-analysis. *Lancet*, *369*(9576), 1876-1882. doi: 10.1016/S0140-6736(07)60854-X

SUPPLEMENTAL INFORMATION

| Formula | g/Kg |
|--|-------------|
| Casein, Vitamin-Free” Test | 195.0 |
| DL-methionine | 3.0 |
| Sucrose | 340.267 |
| Corn Starch | 300.0 |
| Soybean Oil | 70.0 |
| Cellulose | 40.0 |
| Mineral Mix, AIN-93G-MX (94046) | 35.0 |
| Calcium Phosphate, dibasic | 4.0 |
| TBHQ, antioxidant | 0.014 |
| Choline Bitartrate | 2.5 |
| Niacin | 0.03 |
| Calcium Pantothenate | 0.016 |
| Thiamin (81%) | 0.006 |
| Riboflavin | 0.006 |
| Biotin | 0.0002 |
| Vitamin E, DL-alpha tocopheryl acetate (500 IU/g) | 0.15 |
| Vitamin A Palmitate (500,000 IU/g) | 0.008 |
| Vitamin D ₃ , choleclciferol (500,000 IU/g) | 0.002 |
| Vitamin K ₁ , phylloquinone | 0.0008 |
| Succinylsulfathiazole | 10.0 |

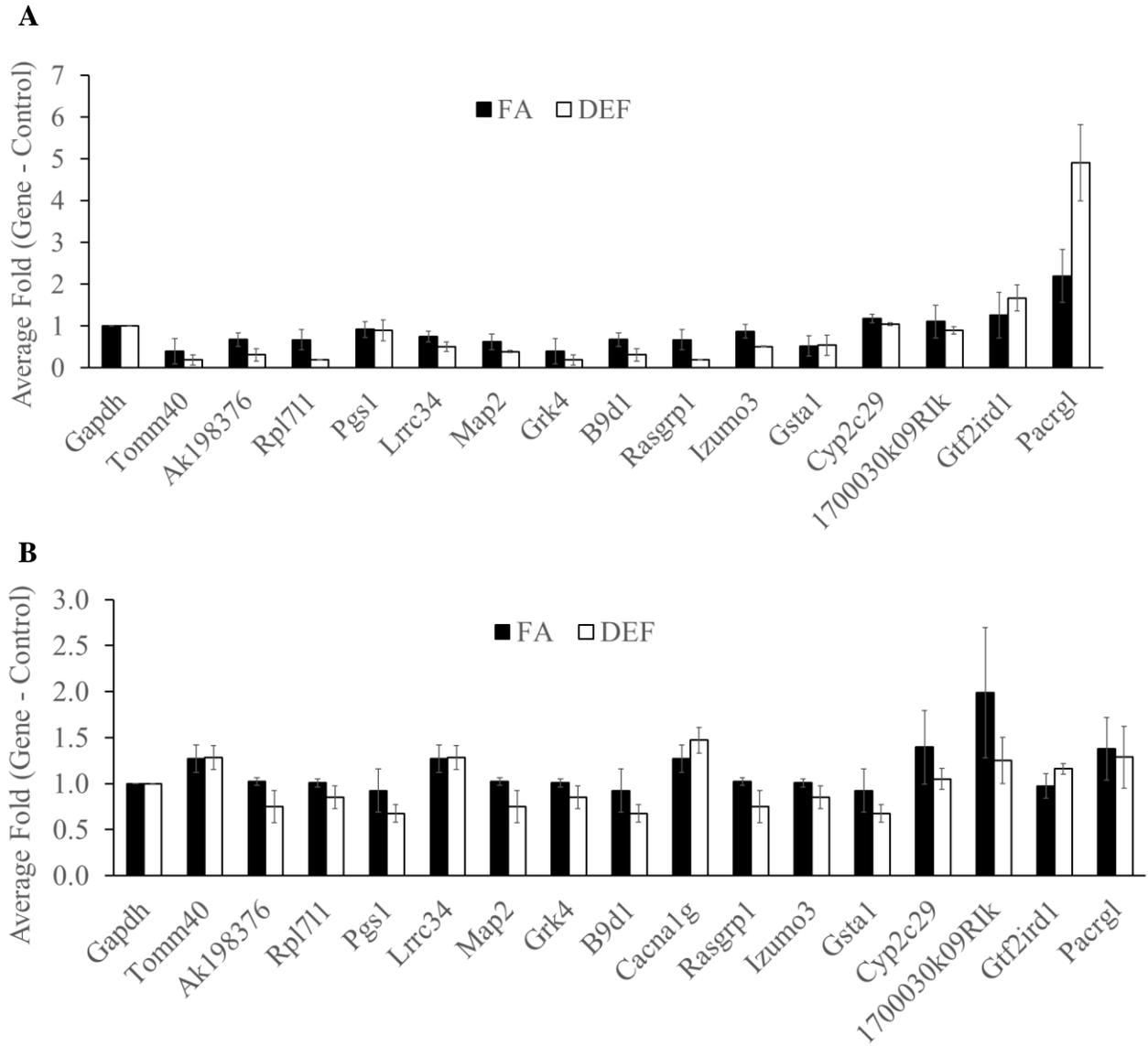
Supplemental Table 2.1. Teklad Envigo Custom folate, B₆, and B₁₂ dietary chow.

A custom dietary chow deficient in folate, B₆, and B₁₂ was obtained from Envigo Teklad Diets (Madison, WI). Methionine levels in the diet were normal. Components of the diet are listed in the table. This chow contained the following composition by percent weight: protein 17.9%, carbohydrate 61.8%, and fat 7.0%.

| Gene name | Forward primer | Reverse Primer |
|------------------|--|------------------------|
| Pgs1 | CTCCTGGACTTCACCAGAGG | CACCACTCAGGATGACGTTG |
| Cyp2c29 | GCATGGGTATGAAGCAGTGA | TTTTTCAGCCATTGGAAAGC |
| Izumo3 | CCTTTCATGGGGTCAGAGG | TACCAGGTTTCCCAACAAGG |
| 1700030k09Rik | GGGGACACAGTCTCCATCAC | TTGATGCATCCCATCTTTGA |
| Gtf2ird1 | TACACGCCTGGGATCCTAAC | ACACCCCAAAAACCAACAAC |
| Pacrgl | TCTGAGGGAGACCAAGCATC | AGTGCGCTCAGTCCTCTTTC |
| Lrrc34 | TTGCACTTTCCAGTCAATG | CGTCCACCATATACGGTTCC |
| Cacna1g | TCTGCTGTGCCTTCTTCATCAT | TCAGCACAGTCGGACTTGTTAG |
| Grk4 | TGAAGCAGAATTCTCCCTCTCAGGCTGGCATTGTGTCTCTTA | |
| Gsta1 | CCCCAATGTGAAGAAGTTCC | TTGAAAGCCTTCCTTGCTTC |
| Rasgrp1 | CTTCGAAGCCACCAGTTGTAG | ATCCTTTGGACGTGTTTGCT |
| Map2 | ATGAAGGAAAGGCACCACAC | CCTGGTCCTTCATCTCTGGA |
| B9d1 | ATCACCGGGCAGGTAGAGA | CAGTCCTGGCCATACACAAA |
| Tomm40 | GCCAGTGAGAGCCGAGTC | CAGTGCTTGACCCAGTGC |
| Ak198376 | GGTCTCTTGCCTTGAGATGG | CCCCTGCCTACTCTCTTCAA |
| Rpl7l1 | GGGGCCTATGGAGAAGTAGC | TTAACCACTGAGCCCACCTT |

Supplemental Table 2.2. qPCR primers for gene expression.

UCSC genome browser was used to BLAST array probes and align primers over probe regions to ensure primers amplified the same transcripts as the probes. Primer3 was then used to design primers around the probe region using the following parameters where possible: GC clamps, T_m 60_±2, and GC content 50_±10%.



Supplemental Figure 2.1. Expression of hippocampus-regulated genes in other tissues.

Each of the genes confirmed to be differentially expressed in hippocampus were also tested in liver tissue of the same mice to determine if differential expression was systemic or tissue specific. (A) Of the 16 genes tested, only one was found to be differentially expressed in the liver ($P = 0.04$), and (B) none were differentially expressed in the heart. *Cacna1g* was omitted from the liver graph as the gene was not expressed in the liver. Error bars represent SEM values calculated from the 3 fold values prepared from three independent biological replicates.

| Regulation | Gene |
|------------|---|
| Down | Socs3, Sec14l4, Csf1r, Tesc, 4930539M17Rik, mouselincRNA1197, Gm12730m, AV099323, AK047272, humanlincRNA2246, Psmc3ip, AK180852, Zfand2b, Clstn3, mouselincRNA0543, Atp13a2, Gm2464, DQ569461, H3f3b, uc.298, AK019720, AK155963, humanlincRNA1516, Cacng4, Gm13716, Pcyt1b, Gm13390, Npr1, humanlincRNA2050, Npepl1, AK133331, AK157393, E030003E18Rik, AK006673, Gm16291, mouselincRNA1097, AK133422, mouselincRNA1592, Cggbp1, Cwh43, Mtag2, AK157627, Dclk1, Herc3, AK080611, Bptf, AK133313, Gm26870, Aim2, M19910, AK040794, Atp9b, AK131777, Itfg3, AK041438, Gm13429, Gm4779, 9330020H09Rik, Gm9939, Col22a1, 1110008P14Rik, humanlincRNA0324, XLOC_015982, AK138628, AK018082, AK040264, Gm13415, 1700006J14Rik, XLOC_015776, Pogk, Kbtbd2, Pdzd7, Pacrgl, Taok1, Jade3, Zfp462, Lce1l, Zfp467, Map2, Ctnnbip1, Col11a1, Ankrd27, Rasgrp1, Syce2, Cacna1h, Gm3238, Id3, Rnls, Cacna1g, Unkl, Sp3, Camsap3, Arih2, Cyp2c29, B9d1, Rad51b, 1700030K09Rik, Olfr960, 4930505A04Rik, Ablim1, Izumo3, Tomm40, Imp4, Grk4, Elavl4, Kctd6, Sod3, Mettl23, Bcl3 |
| Up | XLOC_013198, Gm12069, Gsta1, 4933402D24Rik, Smpd4 |

Supplemental Table 2.3. Genes whose expression was altered at both 6 and 18 mo.

Microarray lists from differentially expressed genes (fold change > 2 and P<0.05) at 6 mo and 18 mo of folic acid deficiency were overlapped to determine the genes whose expression remained altered at both time points. The genes as well as the direction of their altered regulation are listed.

| Gene | Liver | | | Hippocampus | | |
|----------|-----------|-------------|---------|-------------|-------------|---------|
| | Direction | Fold Change | P-value | Direction | Fold Change | P-value |
| Dclk1* | up | 2.02087 | 0.02023 | down | 3.79599 | 0.00244 |
| Ndufa6 | up | 2.4101 | 0.00377 | up | 2.95658 | 0.00022 |
| Ptprs | down | 2.4978 | 0.0454 | down | 5.47737 | 0.00751 |
| Clstn3 | down | 2.06301 | 0.01831 | down | 2.29755 | 0.01925 |
| Zfp467 | down | 2.36251 | 0.01679 | down | 3.38203 | 0.04035 |
| Olfir874 | down | 2.53804 | 0.04679 | down | 3.45702 | 0.04131 |
| Elavl4 | down | 2.43506 | 0.01464 | down | 2.80454 | 0.01396 |

Supplemental Table 2.4. Genes whose expression was altered at 18 mo in both the hippocampus and the liver.

Microarray lists from differentially expressed genes (fold change > 2 and P < 0.05) at 18 mo in liver and hippocampus tissue were compared for overlap. An asterisk represents a gene which was differentially expressed in different directions in the two tissues.

| Gene Set | Gene Ontology | Genes |
|---------------------------------------|---|---|
| Genes Downregulated (317 genes) | Regulation of Nucleobase-Containing Compound | Kat5, Ankrd27, Zfp523, Ferd31, Arrb2, Fgfr4, Prr5, Pth1r, Tnnt2, Cbfa2t3, Dmrta2, Foxp4, Npr1, Lhx6, Foxn1, Mllt10, Bcl3, Pogk, Hpn, Sp3, Pou6f1, Rbfox2, Homez, Gtf2ird1, Zfp462, Zfp467, Hmga1, Taf9, Bptf, Ctnnbip1, Rbpj, Sfrp5, Meg3, Dock7, Snrnp70, Klf1, Msx3, Npb, Hdac10, Ablim1, Cggbp1, Psmc3ip, Id3, Zfp592, Vegfa, Mapk8ip1, Tesc |
| | Transcription from RNA Polymerase II Promoter | Kat5, Arrb2, Cbfa2t3, Foxp4, Foxn1, Polr2i, Mllt10, Bcl3, Sp3, Zfp462, Hmga1, Taf9, Bptf, Ctnnbip1, Rbpj, Sfrp5, Msx3, Hdac10, Ablim1, Psmc3ip, Id3, Vegfa |
| | Intracellular Protein Kinase Cascade | Nrbp1, Arrb2, Fgfr4, Map2k7, Bcl3, Itpkb, Mapk10, Madd, Tlr7, Sfrp5, Gpx1, Csf1r, Ccl19, Cdc42ep5, Socs3, Taok1, Vegfa, Mapk8ip1 |
| | Detection of Stimulus in Sensory Perception | Col11a1, Arrb2, Olfr961, Olfr960, Hpn, Olfr1466, Olfr314 |
| | Ion Channel Activity | Mcoln1, Cacna1g, Tomm40, Grid1, Kctd6, Hpn, Vdac2, Cacng4, Cacna1h |
| | Gated Channel Activity | Cacna1g, Tomm40, Grid1, Kctd6, Hpn, Vdac2, Cacng4, Cacna1h |
| | Double Stranded DNA Binding | Sp3, Foxp4, Aim2, Foxn1, Cggbp1 |
| MAPK Signaling Pathway | Arrb2*, Cacna1g*, Fgfr4*, Mapk10, Sos1, Map2k7, Taok1*, Mapk8ip1, Cacng4, Rasgrp1, Cacna1h* | |
| Genes Upregulated (89 Genes) | Heparin Binding | Lamc2, Fgf14, Vegfa |

Supplemental Table 2.5. Gene ontology with all genes listed.

All genes that were differentially expressed in the deficient condition compared to the folic acid condition according to microarray analysis (fold change > 2 and P<0.05) were analyzed with GeneCodis using all of the microarray probes as the background list. Two random lists of genes of the same size as the upregulated and two of the same size as the downregulated experimental lists were also run as controls for false positives in gene ontologies. Annotations are presented in this table if (a) the experimental annotation showed a greater number of genes or a better p-value than the random lists and (b) the annotations appeared to be relevant to this study. P-values were calculated by GeneCodis using Chi-square tests. Asterisks represent genes in the MAPK category with known implications in Alzheimer's disease.

| Gene Set | Gene Ontology | Genes | P-value |
|---------------------------------|--|--|------------|
| Genes downregulated (109 Genes) | Positive regulation of transcription by RNA polymerase II | Psmc3ip, Bcl3, Zfp462, Ablim1, Bptf, Mettl23, Zfp467 | 0.00960148 |
| | Detection of mechanical stimulus involved in sensory perception of sound | Col11a1, Pdzd7 | 3.44E-15 |
| | Voltage-gated calcium channel complex | Cacna1h, Cacng4, Cacna1g | 6.09E-26 |
| | Voltage-gated sodium channel complex | Cacna1h, Cacna1g | 1.61E-17 |
| | MAPK signaling pathway | Csf1r, Cacna1h, Cacng4, Cacna1g, Rasgrp1, Taok1 | 5.83E-09 |

Supplemental Table 2.6. Gene ontology of genes whose expression was altered at 6 mo and 18 mo of folic acid deficiency.

All genes that were differentially expressed in the deficient condition compared to the folic acid condition according to microarray analysis (fold change > 2 and P<0.05) at both 6 mo and 18 mo were analyzed with GeneCodis using all of the microarray probes as the background list. Annotations are presented in this table if (a) the experimental annotation showed a greater number of genes or a better p-value than the random lists and (b) the annotations were comparable to annotations in the 6 mo only ontological analysis. P-values were calculated by GeneCodis using Chi-square tests.

CHAPTER 3

Chronic folate deficiency induces changes in hepatic site-specific methylation but not gene expression in the aging mouse

ABSTRACT

Folate is an essential B vitamin that plays a vital role in maintaining DNA stability by aiding in nucleotide synthesis and acting as a primary carbon donor to the majority of methylation reactions in the body. As such, folate deficiency has been associated with a number of adverse outcomes such as cardiovascular disease, cognitive defects, and many cancers. In this study we examine the outcomes of folic acid deficiency in mice, beginning post-weaning and continuing through the duration of adult life (18 mo). My goal was to assess DNA methylation and gene expression changes in the liver resulting from chronic folic acid deficiency. A promoter microarray showed no significant changes in global DNA methylation but a number of site-specific loci exhibiting extensive hyper- and hypomethylation ($P \leq 0.05$, fold change > 2.7) were identified. These loci were identified as promoter regions for a number of genes, including three tumor suppressors: *Brcal*, *Gad1*, and *Tet2*. Despite the observed differential methylation patterns, the expression of these genes remained surprisingly unchanged both by RNA microarray analysis (FDR > 0.82) and by qPCR analysis ($P > 0.1$). Bisulfite sequencing of *Gad1* confirmed site-specific differential methylation patterns in promoter regions of the deficient mice compared to the control mice. Since these methylation patterns did not significantly alter gene expression, they may serve to show that the specific portions of the promoter region investigated in this study are not the primary loci responsible for regulating *Gad1* expression. These results indicate that alternative methylation mechanisms are not able to fully compensate for chronic folate deficiency in the liver and that not all methylation changes in hepatic promoter regions are associated with differential gene expression. Thus, they may not be the cause of observed disease states associated with folate deficiency.

INTRODUCTION

Folate is an essential B vitamin which acts as the primary one-carbon donor in purine and thymidine synthesis as well as the formation of S-adenosylmethionine (SAM), which in turn acts as the primary methyl donor for the majority of methylation reactions in the body (Mudd & Poole, 1975). SAM methylates a variety of biological molecules including DNA, RNA, proteins, and lipids (Bauerle, Schwalm, & Booker, 2015). Many of these DNA methylation reactions epigenetically regulate gene expression (Nicolas Veland, 2017). In humans, the majority of DNA methylation occurs at cytosine-guanine (CpG) dinucleotides, with close to 70-80% of human CpGs normally found in a methylated state (Grandjean, Yaman, Cuzin, & Rassoulzadegan, 2007; Kotsopoulos, Sohn, & Kim, 2008). Although CpGs are sparsely spaced throughout the genome, CpG dense regions, known as CpG islands, are located at the 5' end of many human genes, encompassing the promoter region, untranslated region, and exon 1; unlike the majority of the sparse CpGs throughout the genome, CpG islands are typically unmethylated (Esteller, 2007; Jones & Baylin, 2002). Methylation in these islands can epigenetically regulate expression of associated genes, often in an inverse manner with increased methylation resulting in decreased gene expression (Nicolas Veland, 2017).

DNA methylation is a dynamic process that is continually modulated in response to normal fluctuations in cellular needs or abnormal environmental factors (Kim, 2005). Epigenetic changes often progress gradually, sometimes resulting in reversible outcomes and other times resulting in permanent establishment of new patterns depending on timing, dosage, and longevity of environmental influences (Kotsopoulos et al., 2008). One such environmental influence that possesses great potential to impact global and site-specific DNA methylation is folate, the primary carbon donor to SAM (Christman, Sheikhnejad, Dizik, Abileah, & Wainfan, 1993).

Without folate, most cells are not able to produce SAM and thus experience a severe reduction in methylation capabilities. Only the liver and kidney possess significant capabilities to produce SAM in the absence of folate, and even these specialized organs are very limited in their ability to use alternative methyl donors such as betaine (Sunden, Renduchintala, Park, Miklasz, & Garrow, 1997). As such, folate deficiency has been associated with aberrant DNA methylation patterns and changes in DNA expression in a number of tissues, with the liver often being a tissue of focus given its central role in folate metabolism, storage, and excretion (Kim et al., 1997; Pogribny, James, Jernigan, & Pogribna, 2004; Zapisek, Cronin, Lyn-Cook, & Poirier, 1992). Physiological outcomes of folate deficiency include increased incidence of cardiovascular events, various cancers, and cognitive defects (Blom & Smulders, 2011; Kennedy et al., 2011; Smith et al., 2010; X. Xu & Chen, 2009).

Since folate is well established to be of great importance in DNA methylation, stability, and repair, and of particular importance in avoiding neural tube defects in developing embryos, the United States implemented mandatory folate supplementation, in the synthetic form of folic acid (FA), in grains beginning in the early 1990s (Choumenkovitch et al., 2002). While this mandatory supplementation has shown vast success in improving folate levels in the majority of Americans, a number of conditions still result in folate deficiency in a sizable portion of the population. These conditions include smoking, chronic alcoholism, inflammatory bowel diseases, obesity, medications such as methotrexate and anti-epileptic drugs, gene polymorphisms, and exposure to environmental toxins such as fumonisins (Devlin et al., 2000; Gloria et al., 1997; Kremer, Galivan, Streckfuss, & Kamen, 1986; Marasas et al., 2004; Morrell, 2002; Ong, Moreno, & Ross, 2011; Ortega et al., 1994; Ortega et al., 2004; Quinlivan et al., 2005; Russell et al., 1983). Many of these conditions do not arise until early adulthood and often

continue through the duration of an adult's life, yet few studies have examined the long-term effects of folate deficiency beginning in early adulthood and persisting through the duration of adult life. Thus, while it is well documented that folate deficiency early in life greatly impacts methylation and gene expression, it is unknown whether adequate folate intake early in life is sufficient to overcome folate deficiency later in life. One of the few studies examining longer-term effects of folate deficiency was conducted by Kotsopoulos, *et al.* in 2008, which investigated outcomes of mild dietary folate deficiency in rats beginning post-weaning and continuing through 30 weeks, or roughly through puberty (Kotsopoulos et al., 2008). Thus, the aim of this study was to examine FA deficiency in mice beginning post-weaning and persisting through the duration of the adult life (18 mo) to determine outcomes of chronic post-weaning folate deficiency. Given the central role of the liver in folate processing, we focused on the liver in this study, examining DNA methylation and gene expression changes that resulted from an adult lifetime of folate deficiency. We posit that chronic folate deficiency throughout adult life will induce significant methylation changes in the liver, possibly leading to altered hepatic gene expression.

METHODS

Mice and diets.

As described previously (Chapter 2), twelve female outbred CD-1 mice (Charles River Laboratories, Wilmington, MA) were crossed with males of the same strain to produce pups for this study. This strain was chosen for the high genetic diversity, healthy offspring, and large litter size associated with outbred strains. Two weeks prior to breeding half of the mice were placed on a custom chow supplemented with FA. Parental mice were approximately 42 days old at the

time of breeding. Female pups remained on the parental diet with their mothers until they had been weaned. Once the pups were weaned, half of each litter continued on the parental FA diet while the other half began a FA-deficient diet. Each dietary condition contained three independent litters with multiple pups. At least one mouse from each litter was euthanized by carbon dioxide asphyxiation at 18 mo, and all efforts were made to minimize suffering.

Experimental protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at Liberty University (protocol 3.160309). Mice were kept on a 12 hr light/dark cycle in a temperature and moisture-controlled room for the duration of the study.

Parental mice were given a custom vitamin B-deficient chow (Envigo-Teklad Diets, Madison, WI; Supplemental Table 2.1) containing 1% succinylsulfathiazole to inhibit microbial folate production in the gut. Vitamins B6 and B12 were supplemented (7 mg/kg and 25 mg/kg, respectively, Solgar, Leonia, NJ) via water for all mice since the custom chow lacked these two B vitamins and our study wanted to determine the effects of folate (B9) only. FA supplemented mice were given water also fortified with FA (2 mg/kg dissolved in 5 mL of water), following normal dietary recommendations. Initial calculations (data not shown) of average daily water consumption was 5 mL per day and determined the supplemental vitamin concentrations in the water. Fortified water was made fresh and changed every 3 days. Blood and tissue folate levels were not determined in this study since rodent liver folate stores have been demonstrated to be significantly depleted in as little as 25 days without dietary folate (Clifford, Heid, Muller, & Bills, 1990).

Tissue extraction and RNA and DNA Isolation.

After 18 mo on their respective diets, mice were sacrificed to obtain tissue samples. Whole liver samples were quartered and soaked in RNALater (Applied Biosystems, Foster City, CA) for 24 hr at 4°C before storage at -80°C. RNA was extracted from liver tissue using Trizol (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Briefly, tissue was homogenized in 0.1x Trizol volume and incubated at room temperature for 5 min. Chloroform was added up to 0.2x Trizol volume, and samples were incubated an additional 2 min at room temperature before being centrifuged at 12,000 x g at 4°C for 15 min. The aqueous phase was separated, mixed with 0.5x volume of isopropanol, incubated for 10 min at room temperature, then centrifuged at 12,000 x g at 4°C for 10 min. The pellet was resuspended in 75% ethanol, vortexed, and centrifuged at 12,000 x g at 4°C for 5 min. Finally, the sample was dried and resuspended in 30 µl of RNase-free water and incubated at 55°C for 10 min. The sample was quantified and purity was assessed using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA).

DNA was extracted by suspending 100 mg liver tissue in 1x volume Stop Buffer (20% SDS, 5 M NaCl, 0.5 M EDTA), 0.04x volume proteinase K (2.5 mg/ml), and 0.004x volume RNase A (100 mg/mL) and incubating overnight at 42°C. Tissue was homogenized by pipetting, incubated at 65°C for 1 hr, and mixed again by pipetting. 1x volume phenol-cholorform-isoamyl alcohol (PCIAA, Fisher Scientific, Hampton, NH) was added and mixed by vortexing. Samples were separated by centrifugation at 17,000 x g for 5 minutes, and the top aqueous layer was collected into a new tube. This process was repeated once more with PCIAA followed by once with chloroform to ensure DNA purity. DNA was precipitated in 3x volume of cold 100% ethanol, 0.1x volume of 3M sodium acetate, and 0.5 µl of glycogen (20 mg/ml). The samples

were incubated at -20°C for 30 min and centrifuged at 17,000 x g at 4°C for 0 min. Pellets were washed with 75% ethanol and resuspended in 200 µl of RNase-free water (Thermo Fisher Scientific, Waltham, MA). DNA quality and quantity were assessed using a Nanodrop ND-1000 spectrophotometer.

HELP Assay.

To assess site-specific whole genome methylation, the HpaII tiny fragment Enriched by Ligation-Mediated PCR (HELP) assay was used (Figure 3.1) (Oda et al., 2009; M. Suzuki & Greally, 2010). This assay uses a methylation sensitive restriction enzyme (HpaII) and a methylation insensitive enzyme (MspI) to differentially digest DNA for subsequent ligation mediated PCR followed by microarray analysis. Each sample was digested using 1 µg of DNA, 20 µl of 10x CutSmart Buffer (New England Biolabs, Ipswich, MA), and 40 Units of either HpaII or MspI (New England Biolabs, Ipswich, MA) brought up to 200 µl using RNase free water. Each digestion reaction was incubated overnight in a 37°C water bath. Digestion reactions were purified using PCIAA as described above, and DNA was resuspended in 15 µl of 10 mM Tris-HCl (pH 8). Ligation of each digested DNA pool was performed using two sets of oligonucleotides: NHpaII12 (5'-CGGCTTCCCTCG-3'), NHpaII24 (5'-GCAACTGTGCTATCCGAGGGAAGC-3'), JHpaII12 (5'-CGGCTGTTCATG-3'), and JHpaII24 (5'-CGACGTCGACTATCCATGAACAGC-3'). Corresponding 12-mer and 24-mer oligonucleotides were annealed prior to the ligation reaction. Ligation reactions contained 6 µl of 5x T4 ligase buffer (New England Biolabs, Ipswich, MA), 15.5 µl of digested DNA, 4 µl of 50 µM pre-annealed JHpaII linker, 4 µl of 50 mM pre-annealed NHpaII linkers, and 1 µl of T4 DNA ligase (4 U/µl). The ligation mix was incubated overnight

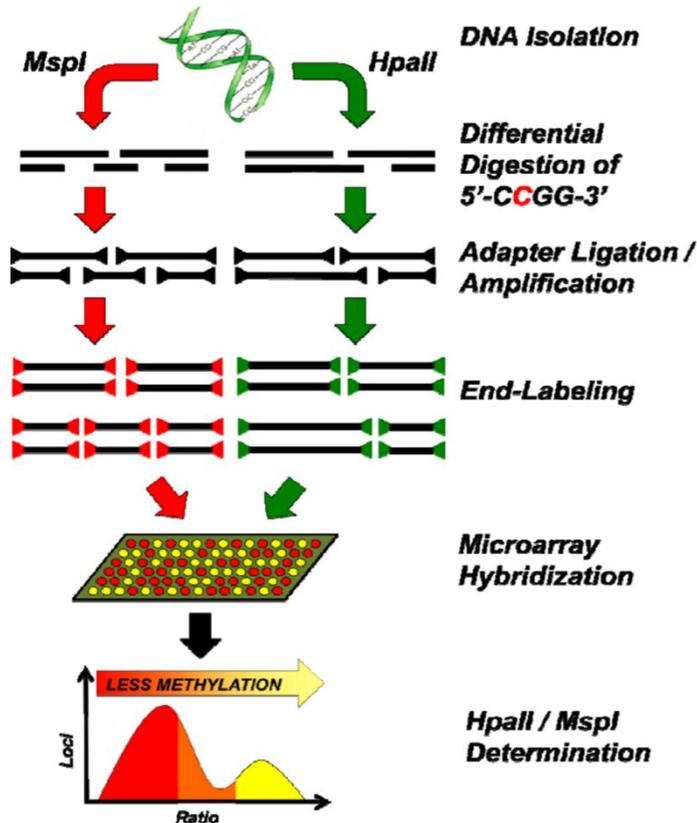


Figure 3.1. HELP assay overview.

The HpaII tiny fragment Enriched Ligation-Mediated PCR (HELP) assay allows for detection of site-specific DNA methylation changes on a global scale. First, DNA is isolated and purified. Next this DNA is differentially digested with HpaII and MspI, isoschizomers that cut the sequence 5'-CCGG-3' in a methylation sensitive and insensitive manner, respectively. Fragments produced by these digestions are ligated to adapter sequences, allowing for amplification of products ranging from 50-2000 bp. Amplified products are purified and sent to ArrayStar for end-labeling and co-hybridization on a promoter microarray. Raw microarray array data is processed to determine HpaII/MspI ratios, signifying altered DNA methylation patterns at each probe in the folic acid deficient versus control conditions. Figure copied with permission from (Taher, 2013).

at 16°C in a thermocycler then diluted to 100 µl using 10 mM Tris-HCl (pH 8). Ligation mediated PCR (LM-PCR) was used to amplify fragments within a range of 50-2000 bp. Each reaction contained the following: 25 µl of 2x SybrGreen PCR mix (ThermoFisher Scientific, Waltham MA), 0.5 µl of 100 µM NHpaII24 oligonucleotide, 0.5 µl of 100 µM JHpaII24 oligonucleotide, either 4µl of HpaII-digested DNA or 2 µl of MspI digested DNA, and up to 50 µl with RNase-free water. Twice as much HpaII-digested DNA was used compared to MspI-digested DNA to account for the difference in complexity of the MspI fragments, which may cause the fragments to amplify and saturate the solution more quickly, causing these fragments to undergo excessive cycles that may result in undesirable artifacts. This difference in DNA quantity follows a previously described technique (Oda et al., 2009; M. Suzuki & Greally, 2010). The LM-PCR scheme was as follows: 72°C for 10 min, 20 cycles of 95°C for 30 sec and 72°C for 3 min followed by a plate read, and 72°C for 10 min. LM-PCR reactions were verified using a 1% agarose gel and purified using PCIAA as described above. Quality and quantity of DNA was verified using a Nanodrop ND-1000 spectrophotometer.

DNA Promoter Microarray.

A 4 x 180K RefSeq Mouse Promoter Array was obtained from ArrayStar for each HELP-treated sample (N=3 for each dietary condition). This array investigated 22,327 gene promoter regions covered by ~180,000 probes with approximately 205 bp spacing. All MspI-digested samples were hybridized to Cy3, and all HpaII-digested samples were hybridized to Cy5. Additionally, dye hybridization was swapped for one sample from each dietary condition to ensure no dye-based bias existed in the arrays.

DNA Promoter Microarray Analysis.

The raw microarray data from ArrayStar was processed using R version 3.1.0 (R Foundation for Statistical Computing, Vienna, Austria) and packages “matrixTests” and “data.table” for analysis and “ggplots” for graphing (Appendix D). Correlation graphs were generated to examine differences between microarrays from the two dietary conditions. The background fluorescence threshold was set as 2.5 mean absolute deviations (MAD) above the median fluorescence signal of random probes (Oda et al., 2009; M. Suzuki & Greally, 2010; Taher, 2013). For a probe to be considered valid, its MspI signals had to be above the threshold in all biological replicates in all conditions; HpaII levels did not need to be above threshold since the HELP assay creates DNA fragments that may be below threshold in methylation extremes. Data from each biological replicate in each dietary condition (N=3 for FA control and deficient) was averaged, and an HpaII/MspI ratio histogram was generated for each condition. Histograms were mean-centered, and the x-axis was converted from linear to \log_2 .

Genomic Regions Enrichment of Annotations Tool (GREAT) software (Stanford University and Bejarano Lab) was used to map the top 1% most changing probe regions (0.5% most hypomethylated and 0.5% most hypermethylated, 135 regions each) to gene symbols (McLean et al., 2010). Mouse genome mm9 was used as the species assembly and the whole genome was used for background regions. Genes were assigned basal regulatory domains as follows: proximal domain from 5 kb upstream to 1 kb downstream of the transcription start site (TSS) and distal regulatory domain up to an additional 1000 kb from the proximal domain, regardless of nearby genes. Additionally, curated regulatory domains were used to override other domain recognition rules if significant experimental evidence demonstrates that an element falls

outside of the assigned regulatory domains. Probes were allowed to be mapped to multiple genes if experimental evidence or relative location did not provide a clear association to a single gene.

RNA Microarray Analysis.

A 60K expression microarray was obtained from ArrayStar to compare each dietary condition (N=3) at each time point. This V3.0 Mouse LncRNA Array had a total of 35,923 lncRNAs and collected from all major public databases and repositories, such as RefSeq, UCSC Known Genes, Ensembl, Fantom, RNAdb, NRED, and scientific publications, and a total of 24,881 protein coding mRNAs. A total of 5000 ng of RNA was provided at a concentration of 200 ng/ μ l. RNA integrity and quantity were verified using standard denaturing agarose gel electrophoresis and the NanoDrop ND-1000.

Agilent Feature Extraction software (version 11.0.1.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed with using the GeneSpring GX v12.1 software package (Agilent Technologies). After quantile normalization of the raw data, LncRNAs and mRNAs that at least 6 out of 12 samples have flags in Present or Marginal (“All Targets Value”) were chosen for further data analysis. Differentially expressed LncRNAs and mRNAs with statistical significance were identified through Volcano Plot filtering between the two groups of samples. Hierarchical Clustering was performed using the R software (version 2.15). The sample preparation and microarray hybridization were performed based on the manufacturer’s standard protocols (Shi & Shang, 2016).

Gene Expression Quantitative PCR.

Gene-specific quantitative PCR (qPCR) was used to confirm microarray results of the three genes (*Brcal*, *Gad1*, and *Tet2*), which were chosen due to their significant change in promoter methylation and physiological importance (described below). A High-Capacity RNA-to-cDNA Kit (Applied BioSystems, Foster City, CA) was used to convert RNA to cDNA according to manufacturer's instructions. Briefly, 2 μg of RNA was incubated for 60 min in a 30 μl reaction containing 15 μl of 2x RT Buffer mix and 1.5 μl of 20x RT enzyme mix brought up with nuclease-free water. The reaction was stopped by heating to 95°C for 5 min. The resulting cDNA was diluted 1:100 with nuclease-free water for use in qPCR reactions. Primers for qPCR were designed using the UCSC BLAT tool targeting ArrayStar probes from the methylation microarray which had been mapped to genes by GREAT analysis (<https://genome.ucsc.edu/cgi-bin/hgBlat>). Primer3 was then used to design primers (<http://bioinfo.ut.ee/primer3-0.4.0/primer3/>). Parameters for Primer3 software were as follows: primer T_m was $59 \pm 2^\circ\text{C}$, primer size was $20 \pm 2\text{bp}$, and GC clamps were used when possible; all other criteria were left as the default software settings. Primer sequences are given in Supplemental Table 3.1. The 25 μl qPCR reactions contained 0.02 μg of cDNA, 12.5 μl of 2x SybrGreen PCR Supermix, and forward and reverse primers with a final concentration of 0.625 μM each. All reactions were performed in duplicate for each cDNA pool (N=3 independent biological replicates for each condition) using *Gapdh* and *ActB* as the control genes; reactions lacking cDNA template were included as controls for primer self-annealing and amplification. Amplification was performed using a BioRad MJ Mini Personal Thermal Cycler. The qPCR cycling parameters were as follows: 95°C for 3 min (1 cycle), 95°C for 10 sec followed by 59°C for 1 min (40 cycles), finished with a melt curve analysis. The amplification graphs were generated using BioRad CFX

manager 2.0 software. The quantification cycle (Cq) values were obtained for all samples and used for quantification with the $2\Delta Cq$ method (Livak & Schmittgen, 2001).

Bisulfite conversion.

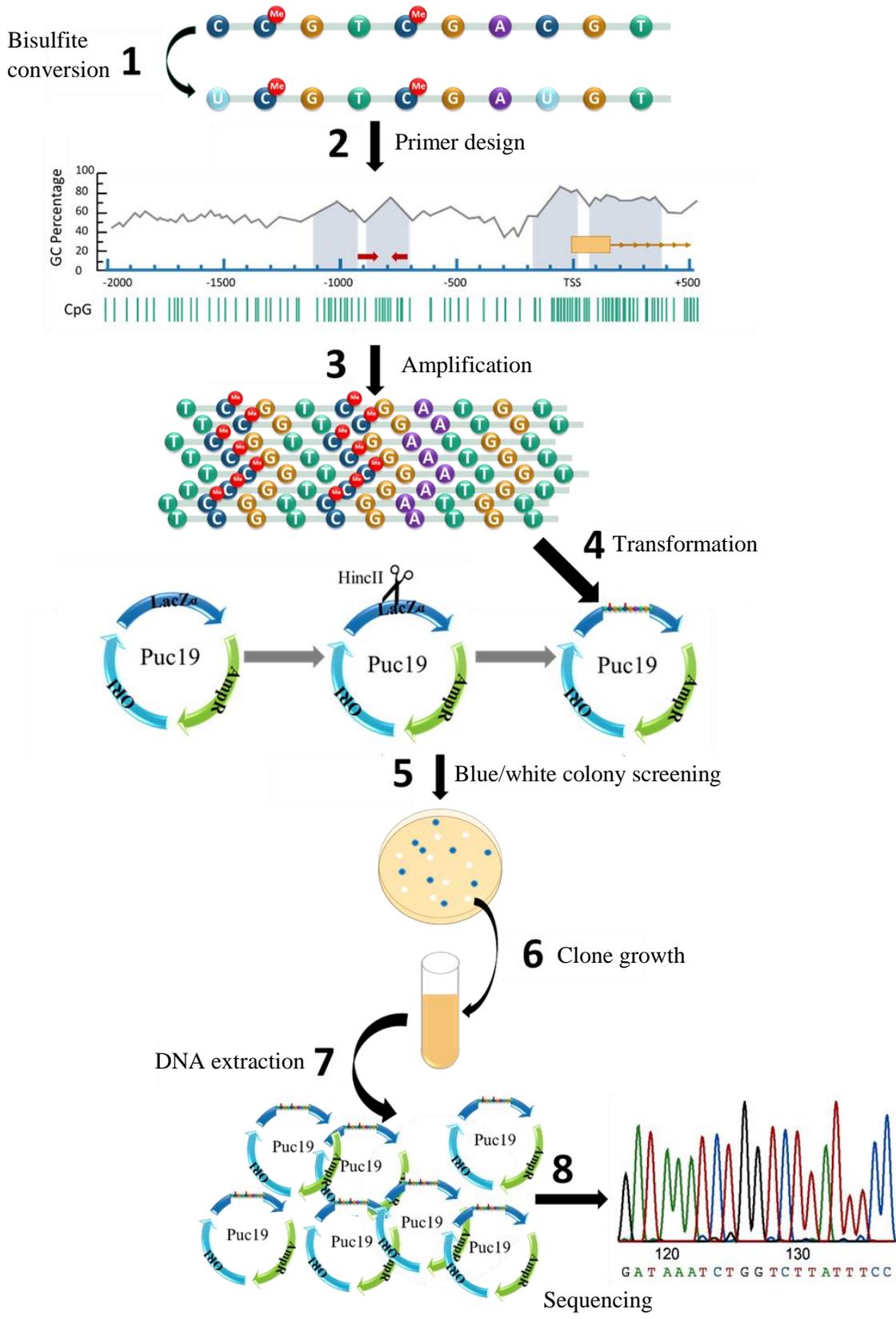
Bisulfite sequencing was used to determine methylation of specific CpGs of some significant results from the HELP promoter microarray (Figure 3.2). To begin, genomic DNA from each of the 3 independent biological replicates from each dietary condition was bisulfite converted using the EZ DNA Methylation-Gold Kit (Zymo Research, Irving, CA) following manufacturer's instructions. Briefly, 130 μ L of CT Conversion Reagent was added to 500 ng of DNA and mixed by flicking. The mixture was incubated in a thermal cycler at 98°C for 10 min followed by 64°C for 2.5 hr. 600 μ L of Binding Buffer was added, and the solution was transferred to a Zymo-Spin IC column and spun at full speed for 30 sec. 100 μ L of Wash Buffer was added and the column was spun for 30 sec. 200 μ L of M-Desulphonation Buffer was added to the column and allowed to incubate at room temperature for 20 minutes. The column was spun at full speed for 30 sec. 200 μ L of M-Wash buffer was added, the column was spun at full speed for 30 sec, and this the wash step was repeated once. All flow-through was discarded, and bisulfite converted DNA was eluted in 10 μ L of M-elution buffer by centrifuging at full speed for 30 sec.

Bisulfite PCR.

Three genes were chosen from the HELP analysis for bisulfite sequencing based on the following criteria: (1) the associated promoter array probe was within the top 1% most changing in methylation, (2) the gene is a known tumor suppressor, and (3) the gene has been studied in

Figure 3.2. Bisulfite sequencing workflow.

In order to examine methylation patterns on specific nucleotides in promoter regions of tumor suppressor genes, bisulfite sequencing was performed. (1) Genomic DNA was bisulfite treated, a process by which unmethylated cytosines are deaminated to form uracils, while methylated (Me) cytosines remain as cytosines. (2) Promotor regions were defined as -2000 bp to +500 bp from the transcription start site (TSS) of a given gene (orange). Within the promotor region, primers (red arrows) were designed, when possible, to amplify DNA within CpG islands (grey blocks). (3) Amplification of bisulfite-treated DNA results in sequences where unmethylated cytosines, which have been converted to uracils, are transcribed as thymines. However, bisulfite treatment often leads to DNA degradation, resulting in a non-homogenous pool of PCR products which often make sequencing results unclear. Therefore, PCR products were inserted into the pUC19 vector to produce clones (4). This vector was digested using HincII, which cuts the plasmid within the lacZ gene and allows for ligation of blunt-end PCR products. (5) The ligated vector was transformed into *E. coli*, which were plated on Luria Bertani (LB) agar containing Ampicillin (Amp) and X-gal, allowing for blue/white screening of the transformant colonies. In the blue white screening method, colonies with uninterrupted lacZ genes interact with X-gal to form blue colonies, while colonies with interrupted lacZ (by the insertion of PCR product) form white colonies. (6) A single white colony from the LB plate was picked and grown in 4 mL of LB Amp broth, resulting in growth of a single clone of ligated plasmid in adequate concentration for DNA sequencing. (7) Plasmid DNA was extracted from LB broth cultures and purified using a plasmid miniprep kit. (8) Purified plasmid DNA was sent for Sanger sequencing analysis. Sequence results were compared against the original non-treated DNA sequence, allowing for the detection of unmethylated CpGs by the presence of thymine rather than cytosine.



the liver. Three genes met this criteria: *Brca1*, *Gad1*, and *Tet2*. Bisulfite primers were designed within the promoter region of these genes, defined as -2000 through +500 bp relative to the TSS, using MethPrimer software <https://www.urogene.org/methprimer/> (Li & Dahiya, 2002). Multiple primer regions were chosen for each gene. Primer sequences are listed in Supplemental Table 3.2. The PCR reaction contained 1 μ L of bisulfite treated DNA (50 ng/ μ L assuming full kit recovery), 10 μ L of 2x PhusionU Hotstart PCR MasterMix (ThermoFisher Scientific, Waltham, MA), and 2 μ L of 2.5 μ M primer mix, brought up to 20 μ L with water. The PCR schematic was as follows: 98°C for 30 sec; 98°C for 10 sec, X°C for 30 sec, 72°C for 2 minutes, repeated for 30 cycles; and 72°C for 15 minutes. The annealing temperature (X) was calculated using the ThermoFisher Tm calculator, which takes into account the unique conditions needed for the high-fidelity, uracil-tolerant PhusionU polymerase <https://www.thermofisher.com/us/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html>. Calculated annealing temperatures are listed in Supplemental Table 3.2. PCR products were verified on a 2% agarose gel then cleaned using a DNA Clean and Concentrator kit (Zymo Research, Irving, CA).

Bisulfite Cloning and Sequencing.

After cleaning, the PCR product was phosphorylated in the following reaction: full amount of clean PCR product (30 μ L), 5 μ L of T4 PNK reaction buffer (ThermoFisher Scientific, Waltham, MA), 5 μ L of 10 mM ATP, 1 μ L of T4 PNK (ThermoFisher Scientific, Waltham, MA), brought up to 50 μ L with water. The reaction was incubated at 37°C for 30 min then heat inactivated at 65°C for 20 minutes. The cloning vector, pUC19, was digested with

HincII, a blunt cutter to compliment the blunt PCR products created by PhusionU polymerase. HincII cuts within the lacZ gene of pUC19, allowing for blue/white screening of subsequent transformants. The digestion was carried out in tandem with dephosphorylation of the vector in the following reaction: 1 µg (0.56 pmol) of pUC19, 5 µL of 10 NEB CutSmart Buffer (New England Biolabs [NEB], Ipswich, MA), 1 µL HincII (NEB), 2 µL (2 units) rSAP (NEB), brought up to 50 µL with water. This reaction was incubated at 37°C for 30 minutes then heat inactivated at 65°C for 20 minutes. The phosphorylated PCR product was inserted into the digested vector in the following ligation reaction: 2 µL of pUC19, 2 µL of PCR product (a roughly 5 molar excess of pUC19), 5 µL of Blunt/TA Ligation Master Mix (NEB, Ipswich, MA), brought up to 10 µL with water. The ligation occurred at room temperature for 15 min, after which it was immediately placed on ice.

Competent GC5 *E. coli* cells (Genesee Scientific, San Diego, CA) were thawed on ice for 10 min in 50 µL aliquots. 5 µL of the ligation reaction was added to each thawed cell aliquot and allowed to rest on ice for 30 min. Cells were then heat shocked in a 42°C water bath for 45 sec, then again rested on ice for 5 min. 450 µL of S.O.C. medium (ThermoFisher Scientific, Waltham, MA) was added, and cells were incubated at 37°C for 1 hr with shaking at 250 rpm. 25 µL of this cell culture was plated on Luria Bertani (LB) agar plates containing 100 µg/mL of Ampicillin (Invitrogen, Carlsbad, CA), 20 µg/ml of 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal, ThermoFisher Scientific, Waltham, MA), and 0.1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG, Invitrogen, Carlsbad, CA). Plates were incubated at 37°C for 24 hrs, and 2 white colonies were picked for each primer region for each biological replicate, resulting in two clones for each region for each replicate. Each colony was incubated in 5 mL of LB broth containing 100 µg/mL of Ampicillin at 37°C with shaking at 250 rpm overnight. DNA

was then extracted and purified using the ZR Plasmid Miniprep Classic Kit (Zymo Research, Irvine, CA). Successful insertion of the PCR product into the vector was confirmed by enzymatic digestion in the following reaction: 4 μ L of extracted plasmid, 0.5 μ L of HindIII (NEB, Ipswich, MA), 0.5 μ L of SspI (NEB), 2.5 μ L of CutSmart Buffer (NEB), brought up to 25 μ L with water. The reaction was incubated at 37°C for 15 min then run on a 1.5% agarose gel. Purity and quantity of each clone was assessed using a Nanodrop ND-1000 spectrophotometer. Purified and confirmed plasmids were then sent to GeneWiz (South Plainfield, NJ) for Sanger Sequencing.

The quality of all sequence trace files was examined using GeneStudio Pro software (GeneStudio, Inc) and using the quality control numbers provided by Gene Wiz. Sequences for all clones of all biological replicates of a given primer region were aligned, DNA methylation patterns were determined, and figures were created using BiQ analyzer software (Bock et al., 2005). For a sequence to be included in methylation analysis, it must have contained 80% alignment to the original DNA sequence for that region and have passed all quality control assessments.

RESULTS

Folic acid deficiency does not significantly alter global hepatic methylation.

In order to determine the effect of FA deficiency on liver methylation status, we employed the HELP assay, which uses differential digestion with methylation sensitive and insensitive enzymes followed by promoter microarray analysis (Figure 3.2). The enzymes used in this assay, HpaII and MspI, are isoschizomers that recognize the sequence 5'-CCGG-3', thus allowing for detection of CpG methylation, which is the primary site of methylation in mammals.

Initial analysis showed relatively stable global methylation levels in the control mice compared to the FA deficient mice. Both conditions exhibited a bimodal distribution of methylation levels, representing log₂ HpaII/MspI ratios that were either hypomethylated or hypermethylated relative to the mean of all of the ratios (Figure 3.3). In the control mice 50.001% of the ratios showed hypermethylation and 49.998% of the ratios showed hypomethylation; in the deficient mice 52.216% of ratios demonstrated hypermethylation and 47.738% of ratios demonstrated hypomethylation (Appendix D). Correlations between independent biological replicates in a given condition were $r \geq 0.96$ for all comparisons. The correlation between average of FA biological replicates and average of DEF biological replicate. The correlation between the conditions indicates that overall global methylation levels were similar across both conditions. Further, a microarray dye swap was used in one biological replicate from each dietary condition and compared to the opposite dye pattern of the same biological replicate to ensure the absence of dye bias; correlation between the two dye patterns was $r = 0.98$ and $r = 0.97$ for the control and deficient conditions, respectively (data not shown), demonstrating a lack of dye bias.

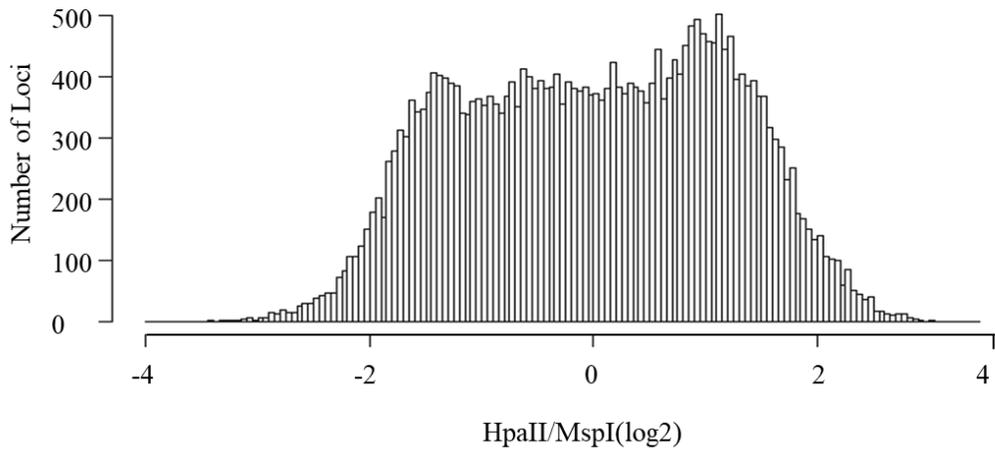
Folic acid deficiency induces significant changes in methylation in a subset of loci.

To determine if FA deficiency induced methylation changes at specific loci, control microarray data was subtracted from FA deficient microarray data, a Welch t-test was performed comparing the control to FA deficient replicates, and results were plotted as a volcano plot demonstrating significant changes in methylation at specific loci ($P \leq 0.05$, log₂ ratio ≥ 1 or ≤ -1 , representing absolute fold change ≥ 2.7) (Figure 3.4a). Further, this same data was plotted as a histogram, with log₂ HpaII/MspI ratios great than 1 representing microarray probes with

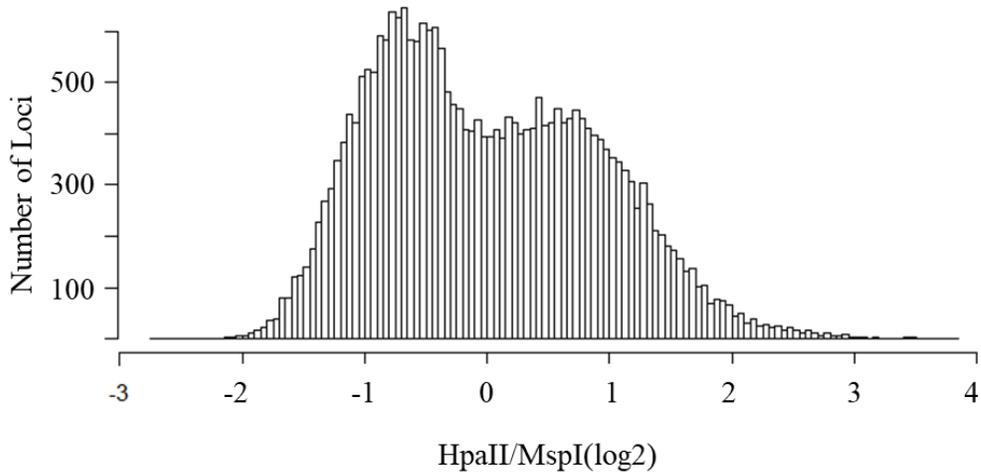
Figure 3.3. Global DNA methylation levels remain constant in folic acid deficiency.

The log₂ HpaII/MspI ratios for the control (A) and deficient (B) condition are plotted as frequency distributions. The data are mean centered at zero for each condition, with negative ratios representing data that is more methylated and positive ratios representing data that is less methylated than the mean of each condition. For control mice 50.001% of the ratios show hypermethylation and 49.998% of the ratios show hypomethylation. For the deficient mice 52.216% of ratios demonstrated hypermethylation and 47.738% of ratios demonstrated hypomethylation. (C) A scatter plot is shown comparing the microarray data of the average HpaII/MspI log₂ ratios for the three folic acid (FA) replicates versus the average HpaII/MspI log₂ ratios for the three deficient (DEF) replicates. The correlation coefficient was calculated using R version 3.1.0.

A



B



C

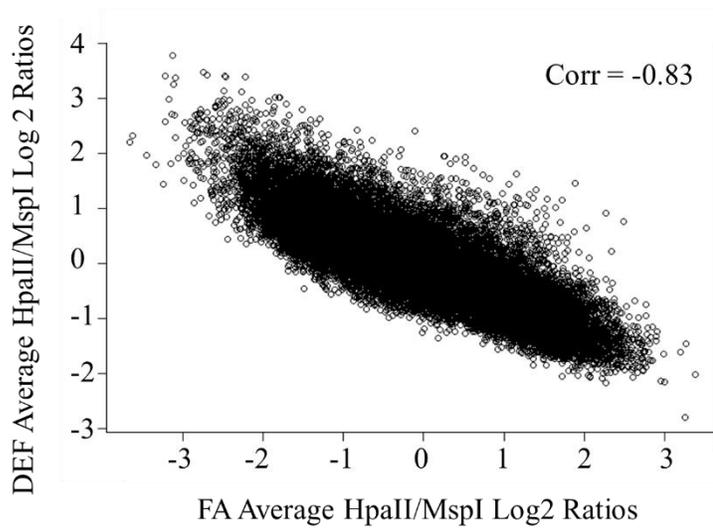
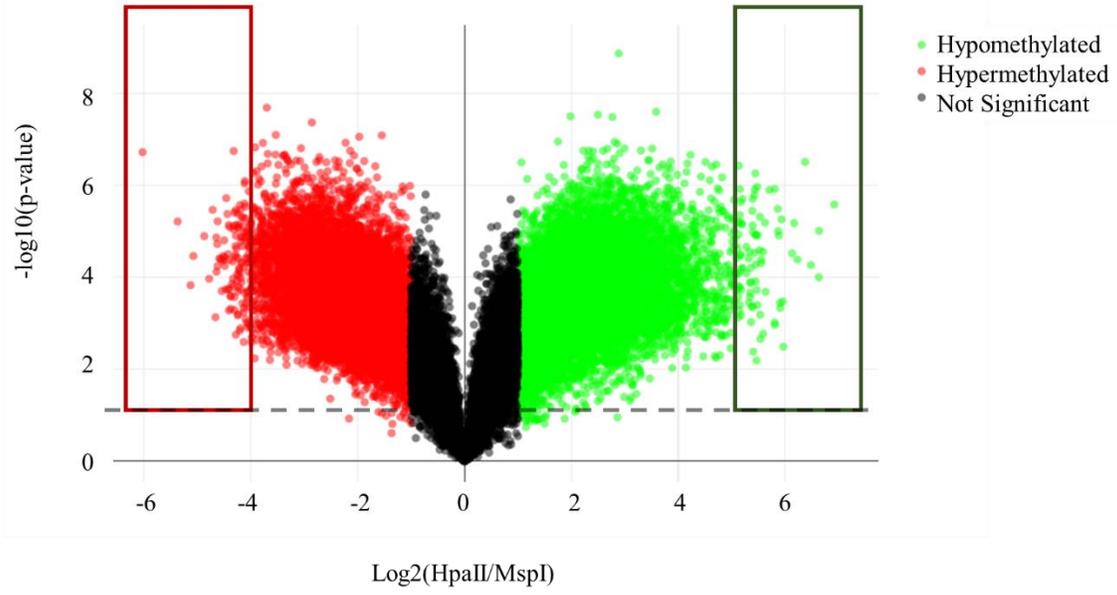
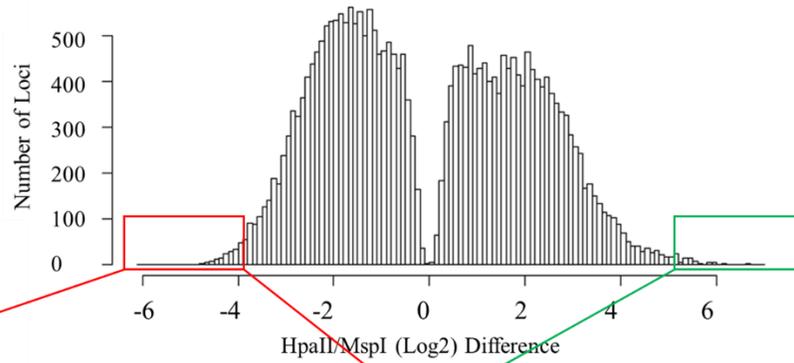
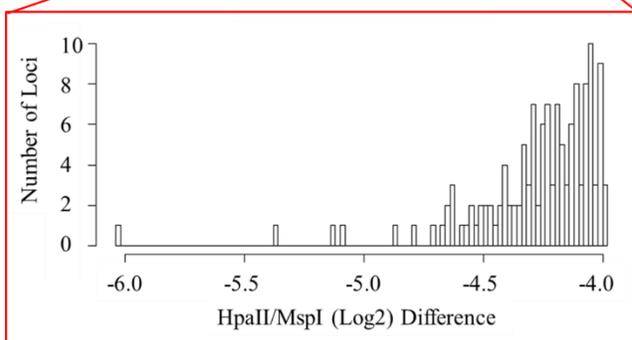
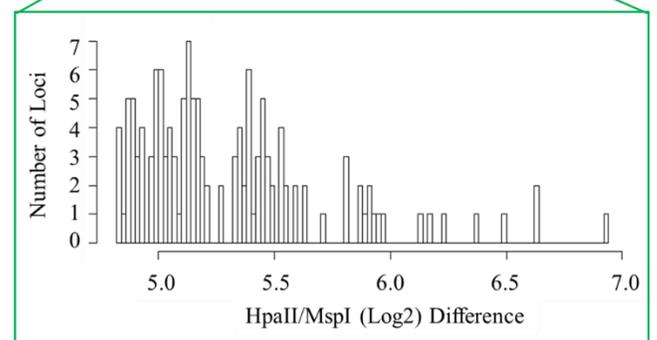


Figure 3.4. Folate deficiency induces significant methylation changes at specific loci.

Control microarray data were subtracted from folic acid deficient microarray data, and results were plotted as a volcano plot as well as a frequency distribution. (A) Colored data points represent loci with significant fold changes in the deficient condition relative to the control condition, noted by absolute values of HpaII/MspI \log_2 ratios ≥ 1 , representing fold changes ≥ 2.7 . Red dots represent hypermethylated loci, and green dots represent hypomethylated loci. The grey dotted line indicates $P \leq 0.05$, representing the threshold for significance for the colored points. Since such a large number of loci were above the threshold for significance, the top 1% most changing were chosen to focus on, noted by the red box (0.5% most hypermethylated) and green box (0.5% most hypomethylated). (B) A frequency distribution shows probes with significant p-values ($P \leq 0.05$). Probes with an HpaII/MspI \log_2 ratio ≥ 1 or ≤ -1 exhibit significant fold change (fold change ≥ 2.7). The frequency of the top 0.5% most hypermethylated (C) and hypomethylated (D) probes are shown.

A**B****C****D**

significant hypomethylation and ratios less than 1 representing probes with significant hypermethylation in the deficient condition ($P < 0.05$, fold change ≥ 2.7) (Figure 3.4b). A total of 20,761 probes showed significant changes, with 9952 probes (36.8% of all probes with $P > 0.05$) exhibiting significant hypomethylation and 10809 probes (40% of all probes with $P > 0.05$) exhibiting significant hypermethylation (fold change ≥ 2.7). Given the large number of probes showing significant change, this study focused on the probes with the greatest absolute change, defined as the top 1% most changing probes (0.5% gaining methylation and 0.5% losing methylation). This group included 135 probes gaining methylation in the FA deficient condition and 135 probes losing methylation in the FA deficient condition compared to the control condition (Figure 3.4 c and d). Of the top 0.5% of hypermethylated loci, the minimum HpaII/MspI log₂ ratio was 4.0, representing a 54.5-fold increase in methylation in the deficient condition. Of the top 0.5% of hypomethylated loci, the minimum HpaII/MspI log₂ ratio was 4.75, representing a 115.6-fold increase in methylation in the deficient condition.

Differentially methylated loci are associated with promoter regions of a number of genes.

To determine genes associated with differential methylation patterns in the FA deficient mice, the top 1% of significant probes were mapped to nearby genes (Supplemental Table 3.3). The majority of the promoter probes fell within 5 kb of the TSS of the gene they were predicted to regulate (60% of hypomethylated probes and 83.2% of hypermethylated probes), and within that proximal probe group, the majority of the probes fell within 1 kb of TSS (77.6% of the hypomethylated and 89.7% of the hypermethylated proximal [within 5 kb] probes) (Figure 3.5). Hypermethylated probes ranged in distance from 8,207,451 bp to 7 bp from the TSS, while the hypomethylated probes ranged from 824,743 bp to 48 bp from the TSS.

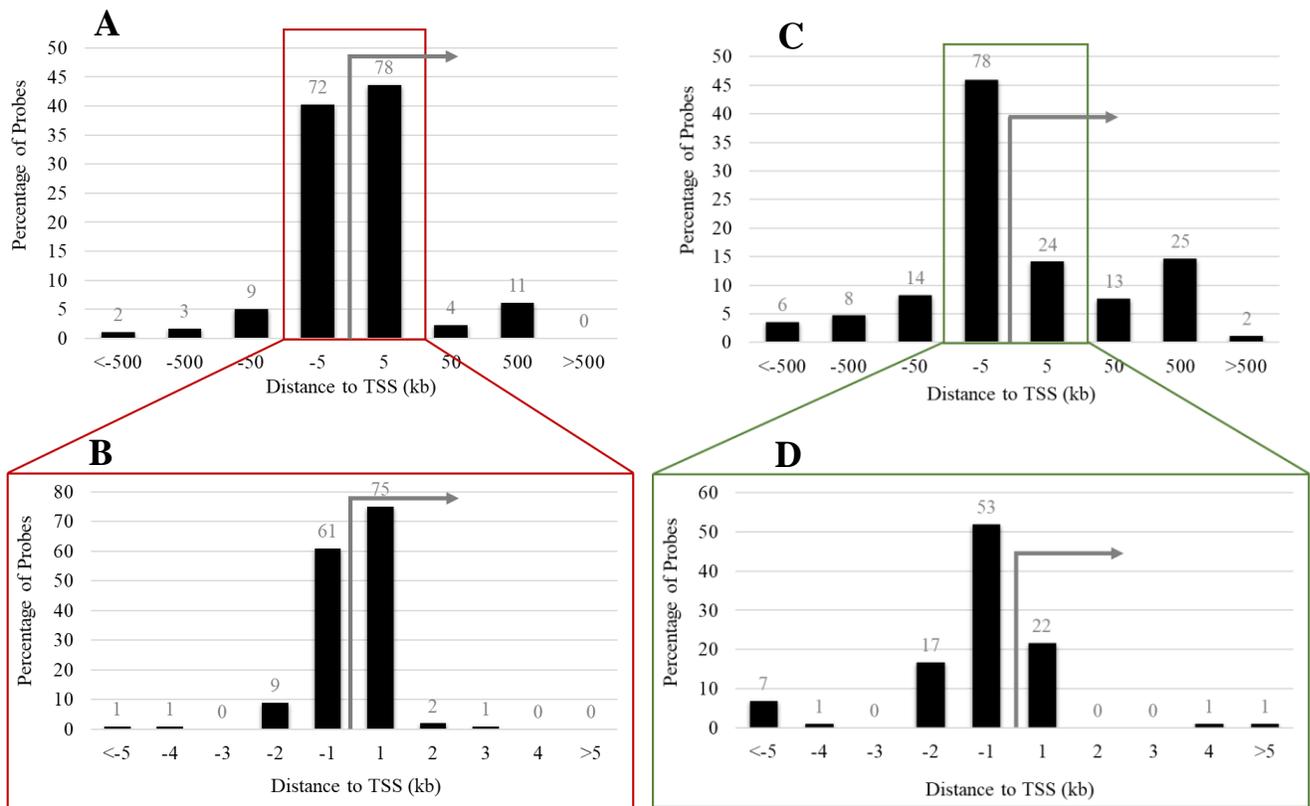


Figure 3.5. Relative gene locations of loci with significant methylation changes.

Differentially methylated loci were mapped to associated genes using GREAT software. The distance from the probe to the TSS (grey arrow) is plotted, with percentage of probes noted on the y-axis and the number of probes noted above each bar. (A) The majority of hypermethylated probes (83.2%) fell within 5 kb of the TSS, and (B) the majority of these probes (89.7%) fell within 1 kb of the TSS. Similarly, (C) the majority of hypomethylated probes (60%) fell within 5 kb of the TSS, and (D) the majority of these probes (77.6%) fell within 1 kb of the TSS. This was expected since the array was a promotor array.

Chronic folic acid deficiency is not associated with significant changes in gene expression.

Since DNA methylation changes in promoter regions are often associated with changes in gene transcription, an RNA microarray was obtained to determine differentially expressed genes in the FA deficient condition compared to the control condition. Initial results indicated increased expression in 160 genes, decreased expression in 192 genes, and no expression changes in the remaining 23,583 genes tested ($P < 0.05$, fold change >2) (Figure 3.6). However, upon further examination, all differentially expressed genes had a high false discovery rate (FDR ≥ 0.823), lending to uncertainty regarding trustworthiness of these results. Although it is likely that a small number of the proposed genes are in fact differentially expressed, the specific genes cannot be determined by microarray analysis alone.

Tumor suppressors with significant methylation changes do not show significant expression changes.

Since the RNA microarray did not give any definitive results for expression changes, expression changes were examined on a subset of genes showing significant methylation changes. Genes were chosen for expression analysis if (1) they were associated with top 1% most changing methylation probes, (2) they have been studied in the liver, and (3) they are known tumor suppressor genes. These criteria were chosen since a significant portion (about one third) of genes associated with differential methylation and potential differential expression in the liver were tumor suppressor genes. Three genes qualified for this analysis: *Brca1*, *Gad1*, and *Tet2*; all of these genes were associated with hypermethylation in the FA deficient condition. Prior to designing primers for qPCR expression analysis, hypermethylation patterns from the HELP

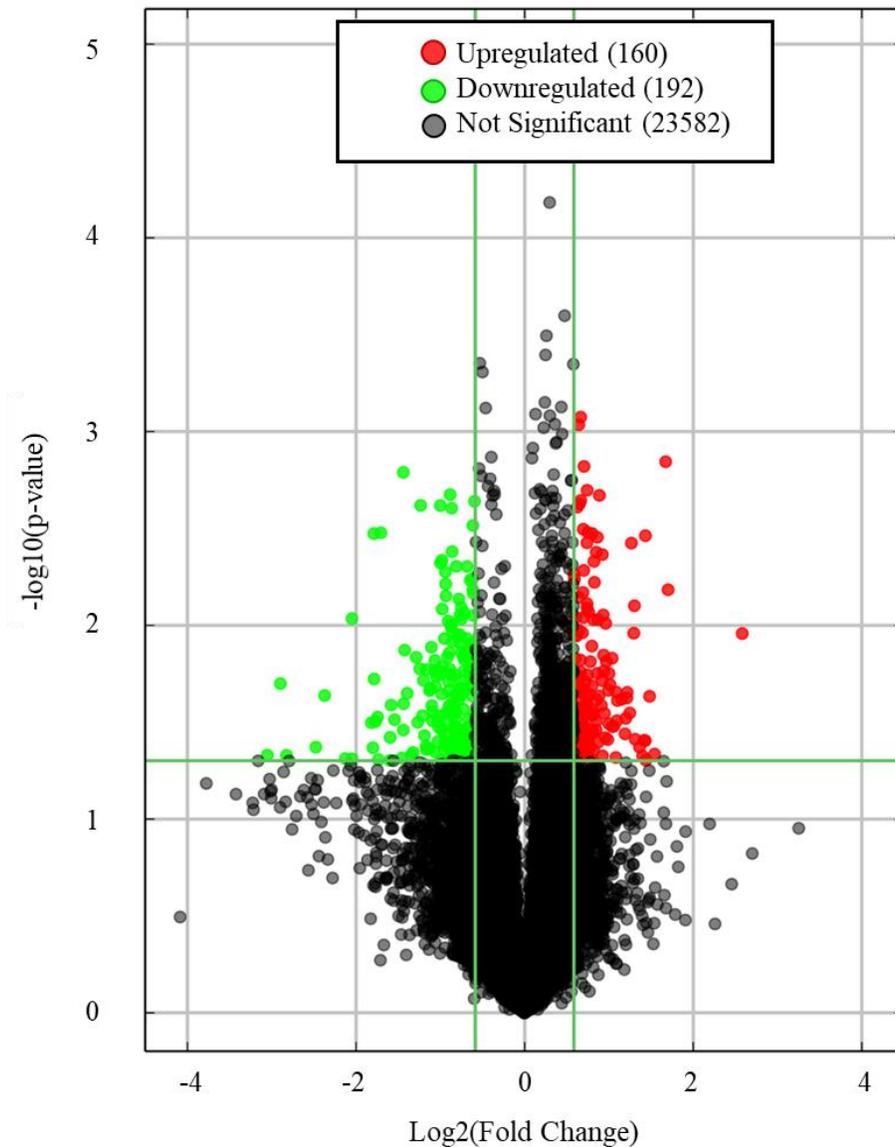


Figure 3.6. Folate deficiency does not induce significant gene expression changes.

RNA microarray analysis showed 352 genes to be differentially expressed in the folic acid deficient compared to control condition. Vertical green lines demonstrate the significance threshold for fold change (fold change ≥ 2), and the horizontal green line demonstrates the threshold of significance for p-value ($P \leq 0.05$). All colored dots are above both thresholds. However, all of these colored dots have a high false discovery rate ($\text{FDR} \geq 0.823$), leading to suspicion regarding the credibility of these results.

microarray were visually confirmed using SignalMap software by Roche NimbleGen (Supplemental Figure 3.1). Each gene was tested by qPCR using N=3 independent biological replicates for each condition. None of the three genes tested showed significant changes in gene expression ($P > 0.11$), confirming results from RNA microarray analysis (Figure 3.7)

Methylation changes in the *Gad1* promoter region are site-specific.

The disparity between expected outcomes of methylation and expression changes may indicate that methylation of specific CpGs in the promoter regions of the genes regulate expression, while other CpGs carry less weight in this regulation. Thus, bisulfite sequencing was used to determine specific CpG methylation patterns in the FA compared to deficient mice. Since time restraints prevented inclusion of bisulfite sequencing analysis of all three tumor suppressors for this dissertation, *Gad1* was chosen for initial observations since this gene exhibited the greatest success in bisulfite PCR amplification (as determined by the greatest number of successful regions amplified by PCR) and since it has been shown to be hypermethylated in liver cancer (Yan et al., 2016). Bisulfite treated *Brcal* and *Tet2* were also successfully amplified in multiple promoter regions; however, sequencing of these regions lies beyond the scope of this dissertation (Supplemental Figure 3.2).

Three segments within the promoter region of *Gad1* (-2000 through +500 bp relative to the TSS) were amplified by bisulfite PCR, cloned into *E. coli*, and sequenced. Two of these segments were upstream of the TSS, and the third region was roughly centered around the TSS (Figure 3.8a). Two clones were sequenced for each individual biological replicate of each dietary condition, and sequences with quality scores below the recommended threshold of the sequencing company were not included in final results. At least one clone of each biological

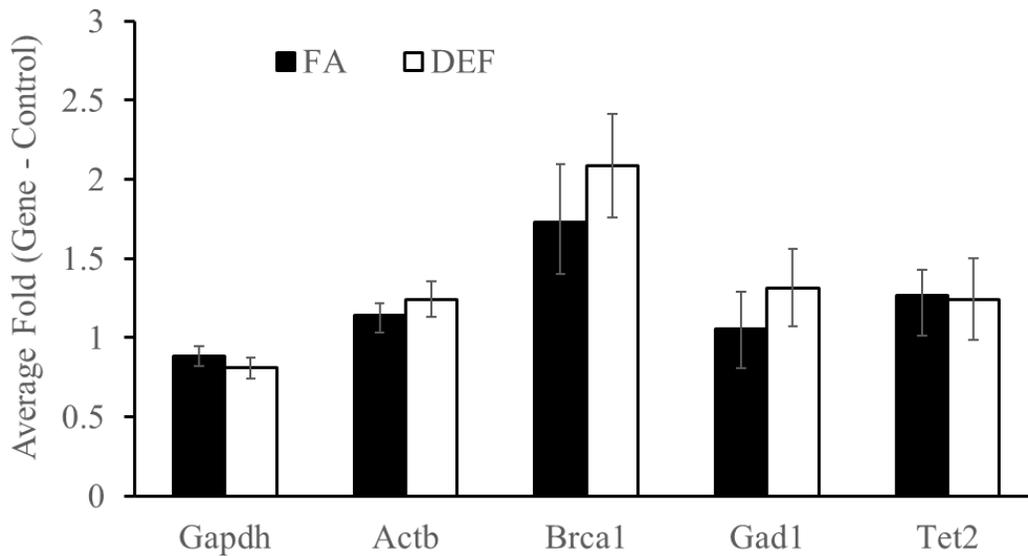


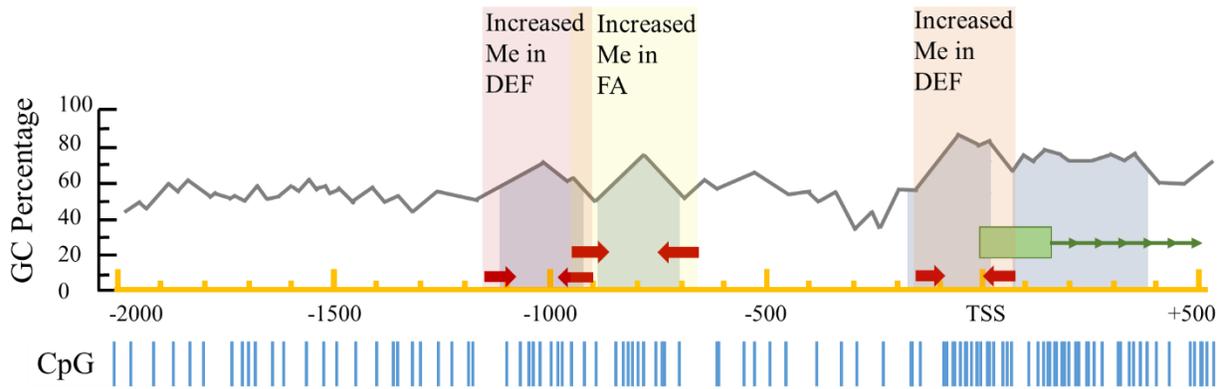
Figure 3.7. qPCR confirms tumor suppressor genes are not differentially expressed.

Three tumor suppressor genes with significant methylation changes in their promoter regions were analyzed by qPCR to determine if those methylation changes resulted in differential expression. None of these three genes exhibited differential expression in the folic acid deficient (DEF) versus control (FA) condition (*Brca1* P= 0.47, *Gad1* P=0.16, *Tet2* P=0.11). Each gene was tested using N=3 independent biological replicates for each condition. *Gapdh* and *ActB* were used as controls, and data was normalized to the average of these two genes. Plotted *Gapdh* and *ActB* levels are relative to the expression of the average of *Gapdh* and *ActB* for one FA replicate to demonstrate the lack of variability between independent replicates and to demonstrate the lack of folate dependence. Error bars represent SEM values calculated from the 3 normalized array values or 3-fold values prepared from three independent biological replicates.

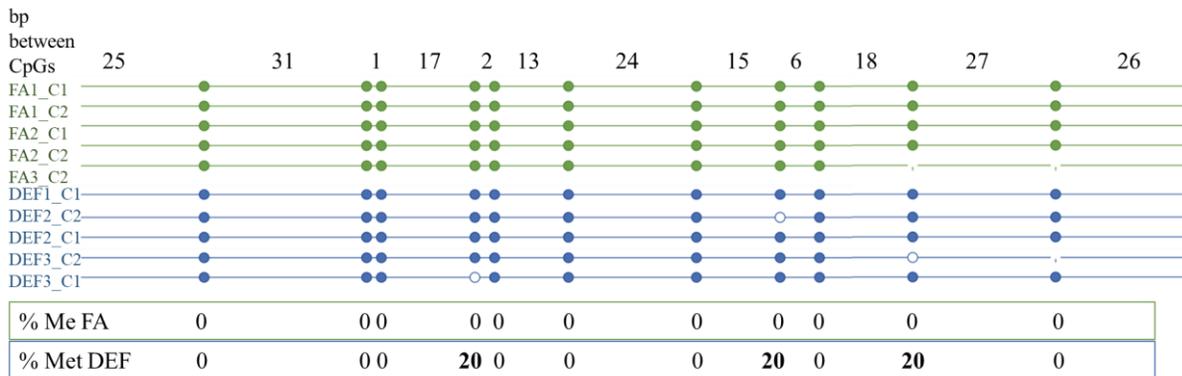
Figure 3.8. Folate-induced differential methylation in *Gad1* promoter regions.

(A) Site-specific methylation changes in the promoter region of *Gad1* were assessed by bisulfite sequencing. Each specific region contained portions of a cytosine-guanine dinucleotide (CpG) island (light grey filling). Locations of CpGs are denoted as blue vertical lines, and CpG percentage of a particular region is noted on the y-axis and by the grey trace line. X-axis locations are relative to the transcription start site (TSS). (B) *Gad1* exon 1 is noted by the green box, and intron 1 is noted by the green arrows. Region 1 (red box) was 1136 – 888 base pairs (bp) upstream of the TSS. Region 2 (yellow box) was 932 – 639 bp upstream of the TSS. Region 3 (orange box) was 144 bp upstream through 84 bp downstream of the TSS. Region 1 exhibited increased methylation (Me) in the deficient (DEF) condition. Methylated CpGs are noted as filled-in circles and unmethylated CpGs are noted as empty circles in the lollipop plot. The percentage of methylated CpGs for the folic acid control (FA) and DEF condition is noted below the lollipop plot. The number of bp between each CpG is noted above the lollipop plot. All FA replicates are noted in green, and all DEF replicates are noted in blue. (C) Region 2 exhibited increased methylation in the FA condition. (D) Region 3 exhibited relatively stable methylation in the FA versus DEF condition upstream of the TSS and increased methylation in the DEF condition downstream of the TSS.

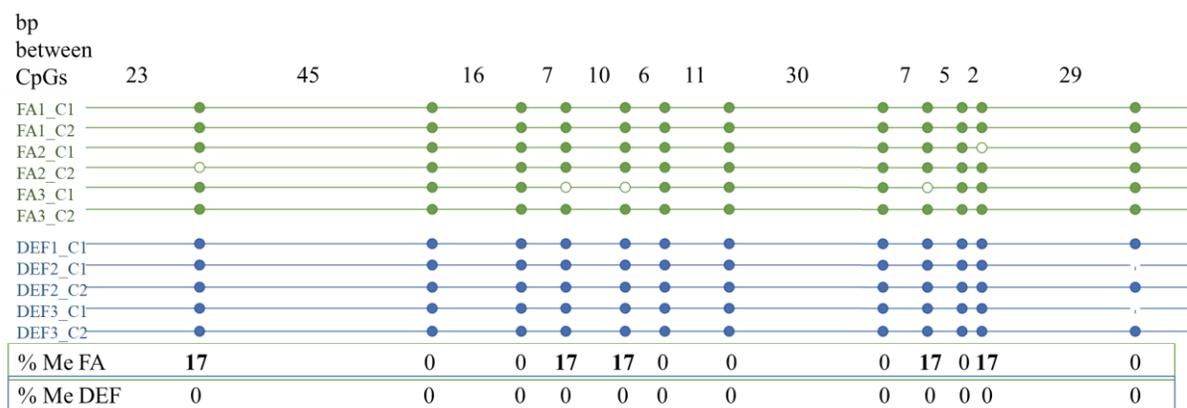
A



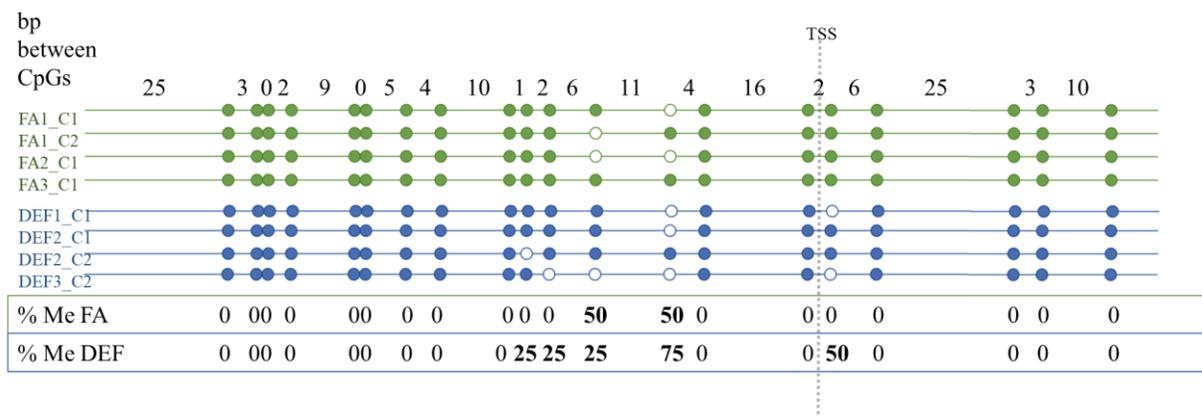
B



C



D



replicate for each condition was included in final results. Primer region 1 demonstrated an increase in methylation of some CpGs in the deficient condition and a complete lack of methylation in the control condition (Figure 3.8b, Supplemental Figure 3.3a). Conversely, primer region 2 demonstrated limited methylation in the control condition and complete lack of methylation in the deficient condition (Figure 3.8c, Supplemental Figure 3.3b). Primer region 3 exhibited some methylation in both the control and deficient conditions with all methylation in the FA condition upstream of the TSS, while methylation in the deficient condition was distributed both upstream and downstream of the TSS (Figure 3.8 d, Supplemental Figure 3.3c).

DISCUSSION

Since the liver is major site of folate storage and metabolism, chronic folate deficiency might be expected to greatly affect the liver (Visentin, Diop-Bove, Zhao, & Goldman, 2014). Further, since folate is a major methyl group donor in the majority of methylation reactions in the body, it is likely that outcomes of folate deficiency are associated with changes in methylation (Mudd & Poole, 1975; Shane, 2008). Indeed, folate deficiencies ranging in length from gestation through early adulthood (4 mo in rodents) induce methylation changes (Kotsopoulos et al., 2008). This study extends the scope of understanding of folate deficiency by following chronic post-natal folate deficiency through the end of adulthood in a mouse model (18 mo) while comparing methylation and expression.

Results of this study indicate that post-weaning dietary folate deficiency through the duration of adult life does not induce significant global methylation changes in the liver, as demonstrated by relative consistency in overall distribution of methylation and high correlation in probe values between control and deficient conditions (Figure 3.3). This is consistent with

more short term (30 week or roughly through puberty) observations that post-weaning folate deficiency does not significantly alter global liver methylation in the adult rat liver (Kotsopoulos et al., 2008). However, this outcome is counterintuitive since folate is known to be the primary methyl group donor in formation of SAM for methylation reactions, so chronic absence of this methyl donor would be expected to result in significant hypomethylation. One possible explanation for this seeming disparity is that secondary methyl group donors such as choline and betaine are able to compensate for folate deficiency and serve as primary methyl group donors (Kim et al., 1994; Niculescu & Zeisel, 2002). Since betaine-homocysteine methyltransferase are predominantly expressed in the liver and kidney, it would be interesting to examine how other tissues lacking this compensatory mechanism react to chronic folate deficiency (Sunden et al., 1997). Perhaps one reason that the liver and kidney are not amongst the most common organs to present with cancer associated with folate deficiency is because they are able to compensate using alternative methyl sources while all other tissue largely lack this ability (Kennedy et al., 2011; Kim, 2018; Locasale, 2013; X. Xu & Chen, 2009; Zhang et al., 2017).

A lack of significant global methylation change does not exclude the possibility of site-specific methylation changes. Indeed, significant changes ($P \leq 0.05$ and fold change > 2.7) were observed in a number of individual loci in the deficient condition compared to the control condition (Figure 3.4). Since significant methylation changes were observed in such a large number of loci (20,761 total), a subset of the top 1% most changing loci, representing 0.5% most hypermethylated and 0.5% most hypomethylated were focused on for further analysis. Each of these loci exhibited at least a 54.5-fold change in methylation in the deficient condition compared to the control condition, demonstrating the great magnitude of impact dietary folate has on the methylation landscape. Again, this is consistent with more short-term rodent studies

demonstrating site-specific changes in DNA methylation due to folate deficiency (Kim et al., 1997).

Given the close connection between DNA methylation and gene expression, we sought to determine the relationship between differentially methylated loci, representing promoter regions, to associated genes they putatively regulate. Therefore, the top 1% most changing loci from the promoter microarray were mapped to genes with which they were most likely associated. These loci ranged in distance from 8,207,451 bp to 7 bp to the TSS of their associated gene, with the majority (60%) of those loci falling within 1 kb of the TSS (Figure 3.5). Although promoter regions are generally located fairly close to the TSS, distal locations have also been well documented (Kulaeva, Nizovtseva, Polikanov, Ulianov, & Studitsky, 2012; Sanyal, Lajoie, Jain, & Dekker, 2012). Extremely distal loci (> 5kb) represented a small portion for genes in this study (<28%), so further analysis was not conducted to verify the relationship between distal loci and associated genes.

Not only were differentially methylated loci mapped to putatively regulated genes, but a separate RNA microarray determined differentially expressed genes in the FA deficient condition. Unexpectedly, this microarray did not show any reliable changes in gene expression. Although 352 genes showed promising p-value and fold change results ($P \leq 0.05$, fold change ≥ 2), all of these genes had a high false discovery rate ($FDR \geq 0.823$), indicating that many of these genes are likely not differentially expressed. It is possible that some of these genes are truly differentially expressed, but it is more likely that many of them are not differentially expressed. This result is surprising since methylation changes in promoter regions are often inversely correlated with gene expression (M. M. Suzuki & Bird, 2008). Yet this study shows that the number of loci with significant methylation changes far outweighs the number of genes showing

significant expression changes (likely ~20% of the 352 genes with $P \leq 0.05$ and fold change ≥ 2 , based on false discovery rate). Since not all CpG methylation directly affects gene expression, it is possible that many of the areas observed to exhibit differential methylation in this study are not significantly influential in gene expression.

Given the uncertain nature of the expression microarray results, the remainder of the study focused on the methylation microarray results. To this end, genes were chosen for further analysis based on the following criteria: (1) association with a top 1% most differentially methylated probes, (2) experimentally associated with the liver, and (3) experimentally determined to be a tumor-suppressor. These criteria were chosen because of the 40 genes exhibiting differential promoter methylation and possible differential expression due to folate deficiency in the liver, 9 have been studied specifically in the liver, and of these 9 genes, 3 are known tumor suppressors. The three genes that met these criteria were *Brcal*, *Gad1*, and *Tet2*. Each of these genes was hypermethylated in the deficient condition compared to the control condition, which is of particular interest given that each of these genes is a known tumor suppressor (Schnepp et al., 2017; Y. P. Xu et al., 2019; Yoshida & Miki, 2004). This result is consistent with other studies showing folate deficiency to be associated with hypermethylation in CpG islands of tumor-suppressor genes (Kim, 2004; Park, Yu, & Shim, 2006). Typically, hypermethylation in a promoter region results in downregulated gene expression; thus, folate deficiency-mediated hypermethylation of these gene may lead to increased risk of tumor growth. Expression changes in each of these genes was assed using qPCR, and all genes were determined to have no differential expression in the control versus deficient conditions (Figure 3.7). Although these results are somewhat surprising based on expectations from the methylation microarray data, they confirm expectations from the expression microarray. Thus, it is it likely

that the differential methylation patterns in these genes are not located in regions critical to gene expression or that additional factors apart from promoter methylation are significantly responsible for regulation of these genes.

Bisulfite sequencing was employed to determine specific methylation patterns in the promoter region (defined to be -2000 bp through +500 bp from the TSS) of the three hypermethylated tumor suppressors. Although multiple domains within the promoter region of each of these genes was successfully amplified by bisulfite PCR, due to time constraints, *Gad1* was chosen to be the single preliminary candidate for bisulfite sequencing analysis. *Gad1* was chosen as the initial gene of interest since it had the greatest success in bisulfite PCR (as indicated by the greatest number of successfully amplified domains, Supplemental Figure 3.2) and because of its association with liver cancer (Yan et al., 2016). Three promoter regions were studied for this gene, two of which were upstream of the TSS and the third of which spanned a short distance on either side of the TSS. Interestingly, the more distal upstream region exhibited increased methylation in the FA deficient condition and complete lack of methylation in the control condition, while the adjacent and more proximal upstream region exhibited the complete opposite methylation pattern (Figure 3.8b and c). Additionally, the portion of the third region that was upstream of the TSS showed no significant difference in methylation between the two dietary conditions, with many of the same CpGs methylated in both conditions (Figure 3.8d). Thus, it is possible that none of these regions contribute significantly to regulation of *Gad1* gene expression. The portion of the third promoter region that was downstream of the TSS demonstrated increased methylation in the deficient condition and no methylation in the control condition. This is consistent with a recent study by Schnepf, *et.al.* which demonstrated differential methylation downstream of the *Gad1* TSS in association with differential expression

of *Gad1* and increased cancer cell metastasis (Schnepp et al., 2017). They demonstrated this differential methylation in the translated region of *Gad1*, while results of this study demonstrate differential methylation in the untranslated region of *Gad1*. Thus, it is possible that methylation of the translated region has a greater impact on *Gad1* gene expression than methylation of the untranslated or upstream promoter regions of this gene.

It is important to note that the nature of the HELP assay used in this study only allowed for detection of changes in CpG methylation in the context of a 4-base palindrome; while our current understanding of methylation shows that the vast majority of mammalian methylation occurs at CpGs, non-CpG methylation has also been reported (Grandjean et al., 2007; Kouidou, Malousi, & Maglaveras, 2006; Ramsahoye et al., 2000; White, Watt, Holt, & Holt, 2002). It is possible that non-CpG methylation was also altered due to folate deficiency. Further, the HELP assay relied on differentially digesting genomic DNA to detect differential methylation patterns. While this method has been proven to be reliable in methylation detection, its application to folate deficiency induced methylation changes may present with additional complications (Oda et al., 2009). Since folate is important not only for methylation reactions but also for nucleotide synthesis, folate deficiency has been associated with single and double strand DNA breaks as well as uracil misincorporation (Han, Hankinson, Zhang, De Vivo, & Hunter, 2004; Kim et al., 1997; Kim et al., 2000). These DNA strand breaks have been shown to accumulate as folate deficiency progresses, leading to the possibility that some of the observed DNA breaks in this study may be due to chronic folate deficiency rather than enzymatic digestion in the HELP assay. All breaks in this study were interpreted as methylation changes since that is the basis of the assay; this interpretation is thought to be fairly trustworthy since a large number of cells were used from 3 independent biological replicates for each condition, and it is unlikely that the same

DNA breaks occurred in the majority of cells in all replicates. However, particular regions of DNA appear to be more prone to DNA breaks, so it is possible that relatively consistent DNA breaks occurred in susceptible regions of all replicates, leading to misinterpretation of the HELP assay data (Kim et al., 1997; Kim et al., 2000).

Further, it is important to note that the nature of folate deficiency may lead to misinterpretation of bisulfite sequencing data. Since folate deficiency can lead to misincorporation of uracil in DNA, and since bisulfite sequencing relies on the conversion of unmethylated cytosines to uracils to detect methylation status, it is possible that some of the results of bisulfite sequencing that were interpreted to be unmethylated cytosines were instead misincorporated uracils. To control for this possibility, these regions of DNA should be sequenced in non-bisulfite treated DNA from each of the biological replicates to account for any nucleotides that were present as uracil prior to bisulfite sequencing. Additionally, these regions should be assayed by differential enzymatic digestions, using enzymes specific to bisulfite treated DNA to assay differences in bisulfite treated samples as well as enzymes specific to the predicted genomic DNA sequences to control for uracil misincorporations.

Finally, it is important to note that folate's role in methylation reactions is a part of the methionine cycle. In this cycle, folate donates a carbon to homocysteine, forming methionine, which is then converted to SAM for methylation reactions (Banerjee & Matthews, 1990). As noted previously, folate is the primary methyl donor for this reaction, but the liver and kidney possess limited ability to use betaine as the carbon donor. Since betaine has not been shown to fully compensate for an absence of folate, folate deficiency typically leads to increased homocysteine levels (Mudd & Poole, 1975). It is possible that this increase in homocysteine, rather than the direct actions of products of folate reactions, is the primary cause for diseases

associated with folate deficiency. Indeed, elevated homocysteine has been associated with all of the major adverse outcomes of folate deficiency including cancers, cardiovascular disease, and cognitive impairment (Hasan et al., 2019).

In summary, the preliminary results of this study indicate that hepatic DNA is differentially methylated at site-specific locations, despite a lack of global methylation changes. Although this differential methylation occurs in promoter regions of genes, gene expression remains largely unchanged. It is possible that the observed methylation changes are not located in portions of the promoter region that are significantly responsible for gene expression, as indicated by *Gad1* bisulfite sequencing data. It is also possible that adverse outcomes of folate deficiency are not directly associated with the actions of molecules to which folate donates carbons, but rather the association lies in a buildup of molecules with which folate would normally react, such as homocysteine. Further, since folate deficiency affects different tissues in different manners, it is possible that the liver possess a unique ability to adapt to folate deficiency, explaining the liver's relative resilience in resisting folate deficiency-induced diseases relative to other tissues such as the pancreas, colon, and breast. Taken together, the results of this study demonstrate a number promoter loci in which differential methylation do not result in differential gene expression, particularly in the three tumor-suppressors examined, perhaps indicating that methylation of these regions may not need to be an area of focus for folate-induced cancer studies. Further the lack of global methylation and expression changes indicate the liver possess mechanisms by which to overcome chronic folate deficiency; identification of these mechanisms may lead to better understanding or treatment options for tissues more susceptible to folate deficiency.

REFERENCES

- Banerjee, R. V., & Matthews, R. G. (1990). Cobalamin-dependent methionine synthase. *FASEB J*, 4(5), 1450-1459. doi: 10.1096/fasebj.4.5.2407589
- Bauerle, M. R., Schwalm, E. L., & Booker, S. J. (2015). Mechanistic diversity of radical S-adenosylmethionine (SAM)-dependent methylation. *J Biol Chem*, 290(7), 3995-4002. doi: 10.1074/jbc.R114.607044
- Blom, H. J., & Smulders, Y. (2011). Overview of homocysteine and folate metabolism. With special references to cardiovascular disease and neural tube defects. *J Inherit Metab Dis*, 34(1), 75-81. doi: 10.1007/s10545-010-9177-4
- Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J., & Lengauer, T. (2005). BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics*, 21(21), 4067-4068. doi: 10.1093/bioinformatics/bti652
- Choumenkovitch, S. F., Selhub, J., Wilson, P. W., Rader, J. I., Rosenberg, I. H., & Jacques, P. F. (2002). Folic acid intake from fortification in United States exceeds predictions. *J Nutr*, 132(9), 2792-2798. doi: 10.1093/jn/132.9.2792
- Christman, J. K., Sheikhnejad, G., Dizik, M., Abileah, S., & Wainfan, E. (1993). Reversibility of changes in nucleic acid methylation and gene expression induced in rat liver by severe dietary methyl deficiency. *Carcinogenesis*, 14(4), 551-557. doi: 10.1093/carcin/14.4.551
- Clifford, A. J., Heid, M. K., Muller, H. G., & Bills, N. D. (1990). Tissue distribution and prediction of total body folate of rats. *J Nutr*, 120(12), 1633-1639. doi: 10.1093/jn/120.12.1633
- Devlin, A. M., Ling, E. H., Peerson, J. M., Fernando, S., Clarke, R., Smith, A. D., & Halsted, C. H. (2000). Glutamate carboxypeptidase II: a polymorphism associated with lower levels of serum folate and hyperhomocysteinemia. *Hum Mol Genet*, 9(19), 2837-2844.
- Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet*, 8(4), 286-298. doi: 10.1038/nrg2005
- Gloria, L., Cravo, M., Camilo, M. E., Resende, M., Cardoso, J. N., Oliveira, A. G., . . . Mira, F. C. (1997). Nutritional deficiencies in chronic alcoholics: relation to dietary intake and alcohol consumption. *Am J Gastroenterol*, 92(3), 485-489.
- Grandjean, V., Yaman, R., Cuzin, F., & Rassoulzadegan, M. (2007). Inheritance of an epigenetic mark: the CpG DNA methyltransferase 1 is required for de novo establishment of a complex pattern of non-CpG methylation. *PLoS One*, 2(11), e1136. doi: 10.1371/journal.pone.0001136
- Han, J., Hankinson, S. E., Zhang, S. M., De Vivo, I., & Hunter, D. J. (2004). Interaction between genetic variations in DNA repair genes and plasma folate on breast cancer risk. *Cancer Epidemiol Biomarkers Prev*, 13(4), 520-524.
- Hasan, T., Arora, R., Bansal, A. K., Bhattacharya, R., Sharma, G. S., & Singh, L. R. (2019). Disturbed homocysteine metabolism is associated with cancer. *Exp Mol Med*, 51(2), 1-13. doi: 10.1038/s12276-019-0216-4

- Jones, P. A., & Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet*, 3(6), 415-428. doi: 10.1038/nrg816
- Kennedy, D. A., Stern, S. J., Moretti, M., Matok, I., Sarkar, M., Nickel, C., & Koren, G. (2011). Folate intake and the risk of colorectal cancer: a systematic review and meta-analysis. *Cancer Epidemiol*, 35(1), 2-10. doi: 10.1016/j.canep.2010.11.004
- Kim, Y. I. (2004). Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? *Cancer Epidemiol Biomarkers Prev*, 13(4), 511-519.
- Kim, Y. I. (2005). Nutritional epigenetics: impact of folate deficiency on DNA methylation and colon cancer susceptibility. *J Nutr*, 135(11), 2703-2709. doi: 10.1093/jn/135.11.2703
- Kim, Y. I. (2018). Folate and cancer: a tale of Dr. Jekyll and Mr. Hyde? *Am J Clin Nutr*, 107(2), 139-142. doi: 10.1093/ajcn/nqx076
- Kim, Y. I., Miller, J. W., da Costa, K. A., Nadeau, M., Smith, D., Selhub, J., . . . Mason, J. B. (1994). Severe folate deficiency causes secondary depletion of choline and phosphocholine in rat liver. *J Nutr*, 124(11), 2197-2203. doi: 10.1093/jn/124.11.2197
- Kim, Y. I., Pogribny, I. P., Basnakian, A. G., Miller, J. W., Selhub, J., James, S. J., & Mason, J. B. (1997). Folate deficiency in rats induces DNA strand breaks and hypomethylation within the p53 tumor suppressor gene. *Am J Clin Nutr*, 65(1), 46-52. doi: 10.1093/ajcn/65.1.46
- Kim, Y. I., Shirwadkar, S., Choi, S. W., Puchyr, M., Wang, Y., & Mason, J. B. (2000). Effects of dietary folate on DNA strand breaks within mutation-prone exons of the p53 gene in rat colon. *Gastroenterology*, 119(1), 151-161. doi: 10.1053/gast.2000.8518
- Kotsopoulos, J., Sohn, K. J., & Kim, Y. I. (2008). Postweaning dietary folate deficiency provided through childhood to puberty permanently increases genomic DNA methylation in adult rat liver. *J Nutr*, 138(4), 703-709. doi: 10.1093/jn/138.4.703
- Kouidou, S., Malousi, A., & Maglaveras, N. (2006). Methylation and repeats in silent and nonsense mutations of p53. *Mutat Res*, 599(1-2), 167-177. doi: 10.1016/j.mrfmmm.2006.03.002
- Kremer, J. M., Galivan, J., Streckfuss, A., & Kamen, B. (1986). Methotrexate metabolism analysis in blood and liver of rheumatoid arthritis patients. Association with hepatic folate deficiency and formation of polyglutamates. *Arthritis Rheum*, 29(7), 832-835.
- Kulaeva, O. I., Nizovtseva, E. V., Polikanov, Y. S., Ulianov, S. V., & Studitsky, V. M. (2012). Distant activation of transcription: mechanisms of enhancer action. *Mol Cell Biol*, 32(24), 4892-4897. doi: 10.1128/MCB.01127-12
- Li, L. C., & Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics*, 18(11), 1427-1431. doi: 10.1093/bioinformatics/18.11.1427
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods*, 25(4), 402-408. doi: 10.1006/meth.2001.1262

- Locasale, J. W. (2013). Serine, glycine and one-carbon units: cancer metabolism in full circle. *Nat Rev Cancer*, *13*(8), 572-583. doi: 10.1038/nrc3557
- Marasas, W. F., Riley, R. T., Hendricks, K. A., Stevens, V. L., Sadler, T. W., Gelineau-van Waes, J., . . . Merrill, A. H., Jr. (2004). Fumonisin disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: a potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. *J Nutr*, *134*(4), 711-716. doi: 10.1093/jn/134.4.711
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., . . . Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*, *28*(5), 495-501. doi: 10.1038/nbt.1630
- Morrell, M. J. (2002). Folic Acid and Epilepsy. *Epilepsy Curr*, *2*(2), 31-34. doi: 10.1046/j.1535-7597.2002.00017.x
- Mudd, S. H., & Poole, J. R. (1975). Labile methyl balances for normal humans on various dietary regimens. *Metabolism*, *24*(6), 721-735. doi: 10.1016/0026-0495(75)90040-2
- Nicolas Veland, Taiping Chen. (2017). Mechanisms of DNA Methylation and Demethylation During Mammalian Development. In T. O. Tollefsbol (Ed.), *Handbook of Epigenetics* (Second ed., pp. 11-24).
- Niculescu, M. D., & Zeisel, S. H. (2002). Diet, methyl donors and DNA methylation: interactions between dietary folate, methionine and choline. *J Nutr*, *132*(8 Suppl), 2333S-2335S. doi: 10.1093/jn/132.8.2333S
- Oda, M., Glass, J. L., Thompson, R. F., Mo, Y., Olivier, E. N., Figueroa, M. E., . . . Grealley, J. M. (2009). High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. *Nucleic Acids Res*, *37*(12), 3829-3839. doi: 10.1093/nar/gkp260
- Ong, T. P., Moreno, F. S., & Ross, S. A. (2011). Targeting the epigenome with bioactive food components for cancer prevention. *J Nutrigenet Nutrigenomics*, *4*(5), 275-292. doi: 10.1159/000334585
- Ortega, R. M., Lopez-Sobaler, A. M., Gonzalez-Gross, M. M., Redondo, R. M., Marzana, I., Zamora, M. J., & Andres, P. (1994). Influence of smoking on folate intake and blood folate concentrations in a group of elderly Spanish men. *J Am Coll Nutr*, *13*(1), 68-72.
- Ortega, R. M., Requejo, A. M., Lopez-Sobaler, A. M., Navia, B., Mena, M. C., Basabe, B., & Andres, P. (2004). Smoking and passive smoking as conditioners of folate status in young women. *J Am Coll Nutr*, *23*(4), 365-371. doi: 10.1080/07315724.2004.10719380
- Park, H. J., Yu, E., & Shim, Y. H. (2006). DNA methyltransferase expression and DNA hypermethylation in human hepatocellular carcinoma. *Cancer Lett*, *233*(2), 271-278. doi: 10.1016/j.canlet.2005.03.017
- Pogribny, I. P., James, S. J., Jernigan, S., & Pogribna, M. (2004). Genomic hypomethylation is specific for preneoplastic liver in folate/methyl deficient rats and does not occur in non-target tissues. *Mutat Res*, *548*(1-2), 53-59. doi: 10.1016/j.mrfmmm.2003.12.014

- Quinlivan, E. P., Davis, S. R., Shelnutt, K. P., Henderson, G. N., Ghandour, H., Shane, B., . . . Gregory, J. F., 3rd. (2005). Methylene tetrahydrofolate reductase 677C->T polymorphism and folate status affect one-carbon incorporation into human DNA deoxynucleosides. *J Nutr*, 135(3), 389-396. doi: 10.1093/jn/135.3.389
- Ramsahoye, B. H., Biniszkiwicz, D., Lyko, F., Clark, V., Bird, A. P., & Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A*, 97(10), 5237-5242. doi: 10.1073/pnas.97.10.5237
- Russell, R. M., Rosenberg, I. H., Wilson, P. D., Iber, F. L., Oaks, E. B., Giovetti, A. C., . . . Press, A. W. (1983). Increased urinary excretion and prolonged turnover time of folic acid during ethanol ingestion. *Am J Clin Nutr*, 38(1), 64-70. doi: 10.1093/ajcn/38.1.64
- Sanyal, A., Lajoie, B. R., Jain, G., & Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature*, 489(7414), 109-113. doi: 10.1038/nature11279
- Schnepf, P. M., Lee, D. D., Guldner, I. H., O'Tighearnigh, T. K., Howe, E. N., Palakurthi, B., . . . Zhang, S. (2017). GAD1 Upregulation Programs Aggressive Features of Cancer Cell Metabolism in the Brain Metastatic Microenvironment. *Cancer Res*, 77(11), 2844-2856. doi: 10.1158/0008-5472.CAN-16-2289
- Shane, B. (2008). Folate and vitamin B12 metabolism: overview and interaction with riboflavin, vitamin B6, and polymorphisms. *Food Nutr Bull*, 29(2 Suppl), S5-16; discussion S17-19. doi: 10.1177/15648265080292S103
- Shi, Y., & Shang, J. (2016). Long Noncoding RNA Expression Profiling Using Arraystar LncRNA Microarrays. *Methods Mol Biol*, 1402, 43-61. doi: 10.1007/978-1-4939-3378-5_6
- Smith, A. D., Smith, S. M., de Jager, C. A., Whitbread, P., Johnston, C., Agacinski, G., . . . Refsum, H. (2010). Homocysteine-lowering by B vitamins slows the rate of accelerated brain atrophy in mild cognitive impairment: a randomized controlled trial. *PLoS One*, 5(9), e12244. doi: 10.1371/journal.pone.0012244
- Sunden, S. L., Renduchintala, M. S., Park, E. I., Miklasz, S. D., & Garrow, T. A. (1997). Betaine-homocysteine methyltransferase expression in porcine and human tissues and chromosomal localization of the human gene. *Arch Biochem Biophys*, 345(1), 171-174. doi: 10.1006/abbi.1997.0246
- Suzuki, M., & Grealley, J. M. (2010). DNA methylation profiling using HpaII tiny fragment enrichment by ligation-mediated PCR (HELP). *Methods*, 52(3), 218-222. doi: 10.1016/j.ymeth.2010.04.013
- Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet*, 9(6), 465-476. doi: 10.1038/nrg2341
- Taher, Noor. (2013). Effects of Beta Amyloid on the DNA Methylation Status of an in Vitro Model of Alzheimer's Disease. *Liberty University Digital Commons*.
- Visentin, M., Diop-Bove, N., Zhao, R., & Goldman, I. D. (2014). The intestinal absorption of folates. *Annu Rev Physiol*, 76, 251-274. doi: 10.1146/annurev-physiol-020911-153251

- White, G. P., Watt, P. M., Holt, B. J., & Holt, P. G. (2002). Differential patterns of methylation of the IFN-gamma promoter at CpG and non-CpG sites underlie differences in IFN-gamma gene expression between human neonatal and adult CD45RO- T cells. *J Immunol*, *168*(6), 2820-2827. doi: 10.4049/jimmunol.168.6.2820
- Xu, X., & Chen, J. (2009). One-carbon metabolism and breast cancer: an epidemiological perspective. *J Genet Genomics*, *36*(4), 203-214. doi: 10.1016/S1673-8527(08)60108-3
- Xu, Y. P., Lv, L., Liu, Y., Smith, M. D., Li, W. C., Tan, X. M., . . . Xiong, Y. (2019). Tumor suppressor TET2 promotes cancer immunity and immunotherapy efficacy. *J Clin Invest*, *130*, 4316-4331. doi: 10.1172/JCI129317
- Yan, H., Tang, G., Wang, H., Hao, L., He, T., Sun, X., . . . Sun, S. (2016). DNA methylation reactivates GAD1 expression in cancer by preventing CTCF-mediated polycomb repressive complex 2 recruitment. *Oncogene*, *35*(30), 4020. doi: 10.1038/onc.2016.28
- Yoshida, K., & Miki, Y. (2004). Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. *Cancer Sci*, *95*(11), 866-871. doi: 10.1111/j.1349-7006.2004.tb02195.x
- Zapisek, W. F., Cronin, G. M., Lyn-Cook, B. D., & Poirier, L. A. (1992). The onset of oncogene hypomethylation in the livers of rats fed methyl-deficient, amino acid-defined diets. *Carcinogenesis*, *13*(10), 1869-1872. doi: 10.1093/carcin/13.10.1869
- Zhang, Z., He, Y., Tu, X., Huang, S., Chen, Z., Wang, L., & Song, J. (2017). Mapping of DNA Hypermethylation and Hypomethylation induced by Folate Deficiency in Sporadic Colorectal Cancer and Clinical Implication Analysis of Hypermethylation Pattern in CBS Promoter. *Clin Lab*, *63*(4), 733-748. doi: 10.7754/Clin.Lab.2016.161018

SUPPLEMENTAL INFORMATION

| Gene Name | Forward Primer | Reverse Primer |
|------------------|-----------------------|-----------------------|
| Brca1 | GAGAAACTCCCCCTCTCAGG | AGATCCCAGGTATGGGGTTC |
| Gad1 | CCTGGAACCCTCACAAGATG | CTGGAAGAGGTAGCCTGCAC |
| Tet2 | TTGTTGTCAGGGTGAGAATCC | CCTTGGTTTCTTGCTTCTGG |

Supplemental Table 3.1. qPCR primers for expression analysis.

UCSC was used to obtain genomic DNA and determine regions where primers could be designed across exon-exon borders. Primer3 was then used to design primers around the probe region using the following parameters where possible: GC clamps, T_m 60+2, and GC content 50+10%. All primers are listed from 5' to 3'.

| Gene | Location | Annealing Temp | Forward Primer | Reverse Primer |
|----------------|-------------------------------|----------------|----------------------------|----------------------------|
| Brca1 region 1 | chr11:101,414,041-101,414,402 | 58.4°C | TGGTTATTGAAAAGATTGGTTG | AAACTTCTATCCCTCCCTTACC |
| Brca1 region 2 | chr11:101,413,421-101,413,641 | 58.3°C | TGTTGGAATTGGTAGTTTTTTT | ATTTCTCTTAAACAACCCCTTC |
| Gad1 region 1 | chr2:70,399,085-70,399,333 | 59.9°C | TTTGTTTTGTTTTTAAAGGATGAGAA | AAATAAAATCCCCCACCATAAAA |
| Gad1 region 2 | chr2:70,399,289-70,399,582 | 59.9°C | TTTTATGGTGGGGGATTTTATTT | AATCTATTTCCCTTTCTCTAAACCCT |
| Gad1 region 3 | chr2:70,400,077-70,400,305 | 56.2°C | TTAGATATTTGTAAAGGAGTTTTAG | CCAAAACACTTCCTTACTTACAC |
| Tet2 region 1 | chr3:133,208,003-133,208,262 | 59°C | GTTTTGAATTTTAGGGAAAATTGAG | ACCCACCTTCTATTTACTTTAACCAC |
| Tet2 region 2 | chr3:133,207,004-133,207,310 | 54.3°C | TTAAATTGTTTTATGAATATTGATG | ATTTATAAATAACATAACAATCCCC |

Supplemental Table 3.2. Bisulfite sequencing primers.

UCSC was used to obtain genomic DNA. MethPrimer was then used to identify CpG islands, which represented significant targets for changes in DNA methylation. Since a uracil-tolerant polymerase with special annealing conditions was used for PCR, annealing conditions specific to this polymerase were determined using a ThermoFisher annealing calculator.

Supplemental Table 3.3a. Genes associated with hypermethylated probes.

| Gene Name | Distance to TSS | Chromosome | Start | End |
|------------------|------------------------|-------------------|--------------|------------|
| 1110008F13Rik | 746 | chr2 | 156689586 | 156689633 |
| 1810043H04Rik | -204 | chr11 | 119960016 | 119960060 |
| 2610318N02Rik | 272 | chr16 | 17124905 | 17124949 |
| Abhd6 | 75 | chr14 | 8835469 | 8835513 |
| Abra | 80041 | chr15 | 41621200 | 41621249 |
| Actr1b | 546 | chr1 | 36766285 | 36766329 |
| Adcy8 | 472 | chr15 | 64753364 | 64753408 |
| Adipor2 | -31436 | chr6 | 119398915 | 119398959 |
| Alg8 | 214 | chr7 | 104520308 | 104520352 |
| Arid4b | -380 | chr13 | 14155657 | 14155701 |
| Ash2l | -245 | chr8 | 26951472 | 26951520 |
| Atm | 260 | chr9 | 53344563 | 53344607 |
| BC030500 | 659 | chr8 | 61391188 | 61391233 |
| BC053749 | 213 | chr7 | 31337056 | 31337100 |
| BC089491 | 516 | chr7 | 29075667 | 29075711 |
| Brcal | -849 | chr11 | 101414020 | 101414064 |
| Brd2 | 967 | chr17 | 34257703 | 34257747 |
| Ccdc134 | 439 | chr15 | 81958769 | 81958813 |
| Ccdc151 | 77 | chr9 | 21806979 | 21807023 |
| Cd9 | 98 | chr6 | 125444689 | 125444733 |
| Chmp2b | -343 | chr16 | 65563263 | 65563307 |
| Cog2 | 324 | chr8 | 127044969 | 127045013 |
| Copa | -71 | chr1 | 174012566 | 174012611 |
| Cryzl1 | 457 | chr16 | 91728740 | 91728785 |
| Cxxc4 | -692899 | chr3 | 133206862 | 133206908 |
| Cyb5d2 | 208 | chr11 | 72609128 | 72609172 |
| Dedd2 | -345 | chr7 | 26006035 | 26006079 |
| Dlgap1 | 425828 | chr17 | 71199940 | 71199984 |
| Dnaic2 | 139 | chr11 | 114588843 | 114588887 |
| Eaf1 | -349 | chr14 | 32307894 | 32307938 |
| Efcab11 | 451274 | chr12 | 100670356 | 100670400 |
| Egr1 | 1713 | chr18 | 35021168 | 35021212 |
| Enthd2 | 7 | chr11 | 119960016 | 119960060 |
| Ephb1 | 34 | chr9 | 102256967 | 102257011 |
| Eral1 | 261 | chr11 | 77893602 | 77893646 |
| Etf1 | 70467 | chr18 | 35021168 | 35021212 |
| Evc2 | -8207451 | chr5 | 29522265 | 29522309 |
| Fam173a | -246 | chr17 | 25929453 | 25929497 |

| | | | | |
|------------|--------|-------|-----------|-----------|
| Fam21 | -117 | chr6 | 116157912 | 116157956 |
| Fam72a | -1180 | chr1 | 133423278 | 133423322 |
| Fam78a | 544 | chr2 | 31938737 | 31938781 |
| Fbxl14 | -30749 | chr6 | 119398915 | 119398959 |
| Fbxl6 | 664 | chr15 | 76368490 | 76368534 |
| Fhdc1 | 174773 | chr3 | 84108618 | 84108662 |
| Foxn3 | -38371 | chr12 | 100670356 | 100670400 |
| Frmd5 | -109 | chr2 | 121632910 | 121632954 |
| Ftsj3 | 341 | chr11 | 106117030 | 106117074 |
| Gabarapl1 | 144 | chr6 | 129483340 | 129483384 |
| Gad1 | -732 | chr2 | 70399345 | 70399389 |
| Galnt16 | -787 | chr8 | 61391188 | 61391233 |
| Ggps1 | -11 | chr13 | 14155657 | 14155701 |
| Gin1 | -490 | chr1 | 99666237 | 99666281 |
| Gm12830 | 231 | chr4 | 114494534 | 114494582 |
| Gm8765 | 402242 | chr13 | 51195901 | 51195945 |
| Gm9900 | -237 | chr11 | 58144056 | 58144100 |
| Gna14 | 202 | chr19 | 16510337 | 16510381 |
| Gnai2 | 229 | chr9 | 107537422 | 107537466 |
| Grin2a | 2223 | chr16 | 9992764 | 9992817 |
| Gtf2f2 | 272 | chr14 | 76410378 | 76410422 |
| Gtpbp10 | 92 | chr5 | 5559424 | 5559468 |
| Gylt11b | -4068 | chr2 | 92215258 | 92215305 |
| Hadha | -599 | chr5 | 30482279 | 30482323 |
| Hadhb | 513 | chr5 | 30482279 | 30482323 |
| Hdhd3 | -71 | chr4 | 62163359 | 62163403 |
| Hist1h2ab | 262 | chr13 | 23843197 | 23843241 |
| Hist1h3b | -917 | chr13 | 23843197 | 23843241 |
| Hist2h2aa2 | 670 | chr3 | 96043524 | 96043568 |
| Hist2h3c2 | -496 | chr3 | 96043919 | 96043963 |
| Hmga2 | -1392 | chr10 | 119914893 | 119914941 |
| Hnrnpul1 | -299 | chr7 | 26540026 | 26540070 |
| Hspb6 | -1209 | chr7 | 31337056 | 31337100 |
| Htr1b | 667 | chr9 | 81525533 | 81525577 |
| Ift122 | -355 | chr6 | 115803171 | 115803215 |
| Ift20 | -584 | chr11 | 78349257 | 78349301 |
| Igsf9b | 209 | chr9 | 27107002 | 27107046 |
| Ik | 110 | chr18 | 36904398 | 36904442 |
| Itsn1 | -835 | chr16 | 91728740 | 91728785 |
| Kbtbd2 | -428 | chr6 | 56748136 | 56748180 |
| Kcnn3 | 20 | chr3 | 89324084 | 89324128 |
| Kdr | -377 | chr5 | 76374838 | 76374882 |
| Khsrp | -287 | chr17 | 57171195 | 57171239 |
| Lbp | 54759 | chr2 | 158186966 | 158187010 |

| | | | | |
|-------------|---------|-------|-----------|-----------|
| Lfng | -493 | chr5 | 141082778 | 141082826 |
| Lpl | 157 | chr8 | 71404525 | 71404569 |
| Lrrc10b | -747 | chr19 | 10532662 | 10532706 |
| Map3k12 | -1105 | chr15 | 102348350 | 102348394 |
| Mark2 | 46037 | chr19 | 7369864 | 7369908 |
| Mbd4 | 196 | chr6 | 115803171 | 115803215 |
| Med31 | 652 | chr11 | 72028420 | 72028464 |
| Mettl6 | 300 | chr14 | 32307894 | 32307938 |
| Mfap1b | 110 | chr2 | 121299671 | 121299715 |
| Mkks | -851 | chr2 | 136717954 | 136717998 |
| Mphosph8 | 169 | chr14 | 57287232 | 57287276 |
| Mycbpap | 87 | chr11 | 94382947 | 94382991 |
| Myt1l | 6907 | chr12 | 30220134 | 30220178 |
| Nab2 | -624 | chr10 | 127104364 | 127104416 |
| Nbr1 | 579 | chr11 | 101414020 | 101414064 |
| Ncbp2 | 347 | chr16 | 31948924 | 31948968 |
| Ncstn | 337 | chr1 | 174012566 | 174012611 |
| Ndufa11 | -592 | chr17 | 56856571 | 56856615 |
| Ndufa2 | -218 | chr18 | 36904398 | 36904442 |
| Nfyb | 48 | chr10 | 82226819 | 82226863 |
| Nipal1 | 254 | chr5 | 73039265 | 73039313 |
| Npas2 | -1205 | chr1 | 39249830 | 39249874 |
| Npat | -567 | chr9 | 53344563 | 53344607 |
| Nrd1 | -249 | chr4 | 108672989 | 108673033 |
| Nrf1 | 127328 | chr6 | 30125336 | 30125381 |
| Nrp1 | -1058 | chr8 | 130881893 | 130881937 |
| Ntf3 | 212 | chr6 | 126116527 | 126116573 |
| Nup133 | -392 | chr8 | 126473535 | 126473579 |
| Olfr718-ps1 | -27326 | chr5 | 143927219 | 143927263 |
| Oxr1 | 342197 | chr15 | 41621200 | 41621249 |
| P4ha1 | -102082 | chr10 | 58683940 | 58683984 |
| Paics | 119 | chr5 | 77380428 | 77380473 |
| Pcf11 | 492 | chr7 | 99817929 | 99817973 |
| Pelp1 | 470 | chr11 | 70223034 | 70223086 |
| Pex16 | -181 | chr2 | 92215258 | 92215305 |
| Pigv | 472 | chr4 | 133228067 | 133228113 |
| Plat | -656 | chr8 | 23867538 | 23867582 |
| Polg | -33207 | chr7 | 86659650 | 86659694 |
| Ppat | 152 | chr5 | 77380428 | 77380473 |
| Ppip5k2 | 729 | chr1 | 99666237 | 99666281 |
| Prkesh | -478 | chr9 | 21806979 | 21807023 |
| Prmt1 | -656 | chr7 | 52242424 | 52242468 |
| Psmc5 | -416 | chr11 | 106117030 | 106117074 |
| Pxdn | -402745 | chr12 | 30220134 | 30220178 |

| | | | | |
|----------|--------|-------|-----------|-----------|
| Rab4b | -135 | chr7 | 27964028 | 27964072 |
| Rbak | 15213 | chr5 | 143927219 | 143927263 |
| Rcor2 | 27992 | chr19 | 7369864 | 7369908 |
| Rhcg | 118088 | chr7 | 86644433 | 86644477 |
| Rnf32 | -250 | chr5 | 29522265 | 29522309 |
| Rps12 | 272 | chr10 | 23506721 | 23506765 |
| Rsb1l1 | 978 | chr5 | 20456640 | 20456684 |
| Sept10 | -26843 | chr10 | 58683940 | 58683984 |
| Sh3bp5l | -174 | chr11 | 58144056 | 58144100 |
| Shc2 | -191 | chr10 | 79100832 | 79100876 |
| Six3 | -1824 | chr17 | 86018337 | 86018381 |
| Slc25a24 | 432 | chr3 | 108926469 | 108926528 |
| Slc35b1 | -94 | chr11 | 95245885 | 95245938 |
| Slc52a2 | -861 | chr15 | 76368490 | 76368534 |
| Slc6a2 | 951 | chr8 | 95484907 | 95484951 |
| Slc8a3 | -150 | chr12 | 82434295 | 82434339 |
| Slx4ip | 72 | chr2 | 136717954 | 136717998 |
| Smarcd2 | -353 | chr11 | 106134617 | 106134661 |
| Smarce1 | -140 | chr11 | 99092449 | 99092493 |
| Smg5 | -383 | chr3 | 88139777 | 88139821 |
| Snhg11 | -14386 | chr2 | 158186966 | 158187010 |
| Sox2 | 232 | chr3 | 34549137 | 34549181 |
| Spin1 | -33794 | chr13 | 51195901 | 51195945 |
| Spred1 | -1363 | chr2 | 116945722 | 116945771 |
| Srgap2 | 638 | chr1 | 133423278 | 133423322 |
| Tagap1 | -396 | chr17 | 7165879 | 7165923 |
| Tarbp2 | -518 | chr15 | 102348350 | 102348394 |
| Tbpl1 | 257 | chr10 | 22450973 | 22451018 |
| Tbx3 | -144 | chr5 | 120120861 | 120120905 |
| Tcam1 | -3423 | chr11 | 106134617 | 106134661 |
| Tcp1l1l | -346 | chr2 | 104552641 | 104552688 |
| Tdrkh | -462 | chr3 | 94216733 | 94216778 |
| Tet2 | -55250 | chr3 | 133206862 | 133206908 |
| Tex2 | 462 | chr11 | 106473760 | 106473804 |
| Tgif1 | 1112 | chr17 | 71199940 | 71199984 |
| Tmem79 | -1373 | chr3 | 88139777 | 88139821 |
| Tnfrsf1 | 494 | chr11 | 78349257 | 78349301 |
| Tob2 | 53 | chr15 | 81688681 | 81688725 |
| Tob2 | 523 | chr15 | 81688681 | 81688725 |
| Tor1a | 477 | chr2 | 30822954 | 30822998 |
| Tpbp | 106 | chr9 | 85736071 | 85736115 |
| Trim2 | -33927 | chr3 | 84108618 | 84108662 |
| Trpc4ap | 470 | chr2 | 155517628 | 155517672 |
| Tubb2b | 264 | chr13 | 34221937 | 34221981 |

| | | | | |
|--------|--------|-------|-----------|-----------|
| Ubap2 | 258 | chr4 | 41221888 | 41221932 |
| Ubb | 627 | chr11 | 62365278 | 62365322 |
| Ube2h | 129180 | chr6 | 30125336 | 30125381 |
| Vmac | -249 | chr17 | 56856571 | 56856615 |
| Zcchc2 | -580 | chr1 | 107886664 | 107886708 |
| Zfp526 | -413 | chr7 | 26005201 | 26005245 |
| Zfp580 | 325 | chr7 | 5003437 | 5003481 |
| Zfp622 | -249 | chr15 | 25913850 | 25913894 |
| Zzef1 | -606 | chr11 | 72609128 | 72609172 |

Supplemental Table 3.3b. Genes associated with hypomethylated probes.

| Gene name | Distance to TSS | Chromosome | Start | End |
|------------------|------------------------|-------------------|--------------|------------|
| 1700016C15Rik | -389 | chr1 | 179659529 | 179659582 |
| 3830403N18Rik | -706112 | chrX | 52683615 | 52683659 |
| 4930503E14Rik | -41673 | chr14 | 44790687 | 44790731 |
| 4930539E08Rik | 367 | chr17 | 29051879 | 29051924 |
| 4931400O07Rik | 59,135 | chrX | 59780275 | 59780327 |
| 4931400O07Rik | 59,746 | chrX | 59780890 | 59780934 |
| 8430423G03Rik | -414 | chr5 | 149761390 | 149761434 |
| 9230009I02Rik | 134,929 | chr11 | 51033496 | 51033540 |
| 9230009I02Rik | 136,593 | chr11 | 51035155 | 51035208 |
| A530032D15Rik | -697 | chr1 | 85107103 | 85107147 |
| A530032D15Rik | -900 | chr1 | 85107298 | 85107357 |
| Acpt | 636 | chr7 | 51512090 | 51512134 |
| Acvr2b | -622 | chr9 | 119310975 | 119311019 |
| Adamts15 | -933 | chr9 | 30730945 | 30730995 |
| Agap1 | -1068 | chr1 | 91350292 | 91350344 |
| Akirin1 | -163 | chr4 | 123427729 | 123427773 |
| Anapc1 | -840 | chr2 | 128513944 | 128513990 |
| Ankrd65 | -3682 | chr4 | 155160841 | 155160890 |
| Ap2a2 | 52436 | chr7 | 148800488 | 148800541 |
| Aplnr | -387 | chr2 | 84976108 | 84976152 |
| Apol10b | -1243 | chr15 | 77427776 | 77427820 |
| Aste1 | 806 | chr9 | 105298501 | 105298555 |
| Atp4a | -1049 | chr7 | 31496156 | 31496201 |
| AW822073 | -7350 | chr10 | 57695782 | 57695826 |
| BC049762 | 42,935 | chr11 | 51033496 | 51033540 |
| BC049762 | 41,271 | chr11 | 51035155 | 51035208 |
| C130026I21Rik | -673 | chr1 | 85267769 | 85267813 |
| C130026I21Rik | -881 | chr1 | 85267969 | 85268028 |
| C3 | 425 | chr17 | 57367112 | 57367156 |
| Cacnb2 | 115975 | chr2 | 14641881 | 14641935 |
| Calr4 | -17801 | chr4 | 108890414 | 108890458 |
| Capzb | -12933 | chr4 | 138736956 | 138737000 |
| Ccdc84 | -1137 | chr9 | 44227198 | 44227255 |
| Chrm4 | -1048 | chr2 | 91761276 | 91761320 |
| Cog6 | 90024 | chr3 | 52731096 | 52731145 |
| Cox18 | -162 | chr5 | 90653137 | 90653181 |
| Cox8c | -130 | chr12 | 104137364 | 104137408 |

| | | | | |
|------------|---------|-------|-----------|-----------|
| Cpa2 | 370 | chr6 | 30491926 | 30491978 |
| Ctnnb1 | -17109 | chr9 | 120825387 | 120825431 |
| D5Ertd577e | -168 | chr5 | 95885632 | 95885682 |
| Dmrtc1c1 | -196 | chrX | 99998084 | 99998128 |
| Fbln7 | 295 | chr2 | 128688982 | 128689041 |
| Foxo1 | 658862 | chr3 | 52731096 | 52731145 |
| Gja3 | 43833 | chr14 | 57633012 | 57633056 |
| Glrx | -1201 | chr13 | 75976093 | 75976137 |
| Gm10922 | -79525 | chrX | 4075076 | 4075120 |
| Gm11756 | 14064 | chr4 | 73552578 | 73552622 |
| Gm11757 | 14,103 | chr4 | 73522891 | 73522935 |
| Gm11757 | -743 | chr4 | 73537737 | 73537781 |
| Gm11757 | -15,584 | chr4 | 73552578 | 73552622 |
| Gm13139 | 43549 | chr4 | 146549670 | 146549722 |
| Gm13151 | -130087 | chr4 | 146549670 | 146549722 |
| Gm14288 | 161,592 | chr2 | 175638477 | 175638524 |
| Gm14288 | 161,311 | chr2 | 175638756 | 175638807 |
| Gm14288 | 64,041 | chr2 | 175736028 | 175736075 |
| Gm14288 | 63,760 | chr2 | 175736307 | 175736358 |
| Gm14305 | -1,000 | chr2 | 176492013 | 176492060 |
| Gm14305 | -724 | chr2 | 176492287 | 176492338 |
| Gm14322 | -4,633 | chr2 | 177492946 | 177492993 |
| Gm14322 | -4,339 | chr2 | 177493236 | 177493292 |
| Gm14325 | -964 | chr2 | 177575958 | 177576016 |
| Gm14326 | -666 | chr2 | 177692638 | 177692697 |
| Gm14403 | -1,104 | chr2 | 177281892 | 177281939 |
| Gm14403 | -815 | chr2 | 177282177 | 177282233 |
| Gm14434 | -4,611 | chr2 | 176265723 | 176265770 |
| Gm14434 | -4,330 | chr2 | 176266002 | 176266053 |
| Gm14525 | -359 | chrX | 26130093 | 26130152 |
| Gm14632 | 761605 | chrX | 30245594 | 30245638 |
| Gm2004 | 93,587 | chr2 | 175638477 | 175638524 |
| Gm2004 | 93,868 | chr2 | 175638756 | 175638807 |
| Gm2004 | 191,138 | chr2 | 175736028 | 175736075 |
| Gm2004 | 191,419 | chr2 | 175736307 | 175736358 |
| Gm2030 | -139 | chrX | 25737281 | 25737337 |
| Gm2030 | -343 | chrX | 25737486 | 25737540 |
| Gm2784 | 47415 | chrX | 30245594 | 30245638 |
| Gm3763 | -523068 | chrX | 4075076 | 4075120 |
| Gm428 | -788 | chr4 | 73328609 | 73328654 |
| Gm4724 | -935 | chr2 | 175262214 | 175262272 |

| | | | | |
|-----------|---------|-------|-----------|-----------|
| Gm4981 | 3,604 | chr10 | 57695782 | 57695826 |
| Gm4981 | -145 | chr10 | 57699531 | 57699575 |
| Gm5169 | -356 | chrX | 24728910 | 24728969 |
| Gm5634 | -868 | chrX | 8540947 | 8540991 |
| Gm5635 | -657 | chrX | 8639537 | 8639581 |
| Gm6710 | -4,823 | chr2 | 175867386 | 175867433 |
| Gm6710 | -4,538 | chr2 | 175867667 | 175867723 |
| Gm7168 | -1008 | chr17 | 14084342 | 14084401 |
| Gm7534 | 465 | chr4 | 133758432 | 133758476 |
| Gm7534 | 48 | chr4 | 133758846 | 133758895 |
| Gm8220 | -72537 | chr14 | 44790687 | 44790731 |
| Gm9 | -145 | chrX | 34751159 | 34751203 |
| Id2 | -31855 | chr12 | 25812790 | 25812834 |
| Ifne | -1165 | chr4 | 88527247 | 88527293 |
| Il22 | 481 | chr10 | 117642456 | 117642501 |
| Ipo7 | -11166 | chr7 | 117150745 | 117150801 |
| Klk14 | -1071 | chr7 | 50944693 | 50944740 |
| Krtap19-1 | -4624 | chr16 | 88874257 | 88874301 |
| Krtap19-2 | 235 | chr16 | 88874257 | 88874301 |
| Lcelj | -825 | chr3 | 92594194 | 92594238 |
| Limk2 | 52406 | chr11 | 3256747 | 3256794 |
| Lrrc73 | -631 | chr17 | 46390461 | 46390505 |
| Minos1 | -49950 | chr4 | 138736956 | 138737000 |
| Mroh1 | 9532 | chr15 | 76220453 | 76220497 |
| Muc6 | 40703 | chr7 | 148800488 | 148800541 |
| Nek11 | -674 | chr9 | 105298501 | 105298555 |
| Nsun6 | 329004 | chr2 | 14641881 | 14641935 |
| Olfr1140 | 369 | chr2 | 87586664 | 87586709 |
| Olfr1423 | 476 | chr19 | 12110732 | 12110776 |
| Olfr1427 | 478 | chr19 | 12173629 | 12173673 |
| Olfr19 | -145735 | chr16 | 16819816 | 16819873 |
| Olfr317 | -211 | chr11 | 58546912 | 58546959 |
| Olfr323 | 436 | chr11 | 58439088 | 58439132 |
| Olfr968 | 199 | chr9 | 39580158 | 39580209 |
| Olfr969 | 219 | chr9 | 39603158 | 39603204 |
| Osbpl9 | -15559 | chr4 | 108890414 | 108890458 |
| Pacrgl | -1048 | chr5 | 48762557 | 48762609 |
| Pea15a | -814 | chr1 | 174137727 | 174137771 |
| Pet2 | 262 | chrX | 86654744 | 86654788 |
| Pik3ip1 | 26367 | chr11 | 3256747 | 3256794 |
| Pitx1 | -374 | chr13 | 55933138 | 55933182 |

| | | | | |
|---------|----------|-------|-----------|-----------|
| Pla2g2f | 20565 | chr4 | 138292952 | 138292999 |
| Plb1 | 120626 | chr5 | 32655689 | 32655736 |
| Plxnb3 | -508 | chrX | 71001908 | 71001952 |
| Pot1a | -122399 | chr6 | 25881595 | 25881654 |
| Ppp1cb | -105634 | chr5 | 32655689 | 32655736 |
| Prf1 | -1269 | chr10 | 60759287 | 60759342 |
| Prx | -595 | chr7 | 28283726 | 28283770 |
| Ptpre | -1047 | chr7 | 142728453 | 142728497 |
| Ptprg | 408 | chr14 | 12386432 | 12386476 |
| Pzp | 108 | chr6 | 128476607 | 128476652 |
| Rasef | -86407 | chr4 | 73522891 | 73522935 |
| Rhox3e | 5378 | chrX | 34795269 | 34795314 |
| Rhox4a | -10041 | chrX | 34795269 | 34795314 |
| Rhox6 | -760 | chrX | 35179300 | 35179353 |
| Rnf122 | -982 | chr8 | 32221288 | 32221332 |
| Rpl14 | 344777 | chr9 | 120825387 | 120825431 |
| Rps4y2 | -518 | chr6 | 148302630 | 148302689 |
| Scgb2b3 | 234 | chr7 | 32146833 | 32146880 |
| Scml4 | -848 | chr10 | 42579303 | 42579353 |
| Scx | -67393 | chr15 | 76220453 | 76220497 |
| Sh3bp1 | -1256 | chr15 | 78728920 | 78728964 |
| Slx11 | -645,440 | chrX | 51834591 | 51834635 |
| Slx11 | -554,437 | chrX | 51925594 | 51925638 |
| Slx11 | 263 | chrX | 52480294 | 52480338 |
| Slx11 | 203,584 | chrX | 52683615 | 52683659 |
| Snrpn | 350,644 | chr7 | 66934233 | 66934279 |
| Snrpn | 350,849 | chr7 | 66934438 | 66934484 |
| Sp110 | -1096 | chr1 | 87496458 | 87496517 |
| Sp140 | -790 | chr1 | 87496458 | 87496517 |
| Spag6 | 9561 | chr16 | 16819816 | 16819873 |
| Speer4e | -807 | chr5 | 14939206 | 14939265 |
| Tbata | -877 | chr10 | 60633833 | 60633881 |
| Tcp10c | 442 | chr17 | 13547855 | 13547905 |
| Tmem41b | -20808 | chr7 | 117150745 | 117150801 |
| Tmem88b | -883 | chr4 | 155160841 | 155160890 |
| Trpv1 | 4424 | chr11 | 73052196 | 73052240 |
| Trpv3 | -28672 | chr11 | 73052196 | 73052240 |
| Ube3a | 450,134 | chr7 | 66934233 | 66934279 |
| Ube3a | 450,339 | chr7 | 66934438 | 66934484 |
| Ubxn10 | -11796 | chr4 | 138292952 | 138292999 |
| Ucma | -302 | chr2 | 4896873 | 4896917 |

| | | | | |
|---------|----------|-------|-----------|-----------|
| Xlr | -733,740 | chrX | 51834591 | 51834635 |
| Xlr | -824,743 | chrX | 51925594 | 51925638 |
| Xlr3a | -929 | chrX | 70343339 | 70343386 |
| Xlr3b | -701 | chrX | 70436838 | 70436882 |
| Xlr3c | -1004 | chrX | 70511707 | 70511758 |
| Zfp874a | -1183 | chr13 | 67553721 | 67553765 |
| Zfp882 | -147 | chr8 | 74432338 | 74432382 |
| Zfp931 | -688 | chr2 | 177813788 | 177813847 |
| Zfp931 | -979 | chr2 | 177814079 | 177814138 |
| Zmym2 | 126403 | chr14 | 57633012 | 57633056 |

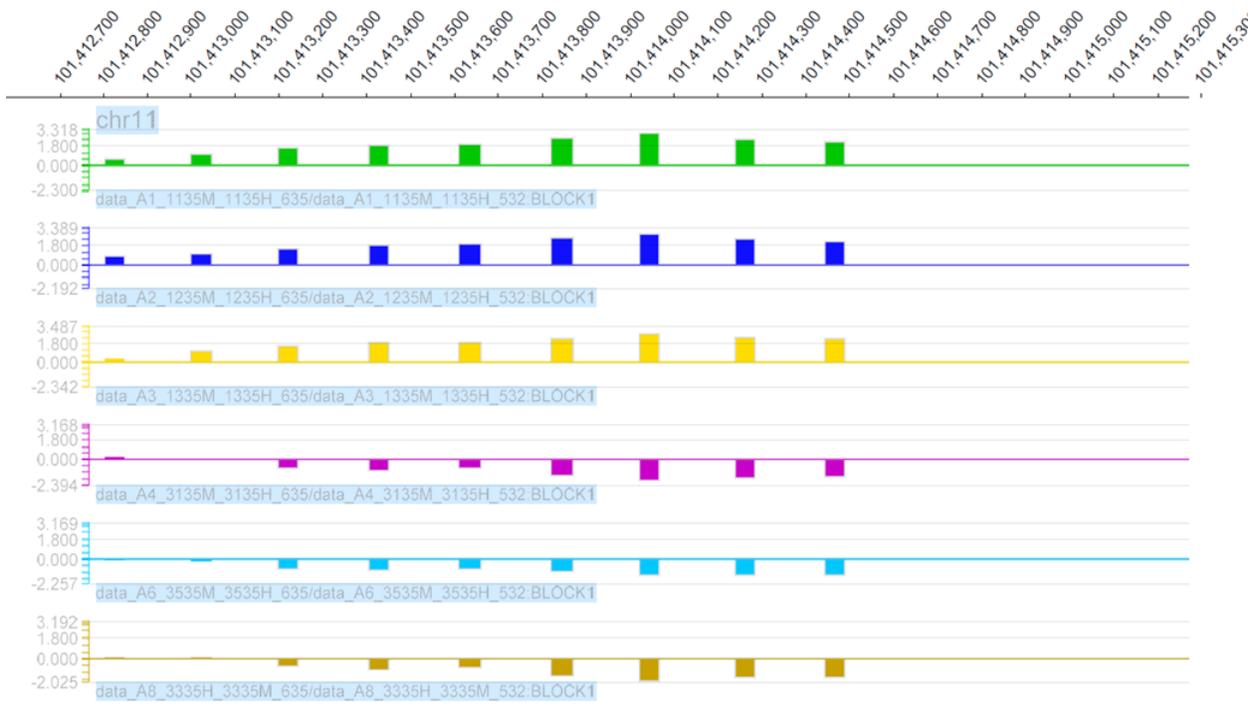
Supplemental Table 3.3. Genes associated with differentially methylated probes.

A list of probes exhibiting the most significant (top 1%) hypermethylation (a) or hypomethylation (b) in the deficient compared to control condition were mapped to associated genes using the GREAT analysis tool. The distance between the probe and the transcription start site (TSS) is listed in the second column. The location of the probe is listed in the third through fifth columns. The name of the probe is listed in the far right column. GREAT analysis allowed for the mapping of a single probe to multiple genes if the distance to the gene's TSS was very large (>5 kb) and no experimental evidence verified the long-distance mapping. Since these distal probes represented a very small percentage (<28%) of mappings, further analysis was not conducted in this study. Probes mapped to multiple genes are noted with an asterisk and a letter; matching letter patterns are used to easily associate identical probes with different gene mappings.

Supplemental Figure 3.1. Verification of differential methylation patterns in microarray probes for tumor suppressor genes.

Three tumor suppressor genes were identified for bisulfite sequencing analysis based on differential methylation patterns observed on the HELP assay promoter microarray. Prior to bisulfite sequencing, the differential methylation patterns of the probes in the promoter region of each gene were verified using SignalMap software to analyze raw GFF files provided by ArrayStar. HpaII/MspI \log_2 ratios are displayed on the x-axis, and chromosomal location is located on the y-axis. Ratios greater than zero (above each respective horizontal line) represent a lesser level of methylation, and ratios less than zero (below each respective horizontal line) represent a greater level of methylation. The top 3 rows (green, royal blue, yellow) represent biological replicates of the folic acid control condition, and the bottom 3 rows (pink, light blue, tan) represent deficient replicates. Differential methylation in promoter regions of each gene were verified by observation of \log_2 ratios in different directions in control versus deficient replicates, as demonstrated by ratios greater than zero for control and less than zero for deficient in each of the genes examined. (A) The *Brcal* promoter region is located at chr11: 101,412,769-101,415,269. The transcription start site (TSS) is located at chr11: 101,413,269, and the most significant differentially methylated probe is located at chr11:101,414,020-101,414,064. (B) The *Gad1* promoter region is located at chr2:70,398,221-70,400,721. The TSS is located at chr2:70,400,221, and the most significant differentially methylated probe is located at chr2:70399345-70399389. (C) The *Tet2* promoter region is located at chr3:133,206,854-133,209,354. The TSS is located at chr3:133,207,354, and the most significant differentially methylated probe is located at chr3:133,206,862-133,206,908.

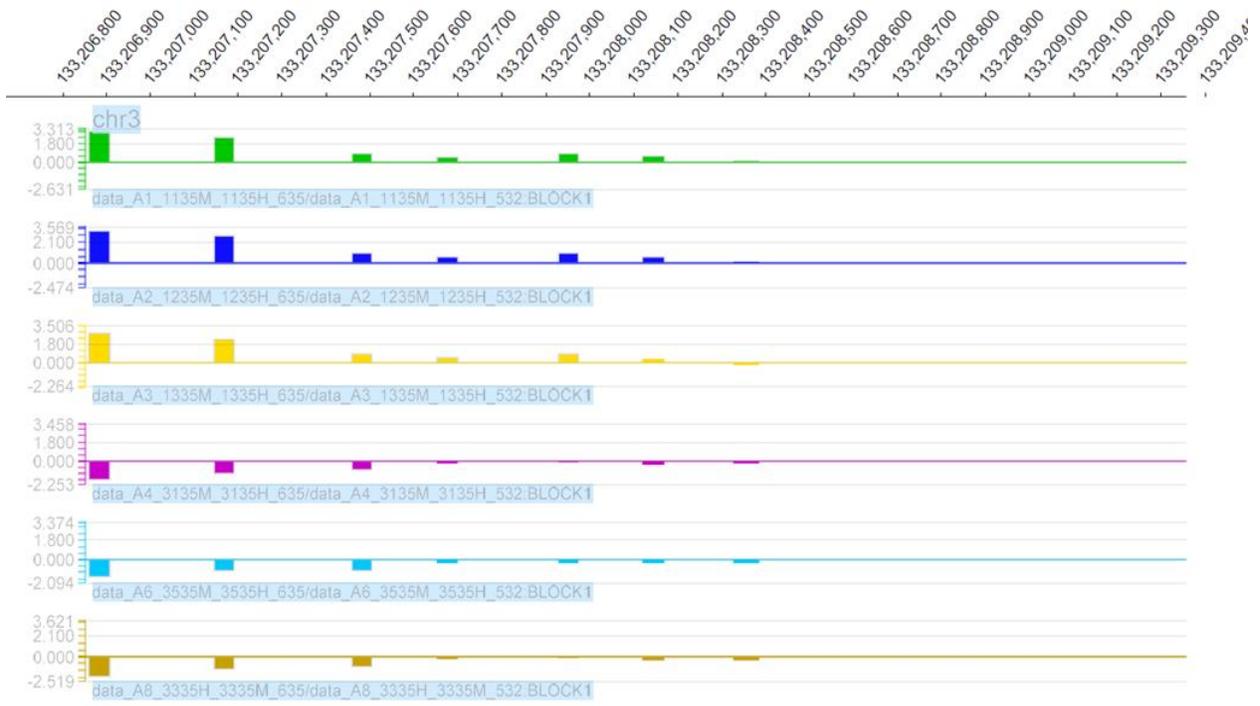
A

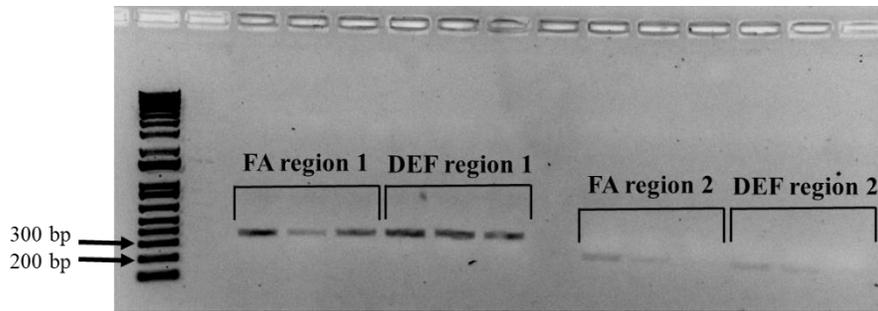
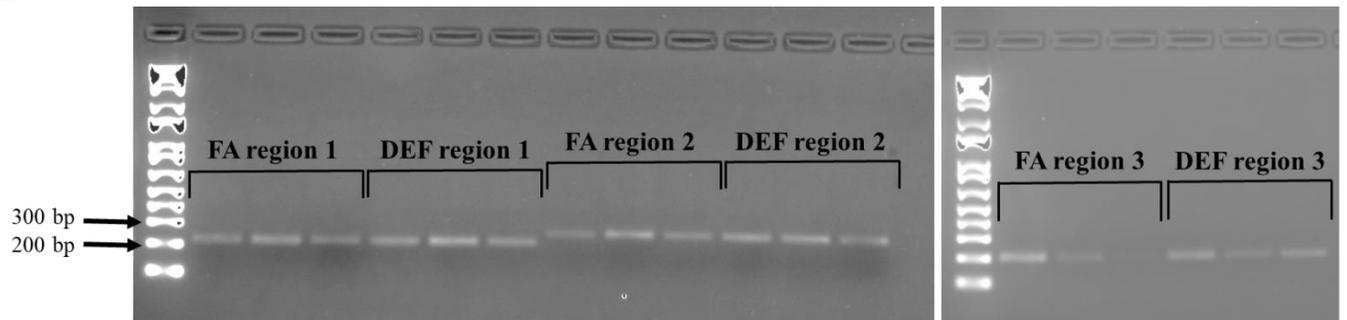
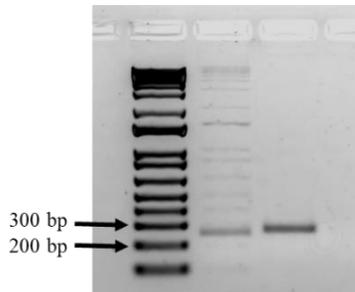


B



C



A**B****C**

Supplemental Figure 3.2. Confirmation of bisulfite PCR.

The initial step for bisulfite sequencing of tumor suppressor genes was PCR amplification of target promoter regions. This amplification is challenging since bisulfite treatment can alter the sequence and integrity of the DNA. Therefore, successful amplification was verified in two regions for *Brca1* (A) and *Tet2* (C) and three regions for *Gad1* (B) on 1% agarose gels. Labels 1-2 or 1-3 represent regions amplified. In A and B the left three wells of each region represent folic acid control mice, and the right three represent deficient mice. For C the left well represents control mice and the right well represents deficient mice.

Supplemental Figure 3.3. Sequencing alignment for Gad1.

Three regions of *Gad1* were analyzed by bisulfite sequencing and aligned using BiQ Analyzer. Two clones were sequenced for each biological replicate for each dietary condition (FA control condition and DEF deficient condition). Sequences were aligned with a modified version of the original genomic DNA where all non-CpG cytosines were converted to thymines, and all CpG cytosines remained as cytosines. Each CpG was highlighted in orange in the genomic sequence along with all clone sequences sites containing methylated CpGs. Each unmethylated CpG is highlighted in purple and noted with thymine replacing cytosine. Clones with less than 80% alignment to the original sequences were considered to be too low quality for subsequent analysis. Figures A, B, and C represent *Gad1* promotor regions 1, 2, and 3, respectively.

CHAPTER 4

Initial characterizations of Zpf410*

* Britton Upchurch contributed to this chapter by assisting with histological staining. Britton Upchurch and Feifan Xu contributed to this chapter by assisting with protein production and isolation.

ABSTRACT

Zfp410 is a zinc-finger protein that has recently been associated with a number of neurological processes including axonal and brain development. Disruption of this protein's function is associated with a number of adverse outcomes including paragangliomas, schizophrenia, and cognitive decline. The role of Zfp410 in health and disease has only recently begun to gain interest, so this protein remains largely uncharacterized. In this study we show the association between Zfp410 binding sites and gene regulation of a number of cell division related genes. We also show association between Zfp410 and genes which are differentially expressed in the hippocampus of mice exhibiting folate-induced cognitive decline. Finally, we produce antibodies for Zfp410 and demonstrate the perinuclear localization of those antibodies in a preserved human hippocampus. These initial findings solidify interest in further characterization of this poorly understood protein.

INTRODUCTION

Zinc finger proteins (ZFP) represent a large class of proteins with a wide variety of functions, including binding to DNA, RNA, and proteins (Gibson, Postma, Brown, & Argos, 1988). As such, many of these proteins are thought to play roles in transcriptional regulation of many genes. ZFPs have been shown to play important roles in a number of neurological diseases including spinal muscular dystrophy, Alzheimer's disease, and Parkinson's disease (Helmken et al., 2003; Ko et al., 2012; Lefebvre et al., 1995; Li, Strohmeyer, Liang, Lue, & Rogers, 2004; Shin et al., 2011). Expression of some of these proteins appears to be localized to the brain or central nervous system, implicating them as neuronal regulators (Cassandri et al., 2017; Pao et al., 2011). Further, this class of protein is able to induce epigenetic modifications through

interactions with histones, implicating them as regulators of gene expression in their target tissues (Borgel et al., 2017). Indeed, the zinc-finger domain is one of the most frequent binding motifs for eukaryotic transcription factors (Cassandri et al., 2017).

Zfp410 is a zinc finger protein which is thought to bind to DNA and act as a regulator of gene expression (Ji & Jaffrey, 2014; Sharp et al., 2011). This transcription factor has recently been associated with a variety of neuronal conditions and features including paragangliomas, schizophrenia, axonal development, and brain development (Bryois et al., 2018; Francis et al., 2014; Ji & Jaffrey, 2014; Smestad & Maher, 2019). Apart from these studies in which Zfp410 is a minor bioinformatics finding in a larger study, this transcription factor remains largely uncharacterized. However, these studies together with our recent observation of an association between Zfp410 and folate deficiency-induced cognitive decline (Chapter 2), indicate that Zfp410 may be of great interest in understanding neuronal functions in disease states.

Because of interesting initial implications of Zfp410, we sought to begin to further characterize this largely mysterious protein. Because the timing of these initial findings did not align with the ability to include a robust characterization as a part of this dissertation, this chapter represents preliminary results that serve to solidify apparent significance of this protein and the need for ongoing characterization of its role in neuronal health and disease.

METHODS

Gene mapping

Zfp410 binding sites were previously determined to be upregulated in the hippocampus of 6 mo old mice with chronic post-weaning folate deficiency (Figure 2.6). The sequence of the enriched motif (AGGGCTGGCTA) was found to be associated with 392 regions in the mouse

genome (Supplemental Table 4.1). These regions were mapped to nearby genes that they putatively regulate using Genomic Regions Enrichment of Annotations Tool (GREAT) software (Stanford University and Bejarano Lab). As previously described (Chapter 3), mouse genome mm9 was used as the species assembly and the whole genome was used for background regions. Genes were assigned basal regulatory domains as follows: proximal domain from 5 kb upstream to 1 kb downstream of the transcription start site (TSS) and distal regulatory domain up to an additional 1000 kb from the proximal domain, regardless of nearby genes. Additionally, curated regulatory domains were used to override other domain recognition rules if significant experimental evidence demonstrates that an element falls outside of the assigned regulatory domains. Probes were allowed to be mapped to multiple genes if experimental evidence or relative location did not provide a clear association to a single gene. Gene ontology was performed as a part of GREAT analysis to determine significance of gene functions (McLean et al., 2010).

Further, this gene list produced by GREAT analysis was cross-referenced with lists of differentially expressed genes associated with folate deficiency at 6 mo and 18 mo in a mouse hippocampus (Figure 4.1). This crossing sought to determine associations between disease states and genes regulated by Zfp410 binding. This cross-referencing was done using R version 3.1.0 (R Foundation for Statistical Computing, Vienna, Austria) and packages “matrixTests” and “data.table” for analysis.

Zfp410 protein isolation and purification

Zfp410 was purchased in a pET45 vector with His tagging and transformed into *E. coli* BL21 cells. Cells were grown up in 200 mL of Luria Bertani broth containing 100 µg/mL of

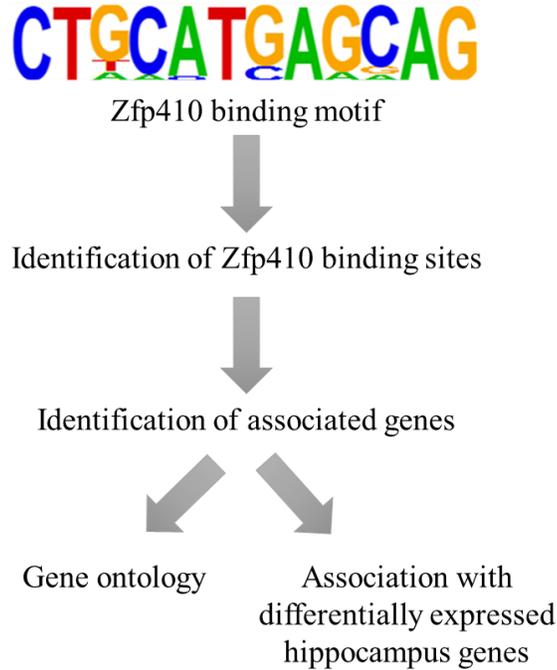


Figure 4.1. Zfp410 study design for binding regions.

Zfp410 binding motifs were previously identified to be enriched in promotor regions of differentially expressed genes in the adult mouse hippocampus following chronic folate deficiency. The primary binding site (CTGCATGAGCAG) was mapped to 392 regions in the mouse genome. These regions were mapped to associated genes by determining proximity to the nearest transcription start site using GREAT software. Gene ontology was performed on this list of associated genes. Further, this list was crossed with lists of differentially expressed genes in the hippocampus of 6 mo and 18 mo mice following chronic post-weaning folate deficiency.

Ampicillin (Invitrogen, Carlsbad, CA) at 37°C with shaking at 250 rpm until the culture reached OD₆₀₀ between 0.4-0.6 (about 3 hours). Cells were then induced with 200 µL of 1M isopropyl β-d-1-thiogalactopyranoside (IPTG, Invitrogen, Carlsbad, CA) and incubated for 2 hr at the same temperature with shaking. Cells were harvest by spinning at 6,000 x g for 15 min at 4°C then resuspended in 10 mL of lysis buffer (10 mM Tris pH 7.5, 200 mM NaCl, 0.5 mM EDTA, 10% glycerol, 0.1% igepal, 10 mM imdazol, 1 mM DTT). Cells were sonicated at 10% amplitude for 3 bursts of 45 sec, with 1 min rest on ice between sonications. Sonicated cultures were spun at 27,000 x g for 10 min at 4°C, and the pellet was resuspended in 5 mL of Buffer A (0.1 M NaCl, 20 mM Tris pH 7.5, 5 mM DTT, 1 mM EDTA, 0.5 M urea) and left rocking overnight at 4°C. The overnight solution was spun at 27,000 x g for 10 min at 4°C, and the pellet was resuspended in 500 µL of Buffer A. 400 µL of Ni-NTA resin (ThermoFisher Scientific, Waltham, MA), and the solution was incubated with rocking for 2 hr at 4°C. The resin was them washed 4 times with 10 mL of His-tag wash buffer (10 mM Tris pH 7.5, 0.2% igepal, 0.3M NaCl, 10% glycerol, 15 mM imidazole, 1 mM DTT). The protein was then eluted from the resin 4 times in 400 µL of His-tag elution buffer (10 mM Tris pH 7.5, 0.1% igepal, 0.2 M NaCl, 10% glycerol, 250 mM imidazol). The eluted protein was placed in dialysis tubing, and imidazole was removed using 400 mL of dialysis buffer (10 mM Tris pH 7.5, 0.1% igepal, 0.2 M NaCl, 10% glycerol). The purified protein was quantified using Bradford and BSA (BioRad, Hercules, CA) and sent to the Pocono Rabbit Farm (Canadensis, PA) for a 91 day protocol of antibody production <http://www.prfal.com/index.cfm?e=inner4&itemcategory=87177>.

Histological staining

Preserved whole hippocampus tissue was harvested from a 77-year old male with no documented nervous system conditions. The tissue was cut into 7.5 mm² sections, rinsed in phosphate buffered saline (PBS, ThermoFisher Scientific, Waltham, MA) and incubated overnight covered in Cryo-Embedding Compound (Ted Pella, Inc. Redding, CA). The tissue was then placed in a cryostat mold and covered in fresh Cryo-Embedding Compound. Embedded tissue was flash-frozen in a dry ice EtOH bath and transferred immediately to the cryostat for cutting. Slices of hippocampus tissue were cut at 7 μm and mounted on slides.

Slides were incubated in phosphate buffered saline and tween 20x (PBT) for 30 min at room temperature. Zfp410 antibody from the Pocono Rabbit Farm was diluted 1:1000 in TBS and added to the slide and incubated in a dark humidified chamber at 4°C for one hour. The slide was then rinsed 3 times with PBT. Alexa Fluor 488 goat anti-rabbit IgG (H+L) (Invitrogen, Carlsbad, CA) secondary antibody was added to the slide and again incubated in a dark humidified chamber at 4°C for 1 hr. The slide was again rinsed 3x with PBT. One drop of DAPI Prolong Gold Antifade Mountant (ThermoFisher Scientific, Waltham, MA) was added, and slides were visualized by fluorescent microscopy.

RESULTS

Zfp410 binding sites associated with promotor regions of numerous genes.

GREAT software was used to map the locations of all Zfp410 binding site to genes that may be regulated by this binding. Although 545 regions contained the AGGGGCTGGCTA binding sequence, a number of these regions were mapped to the same gene, resulting in 51 genes with multiple Zfp410 binding sites in their promoter regions (Table 4.1). The majority of

| Gene | Distance to TSS |
|---------------|--|
| 3830403N18Rik | -186,350, -162,643, -30,923 |
| Amd2 | 282,068; 566,436 |
| Cblb | -302,176; 142,509 |
| Gm20736 | 98,855; 387,266 |
| Gm20738 | -796,944; -568,487; -348,566, -114,428; 174,119; 448,319; 674,399 |
| Gm20747 | -809,685; -529,345; -338,630; -42,516; 251,044 |
| Gm20773 | 26,952; 391,627 |
| Gm20795 | 27,038; 546,277 |
| Gm20806 | -39,631; 253,232 |
| Gm20809 | -854,940; -627,610; -339,041; 1114,377; 174,131; 398,762; 687,187; 911,851 |
| Gm20817 | -143,825; 100,157 |
| Gm20820 | -130,820; 98,777; 413,265; 639,295; 926,401 |
| Gm20823 | -863,780; -632,193; -349,727; -114,294 |
| Gm20843 | -762,663; -400,503; -178,169; 99,740; 166,312; 454,884; 678,799 |
| Gm20852 | -39,797; 378,287; 593,280 |
| Gm20854 | -39,708; 320,853 |
| Gm20865 | 173,313; 403,722 |
| Gm20867 | -113,706; 170,760; 383,563; 544,597 |
| Gm20888 | 91,917; 380,720; 614,940; 897,938 |
| Gm20890 | -361,381; -129,609; 98,715; 383,702; 608,327; 896,791 |
| Gm20911 | -131,304; 99,354; 385,983; 680,160 |
| Gm20924 | -40,569; 238,439; 307,070; 634,328; 863,850 |
| Gm20937 | 99,646; 386,972; 548,845 |
| Gm21095 | -467,385; -128,311; 98,727 |
| Gm21117 | -416,864; -128,782; 99,753; 308,648; 597,074; 826,823 |
| Gm21173 | -947,290; -658,736; -429,949; -143,900; 100,033; 429,022 |
| Gm21209 | -128,527; 99,700 |
| Gm21258 | -476,334; -407,663; -129,607; 98,844 |
| Gm21294 | -419,297; -125,019; 104,164 |
| Gm21394 | -280,877; -39,158 |
| Gm21488 | -880,674; -650,887; -361,245; -129,596; 98,750 |
| Gm21497 | -144,833; 99,952 |
| Gm21518 | -128,588; 99,705 |
| Gm21760 | -426,832; -135,939; 98,819; 371,440 |
| Gm21854 | -272,643; 27,083 |
| Gm21858 | -977,358; -646,550; -419,349; -132,162; 48,112; 328,437; 633,798 |
| Gm21943 | -39,782; 247,822; 475,029; 795,757 |
| Gm28079 | -10,008; 266,266 |
| Gm28102 | -837,839; -630,733; -404,939; -129,286; 99,386; 391,941\ |

| | |
|---------|--|
| Gm28553 | -144,843; 99,997 |
| Gm28827 | -702,033; -421,449; -127,784; 99,110; 372,102 |
| Gm28870 | -137,872; 97,561; 379,753 |
| Gm28897 | -128,827; 98,850 |
| Gm29110 | -888,234; -697,597; -417,046; -128,345; 99,658; 596,737; 824,157 |
| Gm31571 | -938,675; -711,114; -421,192; -128,535; 98,782 |
| Snx7 | -287,281; 52,037 |
| Ssty1 | 144,684; 509,247 |
| Ssty2 | -485,217; -333,004; -39,841 |
| Tbl1xr1 | -541,592; 109,426 |
| Tmem92 | -27,270; -11,683 |
| Zfp516 | -22,198; 482,119 |

Table 4.1 Genes associated with multiple Zfp410 binding sites.

Although 546 regions were associated with the AGGGCTGGCTA binding motif for Zfp410, only 392 genes were associated with this motif. This outcome is due to the finding that 51 genes were associated with multiple Zfp410 binding motifs. These genes may represent targets with significant dependence on Zfp410 for regulation.

the Zfp410 binding regions were located very distally (>500 kb) to the gene with which they were associated (Figure 4.2)

Zfp410 binding sites associated with differentially expression genes in chronic folate deficiency.

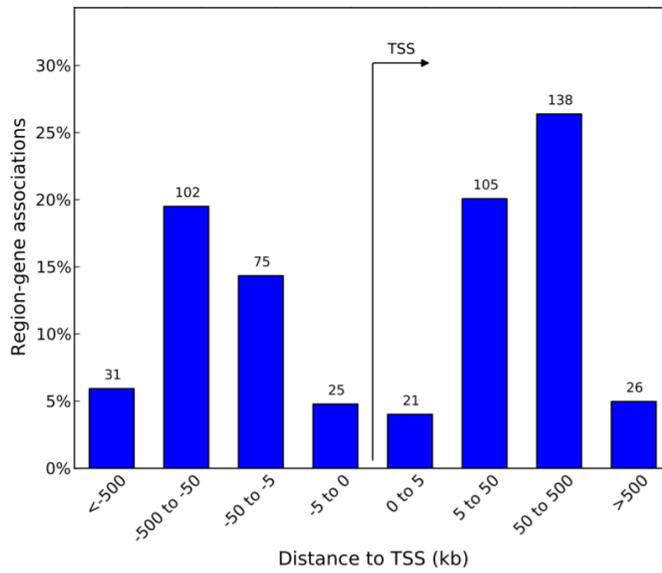
Since Zpf410 was initially examined as a protein of interest due to enrichment of its binding motifs associated with differentially expressed genes in the chronically folate deficient mouse hippocampus, genes from GREAT analysis were cross-referenced with the differentially expressed gene lists from chronically folate deficient mice to determine Zfp410 regulated genes that may be of physiological interest. A total of 6 genes were associated with Zfp410 binding sites and were differentially expressed in the hippocampus mice with chronic post-weaning folate deficiency (Table 4.2). All associated genes were common to both 6 mo and 18 mo of folate deficiency.

Gene ontology was performed on the list of Zfp410 associated genes from GREAT analysis using the same software. All of the ontological categories identified were associated with chromosomal elements or cell division processes (Figure 4.3, Supplemental Table 4.1).

Zfp410 isolated as insoluble protein.

Zpf410 was successfully expressed in *E. coli* cells, isolated and purified. The protein size was expected to be 37 kDa, and this size was confirmed by Western blot analysis (Figure 4.4).

A



B

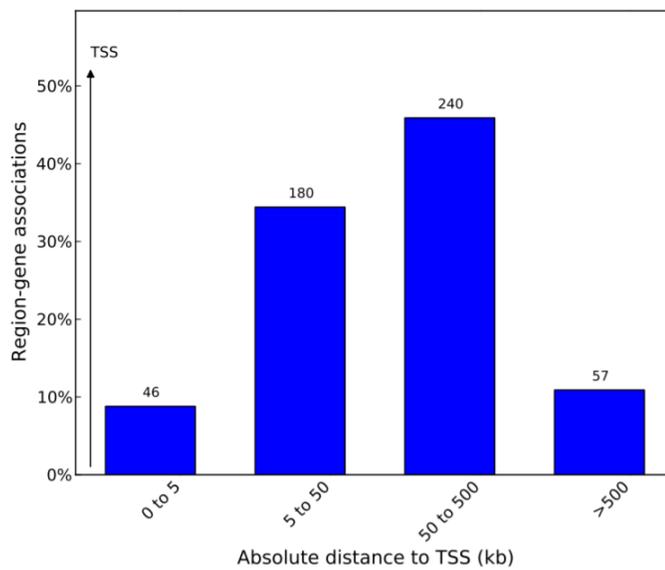


Figure 4.2 Distance between TSS and Zfp410 binding motifs.

Zfp410 binding motifs were linked to associated genes using GREAT software. The motifs are roughly evenly distributed both upstream and downstream of the transcription start site (TSS) of their associated gene (A). The majority of the motifs lie very distal (>5 kb) in relation to the TSS of their associated gene (B).

| Regulation | Gene |
|------------|----------------------------|
| Down | Bhlhb9, Pgs1, Eml1, Pde10a |
| Up | Cbln4, Smyd3 |

Table 4.2. Zfp410 regulated genes associated with differentially expressed genes in cognitive decline.

Six genes were found to be both regulated by Zfp410 binding as associated with differential expression in the hippocampus of a chronically folate-deficient mouse exhibiting memory deficits. All six of these genes were differentially expressed at both 6 mo and 18 mo of folate deficiency. Genes are listed along with their differential regulation in the folate deficient hippocampus.

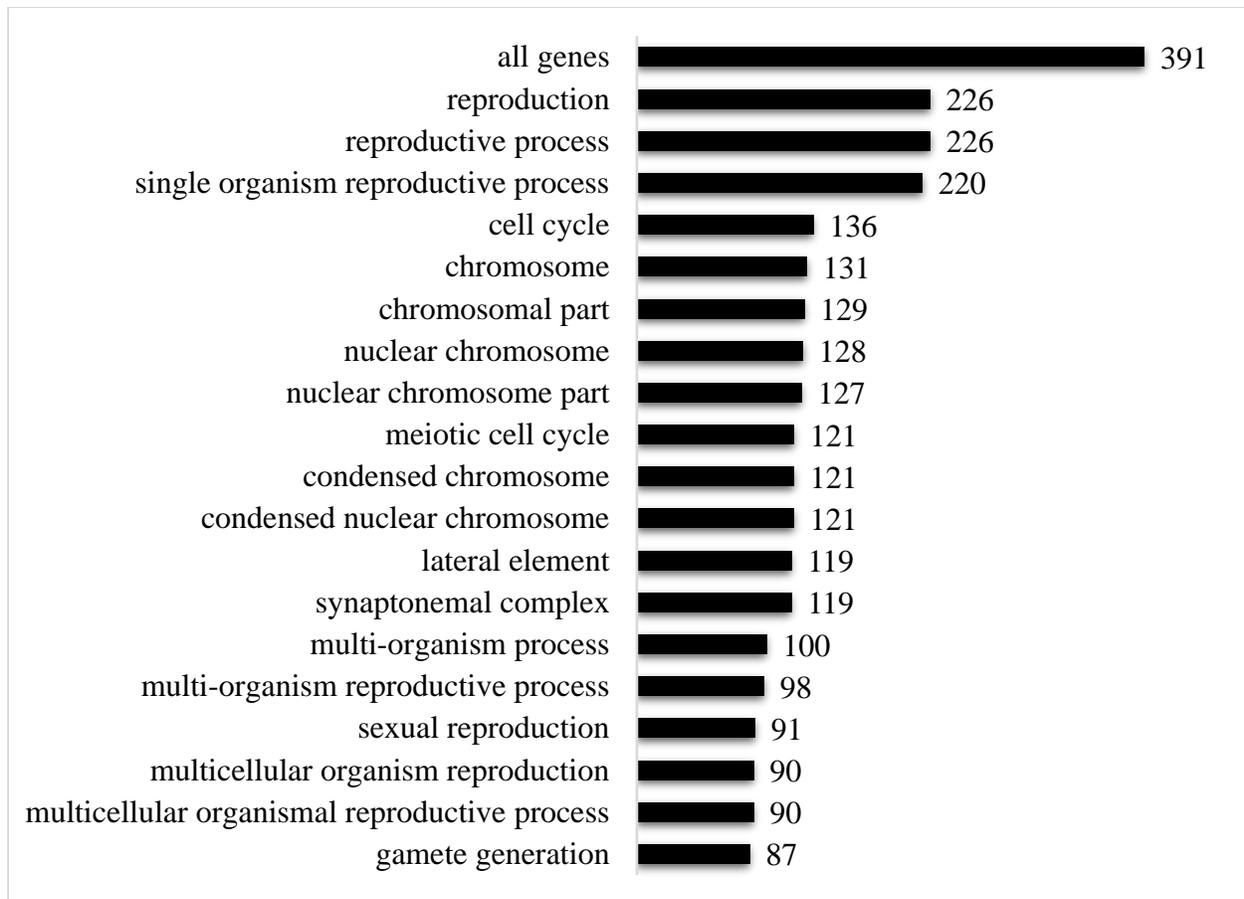


Figure 4.3. Zfp410 binding motifs are associated with cell division processes.

GREAT software was used to determine ontological categories associated with genes regulated by Zfp410 binding. All of these ontological categories are associated with processes involved in cell division. Ontological categories are listed on the left, and the number of associated Zfp410 regulated genes is listed on the right of each category.

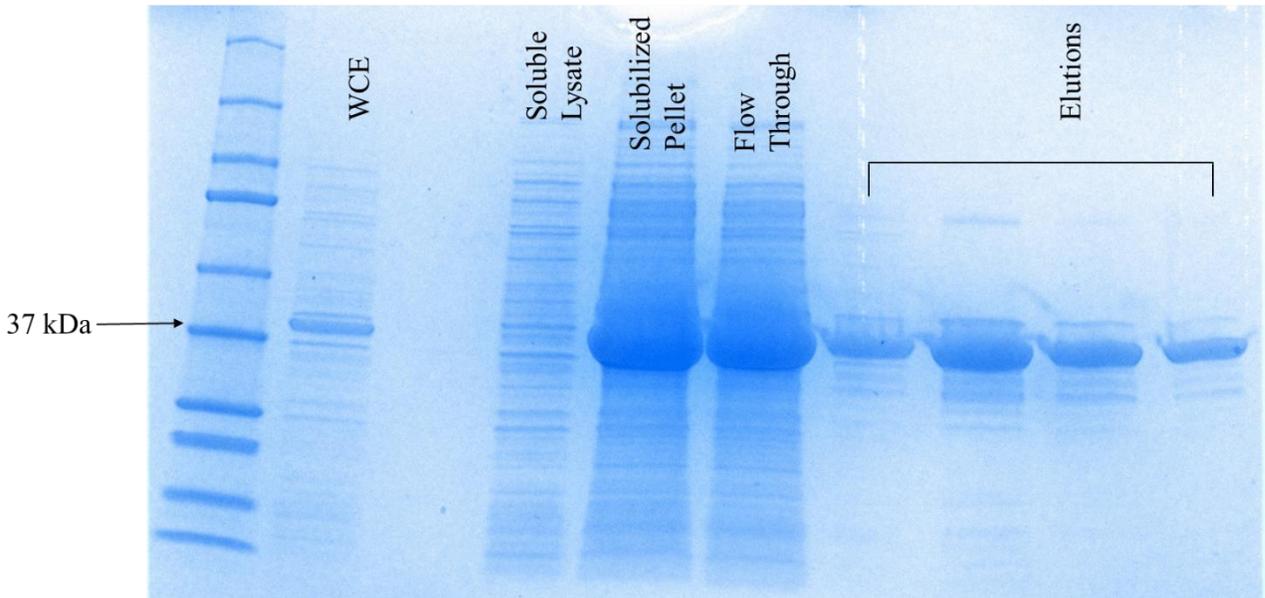


Figure 4.4. Isolation of Zfp410 protein.

Isolation and purification of Zfp410 protein is shown on a 4–20% Mini-protean TGX Stain-Free Protein Gel. The protein was calculated to have an expected size of 38 kDa, and this size was confirmed on the gel. The protein is clearly visible in the whole cell extract (WCE) in which it was expressed, the overnight incubation culture (soluble lysate and pellet), flow through of this culture, and all elutions.

To our knowledge, this represents the first purification as well as the first antibody production for this protein.

Zfp410 antibody localized to perinuclear region in preserved human brains.

Since Zfp410 arose from hippocampal studies, Zfp410 antibody binding was visualized on hippocampus tissue isolated from preserved human brains. Initial results indicate that the antibody is localized to a perinuclear region in hippocampus samples (Figure 4.4). Although this immunostaining was planned to be compared to other tissues, such as unpreserved mouse and rat as well as additional preserved human, unforeseen circumstances did not allow for additional histological samples to be processed.

DISCUSSION

The Zfp410 protein remains largely uncharacterized, with few references available in literature, the majority of which only briefly mention Zfp410 as a secondary finding in a larger study (Bryois et al., 2018; Francis et al., 2014; Smestad & Maher, 2019). Since Zfp410 appears to be of interest in cognitive function based on a handful of recent publications as well as our recent study documenting the outcomes of chronic folate deficiency in the hippocampus (Chapter 2), we sought to undertake initial characterizations of this protein, beginning with data obtained from folate deficiency mouse studies.

It is interesting to note that a significant number of the genes associated with Zfp410 binding sites contain more than one Zfp410 binding site in their promoter region (Table 4.1). Although many of these associations present with a large distance between the gene and the binding site, large distances have been well documented in literature

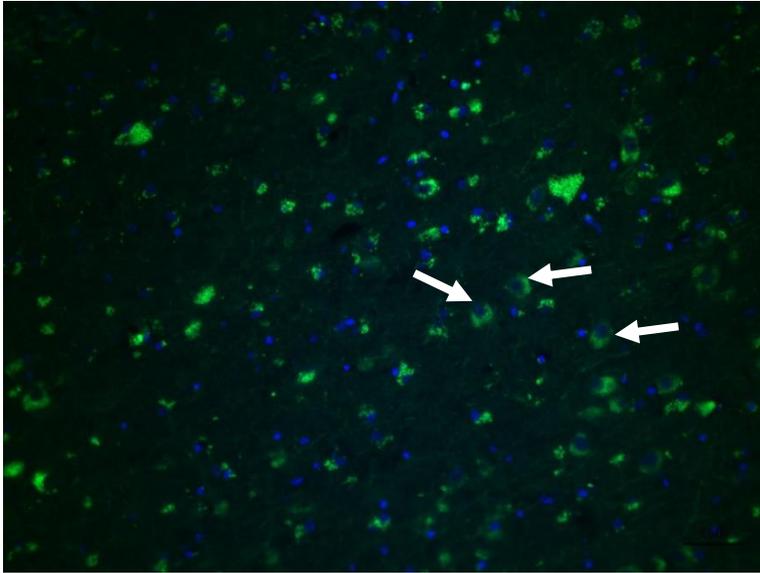


Figure 4.5. Perinuclear localization of Zfp410 antibody in the human hippocampus. Immunostaining with DAPI (blue) and Zfp410 antibody (green) demonstrates a perinuclear localization of Zfp410 in the hippocampus of a preserved human hippocampus with no documented neurological diseases. White arrows represent clear examples of this phenomenon. The image is shown is a composite image at 20x amplification using the Rhod and DAPI filters.

(Kulaeva, Nizovtseva, Polikanov, Ulianov, & Studitsky, 2012; Sanyal, Lajoie, Jain, & Dekker, 2012). Since a large number of genes were associated with multiple Zfp410 binding motifs, it is likely that even though these motifs are fairly distal, these genes are likely associated in some way with Zfp410 and many serve as good targets for initial focus of Zfp410 characterization.

Knowing that Zfp410 was chosen for further characterization analysis due to its association with differentially regulated hippocampal genes in a folate deficient mouse (Chapter 2), it is not surprising to note that the list of genes that Zfp410 putatively regulates intersects with a number of differentially expressed hippocampal genes (Table 4.2). This list represents both upregulated and downregulated genes in the folate deficient hippocampus, demonstrating that Zfp410 binding may not induce the same results with all associated genes.

The observation that Zfp410 is ontologically associated only with chromosomal or cell division related categories is particularly interesting as it appears that the function of this protein is very specific. The ontological observations are further bolstered by histological observation of a perinuclear location for Zfp410 antibody. Since perinuclear proteins are able to bind to cytoskeletal elements which are important in nuclear positioning for cell division, the perinuclear location of this antibody affirms ontological findings of Zfp410's potential involvement in cell division processes (Crisp et al., 2006). However, it is important to note that preserved tissues sometimes present with complications in antibody staining (Paavilainen et al., 2010). Future work will compare staining in unpreserved tissues with this initial data to determine the fidelity of these initial results. Should the results of immunostaining on preserved tissues prove to exhibit similar integrity compared to unpreserved tissues, Zfp410 binding will be compared in human brains with no documented pathologies as well as brains with documented Alzheimer's disease and dementia. Since Zfp410 is associated with cognitive decline, it would be interesting to

determine any differential staining patterns associated with cognitive diseases. Future studies will also compare antibody staining in different brain regions to determine if Zfp410 is localized to a specific brain regions such as the hippocampus or if it is pervasive in all brain tissue.

Although the results of this study are preliminary, they begin to enhance our understanding of a largely uncharacterized transcription factor that appears to be of importance in the central nervous system. Initial findings support the idea that Zfp410 is associated with gene regulation in the hippocampus and that disruptions of this gene's function may result in cognitive disease states. These initial finding warrant further investigations in fully characterizing the structure and function of this relatively novel protein.

REFERENCES

- Borgel, J., Tyl, M., Schiller, K., Pusztai, Z., Dooley, C. M., Deng, W., . . . Bartke, T. (2017). KDM2A integrates DNA and histone modification signals through a CXXC/PHD module and direct interaction with HP1. *Nucleic Acids Res*, *45*(3), 1114-1129. doi: 10.1093/nar/gkw979
- Bryois, J., Garrett, M. E., Song, L., Safi, A., Giusti-Rodriguez, P., Johnson, G. D., . . . Crawford, G. E. (2018). Evaluation of chromatin accessibility in prefrontal cortex of individuals with schizophrenia. *Nat Commun*, *9*(1), 3121. doi: 10.1038/s41467-018-05379-y
- Cassandri, M., Smirnov, A., Novelli, F., Pitolli, C., Agostini, M., Malewicz, M., . . . Raschella, G. (2017). Zinc-finger proteins in health and disease. *Cell Death Discov*, *3*, 17071. doi: 10.1038/cddiscovery.2017.71
- Crisp, M., Liu, Q., Roux, K., Rattner, J. B., Shanahan, C., Burke, B., . . . Hodzic, D. (2006). Coupling of the nucleus and cytoplasm: role of the LINC complex. *J Cell Biol*, *172*(1), 41-53. doi: 10.1083/jcb.200509124
- Francis, C., Natarajan, S., Lee, M. T., Khaladkar, M., Buckley, P. T., Sul, J. Y., . . . Kim, J. (2014). Divergence of RNA localization between rat and mouse neurons reveals the potential for rapid brain evolution. *BMC Genomics*, *15*, 883. doi: 10.1186/1471-2164-15-883
- Gibson, T. J., Postma, J. P., Brown, R. S., & Argos, P. (1988). A model for the tertiary structure of the 28 residue DNA-binding motif ('zinc finger') common to many eukaryotic transcriptional regulatory proteins. *Protein Eng*, *2*(3), 209-218. doi: 10.1093/protein/2.3.209
- Helmken, C., Hofmann, Y., Schoenen, F., Oprea, G., Raschke, H., Rudnik-Schoneborn, S., . . . Wirth, B. (2003). Evidence for a modifying pathway in SMA discordant families: reduced SMN level decreases the amount of its interacting partners and Htra2-beta1. *Hum Genet*, *114*(1), 11-21. doi: 10.1007/s00439-003-1025-2
- Ji, S. J., & Jaffrey, S. R. (2014). Axonal transcription factors: novel regulators of growth cone-to-nucleus signaling. *Dev Neurobiol*, *74*(3), 245-258. doi: 10.1002/dneu.22112
- Ko, C. Y., Chang, L. H., Lee, Y. C., Sterneck, E., Cheng, C. P., Chen, S. H., . . . Wang, J. M. (2012). CCAAT/enhancer binding protein delta (CEBPD) elevating PTX3 expression inhibits macrophage-mediated phagocytosis of dying neuron cells. *Neurobiol Aging*, *33*(2), 422 e411-425. doi: 10.1016/j.neurobiolaging.2010.09.017
- Kulaeva, O. I., Nizovtseva, E. V., Polikanov, Y. S., Ulianov, S. V., & Studitsky, V. M. (2012). Distant activation of transcription: mechanisms of enhancer action. *Mol Cell Biol*, *32*(24), 4892-4897. doi: 10.1128/MCB.01127-12
- Lefebvre, S., Burglen, L., Reboullet, S., Clermont, O., Bulet, P., Viollet, L., . . . et al. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, *80*(1), 155-165. doi: 10.1016/0092-8674(95)90460-3

- Li, R., Strohmeyer, R., Liang, Z., Lue, L. F., & Rogers, J. (2004). CCAAT/enhancer binding protein delta (C/EBPdelta) expression and elevation in Alzheimer's disease. *Neurobiol Aging*, 25(8), 991-999. doi: 10.1016/j.neurobiolaging.2003.10.016
- McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., . . . Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol*, 28(5), 495-501. doi: 10.1038/nbt.1630
- Paavilainen, L., Edvinsson, A., Asplund, A., Hober, S., Kampf, C., Ponten, F., & Wester, K. (2010). The impact of tissue fixatives on morphology and antibody-based protein profiling in tissues and cells. *J Histochem Cytochem*, 58(3), 237-246. doi: 10.1369/jhc.2009.954321
- Pao, P. C., Huang, N. K., Liu, Y. W., Yeh, S. H., Lin, S. T., Hsieh, C. P., . . . Lee, Y. C. (2011). A novel RING finger protein, Znf179, modulates cell cycle exit and neuronal differentiation of P19 embryonal carcinoma cells. *Cell Death Differ*, 18(11), 1791-1804. doi: 10.1038/cdd.2011.52
- Sanyal, A., Lajoie, B. R., Jain, G., & Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature*, 489(7414), 109-113. doi: 10.1038/nature11279
- Sharp, B. M., Chen, H., Gong, S., Wu, X., Liu, Z., Hiler, K., . . . Matta, S. G. (2011). Gene expression in accumbens GABA neurons from inbred rats with different drug-taking behavior. *Genes Brain Behav*, 10(7), 778-788. doi: 10.1111/j.1601-183X.2011.00716.x
- Shin, J. H., Ko, H. S., Kang, H., Lee, Y., Lee, Y. I., Pletinkova, O., . . . Dawson, T. M. (2011). PARIS (ZNF746) repression of PGC-1alpha contributes to neurodegeneration in Parkinson's disease. *Cell*, 144(5), 689-702. doi: 10.1016/j.cell.2011.02.010
- Smestad, J. A., & Maher, L. J., 3rd. (2019). Master regulator analysis of paragangliomas carrying SDHx, VHL, or MAML3 genetic alterations. *BMC Cancer*, 19(1), 619. doi: 10.1186/s12885-019-5813-z

SUPPLEMENTAL INFORMATION

| Ontological Term | P-Value | Fold Enrichment |
|---|----------------|------------------------|
| reproduction | 1.1982E-82 | 4.1737 |
| reproductive process | 1.1879E-82 | 4.1739 |
| single organism reproductive process | 1.5331E-83 | 4.3596 |
| cell cycle | 3.2686E-39 | 3.5816 |
| chromosome | 3.9262E-41 | 3.8716 |
| chromosomal part | 4.5518E-43 | 4.1042 |
| nuclear chromosome | 6.5434E-53 | 5.1257 |
| nuclear chromosome part | 3.0586E-53 | 5.2111 |
| meiotic cell cycle | 4.6402E-67 | 7.4696 |
| condensed chromosome | 1.1279E-66 | 7.4084 |
| condensed nuclear chromosome | 1.7885E-75 | 8.9191 |
| lateral element | 2.5632E-77 | 9.5343 |
| synaptonemal complex | 2.2348E-76 | 9.3468 |
| multi-organism process | 5.8258E-22 | 2.932 |
| multi-organism reproductive process | 4.6068E-28 | 3.6132 |
| sexual reproduction | 4.5872E-28 | 3.8638 |
| multicellular organism reproduction | 1.3385E-23 | 3.3601 |
| multicellular organismal reproductive process | 8.5209E-24 | 3.3829 |
| gamete generation | 1.3176E-29 | 4.2437 |

Supplemental Table 4.1. Gene ontology categories for genes associated with Zfp410 binding motifs.

GREAT software was used to determine ontological categories associated with genes regulated by Zfp410 binding. All of these ontological categories are associated with processes involved in cell division. Ontological categories are listed along with the p-value and fold enrichment calculated by GREAT.

APPENDIX

Appendix A. Hippocampus microarray analysis R code.

```
library(data.table)
hippoarray <- read.csv(file.choose(), header = TRUE)

names(hippoarray)[4] <- "Fold.Change"
#need to change "Fold Change" to Fold.Change since R doesn't like the space

attach(hippoarray)
#now I can type in the names of any of my column headers directly to access them

FA4 <- subset(hippoarray, Arraycond == 'fa4')
#make subset of only FA4 data

MTHF4 <- subset(hippoarray, Arraycond == 'mthf4')

setDT(FA4, keep.rownames = FALSE, key = NULL, check.names = FALSE)
setDT(MTHF4, keep.rownames = FALSE, key = NULL, check.names = FALSE)
#rm() removes names

ALL4 <- fintersect(FA4[,1], MTHF4[,1], all = FALSE)

f4 <- FA4[match(as.character(ALL4$ProbeName), FA4$ProbeName, nomatch = FALSE), ]
#fa probes that intersect with mthf probes
m4 <- MTHF4[match(as.character(ALL4$ProbeName), MTHF4$ProbeName, nomatch =
FALSE), ]
#mthf probes that intersect with mthf probes

library(matrixTests)
row_t_welch(f4[,33:35], f4[,36:38], alternative = "two.sided", mu=0, conf.level = 0.95)
tf4 <- row_t_welch(f4[,33:35], f4[,36:38], alternative = "two.sided", mu=0, conf.level = 0.95)

row_t_welch(m4[,33:35], m4[,36:38], alternative = "two.sided", mu=0, conf.level = 0.95)
tm4 <- row_t_welch(m4[,33:35], m4[,36:38], alternative = "two.sided", mu=0, conf.level = 0.95)

sum(tf4[,12] <= 0.05)
sum(tm4[,12] <= 0.05)

f4t.sig <- cbind.data.frame(tf4, f4, deparse.level = 1, make.row.names = TRUE, stringsAsFactors
= stringsAsFactors)
```

```

m4t.sig<-cbind.data.frame(tm4, m4, deparse.level = 1, make.row.names = TRUE,
stringsAsFactors = stringsAsFactors)

m4t<-subset(m4t.sig, pvalue<=0.05)
f4t<-subset(f4t.sig, pvalue<=0.05)

setDT(f4t.sig, keep.rownames = TRUE, key = NULL, check.names = FALSE)
setDT(m4t.sig, keep.rownames = TRUE, key = NULL, check.names = FALSE)
setDT(f4t, keep.rownames = TRUE, key = NULL, check.names = FALSE)
setDT(m4t, keep.rownames = TRUE, key = NULL, check.names = FALSE)

all.sig<-fintersect(m4t[,27], f4t[,27], all = FALSE)

norms<-cbind.data.frame(f4[,33:38], m4[,33:38])
setcolorder(norms, neworder = c(1,2,3,7,8,9,4,5,6,10,11,12))
setDT(norms, keep.rownames = TRUE, key = NULL, check.names = FALSE)

row_t_welch(norms[,1:6], norms[,7:12], alternative = "two.sided", mu = 0, conf.level = 0.95)
fm4t<-row_t_welch(norms[,1:6], norms[,7:12], alternative = "two.sided", mu = 0, conf.level =
0.95)
#t.test on combined normalized data from fa4 and mthf4

setDT(fm4t, keep.rownames = TRUE, key = NULL, check.names = FALSE)

fm4.all<-cbind.data.frame(fm4t, f4, deparse.level = 1, make.row.names = TRUE)
setDT(fm4.all, keep.rownames = TRUE, key = NULL, check.names = FALSE)

fm4.sig<-subset(fm4.all, pvalue<=0.05)

fmm4.all<-cbind.data.frame(fm4t, m4, deparse.level = 1, make.row.names = TRUE)
setDT(fmm4.all, keep.rownames = TRUE, key = NULL, check.names = FALSE)
fmm4.sig<-subset(fmm4.all, pvalue<=0.05)
#t.test combined with 5mthf data

library(fcros)

library(data.table)
hippoarray <-read.csv(file.choose(), header =TRUE)

```

```

names(hippoarray)[4]<-"Fold.Change"
#need to change "Fold Change" to Fold.Change since R doesn't like the space

attach(hippoarray)
#now I can type in the names of any of my column headers directly to access them

FA18 <- subset(hippoarray, Arraycond == 'fa18')
#make subset of only FA4 data

MTHF18 <- subset(hippoarray, Arraycond == 'mthf18')

setDT(FA18, keep.rownames = FALSE, key= NULL, check.names = FALSE)
setDT(MTHF18, keep.rownames = FALSE, key= NULL, check.names = FALSE)
#rm() removes names

ALL18 <-fintersect(FA18[,1], MTHF18[,1], all = FALSE)

f18 <- FA18[match(as.character(ALL18$ProbeName), FA18$ProbeName, nomatch = FALSE), ]
#fa probes that intersect with mthf probes
m18 <- MTHF18[match(as.character(ALL18$ProbeName), MTHF18$ProbeName, nomatch =
FALSE), ]
#mthf probes that intersect with mthf probes

library(matrixTests)

norms18<-cbind.data.frame(f18[,33:38], m18[,33:38])
setcolorder(norms18, neworder = c(1,2,3,7,8,9,4,5,6,10,11,12))

row_t_welch(norms18[,1:6], norms18[,7:12], alternative = "two.sided", mu = 0, conf.level =
0.95)
fm18t<-row_t_welch(norms18[,1:6], norms18[,7:12], alternative = "two.sided", mu = 0,
conf.level = 0.95)
#t.test on combined normalized data from fa4 and mthf4

setDT(fm18t, keep.rownames = TRUE, key = NULL, check.names = FALSE)

fm18.all<-cbind.data.frame(fm18t, f18, deparse.level = 1, make.row.names = TRUE)
setDT(fm18.all, keep.rownames = TRUE, key = NULL, check.names = FALSE)

fm18.sig<-subset(fm18.all, pvalue<=0.05)
#t.test combined with fa data

fmm18.all<-cbind.data.frame(fm18t, m18, deparse.level = 1, make.row.names = TRUE)

```

```
setDT(fmm18.all, keep.rownames = TRUE, key = NULL, check.names = FALSE)
fmm18.sig<-subset(fmm18.all, pvalue<=0.05)
#t.test combined with 5mthf data
```

Appendix B. Boxplot for weight comparisons of mice R code.

```
mweights<-read.csv(file.choose(), header = TRUE)
#only read in file with FA data

View(mweights)

names(mweights)

attach(mweights)

Age<-as.factor(Age)

Diet<-as.factor(Diet)

#boxplot(Weight~Age,ylim=c(0,100),main="Mouse Weights by Diet and Age",las=1)

#boxplot(Weight[Age=="18 months"]~Diet[Age=="18 months"],ylim=c(0,100),main="Mouse
Weights by Diet and Age",ylab="Grams",xlab="Age",las=1)

#boxplot(Weight~Diet*Age,ylim=c(20,90),main="Mouse Weights by Diet and
Age",ylab="Grams",xlab="Age",col=c(4,5,2,8),las=2)

#don't plot these for now

# ~determines which things are plotted against each other *chooses a discrimination factor

boxplot(Weight~Age*Diet,ylim=c(20,80),main="Mouse Weights by Diet and
Age",ylab="Grams",xlab="Age and Diet",col=c(2,3,4),las=1, axes=F)

axis(1, at=c(2,5), labels=c("FA", "DEF"))

#label x axis

axis(2, at=c(20,40,60,80), las=1)

#label y axis

box()

#boxplot(Weight~Age*Diet,ylim=c(20,90),main="Mouse Weights by
Diet",ylab="Grams",xlab="Folate Diet",col=c(2,3,4),axes=F,las=2)

#axis(2,at=c(20,30,40,50,60,70,80,90),las=1)

#box()

#axis(1,at=c(2,5,8,11),labels=c("FA","FA-def","5MTHF","5MTHF-def"))
```

```

#this legend can specify different symbols and sizes using pch and cex
#legend(x=0.5,y=90,pch=15,cex=2,legend=c("6 months","12 months", "18 months"),bty="n")
legend(x=0.5,y=75,legend=c("6 months", "12 months", "18
months"),bty="n",fill=c(2,3,4),cex=1)

barplot(c(mean(mweights$Weight[Diet=="FA"&
Age=="6"]),mean(mweights$Weight[Diet=="FA"&
Age=="12"]),mean(mweights$Weight[Diet=="FA"&
Age=="18"]),mean(mweights$Weight[Diet=="FA-def"&
Age=="6"]),mean(mweights$Weight[Diet=="FA-def"&
Age=="12"]),mean(mweights$Weight[Diet=="FA-def"& Age=="18"])))

axis(1,at=c(0.6,2,3.4,5,6.4,8),labels=c("FA6", "FA12", "FA18", "FA-def6", "FA-def12", "FA-
def18"))

t.test(Weight[Diet=="FA" & Age=="6"],Weight[Diet=="FA-def" & Age=="6"])
#t = 0.58619, df = 2.3052, p-value = 0.6101

t.test(Weight[Diet=="FA" & Age=="12"],Weight[Diet=="FA-def" & Age=="12"])
#t = 1.2535, df = 3.3826, p-value = 0.2897

t.test(Weight[Diet=="FA" & Age=="18"],Weight[Diet=="FA-def" & Age=="18"])
#t = 0.42224, df = 3.3145, p-value = 0.6988

save.image(file = "mouse_weights")

```

Appendix C. Volcano plots of mRNA and lncRNA R code.

```
F4<-read.csv(file.choose(), header = TRUE)
#FA4 mRNA and lncRNA with Fold.Change.Direction representing down changes as negatives

attach(F4)

#with(F4, "Fold.Change.Direction", "P.value", main="Volcano plot", xlim=c(-4,4))

install.packages("ggplot2")
install.packages("gridExtra")
install.packages("plotly")

library(plotly)

pal <- c("black", "black", "black")

#plot_ly(type = "scatter", data = F4, x = Fold.Change.Direction, y = F4[,3], mode = "markers",
color = F4[,7], colors = pal, opacity = 0.75)

p <- plot_ly(type = "scatter", data = F4, x = Fold.Change.Direction, y = F4[,3], mode =
"markers", color = F4[,7], colors = pal, alpha = 0.5)
layout(p, xaxis = list(showgrid = FALSE), yaxis = list(showgrid = FALSE))

plot(x = log2(Fold.Change.Direction), y = -log10(P.value), main = "FA 4 month", pch = 19,
frame = FALSE, color = F4[,6], colors = pal)
#tried basic plotting but can't get colors to change

library(scatterplot3d)

grps <- as.factor(F4[,6])
colors <-c("black", "springgreen4", "springgreen2")

scatterplot3d(data = F4, x = FDR, y = -log10(P.value), z = Fold.Change.Direction, pch=16, color
= colors[grps], grid = TRUE, box = FALSE, angle = -30)
#legend(legend = c("NotSignificant", "SignificantFDR", "Significant"),col = c("black",
"springgreen4", "springgreen2"), pch = 16)

corr<-plot_ly(type = "scatter", data = F4, x = F4[,25], y = F4[,26], mode = "markers", marker =
list(color = "black"), alpha = 0.5)
layout(corr, xaxis = list(showgrid = FALSE), yaxis = list(showgrid = FALSE))
#correlations plot (Fa norm vs Def norm)
```

Appendix D. HELP assay microarray analysis R code.

```
m1335 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
mb1335 <- subset(m1335, Category == "bcg")
```

```
#background probes for 1335
```

```
me1335 <- subset(m1335, Category == "exp")
```

```
#experimental probes for 1335
```

```
m1135 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
mb1135 <- subset(m1135, Category == "bcg")
```

```
me1135 <- subset(m1135, Category == "exp")
```

```
m1235 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
mb1235 <- subset(m1235, Category == "bcg")
```

```
me1235 <- subset(m1235, Category == "exp")
```

```
m3135 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
mb3135 <- subset(m3135, Category == "bcg")
```

```
me3135 <- subset(m3135, Category == "exp")
```

```
m3335 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
#use A8 instead of A5 since correlations were off with A5 (maybe samples switched?)
```

```
#A8 is the dye swap of A5
```

```
mb3335 <- subset(m3335, Category == "bcg")
```

```
me3335 <- subset(m3335, Category == "exp")
```

```
m3535 <- read.csv(file.choose(), header = TRUE, stringsAsFactors = FALSE)
```

```
mb3535 <- subset(m3535, Category == "bcg")
```

```
me3535 <- subset(m3535, Category == "exp")
```

```
median(mb1335[,13])
```

```
#green median signal = 17
```

```
median(mb1335[,14])
```

```
#red median signal = 18
```

```
mad(m1335[,13])
```

```
#green MAD = 80.0604
```

```
mad(m1335[,14])
```

```
#red MAD = 102.2994
```

```
median(mb1135[,13])
```

```
#green median signal = 17
```

```
median(mb1135[,14])
```

```
#red median signal = 17.5
```

```
mad(m1135[,13])
```

```
#green MAD = 108.2298
```

```
mad(m1135[,14])
```

```

#red MAD = 137.1405
median(mb1235[,13])
#green median signal = 17
median(mb1235[,14])
#red median signal = 17.5
mad(m1235[,13])
#green MAD = 82.2843
mad(m1235[,14])
#red MAD = 111.9363
median(mb3135[,13])
#green median signal = 38.5
median(mb3135[,14])
#red median signal = 20
mad(m3135[,13])
#green MAD = 219.4248
mad(m3135[,14])
#red MAD = 240.9225
median(mb3335[,13])
#green median signal = 33
median(mb3335[,14])
#red median signal = 22
mad(m3335[,13])
#green MAD = 188.2902
mad(m3335[,14])
#red MAD = 217.9422
median(mb3535[,13])
#green median signal = 35.5
median(mb3535[,14])
#red median signal = 23
mad(m3535[,13])
#green MAD = 234.2508
mad(m3535[,14])
#red MAD = 230.5443

all <- cbind(me1135, me1235[,13:14], me1235[,19], me1335[,13:14], me1335[,19],
me3135[,13:14], me3135[,19], me3335[,13:14], me3335[,19], me3535[,13:14], me3535[,19],
deparse.level = 1)
#combine all mouse data and include green and red median signals and log ratios
allQ <- subset(all, all[,13] >= 287.5745 & all[,14] >= 360.35125 & all[,21] >= 222.71075 &
all[,22] >= 297.34075 & all[,24] >= 217.151 & all[,25] >= 275.485 & all[,27] >= 587.062 &
all[,28] >= 622.30625 & all[,30] >= 503.7255 & all[,31] >= 566.811 & all[,33] >= 621.127 &
all[,34] >= 599.36075)

```

```
#subset all data such that all signals are at least 3MAD above background (call these quality signals)
```

```
mean(allQ[,19])
```

```
#1.229728
```

```
mean(allQ[,23])
```

```
#1.401805
```

```
mean(allQ[,26])
```

```
#1.30673
```

```
mean(allQ[,29])
```

```
#-0.3910728
```

```
mean(allQ[,32])
```

```
#-0.2380447
```

```
mean(allQ[,35])
```

```
#-0.5233048
```

```
#find the mean of all the log ratios
```

```
allQ$norm11 <- with(allQ, allQ[,19] - 1.229728)
```

```
allQ$norm12 <- with(allQ, allQ[,23] - 1.401805)
```

```
allQ$norm13 <- with(allQ, allQ[,26] - 1.30673)
```

```
allQ$norm31 <- with(allQ, allQ[,29] + 0.3910728)
```

```
allQ$norm33 <- with(allQ, allQ[,32] + 0.2380447)
```

```
allQ$norm35 <- with(allQ, allQ[,35] + 0.5233048)
```

```
#subtract the logratio mean from the logratio to mean center all the data
```

```
#mean range is now 1.077487e-07 to 3.004203e-09, which I will effectively note as zero
```

```
allQ$controls <- with(allQ, rowMeans(allQ[,36:38]))
```

```
allQ$experimentals <- with(allQ, rowMeans(allQ[,39:41]))
```

```
#create new columns containing averages of the three replicate mice
```

```
cor(allQ[,42], allQ[,43])
```

```
#-0.828129
```

```
plot(allQ[,42], allQ[,43])
```

```
#correlation plot of average FA(column 42) vs DEF (column 43) normalized signal intensities
```

```
cor(allQ[36], allQ[37])
```

```
#correlations between control samples: 0.9787073, 0.9766367, 0.9779943
```

```
#correlations between experimental samples: 0.9732059, 0.9640061, 0.9563756
```

```
length(which(allQ[,42]<0))
```

```
#12545 (50.001%) of control genes were hypermethylated
```

```
length(which(allQ[,42]>0))
```

```

#12535 (49.998%) of control genes were hypomethylated
length(which(allQ[,43]<0))
#13096 (52.216%) of experimental genes were hypermethylated
length(which(allQ[,43]>0))
#11984 (47.738%) of experimental genes were hypomethylated

allQ$difference <- with(allQ, allQ[,43] - allQ[,42])

allOrder <- allQ[order(allQ$difference),]
#sort data in order based on the difference

length(which(allQ$difference >= 1))
#8213 values (32%) >=1 (implies >= 2.7 fold change)
length(which(allQ$difference <= -1))
#8578 values (34%) <= -1
length(which(allQ[,44]<0))
#12554 (50.06%) probes gained methylation (>0 log difference)
length(which(allQ[,44]>0))
#12526 (49.94%) probes lost methylation (<0 log difference)

allOrder$SumSqs <- kmeans(allOrder[,44], centers = k)

hist(allQ$norm11, breaks = 150, main = "M1135 normalized", xlab = "HpaII/Msp1(log2)")
hist(allQ$norm12, breaks = 150, main = "M1235 normalized", xlab = "HpaII/Msp1(log2)")
hist(allQ$norm13, breaks = 150, main = "M1335 normalized", xlab = "HpaII/Msp1(log2)")
hist(allQ$norm31, breaks = 150, main = "M3135 normalized", xlab = "HpaII/Msp1(log2)")
hist(allQ$norm33, breaks = 150, main = "M3335 normalized", xlab = "HpaII/Msp1(log2)")
hist(allQ$norm35, breaks = 150, main = "M3535 normalized", xlab = "HpaII/Msp1(log2)")
#histograms of normalized ratio for individual mice

hist(allQ[,19], breaks = 150, main = "M1135 raw", xlab = "HpaII/Msp1(log2)")
hist(allQ[,23], breaks = 150, main = "M1235 raw", xlab = "HpaII/Msp1(log2)")
hist(allQ[,26], breaks = 150, main = "M1335 raw", xlab = "HpaII/Msp1(log2)")
hist(allQ[,29], breaks = 150, main = "M3135 raw", xlab = "HpaII/Msp1(log2)")
hist(allQ[,32], breaks = 150, main = "M3335 raw", xlab = "HpaII/Msp1(log2)")
hist(allQ[,35], breaks = 150, main = "M3535 raw", xlab = "HpaII/Msp1(log2)")

hist(allQ$controls, breaks = 150, main = "Control (normal folate levels)", xlab =
"HpaII/Msp1(log2)")
hist(allQ$experimentals, breaks = 150, main = "Experimental (folate deficient)", xlab =
"HpaII/Msp1(log2)")
#histograms of average control and average experimental mice

```

```
hist(allQ$difference, breaks = 150, main = "Folate deficiency induced methylation changes",
xlab = "HpaII/MspI(Log2) Difference")
#histogram of normalized ratio data
```

```
hist(allOrder[1:125,44], breaks = 50, main = "High gain of methylation", xlab =
"HpaII/MspI(Log2) Difference")
hist(allOrder[24955:25080,44], breaks = 50, main = "High loss of methylation", xlab =
"HpaII/MspI(Log2) Difference")
#graphs representing the 1% most changed data (top 0.5% and bottom 0.5%)
```

```
allM <- subset(all, all[,13] >= 287.5745 & all[,21] >= 222.71075 & all[,24] >= 217.151 &
all[,27] >= 587.062 & all[,30] >= 503.7255 & all[,33] >= 621.127)
#subset all data such that all MspI but not HpaII signals are at least 3MAD above background
(call these quality signals)
```

```
mean(allM[,19])
#1.203017
mean(allM[,23])
#1.371302
mean(allM[,26])
#1.274316
mean(allM[,29])
#-0.4305832
mean(allM[,32])
#-0.2792487
mean(allM[,35])
#-0.5729359
#find the mean of all the log ratios
```

```
allM$norm11 <- with(allM, allM[,19] - 1.203017)
allM$norm12 <- with(allM, allM[,23] - 1.371302)
allM$norm13 <- with(allM, allM[,26] - 1.274316)
allM$norm31 <- with(allM, allM[,29] + 0.4305832)
allM$norm33 <- with(allM, allM[,32] + 0.2792487)
allM$norm35 <- with(allM, allM[,35] + 0.5729359)
#subtract the logratio mean from the logratio to mean center all the data
#mean range is now 1.077487e-07 to 3.004203e-09, which I will effectively note as zero
```

```
allM$controls <- with(allM, rowMeans(allM[,36:38]))
allM$experimentals <- with(allM, rowMeans(allM[,39:41]))
```

```

#create new columns containing averages of the three replicate mice

length(which(allM[,42]<0))
#15064 (49.08%) of control genes were hypermethylated
length(which(allM[,42]>0))
#15629 (50.92%) of control genes were hypomethylated
length(which(allM[,43]<0))
#16327 (53.19%) of experimental genes were hypermethylated
length(which(allM[,43]>0))
#14366 (46.18%) of experimental genes were hypomethylated

allM$difference <- with(allM, allM[,43] - allM[,42])

allMOrder <- allM[order(allM$difference),]
#sort data in order based on the difference

length(which(allM$difference >= 1))
#10006 values (32.6%) >=1 (implies >= 2.7 fold change)
length(which(allM$difference <= -1))
#10839 values (35.3%) <= -1

hist(allM$controls, breaks = 150, main = "Control (normal folate levels)", xlab =
"HpaII/MspI(log2)")
hist(allM$experimentals, breaks = 150, main = "Experimental (folate deficient)", xlab =
"HpaII/MspI(log2)")
#histograms of average control and average experimental mice

hist(allM$difference, breaks = 150, main = "Folate deficiency induced methylation changes",
xlab = "HpaII/MspI(Log2) Difference")
#histogram of normalized ratio data

hist(allMOrder[1:153,44], breaks = 100, main = "High gain of methylation", xlab =
"HpaII/MspI(Log2) Difference")
hist(allMOrder[30540:30693,44], breaks = 100, main = "High loss of methylation", xlab =
"HpaII/MspI(Log2) Difference")
#graphs representing the 1% most changed data (top 0.5% and bottom 0.5%)

library(matrixTests)
library(data.table)

setDT(allM, keep.rownames = TRUE, key = NULL, check.names = FALSE)

```

```

allMT <- row_t_welch(allM[,37:39], allM[,40:42], alternative = "two.sided", mu = 0, conf.level
= 0.95)
#t.test on normalized data
#range of pvals: 0.0000000-0.9999669

setDT(allMT, keep.rownames = FALSE, key = NULL,check.names = FALSE)
MT <- cbind.data.frame(allM, allMT, deparse.level = 1, make.row.names = TRUE)
#combine tables containing array data and t.test data

MTsig <- subset(MT, pvalue <= 0.05)
#subset to only keep significant data (pvalues <= 0.05)

MTOrder <- MTsig[order(MTsig$difference),]
#sort data in order based on the difference

length(which(MTOrder$difference >= 1))
#9952 values (36.8%) >=1 (implies >= 2.7 fold change)
length(which(MTOrder$difference <= -1))
#10809 values (40%) <= -1

length(which(MTOrder[,43]<0))
#13224 (48.9%) of control genes were hypermethylated
length(which(MTOrder[,43]>0))
#13821 (51.1%) of control genes were hypomethylated
length(which(MTOrder[,44]<0))
#14468 (53.5%) of experimetnal genes were hypermethylated
length(which(MTOrder[,44]>0))
#12577 (46.5%) of experimental genes were hypomethylated

hist(MTOrder$controls, breaks = 150, main = "Control (normal folate levels)", xlab =
"HpaII/Msp1(log2)")
hist(MTOrder$experimentals, breaks = 150, main = "Experimental (folate deficient)", xlab =
"HpaII/Msp1(log2)")
#histograms of average control and average experimental mice

hist(MTOrder$difference, breaks = 150, main = "Folate deficiency induced methylation
changes", xlab = "HpaII/MspI(Log2) Difference")
#histogram of normalized ratio data

MTO <- data.frame(Probename = MTOrder[,2], difference = MTOrder[,45])
#make a new data frame containing probe names and difference data since R doesn't want to
make a histogram of my data table

```

```

hist(MTO[1:135,2], breaks = 100, main = "High gain of methylation", xlab =
"HpaII/MspI(Log2) Difference")
hist(MTO[26910:27045,2], breaks = 100, main = "High loss of methylation", xlab =
"HpaII/MspI(Log2) Difference")
#graphs representing the 1% most changed data (top 0.5% and bottom 0.5%)

library(dplyr)
MTO$class <- case_when(MTO$difference >=1 ~ "dec", MTO$difference <=-1 ~ "inc",
MTO$difference <1 & MTO$difference >-1 ~ "ns")
#assign all probe values a name based on whether they have a greater than or less than 2.7 fold
change
MTO$pval <- MTO$order$pvalue
#add pvalue column
MTO$LogP <- -log10(MTO$pval)
#add a -Log10 pvalue column

library(plotly)

pal <- c("red", "green", "black")

p <- plot_ly(type = "scatter", data = MTO, x = MTO[,2], y = MTO[,5], mode = "markers", color
= MTO[,3], colors = pal, alpha = 0.5)
layout(p, xaxis = list(showgrid = FALSE), yaxis = list(showgrid = FALSE))
#graph that does not include bad pvalues

MT$class <- case_when(MT$difference >=1 ~ "dec", MT$difference <=-1 ~ "inc",
MT$difference <1 & MT$difference >-1 ~ "ns")
MT$LogP <- -log10(MT$pvalue)
MTO$MostName <- c(1:135 == "dec", 136:26909 == "ns", 26910:27045 == "inc")
#tried to add names for top and bottom 0.5% of data, but it didnt work

pa <- plot_ly(type = "scatter", data = MT, x = MT$difference, y = MT$LogP, mode = "markers",
color = MT$class, colors = pal, alpha = 0.5)
layout(pa, xaxis = "Change in HpaII/MspI", yaxis = "-log10(pval)")
#graph that does include bad pvalues

```