A LOW VITAMIN B12 INDUCED TRANSCRIPTIONAL MECHANISM THAT REGULATES METABOLIC ACTIVITY OF THE METHIONINE/S-ADENOSYLMETHIONINE CYCLE IN CAENORHABDITIS ELEGANS

A Dissertation Presented

By

Gabrielle Giese

Submitted to the Faculty of the University of Massachusetts Graduate School of Biomedical Sciences, Worcester in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

July 6, 2021
SYSTEMS BIOLOGY
A LOW VITAMIN B12 INDUCED TRANSCRIPTIONAL MECHANISM THAT REGULATES METABOLIC ACTIVITY OF THE METHIONINE/S-ADENOSYLMETHIONINE CYCLE IN CAENORHABDITIS ELEGANS

A Dissertation Presented
By

Gabrielle Giese

This work was undertaken in the Graduate School of Biomedical Sciences
Interdisciplinary Graduate Program
Under the mentorship of

________________________________________
A.J. Marian Walhout, PhD; Thesis Advisor

________________________________________
Mark Alkema, PhD; Member of Committee

________________________________________
Cole Haynes, PhD; Member of Committee

________________________________________
Dohoon Kim, PhD; Member of Committee

________________________________________
Benjamin Tu, PhD; External Member of Committee

________________________________________
Amy Walker, PhD, Chair of Committee

________________________________________
Mary Ellen Lane, Ph.D.,
Dean of the Graduate School of Biomedical Sciences

July 6, 2021
DEDICATION

I dedicate this thesis to the billions of worms that have involuntarily sacrificed themselves in the name of science throughout my degree.

“*Ihr habt den Weg vom Wurme zum Menschen gemacht,*

*und Vieles ist in euch noch Wurm. ”*

- Friedrich Nietzsche
ACKNOWLEDGEMENTS

None of this, literally and figuratively, would be possible without my parents. Despite my gradual evolution into becoming a crazy cat lady who walks alone, and all places are alike to her, you have continued to support me above and beyond all reasonable expectation. You have been and always shall be my [best] friends.

First, I want to acknowledge my advisor, Marian Walhout, who believed in me when I started out as a tech and helped grow my confidence as a researcher. I am undoubtably a better scientist thanks to her guidance and mentorship. I wanted to get my degree in her lab because I had the opportunity to see how she shaped several graduate students into excellent scientists. I knew that if anyone could teach me how to correctly plan an experiment, communicate my science and most importantly, ask a question, it would be her. No matter what happens I hope we will always be able to ride horses together in the future.

There are far too many people to acknowledge without writing another chapter to this thesis. I could not have even attempted graduate school without Lesley MacNeil, who taught me so many valuable skills in the lab and has been a strong ally to me. I have always admired her sharp wit, creativity and huge heart. I especially miss having Juan Fuxman Bass in the lab to ask advice from and tease about competing in the math Olympics. Ashlyn Ritter and Emma Watson inspired me to apply to graduate school. I am incredibly appreciative of Ashlyn’s ingenuity in her science and also her championship of me. I must also thank Jote Bulcha,
who was a good friend to me for many years. In all that time I never once heard Jote utter a bad word about anyone. I wish him the very best in his new line of research. My relationship with Aurian Garcia-Gonzalez had a rocky start but I’m really grateful that it turned around. Auri is a very rare person who is always willing to give the shirt off her back to anyone who needs help. She made my life in the lab easier and has always been willing to help me and anyone else who asked. I have no doubt she’s going to go on to do great things and save many lives in the process.

I am incredibly grateful to the patience and generosity of Olga Ponomarova, Hefei Zhang and Xuhang Li. The fact that Olga did not strangle me while teaching me how to run the GC-MS is a testament to her character. I also have greatly appreciated all the times she’s lent me her ears and brain when I wanted to discuss my project or horses (which she pretended to care about graciously). Fei has also been generous with his time and molecular biology expertise and has been exceptionally tolerant of my American humor. As if there was any doubt, Hang is going to make an excellent mentor and professor someday thanks to his talent to carefully explain difficult concepts and his ability to point you in the right direction without giving away all the answers.

I want to also thank Amy Holdorf for being a friend and confidant and shoulder to cry on from time to time. I am thankful to work with Melissa Walker, who never says no and always gives her absolute best. In particular I have been so grateful that she has been able to continue the metabolite project while I worked
on this thesis. It’s always a special treat when she makes an unexpected joke. I also wish to acknowledge everyone else in the lab and I’m sorry I don’t have more time to do so in detail: Zeynep Mirza (prettier and smarter than Gal Gadot), Shivani Nanda (thanks for not suffocating me with a pillow when I snored and woke you up that time we shared a room at the CCSB retreat), Brent Horowitz (totally saved my butt with that last lab meeting), Sushila Bhattacharya (the lab’s biggest heart), Yonk Uk (stay weird my friend), Nananana BATMAN and the amazing Amy Levasseur.

This thesis would not have been possible if not for the generosity of several members from other labs. Thomas Leete of the Lee Lab and Sunil Guharajan of the Brewster lab have always been willing to answer my (sometimes idiotic) questions about molecular biology and scriping, respectively. I would not have been able to do the EMS screen and whole genome sequencing without the help of Johan Gibcus from the Dekker lab and Takao Ishidate from the Mello lab. There was nothing to gain from donating their time and expertise to me and yet these generous people did, and I will be forever grateful.

Three men in my life require special commendations. Safak Yilmaz has been an exceptional friend and colleague. He has patiently answered many of my inane questions and taught me so much about metabolic networks. I especially appreciate arguing with Safak because even when you’re right and he’s wrong, he still wins the argument. And thanks for all the Turkish delight and tea! I cannot thank Alejandro Montenegro-Montero enough for being a friend, laughing at my
jokes (good and bad) and answering so many questions patiently. I still won’t marry you though, no matter what kind of awesome food truck you’re driving. Cédric Diot has been a life saver and dearest friend to me since he joined the lab. I’m so grateful that he has been here to talk science to and gossip with over the past few years. Not only is Cédric exceptionally talented and sharp as a tack, he is incredibly generous with his expertise and has helped me many times especially with my writing and presentations. This thesis would not have been possible without him. Obviously. For real.

Finally, I want to thank my barn friends who’ve been a wonderful support network over the past years. Thanks specially to Kathy Cohan for being an amazing auntie to Sol.
Cells must regulate their metabolism in order to grow, adapt to changes in nutrient availability and maintain homeostasis. Flux, or the turnover of metabolites, through the metabolic network can be regulated at the allosteric and transcriptional levels. While study of allosteric regulation is limited to biochemical examination of individual proteins, transcriptional control of metabolism can be explored at a systems level. We endeavored to elucidate transcriptional mechanisms of metabolic flux regulation in the model organism *Caenorhabditis elegans* (*C. elegans*). We also worked to create a visual tool to explore metabolic pathways that will support future efforts in the research of metabolic gene regulation. *C. elegans* is a small, free-living nematode that feeds on bacteria and experiences a high level of diversity in nutrient level and composition. Previously, we identified a mechanism by which the essential cofactor, vitamin B12, regulates the expression of genes involved in the degradation of propionate, referred to as B12-mechanism-I. This mechanism functions to prevent the toxic accumulation of propionate and requires the TFs NHR-10 and NHR-68. Using genetic screens as well as transcriptomic and metabolomic approaches, we discover a second mechanism by which vitamin B12 regulates metabolic gene expression: B12-mechanism-II. Unlike B12-mechanism-I, B12-mechanism-II is independent of propionate, requires the transcription factor NHR-114 and functions to maintain the metabolic activity of the Methionine/S-adenosylmethionine cycle in a tightly
regulated regime. We also present WormPaths, an online resource that allows visualization of *C. elegans* metabolic pathways and enables metabolic pathway enrichment of user-uploaded transcriptomic data.
# TABLE OF CONTENTS

**DEDICATION** ........................................................................................................................................ iii

**ACKNOWLEDGEMENTS** ........................................................................................................................ iv  

**ABSTRACT** ........................................................................................................................................... viii

**LIST OF TABLES** ................................................................................................................................. xiii

**LIST OF FIGURES** ................................................................................................................................. xiv

**LIST OF SUPPLEMENTARY FILES** .................................................................................................... xvi

**COPYRIGHT MATERIAL** ....................................................................................................................... xviii

## CHAPTER I: INTRODUCTION

Preface ...................................................................................................................................................... 1

Introduction ............................................................................................................................................... 1

The nematode *C. elegans* .......................................................................................................................... 4

The *C. elegans* gene regulatory network ................................................................................................. 5

Transcriptional rewiring of *C. elegans* metabolism: a propionate persistence detector .......................................................... 7

How do metabolic networks communicate with gene regulatory networks? 10

The *C. elegans* metabolic network model and metabolic pathways maps ............................................ 12

Conclusions and future perspectives ....................................................................................................... 14

## CHAPTER II: VITAMIN B12 AND ITS ROLE IN METABOLISM

Summary ...................................................................................................................................................... 16

Introduction ............................................................................................................................................... 17

The biosynthesis of vitamin B12 .................................................................................................................. 18

Transport and processing of vitamin B12 .................................................................................................. 24

Metabolic functions of vitamin B12 ........................................................................................................... 27

  *The propionate degradation pathway* .................................................................................................. 29

  *The Met/SAM cycle* ............................................................................................................................... 31

S-adenosylmethionine (SAM) ...................................................................................................................... 32

  *SAM and transmethylation* ....................................................................................................................... 33

  *SAM and the production of phosphatidylcholine* .................................................................................. 35

  *SAM and the transsulfuration pathway* ................................................................................................. 39

Regulation of the Met/SAM cycle ........................................................................................................... 42
Life without vitamin B12

Bacteria

Yeast

Flies

Conclusion

CHAPTER III: CAENORHABDITIS ELEGANS METHIONINE/S-ADENOSYL-METHIONINE CYCLE ACTIVITY IS SENSED AND ADJUSTED BY A NUCLEAR HORMONE RECEPTOR

Preface

Summary

Introduction

Results

Low dietary vitamin B12 activates two transcriptional mechanisms

Met/SAM cycle perturbations activate B12-mechanism-II

B12-mechanism-II activates Met/SAM cycle gene expression in response to Met/SAM cycle perturbations

nhr-114 mediates Met/SAM cycle activation in response to low Met/SAM cycle activity

Methionine and choline supplementation suppress B12-mechanism-II

B12-mechanism-II activates influx and represses efflux of the Met/SAM cycle

Discussion

Materials and methods

C. elegans strains

Bacterial strains

GFP intensity measurement using image analysis

EMS screen

Mutant mapping

Whole genome sequencing

RNAi screen

Expression profiling by RNA-seq

Expression profiling by qRT-PCR

Body size measurement

Gas chromatography-mass spectrometry
CHAPTER IV: WORMPATHS: CAENORHABDITIS ELEGANS METABOLIC PATHWAY ANNOTATION AND VISUALIZATION

Preface................................................................................................................................. 96
Summary.................................................................................................................................. 96
Introduction.............................................................................................................................. 97
Results..................................................................................................................................... 101
Assigning C. elegans metabolic genes to pathways at different levels................................. 101
WormPaths maps – visualization and navigation................................................................. 111
WormPaths advantages.......................................................................................................... 117
WormPaths levels can be used for pathway or gene set enrichment analysis.......................... 123
Discussion and vision............................................................................................................. 127
Materials and methods.......................................................................................................... 128
Design of pathway maps........................................................................................................ 128
Illustration of pathway maps.................................................................................................. 129
Pathway enrichment analysis............................................................................................... 130
Metabolite structures............................................................................................................ 131

CHAPTER V: CONCLUSIONS

Introduction............................................................................................................................ 133
Multiomics approaches and pathway tools to study transcriptional regulation of metabolic flux......................................................................................................................... 134
Transcriptional regulation of Met/SAM cycle activity......................................................... 136
Met/SAM cycle activity is sensed and regulated by a nuclear hormone receptor............... 137
Why is acdh-1 transcriptionally activated by B12-mechanism-II?........................................ 140
Conclusion.............................................................................................................................. 143

BIBLIOGRAPHY.................................................................................................................... 145
LIST OF TABLES

CHAPTER II
2.1: Vitamin B12 related genes and pathways of model organisms and humans................................................................. 46

CHAPTER III
3.1: C. elegans strains used in this project........................................... 87
3.2: qPCR primers used in this project.................................................. 93

CHAPTER IV
4.1: Pathways at levels 1 through 4.......................................................... 104
4.2: All maps and the corresponding level to which each map was drawn 108
# LIST OF FIGURES

## CHAPTER I

1.1: Interaction between transcription and metabolism.............................. 3
1.2: *Caenorhabditis elegans* can be a powerful tool to interrogate metabolic and transcriptional cross talk *in vivo*................................................................. 6
1.3: Activation of the propionate degradation shunt pathway is controlled by a persistence detector..................................................................................... 9
1.4: Metabolism is known to influence transcription by three different mechanisms............................................................................................................. 11

## CHAPTER II

2.1: The chemical structure of vitamin B12................................................ 20
2.2: Biosynthetic pathways of vitamin B12.................................................. 22
2.3: Metabolic pathways connected to vitamin B12...................................... 28
2.4: Enlarged lipid droplets cause by a decrease in phosphatidylcholine due to mutation of sam-1 in *C. elegans*................................................................. 36
2.5: Allosteric regulation of Met/SAM cycle genes by Met/SAM cycle metabolites............................................................................................................. 43

## CHAPTER III

3.1: Two mechanisms of gene regulation by low vitamin B12 dietary conditions......................................................................................................................... 55
3.2: Met/SAM cycle perturbations activate B12-mechanism-II...................... 61
3.3: Genetic mutations in Met/SAM cycle genes activate Met/SAM cycle gene expression............................................................................................................. 64
3.4: *nhr-114* is required for activation of *Pacdh-1::GFP;Δnhr-10* mutant reporter in response to Met/SAM cycle perturbations................................. 68
3.5: nhr-114 is required for B12-mechanism-II in response to Met/SAM cycle perturbations................................................................. 70
3.6: Methionine and choline supplementation suppress vitamin B12-
mechanism-II................................................................................. 74
3.7: Low SAM and not phosphatidylcholine activates B12-mechanism-II.. 78
3.8: nhr-114 transcriptionally regulates Met/SAM cycle influx and efflux... 81
3.9: Model of B12-mechanism-II....................................................... 83

CHAPTER IV
4.1: WormPaths annotation of C. elegans metabolic genes.................. 103
4.2: An example of using WormPaths to search for a specific gene......... 112
4.3: WormPaths examples................................................................ 115
4.4: Tissue-specific expression of pck-1 and pck-2.............................. 116
4.5: WormPaths provides easy to navigate C. elegans-specific maps...... 119
4.6: WormPaths maps provides additional reactions to metabolic
pathways......................................................................................... 121
4.7: WormPaths maps clean up pathway associations for individual
genes............................................................................................... 123
4.8: Pathway enrichment analysis using WormPaths levels................. 126

CHAPTER V
5.1: Proposed mechanisms by which SAM could be sensed by
NHR-114.......................................................................................... 138
5.2: A model of metabolites tuning the transcriptional response............ 140
5.3 Two possibilities that might explain up-regulation of acdh-1 by Met/SAM
cycle perturbation........................................................................ 142
LIST OF SUPPLEMENTARY FILES

Supplementary file 3.1
mRNA fold change and p adjusted values of genes involved in the propionate shunt, Met/SAM cycle and phospholipid metabolism in the following conditions. Wild type on vitamin B12 vs. wild type on untreated; wild type on vitamin B12 and propionate vs. wild type on vitamin B12; \( \Delta nhr-10 \) untreated vs. wild type untreated; \( \Delta nhr-10 \) on vitamin B12 vs. \( \Delta nhr-10 \) untreated; \( \Delta nhr-68 \) untreated vs. wild type untreated; \( \Delta nhr-68 \) on vitamin B12 vs. \( \Delta nhr-68 \) untreated. This table is a subset of data published in Bulcha et al., 2019 (GSE123507).

Supplementary file 3.2
Genes with an mRNA fold change greater or equal to 1.5 and a p adjusted value less than 0.01 in Met/SAM cycle mutants vs. the parental strain \( Pacdh-1::GFP;\Delta nhr-10 \). Genes that change in all four mutant strains are shown in tab 1, while all genes regardless of p adjusted value are shown in tab 2.

Supplementary file 3.3
mRNA fold change and p adjusted values of all detected genes in the following conditions. \( Pacdh-1::GFP \) on \( nhr-114 \) RNAi vs. vector; \( Pacdh-1::GFP;\Delta metr-1 \) on vector vs. \( Pacdh-1::GFP \) on vector; \( Pacdh-1::GFP;\Delta metr-1 \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP \) on vector; \( Pacdh-1::GFP;\Delta metr-1 \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP;\Delta metr-1 \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP;\Delta metr-1 \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP;\Delta metr-1(ww52) \) on vector vs. \( Pacdh-1::GFP \) on vector; \( Pacdh-1::GFP;\Delta nhr-10;\Delta metr-1(ww52) \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP \) on vector; \( Pacdh-1::GFP;\Delta nhr-10;\Delta metr-1(ww52) \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP \) on vector; \( Pacdh-1::GFP;\Delta nhr-10;\Delta metr-1(ww52) \) on \( nhr-114 \) RNAi vs. \( Pacdh-1::GFP;\Delta nhr-10;\Delta metr-1(ww52) \) on vector.

Supplementary file 3.4
mRNA fold change and p adjusted values of all detected genes in the following conditions. \( Pacdh-1::GFP \) on methionine vs. \( Pacdh-1::GFP \) untreated; \( Pacdh-1::GFP;\Delta nhr-114 \) untreated vs. \( Pacdh-1::GFP \) untreated; \( Pacdh-1::GFP;\Delta nhr-114 \) on methionine vs. \( Pacdh-1::GFP \) on methionine; \( Pacdh-1::GFP;\Delta nhr-114 \) on methionine vs. \( Pacdh-1::GFP;\Delta nhr-114 \) untreated.

Supplementary file 4.1
Gene sets per each pathway by level by gene name
Supplementary file 4.2

Gene sets per each pathway by level by WormBase ID

Supplementary file 4.3

All pathway associations listed by gene.
COPYRIGHT MATERIAL

• Chapter I was adapted from the following article and according to the publisher’s copyright policies, as the author I may include it in my thesis:


• Chapter III was adapted from the following article under CC BY 4.0:


• Chapter IV was adapted from the following article under CC BY 4.0:

CHAPTER I: INTRODUCTION

Preface

This introductory chapter is adapted from a perspective written by myself, Shivani Nanda, Amy Holdorf and Marian Walhout titled “Transcriptional regulation of metabolic flux: A Caenorhabditis elegans perspective”, published in the journal *Current Opinion in Systems Biology* in 2019¹.

Introduction

Metabolic reactions form the basis of life to generate biomass, produce energy and eliminate waste. Together, metabolic reactions function in the context of a highly interconnected metabolic network. Flux through the metabolic network is heavily regulated at different levels to ensure homeostasis under changing nutritional or environmental conditions. Changes in metabolism can result in a number of diseases such as obesity and diabetes, and altered metabolism also plays a key role in cancer where the metabolic network is rewired to increase cell proliferation and angiogenesis²–⁴.

The two major regulatory mechanisms of metabolic flux are allosteric regulation of enzyme activity and transcriptional control of enzyme levels (*Figure 1.1*). Allostery has long been appreciated as a major regulatory mechanism of metabolism where metabolic enzyme activity is modulated by physical interactions with metabolites. However, recent analyses reveal that large fractions of the
transcriptome respond to changes in metabolism\textsuperscript{5,6}, and therefore the transcriptional control of metabolic gene expression through the action of specific transcription factors (TFs) is prevalent. For instance, many metabolic genes change in expression in a circadian manner to enable appropriate metabolism in our 24-hour cycle\textsuperscript{7,8}. While allostery is mechanistically relatively simple, the transcriptional control of metabolism is more complicated because metabolism and transcription mostly occur in different cellular compartments; the vast majority of metabolic enzymes are encoded by nuclear genes, while most metabolic reactions occur outside the nucleus, in the cytoplasm, mitochondria, or other organelles. Therefore, a change in metabolism has to be relayed to the relevant TFs, resulting in TF relocalization and/or activation. A well-known example of this process in mammals is the sterol regulatory element-binding protein 2 (SREBP2), a TF that functions to maintain cholesterol levels and is vital to lipid and membrane homeostasis. SREBP2 is integrated within ER membrane. When cholesterol levels are low, SREBP2 is released from the ER and translocates to the nucleus where it activates the transcription of cholesterol biosynthesis genes\textsuperscript{9}. Recent discoveries have been made about metabolic reactions occurring in the nucleus. For instance, nuclear pyruvate dehydrogenase synthesizes acetyl-CoA used to modify histones\textsuperscript{10}. Future studies will determine the relative contribution of acetyl-CoA produced in different cellular compartments to biomass generation (e.g., lipids), energy production via the TCA cycle, and histone modifications in the nucleus.
Figure 1.1 – Interaction between transcription and metabolism. Nutrients are taken in from the environment and converted into energy or biomass. Regulating metabolic flux is essential to maintain metabolite pools in balance. This can be accomplished allosterically or transcriptionally. Metabolites can bind to transcription factors (TFs), which go on to regulate target genes. In addition, metabolism also produces cofactors of chromatin-modifying enzymes that regulate chromatin state and, therefore, transcription.

Both metabolism and gene regulation are orchestrated at the network level. The metabolic network is the collection of metabolic reactions that convert nutrients and cellular metabolites into biomass building blocks and energy. The gene regulatory network is comprised of genes and their cis-regulatory DNA elements, and trans-acting regulatory proteins such as TFs. By physically binding cis-regulatory DNA elements, TFs can activate or repress expression of their target genes. While much progress has been made to understand the transcriptional regulation of individual metabolic genes, little is known about this process at a
systems level. Here, we discuss metabolic and gene regulatory networks and the communication between them. We focus on recent progress made with the nematode C. elegans, which has emerged as a highly tractable multicellular model organism with which it is possible to combine experimental gene regulatory network analysis with computational modeling of metabolic flux.

The nematode C. elegans

The hermaphrodite C. elegans has an invariant body plan of 959 somatic cells in which the lineage of each cell has been determined. It is a bacterivore that can easily be maintained in the laboratory on different bacterial diets. Its small size, short life cycle, and transparent body make it highly tractable for large-scale genetic studies and phenotypic analysis. The C. elegans genome encodes ~20,000 proteins, many of which are conserved from nematode to humans. C. elegans can be fed double-stranded RNA-containing bacteria to induce RNA interference, making large-scale RNAi screens both easy and cost-effective. Transcriptional and translational fluorescent protein reporters can be used to assess gene regulation and protein localization in vivo. When combined with RNAi, such reporters can be used to gain insight into transcriptional regulation of promoter activity, or the metabolic rewiring of gene expression. Although bacteria and yeast are well established as excellent models for studying regulation of metabolism, as a multicellular organism, C. elegans provides the opportunity to study cell non-autonomous regulation as well as tissue-specific gene expression.
Given the tractability of the organism, the homology with human genes, and the availability of genome scale or specific RNAi collections, *C. elegans* makes an ideal model system to elucidate the relationships between metabolic and gene regulatory networks.

**The *C. elegans* gene regulatory network**

Gene regulatory networks are comprised of genes and their regulators\(^\text{17}\). At the transcriptional level, these regulators include TFs, chromatin factors and cofactors. So far, most research in this field, including in *C. elegans*, has focused on the identification of TF binding sites\(^\text{18,19}\), and the mapping of physical TF-DNA interactions with larger regulatory genomic regions such as promoters by chromatin immunoprecipitation\(^\text{20}\) or yeast one-hybrid assays\(^\text{19,21,22}\). Much less is known about regulatory, or functional, interactions between regulatory DNA and TFs, and where in the animal these interactions occur. *C. elegans* lends itself to tissue-relevant mapping of regulatory interactions between genes and TFs wherein the effect of TF perturbation on promoter activity is measured. For instance, 19 intestinal reporter strains have been used to delineate a true regulatory network in which promoter activity changes were recorded specifically in the animal's intestine in response to RNAi of each individual TF\(^\text{15}\). Surprisingly, there was only modest overlap between regulatory and physical TF interactions. This indicates that many TFs affect gene expression indirectly, via the regulation of other types of molecules. Since the intestine is a highly metabolic organ, it is
tempting to speculate that TF knockdown can result in metabolic changes that affect gene expression. For instance, perturbation of a TF could result in a change in the expression of a metabolic gene. This change can then lead to a change in the concentration of a specific metabolite that, in turn, relays this information to another TF that then directly regulates the expression of the reporter used (Figure 1.2). Several examples of metabolites and enzymatic cofactors regulating C. elegans metabolic gene expression are known, including vitamin B12, the short-chain fatty acid propionate, zinc and nitric oxide.

Figure 1.2 – Caenorhabditis elegans can be a powerful tool to interrogate metabolic and transcriptional cross talk in vivo. The effect of transcription factor RNAi can be observed through the use of intestine-specific transcriptional reporters. One possibility is RNAi knockdown of a transcription factor prevents the expression of a metabolic enzyme required to produce a metabolite that activates the transcription of the target reporter either by binding to another transcription factor or by influencing chromatin state. dsRNA, double-stranded RNAi, RNA interference; TF, transcription factor.
Transcriptional rewiring of *C. elegans* metabolism: a propionate persistence detector

Propionate is a short-chain fatty acid that is produced upon the breakdown of the branched chain amino acids valine and isoleucine, the amino acids methionine and threonine, as well as odd-chain fatty acids and cholesterol. In both *C. elegans* and humans, propionate is toxic when it accumulates\(^{27,28}\). Like humans, *C. elegans* uses vitamin B12 as a cofactor in a canonical propionate breakdown pathway, and under B12-replete dietary conditions propionate is converted to succinate, which anaplerotically can enter the TCA cycle. Remarkably, however, *C. elegans* is not only able to grow on bacteria that synthesize vitamin B12, they can also thrive on bacteria that do not\(^{23,29}\). Indeed, the vast majority of research on *C. elegans* has been done with animals fed the standard laboratory diet of *E. coli* OP50, a strain that cannot synthesize vitamin B12. On bacterial diets low in vitamin B12, *C. elegans* transcriptionally activates the expression of five metabolic genes that function together in an alternate propionate breakdown, or shunt pathway\(^{27}\). The first gene in this pathway is virtually off when vitamin B12 levels are sufficiently high. This indicates that the animal favors the canonical, vitamin B12-dependent propionate breakdown pathway over using the shunt. The propionate shunt may be less favorable since the first intermediate in the pathway is acrylyl-CoA, which after removal of the CoA is converted into highly toxic acrylate. This idea is substantiated by the observation that knockdown of the enzyme that consumes acrylyl-CoA, *ech-6*, renders the animals very sick. Double perturbation of *ech-6*
with the enzyme that produces acrylyl-CoA, *acdh-1*, however, rescues this phenotype\(^27\).

Recently, it was found that a specific type of regulatory circuit known as a type I coherent feedforward loop with an AND-logic gate is required for the activation of propionate shunt gene expression\(^24\). This circuit consists of two nuclear hormone receptor TFs: NHR-10 and NHR-68, both of which are required for propionate shunt activation. In addition, NHR-10 activates the expression of NHR-68, and NHR-68 autoactivates, which likely provides tunability to the system (Figure 1.3). Propionate shunt expression is activated with a delay of several hours, and only when propionate supplementation is sustained. This is precisely what modeling of this type of circuit predicted more than 15 years ago\(^30\). Here, the delay is caused by the time it takes to express sufficient levels of NHR-68, which is under the control of NHR-10, and because sufficient levels of both TFs are required. Therefore, this circuit functions as a true persistence detector to rewire propionate breakdown metabolism in an animal only when propionate accumulates persistently. However, the molecular mechanism by which propionate accumulation is relayed to NHR-10 and, perhaps, NHR-68 remains to be elucidated.
Figure 1.3 – Activation of the propionate degradation shunt pathway is controlled by a persistence detector. The nuclear hormone receptor transcription factors NHR-10 and NHR-68 act in a type I feedforward loop with an AND logic gate that causes a delayed response of the propionate shunt pathway only when there is a persistent, not a transient, level of propionate.

Vitamin B12 is also required by another metabolic pathway in the cell: the one-carbon cycle, which is comprised of two sub-cycles: the methionine/S-adenosylmethionine (Met/SAM) cycle and the folate cycle. This pathway was shown to regulate the development acceleration caused by vitamin B12 in *C. elegans*\textsuperscript{23}. We have previously found that in conjunction with phenotypic changes, many gene expression changes also occur in animals fed a diet replete with vitamin B12 compared to those fed a vitamin B12 deficient diet\textsuperscript{29}. These results suggest that there may be additional transcriptional mechanisms by which vitamin B12 affects the expression of metabolic genes.
How do metabolic networks communicate with gene regulatory networks?

There are several mechanisms by which metabolic and gene regulatory networks can communicate (Figure 1.4). First, metabolites can directly interact with TFs. Among the best-studied examples are NHRs, which can regulate gene expression upon physically binding ligands such as fatty acids\(^{31}\). Notably, the *C. elegans* genome encodes more than 270 NHRs, while the human genome only encodes 48\(^{32,33}\). Interestingly, NHRs often physically interact with the promoters of metabolic genes, including those involved in compound detoxification\(^{19,34}\). The transcriptional and physiological function has been characterized for only a small subset of NHRs. For instance, NHR-49 has been shown to modulate the shift between fat consumption for energy and fat composition for storage by activating *acs*-2 and *ech*-1, genes involved in mitochondrial β-oxidation, and *fat*-7, a gene that encodes a fatty acid desaturase\(^{35}\). Other examples include NHR-8, which regulates cholesterol homeostasis, bile acid production and fat metabolism\(^{36}\), and NHR-25, an important factor in *C. elegans* development and molting, which responds to long-chain fatty acids produced by *acs*-3 and regulates fat uptake and storage\(^{37}\). However, only two NHRs, DAF-12 and NHR-33/HIZR-1, have known ligands. Under growth conducive conditions, insulin/TGF-β pathways result in the production of the bile acid-like dafachronic acid, which directly binds as a ligand to DAF-12\(^{38}\). Once dafachronic acid is bound, DAF-12 activates growth programs and represses dauer formation. NHR-33/HIZR-1 directly binds zinc and accumulates in the nucleus where it activates zinc detoxifying and transporting
enzymes. It would be interesting to see if the persistence detector mentioned above is activated by a physical interaction between NHR-10 and propionate. Studying this putative interaction is challenging, however, because propionate is volatile and very small, with only three carbons.

**Figure 1.4 – Metabolism is known to influence transcription by three different mechanisms.** (I) Metabolites produced by the metabolic network might act as ligands that activate or repress transcription factors. (II) Metabolites can also act as cofactors of histone modifiers, causing changes in chromatin state. (III) Metabolic enzymes themselves have also been shown to bind to DNA and cause changes in transcription directly.

The second mechanism by which the metabolic network can communicate with the gene regulatory network is via the modification of chromatin. Examples of these include histone methylation, acetylation and crotonylation. A recent study found that the metabolite S-adenosylmethionine is required for H3K4me3-mediated transcriptional responses to *C. elegans* infection by the pathogenic
bacterium *Pseudomonas aeruginosa*\(^4\). Finally, it has been shown that metabolic enzymes themselves can regulate transcription, for example by binding DNA. For instance, the yeast metabolic enzyme Arg5,6, which acts in the arginine biosynthesis pathway, has been shown to bind a specific DNA sequence and affect transcription\(^4\). While the binding to DNA and regulation of gene expression by metabolic enzymes has not yet been directly shown in *C. elegans*, one metabolic enzyme, ECH-6, has been found to bind the *nhr-68* promoter in yeast one-hybrid assays\(^19\). Future studies are required to validate this finding and to determine whether it provides an additional layer of regulation to the propionate persistence detector.

**The *C. elegans* metabolic network model and metabolic pathway maps**

Metabolic network models mathematically convert nutrients into biomass and energy. A genome-scale metabolic network model links metabolites as nodes via enzymatic reactions as edges. The chemical reactions in this network have been elucidated one reaction at a time, through years of biochemical and genetic studies. Subsequently, specific enzymes have been linked to each metabolic reaction, and the corresponding genes for these enzymes have been identified. However, the experimental connections of genes and proteins to reactions have been done for only a limited number of model organisms, and such connections are extrapolated to other organisms by homology-based computational annotations to generate what is referred to as ‘metabolic network
reconstructions\textsuperscript{42,43}. Genome-scale metabolic network models are more than just gene-protein-reaction associations; when done correctly, they provide mathematical models that can be used with flux balance analysis (FBA) to model metabolism \textit{in silico}. FBA is a constraint-based mathematical method that both utilizes and reinforces genome-scale metabolic models to predict metabolic phenotypes. The \textit{C. elegans} genome encodes more than 3,000 predicted metabolic enzymes. Of these, 1,273 are incorporated in the first genome-scale metabolic network model that contains 1,985 reactions and 887 unique metabolites, and that is fully compatible with FBA\textsuperscript{44}. It was shown that this network is highly predictive of phenotypes associated with metabolic genes. Moreover, it could be effectively combined with metabolic gene mRNA expression profiles to predict metabolic states associated with a hibernating \textit{C. elegans}, known as the dauer state. Hundreds of metabolic genes are not yet associated with a specific reaction, or the reactions they are associated with are not yet connected to the genome-scale model due to other missing reactions or unelucidated pathways. Importantly, since the gene-protein-reaction annotations in \textit{C. elegans} are based on homology to known metabolic enzymes in other organisms, and because many paralogous enzymes share high levels of protein sequence homology, it is likely that the genes associated with a reaction and all reactions associated with a gene will need to be revised continuously based on new experimental evidence.

Another powerful tool are visual metabolic maps showing the relationship between metabolic genes, metabolites and the reactions in which they’re involved.
This type of map can also help demarcate pathway boundaries and thus be used to find categorical enrichment of gene expression data. While online databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG)\textsuperscript{45} and MetaCyc\textsuperscript{46} are extremely useful they have been automatically generated and thus may suffer from a lack of detail. We set out to create a collection of \textit{C. elegans} specific metabolic pathway maps relying heavily on the \textit{C. elegans} metabolic model and literature searches. The resulting tool, WormPaths, provides not only a visual representation of the metabolic network but also a means to analyze metabolic gene expression data.

\textbf{Conclusions and future perspectives}

Metabolism influences gene expression, and \textit{vice versa}. To understand how this works at a network, or systems level, several challenges remain. First, we need to know the TFs that bind to and regulate each metabolic gene, and if and how these TFs respond to changes in metabolism. Second, we need to know the composition of the entire \textit{C. elegans} metabolome, and which metabolites can act as signaling molecules to interact with the gene regulatory network. Third, it will be important to identify ligands for \textit{C. elegans} NHRs and to connect these to gene regulatory effects. Finally, large-scale transcriptomic, proteomic, and metabolomic data need to be experimentally generated using \textit{C. elegans} fed different diets, supplemented with different metabolites or perturbed for different TFs or metabolic enzymes. Computational methods could then be used to integrate these different data types.
For example, machine learning can be used to predict flux levels and metabolite concentrations. Recent work with budding yeast used machine learning to examine the relationships between metabolic flux and enzyme protein levels in a set of kinase mutants, indicating that such approaches will be invaluable for the integration of metabolic and gene regulatory networks\textsuperscript{45}.
CHAPTER II: VITAMIN B12 AND ITS ROLE IN METABOLISM

Summary

Vitamin B12 is an incredibly complex cofactor produced only by a handful of bacteria and archaea, and yet this vitamin is essential for several metabolic processes in many animal cells. In humans, vitamin B12 is required for the operation of two pathways: mitochondrial propionate degradation and recycling methionine in the cytosolic methionine/S-adenosylmethionine (Met/SAM) cycle. Deficiency of vitamin B12 or disruption of either of these two pathways leads to severe diseases such as anemias, neural tube defects, cardiovascular disease and osteoporosis. In fact, the folate cycle, which is linked directly to the Met/SAM cycle through the vitamin B12 dependent enzyme methionine synthase, was the first target of anti-cancer therapies. The Met/SAM cycle in particular is highly anabolic and is responsible for balancing nucleotide and energy production with methylation, protein synthesis and lipogenesis. The Met/SAM cycle also contributes to the redox state of the cell through the transsulfuration of methionine to glutathione. Consequently, allosteric and genetic regulation of the Met/SAM cycle is paramount to organismal health. This chapter reviews the biosynthesis and uptake of vitamin B12 and the function and regulation of the two metabolic pathways in which its involved.
**Introduction**

Vitamin B12 or cobalamin is an essential nutrient that is exclusively made by only a subset of bacteria and archaea. Despite its restricted synthesis, vitamin B12 is required for a functioning metabolism across the phylogenetic tree\(^4^8\) and its deficiency can lead to several serious diseases in humans\(^4^9\). In mammals, vitamin B12 is a cofactor of two enzymes: methylmalonyl-coA mutase (MUT, EC 5.4.99.2), which acts in the degradation of the short-chain fatty acid propionate and methionine synthase (MS, EC 2.1.1.13), which catalyzes the methylation of homocysteine to methionine in the Met/SAM cycle. The Met/SAM cycle is part of a larger and highly anabolic cycle known as the one-carbon cycle (1CC), which produces several important building blocks for cell growth and development such as nucleotides, proteins, energy and ultimately SAM, the primary methyl donor of the cell\(^5^0\).

The biochemistry of vitamin B12 has been studied for over sixty years and as a result much of its transport, processing as well as the allosteric regulation of vitamin B12 related metabolic enzymes in eukaryotic cells have been thoroughly described\(^5^1–5^3\). Its role in transcriptional regulation, however, has been largely understudied. Nevertheless, some evidence for the impact that vitamin B12 and related metabolism have on genetic and epigenetic regulation in several model organisms has recently emerged\(^2^4,5^4,5^5\). Investigating the transcriptional regulation of metabolic genes associated with vitamin B12-dependent pathways will help illuminate the crosstalk between metabolism and gene regulation as well as further
our understanding of related human diseases. In this chapter we discuss the function of vitamin B12 and its impact on cell metabolism, as well as the transcriptional regulation of related metabolic genes.

The biosynthesis of vitamin B12

‘Nature’s most beautiful cofactor’, vitamin B12, wasn’t isolated until 1947 by several research groups in the US and the UK. Its effect on human health, however, has been shown since the mid 19th century when Thomas Addison and several other renown physicians described cases of pernicious anemias. In 1926, George Whipple, George Minot, and William Murphey published their results showing that these anemias could be treated with liver extracts later understood to contain a large concentration of vitamin B12, thus earning the Nobel prize in Medicine and Physiology in 1934. Finally, in 1955, Dorothy Hodgkin solved the chemical structure of vitamin B12 by crystallography and was awarded the Nobel prize in Chemistry in 1964.

Requiring at least 25 enzymes for production, vitamin B12 is the largest and most complex of the known vitamins. It is comprised of three parts: a corrin ring with methyl, acetyl and propionyl sidechains on the periphery and a cobalt atom at the center, a nucleotide loop, and an axial ligand (Figure 2.1). The trivalent cobalt atom is coordinated to four nitrogen atoms in the corrin ring and a fifth nitrogen belonging to a dimethylbenzimidazole of the nucleotide loop that is linked to one of the corrin ring side groups via a phosphorylated sugar. The sixth
coordination site is known as the reaction center and can be one of four different axial ligands: a methyl group, 5-deoxyadenosine, a hydroxyl or a cyanide group. The cobalt atom can shift between its three oxidation states: cob(III)alamin, the most oxidized, cob(II)alamin and cob(I)alamin, the most reduced. These oxidation states correspond with the number of bonds the cobalt atom can coordinate thus altering its ability to donate the axial ligand, allowing vitamin B12 to participate in several different types of chemical reactions. While cyanocobalamin is the most stable form of vitamin B12 and is typically used in industrial compounds and medicinal supplements, humans, like most animals, only utilize the methyl and adenosyl forms of vitamin B12 but have the ability to interconvert these different forms.
Figure 2.1 – The chemical structure of vitamin B12. Vitamin B12 depicted here as adenosylcobalamin in its ‘base-on’ configuration. Additional β-axial ligands are shown at the top right.

The biosynthesis of vitamin B12 can be carried out aerobically or anaerobically\(^6\) (Figure 2.2). First, the corrin ring tetrapyrrole precursor uroporphyrinogen III (UroIII), is made either from glutamate in the C5 pathway or glycine and succinyl-coA in the C4 pathway. UroIII is the precursor not only to corrinoid but also heme, siroheme and chlorophyll. It is hypothesized, however, that this pathway originally synthesized cobalamin and was later adapted to produce heme, siroheme and chlorophyll\(^6\). After conversion of UroIII to precorrin-2
the anerobic and aerobic pathways diverge and then reconverge in the synthesis of adenosyl cobinamide (Ado-Cbi). The anerobic pathways inserts the cobalt ion earlier in the synthesis of Ado-Cbi than the aerobic pathway. Finally, the nucleotide loop is added from threonine and flavin precursors. In gram-negative bacteria and some archaea there is additionally a salvage pathway in which an adenosine triphosphate (ATP)-binding cassette transporter imports cobinamide into the cell to produce Ado-Cbi\(^65\) (Figure 2.2).
Figure 2.2 – Biosynthetic pathways of vitamin B12.
The vitamin B12 biosynthesis pathway is complex and energetically costly and therefore it is unsurprising that its activity is highly regulated. A key regulatory element of vitamin B12 biosynthesis and related metabolism is the vitamin B12 riboswitch. Riboswitches rely on two transcript-domains: an evolutionarily conserved ligand-binding aptamer domain and an expression platform. When a target metabolite binds to the aptamer domain secondary structural changes occur in the expression platform that for instance, prevent transcription by forming a terminator loop or prevent translation by sequestering the ribosome binding site. Many different types of riboswitches have been found in most bacteria species as well as plants, algae and fungi. Riboswitch ligands include certain amino acids, nucleotides, enzyme cofactors, other metabolites and even metal ions such as magnesium.

In many bacteria, the vitamin B12 riboswitch controls the expression of genes involved in the biosynthesis of vitamin B12 including those found in the cobalamin synthesis pathway, cobalamin transport, and cobalt metabolism and transport. Additionally, the vitamin B12 riboswitch regulates the expression of genes that encode vitamin B12-dependent enzymes, for example, those involved in ethanolamine metabolism. Typically, vitamin B12 riboswitches cause translational repression of their downstream genes in the presence of vitamin B12 by sequestration of the ribosome binding site due to conformational change of the expression platform domain. Some types of vitamin B12 riboswitches, however, restrict transcription by preventing an antiterminator loop from forming when
vitamin B12 is present and thus allowing the formation of an intrinsic terminator in the transcript\textsuperscript{79}. Although vitamin B12 riboswitches, like the majority of other riboswitches, have only been found in prokaryotes, plants and fungi, it is tempting to hypothesize that these otherwise conserved regulatory elements might be present in some animals and have simply gone undetected to this point.

**Transport and processing of vitamin B12**

In humans as with other animals, vitamin B12 cannot be produced *de novo* but instead must be taken up in the diet and then transported into cells. In the diet, vitamin B12 can only be found in meats and dairy products. Ruminants such as cows are strict herbivores living off plant matter that does not contain vitamin B12. Due to their four-chambered stomachs that contain many microorganisms however, cows can derive vitamin B12 directly from their gut microbiota\textsuperscript{80}. While non-ruminants also contain gut microbiota that produce vitamin B12, these bacteria primarily inhabit the large intestine and vitamin B12 is absorbed through the small intestine. Because the small intestine precedes the large intestine along the alimentary tract, any vitamin B12 generated by their gut bacteria is excreted and lost\textsuperscript{81}.

Even before full characterization of vitamin B12, another factor was identified as being required for rescue of pernicious anemia\textsuperscript{57,82}. At the time vitamin B12 was called extrinsic factor (because it’s derived from the diet) and this other element was and still is today called intrinsic factor (IF, because it’s produced
internally). IF is a glycoprotein secreted by the parietal cells that line the stomach and then travels to the duodenum where it binds recently digested free cobalamin. Vitamin B12 is bound by protective haptocorrin (transcobalamin I), also a type of glycoprotein, found in the salivary glands so that it may survive the stomach acid\(^{83}\). Haptocorrin-bound cobalamin then travels to the duodenum where the haptocorrin is digested by pancreatic proteases and the now free cobalamin binds IF. The cobalamin-IF complex can then be absorbed by cubilin receptors at the ileum of the small intestine\(^{84}\). Once in the enterocytes, cobalamin is released from IF and exported to the circulation where it binds transcobalamin II and can then be taken up by various tissues.

Just as the vitamin B12 biosynthetic pathways are complex, so too is its cellular transport and processing into mammalian cells. The process begins when vitamin B12 binds to either the ubiquitously expressed transcobalamin-receptor CD320\(^{85}\), or the renal-specific Megalin receptor on the cell surface\(^{86}\). Holotranscobalamin II is then endocytosed to the lysosome. Once in the lysosome, vitamin B12 is released and apo-transcobalamin is reprocessed back to the cell surface through the transcobalamin degradation pathway\(^{87}\). The now free vitamin B12 is transported out of the lysosome via two integral membrane proteins: ATP-binding cassette subfamily D member 4 (ABCD4)\(^{88}\) and lipocalin-1-interacting membrane receptor domain-containing 1 (LMBD1)\(^{89}\). Although it has been shown these two proteins interact, the mechanism of interaction and transport of vitamin B12 remains unknown\(^{90}\). In the cytosol vitamin B12 is bound to cytoplasmic
chaperone methylmalonic aciduria cblC type with homocystinuria (MMACHC), which converts it to cob(II)alamin and removes the axial ligand\(^91\). Methylmalonic aciduria cblD type with homocystinuria (MMADHC) then targets vitamin B12 to MS in the cytosol or MUT in the mitochondria\(^92\). In the cytosol, cobalamin is converted to methylcobalamin (meCbl) by MS and methionine synthase reductase (MSR). The vitamin B12 mitochondrial transporter has yet to be uncovered in human cells, although it has recently been identified in *Caenorhabditis elegans (C. elegans)* as wht-6\(^93\). Finally, the mitochondrial methylmalonic aciduria cblA and cblB type MMAA and MMAB mediate the adenylation of cobalamin and subsequent transfer to MUT\(^94,95\).

One possible explanation for this elaborate route is due to the scarcity of vitamin B12 and its sensitivity to extreme pH. The multiple components of this pathway also advantageously protect the vitamin from degradation and uptake by gut-residing organisms. Of course, gut microorganisms also have evolved ways to scavenge vitamin B12 from their host even when it is bound to IF\(^96\). Many questions still remain regarding the processing and transport of vitamin B12 into human and animal cells such as might one of the ATP-binding cassette type G human orthologs of *wht-6* be responsible for vitamin B12 transport into the mitochondria? How are these processing enzymes and transporters regulated at the transcriptional level? For instance, in *C. elegans* we have found that several of these vitamin B12 processing and transporting genes are transcriptionally
repressed by high levels of vitamin B12 and activated upon perturbation of genes involved in vitamin B12-dependent pathways\textsuperscript{97}.

**Metabolic functions of vitamin B12**

Vitamin B12 acts as a cofactor in two pathways in the eukaryotic cell: the degradation of propionate in mitochondria and the Met/SAM cycle in the cytosol (Figure 2.3). These pathways connect through the production of propionyl-CoA from the breakdown of homocysteine to α-ketobutyrate but otherwise share little in common. Propionate degradation is a catabolic process in the mitochondria that feeds into the TCA cycle\textsuperscript{98} while the Met/SAM cycle is an anabolic process in the cytosol that manufactures essential metabolite building blocks and energy for cell proliferation\textsuperscript{99}. The reactions themselves in which vitamin B12 is required are also different, one being an isomerization and the other a methylation, and thus require different forms of the vitamin with different axial ligands: MUT in propionate degradation requires adenosylcobalamin while MS uses meCbl. Interestingly, bacteria use vitamin B12 in additional pathways such as in ethanolamine degradation, propanediol metabolism and queuosine synthesis. Here, however, we will focus on the two eukaryotic pathways and their regulation.
Figure 2.3 – Metabolic pathways connected to vitamin B12. Metabolic map depicting transport and processing of vitamin B12 into the cell, propionate
degradation in the mitochondria and the Met/SAM and folate cycles in the cytosol. Not all cofactors are drawn. * indicates reactions that occur in both cytosol and mitochondrial compartments. SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine; DHF, dihydrofolate; dTMP, deoxythymidine monophosphate; dUMP, deoxyuridine monophosphate; THF, tetrahydrofolate; 10-fTHF, 10-formylTHF; 5-meTHF, 5-methylTHF; 5,10-meTHF, 5,10-methyleneTHF; Cbl, cobalamin; meCbl, methylcobalamin; adoCbl, adenosylcobalamin; R-Cbl, R-cobalamin; MTs, methyltransferases; GSH, glutathione; GSSG, glutathione disulfide; BCKDHA, branched-chain α-keto dehydrogenase; TCA, tricarboxylic acid cycle.

**The propionate degradation pathway**

Propionate is a short-chain fatty acid and a toxic waste product of odd-chain fatty acid, cholesterol, and branch-chain amino acid degradation. Propionate is also produced and excreted by some human gut microbiota\(^\text{100,101}\). Canonically, the breakdown of propionate begins with the addition of coenzyme A, similar to other carboxylic acids to make propionyl-CoA in the mitochondria. Propionyl-CoA can then be broken down in a few steps to produce succinyl-CoA, which feeds directly into the TCA cycle (Figure 2.3). Vitamin B12 is required by MUT in the radical skeleton rearrangement of methylmalonyl-CoA to succinyl-CoA\(^\text{102}\). MUT binds to vitamin B12 in what is term the ‘base-off/His-on” configuration, in which a histidine residue in a conserved sequence motif of the apoenzyme binds to the cobalt ion, displacing the alpha axial ligand and causing homolytic cleavage of the adenosyl ligand. The reaction continuous through a mechanism known as ‘radical roulette’ that causes the swapping of a substrate hydrogen atom with the adjacent coenzyme A group to produce succinyl-CoA\(^\text{103}\).
Some eukaryotes do not utilize vitamin B12 and instead have alternative methods of degrading propionate such as the acrylate pathway or 2-methylcitrate pathway, which will be discussed in a later section. Interestingly, the free-living nematode *C. elegans*, conditionally operates both the canonical B12-dependent propionate degradation pathway as well as a acrylate-like or ‘shunt’ pathway\(^\text{27}\). As a bacterivore that lives on or nearby rotting vegetation, *C. elegans* is likely to experience a varied diet of different bacterial species with different levels of vitamin B12. The genes involved in the shunt pathway are transcriptionally upregulated by the accumulation of propionate, likely providing the animal with metabolic flexibility when vitamin B12-producing bacteria are scarce\(^\text{27}\).

Unlike the propionate shunt pathway in *C. elegans*, very little is known about the transcriptional regulation of the vitamin B12-dependent propionate degradation pathway in humans and mammals. The enzymes, particularly the vitamin B12-dependent protein MUT, in this pathway are well characterized biochemically and expressed fairly ubiquitously\(^\text{104}\). Nevertheless, there are some known allosteric regulators of MUT and its vitamin B12 cofactor. For example, nitrous oxide binds to and inhibits adenosylcobalamin through oxidation, preventing it from binding MUT\(^\text{105}\). Itaconyl-CoA also inhibits MUT by irreversibly binding in the vitamin B12 pocket\(^\text{106}\). Only one study examined mRNA levels of MUT and found that although these correlate with total enzyme levels in different tissues, they do not correspond to holoenzyme activity, which suggests that post-translation control of the enzyme is integral to its regulation\(^\text{107}\).
**The Met/SAM cycle**

A major biological function of the highly conserved 1CC is to act as a scaffold to relay otherwise unstable one-carbon units to produce the methyl donor SAM. During this process the folate cycle produces nucleotides and NADPH, while the Met/SAM cycle contributes to protein synthesis and the methylation of chromatin and metabolites. These two sub-cycles connect through the shared reaction where a methyl is donated to homocysteine from 5-methyltetrahydrofolate (5-meTHF) to produce methionine, the precursor of SAM (Figure 2.3). In many organisms, this central reaction requires vitamin B12, thus linking nucleotide and energy production with cellular methylation and vitamin B12 biology in a complex and highly regulated coordination. Although it is difficult to dissociate one of these pathways from the other, the primary focus of this section will be the Met/SAM cycle. The Met/SAM cycle interconnects many parts of the metabolic network and therefore has a broad influence on many cellular processes including gene regulation, autophagy and cell death, organismal development, lipogenesis, and stress responses.

The MS reaction, which occurs in the cytosol, is unique in several ways; it is the only reaction that utilizes 5-meTHF producing tetrahydrofolate (THF) and therefore its activity directly impacts flux through the folate cycle; the reaction has no energy barriers *in vivo* and has an unlimited turnover rate, which hints at its importance in methionine regeneration. The mechanism of action requires the
base-off/His-on configuration of meCbl and begins with a negative homocysteine ion grabbing the methyl group from the cobalt atom thus generating methionine and cob(II)alamin. The methyl from a positive 5-meTHF ion is transferred to cobalt to regenerate meCbl and THF\textsuperscript{109}. Roughly every 1000 turn overs meCbl is not regenerated and in that situation another enzyme, MSR along with the cofactors flavin adenine dinucleotide (FAD) and SAM reduce and remethylate cob(II)alamin\textsuperscript{110}. Betaine, a product of choline catabolism and an intracellular osmolyte, can be used as an alternative methyl donor instead of 5-meTHF by the enzyme betaine homocysteine methyltransferase (BHMT) producing methionine and dimethylglycine. This reaction, which also requires meCbl, occurs primarily in the mammalian liver and kidney and is thought to be a means to regulate the concentrations of homocysteine, methionine and betaine in these tissues\textsuperscript{111–113}.

**S-adenosylmethionine (SAM)**

A major function of the Met/SAM cycle is the production of the methyl-donor SAM from the amino acid methionine. Methionine adenosyltransferase (MAT) transfers the adenosyl moiety of ATP to methionine, creating a high-energy sulfonium ion that makes donation of the attached methyl group kinetically favorable\textsuperscript{114}. SAM is a pleiotropic metabolite that participates in four types of reactions: The methylation of DNA, histones, proteins, lipids and other metabolites; The production of polyamines, which are important for cell growth and differentiation; The radicalization of 5'-deoxyadenosyl that is involved in the synthesis of the important
cofactors lipoate and biotin. Finally, SAM is an important step in the transsulfuration pathway, which is responsible for producing glutathione, one of the most abundant metabolites in the cell and a strong antioxidant\textsuperscript{115,116}. This chapter will focus on the transmethylation and transsulfuration pathways exclusively.

\textit{SAM and transmethylation}

SAM-dependent methyl transferases are a large family of enzymes, making up nearly one percent of all human proteins encoded by the genome\textsuperscript{117}. Chromatin methylation depends on the production of SAM as a methyl donor and plays an important role in cell differentiation and stress responses. Methylation marks can change chromatin structure and are therefore one way to link the cellular metabolic state to gene expression. In 2015, Mentch \textit{et al} explored this connection in a study published in \textit{Cell Metabolism} by measuring methylation marks using chromatin immunoprecipitation with sequencing (CHIP-seq) and gene expression by RNA-seq in cells grown in methionine-depleted media\textsuperscript{118}. The authors found that by restricting methionine, SAM and other 1CC cycle metabolite concentrations were also decreased as measured by liquid chromatography coupled with mass spectrometry (LC-MS) and consequently histone 3 trimethylated at the 4th lysine residue (H3K4me3) marks decreased globally and gene expression was altered. In particular, the expression of genes in the 1CC were downregulated suggesting a feedback mechanism to maintain metabolic homeostasis. Additionally, some
cancer-associated genes also had decreased H3K4me3 marks and expression, which emphasizes the role of diet and nutrient status in health and supports other experiments linking methionine restriction to health benefits\textsuperscript{119–121}. 

Although SAM is typically produced in the cytoplasm where the Met/SAM cycle enzymes function and therefore affects global methyl concentration, local production of SAM in the nucleus appears to also be important for tuning gene expression\textsuperscript{122}. In yeast, the SAM producing enzyme, SAM synthetase has been found in the nucleus as part of a large protein complex known as serine and SAM responsive complex (SESAME)\textsuperscript{123}. The SESAME complex is made up of both SAM synthetases, pyruvate kinase M2 (PKM2), enzymes involved in serine metabolism and acetyl-CoA synthetase. SESAME associates with the Set1 methyltransferase complex in an apparent effort to coordinate nutritional status with H3K4 methylation and H3T11 phosphorylation at the site of target genes including PKM2’s own promoter.

Histone methylation is important in stem cell and T cell differentiation\textsuperscript{124,125}, and recently the involvement of the Met/SAM cycle in histone methylation has found to be important in wound healing and stress responses\textsuperscript{126}. In their 2019 Molecular Cell article, Yu et al found that lipopolysaccharide activates the Met/SAM cycle not through one-carbon units but through the production of ATP derived from the pentose phosphate pathway and the serine synthesis pathway\textsuperscript{127}. This synergistic activation of the Met/SAM cycle leads to altered H3K4 trimethylation and upregulates the expression of interleukin-1b, a cytokine, in macrophages. In
C. elegans, changes in gene expression due to histone methylation by SAM are important for acute stress response. In particular, animals were more susceptible to pathogenic bacteria when SAM production was lowered due to disfunction of the sams-1 gene, an ortholog of MAT1A, the gene that encodes MS40.

**SAM and the production of phosphatidylcholine**

Arguably one of the most important products of SAM methylation is phosphatidylcholine (PC) by the ER associated protein phosphatidylethanolamine N-methyltransferase (PEMT). Both phosphatidylethanolamine (PE) and PC are the most abundant phospholipids found in cellular membranes and consequently impact many processes such as metabolite absorption at the intestinal brush boarder128, lipid droplet formation129,130, and mitochondrial fission/fusion dynamics131,132. PC has also been found to exist outside of the membrane, typically in the nucleus and can act as a ligand as well as a signaling molecule in lipogenic biosynthesis pathways133,134.

PC deficiency caused by perturbation of MAT creates enlarged lipid droplets, as easily visualized in the translucent C. elegans harboring a mutation in the sams-1 gene (Figure 2.4)135. Lipid droplets are important organelles heavily involved in lipid homeostasis. When PC production is low and consequently the ratio of PE to PC is high, lipid droplets are enlarged due in part to a change in membrane dynamics as well as an accumulation of triglycerides which lowers the surface area to volume ratio136. Further, higher amounts of PE allow for fusion of
smaller droplets into larger ones\textsuperscript{137}. In mammals, deficiency in PC can lead to fatty liver disease, although at the same time may protect against obesity and type II diabetes\textsuperscript{138,139}. The latter is most likely due to the role of PEMT in downstream production of choline, which can lead to increased acetylcholine levels that might influence energy expenditure through activation of the M\textsubscript{3} receptor\textsuperscript{139}.

**Figure 2.4** – Enlarged lipid droplets caused by a decrease in phosphatidylcholine due to mutation of *sams-1* in *C. elegans*. Differential interference contrast microscopy images of *C. elegans* showing a severe developmental defect and enlarged lipid droplets due to perturbation of *sams-1*. White arrowhead indicates an enlarged lipid droplet.

As part of its role in lipogenesis, PC regulates the activity of the sterol regulatory element binding protein 1 (SREBP1), a homologue of SREBP2, the Nobel prize winning transcription factor responsible for regulating cholesterol biosynthesis\textsuperscript{140}. Low PC activates SREBP1 by causing the relocalization of SREBP-activating proteases from the Golgi to the ER, where SREBP1 is bound.
SREBP1, in turn, regulates lipogenesis genes as well as the Met/SAM cycle in a negative feed-back mechanism\textsuperscript{135}.

As a phospholipid, the lipid tail of PC is subject to modifications which produces many different PC isoforms and some of these have been identified as ligands of peroxisome proliferator-activated receptors (PPARs)\textsuperscript{133} as well as other nuclear receptor transcription factors\textsuperscript{141}. PPARs transcriptionally regulate lipid metabolism in different tissues\textsuperscript{142}. Thus, derivatives of PC, similar to other phospholipids, can transcriptionally regulate metabolic genes by binding and activating transcription factors.

While PC is the main acceptor substrate of SAM-dependent methylation and generates 50\% of plasma homocysteine, the PEMT pathway is not the only nor even the prevailing pathway to produce PC\textsuperscript{143}. In 1956, Eugene Kennedy elucidated a pathway through which diet-derived choline is converted to PC\textsuperscript{144}. Unlike the Kennedy pathway enzymes, PEMT activity is very low in all tissues except the liver and therefore cannot account for the majority of PC production in the mammalian system\textsuperscript{145}. Even in the liver where both PC producing pathways are active, the SAM-dependent pathway produces only 30\% of the total PC\textsuperscript{146}. The purpose of these seemingly redundant pathways has come under some speculation and even heated debate\textsuperscript{147}. There is strong evidence that PEMT activity in the liver keeps PE to PC ratios in an optimal range\textsuperscript{138,148}. Another hypothesis is that under times of starvation when dietary choline is scarce, the SAM-dependent pathway allows the cell not only to continue generating PC but
also regenerate choline\textsuperscript{149}. In support of this, PEMT\textsuperscript{-/-} knock out mice have no obvious phenotypic defects unless fed a diet deficient in choline and then they experience end-stage liver failure\textsuperscript{149}. Similarly, in C. elegans knock down by RNAi of the genes involved in the transmethylation reaction that produces the PC head group prevents animals from developing unless supplemented with exogenous choline\textsuperscript{150}.

Another hypothesis for the existence of two PC-generating pathways is that each produce a different pool of PC isoforms. Indeed, in 1999 DeLong et al found that SAM-dependent methylation of PE produces a broader variety of PC isoforms that may have unique cellular functions as signaling molecules\textsuperscript{151}. PC produced by the Kennedy pathway, however, was found to be much more homogenous and, according to the authors, more likely to function as membrane lipids.

Transmethylation by SAM produces S-adenosylhomocysteine (SAH), which is then hydrolyzed to homocysteine that goes on to either regenerate methionine or produce cystathionine as part of the transsulfuration pathway (Figure 2.3). An important, if often overlooked, function of PE methylation by SAM is for PC to act as a methyl sink in order to drive the SAM to SAH reaction forward and thus produce important sulfonium compounds such as cystathionine and glutathione. When PC synthesis is blocked, cells compensate by using histones as alternative methyl sinks\textsuperscript{126}. This result illustrates the importance of maintaining flux through the Met/SAM cycle not only for transmethylation but also for the producing sulfur containing metabolites via the transsulfuration pathway. Perhaps this extends to
other corners of the Met/SAM cycle. It could be hypothesized that maintaining flux through the MS reaction that converts homocysteine to methionine is also critical for regulating the folate cycle and therefore nucleotide and NADPH pools.

**SAM and the transsulfuration pathway**

In addition to being conjoined to the folate cycle through the vitamin B12-dependent MS reaction, the Met/SAM cycle is part of the transsulfuration pathway beginning with homocysteine. Instead of being remethylated to methionine, homocysteine can be condensed to cystathionine by the pyridoxal-phosphate (PLP)-dependent enzyme, cystathionine-β-synthase (CBS) and ultimately results in the production of glutathione (Figure 2.3). Glutathione is a tripeptide comprised of three amino acids: glutamate, cysteine and glycine and exists in two forms: oxidized (GSSG) and reduced (GSH). Cells maintain intracellular, compartment specific glutathione concentrations that may fluctuate with the cell cycle and stress status of the cell. GSH is used as a reducing agent in a number of reactions resulting in the oxidized GSSG form. NADPH is required to reduce GSSG back to GSH linking the two major antioxidants of the cell in one reaction. Consequently, the GSH to GSSG ratio can be a good indicator of the overall cellular reduction-oxidation (redox) state. Glutathione is also involved in the progression of cell autophagy and cell death. For instance, nutrient starvation induces expulsion of GSH from the cell by a membrane bound transporter thus altering the redox balance and promoting oxidation of protein thiols that aid in signaling autophagy.
Glutathione is also involved in important signaling pathways that induce apoptosis\textsuperscript{155}, necroptosis\textsuperscript{156} and ferroptosis\textsuperscript{157}. Maintaining the GSH to GSSG ratio is yet another reason that Met/SAM cycle metabolites feeding into the transsulfuration pathway must be carefully regulated.

The Met/SAM cycle is connected to redox balance and NADPH indirectly as well through the folate cycle. In fact, the folate cycle generates nearly an equivalent amount of NADPH as the pentose phosphate pathway\textsuperscript{158}. Unlike the Met/SAM cycle, many of the folate cycle reactions occur in both the cytosol and the mitochondria. In the cytosol, trifunctional enzyme methylenetetrahydrofolate dehydrogenase, cyclohydrolase, and formyltetrahydrofolate synthetase 1 (MTHFD1), uses a carbon unit from formate to methylate 10-formylTHF thus generating 5,10-methyleneTHF while oxidizing NADPH to NADP\textsuperscript{+} in the process. In the mitochondria however, this reaction runs in reverse, generating formate and reducing NADP\textsuperscript{+} to NADPH. The flux of these reactions in both compartments was quantified by carbon tracing based fluxomic experiments\textsuperscript{158}. It’s been hypothesized that partitioning of the folate cycle in this way uncouples glycolysis from NADPH production and thereby prevents a buildup of NADHP in the cytoplasm\textsuperscript{50,159}. Another possible explanation is that the generation of formate as a one-carbon unit donor may be important when nutrient levels are low. Interestingly, one study in rats found that when vitamin B12 was depleted in the diet, formate production was greatly increased possibly as a compensatory measure due to low 1CC flux\textsuperscript{160}. 
Finally, methionine directly participates in redox balance. Methionine is an essential amino acid that is important in protein synthesis and feeds directly into the Met/SAM cycle as a precursor of SAM. Although in eukaryotes, methionine can only be obtained through the diet, there are several ways it can be recycled, which perhaps hints at its critical role in the cell. Methionine can of course be produced by methylation of homocysteine in the Met/SAM cycle, but it can also be regenerated through the methionine salvage pathway that feeds off of SAM. Further, protein-bound methionine can be scavenged through the action of sulfoxide reductase and proteases. Similar to cysteine residues, the thiol group of both free and protein-bound methionine can be oxidized or reduced by methionine sulfoxide reductase A (MSRA)\textsuperscript{161}. In fact, it is this very feature of methionine that allows it to act as a buffer to protect proteins from damaging oxidation. A decrease in survival reflecting heightened oxidation damage has been shown when 40% of the methionine residues was exchanged with norleucine, a synthetic methionine mimic that lacks the sulfur atom in \textit{E. coli}\textsuperscript{162}. Although oxidation of methionine residues may be tolerated to a certain extent, after a threshold has been reached, oxidized methionine may signal proteases and ubiquitination to degrade the damaged protein\textsuperscript{163}. There is also some speculation that oxidation and reduction of protein-bound methionine may play a regulatory role on the function of proteins, analogous to protein phosphorylation\textsuperscript{164}. In this scenario methionine oxidation would act like a reversible switch with the ability to deactivate and then reactivate the protein. Mainly due to a lack of adequate tools, there has been little effort to
investigate this phenomenon, but examples have been described in bacteria and plants\textsuperscript{165,166}. As an essential amino acid that is not only integral to protein initiation and oxidation buffering, but also to SAM biosynthesis it must be imperative to balance the metabolic flux of methionine between these two processes.

\textbf{Regulation of the Met/SAM cycle}

Regulation of metabolism is not a novel concept and has historically thought to be carried out post-translationally or by allosteric means. Indeed, the allosteric regulation of the Met/SAM cycle is well documented, most notably in the ratio of SAM to SAH (\textbf{Figure 2.5}). SAH is a potent allosteric inhibitor of all methyl transferases and therefore a high ratio of SAM to SAH promotes transmethylation while a low ratio blocks it\textsuperscript{167–169}. Hence, the conversion of SAH to homocysteine by S-adenosylhomocysteine hydrolase (AHCY) is absolutely crucial to maintain flux through the SAM to SAH reaction\textsuperscript{170}. SAM itself is an allosteric activator of CBS, the enzyme that converts homocysteine to cystathionine\textsuperscript{171}. Both SAM and SAH compete for a binding spot on methylenetetrahydrofolate reductase (MTHFR) but only SAM inhibits the enzyme, allowing MTHFR to ‘sense’ the ratio between SAM and SAH\textsuperscript{172} (\textbf{Figure 2.5}). MTHFR produces 5-meTHF, which is used as a cofactor in the MS reaction. Consequently, when there is a high SAM to SAH ratio, SAM blocks its own accumulation and simultaneously prevents other Met/SAM cycle metabolite unbalances by synergistically decreasing flux through the MS reaction and increasing efflux of homocysteine. Additionally, SAM inhibits some of the MAT
genes but has been shown to activate others in a spatiotemporal specific manner\textsuperscript{173,174}.

\textbf{Figure 2.5 – Allosteric regulation of Met/SAM cycle genes by Met/SAM cycle metabolites.} Black arrows show metabolic reactions. Red repressive arrows indicate inhibition and green arrows indicate activation. SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine; MTs, methyltransferases; meCbl, methylcobalamin; THF, tetrahydrofolate; 5-meTHF, 5-meTHF; 5,10-meTHF, 5,10methyleneTHF.

The main biological function of the mammalian enzyme, glycine N-methyltransferase (GNMT), is to regulate SAM levels\textsuperscript{175}. Using SAM as a methyl donor, GNMT methylates glycine to produce sarcosine, a metabolite with no known essential function, but is quickly reconverted to glycine by sarcosine dehydrogenase (SARDH). GNMT is inhibited by 5-meTHF\textsuperscript{176} (\textbf{Figure 2.5}). Hence, increased SAM levels synergistically decrease folate derived one-carbon unit production and activate GNMT by inhibiting 5-meTHF. GNMT can then decrease SAM levels by methylating glycine to sarcosine. In this manner, GNMT links folate
and Met/SAM cycle regulation to maintain optimal activity through the cycle at the allosteric level.

The mechanistic target of rapamycin complex 1 (mTORC1) is the central complex of mTOR signaling, a nutrient sensing system that, when activated, inhibits autophagy and promotes cell growth\textsuperscript{177}. mTOR signaling can indirectly affect the Met/SAM cycle by activating \textit{de novo} synthesis of serine, a primary source of one carbon units for the 1CC. The mitochondrial folate cycle is also a downstream target of mTOR signaling in response to mitochondrial stress\textsuperscript{178}. Methionine, in turn, is sensed indirectly through the production of SAM which activates mTORC1 by two known mechanisms. SAM binds to SAMTOR, which prevents the formation of the mTORC1 inhibiting SAMTOR-GATOR1 complex\textsuperscript{179}. Sam also methylates protein phosphatase 2A (PP2A) both in yeast and in human cells, which leads to activation of mTORC1\textsuperscript{180–182}.

Like vitamin B12, SAM riboswitches have also been identified in bacteria\textsuperscript{183}. In fact, they make up the largest class of riboswitches and show high affinity for SAM and low affinity for its analogs such as SAH. At least six types of SAM riboswitches have been structurally resolved and fully characterized\textsuperscript{184}. When SAM is bound, these riboswitches repress the expression of genes involved in the biosynthesis and transport of methionine, cysteine and other sulfonium metabolites. Interestingly, a SAM riboswitch in tandem with a vitamin B12 riboswitch was discovered, and both function together as a two input Nor-logic gate to control the expression of MetE, a vitamin B12-independent MS enzyme\textsuperscript{185}. 
Thus, activation of this alternative enzyme will only occur when both SAM and vitamin B12 are low. Although these elements have not been found in animals, there has been some evidence as well as dispute over a potential vitamin B12 responsive internal ribosome entry sequence upstream of the coding sequence in the human MS gene\textsuperscript{186,187}.

The transcriptional regulation of Met/SAM cycle genes remains largely unexplored. In \textit{C. elegans}, many of the Met/SAM cycle genes including \textit{sams-1/MAT} are regulated by SBP-1/SREBP\textsuperscript{135}. Low PC activates SBP-1/SREBP1, which activates Met/SAM cycle gene expression thus creating a negative feedback loop to regulate the production of PC. A recent study found that precise knock out by CRISPR interference (CRISPRi) of the \textit{E. coli} MS gene MetE causes upregulation of other Met/SAM cycle genes in a compensatory mechanism\textsuperscript{188}. Although the transcriptional regulatory network of the Met/SAM cycle remains unclear, the work done so far suggests that the metabolic activity of the cycle must be under both tight transcriptional and allosteric control.

\textbf{Life without vitamin B12}

In most eukaryotic, multicellular life, vitamin B12 is indispensable for two important and in the case of the Met/SAM cycle, highly conserved metabolic pathways. Many organisms, however, are able to grow and thrive without it even though the need to degrade propionate and power the Met/SAM cycle remains. Plants notably don’t utilize vitamin B12 but instead degrade propionate via a β-oxidation-like
pathway\textsuperscript{189} and have a vitamin B12-independent MS enzyme for the recycling of methionine from homocysteine\textsuperscript{190}. MUT is the only enzyme that can catalyze the rearrangement of methylmalonyl-CoA to succinyl-CoA and it requires vitamin B12. In order to degrade propionate without requiring vitamin B12 organisms must employ an entirely different pathway. In the case of the Met/SAM cycle, alternative forms of the MS enzyme exist that do not require vitamin B12. This raises the topic of the evolutionary benefits vs. costs of depending on a low abundance cofactor that must be obtained through the diet. To examine some of the vitamin B12 independent mechanisms this section will focus on several familiar model systems. Both vitamin B12-dependent and independent MS enzymes and propionate degradation pathways in additional model organisms and humans are described in Table 2.1.

<table>
<thead>
<tr>
<th>Species</th>
<th>Makes vitamin B12</th>
<th>Uses vitamin B12</th>
<th>Propionate degradation</th>
<th>MUT gene</th>
<th>MS gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Escherichia coli</em></td>
<td>No</td>
<td>Optional</td>
<td>2-methylcitrate\textsuperscript{191}</td>
<td>N/A</td>
<td>metE, metH</td>
</tr>
<tr>
<td><em>Pseudomonas aeruginosa</em></td>
<td>Yes</td>
<td>Optional</td>
<td>2-methylcitrate\textsuperscript{192}</td>
<td>N/A</td>
<td>metE, metH</td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii</em></td>
<td>Yes</td>
<td>Optional</td>
<td>Unknown</td>
<td>N/A</td>
<td>METE, METH</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>No</td>
<td>No</td>
<td>2-methylcitrate\textsuperscript{193}</td>
<td>N/A</td>
<td>MET6</td>
</tr>
</tbody>
</table>
* Indicates pathway is predicted based on gene homology.

**Bacteria**

Even among bacteria that use vitamin B12 as cofactor for MUT, many species appear to have additional propionate degradation pathways that are vitamin B12 independent\(^{191}\). Propionate can be used as a carbon source as well as a defense mechanism for many bacteria and has its own synthetic pathway in several species\(^{196}\). For example, *Cutibacterium acnes* (formerly *Propionibacterium acnes*) inhabit human skin and secrete significant amounts of propionate that prevent *Staphylococcus epidermidis* biofilm formation and ultimately compete for occupancy of its biological niche\(^{197}\). Interestingly, vitamin B12 supplementation to human hosts has been shown to transcriptionally represses vitamin B12 biosynthesis genes in *Cutibacterium acnes*, which leads to an increase in
porphyrin production that causes inflammation and consequently acne in some individuals\textsuperscript{198}.

Aside from the canonical B12 dependent breakdown pathway, bacteria also utilize the 2-methylcitrate pathway to degrade propionate and feed the TCA cycle\textsuperscript{196}. Likewise, bacteria can synthesize methionine \textit{de novo} as well as recycle it from homocysteine both by a vitamin B12 dependent enzyme and a vitamin B12 independent enzyme, MetH and MetE respectively. Interestingly, these two genes appear to have been produced through convergent evolution based on sequence homology and mechanism of action\textsuperscript{199}. The existence of multiple, semi-redundant pathways may be a strategy of organisms living in an environment with fluctuating availability of shared nutrients. Finally, many bacterial species are bioengineered to produce both vitamin B12 and propionic compounds in the food, cosmetic, plastic and pharmaceutical industries\textsuperscript{200}.

\textbf{Yeast}

\textit{Saccharomyces cerevisiae} otherwise known as baker’s yeast, like plants, does not require vitamin B12. Yeast lack the MUT enzyme and instead degrade propionate through the 2-methylcitrate cycle, which was originally discovered in fungi\textsuperscript{201}. Yeast can also biosynthesize methionine \textit{de novo} although through a different route than bacteria. Yeast are able to metabolize sulfate that can be incorporated into O-acetyl-L-homoserine to produce homocysteine\textsuperscript{202}. Homocysteine is then converted to methionine by the vitamin B12 independent enzyme Met6, which shares
sequence similarity to metH in bacteria\textsuperscript{203}. Unlike humans, yeast have both methionine sulfoxide reductases: MSRA and MSRB, which allows the oxidization and reduction of both S and R epimers of methionine sulfoxide. Studies have shown that the function of these enzymes in maintaining protein and cellular redox balance contribute to lifespan extension under methionine restriction\textsuperscript{204}. Experiments in yeast have been instrumental in elucidating the role of methionine as a nutritional cue for the regulation of metabolic processes such as amino acid biosynthesis, autophagy, cell growth and proliferation\textsuperscript{205}.

\textit{Flies}

Although somewhat debated, there is little evidence that \textit{Drosophila melanogaster} (\textit{Drosophila}), the small fruit fly, requires vitamin B12\textsuperscript{206}. The vitamin is not included in its laboratory diet, however absorption from gut microorganisms cannot be ruled out. In fact, the source of vitamin B12 for other insects that do require it, is primarily the gut microbiota\textsuperscript{206}. \textit{Drosophila} is able to degrade propionate via the vitamin B12 independent acrylate pathway not dissimilar to the propionate shunt pathway found in \textit{C. elegans}\textsuperscript{27,194}. But unlike \textit{C. elegans}, \textit{Drosophila} doesn't have any of the enzymes that act in the vitamin B12-dependent propionate breakdown pathway and consequently they have no way of converting propionate to succinate. While \textit{Drosophila} does appear to have a functional 1CC, the putative MS enzyme among other aspects of the cycle remain to be fully characterized\textsuperscript{207}. On the other hand, \textit{Drosophila} only has one SAM-S enzyme that converts methionine to SAM, making
it a simpler and therefore attractive model for the study of Met/SAM metabolic regulation\textsuperscript{208}. Recent research has proven the value of studying the regulation of metabolism in \textit{Drosophila}, especially in the context of the microbiome, which has been shown to affect fly development\textsuperscript{209}, lifespan\textsuperscript{210}, and even behavior\textsuperscript{211}. Hopefully in the near future \textit{Drosophila} will have its own reconstructed metabolic network to aide in this endeavor.

\textbf{Conclusion}

Vitamin B12, its impact on human health and its role as a cofactor in propionate degradation and one carbon metabolism have been studied since the mid 1900’s. The exquisitely complicated synthesis and route of absorption of this small molecule suggest that, despite its importance as a cofactor, vitamin B12 is a scarce and precious nutrient. As such, many organisms have evolved the means either to completely do without it or have the metabolic plasticity to survive when vitamin B12 is limited via metabolic network rewiring. The function and regulation of these pathways is essential to organismal health. In particular, humans deficient in vitamin B12 exhibit propionic and homocysteine acidemias as well as developmental defects\textsuperscript{49}. Dysregulation of one carbon metabolism is linked to numerous diseases including most famously cancer\textsuperscript{212}. 
CHAPTER III: CAENORHABDITIS ELEGANS METHIONINE/S-ADENOSYLMETHIONINE CYCLE ACTIVITY IS SENSED AND ADJUSTED BY A NUCLEAR HORMONE RECEPTOR

Preface

The research presented in this chapter was originally published in *eLife* in 2020 with the title “*Caenorhabditis elegans* methionine/S-adenosylmethionine cycle activity is sensed and adjusted by a nuclear hormone receptor”\(^9\). I am the first author, Marian Walhout is the corresponding author, and the other authors are: Melissa D. Walker, Olga Ponomarova, Hefei Zhang, Xuhang Li, and Gregory Minevich.

Olga Ponomarova wrote most of the body size analysis MATLAB script and performed the metabolomic measurements by gas chromatography-mass spectrometry (GC-MS) shown in Figure 3.6a. Hefei Zhang and Xuhang Li performed the RNA-seq, cleaned and mapped reads that I used to analyze transcript levels and differential expression shown in Figures 3.5c, 3.5e, and 3.7a, 3.7d, and 3.8 (Supplementary files 3.3 and 3.4). Gregory Minevich ran the CloudMap pipeline on the whole genome sequencing data to identify casual variants from the forward genetic screen shown in Figure 3.2b. I performed all other experiments with technical help from Melissa Walker.

Summary

Vitamin B12 is an essential micronutrient that functions in two metabolic pathways:
the canonical propionate breakdown pathway and the methionine/S-adenosylmethionine (Met/SAM) cycle. In Caenorhabditis elegans (C. elegans), low vitamin B12, or genetic perturbation of the canonical propionate breakdown pathway results in propionate accumulation and the transcriptional activation of a propionate shunt pathway. This propionate-dependent mechanism requires the nuclear hormone receptor transcription factor nhr-10 and is referred to as “B12-mechanism-I”. Here, we report that vitamin B12 represses the expression of Met/SAM cycle genes by a propionate-independent mechanism we refer to as “B12-mechanism-II”. This mechanism is activated by perturbations in the Met/SAM cycle, genetically or due to low dietary vitamin B12. B12-mechanism-II requires nhr-114 to activate Met/SAM cycle gene expression, the vitamin B12 transporter, pmp-5, and adjust influx and efflux of the cycle by activating msra-1 and repressing cbs-1, respectively. Taken together, Met/SAM cycle activity is sensed and transcriptionally adjusted to be in a tight metabolic regime.

Introduction

Metabolism lies at the heart of most cellular and organismal processes. Anabolic metabolism produces biomass during development, growth, cell turnover and wound healing, while catabolic processes degrade nutrients to generate energy and metabolic building blocks. Animals must be able to regulate their metabolism in response to nutrient availability and to meet growth and energy demands. Metabolism can be regulated by different mechanisms, including the allostERIC
modulation of metabolic enzyme activity and the transcriptional regulation of metabolic genes. Changes in metabolic enzyme level and/or activity can result in changes in metabolic flux, which is defined as the turnover rate of metabolites through enzymatically controlled pathways. Metabolic flux can result in the accumulation or depletion of metabolites\textsuperscript{213–215}. These metabolites may interact with enzymes directly to allosterically affect the enzyme’s catalytic properties. Alternatively, metabolites can alter the transcriptional regulation of metabolic enzymes by interacting with transcription factors (TFs) and changing their activity or localization\textsuperscript{1,216}. A classic example of the transcriptional regulation of metabolism is the activation of cholesterol biosynthesis genes in mammals by SREBP that responds to low levels of cholesterol\textsuperscript{217}.

Vitamin B\textsubscript{12} is an essential cofactor for two metabolic enzymes: methylmalonyl-CoA mutase and methionine synthase. Methylmalonyl-CoA mutase (EC 5.4.99.2) catalyzes the third step in the breakdown of the short-chain fatty acid propionate, while methionine synthase (EC 2.1.1.13) converts homocysteine into methionine in the Met/SAM cycle (Figure 3.1a). The Met/SAM cycle is part of one-carbon metabolism, which also includes folate metabolism and parts of purine and thymine biosynthesis\textsuperscript{50}. The one-carbon cycle produces many important building blocks for cellular growth and repair, including nucleotides and SAM, the major methyl donor of the cell. SAM is critical for the synthesis of phosphatidylcholine, an important component of cellular membranes, as well as for the methylation of DNA, RNA, and histones\textsuperscript{126}. Both vitamin B\textsubscript{12}-dependent metabolic pathways
have been well studied at the biochemical level, however little is known about how these pathways are regulated transcriptionally.
Figure 3.1 – Two mechanisms of gene regulation by low vitamin B12 dietary conditions. (A) Cartoon of vitamin B12-related metabolic pathways in *C. elegans*. 

**A**

- **Phospholipid metabolism**
  - Phosphatidyl-ethanolamine
  - Phosphatidyl-choline
  - CDP-choline
  - CTP
  - Choline
- **Alpha-Ketobutyrate**
  - Alpha-Ketobutyrate
  - Cystathionine
- **Homocysteine**
  - Methionine
  - SAM
  - SAH
- **One carbon cycle**
  - Methionine sulfoxide
  - Methyl
  - THF
  - 5,10-meTHF
  - 10-fTHF
  - 5-meTHF
- **Methionine sulfoxide**
  - Methionine sulfoxide
  - MetR-1
  - mtrr-1
  - B1
  - Cbs-1
  - Cbs-2
  - Ahcy-1
  - Mel-32
  - Dhc-1
  - Cbl-1
  - Cth-1
  - Cth-2
- **SAMS-1,3,4,5**
  - SAMS-1,3,4,5
  - Phosphatidyl-ethanolamine
  - Phosphatidyl-choline
  - Choline
  - Mel-32
  - Dhfr-1
  - Tyms-1
  - Dao-3
  - AHCY-1
  - Mthf-1
  - Serine
  - Glycine
  - 3-HP-CoA
  - MSA
  - AHY-1
  - Mthf-1
- **Phospholipid metabolism**
  - Phosphatidyl-ethanolamine
  - Phosphatidyl-choline
  - CDP-choline
  - CTP
  - Choline
  - Methionine
  - SAM
  - SAH
  - Methionine sulfoxide
  - MetR-1
  - mtrr-1
  - B1
  - Cbs-1
  - Cbs-2
  - Ahcy-1
  - Mel-32
  - Dhc-1
  - Cbl-1
  - Cth-1
  - Cth-2
- **One carbon cycle**
  - Methionine sulfoxide
  - Methyl
  - THF
  - 5,10-meTHF
  - 10-fTHF
  - 5-meTHF
- **Methionine sulfoxide**
  - Methionine sulfoxide
  - MetR-1
  - mtrr-1
  - B1
  - Cbs-1
  - Cbs-2
  - Ahcy-1
  - Mel-32
  - Dhc-1
  - Cbl-1
  - Cth-1
  - Cth-2
- **SAMS-1,3,4,5**
  - SAMS-1,3,4,5
  - Phosphatidyl-ethanolamine
  - Phosphatidyl-choline
  - Choline
  - Mel-32
  - Dhfr-1
  - Tyms-1
  - Dao-3
  - AHCY-1
  - Mthf-1
  - Serine
  - Glycine
  - 3-HP-CoA
  - MSA
  - AHY-1
  - Mthf-1

**B**

- **Wild type**
  - Untreated
  - Vitamin B12
  - Vitamin B12 + Propionate
- **Δnhr-10**
  - Untreated
  - Vitamin B12
  - Vitamin B12 + Propionate

**C**

- **acdh-1**
  - TPH
  - Wild type
  - Δnhr-10

**D**

- **Low vitamin B12**
  - Canonical pathway activity
  - B12-mechanism-I
  - Propionate
  - nhr-10
  - acdh-1
  - B12-mechanism-II

*Figure 3.1 – Two mechanisms of gene regulation by low vitamin B12 dietary conditions. (A) Cartoon of vitamin B12-related metabolic pathways in *C. elegans*.**
CDP, cytidine 5'-diphosphocholine; DHF, dihydrofolate; 3-HP, 3-hydroxypropionate; THF, tetrahydrofolate; 5,10-meTHF, 5,10-methylenetetrahydrofolate; 5-meTHF, 5-methyltetrahydrofolate; 10-FTFH, 10-formyltetrahydrofolate; BCKDH, branched-chain α-ketoacid dehydrogenase complex; MM-CoA, methylmalonyl-coenzyme A; *MUT, human methylmalonyl-coenzyme A mutase; *MS, human methionine synthase; MSA, malonic semialdehyde; SAH, S-adenosylhomocysteine; SAM, S-adenosylmethionine; FFL, feed forward loop. Dotted arrows indicate multiple reaction steps. (B) Fluorescence microscopy images of Pacdh-1::GFP reporter animals in wild type and ∆nhr-10 mutant background with different supplements as indicated. Insets show brightfield images. (C) RNA-seq data of acdh-1 mRNA with and without 20 nM vitamin B12 in wild type and ∆nhr-10 mutant animals24. Datapoints show each biological replicate and the bar represents the mean. TPM, transcripts per million. P adjusted values are provided in Supplementary file 3.1. (D) Cartoon illustrating two mechanisms of gene regulation by low vitamin B12 dietary conditions.

The nematode C. elegans is a highly tractable model for studying the relationships between diet, disease and metabolism218,219. C. elegans is a bacterivore that can thrive on both high and low vitamin B12 diets16,23,29. We previously discovered that perturbations of the canonical propionate breakdown pathway, either genetically or by low dietary vitamin B12, results in the transcriptional activation of five genes comprising an alternative propionate breakdown pathway, or propionate shunt (Figure 3.1a)27. Activation of the propionate shunt occurs only with sustained propionate accumulation and absolutely depends on the nuclear hormone receptor (NHR) nhr-1024. nhr-10 functions together with nhr-68 in a type 1 coherent feedforward loop known as a persistence detector24. We refer to the regulation of gene expression by accumulation of propionate due to low vitamin B12 dietary conditions as “B12-mechanism-I”.
Here, we report that vitamin B12 represses the expression of Met/SAM cycle genes by a propionate-independent mechanism we refer to as “B12-mechanism-II”. We find that B12-mechanism-II is activated upon perturbation of the Met/SAM cycle, either genetically or nutritionally, due to low dietary vitamin B12. This mechanism requires another NHR, \textit{nhr-114}, which responds to low levels of SAM. B12-mechanism-II not only activates Met/SAM cycle gene expression, it also activates the expression of the vitamin B12 transporter \textit{pmp-5} and the methionine sulfoxide reductase \textit{msra-1}, and represses the expression of the cystathionine beta synthase \textit{cbs-1}. The regulation of the latter two genes increases influx and reduces efflux of the Met/SAM cycle, respectively. These findings indicate that low Met/SAM cycle activity is sensed and transcriptionally adjusted to be maintained in a tightly controlled regime. Taken together, in \textit{C. elegans} the genetic or nutritional perturbation of the two vitamin B12-dependent pathways is sensed by two transcriptional mechanisms via different NHRs. These mechanisms likely provide the animal with metabolic adaptation to develop and thrive on different bacterial diets in the wild.

**Results**

*Low dietary vitamin B12 activates two transcriptional mechanisms*

As in humans, vitamin B12 acts as a cofactor in two \textit{C. elegans} pathways: the canonical propionate breakdown pathway and the Met/SAM cycle, which is part of one-carbon metabolism (Figure 3.1a). These pathways are connected because
homocysteine can be converted into cystathionine by the cystathionine beta synthase CBS-1, which after conversion into α-ketobutyrate is converted into propionyl-CoA. When flux through the canonical propionate breakdown pathway is perturbed, either genetically or nutritionally, i.e., when dietary vitamin B12 is low, a set of genes comprising an alternative propionate breakdown pathway or propionate shunt is transcriptionally activated\textsuperscript{16,23,27,29}. Low vitamin B12 results in an accumulation of propionate, which, when sustained, activates a gene regulatory network circuit known as a type 1 coherent feedforward loop with an AND-logic gate composed of two TFs, \textit{nhr-10} and \textit{nhr-68}\textsuperscript{24}. The first gene in the propionate shunt, \textit{acdh-1}, acts as a control point: its expression is induced several hundred-fold when vitamin B12 is limiting\textsuperscript{16,23,27,29}.

We have previously used transgenic animals expressing the green fluorescent protein (GFP) under the control of the \textit{acdh-1} promoter as a vitamin B12 sensor\textsuperscript{23,29,34}. In these \textit{Pacdh-1::GFP} animals, GFP expression is high throughout the intestine on an \textit{E. coli} OP50 diet, which is low in vitamin B12, and GFP expression is very low on a \textit{Comamonas aquatica} DA1877 diet that is high in vitamin B12\textsuperscript{23,29}. Low GFP expression resulting from vitamin B12 supplementation to the \textit{E. coli} OP50 diet can be overcome by addition of propionate (\textbf{Figure 3.1b})\textsuperscript{24}. The activation of \textit{acdh-1} expression in response to accumulating propionate is completely dependent on \textit{nhr-10} (\textbf{Figure 3.1b})\textsuperscript{24}. Interestingly, we found that while GFP levels are reduced in the anterior intestine, there is still remaining GFP expression in the posterior intestine in \textit{Pacdh-1::GFP} transgenic animals lacking
*nhr-10* (Figure 3.1b). Since *nhr-10* is absolutely required to mediate the activation of *acdh-1* by propionate, this means that there is another, propionate-independent mechanism of activation. Importantly, the residual GFP expression in *Pacdh-1::GFP; Δnhr-10* was completely repressed by the supplementation of vitamin B12 (Figure 3.1b). This result was confirmed by inspecting our previously published RNA-seq data: in *Δnhr-10* animals there is residual endogenous *acdh-1* expression which is eliminated by the addition of vitamin B12 (Figure 3.1c, Supplementary file 3.1)\(^2^4\). These results demonstrate that there is another mechanism by which low vitamin B12 activates gene expression that is independent of propionate accumulation, which occurs when flux through the canonical propionate breakdown pathway is perturbed. We refer to the activation of gene expression in response to canonical propionate breakdown perturbation as "B12-mechanism-I" and the other, propionate-independent mechanism as "B12-mechanism-II" (Figure 3.1d).

*Met/SAM cycle perturbations activate B12-mechanism-II*

To determine the mechanism by which "B12-mechanism-II" is activated, we used the *Pacdh-1::GFP* vitamin B12 sensor in the *Δnhr-10* mutant background, which cannot respond to B12-mechanism-I. We first performed a forward genetic screen using ethyl methanesulfonate (EMS) to find mutations that activate GFP expression in *Pacdh-1::GFP; Δnhr-10* animals in the presence of vitamin B12 (Figure 3.2a). We screened ~8,000 genomes and identified 27 mutants, 16 of
which were viable and produced GFP-expressing offspring. Seven of these mutants were backcrossed with the *Pacdh-1::GFP;Δnhr-10* parent strain. Single nucleotide polymorphism mapping and whole genome sequencing revealed mutations in *metr-1, mtrr-1, sams-1, mthf-1* and *pmp-5* (Figure 3.2b). The first four genes encode enzymes that function directly in the Met/SAM cycle (Figure 3.1a). *metr-1* is the single ortholog of human methionine synthase; *mtrr-1* is the ortholog of *MTRR* that encodes methionine synthase reductase; *sams-1* is orthologous to human *MAT1A* and encodes a SAM synthetase; and *mthf-1* is the ortholog of human *MTHFR* that encodes methylenetetrahydrofolate reductase. We also found mutations in *pmp-5*, an ortholog of human *ABCD4*, which encodes a vitamin B12 transporter.
Figure 3.2 – Met/SAM cycle perturbations activate B12-mechanism-II. (A) Workflow for the EMS mutagenesis screen using Pacdh-1::GFP;Δnhr-10 reporter animals supplemented with 20 nM vitamin B12. WGS: whole genome sequencing. (B) To-scale cartoons of amino acid changes in the proteins encoded by the genes found in the forward genetic screen. (C) Workflow for RNAi screen using Pacdh-
Next, we performed a reverse genetic RNAi screen using a library of predicted metabolic genes in *Pacdh-1::GFP;∆nhr-10* animals fed *E. coli* HT115 bacteria (the bacterial diet used for RNAi experiments) supplemented with vitamin B12 (*Figure 3.2c*). In these animals, GFP expression is off and we looked for those RNAi knockdowns that activated GFP expression in the presence of vitamin B12. Out of more than 1,400 genes tested, RNAi of only five genes resulted in activation of GFP expression: *metr-1, mtrr-1, sams-1, mthf-1* and *mel-32* (*Figure 3.2d*). Four of these genes were also found in the forward genetic screen (*Figure 3.2b*). The fifth gene, *mel-32*, also functions in one-carbon metabolism (*Figure 3.1a*). It is an ortholog of human *SHMT1* and encodes serine hydroxymethyltransferase that converts serine into glycine thereby producing 5,10-methylenetetrahydrofolate, the precursor of 5-methyltetrahydrofolate (5-meTHF), which donates a methyl group in the reaction catalyzed by METR-1 (*Figure 3.1a*). These results show that genetic perturbations in Met/SAM cycle genes activate B12-mechanism-II, even in the presence of vitamin B12. This indicates that reduced activity of the Met/SAM cycle, either due to genetic perturbations or as a result of low dietary vitamin B12 activates B12-mechanism-II (*Figure 3.2e*). Therefore, genetic perturbations in either pathway that uses vitamin B12 as a cofactor activate the vitamin B12 sensor.
**B12-mechanism-II activates Met/SAM cycle gene expression in response to Met/SAM cycle perturbations**

To ask what other genes are activated by B12-mechanism-II, we performed RNA-seq using four Met/SAM cycle mutants identified in the forward genetic screen. All strains were supplemented with vitamin B12 and gene expression profiles of the Met/SAM cycle mutants were compared to the parental \textit{Pacdh-1::GFP;\textDelta{nhr-10}} strain. We found a set of 110 genes that are upregulated in all four Met/SAM cycle mutants, and a smaller set of 11 genes that are downregulated (Figure 3.3a, Supplementary file 3.2). Importantly, endogenous \textit{acdh-1} was upregulated in each of the Met/SAM cycle mutants, validating the results obtained with the \textit{Pacdh-1::GFP} vitamin B12 sensor (Supplementary file 3.2). Remarkably, we found that most Met/SAM cycle genes are significantly upregulated in each of the Met/SAM cycle mutants, and one of them, \textit{mel-32}, was significantly increased in three of the four mutants (the fourth mutant just missing the selected statistical threshold) (Figure 3.3b, Supplementary file 3.2).
Figure 3.3 – Genetic mutations in Met/SAM cycle genes activate Met/SAM cycle gene expression. (A) Hierarchical clustering of log10-transformed fold change RNA-seq data. Changes are relative to the *Pacdh-1::GFP;Δnhr-10* parent strain with the cutoffs of a fold change of 1.5 and a *P* adjusted value less than 0.01.
Only genes that change in expression in all four Met/SAM cycle gene mutants are shown. (B) Bar graphs of Met/SAM cycle gene expression by RNA-seq in the four Met/SAM cycle mutants and in the *Pacdh-1::GFP;Δnhr-10* parent strain. Datapoints show each biological replicate and the bar represents the mean. TPM, transcripts per million. *P* adjusted values are provided in Supplementary file 3.2.

(C) RNA-seq comparison of Met/SAM cycle (blue) and propionate shunt (red) gene expression in response to 20 nM vitamin B12. Bar represents the mean of two biological replicates. *P* adjusted values are provided in Supplementary file 3.1.

(D) Comparison of Met/SAM cycle and propionate shunt genes in response to 20 nM vitamin B12 plus 40 mM propionate or in *nhr-10* and *nhr-68* mutant animals. Bar represents the mean of two biological replicates. *P* adjusted values are provided in Supplementary file 3.1. (E) Cartoon illustrating Met/SAM cycle gene activation in response to B12-mechanism-II.

The finding that Met/SAM cycle genes are transcriptionally activated in response to genetic Met/SAM cycle perturbations implies that these genes may also be activated by low vitamin B12. Indeed, inspection of our previously published RNA-seq data24 revealed that expression of Met/SAM cycle genes is repressed by vitamin B12 (Figure 3.3c, Supplementary file 3.1). In contrast to propionate shunt genes, however, Met/SAM cycle genes are not induced in response to propionate supplementation, nor are these genes affected in *nhr-10* or *nhr-68* mutants, which are the mediators of the propionate response (B12-mechanism-I, Figure 3.3d, Supplementary file 3.1). Therefore, Met/SAM cycle gene expression is activated by B12-mechanism-II in response to either genetic or nutritional (low vitamin B12) perturbations in the Met/SAM cycle (Figure 3.3e).

*nhr-114 mediates Met/SAM cycle activation in response to low Met/SAM cycle activity*

How do perturbations in the Met/SAM cycle activate B12-mechanism-II? There are
two components to this question: (1) which TF(s), and (2) which metabolite(s) mediate Met/SAM cycle gene induction? We first focused on identifying the TF(s) involved in B12-mechanism-II. Previously, we identified more than 40 TFs that activate the *Pacdh-1::GFP* vitamin B12 sensor in wild type animals\textsuperscript{15}. Subsequently, we found that only a subset of these TFs are involved in B12-mechanism-I, most specifically *nhr-10* and *nhr-68*\textsuperscript{24}. To identify TFs involved in B12-mechanism-II, we performed RNAi of all TFs that regulate the *acdh-1* promoter\textsuperscript{15,24} in *Pacdh-1::GFP;\Delta nhr-10* animals harboring mutations in Met/SAM genes. As mentioned above, these animals express moderate levels of GFP in response to the activation of B12-mechanism-II by Met/SAM cycle mutations. RNAi of several TFs reduced GFP expression, including *elt-2*, *nhr-23*, *cdc-5L*, *lin-26*, *sbp-1* and *nhr-114* (Figure 3.4). Most of these TFs function at a high level in the intestinal gene regulatory network, elicit gross physiological phenotypes when knocked down, and are also involved in B12-mechanism-I\textsuperscript{15,24}. For instance, *elt-2* is a master regulator that is required for the intestinal expression of most if not all genes and is therefore not specific\textsuperscript{15,220}. The TFs *nhr-23* and *lin-26* were not found to be involved in B12-mechanism-I and could therefore potentially be involved in B12-mechanism-II. However, RNAi of these TFs causes severe developmental delay across all strains and thus their contribution could be less specific (Figure 3.4c). Only *nhr-114* RNAi, which we previously found not to be involved in B12-mechanism-I, specifically repressed GFP expression in the *nhr-10* deletion mutant background (Figures 3.4a, 3.4b)\textsuperscript{24}. This indicates that *nhr-114* mediates the
response to B12-mechanism-II. Interestingly, \textit{nhr-114} RNAi greatly slowed development in Met/SAM cycle mutants but not in the parental strain (Figures 3.4a, 3.4b, insets). This indicates that combined \textit{nhr-114} and Met/SAM cycle perturbations produce a synthetic sick phenotype and points to the functional importance of \textit{nhr-114} in Met/SAM cycle metabolism.
Figure 3.4 – *nhr-114* is required for activation of *Pacdh-1::GFP;Δnhr-10*
**mutant reporter strain in response to Met/SAM cycle perturbations.** (A) *nhr-114* RNAi reduces GFP expression in *Pacdh-1::GFP;Δnhr-10* animals harboring Met/SAM cycle gene mutations. Insets show brightfield images. Difference in exposure time is indicated in yellow. (B) *nhr-114* RNAi represses *Pacdh-1::GFP* expression in all Met/SAM cycle mutants. Insets show brightfield images. (C) Diluted RNAi of indicated TFs repress *Pacdh-1::GFP* in Met/SAM cycle mutants harboring a deletion in *nhr-10*. Some TF RNAi experiments were diluted with vector control RNAi to circumvent strong deleterious phenotypes. Dilution ratios of clone to vector by volume are indicated per clone. Insets show brightfield images.

Previously, it has been reported that *nhr-114* loss-of-function mutants develop slowly and are sterile when fed a diet of *E. coli* OP50 bacteria, while they grow faster and are fertile on a diet of *E. coli* HT115\(^{221}\). *E. coli* HT115 bacteria are thought to contain higher levels of vitamin B12 than *E. coli* OP50 cells\(^{23,222}\). Therefore, we asked whether *Δnhr-114* mutant phenotypes could be rescued by vitamin B12 supplementation. Indeed, we found that both sterility and slow development of *Δnhr-114* mutants fed *E. coli* OP50 could be rescued by supplementation of vitamin B12 (Figures 3.5a, 3.5b).
Figure 3.5 – **nhr-114 is required for B12-mechanism-II in response to Met/SAM cycle perturbations.** (A) \(\Delta nhr-114\) mutant growth and fertility phenotypes are rescued by vitamin B12, methionine or choline supplementation. (B) Quantification of body size of wild type and \(\Delta nhr-114\) mutant animals that are
untreated or supplemented with either vitamin B12, methionine or choline. Statistical significance was determined by the Kruskal-Wallis test with post hoc comparison using Dunn’s multiple comparison test. (C) Bar graphs of Met/SAM cycle gene expression by RNA-seq in Pacdh-1::GFP, Pacdh-1::GFP;Δmetr-1 and Pacdh-1::GFP;Δnhr-10;metr-1(ww52) animals treated with vector control or nhr-114 RNAi. Datapoints show each biological replicate and the bar represents the mean. TPM, transcripts per million. *P* adjusted values are provided in Supplementary file 3.3. (D) Relative mRNA level expression (fold change) as determined by qRT-PCR of Met/SAM cycle genes whose *P* adjusted values were high in RNA-seq shown in C. Datapoints show each biological replicate and bar represents the mean. (E) Pie chart showing portion of genes upregulated by Met/SAM cycle perturbation that are nhr-114 dependent. (F) Cartoon illustrating the requirement of nhr-114 in B12-mechanism-II.

The Met/SAM cycle generates SAM, the major methyl donor of the cell that is critical for the synthesis of phosphatidylcholine\(^ {126}\) (**Figure 3.1a**). It has previously been shown that perturbation of sams-1, which converts methionine into SAM (**Figure 3.1a**), leads to a strong reduction in fecundity and large changes in gene expression\(^ {40,223}\). Importantly, these phenotypes can be rescued by supplementation of choline, which supports an alternative route to phosphatidylcholine biosynthesis\(^ {135}\) (**Figure 3.1a**). Therefore, in *C. elegans*, the primary biological function of the Met/SAM cycle is to produce methyl donors that facilitate the synthesis of phosphatidylcholine. We found that the Δnhr-114 mutant phenotypes can also be rescued by either methionine or choline supplementation (**Figures 3.5a, 3.5b**). Although Gracida and Eckmann did not find that methionine rescued nhr-114 RNAi knock down animals, it is possible this was due to methodological differences. For example, they added bulk L-amino acids and at only one concentration. The actual concentration of ingested methionine might have been masked by other amino acids or was simply too low to provide an
observable rescue. Rescue of nhr-114 mutant phenotypes by choline and methionine support the hypothesis that low levels of phosphatidylcholine are the underlying cause of the Δnhr-114 mutant phenotypes on low vitamin B12 diets, when B12-mechanism-II cannot activate Met/SAM cycle gene expression. This result provides further support for the functional involvement of nhr-114 in the Met/SAM cycle, and leads to the prediction that nhr-114 activates Met/SAM cycle gene expression in response to genetic or nutritional perturbation of the cycle’s activity. To directly test this prediction, we performed RNA-seq on Pacdh-1::GFP and Pacdh-1::GFP;Δnhr-10;metr-1(ww52) animals supplemented with vitamin B12 and subjected to nhr-114 or vector control RNAi. We also included Pacdh-1::GFP;Δmetr-1 animals, which have wild type nhr-10. We found that Met/SAM cycle genes are robustly induced in both metr-1 mutants, recapitulating our earlier observation (Figures 3.5c, 3.5d, 3.3b, Supplementary files 3.3 and 3.2). Importantly, we found that this induction is absolutely dependent on nhr-114 (Figures 3.5c, 3.5d, Supplementary file 3.3). Overall, the induction of 32 of the 110 genes that are upregulated by Met/SAM cycle perturbations requires nhr-114 (Figure 3.5e). These genes are candidate modulators of Met/SAM cycle function. Interestingly, in the presence of vitamin B12, basal Met/SAM cycle gene expression is not affected by nhr-114 RNAi (Figures 3.5c, 3.5d, Supplementary file 3.3). Together with the observation that nhr-114 and Met/SAM cycle perturbations produce a synthetic sick phenotype (Figures 3.4a, 3.4b), this indicates that nhr-114 is specifically involved in B12-mechanism-II, which is
activated when the activity of the cycle is hampered. Taken together, perturbations in the Met/SAM cycle elicited either by low dietary vitamin B12 or by genetic perturbations activate Met/SAM cycle gene expression by B12-mechanism-II, which requires the function of nhr-114 (Figure 3.5f).

**Methionine and choline supplementation suppress B12-mechanism-II**

Which metabolites are involved in the activation of B12-mechanism-II? To start addressing this question, we first performed targeted metabolomics by gas chromatography-mass spectrometry (GC-MS) on animals fed a vitamin B12-deplete *E. coli* OP50 diet with or without supplementation of vitamin B12. The enhanced activity of the Met/SAM cycle in the presence of supplemented vitamin B12 is apparent because methionine levels increase, while homocysteine levels decrease (Figure 3.6a). 3-hydroxypropionate levels are also dramatically decreased by vitamin B12 supplementation, because propionate is preferentially degraded by the canonical propionate breakdown pathway\(^{27}\) (Figure 3.6a). We reasoned that either low methionine, low SAM, low phosphatidylcholine or high homocysteine could activate B12-mechanism-II when activity of the Met/SAM cycle is perturbed.
Figure 3.6 – Methionine and choline supplementation suppress vitamin B12-mechanism-II. (A) Box plots showing GC-MS data from wild type animals fed *E. coli* OP50 with or without supplemented 64 nM vitamin B12. Statistical significance determined by two-tailed t-test. 3-HP, 3-hydroxypropionate. (B) *cbs-1* RNAi reduces GFP expression in *Pacdh-1::GFP;Δnhr-10* and *Pacdh-1::GFP;Δnhr-10* mutants harboring Met/SAM cycle gene mutations. Boxplot showing median and interquartile range of normalized GFP intensity measurements of fluorescent images. (C) Met/SAM cycle point mutants expressing *Pacdh-1::GFP* treated with *cbs-1* RNAi. Insets are brightfield images. (D) Methionine or choline supplementation represses GFP expression in *Pacdh-1::GFP;Δnhr-10* animals. Insets show brightfield images. (E) Methionine or choline supplementation represses GFP expression in *Pacdh-1::GFP;Δnhr-10;metr-1(ww52)* and *Pacdh-1::GFP;Δnhr-10;mthf-1(ww50)* animals, while choline but not methionine supplementation represses GFP expression in *Pacdh-1::GFP;Δnhr-10;sams-
1(ww51) animals. This experiment was performed without vitamin B12 supplementation. (F) GC-MS quantification of methionine levels in Met/SAM cycle mutants and in the parental strain. Statistical significance was determined using one-way ANOVA with post-hoc comparison using Dunnett’s T3 test.

We first tested the possibility that the accumulation of homocysteine in Met/SAM cycle mutants may activate B12-mechanism-II. In *C. elegans*, RNAi of *cbs-1* causes the accumulation of homocysteine$^{224}$. Therefore, we reasoned that, if homocysteine accumulation activates B12-mechanism-II, RNAi of *cbs-1*, should increase GFP expression in the *Pacdh-1::GFP* vitamin B12 sensor. Remarkably, however, we found the opposite: RNAi of *cbs-1* repressed GFP expression in *Pacdh-1::GFP;Δnhr-10* animals but not in the Met/SAM cycle mutants (*Figures 3.6b, 3.6c*). This indicates that a build-up of homocysteine is not the metabolic mechanism that activates B12-mechanism-II. The repression of GFP expression by *cbs-1* RNAi in *Pacdh-1::GFP;Δnhr-10* animals could be explained by a decrease in the conversion of homocysteine into cystathionine and an increase in the conversion into methionine resulting in support of Met/SAM cycle activity (*Figure 3.1a*). In sum, B12-mechanism-II is not activated by a build-up of homocysteine.

Next, we explored whether low methionine, low SAM or low phosphatidylcholine activates B12-mechanism-II. We found that either methionine or choline supplementation dramatically repressed GFP expression in *Pacdh-1::GFP;Δnhr-10* animals (*Figure 3.6d*). However, neither metabolite greatly affected GFP levels in wild type reporter animals (*Figure 3.6d*). Since these animals are fed vitamin B12-depleted *E. coli* OP50 bacteria and have functional
*nhr-10* and *nhr-68* TFs, GFP expression is likely high due to propionate accumulation, *i.e.*, B12-mechanism-I. Importantly, either methionine or choline supplementation also repressed GFP expression induced by B12-mechanism-II due to mutations in *metr-1* or *mthf-1* (Figure 3.6e). However, while choline supplementation repressed GFP expression in *sams-1*(*ww51*) mutants, methionine supplementation did not (Figure 3.6e). SAMS-1 converts methionine into SAM, and methionine levels are greatly increased in *sams-1* mutant animals, while being reduced in *metr-1*, *mtrr-1* and *mthf-1* mutants (Figures 3.1a, 3.6f). Since methionine levels are elevated in *sams-1* mutants, and because methionine supplementation cannot suppress GFP expression in these mutants, these results indicate that low methionine is not the direct activator of B12-mechanism-II, but rather that it is either low SAM, or low phosphatidylcholine, both of which require methionine for their synthesis. In *metr-1* and *mthf-1* mutants methionine supplementation supports the synthesis of SAM and phosphatidylcholine, and in these mutants, methionine supplementation would therefore act indirectly. We did observe a mild reduction in GFP levels upon methionine supplementation in *sams-1*(*ww51*) animals likely because it is not a complete loss-of-function allele, and/or functional redundancy with three other SAMS genes (Figure 3.6e).

To distinguish between the possibilities of low SAM or low phosphatidylcholine activating B12-mechanism-II we next focused on *pmt-2*. PMT-2 is involved in the second step of the conversion of phosphatidylethanolamine head group into the phosphatidylcholine head group^{44} (Figure 3.1a). The
expression of pmt-2 is repressed by vitamin B12, but not activated by propionate, is not under the control of nhr-10 or nhr-68, and is activated in Met/SAM cycle mutants\textsuperscript{24} (Figure 3.7a, Supplementary files 3.1, 3.2). We reasoned that pmt-2 RNAi might allow us to discriminate whether low SAM or low phosphatidylcholine activates B12-mechanism-II. Specifically, we would expect pmt-2 RNAi to activate B12-mechanism-II if low phosphatidylcholine is the main cause, and consequently that GFP expression would increase. Due to the severe growth delay caused by RNAi of pmt-2 we diluted the RNAi bacteria with vector control bacteria. We found that pmt-2 RNAi decreased GFP expression particularly in \textit{Pacdh-1::GFP};\Delta nhr-10 animals (Figures 3.7b, 3.7c). Predictably, choline supplementation rescued the growth defect caused by pmt-2 RNAi. While it is possible that RNAi may cause off-target effects, the rescue by choline suggests that the knock down by pmt-2 RNAi is specific. Importantly, in both choline-supplemented and dilute RNAi conditions, pmt-2 RNAi repressed \textit{Pacdh-1::GFP}. Therefore, we conclude that low phosphatidylcholine is not the inducer of B12-mechanism-II. Instead, our results support a model in which low SAM levels activate B12-mechanism-II. pmt-2 RNAi likely represses the \textit{Pacdh-1::GFP} transgene because SAM levels increase when the SAM-dependent methylation reaction converting the head group of phosphatidylethanolamine into the head group of phosphatidylcholine is blocked\textsuperscript{126}. Taken together, our data support a model in which low Met/SAM cycle activity results in low SAM levels, which activates B12-mechanism-II.
Figure 3.7 – Low SAM and not phosphatidylcholine activates B12-mechanism-II. (A) RNA-seq data of pmt-2 mRNA in each of the datasets. PA: propionic acid. Datapoints show each biological replicate and the bar represents the mean. TPM, transcripts per million. P adjusted values are provided in Supplementary files 3.1, 3.2, 3.4. (B) pmt-2 RNAi suppresses GPF expression in Pacdh-1::GFP;Δnhr-10 animals. pmt-2 RNAi experiments were diluted with vector control RNAi to circumvent strong deleterious phenotypes. Boxplot showing median and interquartile range of normalized GFP intensity measurements of fluorescent images. (C) Wild type or Δnhr-10 expressing Pacdh-1::GFP with pmt-2 RNAi with additional dilutions shown. Dilutions by volume of pmt-2 RNAi with vector control are indicated. Insets show brightfield images. (D) Bar graphs of RNA-seq data showing Met/SAM cycle gene expression in wild type and Δnhr-114 animals with and without methionine supplementation. P adjusted values are provided in Supplementary file 3.4. (E) Relative mRNA level expression (fold change) as determined by qRT-PCR of Met/SAM cycle genes whose P adjusted value was high in RNA-seq shown in D. Datapoints show each biological replicate and bar represents the mean. (F) Cartoon illustrating B12-mechanism-II whereby low vitamin B12 reduces Met/SAM cycle activity, leading to the depletion of SAM and the activation of Met/SAM cycle gene expression mediated by nhr-114.

Next, we asked whether nhr-114 is required for the transcriptional response to low SAM. Since SAM is not stable and may not be easily absorbed by C. elegans, we used methionine supplementation, which supports SAM synthesis, except in sams-1 mutant animals. We performed RNA-seq in wild type and Δnhr-114 mutant animals with or without methionine supplementation and found that Met/SAM cycle genes are repressed by methionine supplementation in wild type, but not Δnhr-114 mutant animals (Figures 3.7d, 3.7e, Supplementary file 3.4). Therefore, low SAM levels due to vitamin B12 depletion activate B12-mechanism-II in an nhr-114-dependent manner (Figure 3.7f).

**B12-mechanism-II activates influx and represses efflux of the Met/SAM cycle**

Some of the most strongly regulated vitamin B12-repressed genes include msra-1
and pmp-5$^{24}$ (Figures 3.8a, 3.8b, Supplementary file 3.1). In the forward genetic screen we identified a mutation in pmp-5 that activates the Pacdh-1::GFP transgene in ∆nhr-10 mutant animals in the presence of supplemented vitamin B12. As mentioned above, pmp-5 is an ortholog of human ABCD4, which encodes a vitamin B12 transporter$^{88}$. Thus, increased B12 transport may be used by the animal as a mechanism to increase Met/SAM cycle activity. msra-1 encodes methionine sulfoxide reductase that reduces methionine sulfoxide to methionine (Figure 3.1a). This gene provides an entry point into the Met/SAM cycle by increasing levels of methionine. This observation prompted us to hypothesize that perturbation of Met/SAM cycle activity, either by low dietary vitamin B12 or by genetic perturbations in the cycle, may activate the expression of these genes. Indeed, both genes are induced in the Met/SAM cycle mutants (Figures 3.8a, 3.8b, Supplementary file 3.2). Further, both genes are repressed by methionine supplementation, in an nhr-114 dependent manner (Figures 3.8a, 3.8b, Supplementary file 3.4).
Figure 3.8 – nhr-114 transcriptionally regulates Met/SAM cycle influx and efflux. (A, B, C) RNA-seq data from each of the experiments for msra-1 (A), pmp-5 (B), cbs-1 (C) and nhr-114 (D). Datapoints show each biological replicate and the bar represents the mean. TPM, transcripts per million. P adjusted values are provided in Supplementary files 3.1, 3.2, 3.4.

As a putative transporter of vitamin B12 from the lysosome to the cytosol, pmp-5 is also important for vitamin B12-mechanism-I. Indeed, pmp-5 is upregulated in response to propionate supplementation and is also regulated by nhr-10 and nhr-68 (Figure 3.8b, Supplementary file 3.1). We also noticed that the expression level changes of cbs-1 are opposite of those of msra-1 and pmp-5: cbs-1 is activated by vitamin B12, repressed in Met/SAM cycle mutants, and activated by methionine in an nhr-114-dependent manner (Figure 3.8c, Supplementary files 3.1, 3.2, 3.4). As mentioned above, cbs-1 encodes
cystathionine beta synthase, which converts homocysteine into cystathionine (Figure 3.1a). Reduced cbs-1 expression upon Met/SAM cycle perturbations would therefore likely prevent carbon efflux. Finally, nhr-114 expression itself is repressed by both vitamin B12 and methionine and activated by perturbations in Met/SAM cycle genes (Figure 3.8d, Supplementary files 3.1, 3.2, 3.4). This suggests that nhr-114 activates its own expression, similarly as the auto-activation of nhr-68 in response to propionate accumulation.\textsuperscript{24} nhr-114 expression is not under the control of B12-mechanism-I because it does not change when propionate is supplemented or when nhr-10 is deleted. However, nhr-114 is mildly repressed in \textit{\textDelta}nhr-68 mutants. This suggests that there may be some crosstalk between the two B12 mechanisms\textsuperscript{24} (Figure 3.8d). Taken together, B12-mechanism-II is employed when Met/SAM cycle activity is perturbed to increase Met/SAM cycle gene expression as well as Met/SAM cycle activity and influx, and to decrease Met/SAM cycle efflux (Figure 3.9).
Figure 3.9  – Model of B12-mechanism-II. Vitamin B12 modulates the transcription of Met/SAM cycle genes and controls in/efflux through the sensing of SAM by \textit{nhr-114}. Dashed arrows indicate regulation of metabolic activity. Solid arrows indicate transcriptional regulation. Vitamin B12 increases Met/SAM cycle metabolic activity producing SAM, which represses \textit{nhr-114} transcription thereby reducing the expression of Met/SAM cycle related genes.

Discussion

We have discovered a second mechanism by which vitamin B12 regulates gene expression in \textit{C. elegans}. This B12-mechanism-II is different from B12-mechanism-I, which we previously reported to transcriptionally activate a propionate shunt in response to persistent accumulation of this short-chain fatty acid$^{24}$. B12-mechanism-I is activated under low dietary vitamin B12 conditions, or when the canonical vitamin B12-dependent propionate breakdown pathway is genetically perturbed$^{24,27}$. B12-mechanism-I is elicited by two TFs, \textit{nhr-10} and \textit{nhr-68} that function as a persistence detector in a type I coherent feed-forward loop with AND-logic gate$^{24}$. We unraveled B12-mechanism-II in animals lacking \textit{nhr-10} that cannot employ B12-mechanism-I$^{24}$. B12-mechanism-II targets the other
vitamin B12-dependent metabolic pathway in *C. elegans*: the Met/SAM cycle. It is activated by low activity of this cycle that results in low levels of SAM and the activation of another NHR TF, *nhr-114*, which is not involved in B12-mechanism\(^1\).  

The discovery that Met/SAM cycle activity is sensed by a gene regulatory network resulting in the adjustment of Met/SAM cycle gene expression and in- and efflux modulation indicates that the animal strives to maintain the activity of this cycle in a tight metabolic regime to support development and homeostasis. Too little activity hampers the synthesis of phosphatidylcholine, an essential membrane component required for proliferation and growth, thereby inducing sterility and developmental delay. It is more difficult to assess why excessive Met/SAM cycle activity could be deleterious. One possibility may be the connection between Met/SAM cycle activity and folate status. When homocysteine levels are increased, the *metr-1* reaction is driven toward methionine production and consequently 5-methylenetetrahydrofolate and folate are depleted\(^2\). The folate cycle is essential in maintaining nucleotide and NADPH pools, which are required for biomass and redox homeostasis\(^3\). SAM is not only required for the biosynthesis of phosphatidylcholine; it is also the methyl donor for histone methylation\(^4\). High production of SAM could result in extensive histone modifications, which may result in aberrant overall gene expression\(^5\).

Previously, it has been shown that *sbp-1* is important for the expression of Met/SAM cycle genes. Interestingly, however, *sbp-1* is not specific to B12-
mechanism-II as its perturbation also affects the response to propionate accumulation\(^2\). Since \textit{sbp-1} functions at a high level in the intestinal gene regulatory network, \textit{i.e.}, it influences many gene expression programs\(^1\), it is likely that \textit{sbp-1} responds to multiple different metabolic imbalances and activates each B12-mechanism. In support of this, \textit{sbp-1} activates both \textit{nhr-68} and \textit{nhr-114}, which are important for B12-mechanism-I and B12-mechanism-II, respectively\(^1\)\(^5\)\(^3\)\(^5\). Future studies will reveal the detailed wiring of \textit{sbp-1}, \textit{nhr-10}, \textit{nhr-68} and \textit{nhr-114} into increasingly intricate gene regulatory networks.

The precise molecular mechanism by which the different NHRs mediate the two B12-mechanisms remains to be elucidated. We have previously shown that NHR-10 physically interacts with the \textit{acdhl-1} promoter\(^1\)^\(^9\)^\(^3\)^\(^4\). However, we did not detect any insightful physical promoter-DNA or protein-protein interactions for either NHR-68 or NHR-114, and therefore it is not yet known which promoters or other TFs these TFs bind to\(^1\)^\(^9\)^\(^2\)^\(^2\)\(^2\). It is also not yet clear how high propionate or low SAM levels are sensed. Since NHRs are liganded TFs, an interesting possibility would be that these metabolites function as ligands where propionate would interact with and activate NHR-10 and/or NHR-68, and SAM would interact with and inhibit NHR-114. Alternatively, the ratio between SAM and S-adenosylhomocysteine (SAH, Figure 3.1a) could be sensed by \textit{nhr-114}, either directly or indirectly. Future detailed studies of propionate shunt and Met/SAM cycle gene promoters will likely provide insights into the molecular mechanisms governing both B12-mechanisms.
We have used the *Pacdh-1::GFP* dietary reporter as a crucial tool to elucidate B12-mechanism-I and B12-mechanism-II. While the reason for activating *acdh-1* expression by B12-mechanism-I is clear (it supports an alternate propionate breakdown mechanism), the reason for inducing this gene in response to Met/SAM cycle perturbations remains unclear. There are two possibilities: either perturbations in the Met/SAM cycle produce propionate which requires ACDH-1 to be degraded, or ACDH-1 catalyzes another reaction in addition to the conversion of propionyl-CoA into acrylyl-CoA. Future studies are required to determine if and how *acdh-1* functions in Met/SAM cycle metabolism.

Taken together, we have uncovered a second mechanism of gene regulation by vitamin B12 that ensures the activity of Met/SAM cycle metabolism to be in a tight, homeostatic regime.

**Materials and methods**

*C. elegans strains*

Animals were maintained on nematode growth media (NGM) as described\textsuperscript{228} with the following modifications. Soy peptone (Thomas Scientific) was used in place of bactopeptone and 0.64 nM vitamin B12 was added to maintain strains. N2 (Bristol) was used as the wild type strain. Animals were fed a diet of *E. coli* OP50 unless otherwise noted. The *wwls24[Pacdh-1::GFP + unc-119(+)]* (VL749) strain was described previously\textsuperscript{29,34}. *nhr-114(gk849), metr-1(ok521)*, and *sams-1(ok 2946)* were retrieved from the *C. elegans* Gene Knock-out Consortium (CGC), and *nhr-
10(tm4695) was obtained from the National Bioresource Project, Japan. All mutant strains were backcrossed three times with N2 wild type animals and crossed with VL749 prior to use in experiments. VL868 [wwIs24[Pacdh-1::GFP + unc-119(+)];nhr-10(tm4695)], VL1127 [wwIs24[Pacdh-1::GFP + unc-119(+)];nhr-114(gk849)], VL1102 [wwIs24[Pacdh-1::GFP + unc-119(+)];metr-1(ok521)], and VL1115 [wwIs24[Pacdh-1::GFP + unc-119(+)];sams-1(ok2946)] are referred to in the text as Δnhr-10, Δnhr-114, Δmetr-1 and Δsams-1 respectively. C. elegans strain CB4856 (Hawaiian) strain was obtained from the CGC. All strains and genotypes are listed in Table 3.1.

Table 3.1 – C. elegans strains used in this project.

<table>
<thead>
<tr>
<th>C. elegans strains</th>
<th>Strain</th>
<th>Genotype</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2</td>
<td>wild type</td>
<td></td>
<td>Laboratory reference strain</td>
</tr>
<tr>
<td>TM4695</td>
<td>nhr-10 (tm4695) III</td>
<td></td>
<td>Deletion</td>
</tr>
<tr>
<td>VC1527</td>
<td>nhr-68(gk708) V</td>
<td></td>
<td>Deletion</td>
</tr>
<tr>
<td>VL749</td>
<td>wwIs24[Pacdh-1::GFP;unc-119(+)]</td>
<td></td>
<td>Reporter</td>
</tr>
<tr>
<td>VL868</td>
<td>wwIs24[Pacdh-1::GFP; unc-119(+)]; nhr-10(tm4695)</td>
<td></td>
<td>Reporter; Deletion</td>
</tr>
<tr>
<td>VL1127</td>
<td>wwIs24[Pacdh-1::GFP;unc-119(+)]; nhr-114(gk849)</td>
<td></td>
<td>Reporter; Deletion</td>
</tr>
<tr>
<td>VL1102</td>
<td>wwIs24[Pacdh-1::GFP;unc-119(+)]; metr-1(ok521)</td>
<td></td>
<td>Reporter; Deletion</td>
</tr>
<tr>
<td>VL1115</td>
<td>wwIs24[Pacdh-1::GFP;unc-119(+)]; sams-1(ok2946)</td>
<td></td>
<td>Reporter; Deletion</td>
</tr>
<tr>
<td>CB4856</td>
<td>Wild isolate</td>
<td></td>
<td>Hawaiian strain</td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td>Reporter; Deletion; Point mutant</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>VL1199</td>
<td><code>wwl524[Pacdh-1::GFP;unc-119(+); nhr-10(tm4695);mthf-1(ww50)]</code></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL1200</td>
<td><code>wwl524[Pacdh-1::GFP;unc-119(+); nhr-10(tm4695);sams-1(ww51)]</code></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL1201</td>
<td><code>wwl524[Pacdh-1::GFP;unc-119(+); nhr-10(tm4695);metr-1(ww52)]</code></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VL1205</td>
<td><code>wwl524[Pacdh-1::GFP;unc-119(+); nhr-10(tm4695);mtrr-1(ww56)]</code></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Bacterial strains**

*E. coli* OP50 and *E. coli* HT115 were obtained from the CGC and grown from single colony to saturation overnight in Luria-Bertani Broth (LB) at 37°C, shaking at 200 rpm. *E. coli* HT115 carrying RNAi plasmids were maintained on 50 µg/mL ampicillin.

**GFP intensity measurement using image analysis**

Raw images were analyzed using Fiji/ImageJ v1.53. Animals in brightfield images were outlined manually using the selection tool. Measurements of area, integrated density and mean gray value were redirected to the same animal in the corresponding fluorescent image. Several surrounding background measurements were also selected, and their mean gray values were averaged. Corrected total fluorescence was calculated by subtracting the product of the object’s area and the mean gray value of the background from the object’s integrated density as described previously.
EMS screen

The EMS mutagenesis protocol was adapted from Jorgensen and Mango 2002. Pacdh-1::GFP; nhr-10(tm4695) animals were treated with 50 mM EMS (Sigma) for four hours and then washed five times with M9 buffer. Mutagenized animals were allowed to recover on NGM agar plates seeded with E. coli OP50, and 200 animals were picked and transferred to NGM agar plates containing 20 nM vitamin B12. F2 animals were screened for the presence of GFP. At least 8,000 haploid genomes were screened, and 27 homozygous mutants were selected, 16 of which remained viable.

Mutant mapping

Chromosome assignment was done by crossing EMS mutants into the CB4856 (Hawaiian) strain. Separate pools of GFP positive and GFP negative F2 animals were mapped using single nucleotide polymorphisms as described.

Whole genome sequencing

Mutant strains ww50, ww51, ww52, ww53, ww54, ww55 and ww56 were backcrossed four to five times to the VL868 [wwls24[Pacdh-1::GFP + unc-119(+)]; nhr-10(tm4695)] parental strain prior to sequencing. Genomic DNA was prepared by phenol-chloroform extraction and ethanol precipitation. Fragmentation was carried out on a Covaris sonicator E220 and 300-400 bp size fragments were
collected using AMpure beads. Libraries were prepared and barcoded using the Kapa hyper prep kit (KK8500). Samples were sequenced at the core facility of the University of Massachusetts Medical school on an Illumina HiSeq4000 using 50 bp paired-end reads. After filtering out low-quality reads, 300 million reads were recovered resulting in an 18X average coverage of the genome. Reads were mapped to the *C. elegans* reference genome version WS220 and analyzed using the CloudMap pipeline\textsuperscript{233} where mismatches were compared to the parental strain as well as to the other sequenced mutants. Variants with unique mismatches were validated by restriction fragment length polymorphism PCR and sanger sequencing. Results of variant calling from CloudMap can be seen in **Supplementary file 3.5.**

**RNAi screen**

RNAi screening was carried out as described\textsuperscript{13}. Briefly, RNAi clones were cultured in 96 well deep-well dishes in LB containing 50 µg/ml ampicillin and grown to log-phase at 37°C. Clone cultures were concentrated to 20-fold in M9 buffer and 10 µL was plated onto a well of a 96 well plate containing NGM agar with 2 mM Isopropyl β- d-1-thiogalactopyranoside (IPTG, Fisher Scientific). Plates were dried and stored at room temperature. The next day approximately 15-20 synchronized L1 animals per well were plated, followed by incubation at 20°C. Plates were screened 72 hours later. The metabolic gene RNAi screen using VL868 [wwls24[Pacdh-1::GFP + unc-119(+);nhr-10(tm4695)] animals was performed
twice. The TF RNAi screen using the Met/SAM cycle mutants generated by EMS was performed six times. All final hits were sequence-verified and retested on 35 mm NGM agar plates with approximately 200 animals per condition.

**Expression profiling by RNA-seq**

Animals were treated with NaOH-buffered bleach, L1 arrested and plated onto NGM plates supplemented with 20 nM vitamin B12 and fed *E. coli* OP50. 400 late L4/early young adult animals were picked into M9 buffer, washed three times and flash frozen in liquid nitrogen. Total RNA was extracted using TRIzol (ThermoFisher), followed by DNase I (NEB) treatment and purified using the Direct-zol RNA mini-prep kit (Zymo research). RNA quality was verified by agarose gel electrophoresis and expression of known genes were measured via qRT-PCR for quality control. Two biological replicates were sequenced by BGI on the BGISEQ-500 next generation sequencer platform using 100 bp paired-end reads. A minimum of approximately 40 million reads was obtained per sample. Raw reads were processed on the DolphinNext RSEM v1.2.28 pipeline revision 7. In brief, the reads were mapped by bowtie2 to genome version c_elegans.PRJNA13758 (WormBase WS271), and then passed to RSEM for estimation of TPM and read counts. Default parameters were used for both bowtie2 and RSEM.

For later RNA-seq experiments we have developed a more cost-effective, in-house method for RNA-sequencing. Briefly, multiplexed libraries were prepared using Cel-seq2. Two biological replicates were sequenced with a NextSeq
500/550 High Output Kit v2.5 (75 Cycles) on a Nextseq500 sequencer. Paired end sequencing was performed; 13 cycles for read 1, 6 cycles for the illumina index and 73 cycles for read 2. Approximately 12 million reads per sample was achieved.

The libraries were first demultiplexed by a homemade python script, and adapter sequences were trimmed using trimmomatic-0.32 by recognizing polyA and barcode sequences. Then, the alignment to the reference genome was performed by STAR with the parameters “--runThreadN 4 --alignIntronMax 25000 --outFilterIntronMotifs RemoveNoncanonicalUnannotated”. Features were counted by ESAT\textsuperscript{236} with the parameters “-task score3p -wLen 100 -wOlap 50 -wExt 1000 -sigTest .01 -multimap normal -scPrep -umiMin 1”. Features in “c\_elegans.PRJNA13758.WS271.canonical\_geneset.gtf” were used as the annotation table input for ESAT, but pseudogenes were discarded. The read counts for each gene were used in differential expression analysis by DEseq2 package in R 3.6.3\textsuperscript{237}. A fold change cut off of greater than 1.5 and \( P \) adjusted value cut off of less than 0.01 was used. All the processing procedures were done in a homemade DolphinNext pipeline.

The RNA-sequencing data files were deposited in the NCBI Gene Expression Omnibus (GEO) under the following accession numbers:

- Bulcha et al. 2019: GSE123507
- This study: GSE151848
Expression profiling by qRT-PCR

Animals were grown and harvested, and RNA was extracted as described for the RNA-seq experiments. qRT-PCR was performed as described previously\textsuperscript{24}. cDNA was reverse transcribed from total RNA using oligo(dT) 12-18 primer (Invitrogen) and Mu-MLV Reverse Transcriptase (NEB). qPCR primers were designed using the GETprime database\textsuperscript{238}. qPCR reactions were carried out in technical triplicate using the StepOnePlus Real-Time PCR system (Applied Biosystems) and Fast Sybr Green Master Mix (ThermoFisher Scientific). Relative mRNA transcript abundance was calculated using the $\Delta\Delta$CT methods\textsuperscript{239} and normalized to the geometric mean of $ama$-1 and $act$-1 levels. Primers are listed in Table 3.2.

Table 3.2 – qPCR primers used in this project.

<table>
<thead>
<tr>
<th>Primer</th>
<th>Forward sequence</th>
<th>Reverse sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$act$-1</td>
<td>CTCCTGGCCCATCAACCATG</td>
<td>CTTGCTTGGAGATCCACATC</td>
</tr>
<tr>
<td>$ama$-1</td>
<td>AGTGCCGAGATTGAAGAGA</td>
<td>GTATTGCATGTTACCTTTTTCAACG</td>
</tr>
<tr>
<td>$metr$-1</td>
<td>GGAGCAGCTACTGGTAGAC</td>
<td>CACAGATGGCGAAATTGAGAG</td>
</tr>
<tr>
<td>$mtrr$-1</td>
<td>TACGTTCTTCTCGGTCTCG</td>
<td>AGAGCTGTCAGTTGTTGTC</td>
</tr>
<tr>
<td>$sams$-1</td>
<td>ATTATCAAGGAGCCTCGACCT</td>
<td>ATGGGAACTCAGAGTGACC</td>
</tr>
<tr>
<td>$ahcy$-1</td>
<td>CGATTGGAGAGATTGACGTG</td>
<td>GTGTAACGGTCAACCTGTG</td>
</tr>
<tr>
<td>$mel$-32</td>
<td>TGACTCATGGATTCTCTCACCC</td>
<td>GATCAACCTGTATGGAAGAGAC</td>
</tr>
<tr>
<td>$mthf$-1</td>
<td>GTTGAGACCGAGATGAGAATGC</td>
<td>TTCATAATGCTTTGGTGACCAG</td>
</tr>
<tr>
<td>$pmt$-2</td>
<td>TTCATGCAGAAAATTTACCA</td>
<td>GTCTTTCTCGATGTACCC</td>
</tr>
<tr>
<td>$cbs$-1</td>
<td>GAAGCTAGAGTATCTCAATATTGCG</td>
<td>CCAATCTCCTACGAACCTGG</td>
</tr>
</tbody>
</table>
Body size measurement

Approximately 100 synchronized L1 animals were plated across four wells of a 48 well plate per condition. L4 animals were collected and washed three times in 0.03% sodium azide and transferred to a 96 well plate. Excess liquid was removed, and plates were rested for an hour to allow animals to settle and straighten. Pictures were taken using an Evos Cell Imaging System microscope and image processing was done using a MATLAB (MathWorks) script named “wormFinder.m” written in-house and made available at the following link: https://github.com/shiaway/wormFinder/blob/master/wormFinder.m

Gas chromatography-mass spectrometry

For Figure 5A, gravid adults were harvested from liquid S media cultures supplemented with or without 64 nM vitamin B12 and fed concentrated E. coli OP50. For Figure 5E, gravid adults were harvested from NGM agar plates treated with 64 nM vitamin B12 and seeded with E. coli OP50. Animals were washed in 0.9% saline until the solution was clear and then twice more (3-6 times total). Metabolites were extracted and analyzed as described previously\textsuperscript{240}. Briefly, 50 µL of a semi-soft pellet of animals was transferred to a 2 mL FastPrep tube (MP Biomedicals) and flash frozen with liquid nitrogen. Metabolites were extracted in 80% cold methanol. Acid-washed micro glass beads (Sigma) and a FastPrep-24 5G homogenizer (MP Biomedicals) were used to disrupt animal bodies. After
settling, supernatant was transferred to glass vials (Sigma) and dried by speed-vac overnight. MeOX-MSTFA derivatized samples were analyzed on an Agilent 7890B/5977B single quadrupole GC-MS equipped with an HP-5ms Ultra Inert capillary column (30 m × 0.25 mm × 0.25 μm) using the same method as described\textsuperscript{240}.
CHAPTER IV: WORMPATHS: CAENORHABDITIS ELEGANS METABOLIC PATHWAY ANNOTATION AND VISUALIZATION

Preface

This work was conducted primarily during the Covid19 lockdown in 2020 by the whole lab as a group effort. It has been accepted at the journal Genetics in 2021 with the title “WormPaths: Caenorhabditis elegans metabolic pathway annotation and visualization”241. Melissa Walker, Amy Holdorf and I are co-first authors. Marian Walhout and Safak Yilmaz are co-corresponding authors. The other authors are: Sushila Bhattacharya, Cédric Diot, Aurian García-González, Brent Horowitz, Yong-Uk Lee, Thomas Leland, Xuhang Li, Zeynep Mirza, Huimin Na, Shivani Nanda, Olga Ponomarova, Hefei Zhang and Jingyan Zhang.

Gene to pathway annotations were done by Safak Yilmaz, Amy Holdorf, and Marian Walhout with input from all other authors. All authors participated in WormPaths maps design and drawing. SVG maps were drawn by Melissa Walker with help from Amy Holdorf and myself. Metabolite structures were retrieved by Tom Leland or hand drawn as needed by myself. Computational analysis and website design were performed by Safak Yilmaz.

Summary

In our group, we aim to understand metabolism in the nematode Caenorhabditis elegans (C. elegans) and its relationships with gene expression, physiology and the response to therapeutic drugs. Visualization of the metabolic pathways that
comprise the metabolic network is extremely useful for interpreting a wide variety of experiments. Detailed annotated metabolic pathway maps for C. elegans are mostly limited to pan-organismal maps, many with incomplete or inaccurate pathway and enzyme annotations. Here we present WormPaths, which is composed of two parts: 1) the careful manual annotation of metabolic genes into pathways, categories and levels, and 2) 62 pathway maps that include metabolites, metabolite structures, genes, reactions, and pathway connections between maps. These maps are available on the WormFlux website. We show that WormPaths provides easy-to-navigate maps and that the different levels in WormPaths can be used for metabolic pathway enrichment analysis of transcriptomic data. In the future, we envision further developing these maps to be more interactive, analogous to road maps that are available on mobile devices.

**Introduction**

Metabolism can be broadly defined as the total complement of reactions that degrade and synthesize biomolecules to produce the biomass and generate the energy organisms need to grow, function and reproduce. Metabolic reactions function in metabolic pathways that are interconnected to form the metabolic network. In metabolic networks, the nodes are metabolites and the edges are conversion and transport reactions carried out by metabolic enzymes and transporters.

Genome-scale metabolic network models provide mathematical tools that
are invaluable for the systems-level analysis of metabolism. Such models have been constructed for numerous organisms, including bacteria, yeast, the nematode *C. elegans* and humans\(^42\). Metabolic network models are extremely useful because they can be used with flux balance analysis (FBA) to derive specific insights and hypotheses. For example, gene expression profiling data can be used to gain insight into metabolic network activity at pathway, reaction and metabolite levels under different conditions or in particular tissues\(^{44,242-244}\).

Visualizing the metabolic pathways that together comprise the metabolic network of an organism aids in the interpretation of results from different types of large-scale, systems-level studies such as gene expression profiling by RNA-seq, phenotypic screens by RNAi or CRISPR/Cas9, and genetic interaction mapping. Several resources are available online for the visualization and navigation of metabolic pathways. Probably the most widely used is the Kyoto Encyclopedia of Genes and Genomes (KEGG), a platform that provides pan-organism annotations and metabolic pathway maps\(^{245}\). Other online resources include MetaCyc\(^{246}\), BRENDA\(^{247}\), REACTOME\(^{248}\). While all of these platforms are extremely useful resources for metabolic pathway mapping, enzyme classification, and pathway visualization, they can have incomplete or incorrect pathway and enzyme information due to a lack of extensive manual curations for specific organisms. As a result, map navigation can be rather non-intuitive.

Over the last five decades or so, the free-living nematode *C. elegans* has proven to be an excellent genetic model to gain insights into a variety of biological
processes, including development, reproduction, neurobiology/behavior, and aging\textsuperscript{249–251}. More recently, \textit{C. elegans} has emerged as a powerful model to understand basic metabolic processes\textsuperscript{252,253}. \textit{C. elegans} is a bacterivore that can be fed different bacterial species and strains in the lab\textsuperscript{218,254}. Numerous studies have begun to shed light on the metabolic mechanisms by which different bacterial diets can affect the animal's metabolism\textsuperscript{16,26,29,255–258}. For instance, we have discovered that when fed a diet low in vitamin B12, \textit{C. elegans} adjusts the two metabolic pathways that rely on this cofactor. Specifically, it rewires propionate degradation by transcriptionally activating a propionate shunt and upregulates methionine/S-adenosylmethionine (Met/SAM) cycle genes to adjust cycle activity\textsuperscript{23,24,97}. To enable more global analyses of \textit{C. elegans} metabolism, we previously reconstructed its first genome-scale metabolic network model\textsuperscript{44}. The recently updated version of this model includes 1,314 genes, 907 metabolites and 2,230 reactions, and is referred to as iCEL1314\textsuperscript{244}. Information about this network and all the components involved is publicly available on our WormFlux website (http://wormflux.umassmed.edu).

Over time, we found that we were missing metabolic pathway maps that are easy to navigate and that can be used to help interpret results from phenotypic screens and gene expression profiling experiments. We used KEGG pathways, which provide generic, non-organism-specific visualizations, as a starting point to redraw maps of \textit{C. elegans} metabolism on paper to help us interpret our data. In KEGG, enzymes are indicated by Enzyme Commission numbers and maps are
colored with those enzymes predicted to be present in an organism of interest; however, organism-specific pathways cannot be extracted. Further, many of these maps contain incorrect or partially correct reactions for *C. elegans*. We found that redrawing pathway maps that contain information about metabolites, genes encoding the proteins that catalyze metabolic reactions or transport metabolites between cells or cellular compartments, molecular structures, and cofactors was very helpful to our studies\textsuperscript{24,27,97}.

Here we present WormPaths, a web-based collection of standardized metabolic pathway maps for *C. elegans*. In total, WormPaths contains 62 maps covering major metabolic pathways (glycolysis/gluconeogenesis, TCA cycle, etc.), amino acid metabolism, and pathways fundamental to *C. elegans* physiology (collagen biosynthesis, ascaroside biosynthesis, propionate degradation, etc.). Each map connects to other pathways, thereby covering the entire iCEL1314 network. Importantly, the network was expanded by adding reactions and genes found in the literature that were heretofore missed. Maps were carefully curated, hand-drawn, and then visualized in a standardized Scalable Vector Graphics (SVG) format, which allows interactive usage in web applications. WormPaths annotations and maps are publicly available on the WormFlux website (http://wormflux.umassmed.edu). Our careful gene-to-pathway annotations at different levels (see Results) enable statistical enrichment analyses. Finally, our maps may provide a useful format for the drawing of metabolic pathway maps in
other organisms. In the future, we envision further refining the maps through detailed literature reviews and experiments.

**Results**

*Assigning *C. elegans* metabolic genes to pathways at different levels*

To generate WormPaths, we built on available resources, most notably the iCEL1314 metabolic network model\(^{244}\), KEGG\(^{245}\), MetaCyc\(^{246}\), WormBase\(^{259}\), and literature searches (*Figure 4.1a*). Briefly, we manually curated each of the 1314 genes present in the iCEL1314 model and assigned them to one or more biochemical pathway (see Methods). In addition, we used a “category” annotation for metabolic genes that best fit in groups or enzyme categories rather than specific biochemical pathways (*Figure 4.1b*). Examples of categories include the 40 iCEL1314 genes in complex I of the electron transport chain (ETC), guanylate cyclases that convert guanosine triphosphate (GTP) to cyclic guanosine monophosphate (cGMP), and vacuolar ATPases that maintain proton gradients across organellar plasma membranes. Because all metabolic pathways are connected into a metabolic network and some pathways are embedded, or nested, into larger pathways, we decided to annotate *C. elegans* metabolic pathways at different levels. Categorizing genes into pathways at different levels, enables enrichment analyses at different levels of resolution (see below). Level 1 includes the broadest assignment to ten annotations: amino acids, carbohydrates, cofactors and vitamins, energy, lipids, nucleotides, one-carbon cycle, reactive oxygen
species, other amino acids, and other (Table 4.1). Levels 2, 3 and 4 further refine pathways within Level 1 annotations. For instance, the propionate shunt\(^27\) (Level 4) is part of propionate degradation (Level 3), which is part of short-chain fatty acid degradation (Level 2), which is part of lipids (Level 1) (Figure 4.1c, Table 4.1). Altogether, there are 10 groups of pathways or categories at Level 1, 61 groups at Level 2, 79 groups at Level 3, and 85 groups at Level 4. Not all Levels 2 or 3 can be further subdivided, and therefore there is redundancy at the higher levels (3 and 4) (Table 4.1). For each pathway, we decided as a group which level would be most useful for visualization as a map and a team of two lab members worked together to design and draw a draft map (Table 4.2). For example, Level 2 branched-chain amino acid degradation can be subdivided into three maps at Level 3: isoleucine, leucine, and valine degradation, each of which is visualized separately. Another example is methionine metabolism (Level 3), which can be further refined to methionine salvage and the Met/SAM cycle (Level 4). Other amino acids need no further categorization and maps are drawn at Level 2, such as histidine and lysine degradation.
Figure 4.1 – WormPaths annotation of *C. elegans* metabolic genes. (A) Cartoon outlining resources used to generate WormPaths. (B) Pipeline of gene to pathway/category annotations and map construction. (C) Example of pathway-centered WormPaths annotations. (D) Example of gene-centered WormPaths annotations. SVG, scalable vector graphics; SCFA, short chain fatty acids; BCAA, branched-chain amino acids.
Table 4.1 – Pathways at levels 1 through 4.

<table>
<thead>
<tr>
<th>LEVEL 1</th>
<th>LEVEL 2</th>
<th>LEVEL 3</th>
<th>LEVEL 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMINO ACIDS</td>
<td>ALA, ASP/ASN AND GLU/GLN METABOLISM</td>
<td>ALANINE METABOLISM</td>
<td>ALANINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>ALA, ASP/ASN AND GLU/GLN METABOLISM</td>
<td>ASPARTATE/ASPARAGINE METABOLISM</td>
<td>ASPARTATE/ASPARAGINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>ALA, ASP/ASN AND GLU/GLN METABOLISM</td>
<td>GLUTAMATE/GLUTAMINE METABOLISM</td>
<td>GLUTAMATE/GLUTAMINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>ARG AND PRO METABOLISM</td>
<td>ARGinine METABOLISM</td>
<td>ARGinine METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>ARG AND PRO METABOLISM</td>
<td>PROLINE METABOLISM</td>
<td>PROLINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>BRANCHED-CHAIN AMINO ACID DEGRADATION</td>
<td>ISOLEUCINE DEGRADATION</td>
<td>ISOLEUCINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>BRANCHED-CHAIN AMINO ACID DEGRADATION</td>
<td>LEUCINE DEGRADATION</td>
<td>LEUCINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>BRANCHED-CHAIN AMINO ACID DEGRADATION</td>
<td>VALINE DEGRADATION</td>
<td>VALINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>CYS AND MET METABOLISM</td>
<td>CYSTEINE METABOLISM</td>
<td>CYSTEINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>CYS AND MET METABOLISM</td>
<td>METHIONINE METABOLISM</td>
<td>METHIONINE SALVAGE</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>GLY, SER AND THR METABOLISM</td>
<td>GLYCINE AND THREONINE METABOLISM</td>
<td>GLYCINE AND THREONINE METABOLISM OTHER</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>GLY, SER AND THR METABOLISM</td>
<td>GLYCINE AND THREONINE METABOLISM</td>
<td>GLYCINE CLEAVAGE SYSTEM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>GLY, SER AND THR METABOLISM</td>
<td>SERINE METABOLISM</td>
<td>SERINE METABOLISM</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>HISTIDINE DEGRADATION</td>
<td>HISTIDINE DEGRADATION</td>
<td>HISTIDINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>LYSINE DEGRADATION</td>
<td>LYSINE DEGRADATION</td>
<td>LYSINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>PHENYLALANINE DEGRADATION</td>
<td>PHENYLALANINE DEGRADATION</td>
<td>PHENYLALANINE DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>TRYPHTOPHAN DEGRADATION</td>
<td>TRYPHTOPHAN DEGRADATION</td>
<td>TRYPHTOPHAN DEGRADATION</td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td>TYROSINE METABOLISM</td>
<td>TYROSINE METABOLISM</td>
<td>TYROSINE METABOLISM</td>
</tr>
<tr>
<td>CARBOHYDRATES</td>
<td>AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM</td>
<td>AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM</td>
<td>AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>GLYCOLYSIS/ GLUCONEOGENESIS</td>
<td>GLYCOLYSIS/ GLUCONEOGENESIS</td>
<td>GLYCOLYSIS/ GLUCONEOGENESIS</td>
<td>GLYCOLYSIS/ GLUCONEOGENESIS</td>
</tr>
<tr>
<td>ORGANIC ACIDS</td>
<td>GLYOXYLATE AND DICARBOXYLATE METABOLISM</td>
<td>GLYOXYLATE AND DICARBOXYLATE METABOLISM</td>
<td>GLYOXYLATE AND DICARBOXYLATE METABOLISM</td>
</tr>
<tr>
<td>ORGANIC ACIDS</td>
<td>PYRUVATE METABOLISM</td>
<td>PYRUVATE METABOLISM</td>
<td>PYRUVATE METABOLISM</td>
</tr>
<tr>
<td>PENTOSE AND GLUCURONATE INTERCONVERSIONS</td>
<td>PENTOSE AND GLUCURONATE INTERCONVERSIONS</td>
<td>PENTOSE AND GLUCURONATE INTERCONVERSIONS</td>
<td>PENTOSE AND GLUCURONATE INTERCONVERSIONS</td>
</tr>
<tr>
<td>PENTOSE PHOSPHATE PATHWAY</td>
<td>PENTOSE PHOSPHATE PATHWAY</td>
<td>PENTOSE PHOSPHATE PATHWAY</td>
<td>PENTOSE PHOSPHATE PATHWAY</td>
</tr>
<tr>
<td>SUGARS</td>
<td>FRUCTOSE AND MANNOSE METABOLISM</td>
<td>FRUCTOSE AND MANNOSE METABOLISM</td>
<td>FRUCTOSE AND MANNOSE METABOLISM</td>
</tr>
<tr>
<td>SUGARS</td>
<td>GALACTOSE METABOLISM</td>
<td>GALACTOSE METABOLISM</td>
<td>GALACTOSE METABOLISM</td>
</tr>
<tr>
<td>SUGARS</td>
<td>N-GLYCAN BIOSYNTHESIS</td>
<td>N-GLYCAN BIOSYNTHESIS</td>
<td>N-GLYCAN BIOSYNTHESIS</td>
</tr>
<tr>
<td>SUGARS</td>
<td>STARCH AND SUCROSE METABOLISM</td>
<td>STARCH AND SUCROSE METABOLISM</td>
<td>STARCH AND SUCROSE METABOLISM</td>
</tr>
<tr>
<td>IRON METABOLISM</td>
<td>IRON METABOLISM</td>
<td>IRON METABOLISM</td>
<td>IRON METABOLISM</td>
</tr>
<tr>
<td>NICOTINAMIDE METABOLISM</td>
<td>NICOTINAMIDE METABOLISM</td>
<td>NICOTINAMIDE METABOLISM</td>
<td>NICOTINAMIDE METABOLISM</td>
</tr>
<tr>
<td>PANTOTHENATE AND COA BIOSYNTHESIS</td>
<td>PANTOTHENATE AND COA BIOSYNTHESIS</td>
<td>PANTOTHENATE AND COA BIOSYNTHESIS</td>
<td>PANTOTHENATE AND COA BIOSYNTHESIS</td>
</tr>
<tr>
<td>PORPHYRIN METABOLISM</td>
<td>PORPHYRIN METABOLISM</td>
<td>PORPHYRIN METABOLISM</td>
<td>PORPHYRIN METABOLISM</td>
</tr>
<tr>
<td>RETINOL METABOLISM</td>
<td>RETINOL METABOLISM</td>
<td>RETINOL METABOLISM</td>
<td>RETINOL METABOLISM</td>
</tr>
<tr>
<td>RIBOFLAVIN METABOLISM</td>
<td>RIBOFLAVIN METABOLISM</td>
<td>RIBOFLAVIN METABOLISM</td>
<td>RIBOFLAVIN METABOLISM</td>
</tr>
<tr>
<td>UBIQUINONE METABOLISM</td>
<td>UBIQUINONE METABOLISM</td>
<td>UBIQUINONE METABOLISM</td>
<td>UBIQUINONE METABOLISM</td>
</tr>
<tr>
<td>VITAMIN B12 METABOLISM</td>
<td>VITAMIN B12 METABOLISM</td>
<td>VITAMIN B12 METABOLISM</td>
<td>VITAMIN B12 METABOLISM</td>
</tr>
<tr>
<td>COFACTORS AND VITAMINS</td>
<td>VITAMIN B6 METABOLISM</td>
<td>VITAMIN B6 METABOLISM</td>
<td>VITAMIN B6 METABOLISM</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>COFACTORS AND VITAMINS</td>
<td>FOLATE BIOSYNTHESIS</td>
<td>FOLATE BIOSYNTHESIS</td>
<td>FOLATE BIOSYNTHESIS</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSFER FLAVOPROTEIN</td>
<td>ELECTRON TRANSFER FLAVOPROTEIN</td>
<td>ELECTRON TRANSFER FLAVOPROTEIN</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>ETC-COMPLEX I</td>
<td>ETC-COMPLEX I</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>ETC-COMPLEX II</td>
<td>ETC-COMPLEX II</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>ETC-COMPLEX III</td>
<td>ETC-COMPLEX III</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>ETC-COMPLEX IV</td>
<td>ETC-COMPLEX IV</td>
</tr>
<tr>
<td>ENERGY</td>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>ETC-COMPLEX V</td>
<td>ETC-COMPLEX V</td>
</tr>
<tr>
<td>ENERGY</td>
<td>SULFUR METABOLISM</td>
<td>SULFUR METABOLISM</td>
<td>SULFUR METABOLISM</td>
</tr>
<tr>
<td>ENERGY</td>
<td>TRICARBOXYLIC ACID CYCLE</td>
<td>TRICARBOXYLIC ACID CYCLE</td>
<td>TRICARBOXYLIC ACID CYCLE</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>ETHER LIPID METABOLISM</td>
<td>ETHER LIPID METABOLISM</td>
<td>ETHER LIPID METABOLISM</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>BCFA BIOSYNTHESIS</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>FATTY ACID BIOSYNTHESIS OTHER</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>MUFA BIOSYNTHESIS</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>PUFA BIOSYNTHESIS</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID DEGRADATION</td>
<td>MITOCHONDRIAL FATTY ACID DEGRADATION</td>
<td>MITOCHONDRIAL FATTY ACID DEGRADATION</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID DEGRADATION</td>
<td>PEROXISOMAL FATTY ACID DEGRADATION</td>
<td>PEROXISOMAL FATTY ACID DEGRADATION</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>FATTY ACID DEGRADATION</td>
<td>FATTY ACID DEGRADATION OTHER</td>
<td>FATTY ACID DEGRADATION OTHER</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>GLYCEROLIPID METABOLISM</td>
<td>GLYCEROLIPID METABOLISM</td>
<td>GLYCEROLIPID METABOLISM</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>GLYCEROPHOSPHOLIPID METABOLISM</td>
<td>GLYCEROPHOSPHOLIPID METABOLISM</td>
<td>GLYCEROPHOSPHOLIPID METABOLISM</td>
</tr>
<tr>
<td>Category</td>
<td>Metabolism 1</td>
<td>Metabolism 2</td>
<td>Metabolism 3</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------------------------------</td>
<td>---------------------------------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>SHORT CHAIN FATTY ACID DEGRADATION</td>
<td>PROPIONATE DEGRADATION</td>
<td>PROPIONATE DEGRADATION CANONICAL</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>SHORT CHAIN FATTY ACID DEGRADATION</td>
<td>PROPIONATE DEGRADATION</td>
<td>PROPIONATE DEGRADATION OTHER</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>SHORT CHAIN FATTY ACID DEGRADATION</td>
<td>PROPIONATE DEGRADATION</td>
<td>PROPIONATE DEGRADATION SHUNT</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>SPHINGOLIPID METABOLISM</td>
<td>SPHINGOLIPID METABOLISM</td>
<td>SPHINGOLIPID METABOLISM</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>STEREOID METABOLISM</td>
<td>BIOSYNTHESIS OF BILE ACID-LIKE MOLECULES</td>
<td>BIOSYNTHESIS OF BILE ACID-LIKE MOLECULES</td>
</tr>
<tr>
<td>LIPIDS</td>
<td>STEREOID METABOLISM</td>
<td>STEREOID METABOLISM OTHER</td>
<td>STEREOID METABOLISM OTHER</td>
</tr>
<tr>
<td>NUCLEOTIDES</td>
<td>PURINE METABOLISM</td>
<td>PURINE METABOLISM</td>
<td>PURINE METABOLISM</td>
</tr>
<tr>
<td>NUCLEOTIDES</td>
<td>PYRIMIDINE METABOLISM</td>
<td>PYRIMIDINE METABOLISM</td>
<td>PYRIMIDINE METABOLISM</td>
</tr>
<tr>
<td>ONE CARBON CYCLE</td>
<td>FOLATE CYCLE</td>
<td>FOLATE CYCLE</td>
<td>FOLATE CYCLE</td>
</tr>
<tr>
<td>ONE CARBON CYCLE</td>
<td>MET/SAM CYCLE</td>
<td>MET/SAM CYCLE</td>
<td>MET/SAM CYCLE</td>
</tr>
<tr>
<td>OTHER</td>
<td>AMINOACYL-TRNA BIOSYNTHESIS</td>
<td>AMINOACYL-TRNA BIOSYNTHESIS</td>
<td>AMINOACYL-TRNA BIOSYNTHESIS</td>
</tr>
<tr>
<td>OTHER</td>
<td>ASCAROSIDE BIOSYNTHESIS</td>
<td>ASCAROSIDE BIOSYNTHESIS</td>
<td>ASCAROSIDE BIOSYNTHESIS</td>
</tr>
<tr>
<td>OTHER</td>
<td>CHITIN BIOSYNTHESIS</td>
<td>CHITIN BIOSYNTHESIS</td>
<td>CHITIN BIOSYNTHESIS</td>
</tr>
<tr>
<td>OTHER</td>
<td>CHITIN DEGRADATION</td>
<td>CHITIN DEGRADATION</td>
<td>CHITIN DEGRADATION</td>
</tr>
<tr>
<td>OTHER</td>
<td>COLLAGEN BIOSYNTHESIS</td>
<td>COLLAGEN BIOSYNTHESIS</td>
<td>COLLAGEN BIOSYNTHESIS</td>
</tr>
<tr>
<td>OTHER</td>
<td>GUANYLATE CYCLASE</td>
<td>GUANYLATE CYCLASE</td>
<td>GUANYLATE CYCLASE</td>
</tr>
<tr>
<td>OTHER</td>
<td>INOSITOL PHOSPHATE METABOLISM</td>
<td>INOSITOL PHOSPHATE METABOLISM</td>
<td>INOSITOL PHOSPHATE METABOLISM</td>
</tr>
<tr>
<td>OTHER</td>
<td>KETONE BODY METABOLISM</td>
<td>KETONE BODY METABOLISM</td>
<td>KETONE BODY METABOLISM</td>
</tr>
<tr>
<td>OTHER</td>
<td>METHYLGLOXAL DETOXIFICATION</td>
<td>METHYLGLOXAL DETOXIFICATION</td>
<td>METHYLGLOXAL DETOXIFICATION</td>
</tr>
<tr>
<td>Pathway Name</td>
<td>Map Level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALANINE METABOLISM</td>
<td>LEVEL 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMINO SUGAR AND NUCLEOTIDE SUGAR METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARGinine METABOLISM</td>
<td>LEVEL 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCAROSIDE BIOSYNTHESIS</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPARTATE/ASPARAGINE METABOLISM</td>
<td>LEVEL 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOSYNTHESIS OF BILE ACID-LIKE MOLECULES</td>
<td>LEVEL 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COLLAGEN BIOSYNTHESIS</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYSTEINE METABOLISM</td>
<td>LEVEL 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELECTRON TRANSPORT CHAIN</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETHER LIPID METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FATTY ACID BIOSYNTHESIS</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOLATE BIOSYNTHESIS</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEVALONATE METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOLYBDENUM COFACTOR BIOSYNTHESIS</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELENOCOMPOUND METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAURINE AND HYPOTaurine METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROS METABOLISM</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VITAMIN B12 ENZYME</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VACUOLAR ATP-ASE</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VITAMIN B12 ENZYME</td>
<td>LEVEL 2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 – All maps and the corresponding level to which each map was drawn.
FOLATE CYCLE LEVEL 2
FRUCTOSE AND MANNOSE METABOLISM LEVEL 3
GALACTOSE METABOLISM LEVEL 3
GLUTAMATE/GLUTAMINE METABOLISM LEVEL 3
GLUTATHIONE METABOLISM LEVEL 2
GLYCEROLIPID METABOLISM LEVEL 2
GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE LEVEL 2
GLYCEROPHOSPHOLIPID METABOLISM LEVEL 2
GLYCEINE AND THREONINE METABOLISM LEVEL 3
GLYCOLYSIS/GLUCONEOGENESIS LEVEL 2
GLYOXYLATE AND DICARBOXYLATE METABOLISM LEVEL 3
HISTIDINE DEGRADATION LEVEL 2
INOSITOL PHOSPHATE METABOLISM LEVEL 2
IRON METABOLISM LEVEL 2
ISOLEUCINE DEGRADATION LEVEL 3
KETONE BODY METABOLISM LEVEL 2
LEUCINE DEGRADATION LEVEL 3
LYSINE DEGRADATION LEVEL 2
MET/SAM CYCLE LEVEL 2
METHIONINE SALVAGE LEVEL 4
METHYLGLYOXAL DETOXIFICATION LEVEL 2
MEVALONATE METABOLISM LEVEL 2
MITOCHONDRIAL FATTY ACID DEGRADATION LEVEL 3
MOLYBDENUM COFACTOR BIOSYNTHESIS LEVEL 2
N-GLYCAN BIOSYNTHESIS LEVEL 3
NICOTINATE AND NICOTINAMIDE METABOLISM LEVEL 2
PANTOTHENATE AND COA BIOSYNTHESIS LEVEL 2
PENTOSE AND GLUCURONATE INTERCONVERSIONS LEVEL 2
PENTOSE PHOSPHATE PATHWAY LEVEL 2
PEROXISOMAL FATTY ACID DEGRADATION LEVEL 3
PHENYLALANINE DEGRADATION LEVEL 2
PROLINE METABOLISM LEVEL 3
PROPIONATE DEGRADATION LEVEL 3
PURINE METABOLISM LEVEL 2
PYRIMIDINE METABOLISM LEVEL 2
PYRUVATE METABOLISM LEVEL 3
RETINOL METABOLISM LEVEL 2
RIBOFLAVIN METABOLISM LEVEL 2
ROS METABOLISM LEVEL 2
In WormPaths, 32% of genes are annotated to multiple pathways. While many genes do in fact act in multiple pathways, others may be annotated to multiple pathways because gene-protein-reaction annotations are based on homologies with known enzymes, and the exact participation of each gene in different pathways cannot be resolved without experimentation. For instance, the acyl-CoA dehydrogenase-encoding gene acdh-1 is annotated to different degradation reactions in amino acid and lipid metabolism (Figure 4.1d, Supplementary files 4.1 – 4.3). However, only its role in the propionate shunt has been experimentally characterized\(^{27}\). Importantly, its close paralog acdh-2 is annotated to the same pathways but was experimentally shown not to be involved in the propionate shunt\(^ {27}\). Future biochemical and genetic studies are needed to disentangle which enzymes can catalyze multiple reactions, and which are specific to individual reactions.
WormPaths maps – visualization and navigation

After map level assignments and pathway design, maps were sketched digitally or by hand and electronically uploaded to Google Docs for sharing followed by manual conversion to SVG format, an Extensible Markup Language (XML)-based vector image format for general useability on the Internet by both individual users and computer programs. Metabolite structure images for all products and reactants were downloaded from KEGG and other resources (see Methods) and some that were not available were hand drawn. All reactions on the SVG maps were manually verified and checked for errors. Maps were then uploaded to the WormFlux webpage, where they are available in a drop-down list. All maps are searchable and clickable. For example, a search for the gene *metr-1* will result in the WormFlux gene page for *metr-1*, which has links to the Met/SAM cycle and folate cycle pathways, each of which brings the corresponding map with the *metr-1* gene highlighted (Figure 4.2). In reverse, clicking on a gene in any map leads to the associated WormFlux page, where key identifiers and reactions in which the gene is involved are listed. The same is true when searching and clicking metabolites.
**Figure 4.2** – An example of using WormPaths to search for a specific gene. A search for the gene *metr-1* will lead to the gene overview, followed by the specific pathway maps that *metr-1* is involved in with the gene highlighted in red.

In total, WormPaths provides 62 maps of *C. elegans* metabolic pathways, that connect into the larger iCEL1314 network. **Figure 4.3a** shows an example of the WormPaths map for glycolysis/gluconeogenesis. This is a Level 2 map that is part of carbohydrates (Level 1). In metabolic networks, nodes are metabolites and edges are the reactions in which these metabolites are converted into one another, or transported between cellular compartments, or between the cell and the extracellular environment. The edges in these maps are black for enzymatic reactions and green for transport reactions (**Figure 4.3b**). The genes encoding the enzymes predicted to catalyze the reactions are indicated in blue, and co-reactants are indicated in orange (**Figures 4.3a, 4.3b**). Some reactions have multiple alternative genes associated with them. These “OR” genes are separated by a vertical bar (|). For example, in glycolysis/gluconeogenesis the interconversion between phosphoenolpyruvate (pep) and oxaloacetate (oaa) is associated with *pck-1*, *pck-2*, or *pck-3* (**Figure 4.3a**). None of these genes is associated with any other reaction and therefore they may function in different conditions or in different tissues. Indeed, at the second larval (L2) stage, two of these three genes show very distinct tissue expression patterns, while the mRNA for the third gene (*pck-3*) was undetectable (**Figure 4.4**). For edges where multiple enzymes together catalyze a reaction, an ampersand (&) is used to indicate “AND” genes. For example, *pdha-1* and *pdhb-1* are both required in the pyruvate dehydrogenase
complex that catalyzes the conversion of pyruvate (pyr) into acetyl-CoA (accoa) (Figure 4.3a).
A

CYTOSOL

EXTRACELLULAR SPACE

MITOCHONDRIA

B

Reactions

gene
Enzymatic reaction
rna
Non-enzymatic reaction
Transport
Multiple transports
Transportable between cytosol
and extracellular space
Transportable between cytosol
and mitochondria
Not in iCEL1314

Metabolites

gene
metabolite
Main metabolite
Co-reactant
metabolite
Not in iCEL1314

Number of pathway connections

None
1
2
3
4
> 4

C

D

amp, h

<table>
<thead>
<tr>
<th>Not in iCEL1314</th>
<th>Not in iCEL1314</th>
</tr>
</thead>
</table>

115
**Figure 4.3 – WormPaths examples.** (A) WormPaths map of glycolysis/gluconeogenesis. (B) The key to the reactions, metabolite transportability, and number of pathway connections that appears on the WormPaths website. (C) An example of a web pop-up window from glycolysis/gluconeogenesis that shows the metabolite structure of beta-D-glucose 6-phosphate upon hovering the cursor over g6p-B. (D) Example of a literature-curated reaction highlighted in the gray box.

**Figure 4.4 – Tissue-specific expression of pck-1 and pck-2.** TPM, transcripts per million.

For metabolite names both in WormPaths (Figure 4.3a) and in WormFlux\(^\text{44}\) we used Biochemical Genetic and Genomic (BiGG) database abbreviations where available\(^\text{261}\). The transportability of metabolites between subcellular compartments is indicated by a colored circle (Figure 4.3b), and the number of pathways connected between each metabolite is indicated by a grayscale square. When metabolites are hovered over by the cursor, the full name, formula, and chemical structure of the metabolite are displayed in an image that pops up (Figure 4.3c). For many transport reactions, the transporter is not yet known and only few have associated genes, or the transport gene is not part of the iCEL1314 metabolic
model. We found that, by having multiple people manually evaluate different metabolic genes and pathways, the iCEL1314 metabolic model can be further improved. For example, we found that the conversion of γ-linolenoyl-CoA (lnlncgcoa) to stearidonyl-CoA (strdnccoa) by fat-1 was missing from the model even though this reaction is described in the literature\textsuperscript{262} (Figure 4.3d). Using a combination of KEGG, WormBase, and literature searches, we added six genes to the maps that are not in the iCEL1314 model: two to new reactions and four to existing reactions. We also added 25 genes from the iCEL1314 model to 16 reactions, 5 existing, and 11 new. Finally, we found three pathway connections not identified in the iCEL1314 model.

**WormPaths advantages**

Metabolic maps provided by KEGG are extremely useful and frequently published in the primary literature (e.g. GAO et al. 2018; CHAN et al. 2019)\textsuperscript{263,264}. However, these maps can be non-intuitive for several reasons. First, these ‘pan-organism’ maps display all the chemistry known for a particular pathway based on enzymes identified by Enzyme Commission number. However, many reactions can be found in some organisms but not others. For instance, many reactions are specific to prokaryotes. By selecting an organism of choice, here *C. elegans*, KEGG colors the boxes representing enzymes in green if the enzyme is predicted to occur in that organism (Figure 4.5a). Second, one has to hover over the enzyme box to visualize the associated gene(s). Third, there can be a lot of overlap between
different pathways, and pathways in WormPaths have been greatly simplified without losing critical information (Figure 4.5b). For example, the C. elegans pantothenate and CoA biosynthesis map in KEGG looks extremely complicated, but many of the boxes in the KEGG map are white, indicating that there is no known gene for the pertaining reactions in C. elegans. Further, the KEGG map contains components of cysteine and methionine metabolism, arginine and proline metabolism, propionate degradation, glycolysis, and other overlapping pathways. The WormPaths map strips away these excess genes and pathways and focuses solely on pantothenate and CoA formation (Figure 4.5b). In this specific example connections to other pathways from the terminal metabolites are not indicated by boxes due to the fact that cys-L, ctp, cmp, and coa all connect to more than four other pathways, making the map cumbersome to navigate. The connecting pathways can be viewed on the WormFlux website by clicking the metabolite of interest.
PANTOTHENATE AND CoA BIOSYNTHESIS

A

B

EXTRACELLULAR SPACE

MITOCHONDRIA

CYTOSOL

Pantothenate

CoA

Pantothenic Acid

Coenzyme A

Pantothenate

CoA

Pantothenic Acid

Coenzyme A

Pantothenate

CoA

Pantothenic Acid

Coenzyme A

Pantothenate

CoA

Pantothenic Acid

Coenzyme A

Pantothenate

CoA

Pantothenic Acid

Coenzyme A
In addition to simplifying metabolic pathway maps, we also extended several WormPaths maps relative to KEGG. For instance, the WormPaths ketone body metabolism map has additional conversions with associated genes, relative to the map available in KEGG (Figure 4.6). More precise connections to other pathways, transport reactions, and subcellular localization of the reactions are visualized in WormPaths.
Figure 4.6 – WormPaths maps provide additional reactions to metabolic pathways. (A) Ketone body metabolism map in KEGG. Green boxes indicate enzymes found in *C. elegans*. (B) Ketone body metabolism map in WormPaths.

In KEGG, genes are associated with any pathway assigned to that gene by gene-protein-reaction associations. However, sometimes these reactions can be isolated because surrounding reactions are not found in the organism of interest, thus the isolated reaction does not connect to the larger pathway or network of said organism. The isolated reactions may be incorrect annotations that are not likely to exist in the organism, or they may have been incorrectly inserted into the pathway based on homology to another organism\(^{42}\). For instance, the aldehyde dehydrogenase *alh*-2 is associated with 15 KEGG pathways (Figure 4.7a). However, in several of these KEGG reactions, *alh*-2 is associated with one or more isolated reactions that are not connected to iCEL1314\(^{244}\) (Figure 4.7b). This can be further visualized in the KEGG pantothenate and CoA biosynthesis map from Figure 4.5a; enzyme EC1.2.1.3 on the lower left is not connected to the rest of the pathway. Further, only five of the 15 KEGG pathways associated with *alh*-2 have the enzyme connected to the rest of the pathway via other *C. elegans* enzymes (glycolysis/gluconeogenesis, glycerolipid metabolism, leucine degradation, isoleucine degradation, and valine degradation). Altogether, WormPaths identifies four pathways for *alh*-2, and all are shared with KEGG (Figure 4.7a). Refining gene-to-pathway annotations in WormPaths is especially important for statistical analyses; when a gene is incorrectly associated with different pathways, this can affect the significance of detected enrichments.
Figure 4.7 – WormPaths maps clean up pathway associations for individual genes. (A) Gene-to-pathway annotations for *alh-2* in KEGG and WormPaths. (B) KEGG annotation for *alh-2* (green box with red text) in ascorbate and aldarate metabolism. White boxes indicate no known enzyme in *C. elegans*.

*WormPaths levels can be used for pathway or gene set enrichment analysis*

To demonstrate how the levels in WormPaths can be used to identify high-resolution metabolic pathway enrichment in transcriptomic data, we analyzed a previously published RNA-seq dataset measuring the transcriptomes of untreated
animals, animals treated with 20 nM vitamin B12, or 20 nM vitamin B12 and 40 mM propionate. We selected all of the genes from the genome scale experiment with a p-adjusted value of ≤ 0.05 and a fold change of ± 1.5. and performed pathway enrichment analysis using the pathway enrichment analysis (PEA) tool in WormFlux (see Methods). This approach confirmed our previous findings that propionate degradation by the shunt pathway and the Met/SAM cycle are enriched in this dataset.

In collaboration with the Walker lab, we previously developed WormCat, an online tool for identifying genome-scale coexpressed gene sets (Figure 4.8). In WormCat, genes are assigned to a single functional annotation, while in WormPaths, metabolic genes can be assigned to multiple reactions and, therefore, pathways. This, together with the inclusion of different Levels of metabolism, allows gene enrichment analysis at greater resolution (Figure 4.8). In contrast to WormCat, however, WormPaths is limited to the genes included in the iCEL1314 model. Given the advanced curation of the genes in WormPaths, using these gene sets provides a complementary level of resolution for the analysis of metabolic pathways, relative to WormCat. Thus, we suggest that researchers first use WormCat for gene set enrichment analysis and that they include WormPaths in their analyses when they find an enrichment for metabolic genes. Finally, our high-resolution metabolic pathway annotations can be integrated as custom gene-sets while performing other kinds of enrichment analysis, for example using...
classical Gene Set Enrichment Analysis\textsuperscript{266} to extract specific desired information from gene expression profiling data.
### Vitamin B12 vs Untreated

<table>
<thead>
<tr>
<th>Category</th>
<th>Genes</th>
<th>P-value</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIPIDS</td>
<td></td>
<td></td>
<td>Level 1</td>
</tr>
<tr>
<td>Fatty Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mitochondrial fatty acid degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short Chain Fatty Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propionate degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propionate degradation shunt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycerophospholipid Metabolism**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONE CARBON CYCLE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Folate Cycle**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met/SAM Cycle**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AMINO ACIDS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branched-Chain Amino Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isoleucine degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>valine degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>serine metabolism*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ketone Body Metabolism**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin B12 Enzyme**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Vitamin B12 + Propionate vs Vitamin B12

<table>
<thead>
<tr>
<th>Category</th>
<th>Genes</th>
<th>P-value</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIPIDS</td>
<td></td>
<td></td>
<td>Level 1</td>
</tr>
<tr>
<td>Fatty Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mitochondrial fatty acid degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short Chain Fatty Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propionate degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>propionate degradation shunt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ONE CARBON CYCLE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met/SAM Cycle**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Branched-Chain Amino Acid Degradation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>isoleucine degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>valine degradation*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHER AMINO ACIDS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ketone Body Metabolism**</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### WormPaths

- **LEVEL 1**: 
  - > 20 genes: 10^-6
  - 15-20 genes: 10^-5
  - 11-15 genes: 10^-4
  - 6-10 genes: 0.01
  - 2-5 genes: 0.05

### WormCat

- **CATEGORY 1**
  - > 50 genes: 10^-10
  - 25-49 genes: 10^-5
  - 10-24 genes: 0.001
  - 5-9 genes: 0.05
  - 2-4 genes: 0.05

* = levels 3 and 4 are redundant
** = levels 2, 3 and 4 are redundant
Figure 4.8 – Pathway enrichment analysis using WormPaths levels. Pathway enrichment analysis using a previously published RNA-seq dataset of \textit{C. elegans} untreated, treated with 20nM vitamin B12, or treated with 20nM vitamin B12 and 40mM propionate with a $p$-adjusted of $\leq 0.05$ and a fold change of $\pm 1.5$ shows enrichment of lipids and one-carbon cycle pathways (left, blue). The arrows indicate the directionality of differentially expressed genes. No arrow indicates both increased and decreased gene expression. WormPaths enrichment for curated metabolic genes complements and adds resolution to the genome scale enrichment metabolic results from WormCat (right, orange).

Discussion and vision

We have developed WormPaths, an expandable online catalog of \textit{C. elegans} metabolic pathway maps and gene annotations. Our overall annotations predict a total of more than 3,000 metabolic genes in \textit{C. elegans}, based on homologies with metabolic enzymes or protein domains\textsuperscript{44}. Therefore, metabolic network models such as iCEL\textsuperscript{1314} continue to grow and evolve as more experimental data becomes available. We encourage \textit{C. elegans} researchers to contact us and help with updates and additions, and to point out any errors they may find. We expect that new metabolic reactions and metabolites will continue to be added to future versions of iCEL as they are discovered. For instance, the iCEL\textsuperscript{1314} model incorporates the relatively recently discovered ascaroside biosynthesis pathway\textsuperscript{244}. In the future, we hope to visualize changes in gene expression, metabolite concentrations, and potentially metabolic rewiring, which can occur under different dietary or environmental conditions. Altogether, WormPaths builds on and provides advantages over KEGG and the visualization strategy used to develop WormPaths should be applicable to other model organisms.
Materials and methods

Design of pathway maps

The design of pathway maps aimed at capturing and visualizing metabolic functions in such a way that would be broadly useful for both statistical analyses and navigation purposes. The starting point for pathway definitions was the pathway annotations of reactions and genes of iCEL1314 in WormFlux and in KEGG. Existing pathways were then split and/or modified such that the functional resolution of pathways was increased without disrupting the coherence of reactions, while the number of overlapping reactions was minimized. For example, the valine, leucine and isoleucine degradation pathway (one map in KEGG) was first divided into three maps to increase pathway resolution: valine degradation, leucine degradation, and isoleucine degradation. Then, a reaction that existed in the original pathway that converts propionyl-CoA to methylmalonyl-CoA (i.e., RM01859 in iCEL1314 and R01859 in KEGG) was removed from valine degradation and isoleucine degradation maps to avoid a redundant overlap with propionate degradation, where this reaction defines a critical step. In KEGG, R01859 is associated with glyoxylate and dicarboxylate metabolism in addition to valine, leucine, and isoleucine degradation and propionate metabolism, thus appearing in three places. However, propionyl-CoA to methylmalonyl-CoA conversion is clearly the first step of canonical propionate degradation.

Typically, pathways were designed to start or end with three types of metabolites: (i) the main substrate or product by definition (e.g. histidine is the
starting point in histidine degradation, and collagen is the endpoint in collagen biosynthesis), (ii) a connection to other pathways (e.g., valine degradation ends with propionyl-CoA through which it is connected to propionate metabolism), and (iii) an endpoint that can be transported to or from extracellular space (e.g., histamine is produced in histidine degradation pathway and exported). The connections of terminal metabolites to other pathways are indicated in maps by clickable pathway boxes as in KEGG, unless the metabolite is associated with more than two other pathways. When a terminal metabolite is not associated with any other pathway, a proper transport that explains the source or fate of the metabolite is included. If a transport is not available either, then it follows that the metabolite is associated with reactions not included in WormPaths maps yet, which is indicated by a box labeled “other”. In any case, the number of pathways and the types of transports (cytosol-extracellular space or mitochondria-cytosol) a metabolite is associated with are indicated by colored squares and circles, respectively, as shown by a legend appended to every map. Furthermore, clicking a metabolite brings the page of that metabolite in WormFlux, which shows all pathways and reactions it is associated with. Thus, information about the pathway associations and transportability of not just terminals, but every metabolite in a pathway, is reachable from the pathway map.

Illustration of pathway maps

Draft maps were drawn as SVG files in the open-source vector graphics editor
Inkscape (http://inkscape.org) following a template. Genes from each map were extracted from the SVG files and cross-referenced to the master levels spreadsheet (Table 4.1). After correction of errors the final SVG maps were wrapped with HTML format and uploaded to the WormFlux website (http://wormflux.umassmed.edu). Maps were blended with WormFlux pages and made interactive using PHP language for server-side processes (e.g., search) and JavaScripting language for the client-side actions (e.g., metabolite image display).

Pathway enrichment analysis
To facilitate PEA, we developed an interactive webtool in the WormFlux website (http://wormflux.umassmed.edu/WormPaths/pea.php). This tool takes a list of genes from the user as the input. First, the overlap between the input and the background (universal) set, i.e., the entire gene list of iCEL1314, is determined. This overlap defines the metabolic hits. For each gene category at each level, the strength and significance of the intersection between category genes and metabolic hits are evaluated. The strength is the enrichment score calculated as $k/n$, where $k$ is the number of genes in the intersection set, and $n$ stands for the total number of genes in that category. The significance was based on the hypergeometric test performed using the hypergeom function in scipy package of Python 2.7. Briefly, a hypergeometric distribution is derived for the size of the intersection set between the category genes and the metabolic hits using $hypergeom(M,N,n)$, where $M$ is the size of the universal set, which is 1314, and
$N$ is the number of metabolic hits. Then, the probability mass function of this distribution is used to calculate the enrichment $p$-value as the sum of probabilities for $k$ and greater integers, which corresponds to the total probability of having $k$ or more genes in the intersection. Similarly, the depletion $p$-value is calculated as the total probability of having $k$ or less genes in the intersection. The PEA tool provides an interactive color-coded table that illustrates enrichment scores and enrichment or depletion $p$-values for every category at every level. The results are also provided in greater detail as a downloadable tab-separated text file. Importantly, a multiple hypothesis testing correction is not done by PEA to avoid the underestimation of enrichment strength. The users are encouraged to apply Bonferroni correction when searching enrichments or depletions at a significance level of $\alpha$, such that $p < \alpha / n_{test}$ must be satisfied, where $n_{test}$ is the number of tests (e.g., the number of categories at a level of interest).

**Metabolite structures**

Out of the 907 metabolites in iCEL1314, 777 are represented in WormPaths maps by abbreviations that are linked to images with metabolite name, formula and structure. The image of a metabolite is displayed when its abbreviation is hovered over. Names and formulas follow from iCEL1314\textsuperscript{244}. Structures were based on mol file representations\textsuperscript{267} or hand drawings. Mol files were readily obtained from KEGG\textsuperscript{245} for 559 metabolites, and from other public resources including Virtual Metabolic Human Database\textsuperscript{268}, PubChem, and ChEBI for 43 more. All mol files
were converted to PNG format using Open Babel\cite{269}. The structures of 160 metabolites were created based on mol files and shapes of similar molecules using a commercial vector-based graphics software when necessary. These drawings were also saved as PNG. No definitive structures were found for the remaining 15 metabolites (mostly proteins), for which a “Structure not available” sign was used instead of a molecular structure. Finally, each structure image was stacked with the corresponding metabolite name and formula using Inkscape to obtain the images of metabolites displayed in WormPaths.
CHAPTER V: CONCLUSIONS

Introduction

Until recently, metabolism was thought to be regulated primarily at the post-translational and allosteric level due in part to the speed at which enzymes need to respond to the changing metabolic demands of the cell. Consequently, most of what we know about metabolism stems from decades of careful biochemical research isolating and functionally characterizing metabolic enzymes and their kinetic properties. While there are some notable examples of transcription factors (TFs) controlling the expression of metabolic genes, it wasn’t until advances in transcriptomic analysis revealed that in fact, metabolic gene expression is dynamic at the genome scale. In fact, metabolic gene expression changes across cell types, developmental stages, nutrient status and disease states. The next challenge is to understand how metabolic pathways are transcriptionally regulated as a systems level.

C. elegans provides an excellent model to study the crosstalk between metabolism and gene regulation in whole animals across multiple tissues\textsuperscript{15,16,219}. It’s genetically tractable and many genes in the metabolic network are conserved with humans\textsuperscript{42}. C. elegans is amenable to large scale genetic screens and because animals are transparent, metabolic gene expression can be monitored by use of proxy reporters simply under the microscope. By examining metabolic reporter expression when regulatory elements are perturbed, we can delineate the
gene regulatory networks (GRNs) that control metabolic gene expression \textit{in vivo}^{15}. Additionally, by employing both transcriptomic and metabolomic techniques we can begin to understand how metabolic pathway activity is regulated at the transcriptional level. Reconstructed metabolic network models informed by transcriptional data and metabolic pathway maps help generate testable hypotheses by predicting metabolic flux when pathway genes are perturbed\textsuperscript{44}. Further, they can be used in metabolic pathway enrichment analysis of transcriptomic data\textsuperscript{241,270}.

In this work we have used these tools to identify a second mechanism by which vitamin B12 controls metabolic gene expression: B12-mechanism-II. Unlike B12-mechanism-I, this second mechanism is propionate-independent and requires the nuclear hormone receptor TF NHR-114. B12-mechanism-II regulates genes involved in the methionine/S-adenosylmethionine (Met/SAM) cycle in response to changes in Met/SAM cycle activity to maintain the metabolic homeostasis of the cycle. To assist in future projects researching the transcriptional regulation of metabolic pathways using \textit{C. elegans}, we developed an online tool titled WormPaths. WormPaths presents carefully annotated metabolic genes and metabolites in easy-to-navigate, hand drawn, visual pathway maps.

\textbf{Multiomics approaches and pathway tools to study transcriptional regulation of metabolic flux}

Comprehensive study of transcriptional regulation of metabolic genes requires
both a broad view of gene expression and metabolism as well as a detailed understanding of metabolic pathways. To this end transcriptomic analysis by RNA-seq and metabolomic analysis by mass spectrometry are invaluable tools to examine changes in global transcript and metabolite levels due to genetic or nutritional perturbation\textsuperscript{97,271,272}. Systematic integration of the two however, is not trivial in part because there is no direct relationship between metabolite and transcript\textsuperscript{273,274}. Nevertheless, some approaches and online tools are available such as INDEED\textsuperscript{275}, OmicsTIDE\textsuperscript{276}, MetaboAnalyst\textsuperscript{277} and GAIT-GM\textsuperscript{278}. One technique, termed pathways-based integration, depends on \textit{a priori} knowledge of metabolic pathways to map both transcriptomic and metabolomic data in order to identify pathway involvement in response to a perturbation\textsuperscript{279}.

In addition to these -omic techniques, careful scrutiny of metabolism at the gene and pathway level is useful to gain more mechanistic understanding of metabolic gene regulation. WormPaths provides a tool to visualize metabolic pathways in order to generate testable hypotheses at the local level. The genes and reactions were curated primarily from the \textit{C. elegans} reconstructed metabolic network iCel1314 and literature. Searchable, hand-drawn maps allow the user to easily find a gene or metabolite of interest and by examining the pathways, predict metabolic flux. WormPaths provides a framework on which to examine both transcriptomic and metabolomic data and thereby gain a more meaningful and contextual comprehension of these analyses. In fact, by uploading RNA-seq data the user can now perform metabolic pathway enrichment analysis, augmenting
available gene set enrichment tools such as WormCat. In the future, we hope to add a feature that allows the upload of metabolomic data for pathway-based integration analysis.

Transcriptional regulation of Met/SAM cycle activity

While B12-mechanism-I serves to activate an alternative vitamin B12-independent propionate pathway when vitamin B12 is scarce, B12-mechanism-II functions to maintain the metabolic activity of the Met/SAM cycle in a tightly controlled regime. Too large a fluctuation in either direction of metabolic activity could disturb several aspects of the Met/SAM cycle: the ratio of phosphatidylethanolamine (PE) to phosphatidylcholine (PC), histone methylation, the transsulfuration pathway, protein synthesis, as well as energy and NADPH production by the folate cycle. Many of these metabolic components operate in critical relationships to each other. For example, the ratio of SAM to S-adenosylhomocysteine (SAH) is integral to sustain transmethylation; the ratio of PE to PC is important for lipid homeostasis and cellular membrane dynamics; the ratio of reduced to oxidized glutathione produced by the transsulfuration pathway functions in autophagy and cell death pathways. Distortion of these ratios might lead to developmental defects, or in the case of mammals, cancer and other diseases. It is therefore perhaps unsurprising that the Met/SAM cycle metabolic activity must be strictly regulated not only at the allosteric level but also transcriptionally.
**Met/SAM cycle activity is sensed and regulated by a nuclear hormone receptor**

In B12-mechanism-I, the TFs NHR-10 and NHR-68 function in a feed forward loop with an AND-logic gate that operates as a propionate persistence detector\textsuperscript{24}. Metabolic activity of the Met/SAM cycle is monitored through the sensing of low SAM by NHR-114, which consequently regulates the transcription of genes involved in the Met/SAM cycle. How SAM is sensed mechanistically by NHR-114 remains to be resolved. One possibility is that SAM acts as a ligand and binds NHR-114 resulting in sequestration or deactivation of the transcription factor (Figure 5.1). NHRs are ligand-gated TFs, making them ideal candidates to respond to changes in small molecule availability due to altered metabolism or stress levels. Several NHR-ligand pairs have been found in humans, only two, however, have been identified in *C. elegans*: DAF-12 and dafachronic acid\textsuperscript{38}, and NHR-33/HIZR-1 and zinc\textsuperscript{25}. Categories of known NHR ligands in mammals include vitamins, fatty acids, terpenoids, and interestingly phospholipids\textsuperscript{285}.  

Figure 5.1 – Proposed mechanisms by which SAM could be sensed by NHR-114. Sam might directly bind NHR-114 and prevent its activation or translocation to the nucleus. SAM might bind another transcription factor (TF) that regulates nhr-114. Alternatively, SAM might alter the methylation of histones controlling the expression of nhr-114 at the chromatin level. SAM, S-adenosylmethionine; TF, transcription factor.

Although the results point to a model in which SAM is sensed in B12-mechanism-II, the biology is undoubtedly more complex and therefore it wouldn’t be unreasonable to hypothesize that a phospholipid such as PE, PC or another related species binds NHR-114. Although there are several published methods to identify protein-small molecule interactions, one appealing approach to screen for putative receptor ligands is quantitative microdialysis\(^{286}\). This technique involves placing purified protein on one side of a semi-permeable membrane and either a single or a mixture of small molecules on the other and then measuring the
composition of the protein partition via liquid chromatography-mass spectrometry. An increase in the small molecule indicates binding to the protein.

Another possibility is that the transcriptional expression of *nhr-114* is controlled either directly or indirectly through the methylation of histones by SAM (Figure 5.1). For example, an increase in SAM might increase histone methylation and cause a change in binding dynamics either on the promoter of *nhr-114* or of an upstream transcription factor required for *nhr-114* regulation. Although the regulation of histone methylation is commonly thought to be controlled by the expression of methyltransferases, *in vivo* SAM concentrations are within a physiological range that could be expected to affect methylation reaction stoichiometry. Indeed, several studies have shown that methionine restriction leads to a decrease in SAM and alters histone methylation, which leads to a change in histone binding dynamics and thus affects gene expression programs.

Unlike *nhr-114* which is only involved in B12-mechanism-II, the *C. elegans* ortholog of human *SREBP1*, *sbp-1* is important for both B12 mechanisms. *sbp-1* operates at a high level is the intestinal GRN, which means it regulates many downstream TFs under different contexts. Most likely, *sbp-1* responds to different metabolic cues to activate each B12 mechanism. An interesting question is how metabolites modulate the function of SBP-1 to affect the expression of different target genes. There are several ways in which the GRN might be tuned based on changing metabolite levels. For instance, either propionate of SAM might bind
directly or otherwise induce post-translational modifications to SBP-1 or another upstream regulator and change its binding affinity for particular promoters (Figure 5.2).

**Figure 5.2 – A model of metabolites tuning the transcriptional response of SBP-1 to different metabolites.** A metabolite such as propionate (PA) or S-adenosylmethionine (SAM) might directly or indirectly alter the specificity of a transcription factor such as SBP-1 for certain target sequences. Consequently, the transcription factor will regulate the expression of a set of targets specific to that metabolite cue resulting in a specialized response.

**Why is acdh-1 transcriptionally activated by B12-mechanism-II?**

We used the promoter of acdh-1 driving GFP expression as a tool to investigate both B12 mechanisms. In B12-mechanism-I, acdh-1 is transcriptionally activated in response to accumulated propionate in order to degrade it when vitamin B12 is low\(^\text{27}\). The role of acdh-1 in B12-mechanism-II, however remains unclear. There are two possible explanations. One, ACDH-1 may have a second function in addition to catalyzing the conversion of propionyl-CoA to acrylyl-CoA in the propionate shunt (Figure 5.3).
Multifunctional or promiscuous enzymes are enzymes that may have more than one function or may catalyze the same reaction in different pathways. There are many known multifunctional enzymes in both human and C. elegans genomes. For example, FASN-1 catalyzes three different reactions in fatty acid biosynthesis in C. elegans and other animals. Interestingly, MTHFD1, the human ortholog of C. elegans gene K07E3.4 (predicted), channels formate to 10-formylTHF to 5,10-methyleneTHF in the folate cycle. In fact, there’s evidence to suggest that promiscuity among enzymes may be more common than previous expected.

One way a multifunctional enzyme can switch between activities is if its function is dependent on cellular location. For example, the E. coli enzyme PutA acts as a proline dehydrogenase when associated with the plasma membrane but loses its catalytic activity and binds DNA as a repressor when free in the cytoplasm. Propionate is degraded in the mitochondria while the Met/SAM cycle operates in the cytosol. ACDH-1 might function in propionate degradation in the mitochondria but have a second function involved in Met/SAM cycle activity in the cytosol.

When Met/SAM cycle activity is perturbed, genes involved in influx of metabolites such as methionine are upregulated, while those involved in efflux are downregulated as a way to increase metabolic activity of the Met/SAM cycle. An intriguing prospect is that ACDH-1 might participate in a reaction that produces a metabolite that feeds into the Met/SAM cycle. If true, then knock down of acdh-1 in a Met/SAM cycle mutant background should cause an aberrant phenotype such as severe developmental delay or lethality. Future experiments looking for such
synthetic interactions between *acdh-1* and Met/SAM cycles genes may help test this hypothesis.

**Figure 5.3** – Two possibilities that might explain up-regulation of *acdh-1* by Met/SAM cycle perturbation. 1 – ACDH-1 might have a secondary function in a pathway that feeds into the Met/SAM cycle. 2 – Propionate might accumulate due to Met/SAM cycle perturbation and B12-mechanism-II acts to bypass the persistence detector. SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine; THF, tetrahydrofolate; MM-CoA, methylmalonic acid coenzyme A; 3-HP, 3-hydroxypropionate. BCKDH, branched chain α-ketodehydrogenase complex.
Two, perturbation of Met/SAM activity leads to production of propionate, which needs to be degraded by ACDH-1 (Figure 5.3). In fact, although not a major source, methionine can eventually be degraded to propionate as measured by isotope tracing experiments in patients with methylmalonic acidemia\textsuperscript{294,295}. Cystathionine is produced when homocysteine enters the transsulfuration pathway instead of being recycled to methionine. As part of the next step in the transsulfuration pathway, cystathionine is converted to cysteine and \(\alpha\)-ketobutyrate. While cysteine proceeds to the transsulfuration pathway, \(\alpha\)-ketobutyrate can be transported to the mitochondria where it is converted to the CoA analog of propionate, propionyl-CoA (Figure 5.3). Therefore, potentially a decrease in Met/SAM cycle activity increases production of \(\alpha\)-ketobutyrate through cystathionine production and ultimately increases propionyl-CoA concentration. In this scenario, transcriptional activation of \textit{acdh-1} by SAM would act as a regulatory cue, thus bypassing B12-mechanism-I in \textit{acdh-1} upregulation.

**Conclusion**

When a metabolic gene is transcriptionally upregulated, more copies of the enzyme product are made and thus it is generally accepted that an increase in metabolic gene expression leads to an increase in metabolic activity of the associated reaction. In fact, many system biology and biochemistry approaches depend on this assumption. We found, however, that Met/SAM cycle genes are
downregulated when metabolic activity is increased and upregulated when metabolic activity decreases. Is this the only metabolic pathway to be regulated in this seemingly paradoxical fashion? If not, this might help explain some instances when transcriptomic and fluxomic data do not coincide and simultaneously underscores the complexity that must be captured\textsuperscript{296}. It should be emphasized that the discovery of B12-mechanism-II and its function in Met/SAM cycle gene regulation would not be possible without both a systems level view and a deep interrogation of genetic mechanisms.
BIBLIOGRAPHY


33. Maglich, J. M. *et al.* Comparison of complete nuclear receptor sets from the


Quadros, E. V., Nakayama, Y. & Sequeira, J. M. The protein and the gene


100. Salonen, A. *et al.* Impact of diet and individual variation on intestinal


102. Mancia, F. & Evans, P. R. Conformational changes on substrate binding to methylmalonyl CoA mutase and new insights into the free radical mechanism. Structure 6, 711–720 (1998).


115. Lu, S. C. & Mato, J. M. S-adenosylmethionine in liver health, injury, and


120. Han, L. et al. Dietary methionine restriction improves the impairment of cardiac function in middle-aged obese mice. Food Funct. 11, 1764–1778 (2020).


163. Hershko, A., Heller, H., Eytan, E. & Reiss, Y. The protein substrate binding


177. Liu, G. Y. & Sabatini, D. M. mTOR at the nexus of nutrition, growth, ageing


224. Vozdek, R., Hnizda, A., Krijt, J., Kostrouchová, M. & Kožich, V. Novel structural arrangement of nematode cystathionine β-synthases:


243. Opdam, S. *et al.* A systematic evaluation of methods for tailoring genome-


Hum, D. W., Bell, A. W., Rozen, R. & MacKenzie, R. E. Primary structure of a human trifunctional enzyme. Isolation of a cDNA encoding methylenetetrahydrofolate dehydrogenase-methenyltetrahydrofolate


