Western Graduate&PostdoctoralStudies

Western University
Scholarship@Western

Electronic Thesis and Dissertation Repository

12-12-2017 2:30 PM

Role of Asparagine as a Nitrogen Signal and Characterization of a Nitrogen Responsive Glutamine Amidotransferase, GAT1_2.1 in Arabidopsis thaliana

Shrikaar Kambhampati, *The University of Western Ontario*

Supervisor: Dr. Frédéric Marsolais, *The University of Western Ontario*
Co-Supervisor: Dr. Jim Karagiannis, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Biology
© Shrikaar Kambhampati 2017

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Biochemistry Commons

Recommended Citation

Kambhampati, Shrikaar, "Role of Asparagine as a Nitrogen Signal and Characterization of a Nitrogen Responsive Glutamine Amidotransferase, GAT1_2.1 in Arabidopsis thaliana" (2017). *Electronic Thesis and Dissertation Repository*. 5092.
https://ir.lib.uwo.ca/etd/5092

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

ABSTRACT

Maintaining the proper balance between carbon (C) and nitrogen (N) metabolism is critical to the sustained growth of organisms. In plant leaves, this balance is achieved by photoperiod dependent cross-talk between the processes of photosynthesis, respiration, and amino acid metabolism. A crucial mechanism in maintaining C/N balance is the glutamine synthetase/glutamine oxoglutarate aminotransferase (GS/GOGAT) cycle, which is well known to serve as a cross-road between C and N metabolism. Importantly, non-photosynthetic tissues (e.g. roots, germinating seeds) lack a sufficient supply of carbon skeletons under high N conditions and hence may resort to other mechanisms, along with the GS/GOGAT cycle, to achieve proper C/N balance. Our understanding of the pathways involved in this aspect of plant regulation is limited. Considering the importance of asparagine as a major storage form of N, this study examines C and N partitioning within Arabidopsis roots upon asparagine treatment. Based on this work, I propose a role for the enzyme GAT1_2.1 in hydrolyzing excess glutamine to glutamic acid (Glu), which may serve as a carbon skeleton for channeling C to the TCA cycle under high N conditions. *GAT1_2.1*, a gene coding for a class I glutamine amidotransferase of unknown substrate specificity, was shown to be highly responsive to N status and has a root specific expression in *Arabidopsis*. The protein localizes to the mitochondria and the gene is found to be highly co-expressed with *Glutamate Dehydrogenase 2* (*GDH2*). Metabolite profiling data using a *gat1_2.1* mutant of *Arabidopsis* suggests that, in the absence of GAT1_2.1, the GABA shunt pathway is activated to replenish the depleted levels of Glu. This Glu may then be deaminated to 2-oxoglutarate by GDH2 and channeled into the TCA cycle, thus providing a cross-roads between C and N metabolism in root mitochondria. In addition to this work, I also elucidate optimal methods for reliable metabolomics experiments and propose the use of isotopic labelling for the detection of unknown pathways.

KEYWORDS

Asparagine, Glutamine, Metabolism, C/N partitioning, Metabolomics, Transcriptomics, GAT1_2.1, Glutamine amidotransferase, Isotope labelling, *Arabidopsis thaliana*.

CO-AUTHORSHIP STATEMENT

The following thesis contains material from manuscripts either published or in preparation which are co-authored by Shrikaar Kambhampati (SK), Ebenezer Ajewole (EA), Sudhakar Pandurangan (SP), Justin B. Renaud (JR), Ryan S. Austin (RA), Mark W. Sumarah (MS) and Frédéric Marsolais (FM).

My supervisor Frédéric Marsolais provided insight and strategic direction for the projects and edited the manuscripts (Chapters 1-5).

Chapter 1. Author's contributions

SK drafted the manuscript. EA and FM edited the manuscript. A part of this chapter is published in Advances in asparagine metabolism; Progress in Botany (2017)

Chapter 2. Author's contributions

SK designed the research, performed the experiments, analysed the data and drafted the manuscript. SP performed the RNAseq experiments. RS analysed RNAseq data. JR helped in running the metabolite samples. MS participated in the design of study. FM conceived the study, participated in its design and edited the manuscript.

Chapter 3. Author's contributions

SK designed the research, performed the experiments, analysed the data and drafted the manuscript. JR helped in running the metabolite samples and assisted in data analysis. MS participated in the design. FM conceived the study, participated in its design and edited the manuscript.

Chapter 4. Author's contributions

SK designed the research, performed the experiments, analysed the data and drafted the manuscript. JR helped in running the metabolite samples and assisted in data analysis. MS

participated in the design. FM conceived the study, participated in its design and edited the manuscript.

To my parents

For planting in me the seed to succeed

My brother, Ohmkaar

For his support and encouragement throughout my PhD

ACKNOWLEDGMENTS

I vividly remember a conversation with my brother, Ohmkaar Kambhampati, one pleasant Sunday morning. It was during a break from what seemed like endless lab work, as I felt lost in my thoughts of what I want to do with my life. “Why are you a software engineer? Why do you do what you do? And do you enjoy doing what you do?” I asked. With a grin, as if he read my mind like a book and knew exactly what I needed to hear, he said, “Humans relinquished the life of hunters and gatherers, living in caves, suffering from disease and death because of a few people who dreamt of a much needed ‘upgrade’ to our way of life and made it a reality. We are where we are today, due to the tireless efforts of scientists, engineers, policy makers etc., who made their dreams of improving human condition come to life. There is so much more that can be done and I want to be a part of taking it further.” This is of course a short paraphrased version of a half-hour, unimaginably enthusiastic, spiel filled with words of encouragement and wisdom. At the end of our conversation, I remember thinking, throughout my undergrad, masters and thus far in my PhD, the only thing I ever wanted to do was to be a part in achieving food security for the future generations. Although tiny as it may seem that my part is, I am doing it. With a self-reassuring confidence, I went back into the lab and continued making solutions with pleasure and ejecting pipette tips with a great deal of satisfaction. Thank you Ohmkaar, for being older than me, you are the perfect role model.

Elevator pitches, three minute theses, radio appearance and every other possible way of communicating your research to non-scientific audience are skills and opportunities that grad school has bestowed upon me. For some puzzling reason, I couldn’t explain for the life of me, the kind of work that I do to my parents. There were times when my dad sat down with a pen and paper, putting in a huge effort to try and understand what I do, and yet I failed. Despite this, my parents have always believed in me- better yet, took great pride in any little accomplishment during my grad school. I can’t even begin to comprehend the pain they have experienced while sending me away to a place ~13,000 km away from home and not seeing me all through my PhD. I couldn’t have done it without you. Thank you for your constant support and faith in me, I am forever indebted to you. I promise you

that one day I will find a way to explain what I do to understand plants, the truly mystical creatures on earth, to you. My sincere gratitude also to my cousin Ravi Kiran and his family for taking care of me during my grad school and treating me as your own brother.

Grad school can be a blessing or a curse, and the way it goes is almost entirely dependent on your supervisor. Grad school is a common place for complaining and it is not unusual to see or hear the horrors that grad students go through. Oh, have I been so lucky to have Dr. Frédéric Marsolais as my supervisor. Always calm and collected, even when you are upset with my progress, you make me want to work more, achieve more and get things done. Your door has always been open for me and every meeting we had, some weeks almost every day, has been such a delight. I thoroughly enjoyed all the support you have provided, even the soft verbal nudges that got me accelerating the effort in lab. Thank you sincerely, and I hope to find a way to continue working with you in the future. My co-supervisor Dr. Jim Karagiannis, advisors Dr. Mark Bernards and Dr. Ryan Austin have been a great scientific support through this journey. Thank you to all of you.

Life in the lab has always been a great pleasure. Everyday though my grad school, I woke up with a big smile and an urge to go to work as quickly as possible. The company that I had, the inappropriate conversations we had, the times we complained about other people, the times we fought with each other, helped each other, picked up each other's slack, I enjoyed every bit of it. Thank you to Aga Pajak, Sudhakar Pandurangan, Jaya Joshi, Ebenezer Ajewole, and Marwan Diapari. Special thanks to Dr. Justin Renaud, Dr. Mark Sumarah and Tim McDowell. Surprisingly, I took pleasure in your harassment of me. Every conversation with you is a motivation, you always intended the best for me and taught me to become a better scientist in the way I think and design experiments. Justin, I have to say "YOU DA MAN!" I would also like to extend my gratitude to Dr. Christopher Garnham for his encouragement and scientific support.

Social life at AgCan has always been that comedy that starts with "an Indian, an Iranian, a Chinese and a Caucasian went to…" Good times throughout grad school aren't possible without the diverse group of friends that I have, Thank you to Arun, Reza, Mandana, Preetam, Emily, Kelsey, Zayn, Scott, Coby, Sanjay, Craig, Alex, Johnny and others.

TABLE OF CONTENTS

ABSTRACT i

CO-AUTHORSHIP STATEMENT iii

ACKNOWLEDGMENTS vi

TABLE OF CONTENTS viii

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF APPENDICES xvi

LIST OF ABBREVIATIONS xvii

1 **GENERAL INTRODUCTION** 1

1.1 Nitrogen metabolism in plants 1

1.2 Nitrate reduction, signaling and amino acid biosynthesis 2

1.3 Metabolism and regulation of asparagine, an amide amino acid, in plants 4

1.3.1 Asn biosynthesis and degradation 4

1.3.2 Asn as a transport and storage form of N 6

1.3.3 Role of Asn in N metabolism and C/N relations 7

1.3.4 Asn as an N signal in plants 8

1.4 Glutamine amidotransferases in plants 9

1.5 Metabolomics: A tool for advancing our understanding of 'the central dogma' of biology 13

1.6 Objectives 14

1.7 Literature cited 15

2 **COMBINED TRANSCRIPTOMIC AND METABOLITE DATA FROM ASN TREATED ARABIDOPSIS ROOTS PROVIDES NEW INSIGHTS INTO N METABOLISM** 21

2.1 Introduction 21

2.2 Results and discussion 23

2.2.1 Differentially expressed list includes carbon/nitrogen transport and storage related genes 23

2.2.2 Over-represented GO categories indicate C and N reprogramming 26

2.2.3 A significant accumulation of N-rich amino acids indicates N storage ... 30

2.2.4 Lysine and proline catabolism along with the GABA shunt recuperate C/N balance 30

2.2.5 Flux towards aromatic amino acid biosynthesis 33

2.2.6 A class I glutamine amidotransferase (GAT1_2.1) may act as a potential source for C/N partitioning 34

2.3 Concluding remarks and future prospects: 35

2.4 Materials and Methods: 36

2.4.1 Plant material and growth conditions 36

2.4.2 RNA extraction and transcript profiling 36

2.4.3 Total metabolite extraction 37

2.4.4 Data acquisition and analysis 37

2.5 Literature cited 39

3 **THE ROLE OF A CLASS I GLUTAMINE AMIDOTRANSFERASE IN C/N PARTITIONING IN ROOTS OF *ARABIDOPSIS THALIANA*** 42

3.1 Introduction 42

3.2 Results 44

3.2.1 Clustering of class I glutamine amidotransferase domain containing proteins 44

3.2.2 Glutaminase activity of the GAT1_2.1 domain 46

3.2.3 Rescue of the growth phenotype of *E. coli* glutaminase mutant with GAT1_2.1 domain 46

3.2.4 Sub-cellular localization of GAT1_2.1 50

3.2.5 Metabolic profiling of the *gat1_2.1* mutant 50

3.2.6 GAT1_2.1 may not have an acceptor for NH_3 upon glutamine hydrolysis 53

3.3 Discussion 55

3.3.1 *GAT1_2.1* is unique to plants and has a root specific expression 55

3.3.2 The N-terminal GAT1_2.1 domain functions as a glutaminase 55

3.3.3 *gat1_2.1* is non-responsive to Gln and has a shoot branching phenotype 57

3.3.4 *gat1_2.1* uses the GABA shunt pathway to provide carbon skeletons to the TCA cycle 59

3.4 Concluding remarks 61

3.5 Materials and Methods 62

3.5.1 Phylogenetic analysis of class I glutamine amidotransferases 62

3.5.2 Cloning of GAT1_2.1 for sub-cellular localization study 62

3.5.3 Transient expression confocal microscopy 63

3.5.4 Cloning, expression and purification of recombinant GAT1_2.1 domain in *E. coli* 64

3.5.5 Enzymatic assay 65

3.5.6 Rescue of *E. coli* glutaminase mutant *YneH* 66

3.5.7 Plant growth conditions 66

3.5.8 Plant material, genotyping and RT-PCR 67

3.5.9 Total metabolite extraction 69

3.5.10 GAT1_2.1 assay with plant extracts 69

3.5.11 Data acquisition and metabolite analysis 69

3.5.12 Metabolomic data analysis 70

3.5.13 Ammonium quantification 71

3.5.14 Cloning, expression and purification of full length recombinant GAT1_2.1 in plant system 71

3.6 Literature cited 73

4 **EVALUATION OF EXTRACTION AND CHROMATOGRAPHIC METHODS SUITABLE FOR METABOLOMICS IN ARABIDOPSIS THALIANA** 76

4.1 Introduction 76

4.2 Results 78

4.2.1 Comparison of metabolite profiles obtained from different extraction methods 78

4.2.2 Chromatographic resolution of amino acids is achieved with HILIC based chromatography 83

4.2.3 Integration of steps to obtain optimized metabolomics data 83

4.2.4 HRMS is essential for detection of dual labels for LC-MS based metabolomics 87

4.3 Discussion 89

4.4 Materials and Methods 91

4.4.1 Plant growth conditions 91

4.4.2 Extraction using acidified methanol 91

4.4.3 Extraction by phase separation 92

4.4.4 Extraction with chloroform and methanol 92

4.4.5 A simple methanol:water extraction 92

4.4.6 Acidified methanol: water extraction 93

4.4.7 Data acquisition by targeted and untargeted metabolite analysis 93

4.4.8 Optimized chromatography using HILIC 94

4.5 Literature cited 95

5 **GENERAL DISCUSSION** 97

5.1 Excess Asn accumulation results in C/N reprogramming 97

5.2 GAT1_2.1 is a novel plant glutaminase that may regulate C/N partitioning 98

5.3 A new and improved pipeline for LC-MS based metabolomics with a focus on primary metabolism 100

5.4 Use of isotopic label incorporation followed by HRMS for novel pathway identification 101

5.5 Future prospects in understanding N metabolism and improving NUE in plants 101

5.6 Literature cited 103

APPENDICES 105

Curriculum Vitae 130

LIST OF TABLES

Table 1.1: Glutamine amidotransferases (GATs) for which crystal structures have been determined. 11

Table 2.1: Selected list of genes differentially expressed upon Asn treatment based on DAVID functional categorization..... 24

Table 2.2: Over-represented GO categories based on biological process with > 4 fold enrichment. 27

Table 2.3: Amino acids and TCA cycle intermediates identified using LC-MS. 31

LIST OF FIGURES

Figure 1.1: Asparagine biosynthesis and degradation pathway representing all known and unknown enzymes involved in its metabolism. 3

Figure 1.2: Glutamine amidotransferase transfers an amide group from glutamine to a substrate forming aminated product and releasing glutamate. 10

Figure 2.1: DAVID functional categorization. 29

Figure 2.2: Schematic representation of the metabolite changes upon Asn treatment in the context of a pathway. 32

Figure 3.1: GATase1 gene family in *Arabidopsis*. 45

Figure 3.2: Recombinant protein purification and enzyme assay. 47

Figure 3.3: Enzyme kinetics. 48

Figure 3.4: Rescue of *E. coli* glutaminase mutant. 49

Figure 3.5: Subcellular localization of GAT1_2.1. 51

Figure 3.6: Metabolite analysis using wild-type *Arabidopsis* and *gat1_2.1*. 52

Figure 3.7: Recombinant protein production in *N. benthamiana* and enzyme spike assay. 54

Figure 3.8: WT and *gat1_2.1* phenotypic analysis. 58

Figure 3.9: Schematic representation of the proposed model for C/N partitioning. 60

Figure 3.10: *gat1_2.1* T-DNA insertion line, RT-PCR and genotyping. 68

Figure 4.1: Strip charts representing total metabolomic data from different extraction methods. 79

Figure 4.2: Scatter plots with m/z of features and intensities from positive mode. 80

Figure 4.3: Scatter plots with m/z of features and intensities from negative mode..... 81

Figure 4.4: Targeted metabolite analysis showing a few amino acids and organic acids. .. 82

Figure 4.5: Gradient and chromatography using HILIC..... 85

Figure 4.6: Summed intensity plot highlighting isotopically labelled compounds from $^{13}C_5$, $^{15}N_2$ Gln treated *Arabidopsis* roots. 86

Figure 4.7: Comparison of resolving power between different instruments. 88

LIST OF APPENDICES

Appendix A: Differentially expressed gene list FDR < 0.001 105

Appendix B: Differentially expressed gene list FDR < 0.01 113

Appendix C: Diagnostic fit of variance function 126

Appendix D: M-A plot of fold change vs base means for each gene 127

Appendix E: Copyright permission to include figure previously published 128

Appendix F: Copyright permission to include text excerpts previously published 129

LIST OF ABBREVIATIONS

2OG	2-Oxoglutarate
AAT	Aspartate aminotransferase
AS	Asparagine synthetase
Asn	Asparagine
Asp	Aspartate
ASPG	Asparaginase
ATP	Adenosine triphosphate
BSAS	β-substituted alanine synthase
C	Carbon
CFP	Cyan Fluorescent Protein
COX1	Cytochrome oxidase 1
DTT	Dithiothreitol
E4-P	Erythrose 4-phosphate
FDR	False discovery rate
GABA	γ-aminobutyric acid
GAT	Glutamine amidotransferase
GDH	Glutamate dehydrogenase
Gln	Glutamine
Glu	Glutamate

GOGAT	Glutamate synthase/Glutamine oxoglutarate aminotrasnferase
GS	Glutamine synthetase
HESI	Heated electrospray ionization
HILIC	Hydrophilic interaction chromatography
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Km	Michaelis-Menten constant
LB	Luria-Bertani medium
LC-DDA	Liquid chromatography-data dependent acquisition
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
MTBE	Methyl tert-butyl ether
N	Nitrogen
NADH	Nicotinamide adenine dinucleotide
NiR	Nitrite reductase
NR	Nitrate reductase
NUE	Nitrogen Use Efficiency
NupE	Nitrogen Uptake Efficiency
NUtE	Nitrogen Utilization Efficiency
PCR	Polymerase Chain Reaction

PEP	Phosphoenolpyruvate
ROS	Reactive Oxygen Species
RRHD	Rapid resolution high definition
SDS-PAGE	Sodium dodecyl sulphate – Polyacrylamide gel electrophoresis
TAIR	The Arabidopsis Information Resource
TCA	Tricarboxylic acid (cycle)
TOR	Target of rapamycin
Vmax	Maximum reaction rate
WT	Wild-type
YFP	Yellow Fluorescent Protein

1 GENERAL INTRODUCTION

1.1 Nitrogen metabolism in plants

Nitrogen (N) is one of the fundamental inputs for increased crop productivity. In general, plants obtain N from the soil in the form of nitrate (NO_3) or ammonium (NH_4^+). However, some N fixing plants that belong to a family referred to as legumes have the ability to use atmospheric N_2 by symbiotic association with N fixing bacteria. Conventional farming involves supplying N to crop plants in the form of synthetic fertilizers that contain NO_3 or NH_4^+. The global annual amount of synthetic N fertilizer applied to crops has seen a dramatic increase during the last few decades (FAOSTAT). Although crop productivity has been shown to increase significantly with increased N fertilization, contributing to the decrease in world hunger (Good et al., 2004), only 30 - 50% of supplied N fertilizer is utilized by the plants owing to inefficiency in N uptake and metabolism (McAllister et al., 2012). For example, maize yields have been shown to correlate with higher Nitrogen Use Efficiency (NUE) and Nitrogen Uptake Efficiency (NupE) as opposed to increased N fertilization itself (Haegele et al., 2013). To date, several studies have established that improving NUE in plants is of the utmost importance in achieving the combined goals of higher crop productivity and sustainable agricultural development; goals crucial for feeding the growing world population (Hawkesford, 2001; McAllister et al., 2012; Rothstein, 2007). A comprehensive understanding of the biochemical, physiological and molecular mechanisms involved in N uptake, assimilation, transport and storage is thus necessary to improve NUE in plants.

The basic processes involved in N metabolism include 1) N uptake by plant roots in an inorganic form as NO_3, 2) Conversion of NO_3 to NH_4^+ by nitrate and nitrite reductases, and 3) NH_4^+ assimilation to organic forms like Gln and Glu via the glutamine synthetase/glutamine:oxoglutarate aminotransferase (GS/GOGAT) cycle. These processes are very well elucidated at both the biochemical and molecular levels (Hodges, 2002; Lea and Ireland, 1999; Wang et al., 2003).

Despite advances in this field, most of the downstream N regulatory processes remain largely unknown due to the complexity of C/N relations in flowering plants. The current chapter highlights the importance of source-sink relations in determining the fate of organic N and describes the pathways that operate as a function of both N form and level. In addition, aspects of N metabolic pathways relevant to this work are introduced. Finally, a detailed description of the research objectives is put forward.

1.2 Nitrate reduction, signaling and amino acid biosynthesis

As mentioned above, the first step of N metabolism involves the uptake of NO_3 followed by its reduction into NO_2 and NH_4^+. This reduced N is used for the synthesis of a suite of nitrogenous compounds; the largest requirement being for the synthesis of the amino acids that form the building blocks of proteins and which act as precursors to a variety of compounds (e.g. nucleic acids, cofactors, phyto-hormones, and chlorophyll). Besides being a nutrient, NO_3 also acts as a signal molecule and regulates its own uptake and assimilation. With the application of sophisticated microarray technologies, and using *Arabidopsis thaliana* as a model, a number of N responsive genes that are likely involved in the key processes of N uptake and utilization have been identified (Gutiérrez et al., 2007; Wang et al., 2003; Wang et al., 2004). To coordinate the production of carbon (C) skeletons required for amino acid biosynthesis, NO_3 also regulates the expression of genes involved in C metabolism. Upon higher NO_3 availability, plants redirect C from starch biosynthesis to the production of amino and organic acids by modulating the expression of key enzymes involved in these processes (Wang et al., 2003).

The initial assimilation products of reduced N are glutamate (Glu) and glutamine (Gln), which constitute the first amino acids produced by the GS/GOGAT cycle (Lea and Ireland, 1999) (Figure 1.1). Besides the synthesis of the nitrogenous compounds mentioned above, N assimilated into Glu and Gln is incorporated into aspartate (Asp) and asparagine (Asn), via the function of aspartate aminotransferase (AAT) and asparagine synthetase (AS), respectively (Lea et al., 2007) (Figure 1.1). Asp is a metabolically reactive amino acid and serves as an N donor in numerous aminotransferase reactions and is the precursor of a large family of amino acids. Asn, however, is relatively inert and serves primarily as an N storage and transport compound. While the GS/GOGAT cycle,

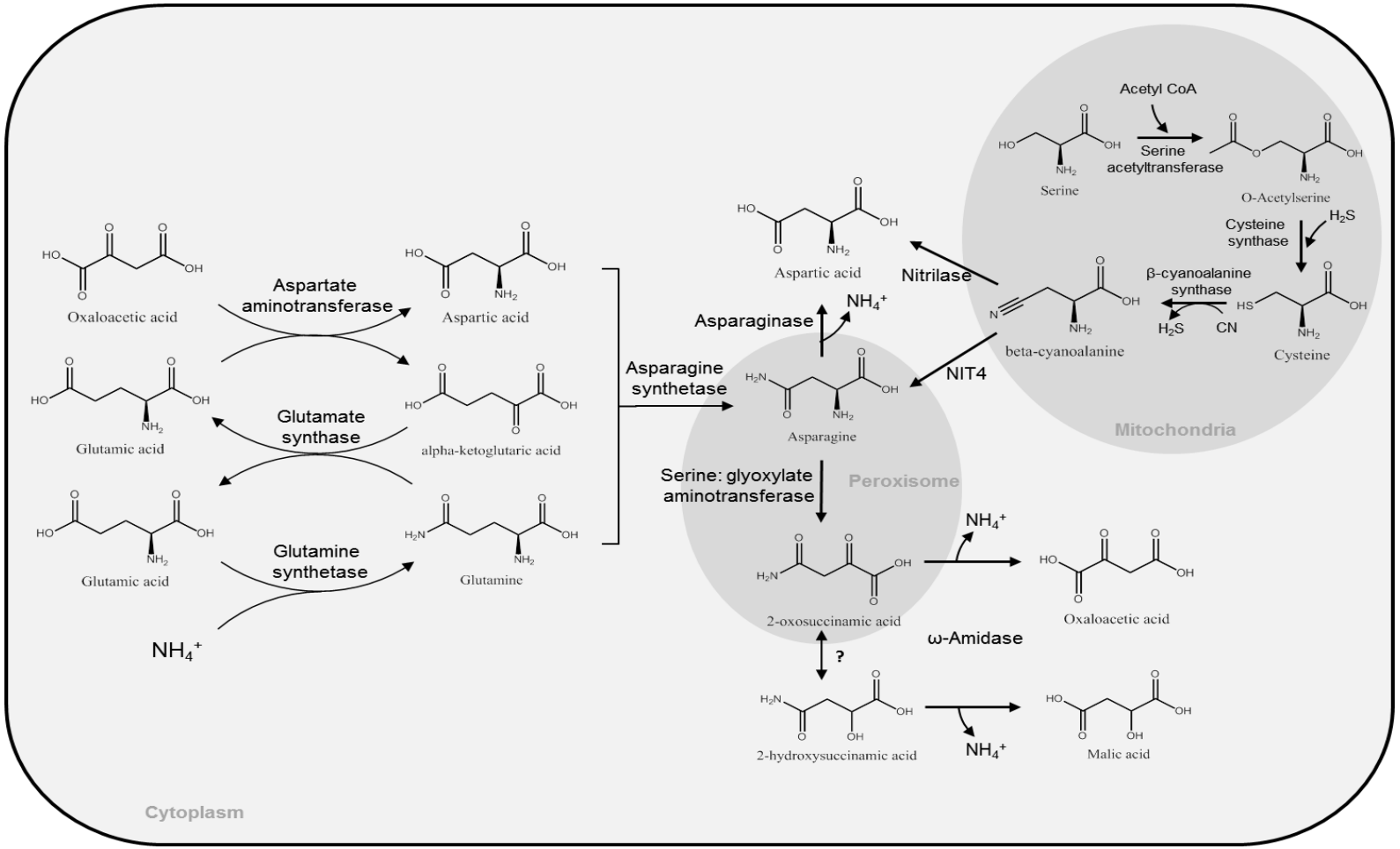


Figure 1.1: Asparagine biosynthesis and degradation pathway representing all known and unknown enzymes involved in its metabolism.

together with the metabolic pathways downstream of Glu and Gln have been clearly described in the literature (Coruzzi, 2003; Noctor et al., 2004), many aspects of Asn metabolism and signalling have yet to be elucidated. To that end, the next section of this chapter provides a detailed review of 1) the key metabolic pathways involved in Asn biosynthesis and catabolism, 2) the importance of Asn in C/N inter-relationships, and 3) the potential role of Asn as a signalling compound.

1.3 Metabolism and regulation of asparagine, an amide amino acid, in plants

1.3.1 Asparagine biosynthesis and degradation

A major route for Asn biosynthesis is via the ATP-dependent transfer of the amide group of glutamine to the β-carboxyl group of aspartate by the action of AS (Figure 1.1). So far, two types of ASs, AsnA and AsnB, have been identified. While prokaryotes utilize AsnA type ASs, which require ammonia as an amide donor, as well as AsnB type ASs, which can catalyze the reaction using either ammonium or glutamine as an amide donor, most eukaryotes use only AsnB type ASs (Duff, 2015; Gaufichon et al., 2010).

AsnB type ASs are members of the N-terminal nucleophile hydrolase (Ntn) group of glutamine amidotransferases (Larsen et al., 1999; Massiere and Badet-Denisot, 1998). They are characterized by an N-terminal cysteine nucleophilic residue producing a cysteinyl-glutamine tetrahedral intermediate from which ammonium is abstracted. Glutamate is released by hydrolysis of the resulting γ-glutamyl thioester intermediate. The ammonia is tunneled to a C-terminal transferase domain. This domain activates Asp through ATP hydrolysis as a β-aspartyl AMP intermediate. Nucleophilic attack by the ammonia results in cleavage and release of Asn.

Two other pathways have been reported for Asn biosynthesis, one from 2-oxosuccinamic acid as a reverse reaction for Asn aminotransferase, and the other by the hydrolysis of β-cyanoalanine using nitrilase/nitrile hydratase (Sieciechowicz et al., 1988). The reverse reaction for Asn aminotransferase is not expected to occur at any great extent in planta as the reaction rates detected are very low (Ireland and Joy, 1983). The pathway involving

β-cyanoalanine requires the detoxication of cyanide by β-cyanoalanine synthase (Figure 1.1), a member of the β-substituted alanine synthases (BSASs), sharing similarity with O-acetylserine sulfhydrylase synthesizing cysteine. β-Cyanoalanine synthase, a mitochondrial localized enzyme, catalyzes the exchange of the thiol group of cysteine for cyanide, producing β-cyanoalanine. The formation of β-cyanoalanine is of particular significance in certain tissues, especially during root development and seed germination. The nitrilase NIT4 acts both as a hydratase and nitrilase for β-cyanoalanine catabolism producing Asn or Asp and ammonia, respectively (Piotrowski et al., 2001; Piotrowski and Volmer, 2006).

Two major routes exist for Asn catabolism in plants. Deamidation and release of Asp, as well as ammonium, is a necessary step for transformation of Asn from its storage N form into other amides and amino acids. This reaction is catalyzed by an enzyme asparaginase (ASPG) and represents the first route of Asn catabolism (Figure 1.1). ASPG are commonly present in all organisms and are classified as bacterial (type I and II) and plant (type III) – type ASPGs. Plant type ASPGs are further subdivided in K^+ dependent and K^+ independent ASPGs, due to their dependency on K^+ for activation (Bruneau et al., 2006). While K^+ dependent ASPGs have a preference of substrate toward Asn, K^+ independent ASPGs are known to have catalytic activity towards both Asn and isoaspartyl dipeptides (Gabriel et al., 2012). The second route of Asn catabolism involves transamination of the α-amino group of Asn to a 2-oxo acid acceptor producing 2-oxosuccinamate followed by its hydrolysis to form oxaloacetate and ammonia via ω-amidase (Zhang and Marsolais, 2014) (Figure 1.1). 2-Oxosuccinamate can also be rapidly converted to 2-hydroxysuccinamate, however, the enzyme involved in this reaction is yet to be characterized. In *Arabidopsis*, the transamination reaction is shown to be catalysed by a serine:glyoxylate aminotransferase, AGT1, which has a higher catalytic activity towards Asn than its prototypical substrates serine or alanine (Zhang and Marsolais, 2014). A comprehensive review of the enzymes involved in Asn biosynthesis and degradation and their gene family members in Arabidopsis was recently provided by Gaufichon et al. (2016b).

1.3.2 Asn as a transport and storage form of N

The N requirements of actively growing roots and leaves are mostly met by 1) inorganic sources (NO_3 and NH_4^+) that are primarily transported via xylem, and 2) a restricted group of organic sources that is largely comprised of Asn (transported through both xylem and phloem) (Sieciechowicz et al., 1988). Asn is very well established as a primary nitrogen transport compound as suggested by its relative abundance in the xylem and phloem sap of lupins and its high N:C ratio (1:2) (Atkins et al., 1983). In its soluble form, Asn is a substrate for only a few enzymatic reactions and has low net charge under physiological conditions making it an ideal storage compound (Lea et al., 2007). Early work by Pate et al. (1981) in lupin showed that Asn was the major amino acid in all plant parts, especially in nodulated roots, leaves and pods where it could account for 60-80% of total amino acid content.

A thorough investigation of nitrogen sources in the xylem sap of different legume species revealed that although ureides form majority of nitrogen transport compounds in nodulated *Glycine max*, *Vigna unguiculata* and *Phaseolus aureus*, amino acids and nitrate predominate under non-nitrogen fixing conditions (Amarante et al., 2006). Only amino acids and nitrate were found in the xylem sap of other legume species viz., *Crotalaria juncea*, *Pisum sativum* and *Lupinus albus* under both nodulated and non-nodulated conditions. Further analysis of amino acid fractions of these six plants also revealed that Asn was the major form of N transport under most conditions (Amarante et al., 2006). The same study also highlighted the importance of Asn as an indicator of N status regardless of whether nitrogen fixation was the source of nitrogen. Furthermore, a comprehensive review of Asn in plants highlighted the role of Asn as a storage and transport form of nitrogen in several perennials (Lea et al., 2007). It is also noteworthy that glutamine and arginine may carry out the same role as Asn depending on species and type of tissue (Lea et al., 2007). While most of the studies on Asn have been conducted in legumes, *Arabidopsis thaliana*, a well-established model organism, was also shown to utilize Asn as one of the primary N storage and transport forms in a photoperiod dependent manner (Lam et al., 1995).

1.3.3 Role of Asn in N metabolism and C/N relations

Biosynthesis of Asn is primarily dependent on the N status and C to N ratio of the plant and is regulated through the expression of *AS* genes. In nodulated soybean roots, AS activity under N stress is decreased, resulting in reduced Asn concentration in the xylem sap (Lima and Sodek, 2003). The change in Asn concentration is reversible and the recovery closely parallels AS activity in nodules (*Glycine max*) (Lima and Sodek, 2003). All three soybean *AS* genes are repressed in root under N deficit (Antunes et al., 2008). Nitrate is involved as a signal in the up-regulation of *AS* genes in Arabidopsis (Scheible et al., 2004; Wang et al., 2003; Wang et al., 2004). Cytokinins, a group of plant hormones, act as a secondary signal of nitrate availability. The expression of cytokinin biosynthetic and signaling genes is up-regulated by nitrate (Krouk, 2016; Sakakibara et al., 2006). In turn, cytokinins up-regulate the expression of *ASN1*, *GDH1*, and *GDH2* (Brenner et al., 2005). Transcript levels of *ASN1* in Arabidopsis are regulated by the C:N ratio. Sucrose suppresses the accumulation of *ASN1* transcript in dark grown plants, however, Asn, glutamine or glutamate prevent this effect (Lam et al., 1994). Similar results were obtained in maize root tips (Chevalier et al., 1996). Arabidopsis *ASN2* is regulated differently. It is induced by sucrose, and repressed by Asn, glutamine or glutamate (Lam et al., 1998). However, by profiling transcripts and performing network analysis under treatments with inorganic N as nitrate and ammonium, or organic N as glutamate or glutamine, in the presence or absence of the GS inhibitor methionine sulfoximine, Gutiérrez et al. (2008) showed that *ASN1* responds to glutamate or a downstream metabolite, in concert with *GDH1*, whereas the K^+-dependent *ASPGB1* had an opposite regulation. *ASN2* positively responds to inorganic N. The transcription factor *BZIP1* regulates *ASN1* (Hanson et al., 2008; Para et al., 2014).

Photoperiod is an important regulator of Asn metabolism. Transcripts of pea *AS1*, encoding a class I AS, are induced in leaves in the dark (Tsai and Coruzzi, 1990). This response results in transient accumulation of Asn at night, under conditions of low C:N ratio (Harmer et al., 2005). The opposite happens with starch reserves, which are mobilized during the night. A similar situation is observed in Arabidopsis, where expression of the orthologous *ASN1* is induced in leaves at night and correlates with Asn

concentration (reviewed in Coruzzi, 2003; Lam et al., 1995). Arabidopsis *ASN2* is associated with ammonium assimilation under light conditions (Wong et al., 2004). An *ASN2* mutant was impaired in salt stress induced, ammonium-dependent Asn accumulation (Maaroufi-Dguimi et al., 2011). Careful analysis of *ASN2* loss-of-function mutants revealed that the gene is involved in N assimilation and Asn export to the phloem. The mutant plants were characterized by impaired growth, increased ammonium accumulation, reduced Asn, and increased γ-aminobutyric acid and alanine concentration in leaves (Gaufichon et al., 2013). This was interpreted as a consequence of increased flux through the γ-aminobutyric acid shunt, leading to alanine biosynthesis via the transamination of pyruvate. The authors indicated that decreased aspartate utilization for Asn biosynthesis was likely to reduce the flux from phosphoenolpyruvate to oxaloacetate, resulting in higher availability of pyruvate and fumarate. *ASN2* expression is localized in leaf phloem companion cells, and the *ASN2* mutants displayed reduced Asn concentration in phloem exudates, and delayed senescence. Loss of function of the closely related *ASN3* led to a relatively milder phenotype (Gaufichon et al., 2016a). The general picture emerging from these data is that class I AS is involved in the regulation of Asn levels in response to the C:N ratio, likely in concert with a glutamate dehydrogenase breaking down the glutamate generated by the AS reaction, generating α-ketoglutarate to balance carbon metabolism.

1.3.4 Asn as an N signal in plants

Whether Asn acts as a metabolite signal, and if so how it is perceived, is an open area of investigation. Beside glutamate and glutamine discussed above, histidine has been recently implicated in the regulation of abscisic acid biosynthesis and fatty acid oxidation (Ma and Wang, 2016). An understanding of amino acid signaling in plants lags behind that of other organisms. Proteins involved in amino acid signaling in plants include 1) the chloroplastic PII, involved in the regulation of Arg biosynthesis and lipid accumulation, and a glutamine sensor, except in the Brassicaceae (Chellamuthu et al., 2014; Feria Bourrellier et al., 2010), 2) the GCN2 protein kinase involved in the perception of amino acid deficiency through binding of uncharged tRNAs, which phosphorylates eIF2α (α-subunit of eukaryotic translation initiation factor (Li et al., 2013; Wang et al., 2017), 3)

the protein kinase target of rapamycin (TOR), a positive regulator of translation and growth processes in response to carbon and N nutrients (Dobrenel et al., 2016), 4) an aspartyl-tRNA synthetase involved in the perception of the immunomodulatory β-aminobutyric acid (Luna et al., 2014), and 5) glutamate receptor homologs, one of which was recently implicated in the control of stomatal opening by acting as a calcium channel in response to methionine binding (Kong et al., 2016).

The activation of mammalian TOR Complex 1 (mTORC1) requires amino acid sensors such as CASTOR1 for arginine and Sestrin2 for leucine (Saxton et al., 2016; Wolfson et al., 2015). Asn was recently identified in a group of amino acids involved in priming mTORC1 for activation by other amino acids including leucine (Dyachok et al., 2016). In mammalian cancer cells, Asn was recently shown to act as a regulator of amino acid uptake, particularly serine, arginine, and histidine, to coordinate nucleotide and protein synthesis and to be an essential metabolite for cell proliferation (Krall et al., 2016). Clearly, more work is required to understand the role and perception of Asn as a potential metabolite signal of N status in higher plants. To that end, Chapter 2 of the current thesis explores the responses of Arabidopsis roots to Asn treatment at both the transcript and metabolite levels and provides a comprehensive set of changes that occur in roots upon increased Asn accumulation.

1.4 Glutamine amidotransferases in plants

As pointed out in section 1.2, the N from Gln is used in the production of a number of nitrogenous compounds including nucleobases, vitamins, and amino sugars. Glutamine amidotransferases (GATs) belong to a family of enzymes that typically consist of two domains or subunits and catalyze a two-step reaction. The first is a glutaminase that hydrolyses glutamine to release ammonia and the second is a synthase that catalyzes the transfer of ammonia to an acceptor substrate to form product (Zalkin and Smith, 2006). The catalytic mechanism of GAT consists of the nucleophilic attack of the thiol group of a catalytic cysteine residue on the δ-carbonyl group of glutamine releasing ammonia that is channeled to act as a nucleophile on ammonia accepting substrate (Zalkin and Smith, 2006) (Figure 1.2). Asparagine synthetase, described in section 1.3.1, is an example of a GAT superfamily protein. GATs are categorized into two major classes (Table 1.1)

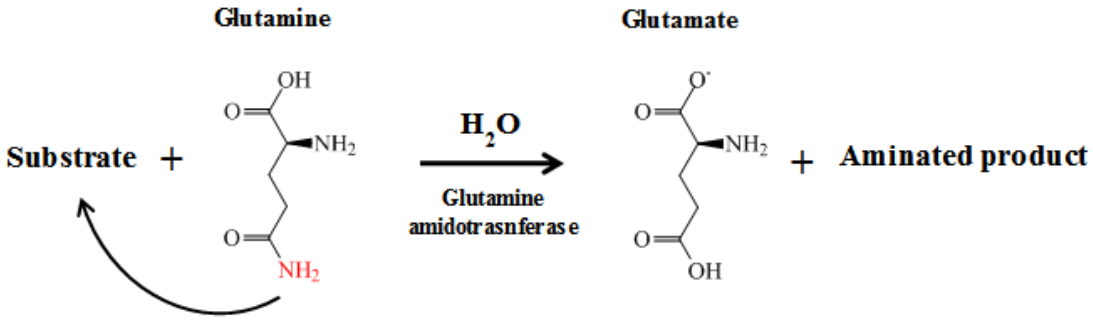


Figure 1.2: Glutamine amidotransferase transfers an amide group from glutamine to a substrate forming aminated product and releasing glutamate.

Table 1.1: Glutamine amidotransferases (GATs) for which crystal structures have been determined.

Enzyme	Abbreviation	Acceptor substrate	cofactor	Product	Reference
Triad GATs					
Anthranilate synthase	AS	Chorismate	$Mg2^{+}$	Anthranilate	Morollo and Eck (2001)
Carbamoyl phosphate synthetase	CPS	Bicarbonate	ATP/$Mg2^{+}$	Carbamoyl phosphate	Thoden et al. (1997)
Cytidine triphosphate synthetase	CTPS	Uridine triphosphate	ATP	Cytidine triphosphate	Endrizzi et al. (2004)
Phosphoribosylformyl glycinamide synthetase	FGAR-AT	Formyl phosphoribosyl glycinamide	ATP	Formamido phosphoribosyl acetamidine	Anand et al. (2004)
Guanosine monophosphate synthetase	GMPS	Xanthosine monophosphate	ATP/$Mg2^{+}$	Guanosine monophosphate	Tesmer et al. (1996)
Imidazole glycerol phosphate synthase	IGPS	Phosphoribulosyl formimino aminoimidazole carboxamide ribonucleotide	-	Imidazole glycerol phosphate	Knöchel et al. (2002)
Pyridoxal 5'-phosphate synthase	PLPS	Ribose 5'-phosphate + Glyceraldehyde 3'-phosphate	-	Pyridoxal 5'-phosphate	Zein et al. (2006)
N terminal nuclophile GATs					
Asparagine synthetase	AsnB	Aspartate	ATP	Asparagine	Nakatsu et al. (1998)
Glucosamine 6'-phosphate synthase	GlmS	Fructose 6'-phosphate	-	Glucosamine 6'-phosphate	Oliva et al. (1995)
Glutamate synthase	GltS	2-Oxoglutarate	FMN/FAD and ferridoxin (or) NAD	Glutamate	van den Heuvel et al. (2002)
Amidophosphoribosyl transferase	GPAT	Phosphoribosyl pyrophosphate	-	Phosphoribosyl amine	Muchmore et al. (1998)

depending on the active site residue. While the class I GATs, also called triad GATs, utilize histidine and glutamate residues that activate the cysteine thiol group, class II GATs use an N-terminal catalytic cysteine giving them their characteristic name, N-terminal nucleophile (Ntn) GATs (Mouilleron and Golinelli-Pimpaneau, 2007). The nomenclature of enzymes that contain glutamine amidotransferase emphasizes the acceptor substrate (Table 1.1) (Massiere and Badet-Denisot, 1998). Both class I and class II glutamine amidotransferases are widely dispersed across all organisms and the structures and functions of most of these enzymes are well characterized. A comprehensive review of these enzymes, along with their mode of action and structural properties, can be found in Mouilleron and Golinelli-Pimpaneau (2007).

In *Arabidopsis thaliana*, an uncharacterized class I glutamine amidotransferase domain containing protein, GAT1_2.1, was recently identified and found to be highly responsive to nitrogen status in roots (Zhu and Kranz, 2012). A null mutant of *A. thaliana* lacking *GAT1_2.1* showed an extensive shoot branching phenotype indicating that this nitrogen regulated gene might represent a link between nitrogen stress responses and shoot branching. Unlike most other GATs, GAT1_2.1 consists only of a glutamine amidotransferase domain and an unannotated C-terminal extension. A BLASTP search of the C-terminal extension of GAT1_2.1 showed that the domain is highly orthologous to sequences found only in the plant kingdom, indicating that this protein has a plant specific function (Zhu and Kranz, 2012). Due to its nitrogen responsiveness and the lack of an identifiable synthase domain, we hypothesized that GAT1_2.1 might be involved in primary metabolism as a glutaminase, hydrolyzing glutamine, and releasing ammonia. The glutamate produced may then be channeled into the TCA cycle via a highly co-expressed glutamate dehydrogenase (*GDH2*) and hence regulate C/N partitioning in roots. Chapter 3 of this thesis describes the functional characterization of GAT1_2.1 and provides evidence for the lack of an acceptor substrate for ammonia. Its possible role in C/N partitioning is also highlighted.

1.5 Metabolomics: A tool for advancing our understanding of 'the central dogma' of biology

To identify the metabolic responses of GAT1_2.1 to N status and its role in C/N partitioning, we decided to pursue a metabolomics approach. The metabolome, which consists of all the small molecules in a given cell/tissue, provides a functional readout of cellular state and represents the outcome of perturbations that occur at the genetic, transcriptomic and proteomic levels (Goodacre et al., 2004). Since metabolites provide a direct signature of biochemical activity – unlike genes that are subject to epigenetic regulation or proteins that are subject to post-translational modifications – they can be correlated with phenotypes and hence can be used to infer mechanistic relationships between genes and phenotypes (Patti et al., 2012). The application of metabolomics strategies, together with the other well studied "omics" technologies, have been shown to deliver a systems level understanding of complex phenotypes (Bino et al., 2004; Fiehn et al., 2000; Saito and Matsuda, 2010; Chen et al., 2017; Hirai et al., 2004; Savoi et al., 2017).

The plant metabolome consists of a wide range of chemical species with diverse physical properties ranging from hydrophilic carbohydrates and amino acids, to hydrophobic lipids, to ionic inorganic compounds (Jorge et al., 2016). An estimated 200,000 distinct metabolites – with a large dynamic range of concentrations from femtomolar to millimolar – exist in the plant kingdom (Fiehn, 2002; Fernie, 2003). This enormous complexity presents huge challenges for the analytical technologies used in plant metabolomics. Hence, well thought out strategic tools are needed to separate and characterize these diverse compounds. No single analytical technique can cover the whole metabolome due to the diversity and broad concentration range of metabolites. To that end, Chapter 4 provides an evaluation of existing extraction methods to identify one suitable for optimal metabolite coverage and provides strategies for data analysis that can improve the reliability and application of this very important 'omics technology in the field of plant science.

1.6 Objectives

Asn is a major storage and transport form of N. However, its signaling roles and mode of action are not fully understood. The long term goal of this research is thus to fully elucidate Asn metabolism, as well as its function as a metabolite signal in plants. This could uncover its potential role in source-sink relationships as well as major interlinks between C and N metabolic pathways. Considering the central role of asparagine in N metabolism, knowledge obtained from this work might also be used to help improve NUE of other agronomically important crops, and hence yield.

The objectives of the current study are as follows:

1. **To identify the transcriptional and metabolic responses to Asn in roots using the model plant, *Arabidopsis thaliana*.** Through the use of RNA-seq and metabolomics we aim to detect the responses to asparagine and identify genetic and metabolic players that contribute to C/N balance upon excess organic N supply.
2. **To characterize a N responsive class I glutamine amidotransferase, GAT1_2.1.** A major N responsive gene identified through my first objective was found to encode a plant specific GAT1_2.1. In this study, we aim to perform a biochemical characterization of this protein and identify its role in C/N partitioning.
3. **Establish an optimized metabolomics platform for detecting primary metabolites.** In order to determine changes in primary metabolism, an optimized method of extraction, chromatographic separation and data analysis are necessary. Hence, the third objective of this study is to establish a standardized pipeline from lab bench to results that can be used for metabolomics studies.

1.7 Literature cited

Amarante, L.d., Lima, J.D., and Sodek, L. (2006). Growth and stress conditions cause similar changes in xylem amino acids for different legume species. Environmental and Experimental Botany 58:123-129.

Anand, R., Hoskins, A.A., Stubbe, J., and Ealick, S.E. (2004). Domain organization of salmonella typhimurium formylglycinamide ribonucleotide amidotransferase revealed by x-ray crystallography. Biochemistry 43:10328-10342.

Antunes, F., Aguilar, M., Pineda, M., and Sodek, L. (2008). Nitrogen stress and the expression of asparagine synthetase in roots and nodules of soybean (*Glycine max*). Physiologia plantarum 133:736-743.

Atkins, C.A., Pate, J.S., Peoples, M.B., and Joy, K.W. (1983). Amino acid transport and metabolism in relation to the nitrogen economy of a legume leaf. Plant physiology 71:841-848.

Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B.J., Mendes, P., Roessner-Tunali, U., Beale, M.H., et al. (2004). Potential of metabolomics as a functional genomics tool. Trends in Plant Science 9:418-425.

Brenner, W.G., Romanov, G.A., Köllmer, I., Bürkle, L., and Schmülling, T. (2005). Immediate-early and delayed cytokinin response genes of *Arabidopsis thaliana* identified by genome-wide expression profiling reveal novel cytokinin-sensitive processes and suggest cytokinin action through transcriptional cascades. The Plant Journal 44:314-333.

Bruneau, L., Chapman, R., and Marsolais, F. (2006). Co-occurrence of both l-asparaginase subtypes in *Arabidopsis*: At3g16150 encodes a K+-dependent l-asparaginase. Planta 224:668-679.

Chellamuthu, V.-R., Ermilova, E., Lapina, T., Lüddecke, J., Minaeva, E., Herrmann, C., Hartmann, Marcus D., and Forchhammer, K. (2014). A widespread glutamine-sensing mechanism in the plant kingdom. Cell 159:1188-1199.

Chen, C., Li, C., Wang, Y., Renaud, J., Tian, G., Kambhampati, S., Saatian, B., Nguyen, V., Hannoufa, A., Marsolais, F., et al. (2017). Cytosolic acetyl-CoA promotes histone acetylation predominantly at H3K27 in *Arabidopsis*. Nature Plants.

Chevalier, C., Bourgeois, E., Just, D., and Raymond, P. (1996). Metabolic regulation of asparagine synthetase gene expression in maize (*Zea mays* L.) root tips. The Plant Journal 9:1-11.

Coruzzi, G.M. (2003). Primary N-assimilation into amino acids in *Arabidopsis*. The Arabidopsis book 2:e0010.

Dobrenel, T., Caldana, C., Hanson, J., Robaglia, C., Vincentz, M., Veit, B., and Meyer, C. (2016). TOR signaling and nutrient sensing. Annual Review of Plant Biology 67:261-285.

Duff, S.M.G. (2015). Asparagine synthetase. In: Amino acids in higher plants--D'Mello, J.P.F., ed. Monsanto, Chesterfield, Missouri, USA: CAB ebooks. 100-128.

Dyachok, J., Earnest, S., Iturraran, E.N., Cobb, M.H., and Ross, E.M. (2016). Amino acids regulate mTORC1 by an obligate two-step mechanism. Journal of Biological Chemistry 291:22414-22426.

Endrizzi, J.A., Kim, H., Anderson, P.M., and Baldwin, E.P. (2004). Crystal structure of *Escherichia coli* cytidine triphosphate synthetase, a nucleotide-regulated

glutamine amidotransferase/ATP-dependent amidoligase fusion protein and homologue of anticancer and antiparasitic drug targets. Biochemistry 43:6447-6463.

Feria Bourrellier, A.B., Valot, B., Guillot, A., Ambard-Bretteville, F., Vidal, J., and Hodges, M. (2010). Chloroplast acetyl-CoA carboxylase activity is 2-oxoglutarate–regulated by interaction of PII with the biotin carboxyl carrier subunit. Proceedings of the National Academy of Sciences 107:502-507.

Fernie, A.R. (2003). Review: Metabolome characterisation in plant system analysis. Functional Plant Biology 30:111.

Fiehn, O. (2002). Metabolomics – the link between genotypes and phenotypes. Plant Molecular Biology 48:155-171.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R.N., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nature Biotechnology 18:1157-1161.

Gabriel, M., Telmer, P.G., and Marsolais, F. (2012). Role of asparaginase variable loop at the carboxyl terminal of the alpha subunit in the determination of substrate preference in plants. Planta 235:1013-1022.

Gaufichon, L., Marmagne, A., Yoneyama, T., Hase, T., Clément, G., Trassaert, M., Xu, X., Shakibaei, M., Najihi, A., and Suzuki, A. (2016a). Impact of the disruption of ASN3-encoding asparagine synthetase on *Arabidopsis* development. Agronomy 6:12.

Gaufichon, L., Masclaux-Daubresse, C., Tcherkez, G., Reisdorf-Cren, M., Sakakibara, Y., Hase, T., clement, G., Avice, J.-C., Grandjean, O., Marmagne, A., et al. (2013). *Arabidopsis thaliana* ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. Plant, Cell & Environment 36:328-342.

Gaufichon, L., Reisdorf-Cren, M., Rothstein, S.J., Chardon, F., and Suzuki, A. (2010). Biological functions of asparagine synthetase in plants. Plant Science 179:141-153.

Gaufichon, L., Rothstein, S.J., and Suzuki, A. (2016b). Asparagine metabolic pathways in *Arabidopsis*. Plant & cell physiology 57:675-689.

Good, A.G., Shrawat, A.K., and Muench, D.G. (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9:597-605.

Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. Trends in Biotechnology 22:245-252.

Gutiérrez, R.A., Lejay, L.V., Dean, A., Chiaromonte, F., Shasha, D.E., and Coruzzi, G.M. (2007). Qualitative network models and genome-wide expression data define carbon/ nitrogen-responsive molecular machines in *Arabidopsis*. Genome Biology 8.

Gutiérrez, R.A., Stokes, T.L., Thum, K., Xu, X., Obertello, M., Katari, M.S., Tanurdzic, M., Dean, A., Nero, D.C., McClung, C.R., et al. (2008). Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene *CCA1*. Proc Natl Acad Sci U S A 105:4939-4944.

Haegele, J.W., Cook, K.A., Nichols, D.M., and Below, F.E. (2013). Changes in nitrogen use traits associated with genetic improvement for grain yield of maize hybrids released in different decades. Crop Science 53:1256-1268.

Hanson, J., Hanssen, M., Wiese, A., Hendriks, M.M.W.B., and Smeekens, S. (2008). The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of *Asparagine synthetase1* and *Proline dehydrogenase2*. The Plant Journal 53:935-949.

Harmer, S.L., Covington, M.F., Blasing, O., and Stitt, M. (2005). Circadian regulation of global gene expression and metabolism. In: Endogenous plant rhythms, Annual plant reviews--Hall, A.J.W., McWatters, H.G., ed.: Blackwell, Oxford. 133-165.

Hawkesford, M.J. (2001). Improving nutrient use efficiency in crops. In: eLS: John Wiley & Sons, Ltd.

Hirai, M.Y., Yano, M., Goodenowe, D.B., Kanaya, S., Kimura, T., Awazuhara, M., Arita, M., Fujiwara, T., and Saito, K. (2004). Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 101:10205-10210.

Hodges, M. (2002). Enzyme redundancy and the importance of 2-oxoglutarate in plant ammonium assimilation. Journal of Experimental Botany 53:905-916.

Ireland, R.J., and Joy, K.W. (1983). Purification and properties of an asparagine aminotransferase from Pisum sativum leaves. Archives of Biochemistry and Biophysics 223:291-296.

Jorge, T.F., Rodrigues, J.A., Caldana, C., Schmidt, R., van Dongen, J.T., Thomas-Oates, J., and António, C. (2016). Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress. Mass Spectrometry Reviews 35:620-649.

Knöchel, T., Pappenberger, A., Jansonius, J.N., and Kirschner, K. (2002). The crystal structure of indoleglycerol-phosphate synthase from *Thermotoga maritima* : kinetic stabilization by salt bridges. Journal of Biological Chemistry 277:8626-8634.

Kong, D., Hu, H.-C., Okuma, E., Lee, Y., Lee, H.S., Munemasa, S., Cho, D., Ju, C., Pedoeim, L., Rodriguez, B., et al. (2016). L-Met Activates Arabidopsis GLR Ca2+ Channels Upstream of ROS Production and Regulates Stomatal Movement. Cell Reports 17:2553-2561.

Krall, A.S., Xu, S., Graeber, T.G., Braas, D., and Christofk, H.R. (2016). Asparagine promotes cancer cell proliferation through use as an amino acid exchange factor. 7:11457.

Krouk, G. (2016). Hormones and nitrate: a two-way connection. Plant Molecular Biology 91:599-606.

Lam, H.-M., Hsieh, M.-H., and Coruzzi, G. (1998). Reciprocal regulation of distinct asparagine synthetase genes by light and metabolites in *Arabidopsis thaliana*. The Plant Journal 16:345-353.

Lam, H.-M., Sheila, S.Y.P., and Coruzzi, G.M. (1994). Metabolic regulation of the gene encoding glutamine-dependent asparagine synthetase in *Arabidopsis thaliana*. Plant physiology 106:1347-1357.

Lam, H.M., Coschigano, K., Schultz, C., Melo-Oliveira, R., Tjaden, G., Oliveira, I., Ngai, N., Hsieh, M.H., and Coruzzi, G. (1995). Use of *Arabidopsis* mutants and genes to study amide amino acid biosynthesis. The Plant cell 7:887-898.

Larsen, T.M., Boehlein, S.K., Schuster, S.M., Richards, N.G.J., Thoden, J.B., Holden, H.M., and Rayment, I. (1999). Three-dimensional structure of *Escherichia coli Asparagine synthetase b*: a short journey from substrate to product. Biochemistry 38:16146-16157.

Lea, P.J., and Ireland, R.J. (1999). Nitrogen metabolism in higher plants. In: Plant Amino Acids: Biochemistry and Biotechnology--Singh, B.K., ed. New York: Marcel Dekker Inc. 1-44.

Lea, P.J., Sodek, L., Parry, M.A.J., Shewry, P.R., and Halford, N.G. (2007). Asparagine in plants. Annals of Applied Biology 150:1-26.

Li, M.W., AuYeung, W.K., and Lam, H.M. (2013). The GCN2 homologue in *Arabidopsis thaliana* interacts with uncharged tRNA and uses Arabidopsis eIF2α molecules as direct substrates. Plant biology 15:13-18.

Lima, J.D., and Sodek, L. (2003). N-stress alters aspartate and asparagine levels of xylem sap in soybean. Plant Science 165:649-656.

Luna, E., van Hulten, M., Zhang, Y., Berkowitz, O., López, A., Pétriacq, P., Sellwood, M.A., Chen, B., Burrell, M., van de Meene, A., et al. (2014). Plant perception of β-aminobutyric acid is mediated by an aspartyl-tRNA synthetase. Nat Chem Biol 10:450-456.

Ma, H., and Wang, S. (2016). Histidine regulates seed oil deposition through abscisic acid biosynthesis and β-oxidation. Plant physiology 172:848.

Maaroufi-Dguimi, H., Debouba, M., Gaufichon, L., Clément, G., Gouia, H., Hajjaji, A., and Suzuki, A. (2011). An *Arabidopsis* mutant disrupted in *ASN2* encoding asparagine synthetase 2 exhibits low salt stress tolerance. Plant Physiology and Biochemistry 49:623-628.

Massiere, F., and Badet-Denisot, M.A. (1998). The mechanism of glutamine-dependent amidotransferases. Cellular and molecular life sciences : CMLS 54:205-222.

McAllister, C.H., Beatty, P.H., and Good, A.G. (2012). Engineering nitrogen use efficient crop plants: the current status. Plant biotechnology journal 10:1011-1025.

Morollo, A.A., and Eck, M.J. (2001). Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*. Nat Struct Mol Biol 8:243-247.

Mouilleron, S., and Golinelli-Pimpaneau, B. (2007). Conformational changes in ammonia-channeling glutamine amidotransferases. Current Opinion in Structural Biology 17:653-664.

Muchmore, C.R., Krahn, J.M., Kim, J.H., Zalkin, H., and Smith, J.L. (1998). Crystal structure of glutamine phosphoribosylpyrophosphate amidotransferase from *Escherichia coli*. Protein Science : A Publication of the Protein Society 7:39-51.

Nakatsu, T., Kato, H., and Oda, J.i. (1998). Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase. Nat Struct Mol Biol 5:15-19.

Noctor, G., Dutilleul, C., De Paepe, R., and Foyer, C.H. (2004). Use of mitochondrial electron transport mutants to evaluate the effects of redox state on photosynthesis, stress tolerance and the integration of carbon/nitrogen metabolism. Journal of Experimental Botany 55:49-57.

Oliva, G., Fontes, M.R.M., Garratt, R.C., Altamirano, M.M., Calcagno, M.L., and Horjales, E. (1995). Structure and catalytic mechanism of glucosamine 6-

phosphate deaminase from *Escherichia coli* at 2.1 Å resolution. Structure 3:1323-1332.

Para, A., Li, Y., Marshall-Colón, A., Varala, K., Francoeur, N.J., Moran, T.M., Edwards, M.B., Hackley, C., Bargmann, B.O.R., Birnbaum, K.D., et al. (2014). Hit-and-run transcriptional control by *bZIP1* mediates rapid nutrient signaling in Arabidopsis. Proceedings of the National Academy of Sciences 111:10371-10376.

Pate, J.S., Atkins, C.A., Herridge, D.F., and Layzell, D.B. (1981). Synthesis, Storage, and Utilization of Amino Compounds in White Lupin (*Lupinus albus* L.). Plant physiology 67:37-42.

Patti, G.J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the apogee of the omic triology. Nature reviews. Molecular cell biology 13:263-269.

Piotrowski, M., Schönfelder, S., and Weiler, E.W. (2001). The *Arabidopsis thaliana* Isogene NIT4 and Its Orthologs in Tobacco Encode β-Cyano-l-alanine Hydratase/Nitrilase. Journal of Biological Chemistry 276:2616-2621.

Piotrowski, M., and Volmer, J.J. (2006). Cyanide metabolism in higher plants: cyanoalanine hydratase is a NIT4 homolog. Plant Molecular Biology 61:111-122.

Rothstein, S.J. (2007). Returning to Our Roots: Making plant biology research relevant to future challenges in agriculture. The Plant cell 19:2695.

Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. Annual Review of Plant Biology 61:463-489.

Sakakibara, H., Takei, K., and Hirose, N. (2006). Interactions between nitrogen and cytokinin in the regulation of metabolism and development. Trends in Plant Science 11:440-448.

Savoi, S., Wong, D.C.J., Degu, A., Herrera, J.C., Bucchetti, B., Peterlunger, E., Fait, A., Mattivi, F., and Castellarin, S.D. (2017). Multi-omics and integrated network analyses reveal new insights into the systems relationships between metabolites, structural genes, and transcriptional regulators in developing grape berries (*Vitis vinifera* L.) exposed to water deficit. Frontiers in plant science 8:1124.

Saxton, R.A., Chantranupong, L., Knockenhauer, K.E., Schwartz, T.U., and Sabatini, D.M. (2016). Mechanism of arginine sensing by *CASTOR1* upstream of *mTORC1*. Nature 536:229-233.

Scheible, W.R., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D., Palacios-Rojas, N., Schindelasch, D., Thimm, O., Udvardi, M.K., and Stitt, M. (2004). Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. Plant Physiol 136:2483-2499.

Sieciechowicz, K.A., Joy, K.W., and Ireland, R.J. (1988). The metabolism of asparagine in plants. Phytochemistry 27:663-671.

Tesmer, J.J.G., Klem, T.J., Deras, M.L., Davisson, V.J., and Smith, J.L. (1996). The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families. Nat Struct Mol Biol 3:74-86.

Thoden, J.B., Holden, H.M., Wesenberg, G., Raushel, F.M., and Rayment, I. (1997). Structure of Carbamoyl Phosphate Synthetase: A journey of 96 Å from substrate to product. Biochemistry 36:6305-6316.

Tsai, F.Y., and Coruzzi, G.M. (1990). Dark-induced and organ-specific expression of two asparagine synthetase genes in *Pisum sativum*. The EMBO Journal 9:323-332.

van den Heuvel, R.H.H., Ferrari, D., Bossi, R.T., Ravasio, S., Curti, B., Vanoni, M.A., Florencio, F.J., and Mattevi, A. (2002). Structural studies on the synchronization of catalytic centers in glutamate synthase. Journal of Biological Chemistry 277:24579-24583.

Wang, L., Li, H., Zhao, C., Li, S., Kong, L., Wu, W., Kong, W., Liu, Y., Wei, Y., Zhu, J.-K., et al. (2017). The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in *Arabidopsis thaliana*. Plant, Cell & Environment 40:56-68.

Wang, R., Okamoto, M., Xing, X., and Crawford, N.M. (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132:556-567.

Wang, R., Tischner, R., Gutierrez, R.A., Hoffman, M., Xing, X., Chen, M., Coruzzi, G., and Crawford, N.M. (2004). Genomic analysis of the nitrate response using a nitrate reductase-null mutant of *Arabidopsis*. Plant Physiology 136:2512-2522.

Wolfson, R.L., Chantranupong, L., Saxton, R.A., Shen, K., Scaria, S.M., Cantor, J.R., and Sabatini, D.M. (2015). *Sestrin2* is a leucine sensor for the *mTORC1* pathway. Science 351:43.

Wong, H.-K., Chan, H.-K., Coruzzi, G.M., and Lam, H.-M. (2004). Correlation of *ASN2* gene expression with ammonium metabolism in *Arabidopsis*. Plant physiology 134:332-338.

Zalkin, H., and Smith, J.L. (2006). Enzymes utilizing glutamine as an amide donor. In: Advances in enzymology and related areas of molecular biology: John Wiley & Sons, Inc. 87-144.

Zein, F., Zhang, Y., Kang, Y.-N., Burns, K., Begley, T.P., and Ealick, S.E. (2006). Structural insights into the mechanism of the plp synthase holoenzyme from *Thermotoga maritima*. Biochemistry 45:14609-14620.

Zhang, Q., and Marsolais, F. (2014). Identification and characterization of omega-amidase as an enzyme metabolically linked to asparagine transamination in Arabidopsis. Phytochemistry 99:36-43.

Zhu, H., and Kranz, R.G. (2012). A nitrogen-regulated glutamine amidotransferase (gat1_2.1) represses shoot branching in *Arabidopsis*. Plant physiology 160:1770-1780.

2 COMBINED TRANSCRIPTOMIC AND METABOLITE DATA FROM ASN TREATED ARABIDOPSIS ROOTS PROVIDES NEW INSIGHTS INTO N METABOLISM

2.1 Introduction

In the past decade, there has been a tremendous progress in the field of nitrogen (N) metabolism, sensing and signaling in plants. The basic processes of N metabolism, including N uptake by plant roots in an inorganic form as NO_3^- or NO_2^-, their conversion to NH_4^+ by nitrate and nitrite reductase (NR, NiR), and NH_4^+ assimilation to organic forms like glutamine and glutamate via the GS/GOGAT cycle are well elucidated at both the biochemical and molecular levels (Hodges, 2002; Lea and Ireland, 1999; Wang et al., 2003). With the development of sophisticated microarray technologies, a number of N responsive genes that are likely involved in the key processes of N uptake and utilization have been identified (Gutiérrez et al., 2007; Wang et al., 2004). Upon nitrate treatment, these genes include, but are not limited to, nitrate transporters, ammonium transporters, nitrate reductase, glutamine and asparagine synthetases and a few genes involved in glycolysis, trehalose-6-phosphate metabolism, iron transport and sulphate uptake (Wang et al., 2003). This indicates that there is an interaction between carbon (C) and nitrogen (N) utilization pathways to achieve a desired C/N balance based upon the availability of high nitrate. Wang et al. (2003) also reported that the responses in roots were much more extensive compared to shoots based on the number of differentially expressed genes within these tissues. A study with a NR null mutant to differentiate responses to nitrate and ammonium revealed that starch mobilization occurs in response to downstream reduced N, as opposed to nitrate, highlighting that the coordination between C and N metabolism occurs beyond the level of N uptake (Wang et al., 2004).

Furthermore, the use of methionine sulphoximine, an inhibitor of glutamine synthetase (GS), to identify responses to reduced organic N as glutamine (Gln) as well as treatment with glutamate (Glu), revealed that a balance between organic and inorganic N controls the expression of genes involved in N-reduction, assimilation and amino acid biosynthesis (Gutiérrez et al., 2008). This study also highlighted a number of

transcription factors that respond to N status along with the involvement of the circadian clock in controlling nitrogen assimilation. Nitrogen-limiting conditions, however, were identified to show a substantially different transcriptome response compared to nitrate induction (Peng et al., 2007). These responses include, but are not limited to, activation of genes involved in protein degradation via the ubiquitin-proteosome pathway, signal transduction pathways, and phenylpropanoid biosynthesis (Peng et al., 2007). Transcriptomic data combined with metabolic profiles were also demonstrated to provide a link between gene expression and metabolism (Hirai et al., 2004). This allowed interaction maps between genes and metabolites in response to C and N treatments to be generated (Gutiérrez et al., 2007; Hirai et al., 2004; Scheible et al., 2004). Despite these advances, the identity of amino acid signals and the mechanisms involved in the perception of the internal organic N status remain to be elucidated.

Asparagine (Asn) is another source of amino and amide nitrogen that is used for amino acid interconversion, intra and intercellular transport and is an efficient source of nitrogen storage and partitioning compounds (Gaufichon et al., 2016; Lea et al., 2007). Asn also elicits several growth responses in *Arabidopsis* when supplied at high concentrations such as delayed seed germination and inhibition of root elongation and root hair formation (Ivanov et al., 2012). Considering the importance of Asn as a major storage and transport form of N to the sink tissues, the current study aims to elucidate the responses of Arabidopsis roots upon Asn treatment using a combination of transcriptomic and metabolomics approaches.

2.2 Results and discussion

In order to identify the transcriptional responses upon Asn treatment, we used 10 day old *Arabidopsis* seedlings transferred to plates with 20 mM Asn or no nitrogen (control) for 2 hours (described in Materials and Methods). A total of 177 genes were found to be differentially expressed with an FDR < 0.001, while a total of 317 genes were differentially expressed with an FDR < 0.01 (Appendix A, B). A diagnostic fit of variance showing a tight spread of per gene variance and a comparison of gene fold change in relation to the average number of reads mapped is provided in the appendix (Appendix C & D). Out of the 177 genes differentially expressed with FDR < 0.001, 12 were found to be repressed and 96 up-regulated with a threshold value of two fold. We used this list of genes for further analysis at the single gene level and for GO categorization.

2.2.1 Differentially expressed list includes carbon/nitrogen transport and storage related genes

Among the top up-regulated genes were the mitochondrial 5S and cytosolic 18S rRNAs and genes related to N storage, including cruciferins, *CRU1* and *CRU3* although these were expressed at relatively low levels. A number of genes involved in cytokinin biosynthesis, iron transport and sulphate uptake which are typical of nitrate induction response were also found to be upregulated (Table 2.1). Among the genes most repressed by Asn were an aluminum-activated malate transporter (*ALMT2*), and three nitrate transporters, *NRT1.8* and *NRT1.5*, associated with nitrate export from the xylem and cadmium tolerance (Chen et al., 2012; Li et al., 2010) and *NRT2.6* associated with the generation of reactive oxygen species and pathogen resistance (Dechorgnat et al., 2012). *ALMT2 (At1g08440)* is a gene found in tandem with *ALMT1* (*At1g08430*) which was previously shown to be involved in malate efflux from the roots under aluminum and low pH toxicity (Hoekenga et al., 2006). Expression of *ALMT1* is regulated by a Cys2-His2 zinc-finger domain containing transcription factor, *STOP1* (Iuchi et al., 2007). *STOP2* (*At5g22890*) is a physiologically minor isoform of *STOP1* that was shown to regulate the expression of most genes in the *ALMT1* network that play a key role in tolerance to aluminum and low pH toxicity in Arabidopsis, including *CIPK23*, involved in the a

Table 2.1: Selected list of genes differentially expressed upon Asn treatment based on DAVID functional categorization.

	TAIR ID	Gene name	Log2 Fold change	*p*-value
Nitrate transport	AT4G21680	*NRT 1.8*	-1.783	3.25E-15
	AT3G45060	*NRT 2.6*	-1.581	1.17E-13
	AT1G32450	*NRT 1.5*	-0.950	2.99E-07
Organic acid transport	AT1G08440	*aluminium acitivated malate transporter (AlMT)*	-2.918	4.44E-31
Ion transport	AT1G60960	*IRT3 (Fe2+) transporting protein*	0.748	1.53E-06
	AT5G24660	*LSU2 response to low sulphur 2*	-1.184	1.71E-06
Lignin degradation	AT3G09220	*LAC7*	1.008	4.81E-07
	AT5G01040	*LAC8*	0.791	2.45E-06
	AT5G01050	*LAC9*	1.224	3.01E-10
Oxidoreductase	AT2G48080	*2OG-Fe(II) oxygenase family protein*	1.235	8.10E-07
	AT3G13610	*2OG-Fe(II) oxygenase family protein*	1.299	3.25E-13
	AT4G10500	*2OG-Fe(II) oxygenase family protein*	1.561	1.32E-08
	AT2G19800	*MIOX2 inositol oxygenase*	1.795	1.85E-17

	AT1G48130	*AtPER1 1-Cys peroxiredoxin PER1*	0.000	2.76E-07
Glycosidase	AT4G25810	*XTR6 xyloglucan:xyloglucosyl transferase*	1.121	1.35E-09
	AT4G30270	*MERI5B xyloglucan:xyloglucosyl transferase*	1.239	1.80E-06
	AT5G57560	*TCH4 xyloglucan endotransglucosylase/hydrolase protein 22*	1.133	4.38E-10
	AT2G43890	*putative polygalacturonase /pectinase*	-1.186	3.52E-06
Other related genes	AT1G67110	*CYP735A2 cytokinin trans-hydrolase*	1.221	6.48E-14
	AT3G63110	*IPT3*	0.638	4.31E-07
	AT5G44120	*CRU1*	3.380	7.64E-30
	AT4G28520	*CRU3*	Inf	1.02E-07
	AT3G30775	*proline dehydrogenase 1*	1.562	6.23E-06
	AT5G38710	*proline dehydrogenase 2*	1.575	2.62E-09
	AT5G18840	*sugar transporter ERD6-like*	1.299	3.25E-13
	AT1G71880	*SUC1*	1.006	3.20E-06
	AT1G15040	*class I glutamine amidotransferase*	2.108	2.65E-19

regulation of K^+ transport, *SULTR3;5*, a sulphate transporter involved in ion-homeostasis, *PGIP1*, cell-wall stabilizing polygalacturonase inhibiting protein-1, *GDH1*, involved in glutamate metabolism and *At5g38200*, a class-I glutamine amidotransferase containing protein of unknown function (Kobayashi et al., 2014). Our data shows an increase in expression of STOP2 along with *CIPK15*, a homolog of *CIPK23* and *GAT1_2.1*, another class-I glutmine amidotransferase, a homolog of *At5g38200*. Here we suggest a potential role of *ALMT2*, regulated by *STOP2*, as a signal for malate levels in cells and a downregulation of this gene could result in decreased transport of malate outside of roots ensuring a sufficient supply of TCA intermediates. A recent review by Gilliham and Tyerman (2016) highlighted a potential membrane signaling role of ALMT family proteins in response to malate and γ-aminobutyric acid (GABA) levels and trigger changes in TCA cycle activity. This led to the characterization of metabolic changes via the metabolite profiling approach discussed in the later sections.

2.2.2 Over-represented GO categories indicate C and N reprogramming

Another level of analysis was performed using Gene Ontology categorization based on biological process with the 177 differentially expressed genes with an FDR < 0.001 using PANTHER database *version 10.0* (Mi et al., 2016). Nitrate transport, organic acid transport and anion transport were found to be among the over represented GO categories with over 5-fold enrichment (Table 2.2) suggesting a relocation of C and N resources. Table 2.2 represents GO categories with over 4-fold enrichment. To further inspect the list of genes at a functional level, we used the DAVID bioinformatics resource. DAVID uses over 40 annotation categories such as GO terms, protein-protein interactions, protein functional domains, bio-pathways, and sequence features to perform functional categorizations (Huang et al., 2008) (Figure 2.1). This revealed an upregulation of many oxidoreductases including three 2-OG (FeII) oxygenase family proteins which have been previously implicated in using α-ketoglutarate as a co-factor in synthesis of gibberellic acids as well as some amino acids (Araujo et al., 2014). A few other categories include, but are not limited to, lignin degradation, glycosidase, and signaling (Figure 2.1). We speculate that lignin degradation may be another source for carbon supply to achieve the aforementioned C/N balance. Overall, the data suggests a major reprogramming of C and

Table 2.2: Over-represented GO categories based on biological process with > 4 fold enrichment.

GO biological process complete	Arabidopsis thaliana - REFLIST (26684)	No. of Genes	Expected	Fold Enrichment	*p*-value
respiratory burst (GO:0045730)	86	8	0.52	> 5	1.38E-04
respiratory burst involved in defense response (GO:0002679)	86	8	0.52	> 5	1.38E-04
peptide transport (GO:0015833)	104	8	0.63	> 5	5.75E-04
oligopeptide transport (GO:0006857)	104	8	0.63	> 5	5.75E-04
amide transport (GO:0042886)	111	8	0.67	> 5	9.35E-04
response to nitrate (GO:0010167)	159	9	0.97	> 5	1.31E-03
amino acid transport (GO:0006865)	222	12	1.35	> 5	2.91E-05
carboxylic acid transport (GO:0046942)	255	13	1.55	> 5	1.43E-05
organic anion transport (GO:0015711)	275	14	1.67	> 5	3.77E-06
organic acid transport (GO:0015849)	258	13	1.57	> 5	1.64E-05
nitrate transport (GO:0015706)	162	8	0.98	> 5	1.48E-02
response to chitin (GO:0010200)	306	15	1.86	> 5	1.63E-06

response to organonitrogen compound (GO:0010243)	321	15	1.95	> 5	3.08E-06
anion transport (GO:0006820)	518	21	3.14	> 5	1.91E-08
response to nitrogen compound (GO:1901698)	593	23	3.6	> 5	4.42E-09
nitrogen compound transport (GO:0071705)	670	22	4.07	> 5	3.19E-07
ion transport (GO:0006811)	1031	30	6.26	4.79	2.09E-09
cellular response to acid chemical (GO:0071229)	517	14	3.14	4.46	7.39E-03
cellular response to oxygen-containing compound (GO:1901701)	642	16	3.9	4.11	4.24E-03

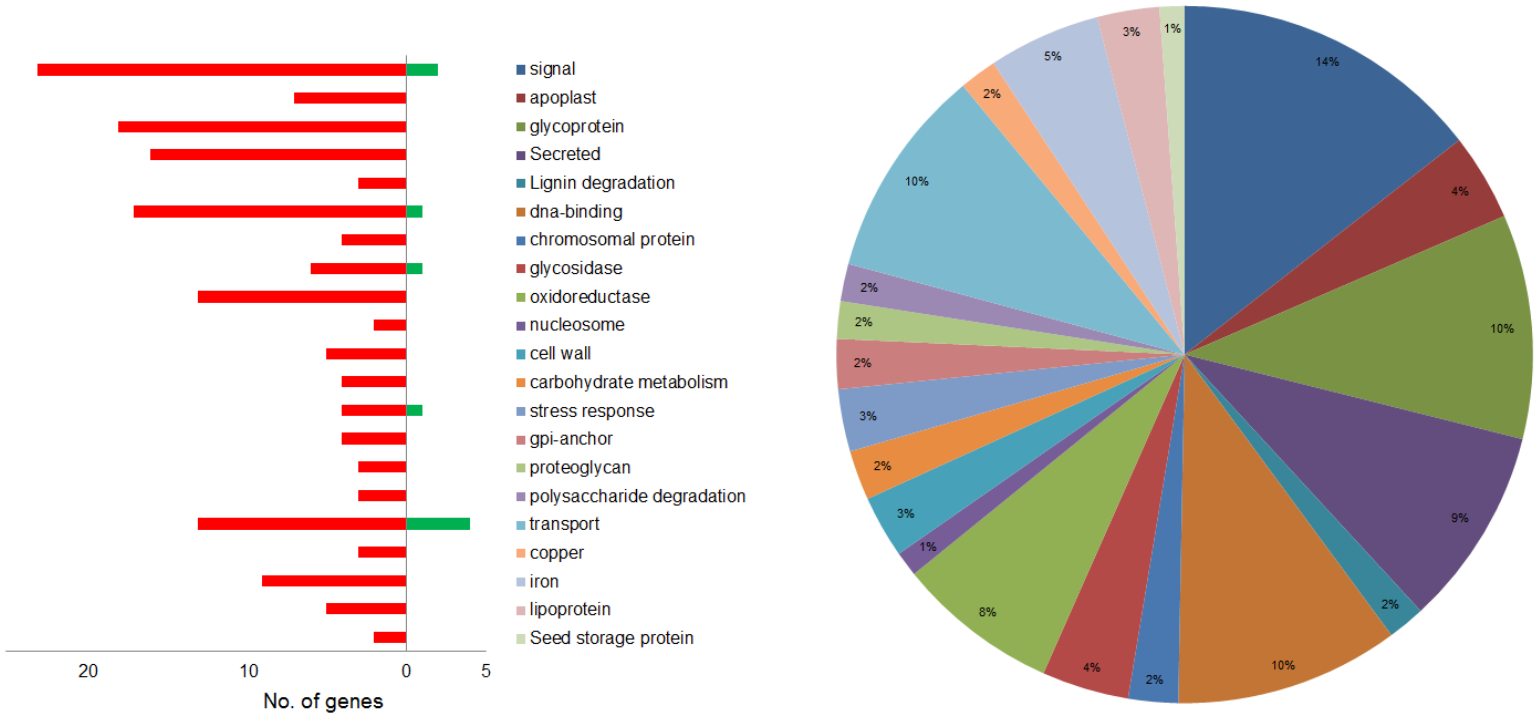


Figure 2.1: DAVID functional categorization.

Pie chart showing the functional categories assigned by DAVID bioinformatics resource. Numbers on the chart represent the percentage of genes that fall under each category among the differentially expressed genes with FDR < 0.001. Bar graph represents the number of genes within each GO category with red bars indicating the upregulated genes and green bars indicating the down-regulated genes.

N metabolites to restore the altered C/N balance. A selected list of genes that are discussed in this manuscript, within these categories, is provided in Table 2.1. A comparison of genes differentially regulated upon Asn treatment with the genes previously indicated to be responsive to inorganic and organic N by Gutiérrez et al. (2008) did not reveal any common responses between the Asn and other sources of N, indicating that the changes shown here are unique to Asn.

2.2.3 A significant accumulation of N-rich amino acids indicates N storage

We further tested the metabolic responses of Arabidopsis roots upon Asn treatment to complement the transcriptome data. Metabolite profiling upon Asn treatment focusing on a set of 35 metabolites, primarily, amino acids, organic acids and intermediates of TCA cycle (Table 2.3), provided further evidence in support of the transcriptome data. A significant accumulation of N-rich amino acids including histidine, ornithine, citrulline and arginine were identified upon Asn treatment (Table 2.3, Figure 2.2). Among the proteinogenic amino acids, arginine, which is primarily synthesized from ornithine and citrulline, has the highest N to C ratio and plays an important role in nitrogen distribution and recycling in plants (Winter et al., 2015). These results suggest a flux through arginine biosynthetic pathway via ornithine and citrulline, leading to increased arginine synthesis and nitrogen storage. Furthermore, the fact that the arginine biosynthetic pathway utilizes aspartate produced from Asn catabolysis and releases 2-oxoglutarate (2OG) to channel back to the TCA cycle (Bender, 2012) supports our argument that under high N availability, potentially, a major redistribution of C and N resources occurs to maintain C/N balance.

2.2.4 Lysine and proline catabolism along with the GABA shunt recuperate C/N balance

No significant difference was identified in the accumulation of organic acids with the exception of isocitrate. This may be because of the continued replenishment of the C pools via lysine anaplerotic pathway, which was previously suggested to be a respiratory

Table 2.3: Amino acids and TCA cycle intermediates identified using LC-MS.

Metabolite	Log 2 fold change
ILE	0.47
LEU	0.32
LYS	-0.59**
TRP	0.66**
THR	0.02
MET	0.42
ASN	5.78***
TYR	0.44*
SER	-0.36*
PRO	-0.44*
HIS	0.75**
VAL	0.20
PHE	1.01**
GLN	0.32
ARG	0.54*
Guanine	-0.27
Adenine	-0.65
CYS	0.00
GABA	-0.97**
ORN	3.95**
CIT	1.20*
Glucose 6-Phosphate	-0.66*
α-ketoisovaliric acid	1.50**
β-D-Glucosamine	-0.45**
Aminoadipate	0.67*
ASP	0.71*
GLU	0.29
Citric Acid	0.53
Isocitric Acid	0.87*
Succinic Acid	0.34
Malic Acid	0.66
2-OG	0.63
OAA	0.49
Fumaric Acid	0.57
Hydroxypyruvic acid	0.64

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

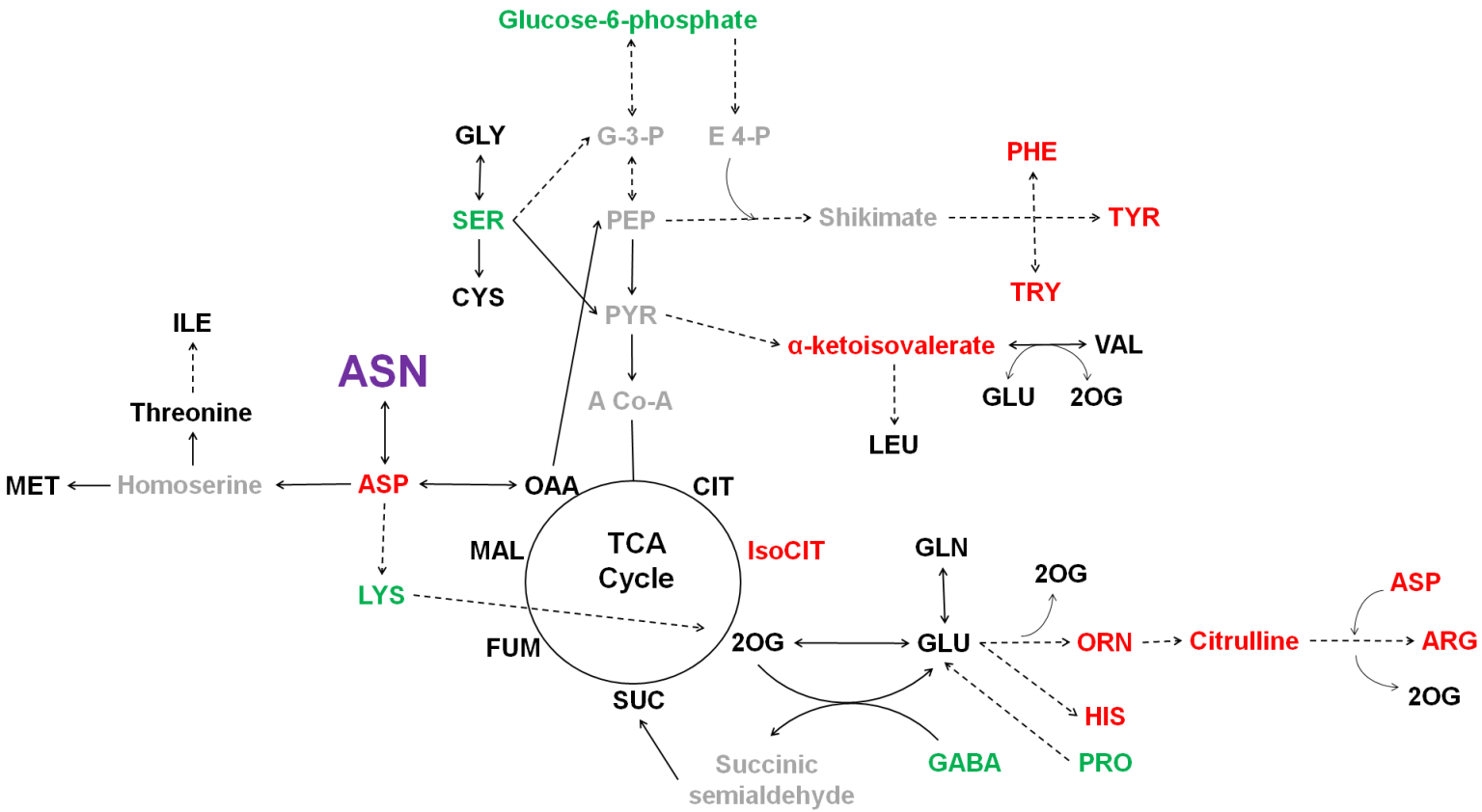


Figure 2.2: Schematic representation of the metabolite changes upon Asn treatment in the context of a pathway.

Red – increased accumulation; green – reduced accumulation; black – no significant difference; grey – not identified.

bypass sustaining 2OG production (Boex-Fontvieille et al., 2013). A 1.5-fold decrease in the accumulation of lysine and a corresponding increase in the level of aminoadipate (1.6 fold), an intermediate in this pathway, also suggest a possible flux of C in the direction of 2OG. Although less efficient, in terms of energy requirement, the lysine pathway was considered to be a preferential, alternative respiratory pathway and its activation may lead to a decreased flux of 2OG synthesis from isocitrate (via I(C)DH, isocitrate dehydrogenase) and hence the increased accumulation of isocitrate (Figure 2.2). Another possible source for maintaining steady state levels of TCA cycle metabolites and ensuring a continuous supply of Glu is proline catabolism via *proline dehydrogenase 1* and *2* (3- and 2-fold up-regulation respectively (Figure 2.2, Table 2.1). Consistent with this prediction, a significant decrease (~1.4-fold) in the accumulation of proline was identified upon Asn treatment. A significant decrease in the accumulation of γ-amino butyric acid (GABA) was also noted which may yet be another source of carbon supply via the activation of the GABA shunt pathway. In plants, the GABA shunt pathway starts with the production of GABA via Ca^{2+}/Calmodulin dependent decarboxylation of Glu via glutamate decarboxylase, followed by its conversion to succinic semialdehyde (SSA) by the action of GABA transaminase. This SSA is then oxidized to succinic acid by succinic semialdehyde dehydrogenase and is fed into the TCA cycle. The GABA shunt pathway was suggested to represent a key regulatory factor for C and N partitioning and a link between amino acid metabolism and TCA cycle (Fait et al., 2008).

2.2.5 Flux towards aromatic amino acid biosynthesis

Aromatic amino acids, phenylalanine in particular, also showed a significant accumulation, consistent with the transcriptome data considering phenylalanine is the precursor for phenylpropanoid biosynthesis. In plants, phenylalanine, tyrosine and tryptophan are primarily produced via the shikimate pathway which begins by the conversion of phosphoenolpyruvate (PEP) and erythrose 4-phosphate (E4-P) into chorismate (Tzin and Galili, 2010). A 2-fold induction of *phosphoenolpyruvate carboxykinase 2* (At5g65690) suggests a flux from aspartate towards PEP via oxaloacetate followed by activation of the shikimate pathway leading to increased accumulation of aromatic amino acids. A significant decrease in glucose-6-phosphate

(~1.6-fold), which is the precursor for E4-P via pentose phosphate pathway, was also identified, further supporting our hypothesis.

2.2.6 A class I glutamine amidotransferase (GAT1_2.1) may act as a potential source for C/N partitioning

One gene of particular interest in the transcriptome data is a class I glutamine amidotransferase, *GAT1_2.1*, of unknown substrate specificity which was up-regulated 4.3-fold. The protein has a class I Gln amidotranferase domain (with a conserved Cys, His, Glu catalytic triad necessary for enzyme function) and a C-terminal extension. Arabidopsis has three homologues that belong to the GAT1_2 subfamily, however, it was previously indicated that *GAT1_2.1* (*At1g15040*) may not have a redundant biological function compared to the other two members (*At5g38200* and *At1g66860*) since a null mutant of *At1g15040* is the only one that carries a shoot branching phenotype (Zhu and Kranz, 2012). Furthermore, *GAT1_2.1* has a root specific expression in *Arabidopsis* while *At5g38200* has a ubiquitous expression pattern with expression levels being very low and the tissue expression of *At1g66860* was unknown due to its absence on the ATH1 microarray (Arabidopsis eFP browser available in BAR database) (Winter et al., 2007).

GAT1_2.1 is highly responsive (50-fold repressed) to N limiting conditions and is known to be involved in shoot branching control in Arabidopsis (Zhu and Kranz, 2012). An *in silico* analysis of its protein sequence, based on domain structure and gene co-localization in bacteria (Hanson et al., 2010), using the SEED database (Overbeek et al., 2013) revealed the potential function of GAT1_2.1 as a glutaminase involved in the hydrolysis of glutamine, catalysing the formation of Glu and NH_3. The glutamate may then be deaminated to 2-oxoglutarate by Glu dehydrogenase (GDH) and channeled into the TCA cycle, thus providing another cross-roads between amino acid and organic acid metabolism. An NAD(H)-dependent GDH, GDH2, localized in mitochondria was shown to perform this function, using a *gdh2-1* mutant of *Arabidopsis* (Miyashita and Good, 2008). Interestingly, a co-expression search using atted-II (Version 7.1) (Obayashi et al., 2014), revealed that *GDH2* is the gene most highly co-expressed with *GAT1_2.1* in relation with tissue specificity and dependence on photoperiod. Although *GAT1_2.1*

showed an increase in transcript levels upon Asn treatment, no significant difference was identified in the accumulation of Glu at the metabolite level. Since Glu is a central metabolite involved in C/N interactions, its biological levels are generally known to be steady (Bender, 2012). To that end, the function GAT1_2.1 requires further investigation and hence Chapter 3 of this thesis aims to characterize its function at both the biochemical and metabolic levels.

2.3 Concluding remarks and future prospects:

In the present study we used a combination of transcriptomics and metabolomics data to identify the responses of Arabidopsis root to Asn treatment. The data does not provide sufficient evidence for a broad signaling role of Asn in plants. However, it does highlight the reallocations that are necessary in order to maintain C and N balance upon increase in N metabolites. It is important, however, to keep in mind that the data represents changes in both transcripts and metabolites upon a 2 hour Asn treatment. Along with the GS/GOGAT cycle, which plays a predominant role in glutamine metabolism and a link between C/N partitioning in shoots, we predict the existence of a potential new pathway, under N nutrient perturbations, for glutamine metabolism and C/N partitioning in roots via *GAT1_2.1*.

2.4 Materials and Methods:

2.4.1 Plant material and growth conditions:

Arabidopsis thaliana ecotype *Columbia* was used for both transcript and metabolite profiling experiments. Plants were grown on vertical plates at 22°C under continuous light (ca. 70 μmol m^{-2} s^{-2}), as previously described by Ivanov et al. (2012) on a defined nutrient medium containing a final concentration of 10 mM potassium phosphate (pH 6.5), 5 mM KNO_3, 2 mM $MgSO_4$, 1 mM $CaCl_2$, 125 μg FeNaEDTA, micronutrients (50 mM H_3BO_3, 12 mM $MnSO_4$, 1 mM $ZnCl_2$, 1 mM $CuSO_4$ and 0.2 mM Na_2MoO_4), 1% sucrose and 1% agar (Wang et al., 2003). Ten-day old seedlings were transferred to plates containing the same medium without nitrogen as control or 10 mM Asn as sole N source. After 2 h, root tissue was harvested, frozen in liquid N_2 and stored at -80°C until RNA or total metabolite extractions was carried out.

2.4.2 RNA extraction and transcript profiling:

Total RNA was extracted from root tissue (100 mg) of wild-type, Asn treated and untreated seelings, using the RNeasy Plant Mini Kit (Qiagen) following the manufacturers protocol. RNA was quantified with a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Burlington, ON). RNA (5 μg) was treated with amplification grade DNase I (Life Technologies) to minimize DNA contamination that may have occurred during the RNA extraction process. A final concentration of 100 ng/μl was used for transcript profiling. Transcripts were profiled by Illumina Gaiix/HiSeq2000 sequencing using single reads of 36 bp and a single channel per replicate (n = 4). Read sets for each condition were mapped to the Arabidopsis TAIR10 c*DNA* gene model using BWA and the resulting BAM files filtered for reads possessing a mapping quality ≥ 20 and a single best hit against a gene model using Samtools (Li and Durbin, 2010; Li et al., 2009). Read counts from each condition were normalized by their effective library sizes and differential expression evaluated for each gene. The *p*-values were adjusted for false discovery using the method of Benjamini and Hochberg (1995). Diagnostic plots for the variance fitting was performed by the DESeq package in R (Anders and Huber, 2010). GO categorization was performed using PANTHER database *v. 10.0* (Mi et al., 2016)

applying Bonferroni correction for multiple hypothesis testing and functional categorization was performed using DAVID bioinformatics resource *v. 6.7* (Huang et al., 2008).

2.4.3 Total metabolite extraction:

One hundred mg of frozen root tissue was homogenized in 2 ml Eppendorf tubes twice for 25 seconds at maximum speed using a tissue lyser. The metabolites were isolated in 2 ml of homogeneous mixture of -20°C methanol:methyl-tert-butyl-ether:water (1:3:1), with shaking for 30 min at 4°C, followed by another 10 min of incubation in an ice-cooled ultra-sonication bath. A volume of 650 µl of methanol:water (1:3) was then added and the homogenate was vortexed and spun for 5 min at 4°C in a table top centrifuge. This allowed phase separation, providing the upper organic phase, containing the lipids, a lower aqueous phase containing the polar and semi polar metabolites, and a pellet of starch and proteins at the bottom of tube. The aqueous phase was then collected and dried down in a speed-vac centrifuge and stored at -80°C until further use. During the day of use, the dried down sample was re-suspended in 500 µl of HPLC grade water and used for data acquisition. This method was adapted from Giavalisco et al. (2011).

2.4.4 Data acquisition and analysis:

All MS data were acquired on a Q-Exactive Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an Agilent 1290 high performance liquid chromatography (HPLC) system. A Zorbax Eclipse Plus RRHD C18 column (2.1 X 50 mm, 1.8 µm) was maintained at 35°C. Mobile phases were comprised of water with 0.1% formic acid (A), and acetonitrile with 0.1% formic acid (B). Mobile phase A was maintained at 100% for 1.25 min and mobile phase B was increased to 50% over 2.25 min, and 100% over 0.5 min. The following heated electrospray ionization (HESI) conditions were optimized for the analysis of amino and organic acids: spray voltage, 3.9 kV (ESI+), 3.5 kV (ESI-); capillary temperature, 250 °C; probe heater temperature, 100 °C; sheath gas, 30 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 60%. Injections of 5 µl were used with a flow rate of 0.3 mL min^{-1}. Compounds were

detected and monitored using a targeted MS/MS and non-targeted, data-dependent MS2. Top 5 ddMS2 experiments were performed in both positive and negative ionization modes by first acquiring a full MS spectrum between m/z 70-650 at 70,000 resolution, automatic gain control (AGC) target of 3e6, maximum injection time (IT) of 250 ms and intensity threshold of 3.8e4. MS/MS spectra were collected at 17,500 resolution, AGC target 5e5, maximum IT 65 ms and isolation window of 1 m/z. Normalized collision energy of 27 was used for both MS/MS and ddMS2 methods.

Data were analyzed and all theoretical masses were calculated with Xcalibur™ software. Full MS peak areas were normalized with a $^{13}C_6$ phenylalanine internal standard for positive ionization and $^{13}C_3$ pyruvic acid for negative ionization and compared across treatments. Compounds were identified using commercial standards when possible. When standards were unavailable, compounds were putatively identified by comparing the product ion spectra observed with those published in the literature and spectral databases (Smith et al., 2005).

2.5 Literature cited

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology 11:R106-R106.

Araujo, W.L., Martins, A.O., Fernie, A.R., and Tohge, T. (2014). 2-Oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. Fronteirs in Plant Science 5:552.

Bender, D.A. (2012). Amino acids synthesized from glutamate: glutamine, proline, ornithine, citruline and arginine. In: Amino acid metabolism West Sussex, UK: John Wiley and Sons, Ltd. 157-223.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57:289-300.

Boex-Fontvieille, E.R., Gauthier, P.P., Gilard, F., Hodges, M., and Tcherkez, G.G. (2013). A new anaplerotic respiratory pathway involving lysine biosynthesis in isocitrate dehydrogenase-deficient *Arabidopsis mutants*. New Phytol. 199:673-682.

Chen, C.Z., Lv, X.F., Li, J.Y., Yi, H.Y., and Gong, J.M. (2012). *Arabidopsis* NRT1.5 is another essential component in the regulation of nitrate reallocation and stress tolerance. Plant Physiol. 159:1582-1590.

Dechorgnat, J., Patrit, O., Krapp, A., Fagard, M., and Daniel-Vedele, F. (2012). Characterization of the *NRT2.6* gene in *Arabidopsis thaliana*: a link with plant response to biotic and abiotic stress. PloS One 7:e42491.

Fait, A., Fromm, H., Walter, D., Galili, G., and Fernie, A.R. (2008). Highway or byway: the metabolic role of the GABA shunt in plants. Trends in Plant Science 13:14-19.

Gaufichon, L., Rothstein, S.J., and Suzuki, A. (2016). Asparagine metabolic pathways in *Arabidopsis*. Plant Cell Physiology. 57:675-689.

Giavalisco, P., Li, Y., Matthes, A., Eckhardt, A., Hubberten, H.M., Hesse, H., Segu, S., Hummel, J., Kohl, K., and Willmitzer, L. (2011). Elemental formula annotation of polar and lipophilic metabolites using (13) C, (15) N and (34) S isotope labelling, in combination with high-resolution mass spectrometry. The Plant journal : for cell and molecular biology 68:364-376.

Gilliham, M., and Tyerman, S.D. (2016). Linking metabolism to membrane signaling: the GABA-malate connection. Trends in Plant Science 21:295-301.

Gutiérrez, R.A., Lejay, L.V., Dean, A., Chiaromonte, F., Shasha, D.E., and Coruzzi, G.M. (2007). Qualitative network models and genome-wide expression data define carbon/ nitrogen-responsive molecular machines in *Arabidopsis*. Genome Biology 8.

Gutiérrez, R.A., Stokes, T.L., Thum, K., Xu, X., Obertello, M., Katari, M.S., Tanurdzic, M., Dean, A., Nero, D.C., McClung, C.R., et al. (2008). Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene *CCA1*. Proc Natl Acad Sci U S A 105:4939-4944.

Hanson, A.D., Pribat, A., and de Creécy-Lagard, V. (2010). 'Unknown' proteins and 'orphans' enzymes: The mising half of the engineering part list - and how to find it. Biochemistry Journal. 425:1-11.

Hirai, M.Y., Yano, M., Goodenowe, D.B., Kanaya, S., Kimura, T., Awazuhara, M., Arita, M., Fujiwara, T., and Saito, K. (2004). Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A 101:10205-10210.

Hodges, M. (2002). Enzyme redundancy and the importance of 2-oxoglutarate in plant ammonium assimilation. Journal of Experimental Botany 53:905-916.

Hoekenga, O.A., Maron, L.G., Piñeros, M.A., Cançado, G.M.A., Shaff, J., Kobayashi, Y., Ryan, P.R., Dong, B., Delhaize, E., Sasaki, T., et al. (2006). *AtALMT1*, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in *Arabidopsis*. Proceedings of the National Academy of Sciences 103:9738-9743.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols 4:44-57.

Iuchi, S., Koyama, H., Iuchi, A., Kobayashi, Y., Kitabayashi, S., Kobayashi, Y., Ikka, T., Hirayama, T., Shinozaki, K., and Kobayashi, M. (2007). Zinc finger protein *STOP1* is critical for proton tolerance in Arabidopsis and coregulates a key gene in aluminum tolerance. Proceedings of the National Academy of Sciences 104:9900-9905.

Ivanov, A., Kameka, A., Pajak, A., Bruneau, L., Beyaert, R., Hernández-Sebastià, C., and Marsolais, F. (2012). *Arabidopsis* mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. Amino Acids 42:2307-2318.

Kobayashi, Y., Ohyama, Y., Kobayashi, Y., Ito, H., Iuchi, S., Fujita, M., Zhao, C.-R., Tanveer, T., Ganesan, M., Kobayashi, M., et al. (2014). *STOP2* activates transcription of several genes for al- and low ph-tolerance that are regulated by *STOP1* in *Arabidopsis*. Molecular Plant 7:311-322.

Lea, P.J., and Ireland, R.J. (1999). Nitrogen metabolism in higher plants. In: Plant Amino Acids: Biochemistry and Biotechnology--Singh, B.K., ed. New York: Marcel Dekker. 1-47.

Lea, P.J., Sodek, L., Parry, M.A.J., Shewry, P.R., and Halford, N.G. (2007). Asparagine in plants. Annals of Applied Biology 150:1-26.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26:589-595.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, G.P.D.P. (2009). The Sequence Alignment/Map format and SAM tools. Bioinformatics 25:2078-2079.

Li, J.Y., Fu, Y.L., Pike, S.M., Bao, J., Tian, W., Zhang, Y., Chen, C.Z., Li, H.M., Huang, J., Li, L.G., et al. (2010). The Arabidopsis nitrate transporter *NRT1.8* functions in nitrate removal from the xylem sap and mediates cadmium tolerance. Plant Cell 22:1633-1646.

Mi, H., Poudel, S., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2016). PANTHER version 10: expanded protein families and functions, and analysis tools. Nucleic Acids Research 44:D336-D342.

Miyashita, Y., and Good, A.G. (2008). NAD(H)-dependent glutamate dehydrogenase is essential for the survival of *Arabidopsis thaliana* during dark-induced carbon starvation. Journal of Experimental Botany 59:667-680.

Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirota, M., and Kinoshita, K. (2014). ATTED-II in 2014: evaluation of gene coexpression in agriculturally important plants. Plant Cell Physiology 55:e6.

Overbeek, R., Olson, R., Pusch, G.D., Olsen, G.J., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Parrello, B., Shukla, M., et al. (2013). The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Research.

Peng, M., Bi, Y.M., Zhu, T., and Rothstein, S.J. (2007). Genome-wide analysis of *Arabidopsis* responsive transcriptome to nitrogen limitation and its regulation by the ubiquitin ligase gene NLA. Plant Mol Biol 65:775-797.

Scheible, W.R., Morcuende, R., Czechowski, T., Fritz, C., Osuna, D., Palacios-Rojas, N., Schindelasch, D., Thimm, O., Udvardi, M.K., and Stitt, M. (2004). Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of Arabidopsis in response to nitrogen. Plant Physiology 136:2483-2499.

Smith, C.A., Maille, G.O., Want, E.J., Qin, C., Trauger, S.A., Brandon, T.R., Custodio, D.E., Abagyan, R., and Siuzdak, G. (2005). METLIN: A metabolite mass spectral database. Therapeutic Drug Monitoring 27:747-751.

Tzin, V., and Galili, G. (2010). The biosynthetic pathways for shikimate and aromatic amino acids in *Arabidopsis thaliana*. The *Arabidopsis* Book / American Society of Plant Biologists 8:e0132.

Wang, R., Okamoto, M., Xing, X., and Crawford, N.M. (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiology 132:556-567.

Wang, R., Tischner, R., Gutierrez, R.A., Hoffman, M., Xing, X., Chen, M., Coruzzi, G., and Crawford, N.M. (2004). Genomic analysis of the nitrate response using a nitrate reductase-null mutant of *Arabidopsis*. Plant Physiology 136:2512-2522.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "electronic fluorescent pictograph" browser for exploring and analyzing large-scale biological data sets. PLoS One 2:e718.

Winter, G., Todd, C.D., Trovato, M., Forlani, G., and Funck, D. (2015). Physiological implications of arginine metabolism in plants. Fronteirs in Plant Science 6:534.

Zhu, H., and Kranz, R.G. (2012). A nitrogen-regulated glutamine amidotransferase (*GAT1_2.1*) represses shoot branching in arabidopsis. Plant physiology 160:1770-1780.

3 THE ROLE OF A CLASS I GLUTAMINE AMIDOTRANSFERASE IN C/N PARTITIONING IN ROOTS OF *ARABIDOPSIS THALIANA*

3.1 Introduction

The maintenance of a proper balance between carbon and nitrogen metabolism is a requirement for the sustained growth of organisms. From a physiological perspective, C-metabolites derived from photosynthesis and respiration accumulate during the light period and serve as skeletons for the assimilation of N metabolites (amino acids) via the action of the glutamine synthetase/glutamate synthase (GS/GOGAT) cycle in leaf tissue (Stitt et al., 2002). A major interaction point between C and N metabolism is the synthesis of Glu from Gln and 2-oxoglutarate (2-OG), utilizing ammonium, via the GS/GOGAT cycle in chloroplasts (Stitt, 1999). Glutamate serves as a precursor for the synthesis of most amino acids downstream (Forde and Lea, 2007). Furthermore, the NADH and ATP produced via respiration are utilized by NR and GS reactions, respectively, where NH_4^+ and Gln are synthesized (Stitt et al., 2002). Hence, N assimilation depends on mitochondrial respiration and the production of 2-OG which provides carbon skeleton for Glu synthesis. These and several other lines of evidence (Coruzzi, 2003; Noctor et al., 2004) emphasize the tight interlinks between C produced via photosynthesis and glycolysis, and the regulation of N status and amino acid pools in relation to light and dark cycles.

The GS/GOGAT cycle, however, is more pronounced in leaves where there is a continuous supply of C skeletons via the action of photosynthesis and glycolysis. The absence of photosynthesis in roots calls for a different mechanism to achieve C/N balance. A transcriptome analysis of the roots of *Arabidopsis* seedlings treated with or without asparagine (Asn), another organic source of N, revealed a 4-fold up-regulation of a gene coding for a class I glutamine amidotransferase, *GAT1_2.1* (At1G15040) (Section 2.4.4). *GAT1_2.1* was also previously suggested to be highly responsive to N status in *Arabidopsis* with a 50-fold repression under long term and 4- to 6-fold repression under short term N limitation (Zhu and Kranz, 2012). We propose a potential role of *GAT1_2.1*

in achieving the aforementioned C/N balance in roots of *Arabidopsis* in addition to the GS/GOGAT cycle. Here, we hypothesize that GAT1_2.1 functions as a glutaminase and catalyses the hydrolysis of Gln to Glu releasing NH_3.

Hydrolysis of Gln in non-photosynthetic organisms, especially in mammalian liver and kidney, is mostly reported to be catalysed by glutaminases for the production of NH_3 (Bender, 2012). Glutaminase was also reported to play a major role in regulation of cancer cell metabolism to compensate for changes in glycolytic cycle (Warburg effect) by elevated glutamine metabolism (Erickson and Cerione, 2010). No known glutaminases are reported in plants to date. However, several proteins belong to the Gln amidotransferase superfamily, for example, a pyridoxal 5`-phosphate (PLP) synthase in *Arabidopsis* is reported to have a Gln amidotransferase domain (PDX2) and a synthase domain (PDX1) and is involved in biosynthesis of vitamin B6 in plants (Tambasco-Studart et al., 2007). The NH_3 produced by hydrolysis of Gln via the PDX2 domain is tunnelled to the PDX1 domain which catalyses the formation of PLP (vitamin B6) using ribulose 5-phosphate (Rub-5-P) or a combination of Rub-5-P and glyceraldehyde 3-phosphate (Tambasco-Studart et al., 2007). Here, we provide evidence for the absence of an acceptor substrate for NH_3 and highlight a possible role of *GAT1_2.1* in primary metabolism.

3.2 Results

3.2.1 Clustering of class I glutamine amidotransferase domain containing proteins

The GATase1 super family in *Arabidopsis* consists of 30 genes that were identified through BLAST searches. The genes in this superfamily have a highly conserved Cys-His-Glu triad that is characteristic of the GATase1 domain (approximately 250 residues) and is involved in removal of the amide group from glutamine. The ammonia released is known to act as a nucleophile on several acceptors (Mouilleron and Golinelli-Pimpaneau, 2007). The GATase1 superfamily is further clustered into six subfamilies and six individual genes based on phylogenetic analysis (Figure 3.1). Supporting information obtained from bioinformatics programs, as described in Section 3.2.1, was also used for functional categorization of GATase1 gene family members. The intron/exon structures of the genes, size of proteins, and predicted localization is highly conserved among subfamilies. The genes that clustered in anthranilate synthase and GATase1-1 subfamilies are the smallest with highly conserved exon/intron structures and protein sizes between 24-29 kDa and only contained the GATase1 domain. The subfamilies with GATase DJ-1 and GATase Pfp-1 consisted of proteins that are 41-51 kDa in size and contain an N-terminal GATase1 domain for glutamine hydrolysis and a C-terminal domain for a potential unknown acceptor of the amide group (Mouilleron and Golinelli-Pimpaneau, 2007). The CTP synthase subgroup consisted of proteins that are 60-67 kDa in size with an N-terminal synthase domain and a C-terminal GATase1 domain and are predicted to function in the addition of an amino group to UTP to synthesize CTP (Endrizzi et al., 2004).

The subgroup GATase1-2 consists of three members and is unique to the plant kingdom. However, *At1g15040* is the only member that, when mutated, resulted in a shoot branching phenotype (discussed in section 3.4) indicating that the function of this protein is not redundant to the other two (*At5g38200* and *At1g66860*) (Zhu and Kranz, 2012). A BLASTp search of the C-terminal extension of this protein only resulted in orthologous sequences in the plant kingdom with > 60% identity and a few algal species, from which land plants evolved (Lewis and McCourt, 2004), with <60% identity further indicating

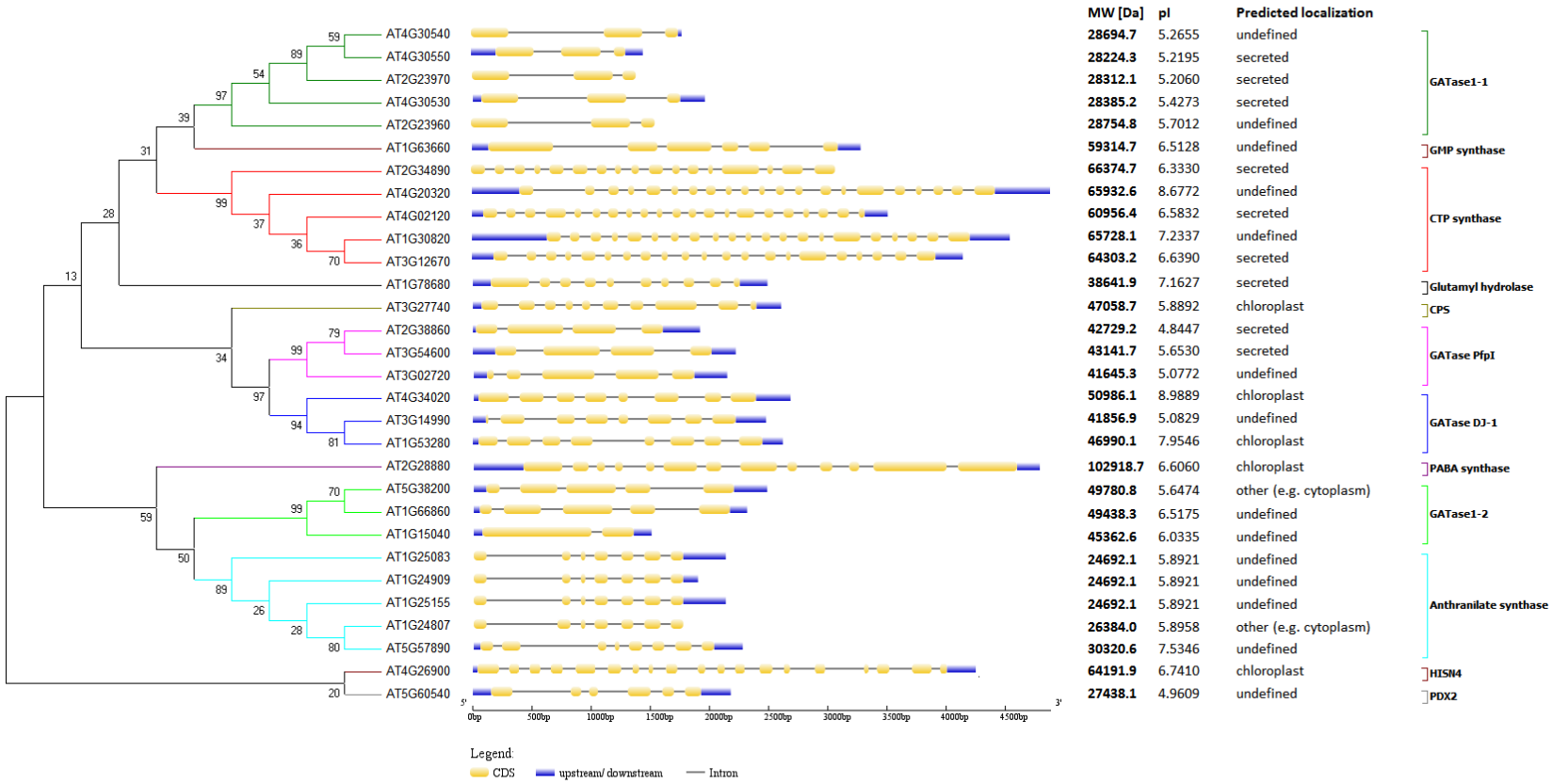


Figure 3.1: GATase1 gene family in *Arabidopsis*.

GATase1 superfamily in *Arabidopsis*. Figure depicts a neighbor joining tree with 1000 bootstrap replicates, exon/intron structures, predicted molecular mass, pI of the protein and their predicted localizations.

that this protein is a plant specific glutamine amidotransferase.

3.2.2 Glutaminase activity of the GAT1_2.1 domain

The full length version of the recombinant protein, GAT1_2.1, was found to be insoluble, as determined by SDS-PAGE, and hence the functional characterization was carried out using the GATase domain coding region (282 amino acids) of the protein (Figure 3.2A). To determine whether GAT1_2.1 has glutaminase catalytic activity, an N-terminal histidine tagged recombinant protein was expressed in *Escherichia coli* and purified using affinity chromatography using a Ni-sepharose column followed by size exclusion chromatography (SEC) to remove further impurities. The molecular mass of the protein was determined to be 46 kDa based on comparison with size standards that were separated using SEC (Figure 3.2B). SDS-PAGE was performed to determine the purity of the recombinant protein (Figure 3.2C). Enzymatic assays were performed with glutamine as substrate, using the auxiliary enzyme, glutamate dehydrogenase (Figure 3.2D) at pH 7.5. The activity was linear with increasing concentrations of enzyme (0-10 µg) (Figure 3.3A) and varying incubation times of primary reaction (1-60 min) (Figure 3.3B). A mutation of the active residue cysteine (C134S) resulted in the complete loss of function of the glutaminase function of the recombinant protein (Figure 3.3C). The apparent K_m of the GATase1 domain of GAT1_2.1 was determined to be 3.22 ± 0.47 mM (n=3) and the V_{max} was 0.27 ± 0.007 nmol sec^{-1} mg^{-1} (n=3). The k_{cat} was calculated to be 9.5 x 10^3 sec^{-1}.

3.2.3 Rescue of the growth phenotype of *E.Coli* glutaminase mutant with GAT1_2.1 domain

The *E. coli* glutaminase deletion mutant Δ*YneH* showed a reduced growth rate in M9 medium with glutamine as the sole source of carbon and nitrogen, when compared with the wild-type strain, BW25113 (Figure 3.4A) (Brown et al., 2008). This growth phenotype was rescued when transformed with an expression plasmid containing the GATase1 domain of GAT1_2.1 protein, further confirming that GATase1 has a glutaminase function (Figure 3.4B).

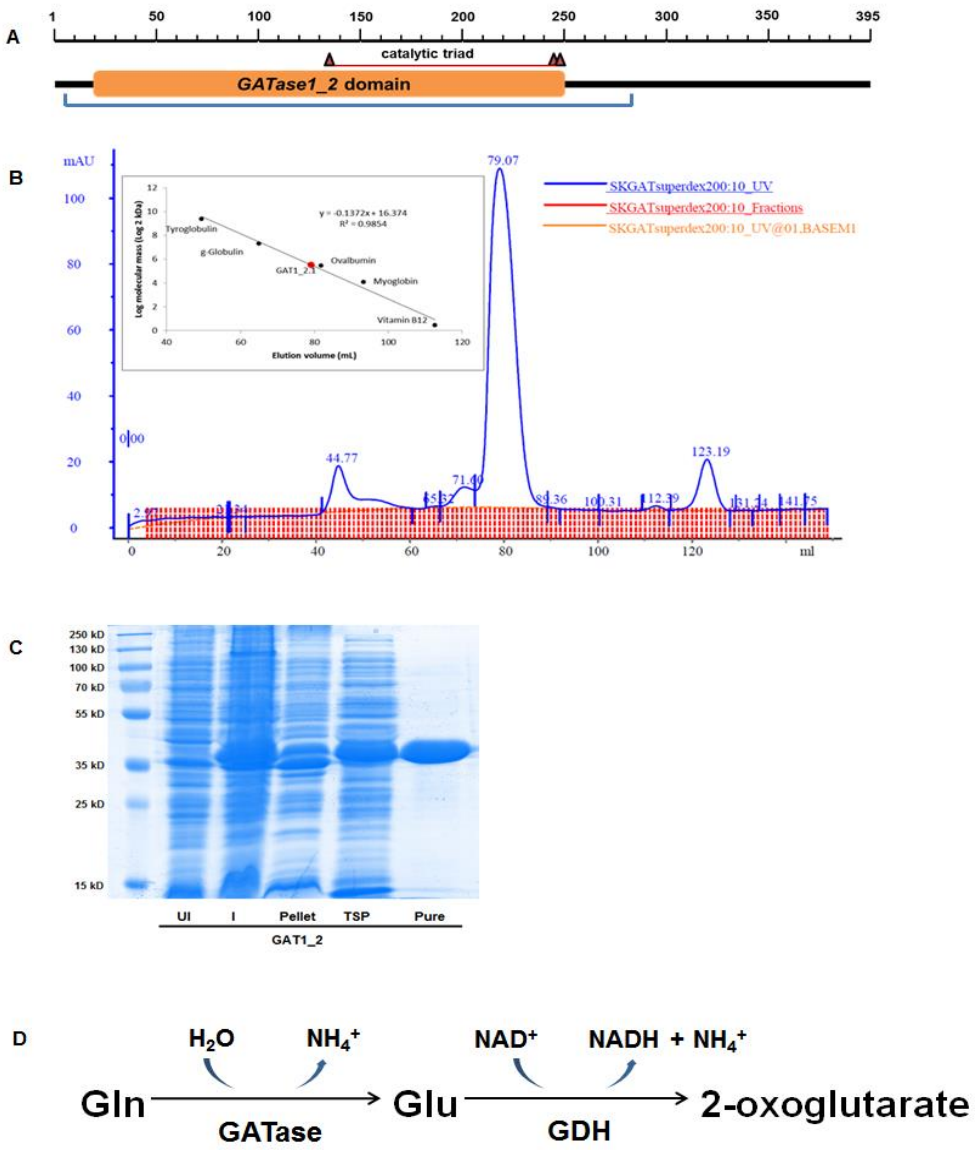


Figure 3.2: Recombinant protein purification and enzyme assay.

A. Schematic representation of GAT1_2.1 protein depicting the GATase1 domain (orange) and the truncated version used for recombinant protein purification. B. Size exclusion chromatography of GAT1_2.1 and the standard curve used for molecular weight determination. C. SDS-PAGE showing the quality of pure protein. D. Representation of the primary and auxiliary reactions used for glutaminase assay.

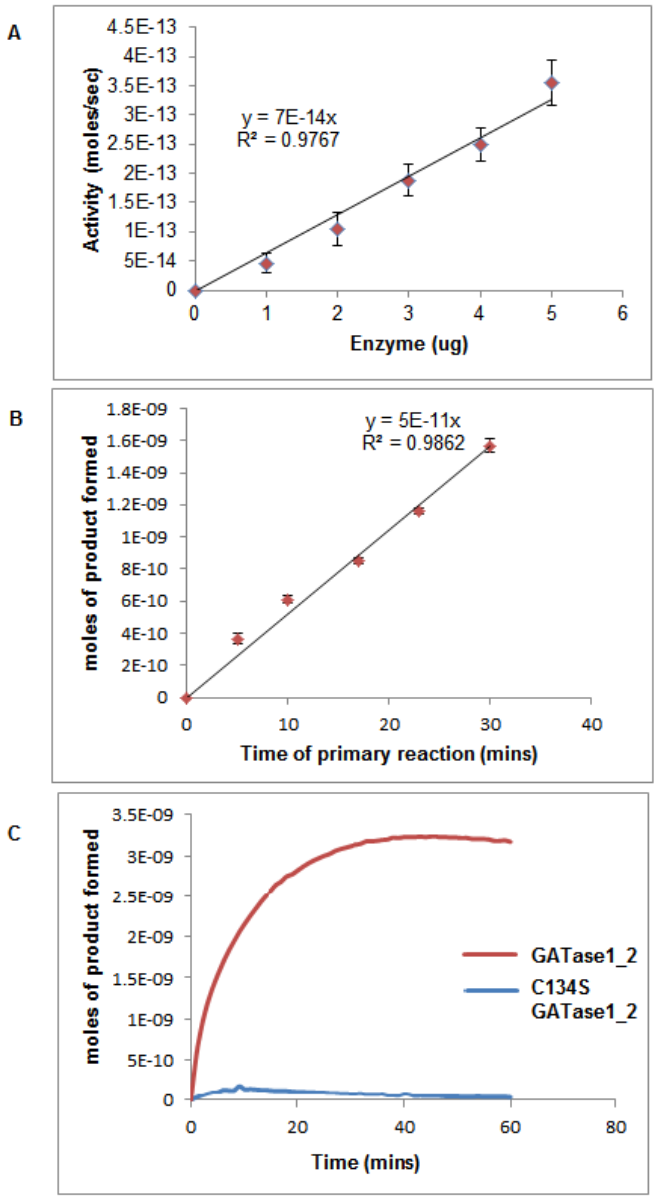


Figure 3.3: Enzyme kinetics.

A. Linear fit showing average of 3 replicates used for increasing concentrations of enzyme from 0-5 µg. 3 µg enzyme was used for kinetic measurements, B. Linear fit average of three replicates used for increasing incubation times of primary reaction from 0-60 mins. 30 mins of incubation time was used for the primary reaction. C. accumulation of product over time as monitored by measuring NADH, in the secondary reaction, at 340 nm. Active site mutant version (C134S) of the enzyme shows complete lack of function.

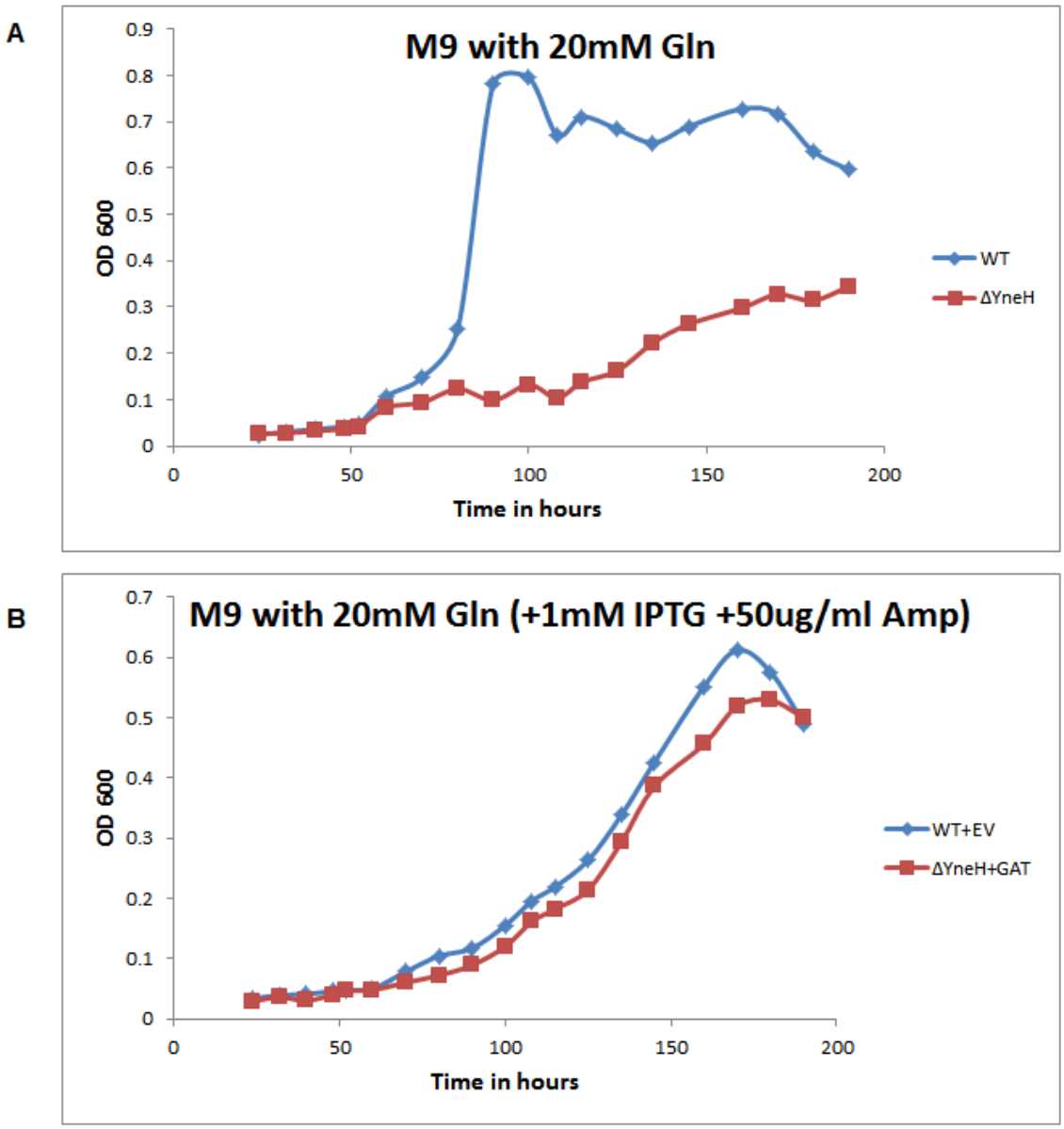


Figure 3.4: Rescue of *E. coli* glutaminase mutant.

E. coli growth curve using A. WT (BW25113) and Δ*YneH* mutant strains showing a reduced rate of growth in the mutant when grown in M9 medium with 20 mM Gln as a sole source of carbon and nitrogen. B. Growth curve with WT (BW25113) transformed with empty pET23a vector as control and the mutant Δ*YneH* transformed with GAT1_2.1 domain expressed in pET23a, showing a phenotypic rescue when grown in M9 medium with 20 mM Gln as carbon and nitrogen source with 1 mM IPTG to induce the expression and 50 µg/mL ampicillin for antibiotic selection.

3.2.4 Sub-cellular localization of GAT1_2.1

A translational fusion of full length GAT1_2.1 and a YFP reporter gene expressed transiently in epidermal cells of *Nicotiana Benthamiana* leaf was used to determine the sub-cellular localization of the protein. Co-localization of GAT1_2.1 with a known CFP tagged mitochondrial marker protein (*COX1*) confirmed that GAT1_2.1 was localized in mitochondria (Figure 3.5).

3.2.5 Metabolic profiling of the *gat1_2.1* mutant

To identify the metabolic responses of GAT1_2.1 to glutamine in roots, a targeted LC-MS/MS method on a Q-Exactive Orbitrap mass spectrometer, was performed and the changes in amino acids and organic acids were quantified. Wild-type *Arabidopsis* and *gat1_2.1* plants were grown on 5 mM KNO_3 and were used as an upstream pathway control, while the ones transferred to 10 mM Gln for 2 hours were used to identify responses to short term Gln treatment. Both genotypes grown in 2 mM Gln represented *GAT1_2.1* responses to long term Gln while the ones grown in 2 mM Glu represented a downstream pathway control. No significant difference in the accumulation of Gln was identified among the two genotypes, potentially owing to the continuous supply of Gln from the media (Figure 3.6). As expected, a significant decrease in the accumulation of Glu was observed in the mutant upon both short and long term Gln treatments (Figure 3.6), while no difference was observed in Glu treatment. Furthermore, the NH_3 content, as measured with HPLC showed a significant over-accumulation in both KNO_3 and short term Gln but not Glu treatment, in WT compared to *gat1_2.1* (Figure 3.6) suggesting that GAT1_2.1 may indeed act as a glutaminase, hydrolyzing Gln to Glu and releasing NH_3. However, no significant difference was observed in long term Gln treatment and this was predicted to be due to the fact that high levels of NH_3 are toxic and hence alternate measures of reducing excess NH_3 in WT were being employed. 2-Oxoglutarate (2-OG) was found to be significantly lower in *gat1_2.1*, in both short and long term Gln treatments, indicating that the supply of carbon skeletons from Glu to 2-OG was affected in the mutant. γ-Aminobutyric acid (GABA) and succinic acid were two other

compounds that showed significantly lower and higher accumulation in the mutant, respectively, upon Gln treatment, suggesting the activation of GABA shunt pathway.

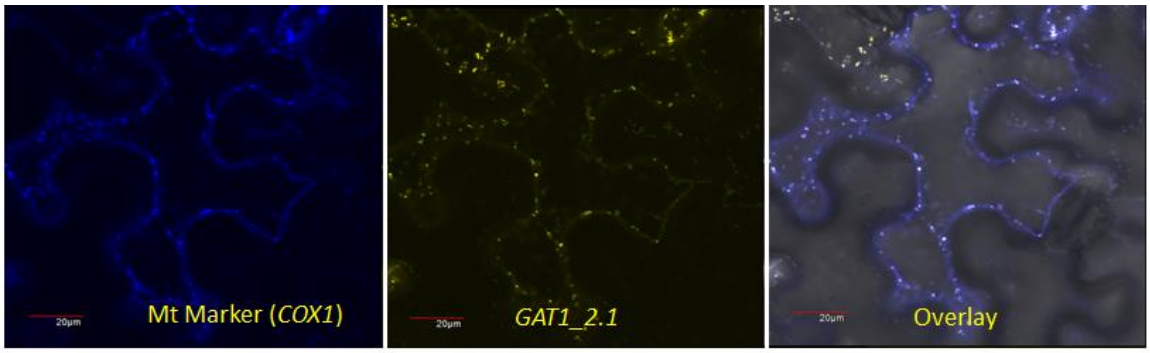


Figure 3.5: Subcellular localization of GAT1_2.1.

Full length GAT1_2.1 was translationally fused upstream of the reporter gene YFP, transformed into *N. benthamiana* by *A. tumefaciens* mediated transformation and visualized in the leaf epidermal cells by confocal laser-scanning microscopy. A CFP-tagged mitochondrial marker was co-expressed with GAT1_2.1. Left pane shows CFP tagged mitochondrial marker (COX1) middle pane shows YFP tagged GAT1_2.1, and right pane shows co-localization of *GAT1_2.1* with *COX*-CFP. Scale bar indicates 20 μm. YFP: Yellow fluorescent protein; CFP, cyan fluorescent protein.

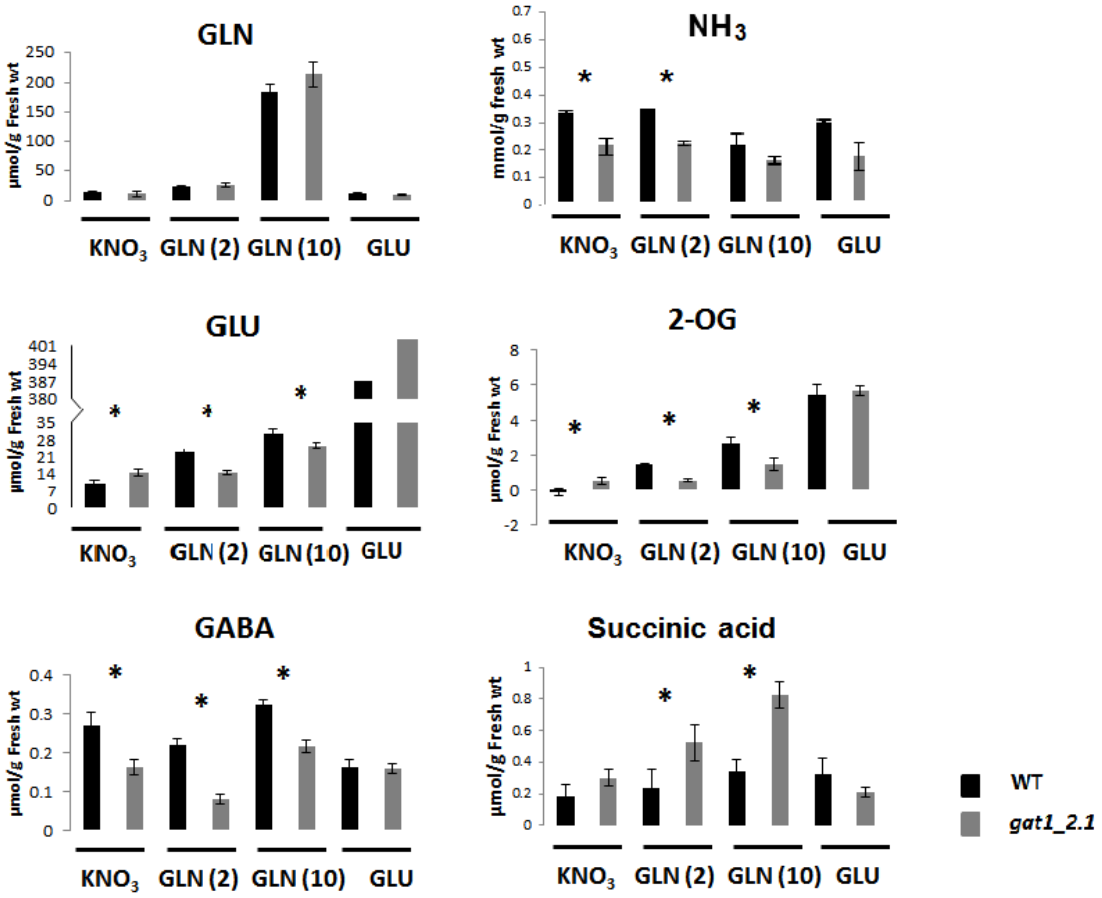


Figure 3.6: Metabolite analysis using wild-type *Arabidopsis* and *gat1_2.1*.

Metabolite quantities shown in µmoles/g fresh wt of tissue used. X-axis represents the treatment. KNO_3 : grown in 5 mM KNO_3 for 10 days, GLN (2) : grown in 5 mM KNO_3 followed by treatment with 10 mM Gln for 2 hours, GLN (10) : grown in 2 mM Gln for 10 days and GLU : grown in 2 mM Glu for 10 days. * represents significant difference ($p < 0.05$) between the two genotypes as assessed using students *t-test*. Error bars indicate standard error ($n = 3$, biological replicates).

3.2.6 GAT1_2.1 may not have an acceptor for NH_3 upon glutamine hydrolysis

In order to identify the presence of a possible acceptor substrate for the amide group of Gln upon Gln hydrolysis, a C-terminal *c-myc* tagged, full length GAT1_2.1 recombinant protein was expressed and purified from *N. benthamiana* using anti *c-myc* antibody (Figure 3.7A). No kinetic measurements were obtained using this protein since the yield of pure protein was limited to 5 µg per 10 grams of leaf tissue. Total *Arabidopsis* root extracts were spiked with the full length version of GAT1_2.1 as well as the GATase1 domain version of protein described in section 3.3.2. A ^{15}N labelled Gln, with the label on the amide N, was also used to spike the plant extracts. After two hours of incubation with the recombinant protein and labelled substrate, the protein was filtered out, using a 3K micro centrifuge filter (Millipore, Billerica, MA), and an untargeted LC-ddMS2 method was performed on a Q-Exactive Orbitrap mass spectrometer, to identify the product that accepted the labelled amide group of Gln. No acceptor was identified, but rather the amide nitrogen was just removed from Gln by both full length and GATase1 domain versions of the protein, as indicated by NH_3 quantified from the samples using HPLC. Furthermore, the amount of Gln (spiked + internal) decreased in protein spiked extract, was equivalent to the amount of NH_3 accumulated indicating a mass balance (Figure 3.7B). This suggests that the C-terminal extension of GAT1_2.1 may not function as an acceptor domain for substrate binding, and that *GAT1_2.1* may just be a glutaminase that hydrolyses Gln releasing NH_3.

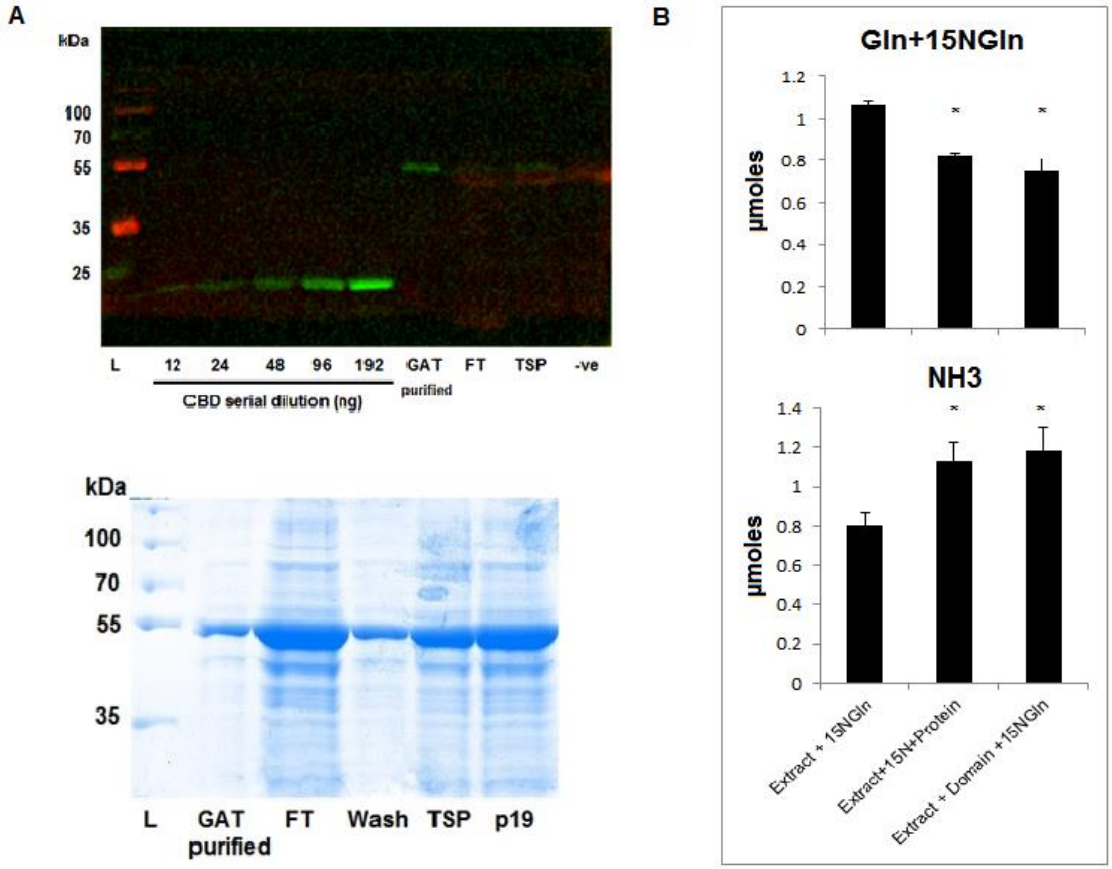


Figure 3.7: Recombinant protein production in *N. benthamiana* and enzyme spike assay.

A. Top pane shows the western blot using *anti*-c-myc antibody. The first five lanes represent serial dilution of CBD protein at known concentrations used for quantification of GAT1_2.1 purified from *N. benthamiana*. Subsequent lanes represent the pure protein, flow-through and total soluble protein used for purification followed by the TSP from negative control which was infiltrated with p19 alone. The bottom pane shows SDS-PAGE with indicated loadings used to assess the quality of purification. B. Top pane shows Gln+^{15}NGln quantity in µmoles indicating a ~30% decrease in accumulation in extracts spiked with both full length and domain versions of recombinant protein. Bottom pane represents NH_3 quantity in nmoles showing a ~30% increase. * represents a significant difference ($p < 0.05$, n = 3 biological replicates) in comparison with extract spiked with ^{15}N Gln alone, as assessed using students *t-test*.

3.3 Discussion

3.3.1 *GAT1_2.1* is unique to plants and has a root specific expression

As the name suggests, glutamine amidotransferases are enzymes that catalyze the removal of an amide group from Gln and transfer it to an acceptor substrate. Two classes of enzymes were identified, one that uses a Cys-His-Glu catalytic triad (Class I) for release of NH_3, and the other that uses an N-terminal cysteine residue (Class II). Collectively, GATases play a central role in metabolism as they are involved in incorporating nitrogen into amino acids, nucleotides, amino sugars, coenzymes, and antibiotics (Mouilleron and Golinelli-Pimpaneau, 2007). GAT1_2.1 belongs to class I GATases and unlike other members of this superfamily, *GAT1_2.1* and the other two members that clustered together (Figure 3.1) have a C-terminal extension that appears to only occur in the plant kingdom, although no functional role for this domain was recognized. Here, we propose a potential role of *GAT1_2.1* in glutamine hydrolysis itself, as a glutaminase release NH_3 instead of transferring it to an acceptor. No known glutaminases have been reported in plants to date, yet their importance in mammals is very well characterized (Erickson and Cerione, 2010).

The function of GAT1_2.1 has not been identified so far; however, it was implicated as a gene highly responsive to N status in plants and has a root specific expression (Arabidopsis eFP browser available in BAR database) (Winter et al., 2007). Zhu and Kranz (2012) reported that under N limiting conditions, the expression of *GAT1_2.1* is repressed 50-fold under long term stress (15 days) while microarray dataset from Hruz et al. (2008) with short term (2 hours) N limitation showed 4- to 6-fold downregulation. Our Asn transcriptomic data set revealed a 4-fold upregulation (section 2.2.6) of this transcript, under short term treatment (2 hours) suggesting that GAT1_2.1 responds to both inorganic and organic N.

3.3.2 The N-terminal GATase1 domain functions as a glutaminase

Although producing a full length version of recombinant GAT1_2.1 protein in *E. coli* was unsuccessful, a truncated version comprised of the GATase1 domain was purified and tested for glutaminase activity. The K_m (3.22 mM) and k_{cat} (9.5×10^3 s^{-1}) were found

to be within the range of biochemically characterized glutaminases from various organisms (Brown et al., 2008) (Brenda database; http://www.brenda-enzymes.org/). Glutaminases in general are known to have low affinity to Gln, however, they have a very high selectivity exhibiting no activity towards D-glutamine or L-asparagine which is also a characteristic of glutamine synthetase enzymes (EC 6.3.1.2, K_m for Glu up to 50 mM, Brenda database). Furthermore, the slow growth phenotype of *E. coli* glutaminase deletion strain, Δ*YneH*, was rescued using the same recombinant protein, suggesting that GAT1_2.1 indeed acts as a glutaminase (Figure 3.4).

GATases, however, are known to adopt mechanisms to avoid glutaminase activity in the absence of the acceptor substrate for NH_3 upon glutamine hydrolysis. For example, a tyrosine residue in the synthase domain of CTPS participates in the active site formation upon UTP binding to initiate GATase activity (Goto et al., 2004) and parts of the glutaminase site are disordered in anthranilate synthase of *Salmonella typhimurium* resorting to conformational changes and active site formation only upon substrate binding (Morollo and Eck, 2001). Since the C-terminus of GAT1_2.1 was removed in our experiment, we are unable to determine its function, or identify the presence of a different acceptor using the recombinant protein produced in *E. coli*. Hence, we expressed and purified a full length c-myc tagged recombinant GAT1_2.1 in *N. benthamiana*. The yield of recombinant protein in plant systems is limited and is not suitable for obtaining kinetic measurements; hence we resorted to a labelled Gln spiking approach to identify the presence of an alternate substrate. Plant extracts were spiked with Gln with ^{15}N labelled amide group along with the full length and truncated versions of recombinant protein followed by metabolomics analysis, and no acceptor for $^{15}NH_3$ was detected. Furthermore, upon measuring the NH_3 content in the spiked samples compared to control, we detected an equivalent amount of increase in NH_3 when compared to the amount of Gln decreased (Figure 3.7). This suggests that GATase may in fact be the plant glutaminase with no acceptor for NH_3. It is however, worth noting that the C-terminal extension contains two phosphorylation sites (S298 and S300), as determined using PhosPhAt 4.0 (Roitinger et al., 2015) that may regulate the activity or stability of GAT1_2.1.

3.3.3 *gat1_2.1* is non-responsive to Gln and has a shoot branching phenotype

To identify the physiological role of *GAT1_2.1* we first tested the Gln responsiveness of the T-DNA insertion mutant (SALK_031983) in comparison with wild-type *Arabidopsis*. 10 day old mutant plants grown with 2 mM Gln as sole N source did not show an enhanced growth compared to the ones grown with KNO_3 as N source, while the wild-type plants showed a robust growth in Gln medium (Figure 3.8A, B). This suggests that the Gln metabolism is impaired to a certain degree in the mutant indicating a predominant role of GAT1_2.1 in primary N metabolism in plants. The mutant also has an enhanced branching phenotype compared to the wild-type (Figure 3.8C). Zhu and Kranz (2012) showed that the phenotype was rescued upon stably transforming the mutant lines with GAT1_2.1 cDNA driven by a constitutive promoter, indicating that the gene is responsible for this phenotype. A potential role of GAT1_2.1 in strigolactone biosynthesis or presence of an unknown substrate involved in shoot branching was suggested by Zhu and Kranz (2012), however based on our biochemical characterization (Section 3.3.2), no such acceptor was identified. To that end, we resorted to a metabolite analysis approach to identify changes in primary metabolism in *gat1_2.1* upon Gln treatment that may be responsible for this branching phenotype. Since the gene is primarily expressed in the roots, we restricted this analysis to roots and hypothesized that the shoot branching phenotype is an indirect consequence of changes in metabolism in roots.

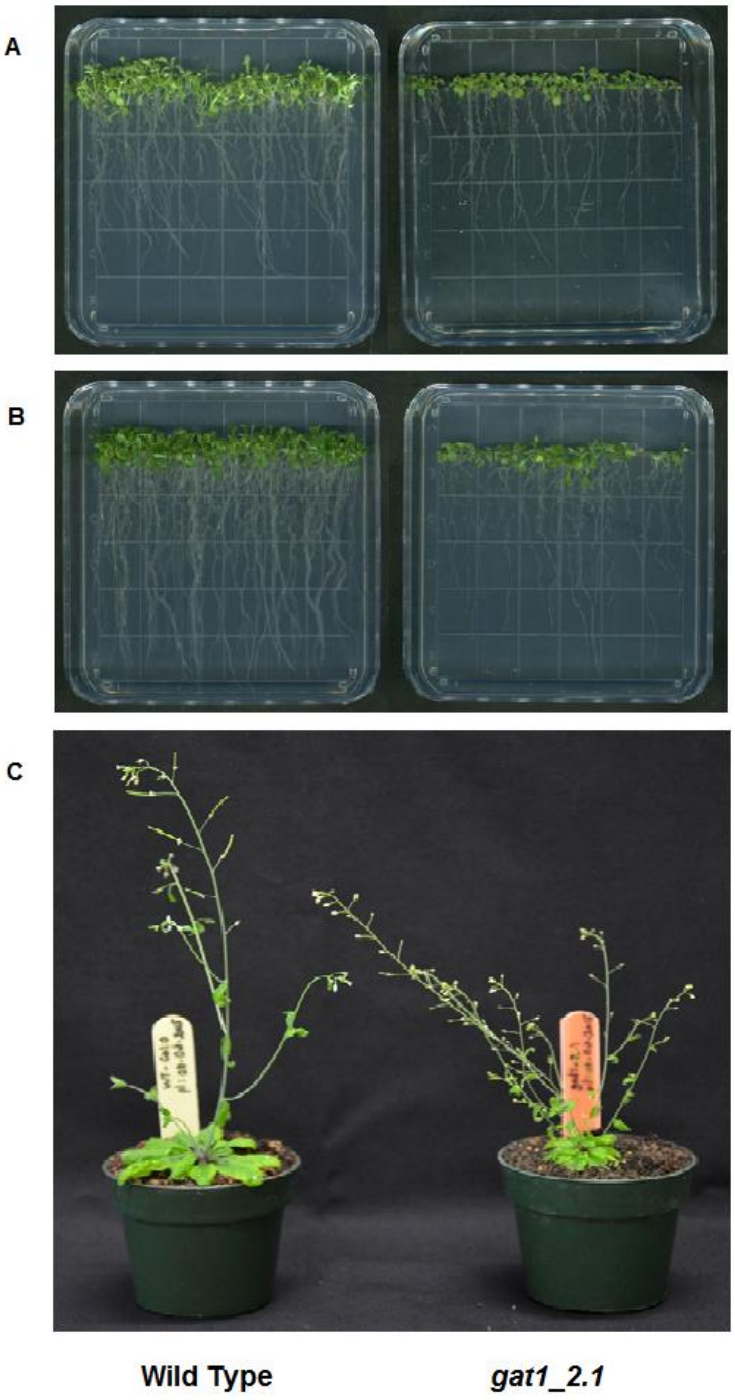


Figure 3.8: WT and *gat1_2.1* phenotypic analysis.

WT (left) and *gat1_2.1* seedlings (right) grown in A. 5 mM KNO3 or B. 2 mM Gln for 10 days C. WT and *gat1_2.1* after 7 weeks of growth.

3.3.4 *gat1_2.1* uses the GABA shunt pathway to provide carbon skeletons to the TCA cycle

To identify the Gln specific responses of GAT1_2.1, primary metabolites including amino acids and organic acids were quantified in roots of *gat1_2.1* and compared to the wild-type upon short term (2 hour) and long term (10 days) Gln treatments. Since, GAT1_2.1 has a glutamine hydrolysis function, we did not expect to see changes with Glu and hence both the genotypes treated with Glu (10 days) were used as control. Our results indicate a reduced accumulation of Glu in the mutants due to the absence of GAT1_2.1. Subsequently, lower 2-OG content was also detected in the mutant suggesting that the supply of carbon skeletons from Glu to the TCA cycle was impaired. An NAD(H)-dependent GDH, *GDH2*, localized in mitochondria was shown to be involved in this function using a *gdh2-1* mutant of *Arabidopsis* (Miyashita and Good, 2008). Interestingly, a co-expression search using atted-II (Version 7.1) (Obayashi et al., 2014), revealed that *GDH2* is the gene most highly co-expressed with *GAT1_2.1* in relation with tissue specificity and dependence on photoperiod. A GAT1_2.1 fused to YFP and transiently expressed in *Nicotiana benthamiana* leaves revealed a mitochondrial localization (Figure 3.5) suggesting a potential co-localization with GDH2 and a possible interlink between GAT1_2.1 and GDH2 function.

Furthermore, a decrease in accumulation of GABA and an increase in succinic acid levels were identified in the mutant compared to the wild-type upon Gln treatment. This suggests that in the absence of GAT1_2.1, when the supply of carbon from Glu to 2-OG was impaired, an alternate mode of C supply is activated via the GABA shunt pathway (Fait et al., 2008) when C from GABA is channeled through succinic acid into the TCA cycle (Figure 3.9). As such, no metabolites that are previously linked to a branching phenotype have been identified in our data and hence, this is yet to be determined. A thorough analysis of the changes using metabolomics is necessary to link GAT1_2.1 to the phenotype. Alternatively, efforts are underway to supply $^{13}C_5$, $^{15}N_2$ labelled Gln to track the distribution of C and N in the mutant in comparison with the wild-type with a goal of identifying the cause of the phenotype.

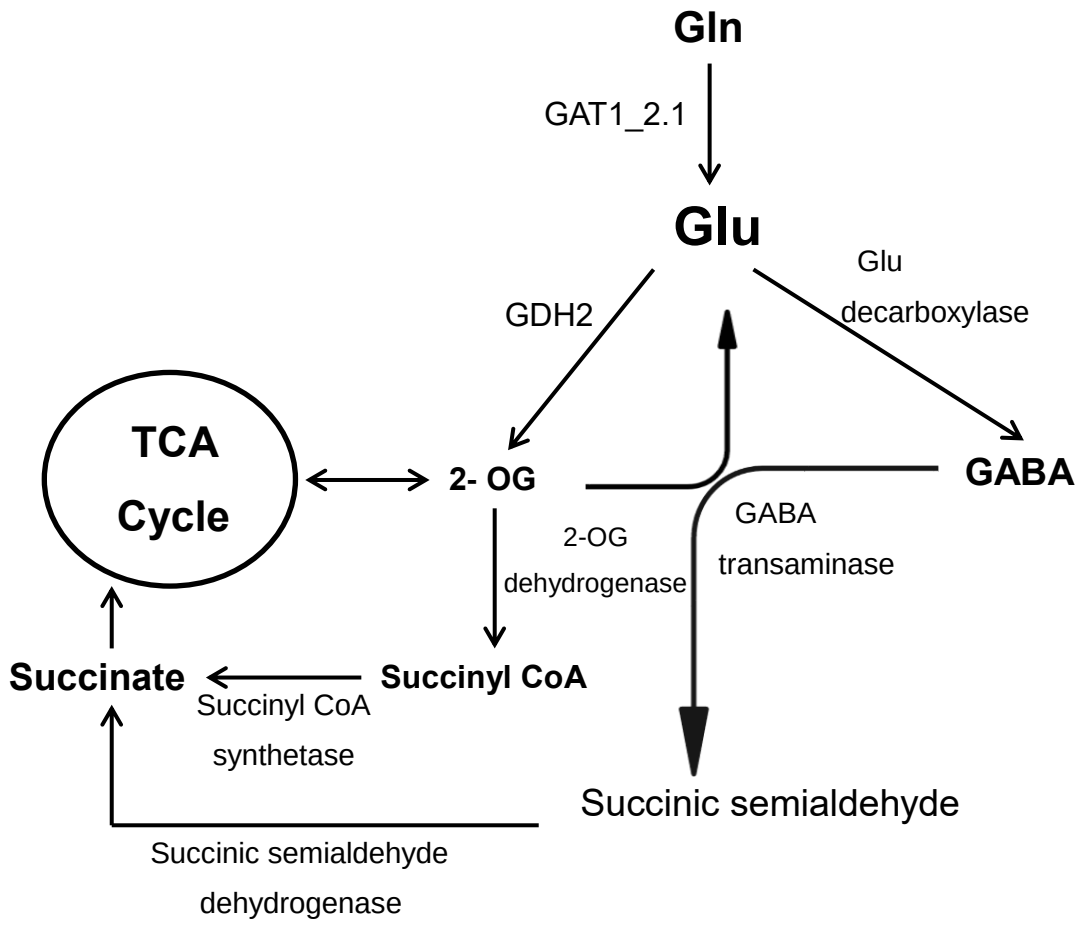


Figure 3.9: Schematic representation of the proposed model for C/N partitioning.

3.4 Concluding remarks

In this study, we characterized a plant specific class I glutamine amidotransferase, GAT1_2.1. GAT1_2.1 has a glutaminase activity and unlike other proteins in the GAT1 superfamily, no acceptor substrate was identified. Glutaminases in bacteria and mammals are well characterized and are known to play a major role in N assimilation. Since plants lack a functional glutaminase ortholog, it is possible that GAT1_2.1 performs this key function under organic N excess to distribute the carbon skeletons from Gln to TCA cycle intermediates. Despite the evidence provided for GAT1_2.1 to function as a glutaminase, the cause for a shoot branching phenotype in the null mutant is not fully understood. Future prospects include determining the importance of GAT1_2.1 in N assimilation, the N responsiveness of its promoter as well as the role of phosphorylation in N sensing. GAT1_2.1 holds potential for applications in enhancing Nitrogen Use Efficiency (NUE).

3.5 Materials and Methods

3.5.1 Phylogenetic analysis of class I glutamine amidotransferases

The full length protein sequence of GAT1_2.1 (At1g15040) was used as query for a BLASTp search against Arabidopsis TAIR 10 database (Altschul et al., 2005). The retrieved protein sequences were aligned using the ClustalW interface in MEGA program with a gap open penalty of 10 and gap extension penalty of 0.2 (Kumar et al., 2016; Thompson et al., 1994). Neighbor-joining was used to construct the phylogenetic tree. Pair-wise deletion was used to circumvent the gaps and Poisson correction distance was used to estimate the evolutionary distance. Bootstrap tests with 1000 bootstrap replicates were conducted to check the reliability of the clusters. The exon/intron structures of the genes were obtained using the Gene Structure Display Server (GSDS 2.0) (Hu et al., 2015). Molecular weight and pI of the proteins were collected from ExPASy -compute Mw/pI tool (http://web.expasy.org/compute_pi/).

3.5.2 Cloning of GAT1_2.1 for sub-cellular localization study

Full length cDNA of GAT1_2.1 was amplified from plasmid stock, C00162M obtained from ABRC, by PCR using attB1 and attB2 site-containing Gateway primers, cDNA-GATaatB1-F, 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTTGTCGCCAATGATCT-3' and cDNA-GATattB2-R, 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTCATAGTTGAGAAAAAAGGAGGACTC-3', following the cycling conditions described above. GAT1_2.1 amplicon was verified by electrophoretic separation and PCR purified using Gel/PCR DNA Fragments Extraction Kit (Geneaid Biotech Ltd, Taiwan) following manufacturers protocol. The purified fragment containing attB sites was recombined with the entry vector pDONR-Zeo (Invitrogen, USA) containing attP sites using BP clonase reaction mix (Invitrogen, USA) following Gateway™ cloning method. The BP reaction product was transformed into *E. coli* XL10 Gold cells (Agilent Technologies, Mississauga, ON) and grown on zeocin (25 µg/ml) supplemented LB plates. Colonies were screened by restriction digest plasmid isolation using High-Speed Plasmid Mini Kit (Geneaid Biotech Ltd, Taiwan) and

sequenced using M13 forward and reverse primers. After confirmation, an LR recombination reaction was performed using entry clone pDONR-Zeo-GAT1_2.1 containing attL sites and destination vector pEarleyGate101 (Earley et al., 2006) containing attR sites (Invitrogen, USA). *E. coli* strain XL10-Gold cells were transformed with destination clone. Cells were grown on kanamycin (50 µg/mL) supplemented LB medium and screened for positive clones by restriction digest. The plasmid DNA carrying the 'destination clone', pEG101-GAT1_2.1 was transformed into *Agrobacterium tumefaciens* strain GV3101 for infiltration.

3.5.3 Transient expression confocal microscopy

To monitor the transient expression of GAT1_2.1 fused to YFP, the plasmid (pEG101-GAT1_2.1) in *A. tumefaciens* strain GV3101 was transiently expressed into *N. benthamiana* leaf epidermal cells by infiltration (Sparkes et al., 2006). Briefly, a single colony was used to inoculate a medium (LB broth containing 10 mM 2-(N-morpholino) ethanesulfonic acid [MES] pH 5.6, and 100 µM acetosyringone) supplemented with kanamycin (50 µg/mL), rifampicin (25 µg/mL), and gentamycin (50 µg/mL) and grown at 28 ºC with shaking (250 rpm) until the OD600 reached 0.5-0.6. The culture was centrifuged at 3,000 g for 30 minutes at room temperature. The pellet was re-suspended in Gamborg's solution (3.2 g/L Gamborg's B5 medium with vitamins, 20 g/L sucrose, 10 mM MES pH 5.6, and 200 µM acetosyringone) to a final OD600 of 1 and incubated at room temperature for 1 h with gentle agitation to activate the virulence gene required for transformation. To verify subcellular localization of GAT1_2.1 a translationally fused mitochondrial protein (cytochrome oxidase with CFP) was mixed with pEG101-GAT1_2.1 in a 1:1 ratio and coexpressed in *N.benthamiana* leaves (Nelson et al., 2007). The leaves of 4-5-week-old *N. benthamiana* plants were transformed using a 1 mL syringe. Bacteria were slowly injected into the abaxial side of the leaf. The infiltrated leaves were labelled and plant growth was continued for 48 h.

Epidermal cell layers of the transfected region of *N. benthamiana* leaves were visualized using an OLYMPUS confocal microscope. A 60 X water immersion objective was used at excitation wavelengths of 514 and 458 nm, and emission spectra of 530-560 nm and 470-500 nm, for YFP and CFP respectively. The 'Sequential Scan Tool', which records

fluorescence in a sequential fashion, was used for studying co-localization of GAT1_2.1 with marker protein.

3.5.4 Cloning, expression and purification of recombinant GAT1_2.1 domain in *E. coli*

Coding sequence for the GATase domain (846 nt) was amplified from the cDNA obtained by the method described in section 3.5.2 using *Pfx* DNA polymerase (Life technologies, Burlington, ON). The primers used in the amplification were designed to introduce vector specific restriction sites, *KpnI* and *SalI* on the N- and C- terminal ends respectively (Fw: 5`-CTCGGTACCGTTGTCGCCAATGATCTCTCGTCC-C` and Rw: 5`- CAGGTCGACGTTTACTTGTTTCTTCTGAAATGCG-3`). The PCR product was cloned into the pSC-B-amp/kan vector as per the manufacturer's protocol (StrataClone Blunt PCR Cloning Kit, Agilent Technologies) and transformed in *E. coli*. Positive recombinant clones selected on ampicillin plates were confirmed by restriction digests and DNA sequencing using the M13 sequencing primers and verified against the TAIR 10 genome database. Confirmed plasmid DNA was digested with *KpnI* and *SalI* and subcloned in the bacterial expression vector pQE30 (Qiagen, Toronto, ON) by restriction enzyme-mediated cloning using *KpnI* and *SalI* restriction endonucleases and T4 DNA ligase following manufacturers' guidelines. The expression constructs were transformed in *E. coli* XL10-Gold (Agilent Technologies, Mississauga, ON). Alternatively, the same coding sequence was also cloned into pET23a (Novagen, USA) following the procedure described above with the exception of restriction sites introduced, *EcoRI* and *XhoI* on the N- and C- terminals respectively, using the primers, Fw: 5`-GCTTGAATTCCTATGGTTGTCGCCAATGATCTCTCG-3` and Rw: 5`-GTAGCTCGAGGTTTACTTGTTTCTTCTGAAATGCG-3`.

XL10-Gold cells with pQE30 vectors containing GAT1_2.1 domain coding sequence were grown in NZY media with ampicillin (100 µg/mL) as antibiotic at 37^{o}C. The cells were induced for expression by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) after reaching an OD600 of 0.6 and continued to grow at room temperature for 16 h. Cells were harvested by centrifugation (6,000 *g*, 30 min at 4^{o}C), suspended in lysis buffer (50 mM sodium phosphate, pH 8.0, 500 mM NaCl and 20 mM imidazole, 2 mM DTT,

pH 7.5) and treated with 1 mg/mL lysozyme for 30 min. Cells were lysed using French press, centrifuged at 18,000 *g* for 45 minutes at 4ºC and the supernatant was collected. This supernatant was purified using immobilized metal affinity chromatography on a 5 mL HisTrap Ni^{2+}-Sepharose column on an ÄKTApurifier system (GE Healthcare). Column was washed with five column volumes of wash buffer (50 mM sodium phosphate, 500 mM NaCl, 40 mM imidazole, 2 mM DTT, pH 7.5) and eluted with a linear gradient varying from 0 to 100% of elution buffer (50 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, 2 mM DTT, pH 7.5). Purified protein was concentrated using Amicon Ultra-15 Ultracel 30 K filter unit (Millipore, Billerica, MA) and buffer exchanged with 100 mM Tris-Cl pH 7.5, 50 mM NaCl and 2 mM DTT using a PD10 column (GE Healthcare) following manufacturer's protocol.

The purified protein was further separated using size exclusion chromatography on ÄKTApurifier system (GE Healthcare) using HiLoad Superdex 200 prep grade column with FPLC buffer (100 mM Tris-Cl pH 7.5, 50 mM NaCl and 2 mM DTT). Molecular weight of eluted protein was calculated based on a standard curve generated with Gel Filtration Standard (Bio-Rad). The protein quantification was performed using the Bio-Rad Protein Assay solution (Mississauga, ON), and bovine serum albumin (BSA) as standard. The yield of protein was determined to be 15 mg L^{-1} of culture volume. Purified protein was analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) using a 12% polyacrylamide gel, and the protein bands were visualized with Coomassie staining as described (Laemmli, 1970). The final protein product was aliquoted and stored at -80ºC in 20% glycerol (v/v) after flash freezing in liquid N_2.

3.5.5 Enzymatic assay

Glutaminase activity of GAT1_2.1 domain was assayed following the method described by Curthoys and Weiss (1974) with a few modifications. In brief, the primary reaction was performed in 50 µL of 50 mM HEPES pH 7.5 with 5 mM $MgCl_2$, 2 mM DTT and 5 mM Gln as substrate. Reactions were started with 3 µg of pure protein and incubated for 30 mins at 30ºC followed by termination with 3 N HCL. A volume of 150 µL of auxillary reaction mixture (100 mM Tris-Cl pH 9.6, 2 mM NAD, 0.4 mM ADP, 0.03% v/v H_2O_2 and 1 unit GDH) was then added and the production of NADH was monitored by

spectrophotometry at 340 nm with a Power Wave XS 96-well plate reader (Bio-Tek instruments, Winooski, VT). The path length was measured to correct the calculated enzyme activities using the path length correction feature in Gen5 software.

3.5.6 Rescue of *E. coli* glutaminase mutant *YneH*

The *E. coli* glutaminase deletion mutant Δ*YneH* with the kanamycin resistance cassette excised, and the corresponding wild-type (BW25113) strains were kindly provided by Dr. Alexander Yakunin at University of Toronto (Brown et al., 2008). The genotype of the strain was confirmed using the primers described by Brown et al. (2008). The mutant strain was transformed with GAT1_2.1 domain coding sequence in pET23a with ampicillin resistance cassette and the wild-type was transformed with an empty pET23a vector as control. Growth experiments with these strains were carried out as described in Brown et al. (2008). In detail, the bacterial strains, wild-type and Δ*YneH* were grown in LB media in 5 mL cultures to an OD600 of 0.6 and then expanded into M9 minimal medium (pH 7.0) with glutamine as the sole source of carbon and nitrogen at 1:1000 dilution and the OD600 was monitored. Alternatively, the strains transformed with plasmids were grown in the same manner with the exception of M9 medium being supplemented with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) for induction of gene expression and 100 μg/mL ampicillin for antibiotic selection.

3.5.7 Plant growth conditions

Wildtype *Arabidopsis* ecotype Columbia and T-DNA insertion lines were used for Gln and Glu treatments. Plants were grown on vertical plates at 22^{o}C under continuous light (ca. 70 μmol m^{-2} s^{-2}), as previously described by Ivanov et al. (2012) on a defined nutrient medium containing a final concentration of 10 mM potassium phosphate (pH 6.5), 5 mM KNO_3, 2 mM $MgSO_4$, 1 mM $CaCl_2$, 125 μg FeNaEDTA, micronutrients (50 mM H_3BO_3, 12 mM $MnSO_4$, 1 mM $ZnCl_2$, 1 mM $CuSO_4$ and 0.2 mM Na_2MoO_4), 1% sucrose and 1% agar (Wang et al., 2003). Ten-day old seedlings were transferred to plates containing the same medium without nitrogen as control or 10 mM Gln as sole N source. After 2 h, root tissue was harvested, frozen in liquid N_2 and stored at -80^{o}C until total metabolite extractions was carried out. For growth in Gln and Glu, the same media and growth

conditions were used with the exception of 5 mM KNO_3 being substituted with either 2 mM Gln or 2 mM Glu and tissue was collected after 10 days.

3.5.8 Plant material, genotyping and RT-PCR

Seeds of homozygous GAT1_2.1 mutants (SALK_031983) were obtained from the Arabidopsis Biological Resource Center, Ohio State University (http://www.arabidopsis.org/abrc/) (Sessions et al., 2002). To confirm homozygosity, gene specific and T-DNA specific primers were used for PCR amplification. One hundred mg of leaf tissue was used for extraction of genomic DNA using GenElute™ Plant Genomic DNA Miniprep Kit (Sigma-Aldrich) following manufacturer's protocol. Primers used for genotyping were as follows; G-At1g15040-F, 5'- CCAAGATTCTCCCCAGAGTTC-3', G-At1g15040-R, 5'-ACACATGAGTTCCTCACCGTC-3' and LBb1, 5'-GCGTGGACCGCTTGCTGCAACT-3'. The PCR reaction was performed using *Taq* DNA polymerase (Life Technologies) with the following conditions; denaturing at 95^{o}C for 2 min, 35 cycles at 95^{o}C for 2 min, 60^{o}C for 30 s and 68^{o}C for 60 s, and a final extension at 68^{o}C for 10 min. Transcript abundance in homozygous GAT1_2.1 mutants in reference to the WT was analyzed using RT-PCR. 100 mg of leaf tissue from WT and GAT1_2.1 mutants was used for RNA extraction using RNeasy Plant Mini Kit (Qiagen) following manufacturers protocol, quantified using NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Burlington, ON) and treated with amplification grade DNAse I (Life Technologies) to minimize DNA contamination. RNA quality was verified by electrophoretic separation on a 1% (w/v) agarose gel prior to cDNA synthesis which was performed using a Thermoscript RT-PCR system (Life Technologies) with 5 ug of total RNA in a 20 µl reaction. Following this, a PCR was carried out using 1 µl of cDNA as template in 25 µl of final volume using the cycling conditions described above and electrophoretically separated on a 1.5% (w/v) agarose gel (Figure 3.10).

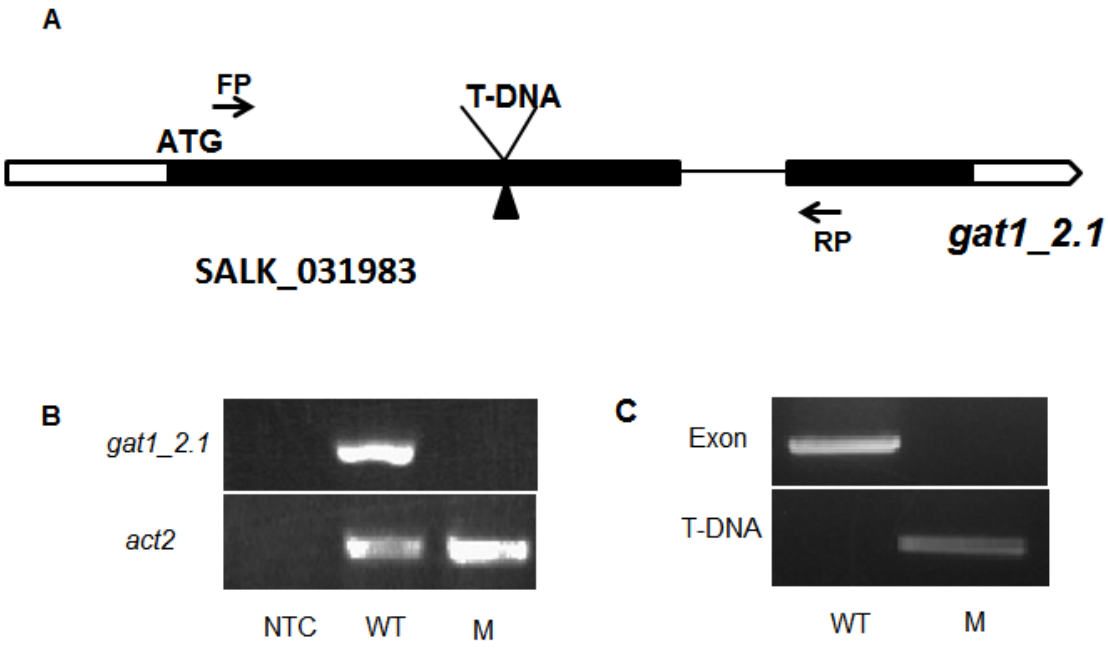


Figure 3.10: *gat1_2.1* T-DNA insertion line, RT-PCR and genotyping

A. Schematic representation of GAT1_2.1 showing the position of T-DNA insertion. Bars represent the exons and triangle indicated the site of T-DNA insertion. B. RT-PCR analysis of the T-DNA insertion line. Top panel shows the RT-PCR with primers spanning the exon region (FP and RP) indicating the absence of transcript. Bottom panel represents the control gene *act2*. C. Genotyping T-DNA insertion line. Top panel shows the PCR amplified product for exon and the bottom panel represents PCR amplified product for T-DNA (see section 3.5.8 for details).

3.5.9 Total metabolite extraction

Fifty mg of root tissue was excised from 10 day old seedlings of WT or *gat1_2.1* grown under conditions described above, collected in 2 ml Eppendorf tubes and flash frozen in liquid N_2. Frozen tissue was homogenized using a tissue lyser and metabolites were isolated using 1 ml of methanol: water (4:1) with incubation in an ultra-sonication bath for 20 min followed by shaking for 30 min at 4ºC. The mixture was centrifuged at 11,000 *g* for 10 min at 4ºC and 700 µl of the supernatant was transferred into fresh tubes and evaporated to dryness using a Vacufuge at ambient temperature. The residue was re-dissolved in 500 µl of 1:1 methanol: water and the samples were filtered using a 0.2 µm PTFE microfuge filter (Whatman). Five µl of 1 µg/ml $^{13}C_6$ Phe was added to the samples for monitoring the quality of LC-MS runs.

3.5.10 GAT1_2.1 assay with plant extracts

For assays with plant extracts, total metabolite extract from root tissue of wild-type *Arabidopsis* obtained from the method described in section 3.5.9 was re-suspended in HEPES buffer pH 7.5 instead of 1:1 methanol: water. Samples were spiked with a final concentration of 1 µM ^{15}N Gln and 5 µg of the full length (Section 3.5.14) or domain versions of GAT1_2.1 protein along with 2 mM DTT and 5 mM $MgCl_2$. Following this, samples were incubated at 37ºC for 2 hours and then filtered through a 3K micro centrifuge filter (Millipore, Billerica, MA) to remove the protein. Samples were then evaporated to dryness using a vacufuge at ambient temperature and the residue was re-dissolved in 1:1 methanol: water, filtered with a 0.2 µm PTFE microfuge filters (Whatman) and subjected to LC-MS analysis and ammonium quantification.

3.5.11 Data acquisition and metabolite analysis

MS data was obtained from four replicates of WT or *gat1_2.1* on a Q-Exactive Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an Agilent 1290 high performance liquid chromatography (HPLC) system. Compounds were resolved using a SeQuant® ZIC®-HILIC column; 3.5µm, 100 Å, 100 × 2.1 mm (EMD Millipore) with mobile phase 5 mM ammonium acetate, pH = 4.00 (A); 90% acetonitrile, 0.1% formic acid (B) and the following gradient: 87% B for 5 min, decreased to 55%

over 8 min and held for 4 minutes before returning to 87% over 3 min. The following heated electrospray ionization (HESI) conditions were optimized for the analysis of amino and organic acids: spray voltage, 3.9 kV (ESI+), 3.5 kV (ESI-); capillary temperature, 250 °C; probe heater temperature, 450 °C; sheath gas, 30 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 60%. Injections of 5 µl were used with a flow rate of 0.3 mL min^{-1}. Compounds were detected and monitored using targeted MS/MS, spectra were collected at 17,500 resolution, AGC target 1e6, maximum IT 65 ms, isolation window of 1 m/z, normalized collision energy of 30, intensity threshold of 1.6e5 and 10s dynamic exclusion. Data analysis and calculation of all theoretical masses was carried out using Xcalibur™ software. Compounds were identified and quantified using commercial standards.

3.5.12 Metabolomic data analysis

For metabolomics analysis, LC-MS data was obtained with the above mentioned chromatographic and HESI conditions but full MS measurements were collected from mass ranges of 75-1100 m/z and 65-900 m/z in positive and negative ionization modes respectively at 140,000 resolutions. The AGC target and maximum IT was set to 3 e6 and 524 ms respectively. Following data acquisition, Thermo .RAW files were converted to .mzml format using ProteoWizard (Kessner et al., 2008) with peak picking filter applied. Files were imported into R using the XCMS package (Smith et al., 2006) and features were detected using the centWave (Tautenhahn et al., 2008) method and a ppm tolerance of 1.0. Pre-filter was set to 6 scans with minimum 5000 intensity, signal to noise threshold was 5 and noise was set to 3×10^6 and 1×10^6 for positive and negative mode respectively. Retention time correction was conducted using the obiwarp (Prince and Marcotte, 2006) method, grouping included features present in at least 25% of all samples, allowable retention time deviation was 10 seconds, and *m/z* width set to 0.015. The “fillPeaks” function with default settings and remaining zeros imputed with two-thirds the minimum value on a per mass basis. Salt clusters and other ionization artefacts were removed from the feature list using the McMillan correction without retention time filter applied (McMillan et al., 2016).

3.5.13 Ammonium quantification

Fifty mg of root tissue was homogenized using a tissue lyser followed by the addition of 1 mL of ice cold extraction medium (10 mM HCOOH in 60% methanol). The homogenate was centrifuged at 16,000 rpm at 4ºC and the supernatant was transferred to a 0.45 µM polysulphone centrifuge tube. This was followed by centrifugation at 5000 *g* at 2ºC. OPA reagent diluted in borate buffer (1:5 v/v) containing 2-mercaptoethanol as reducing agent (Agilent Technologies), was added to the clarified supernatant at 1:1 ratio and placed at 63ºC for 15 min (Husted et al., 2000). NH_3 was detected using a fluorescence detector equipped on an Agilent 1260 high proficiency liquid chromatography (HPLC) system with step excitation wavelengths (10 nm increments) scanning from 220 – 300 nm and collecting the emission at a wavelength of 450 nm. A Zorbax Eclipse Plus RRHD C18 column (4.6 X 100 mm, 3.5 µm) was maintained at 40°C. Mobile phases were comprised of 40 mM NaH_2PO_4 pH 7.8 (A), ACN: MeOH: Water (45:45:10 v/v/v) (B) and MeOH: Water (90:10, v/v) (C). Mobile phase A was maintained at 100% for 1 min and mobile phase B was increased to 50% over next 4 min. Following this, mobile phase B was increased to 80% and C to 20% and maintained for 50 sec. The residual from the column was removed with 100% A over the next 2 min 10 sec. A flow rate of 2 mL min^{-1} was used.

3.5.14 Cloning, expression and purification of full length recombinant GAT1_2.1 in plant system

Full length cDNA of GAT1_2.1 was amplified by PCR (described above) primers designed to introduce Golden Gate cloning sequences at both 5` and 3` ends (Fw: 5`-CATATGGGTCTCCCAAGATGGTTGTCGCCAATGATCTC-3`, Rw: 5`-AAGCTTCGGTCTCTGCTCATAGTTGAGAAAAAAGGAGGAC-3`). The amplicon was cloned into pCamGate-Cyt expression vector (Conley et al., in prep) under the control of double-enhanced cauliflower mosaic virus (CaMV) 35S promoter and a *c-myc* tag sequence at the 3` end, modified for compatibility with Golden Gate system. Final expression vectors that were confirmed with restriction digest and sequencing were transformed into *A. tumefaciens* strain GV 3101. Transformed bacteria were grown to an OD600 of 0.6, and collected by centrifugation at 1,000 *g* for 30 min. The pellets were

resuspended in Agro-infiltration solution (3.2 g/L Gamborg's B5 medium and vitamins, 20 g/L sucrose, 10 mM MES pH 5.6, 200 µM 4'-Hydroxy-3'-dimethoxyacetophenone) to a final OD600 of 0.3, followed by incubation at room temperature with gentle agitation for 1 hour. The suspension was then used for infiltration of the abaxial leaf epidermis through the stomata of *N. benthamiana* plants with a 1 ml syringe (Kapila et al., 1997). Infiltrated plants were continued to grow and leaf tissue was collected after 4 days.

Total soluble protein was obtained from 5 mg of leaf tissue after homogenization with the extraction buffer (1X PBS/Tween-20, 2% PVPP, 1 mM EDTA pH 8.0, 1 mM PMSF, 1 µg/mL leupeptin and 100 mM sodium L-ascorbate) followed by centrifugation at 17,500 *g* at 4^{o}C. The recombinant GAT1_2.1 was purified using a c-myc tagged protein mild purification kit ver. 2 (MBL International) following manufacturers protocol. Quality of the pure protein was verified with SDS-PAGE and quantified with western blotting using *anti*-cmyc antibody and a serial dilution of CBD protein (Figure 3.7).

3.6 Literature cited

Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schäffer, A.A., and Yu, Y.-K. (2005). Protein database searches using compositionally adjusted substitution matrices. FEBS Journal 272:5101-5109.

Bender, D.A. (2012). Amino acids synthesized from glutamate: glutamine, proline, ornithine, citruline and arginine. In: Amino acid metabolism West Sussex, UK: John Wiley and Sons, Ltd. 157-223.

Brown, G., Singer, A., Proudfoot, M., Skarina, T., Kim, Y., Chang, C., Dementieva, I., Kuznetsova, E., Gonzalez, C.F., Joachimiak, A., et al. (2008). Functional and structural characterization of four glutaminases from *Escherichia coli* and *Bacillus subtilis*. Biochemistry 47:5724-5735.

Coruzzi, G.M. (2003). Primary N-assimilation into amino acids in *Arabidopsis*. The *Arabidopsis* book:e0010.

Curthoys, N.P., and Weiss, R.F. (1974). Regulation of renal ammoniagenesis: subcellular localization of rat kidney glutaminase isoenzymes. Journal of Biological Chemistry 249:3261-3266.

Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. The Plant Journal 45:616-629.

Endrizzi, J.A., Kim, H., Anderson, P.M., and Baldwin, E.P. (2004). Crystal structure of *Escherichia coli* cytidine triphosphate synthetase, a nucleotide-regulated glutamine amidotransferase/ATP-dependent amidoligase fusion protein and homologue of anticancer and antiparasitic drug targets. Biochemistry 43:6447-6463.

Erickson, J.W., and Cerione, R.A. (2010). Glutaminase: A hot spot for regulation of cancer cell metabolism? Oncotarget 1:734-740.

Fait, A., Fromm, H., Walter, D., Galili, G., and Fernie, A.R. (2008). Highway or byway: the metabolic role of the GABA shunt in plants. Trends Plant Sci 13:14-19.

Forde, B.G., and Lea, P.J. (2007). Glutamate in plants: metabolism, regulation, and signalling. Journal of Experimental Botany 58:2339-2358.

Goto, M., Omi, R., Nakagawa, N., Miyahara, I., and Hirotsu, K. (2004). Crystal structures of CTP synthetase reveal ATP, UTP, and glutamine binding sites. Structure 12:1413-1423.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., and Zimmermann, P. (2008). Genevestigator V3: A reference expression database for the meta-analysis of transcriptomes. Advances in Bioinformatics 2008:5.

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31:1296-1297.

Husted, S., Hebbern, C.A., Mattsson, M., and Schjoerring, J.K. (2000). A critical experimental evaluation of methods for determination of NH_4^+ in plant tissue, xylem sap and apoplastic fluid. Physiologia plantarum 109:167-179.

Ivanov, A., Kameka, A., Pajak, A., Bruneau, L., Beyaert, R., Hernández-Sebastià, C., and Marsolais, F. (2012). *Arabidopsis* mutants lacking asparaginases develop

normally but exhibit enhanced root inhibition by exogenous asparagine. Amino Acids 42:2307-2318.

Kapila, J., De Rycke, R., Van Montagu, M., and Angenon, G. (1997). An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant Science 122:101-108.

Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534-2536.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for bigger datasets. Molecular Biology and Evolution 33:1870-1874.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680-685.

Lewis, L.A., and McCourt, R.M. (2004). Green algae and the origin of land plants. American Journal of Botany 91:1535-1556.

McMillan, A., Renaud, J.B., Gloor, G.B., Reid, G., and Sumarah, M.W. (2016). Post-acquisition filtering of salt cluster artefacts for LC-MS based human metabolomic studies. Journal of Cheminformatics 8:44.

Miyashita, Y., and Good, A.G. (2008). NAD(H)-dependent glutamate dehydrogenase is essential for the survival of Arabidopsis thaliana during dark-induced carbon starvation. J Exp Bot 59:667-680.

Morollo, A.A., and Eck, M.J. (2001). Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*. Nat Struct Mol Biol 8:243-247.

Mouilleron, S., and Golinelli-Pimpaneau, B. (2007). Conformational changes in ammonia-channeling glutamine amidotransferases. Current Opinion in Structural Biology 17:653-664.

Nelson, B.K., Cai, X., and Nebenführ, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in *Arabidopsis* and other plants. The Plant Journal 51:1126-1136.

Noctor, G., Dutilleul, C., De Paepe, R., and Foyer, C.H. (2004). Use of mitochondrial electron transport mutants to evaluate the effects of redox state on photosynthesis, stress tolerance and the integration of carbon/nitrogen metabolism. Journal of Experimental Botany 55:49-57.

Obayashi, T., Okamura, Y., Ito, S., Tadaka, S., Aoki, Y., Shirota, M., and Kinoshita, K. (2014). ATTED-II in 2014: evaluation of gene coexpression in agriculturally important plants. Plant Cell Physiology 55:e6.

Prince, J.T., and Marcotte, E.M. (2006). Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. Analytical Chemistry 78:6140-6152.

Roitinger, E., Hofer, M., Köcher, T., Pichler, P., Novatchkova, M., Yang, J., Schlögelhofer, P., and Mechtler, K. (2015). Quantitative phosphoproteomics of the ataxia telangiectasia-mutated (ATM) and ataxia telangiectasia-mutated and rad3-related (ATR) dependent DNA damage response in *Arabidopsis thaliana*. Molecular & Cellular Proteomics 14:556-571.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., et al. (2002). A high-throughput *Arabidopsis* reverse genetics system. The Plant cell 14:2985-2994.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry 78:779-787.

Sparkes, I.A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nature Protocols 1:2019-2025.

Stitt, M. (1999). Nitrate regulation of metabolism and growth. Current Opinion in Plant Biology 2:178-186.

Stitt, M., Müller, C., Matt, P., Gibon, Y., Carillo, P., Morcuende, R., Scheible, W.R., and Krapp, A. (2002). Steps towards an integrated view of nitrogen metabolism. Journal of Experimental Botany 53:959-970.

Tambasco-Studart, M., Tews, I., Amrhein, N., and Fitzpatrick, T.B. (2007). Functional analysis of PDX2 from *Arabidopsis*, a glutaminase involved in vitamin B6 biosynthesis. Plant physiology 144:915-925.

Tautenhahn, R., Böttcher, C., and Neumann, S. (2008). Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673-4680.

Wang, R., Okamoto, M., Xing, X., and Crawford, N.M. (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiol. 132:556-567.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. PLoS One 2:e718.

Zhu, H., and Kranz, R.G. (2012). A nitrogen-regulated glutamine amidotransferase (GAT1_2.1) represses shoot branching in *Arabidopsis*. Plant physiology 160:1770-1780.

4 EVALUATION OF EXTRACTION AND CHROMATOGRAPHIC METHODS SUITABLE FOR METABOLOMICS IN ARABIDOPSIS THALIANA

4.1 Introduction

The four central "omics" platforms (genomics, transcriptomics, proteomics and metabolomics) are clearly invaluable to systems biology, not only for hypothesis generation, but also for elucidating complex phenotypes at the systems level (Bino et al., 2004; Fiehn et al., 2000; Saito and Matsuda, 2010). Metabolites are small molecules that undergo chemical transformations during metabolism and provide the final readout of cellular state (Patti et al., 2012). The metabolome consists of the total complement of low molecular weight molecules, referred to as metabolites, in a cell, tissue, or entire organism at a specific physiological state (Goodacre et al., 2004). Metabolomics, being a relatively new "omics" platform, has not seen widespread use in the field of plant primary metabolism due to major challenges in reliable compound identification and data processing techniques (Fiehn et al., 2008). Since metabolic perturbations represent the final outcome of changes at different levels of "the central dogma" of biology, the development of tools and techniques for identification and quantification of metabolomics data is of utmost importance.

Methods in metabolomics studies generally fall into two categories: targeted metabolomics for the measurement and quantification of groups of chemically characterized and annotated metabolites, and untargeted metabolomics, which focuses on the comprehensive analysis of all measurable analytes, including ones that are chemically unknown (Patti et al., 2012). Untargeted approaches provide the potential of identifying novel targets, however, the coverage of the metabolome is often limited by the methodologies used for sample preparation, sensitivity and selectivity of analytical techniques (Roberts et al., 2012). In the past decade, studies in plant primary metabolism using the model plant, *Arabidopsis*, have employed several extraction procedures suitable for the identification and quantification of highly polar and semi-polar analytes using targeted approaches employing liquid or gas chromatographic separations (Giavalisco et

al., 2011; Lisec et al., 2006). Gas chromatography has been a preferred choice for the analysis of primary metabolites due to the low reproducibility of retention times in liquid chromatography (despite its advantages in minimal sample preparation and broader coverage of metabolites). However, extensive sample preparation for derivatization, longer runs, and the minimal range of metabolites detectable through gas chromatography makes it an inefficient technique for untargeted metabolomics. In this chapter, we provide a comprehensive evaluation of five previously described extraction procedures with respect to their suitability of use with untargeted metabolomics approaches. In addition, a liquid chromatography based method for the reliable and reproducible separation of primary metabolites is presented.

Processed metabolomics data typically consist of a list of features that are defined by an accurate mass, retention time and peak areas that are representative of their relative quantities. The tremendous chemical diversity of metabolites, experimental artifacts and the complexity of sample matrices often result in metabolomics data that are not a true representation of metabolites in biological samples (Bowen and Northen, 2010). A typical feature list often contains compounds that represent isotopes, adducts, in-source fragments, artifacts and contaminants, therefore, noise reduction and prioritization of data is essential to reduce the cost and effort of data processing in metabolomics experiments (Mahieu and Patti, 2017). To address the challenge of data processing in untargeted metabolomics, we propose an optimized workflow and demonstrate its application to isotopically enriched plant samples treated with $^{13}C_5$, $^{15}N_2$ Gln. Here, we take advantage of the resolving power of Q-Exactive Orbitrap mass-spectrometer for detection and identification of primary metabolites and their labelled counterparts.

4.2 Results

4.2.1 Comparison of metabolite profiles obtained from different extraction methods

To test the recovery of primary metabolites from plant roots, total metabolite extracts (obtained from the five methods described below) were used. A total of 1449 features in the positive mode and 792 features in the negative mode were obtained following the XCMS based feature detection in R (Section 4.2.7). Figure 4.1 represents strip charts showing normalized peak intensities (y-axis) of features from both positive (A) and negative (B) mode. Although not all features detected are metabolites of biological origin, the strip charts allow a comparison of the metabolite signal recovery of. Acidified methanol, chloroform:methanol and acidified methanol:water extractions were found to result in larger number of features with low normalized intensities (< 6) in both ionization modes. High intensities of maximum number of features were detected in the methanol:water extracts and both aqueous and organic phases of methanol:MTBE:water extracts. Since most of the primary metabolites have a mass to charge ratio (*m/z*) between 100-500 amu, a comparison of the recovery based on *m/z* of features detected in both positive (Figure 4.2) and negative (Figure 4.3) mode was performed (Figure 4.2 and 4.3). Once again, maximum recovery of features between desired m/z range was observed in the methanol:water extracts as well as both phases of methanol:MTBE water extraction. We then used a targeted approach to detect a few amino acids and organic acids (Figure 4.4). Acidified methanol extracts resulted in the lowest signals with high variability between replicates. Variation between replicates is also high for most metabolites identified in chloroform:methanol extracts as well as methanol:MTBE:water extracts. The signal intensities of target metabolites, which include Glu, GABA and malic acid, showed a lower recovery in acidified methanol:water extracts. High quantities with consistent signals from all the metabolites were resulted from the methanol:water extracts (Figure 4.4). Considering all the presented evidence, the methanol:water extraction technique was determined to be the best for extraction and detection of primary metabolites.

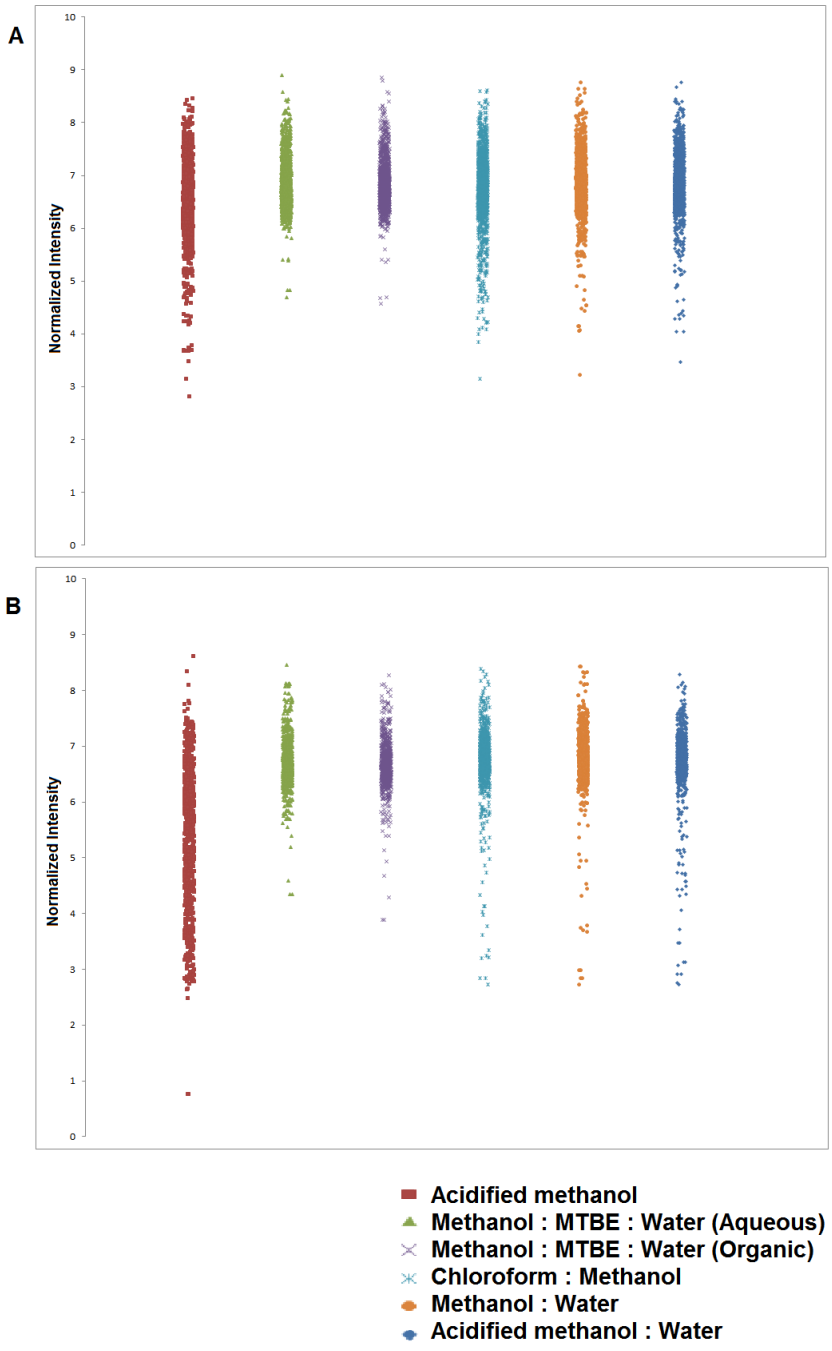


Figure 4.1: Strip charts representing total metabolomic data from different extraction methods.

A. Metabolite features obtained from positive ionization mode and B. Features obtained from negative ionization mode. Y-axis represents normalized intensities that were log transformed peak areas.

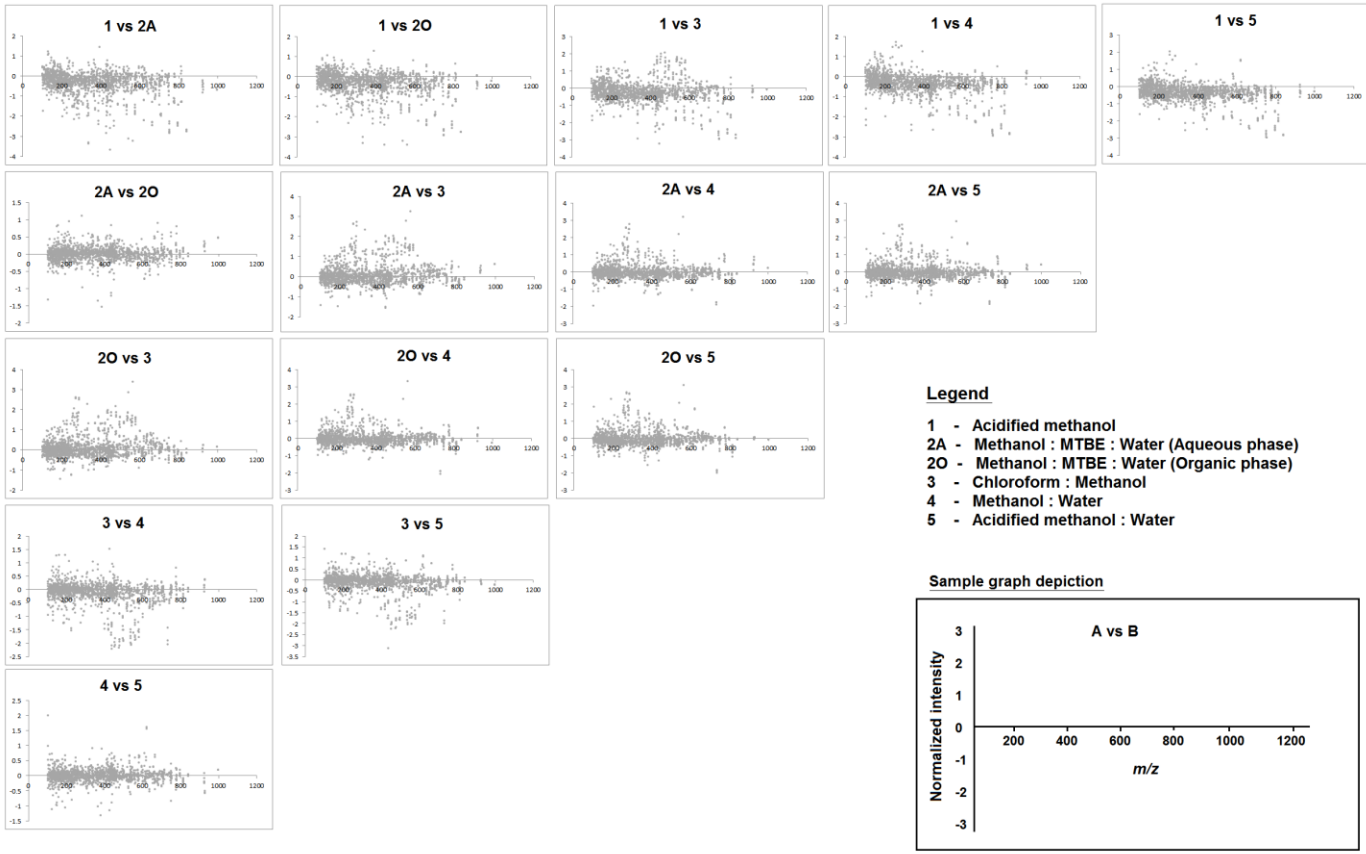


Figure 4.2: Scatter plots with m/z of features and intensities from positive mode.

The individual panels represent comparisons of total metabolomic feature lists between two extraction methods. Labels are formatted as A vs B where A is presented in the positive Y-axis and B is presented in the negative Y-axis. A higher value in positive Y and lower value in negative Y-axes represents higher recovery of that particular feature. The mass range is presented on the X-axis.

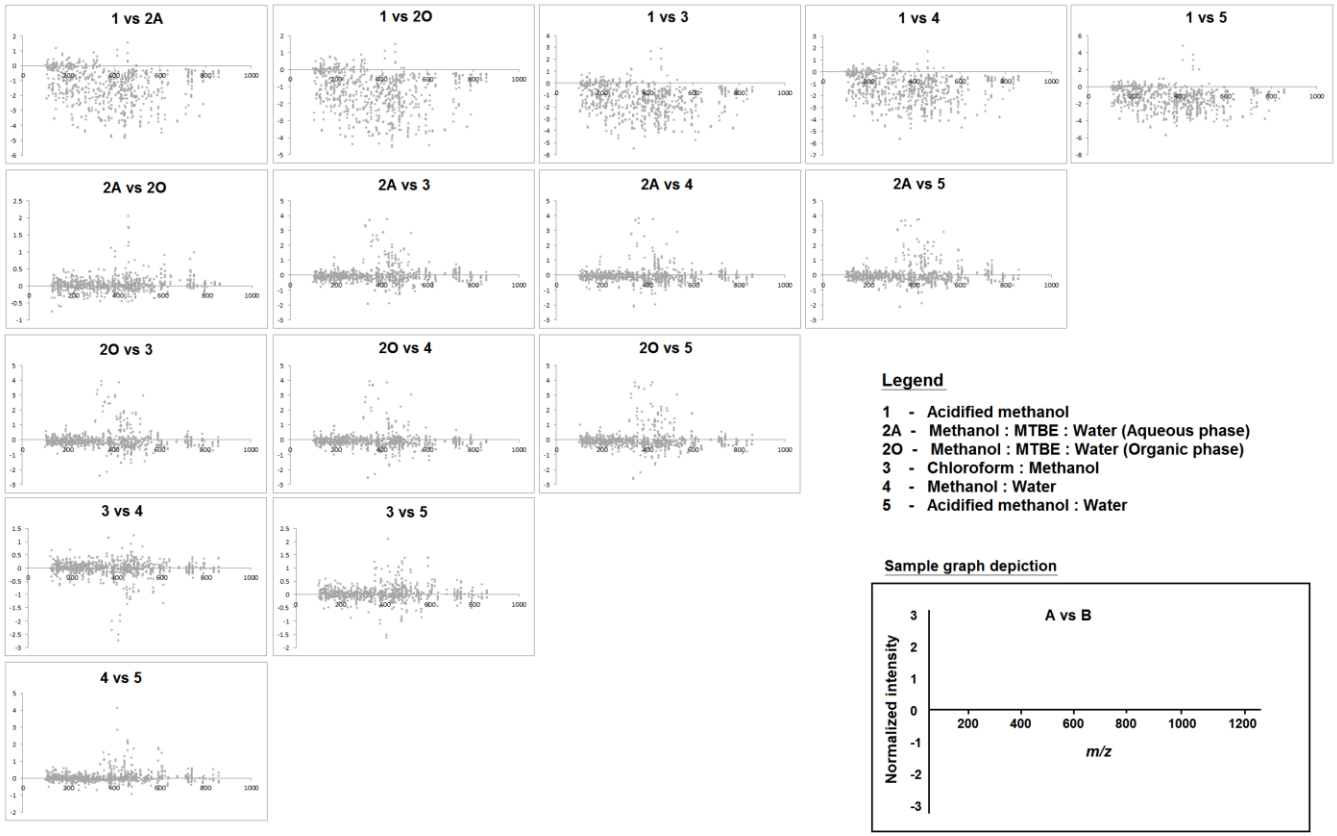


Figure 4.3: Scatter plots with m/z of features and intensities from negative mode.

The individual panels represent comparisons of total metabolomic feature lists between two extraction methods. Labels are formatted as A vs B where A is presented in the positive Y-axis and B is presented in the negative Y-axis. A higher value in positive Y and lower value in negative Y-axes represents higher recovery of that particular feature. The mass range is presented on the X-axis.

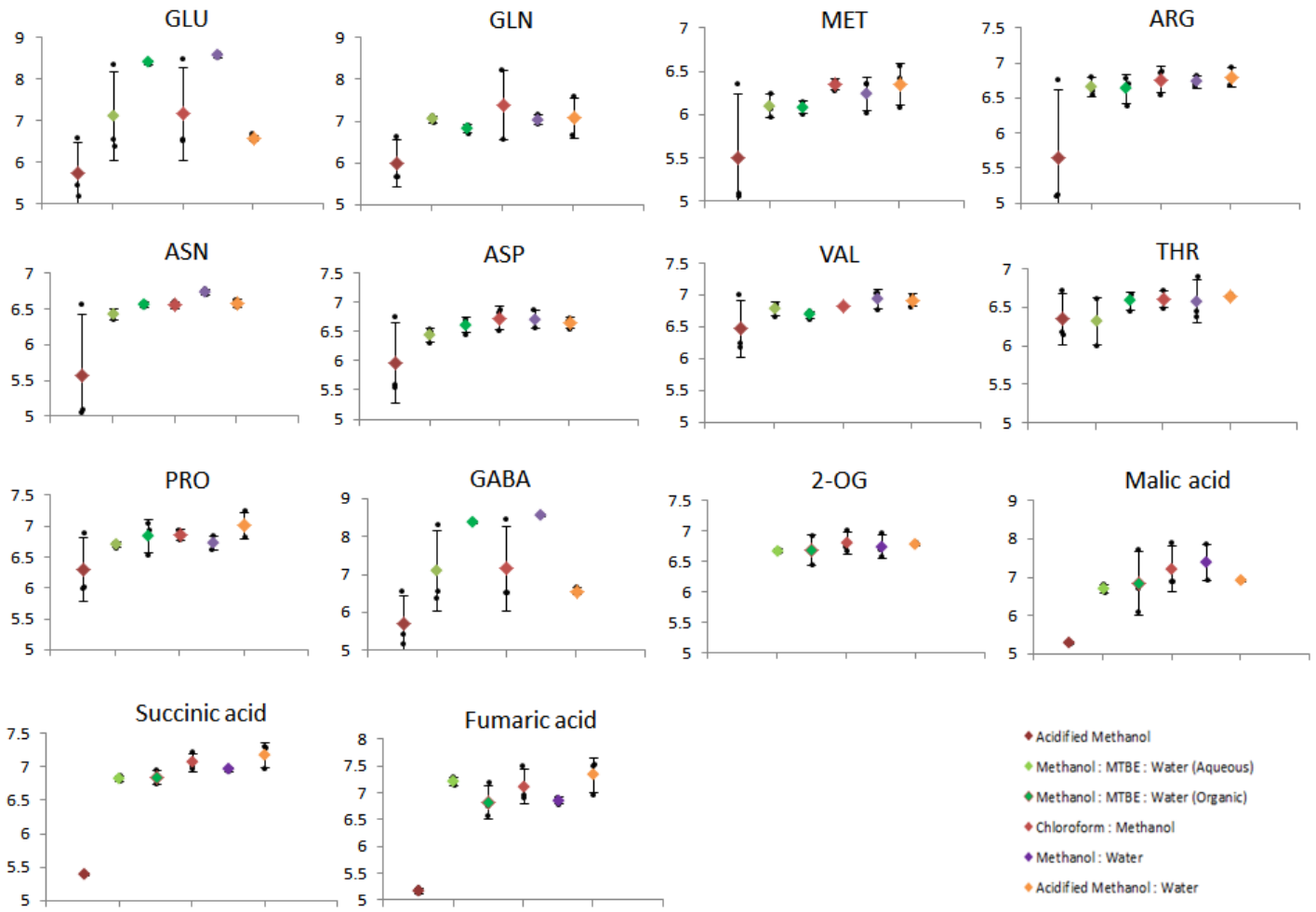


Figure 4.4: Targeted metabolite analysis showing a few amino acids and organic acids.

Y-axis represents normalized intensities that were log transformed peak areas.

4.2.2 Chromatographic resolution of amino acids is achieved with HILIC based chromatography

Quantification of metabolites detected by mass spectrometry is performed by the use of peak areas obtained upon separation of compounds by chromatography using a partitioning column. A common partitioning column using for liquid chromatography is a reverse phase C18 column. However, highly polar compounds such as amino acids and organic acids do not retain well on conventional C18 columns. Hence, we used a hydrophilic interaction column (HILIC). A good separation of all the proteinogenic and non-proteinogenic amino acids including isobaric compounds (Leu/Ile, Thr/Homoser) was achieved by taking advantage of the HILIC water partitioning mode by first separating compounds isocratically with 13% A (5 mM ammonium acetate pH 4.0) and 87% B (acetonitrile with 0.1% formic acid) for 5 mins followed by gradient elution by decreasing % B to 55% over the next 8 minutes (Figure 4.5). Most of the amino acids were eluted during this time and the highly polar ones (Lys, His, Arg) eluted over the next 4 minutes by holding at 55% B (Figure 4.5).

4.2.3 Integration of steps to obtain optimized metabolomics data

A combination of packages in R (described in section 4.2.7) were used to obtain an optimum feature list that has reduced noise with salt clusters filtered and a higher number of true biological compounds retained. The isotopologue parameter optimization (IPO) package in R was first used to obtain a set of XCMS peak picking parameters to excise features that mostly represent biological metabolites based on their associated natural isotopic signatures (Libiseller et al., 2015). Parameters for retention time correction and grouping were also optimized by IPO by minimizing relative retention times within peak groups and maximizing the number of peak groups that show one peak per injection, respectively. Optimized parameters for peak picking, retention time correction, and grouping delivered by IPO were then used for building feature listing with the XCMS package in R (Smith et al., 2006a). A total of 1835 and 916 features were detected in the positive and negative modes respectively following this procedure using the data from different extraction methods described in section 4.3.1. From these feature lists, 21% and

13% of the features that represent salt clusters and other ionization artifacts were filtered out by applying McMillan correction in R (McMillan et al., 2016) from positive and negative ionization modes respectively, resulting in final feature lists that contained 1449 and 792 features (Section 4.3.1).

The workflow described above, however, cannot be used for obtaining optimum feature lists from total metabolite extracts derived from plants treated with stable labelled isotopes since all of the isotopically labelled metabolites are not recognized by IPO due to the peak picking algorithm used by the package (Libiseller et al., 2015). To avoid this problem, a two-step feature detection was used, in which, optimized feature lists were first built from unlabeled controls following the procedure described above followed by the identification of a set of target metabolites for which label enrichment needs to be determined. Feature lists from label enriched samples were then obtained by using relaxed parameters in XCMS (default) following which a custom in-house python script was used to extract labelled compounds based on mass difference compared to their unlabeled counterparts. The python script also includes flexibility for inputting user determined retention time and mass accuracy errors. This workflow was successfully demonstrated in extracting isotopic label enriched features from wild-type *Arabidopsis* seedlings treated with dual labelled ($^{13}C_5$, $^{15}N_2$) glutamine for two hours (Figure 4.6).

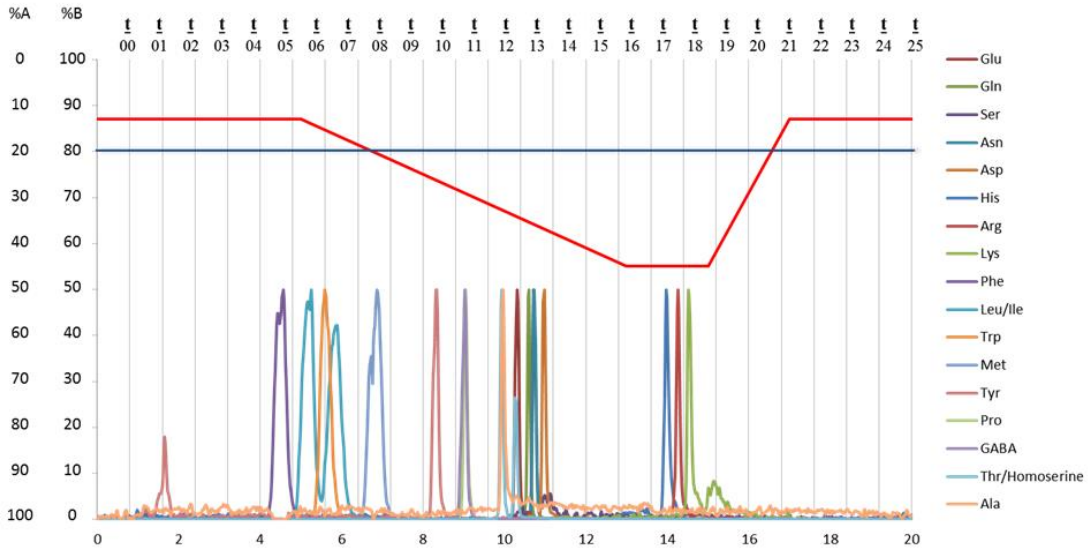


Figure 4.5: Gradient and chromatography using HILIC.

Red line across the figure indicates the gradient used for chromatographic separation. Blue line indicates the 20% A where hydrophilic interactions are at the highest. Y-axis shows the mobile phase. X-axis labels indicate the column volumes on top and retention time at the bottom.

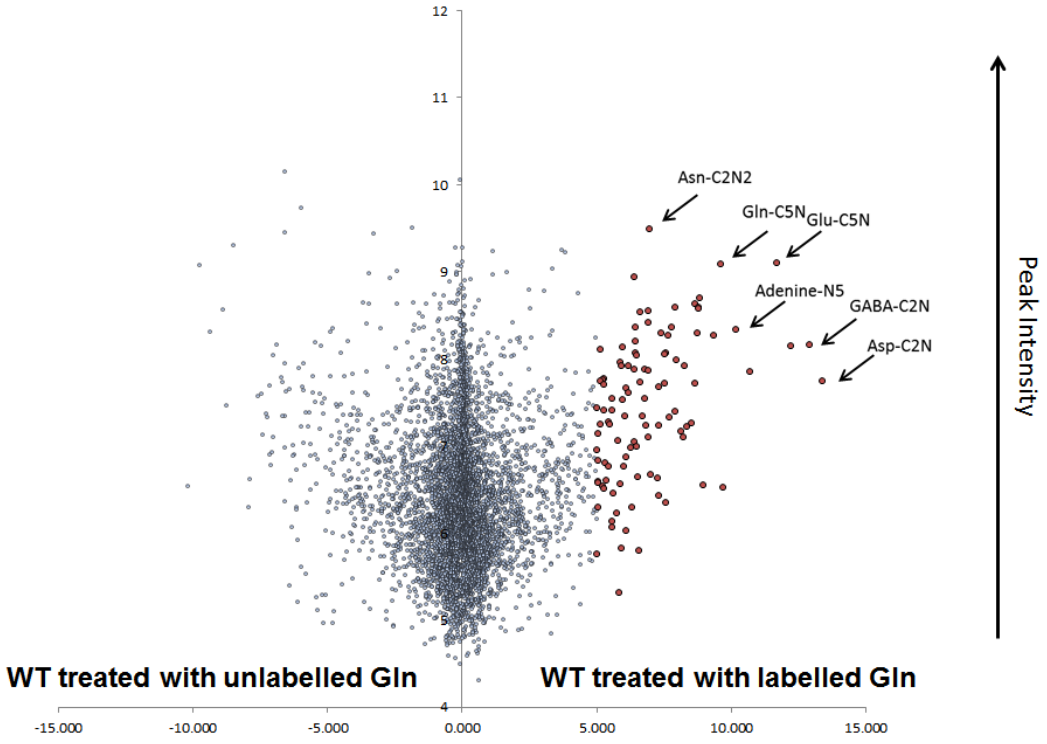


Figure 4.6: Summed intensity plot highlighting isotopically labelled compounds from $^{13}C_5$, $^{15}N_2$ Gln treated *Arabidopsis* roots.

Total metabolomics feature list obtained from non-labeled and labeled Gln treated samples were used to plot the graph. X-axis represents fold change and Y-axis represents normalized intensities obtained by log transformation of peak areas. Compounds labelled within the feature were identified using the in-house python script.

4.2.4 HRMS is essential for detection of dual labels for LC-MS based metabolomics

Use of the above described optimized parameters for detection of dual labelled isotopic features is dependent on the quality of the data obtained. With the availability of high resolution instruments, it is now possible to achieve mass accuracies within < 1 ppm error of exact mass. However, the ability to image a detected ion in a mass spectrometer and differentiate it from any other depends on the resolving power of the instrument. Figure 4.7 demonstrates the resolving power of a Q-Exactive Orbitrap mass spectrometer at a resolution of 140,000 FWHM (full width at half maximum) in comparison with a Q-TOF at a resolution of 20,000 FWHM and a triple-quadrupole at a resolution of 500 FWHM that are two other leaders in the market for mass spectrometry instrumentation. Resolutions for Q-TOF and triple-quadrupole were simulated using the Xcalibur™ software (Thermo Scientific). The figure uses Gln $[M]^+$ ion with accurate mass, 147.07642 m/z, and its associated natural ^{13}C isotopic peak $[M+1C]^+$ with accurate mass, 148.07912 m/z (Figure 4.7A). A closer look at the isotopic peak (Figure 4.7B), however, at higher resolution with a Q-Exactive, reveals a second ^{15}N natural isotopic peak $[M+1N]^+$ with accurate mass, 149.07341 m/z. This peak is masked by the $[M+1C]^+$ peak at lower resolutions. The resolving power of Q-Exactive can be used to our advantage when extracting isotopic label information from total metabolomic feature lists obtained from plant tissues enriched with a dual labelled compound, e.g. ^{13}C, ^{15}N Gln.

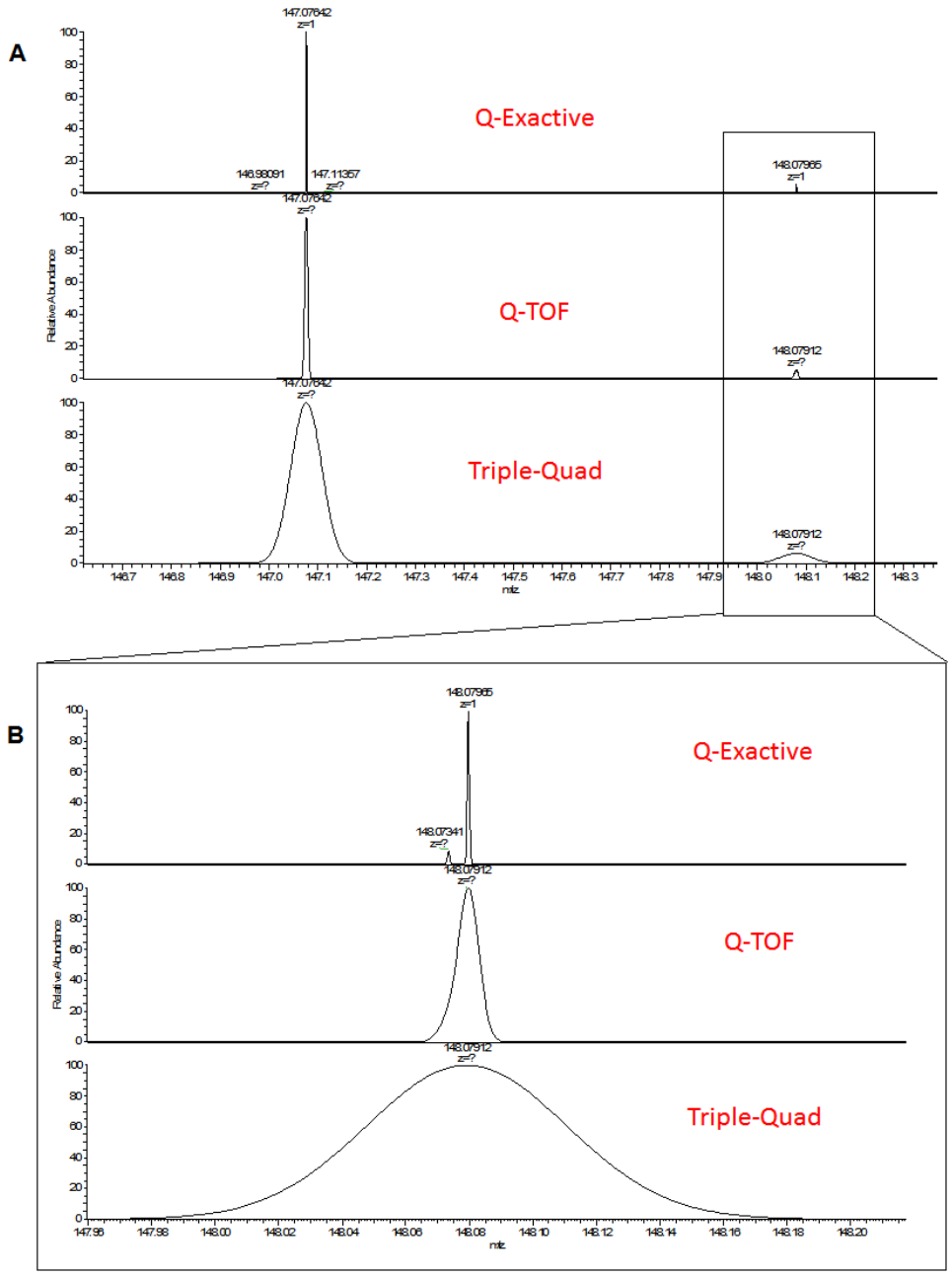


Figure 4.7: Comparison of resolving power between different instruments.

A. presents the peak detected at m/z 147.07642 that corresponds to Gln along with its natural isotopic counterpart at m/z 148.07912. B. is a zoomed in version of the isotopic peak shown in A and depicts the two resolved m/z 148.07341 and 148.07912 that correspond to ^{15}N and ^{13}C natural isotopes respectively in the top panel and the unresolved versions in the bottom two panels.

4.3 Discussion

The terminal downstream product of changes at any level of "the central dogma" of biology is the metabolome (Goodacre et al., 2004). The primary goal of any metabolomics experiment is to determine the total complement of metabolites (Nicholson et al., 1999), however, several challenges exist in this relatively new field that complicate metabolite detection as well as identification (Fiehn et al., 2008). With a focus on primary metabolites, we aimed to evaluate and develop a few methods for optimal extraction of the metabolite pool from plant tissues, chromatographic separation as well as data processing. A total of five extraction procedures, previously described (Giavalisco et al., 2011; Lisec et al., 2006), were evaluated and a simple methanol:water extraction was determined to be best suited for the extraction of maximum metabolite pool with high recovery from plant tissue samples.

Although use of HILIC in liquid chromatography for amino acids and organic acids was demonstrated previously, methods longer than 50 minutes were used to achieve these separations (Schiesel et al., 2010; Zangrando et al., 2010). The HILIC method employs a zwitterionic bonded silica column and enables separation of hydrophilic compounds without prior derivatization and is compatible with mass spectrometric detection (Langrock et al., 2006). Here, we achieved better separation and shorter run time as demonstrated with amino acids (Figure 4.5) by taking advantage of the HILIC water partitioning mode between 10-30% A in the mobile phase. An initial 5-minute isocratic hold at 13% A before initiating the gradient elution resulting in complete separation of all the amino acids. Following successful separation, we demonstrated the use of an optimized workflow for untargeted metabolomics experiments upon isotopic label enrichment by taking advantage of high mass accuracy and resolving power of Q-Exactive Orbitrap mass spectrometer.

Advances in instrumentation have resulted in achieving high mass accuracy that can be used for compound identification (Kind and Fiehn, 2006). However, mass accuracy alone is insufficient if isotopic fine structures of compounds need to be evaluated, as is the case for our example where plant tissues are enriched with $^{13}C_5$, $^{15}N_2$ Gln. Hence, we demonstrated the use of optimized chromatographic conditions together with high

resolution mass spectrometry (HRMS) (Marshall and Hendrickson, 2008) for separation and detection of labelled compounds. Our pipeline for extraction of feature lists from metabolomics experiments integrates IPO, XCMS and McMillan correction packages in R for noise reduction and filtering of contaminants among metabolite features. Further statistical analyses can be applied in R or Excel for prioritization of compound identification based on experimental design. For isotopic feature detection, an in-house python script, with flexibility for errors in mass accuracy and deviations in retention times, was developed and successfully utilized. This can be applied for elucidation of novel pathways by isotope tracking experiments using both targeted and untargeted metabolomics approaches.

4.4 Materials and Methods

4.4.1 Plant growth conditions

Arabidopsis thaliana ecotype *Columbia* was used for both transcript and metabolite profiling experiments. Plants were grown on vertical plates at 22°C under continuous light (ca. 70 µmol m^{-2} s^{-2}), as previously described by Ivanov et al. (2012) on a defined nutrient medium containing a final concentration of 10 mM potassium phosphate (pH 6.5), 5 mM KNO_3, 2 mM $MgSO_4$, 1 mM $CaCl_2$, 125 µg FeNaEDTA, micronutrients (50 mM H_3BO_3, 12 mM $MnSO_4$, 1 mM $ZnCl_2$, 1 mM $CuSO_4$ and 0.2 mM Na_2MoO_4), 1% sucrose and 1% agar (Wang et al., 2003). Roots were excised from ten-day old seedlings, flash frozen in liquid N_2 and stored at -80°C until total metabolite extractions was carried out. Three replicates were used for extraction using all five methods (described below). The methods were optimized under these plant growth conditions so they were suitable for the experiments described in Chapters 2 and 3. For label enrichment comparison, ten-day old seedling were transferred into plates containing the above described media supplemented with 10 mM Gln or 10 mM $^{13}C_5$, $^{15}N_2$ Gln as sole source of N for unlabeled and labelled treatments respectively and roots were excited after 2 hours of incubation.

4.4.2 Extraction using acidified methanol

Fifty mg of frozen root tissue was homogenized using a ball-mill. Ice-cold sample extraction solution (99.875% methanol acidified with 0.125% FA) in a volume/fresh weight ratio of 3 was added to the homogenized root tissue followed by immediate vortexing for 10 s. Assuming a tissue water content of about 95%, this resulted in a final concentration of 75% methanol and 0.1% FA. Extracts were stored on ice until all samples were ready. Samples were sonicated for 15 min at maximum frequency (40 kHz) continuously, in a water bath at room temperature (22°C) followed by centrifugation for 10 min at maximum speed at room temperature. The supernatants from each sample were transferred to a fresh tube and filtered through a 0.2-µm PTFE centrifugation filter.

4.4.3 Extraction by phase separation

Fifty mg of the frozen root tissue was homogenized using a ball-mill. The metabolites were isolated in 1 ml of homogeneous mixture of -20°C methanol:methyl-tert-butyl-ether:water (1:3:1), with shaking for 30 min at 4°C, followed by another 10 min of incubation in an ice-cooled ultra-sonication bath. 650 µl of methanol:water (1:3) is then added and the homogenate was vortexed and spun for 5 min at 4°C in a table top centrifuge. This allowed phase separation, providing the upper organic phase, containing the lipids, a lower aqueous phase, containing the polar and semi polar metabolites, and a pellet of starch and proteins at the bottom of tube. The aqueous phase and organic phase were then collected separately and dried down in a speed vac and then stored at -80°C until required. The dried down sample was re-suspended in 800 µl of 50% methanol and used for further analysis.

4.4.4 Extraction with chloroform and methanol

Fifty mg of frozen root tissue was ground to a fine powder using ball mill and quenched by adding 250 µl of ice-cold CHCl3/CH3OH (3:7, v/v). The frozen mixture was incubated at -20°C for 2 h with occasional mixing. Water soluble components were extracted from the $CHCl_3$ phase by adding 200 µl of water at 4°C with repeated shaking. After centrifugation at 420 *g* for 4 min, the upper, aqueous-CH_3OH phase was transferred to a new tube, and kept at 4°C. The lower, $CHCl_3$ phase was used for re-extraction with 200 µl of cold water, centrifuged as described above, and the second aqueous-CH_3OH extract was combined with the first. The combined aqueous-CH_3OH extract was evaporated to dryness using a centrifugal vacuum dryer at 20°C and re-dissolved in 800 µl of 50% methanol. Viscous, high molecular-mass components were removed from the samples by filtering the extracts by centrifugation at 2300 *g* using a 0.2 µm PTFE filter.

4.4.5 A simple methanol:water extraction

Fifty mg root tissue was homogenized in 1,000 μL of methanol:water (4:1, v/v) using a ball-mill. Samples were sonicated for 30 min at room temperature followed by centrifugation at 11,000 x *g* for 10 min, and 700 μL of the supernatant was transferred to

a fresh tube and evaporated to dryness in a vacuum centrifuge at ambient temperature. The remaining residue was re-dissolved in 400 μL of methanol by vigorous vortexing and diluted with 400 μL of water. Prior to LC/MS analysis, the extracts were filtered through a 0.2 μm PTFE centrifugation filter.

4.4.6 Acidified methanol: water extraction

Total metabolite extraction was carried out following the same method described in section 4.2.5 with the exception of methanol: water (4:1, v/v) extraction buffer being substituted with 79:20:1 (methanol: water: formic acid, v/v/v).

4.4.7 Data acquisition by targeted and untargeted metabolite analysis

All MS data were acquired on a Q-Exactive Quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled to an Agilent 1290 high performance liquid chromatography (HPLC) system. A Zorbax Eclipse Plus RRHD C18 column (2.1 x 50 mm, 1.8 µm) was maintained at 35°C. Mobile phases were comprised of 5 mM ammonium acetate, pH 4 (A), and acetonitrile with 0.1% formic acid (B). Mobile phase A was maintained at 100% for 1.25 min and mobile phase B was increased to 50% over 2.25 min, and 100% over 0.5 min. The following heated electrospray ionization (HESI) conditions were optimized for the analysis of amino and organic acids: spray voltage, 3.9 kV (ESI+), 3.5 kV (ESI-); capillary temperature, 250°C; probe heater temperature, 100°C; sheath gas, 30 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 60%. Injections of 5 µl were used with a flow rate of 0.3 mL min^{-1}. Compounds were detected and monitored using a targeted MS/MS approach with spectra collected at 17,500 resolution, AGC target 5e5, maximum IT 65 ms and isolation window of 1 m/z. Normalized collision energy of 27 was used for the targeted method. Data were analyzed and all theoretical masses were calculated with Xcalibur™ software. Full MS peak areas were normalized with a $^{13}C_6$ phenylalanine internal standard for positive ionization and $^{13}C_3$ pyruvic acid for negative ionization and compared across treatments. Compounds were identified using commercial standards.

For metabolomics analysis, LC-MS data was obtained with the above mentioned chromatographic and HESI conditions but full MS measurements were collected from

mass ranges of 75-1100 m/z and 65-900 m/z in positive and negative ionization modes respectively at 140,000 resolutions. The AGC target and maximum IT was set to 3 e6 and 524 ms respectively. Following data acquisition, Thermo .RAW files were converted to .mzml format using ProteoWizard (Kessner et al., 2008) with peak picking filter applied. Files were imported into R using the XCMS package (Smith et al., 2006b) and features were detected using the centWave (Tautenhahn et al., 2008) method and a ppm tolerance of 1.0. Pre-filter was set to 6 scans with minimum 5000 intensity, signal to noise threshold was 5 and noise was set to 3×10^6 and 1×10^6 for positive and negative mode respectively. Retention time correction was conducted using the obiwarp (Prince and Marcotte, 2006) method, grouping included features present in at least 25% of all samples, allowable retention time deviation was 10 seconds, and *m/z* width set to 0.015. The "fillPeaks" function was used with default settings and remaining zeros were imputed with two-thirds the minimum value on a per mass basis. Salt clusters and other ionization artefacts were removed from the feature list using the McMillan correction without retention time filter applied (McMillan et al., 2016). All features were exported to excel to test for statistical significance between treatments using a student's t-test and to generate volcano plots. An in-house python script was used for extraction of isotopic features from the total feature list obtained.

4.4.8 Optimized chromatography using HILIC

For chromatography using the Hydrophilic interaction column (HILIC), compounds were resolved using a SeQuant® ZIC®-HILIC column; 3.5µm, 100 Å, 100×2.1 mm (EMD Millipore) with mobile phase 5 mM ammonium acetate, pH = 4.00 (A); B = 90% acetonitrile, 0.1% formic acid) and the following gradient: 87% B for 5 minutes, decreased to 55% over 8 minutes and held for 4 minutes before returning to 87% over 3 minutes. The following heated electrospray ionization (HESI) conditions were optimized for the analysis of amino and organic acids: spray voltage, 3.9 kV (ESI+), 3.5 kV (ESI-); capillary temperature, 250 °C; probe heater temperature, 450 °C; sheath gas, 30 arbitrary units; auxiliary gas, 8 arbitrary units; and S-Lens RF level, 60%.. Injections of 5 µl were used with a flow rate of 0.3 mL min^{-1}. Following chromatography full MS data was obtained from the above described method.

4.5 Literature cited

Bino, R.J., Hall, R.D., Fiehn, O., Kopka, J., Saito, K., Draper, J., Nikolau, B.J., Mendes, P., Roessner-Tunali, U., Beale, M.H., et al. (2004). Potential of metabolomics as a functional genomics tool. Trends in Plant Science 9:418-425.

Bowen, B.P., and Northen, T.R. (2010). Dealing with the unknown: metabolomics and metabolite atlases. Journal of the American Society for Mass Spectrometry 21:1471-1476.

Fiehn, O., Kopka, J., Dörmann, P., Altmann, T., Trethewey, R., and Willmitzer, L. (2000). Metabolite profiling for plant functional genomics. Nature Biotechnology 18.

Fiehn, O., Wohlgemuth, G., Scholz, M., Kind, T., Lee, D.Y., Lu, Y., Moon, S., and Nikolau, B. (2008). Quality control for plant metabolomics: reporting MSI-compliant studies. The Plant Journal 53:691-704.

Giavalisco, P., Li, Y., Matthes, A., Eckhardt, A., Hubberten, H.M., Hesse, H., Segu, S., Hummel, J., Kohl, K., and Willmitzer, L. (2011). Elemental formula annotation of polar and lipophilic metabolites using (13) C, (15) N and (34) S isotope labelling, in combination with high-resolution mass spectrometry. The Plant journal : for cell and molecular biology 68:364-376.

Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B. (2004). Metabolomics by numbers: acquiring and understanding global metabolite data. Trends in Biotechnology 22:245-252.

Ivanov, A., Kameka, A., Pajak, A., Bruneau, L., Beyaert, R., Hernández-Sebastià, C., and Marsolais, F. (2012). *Arabidopsis* mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. Amino Acids 42:2307-2318.

Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008). ProteoWizard: open source software for rapid proteomics tools development. Bioinformatics 24:2534-2536.

Kind, T., and Fiehn, O. (2006). Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. BMC Bioinformatics 7:234-234.

Langrock, T., Czihal, P., and Hoffmann, R. (2006). Amino acid analysis by hydrophilic interaction chromatography coupled on-line to electrospray ionization mass spectrometry. Amino Acids 30:291-297.

Libiseller, G., Dvorzak, M., Kleb, U., Gander, E., Eisenberg, T., Madeo, F., Neumann, S., Trausinger, G., Sinner, F., Pieber, T., et al. (2015). IPO: a tool for automated optimization of XCMS parameters. BMC Bioinformatics 16:118.

Lisec, J., Schauer, N., Kopka, J., Willmitzer, L., and Fernie, A.R. (2006). Gas chromatography mass spectrometry-based metabolite profiling in plants. Nature Protocols 1:387+.

Mahieu, N.G., and Patti, G.J. (2017). Systems-level annotation of a metabolomics data set reduces 25 000 features to fewer than 1000 unique metabolites. Analytical Chemistry.

Marshall, A.G., and Hendrickson, C.L. (2008). High-Resolution Mass Spectrometers. Annual Review of Analytical Chemistry 1:579-599.

McMillan, A., Renaud, J.B., Gloor, G.B., Reid, G., and Sumarah, M.W. (2016). Post-acquisition filtering of salt cluster artefacts for LC-MS based human metabolomic studies. Journal of Cheminformatics 8:44.

Nicholson, J.K., Lindon, J.C., and Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica 29:1181-1189.

Patti, G.J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the apogee of the omic triology. Nature reviews: Molecular cell biology 13:263-269.

Prince, J.T., and Marcotte, E.M. (2006). Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. Analytical Chemistry 78:6140-6152.

Roberts, L.D., Souza, A.L., Gerszten, R.E., and Clish, C.B. (2012). Targeted metabolomics. current protocols in molecular biology chapter : Unit 30 - 32.

Saito, K., and Matsuda, F. (2010). Metabolomics for functional genomics, systems biology, and biotechnology. Annual Review of Plant Biology 61:463-489.

Schiesel, S., Lämmerhofer, M., and Lindner, W. (2010). Multitarget quantitative metabolic profiling of hydrophilic metabolites in fermentation broths of β-lactam antibiotics production by HILIC–ESI–MS/MS. Analytical and Bioanalytical Chemistry 396:1655-1679.

Smith, C., Want, E., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006a). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Analytical Chemistry 78.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006b). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry 78:779-787.

Tautenhahn, R., Böttcher, C., and Neumann, S. (2008). Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504.

Wang, R., Okamoto, M., Xing, X., and Crawford, N.M. (2003). Microarray analysis of the nitrate response in *Arabidopsis* roots and shoots reveals over 1,000 rapidly responding genes and new linkages to glucose, trehalose-6-phosphate, iron, and sulfate metabolism. Plant Physiol. 132:556-567.

Zangrando, R., Piazza, R., Cairns, W.R.L., Izzo, F.C., Vianello, A., Zendri, E., and Gambaro, A. (2010). Quantitative determination of un-derivatised amino acids in artistic mural paintings using high-performance liquid chromatography/electrospray ionization triple quadrupole mass spectrometry. Analytica Chimica Acta 675:1-7.

5 GENERAL DISCUSSION

5.1 Excess Asn accumulation results in C/N reprogramming

Asparagine is clearly an ideal storage and transport form of N (described in section 1.2) due to its high N:C ratio and low net charge under physiological conditions (Lea et al., 2007). The N status of a plant, the overall C:N ratio, as well as photoperiod are major determining factors for Asn biosynthesis via the modulation of AS trasncripts (Gutiérrez et al., 2008; Lima and Sodek, 2003; Tsai and Coruzzi, 1990). Asn biosynthesis is upregulated upon high N availability and in the dark period as a means for N storage. Excess Asn can be catabolized and used as an N source for recuperation of C/N balance upon perturbations due to environmental and genetic factors, however, the mechanism and pathways involved in this process were unknown. To address this issue, we conducted transcriptome and metabolome profiling upon Asn treatment to get a comprehensive picture of the mechanism of C/N reprogramming in the roots of the model plant, *Arabidopsis* (Chapter 2).

A number of genes involved in nitrate transport (*NRT1.8*, *NRT1.5*, *NRT2.6*) were repressed, indicating a sufficiency of N compounds within the cell and suppression of further N uptake/influx into the cells (Li et al., 2010). Carbon efflux was lowered by repressing *AlMT2*, a malate transporter, and the supply of C skeletons was achieved by increased expression of genes involved in lignin degradation, polysaccharide degradation and carbohydrate metabolism as identified by GO categorization (Section 2.2.2). Since the transcriptomic data highlighted a major redistribution of C and N resources, we then investigated the metabolic responses upon Asn over accumulation to identify specific pathways through which C/N balance is achieved. To that end, we identified lysine and proline catabolic pathways as well as the GABA shunt to be the primary pathways linking amino acid and carboxy acid metabolism (Section 2.2.4) to achieve the aforementioned C/N balance. Lysine catabolism and the GABA shunt pathway have been previously reported to play a role in C/N partitioning (Boex-Fontvieille et al., 2013; Fait et al., 2008), however, Asn dependent activation of these pathways is being reported here for the first time. Although, no clear mechanism for Asn signal perception was revealed

in this study, a few phytohormone biosynthetic and signaling genes (*IPT3*, *SAUR55*, *AIR12*, *ERF1*, *ERF4*) were found to be upregulated (Appendix A). Here, we suggest that a hormone mediated signaling mechanism of Asn, similar to that of cytokinin dependent nitrate signaling previously described (Gaufichon et al., 2013; Sakakibara et al., 2006), is possible by regulating the genes involved in the above described primary metabolic pathways. Further characterization of Asn signaling machinery is necessary to fully elucidate its mode of function. Although, the data presented in this work does not provide sufficient evidence for a broad signaling role of Asn in plants, it highlight the reallocations that are necessary in order to maintain C and N balance upon increase in N metabolites. It is important, however, to keep in mind that the data represents changes in both transcripts and metabolites upon a 2 hour Asn treatment.

One gene of particular interest among the transcriptomic data is a class I glutamine amidotransferase (*GAT1_2.1*) of unknown substrate specificity (4.3-fold upregulation). The gene was previously found to be highly responsive to N status in plants (Zhu and Kranz, 2012). The protein, GAT1_2.1 has a glutaminase domain responsible for glutamine hydrolysis and a C-terminal extension of unknown homology and the transcript was found to be highly co-expressed with a mitochondrial glutamate dehydrogenase (*GDH2*) that deaminates glutamate to produce 2-oxoglutarate (Section 2.2.6). We suggest a potential new route of C/N interrelationships via glutamine hydrolysis by GAT1_2.1 and the subsequent channeling of C skeletons from glutamate produced into the TCA cycle through the action of GDH2. To test this hypothesis, we proceeded to characterize the function of GAT1_2.1 at both the biochemical and metabolic levels (Chapter 3).

5.2 GAT1_2.1 is a novel plant glutaminase that may regulate C/N partitioning

Cancer-cells are excellent examples of C/N reprogramming in mammals, where the metabolic machinery is altered to achieve the bioenergetics and biosynthetic demands of the cell. In order to meet the carbon requirement, cancer cells adapt to an elevated functioning of glycolytic pathways and hence citric acid cycle activity is lowered (Vander Heiden et al., 2009). To compensate for this, and to help maintain citric acid cycle

functioning, cancer cells resort to increased glutaminase activity in mitochondria thus hydrolyzing glutamine to glutamate followed by the supply of C skeletons from glutamate to 2-oxoglutarate, via *GDH* (Erickson and Cerione, 2010). Two isoforms of mammalian glutaminases have been reported, a liver type and a kidney type, both of which have a high degree of regulation and play an important role in C/N partitioning during starvation, diabetes, high protein diet and metabolic acidosis (Curthoys and Watford, 1995).

Plants, on the other hand, have no known glutaminases reported to date. GS/GOGAT cycle (described in Sections 1.1 and 3.1) is thought to be the major interaction point between C and N metabolism (Stitt et al., 2002). Although this is true, the GS/GOGAT cycle is more pronounced in leaves where there is a continuous supply of C skeletons via photosynthesis. The absence of photosynthesis and the lower activity of GS/GOGAT isoforms in roots (Coschigano et al., 1998; Oliveira and Coruzzi, 1999) suggest a possible alternate mechanism to achieve C/N balance upon perturbation. In Chapter 3, we showed evidence that the nitrogen regulated *GAT1_2.1*, is localized in mitochondria, acts as a glutaminase and may potentially function in conjunction with a mitochondrial GDH to supply C skeletons to TCA cycle. The glutaminase function of GAT1_2.1 may also be regulated post-translationally, similar to the kidney type glutaminase in mammals (Curthoys and Watford, 1995), as noted by the presence of phosphorylation sites (S298, S300) on the C-terminal extension of the protein sequence, that are identified using PhosPhAt 4.0 (Roitinger et al., 2015). Further investigation is necessary to identify the mechanism by which GAT1_2.1 is post-translationally regulated.

The loss of function mutant, *gat1_2.1*, of *Arabidopsis* displays an excessive shoot branching phenotype. Understanding the exact mechanism through which GAT1_2.1 regulates shoots branching, requires further experimentation. However, due to its root specific expression (discussed in Chapter 3), we hypothesized that the phenotype in shoots is an indirect consequence of metabolic perturbations leading to C/N imbalance in the roots. To that end, we decided to take a metabolomics approach with stable isotope labelling to track the mechanism of C/N partitioning in roots of *gat1_2.1* compared to wild-type *Arabidopsis*. The lack of standardized and reliable methods for metabolomics,

however, led us to pursue and establish optimum protocols for stable isotope labelling followed by metabolomics measurements and data analysis suitable for detection and quantification of at least the primary metabolites in *Arabidopsis* (Chapter 4). We demonstrated the use of an optimized workflow for untargeted metabolomics experiments upon isotopic label enrichment by taking advantage of high mass accuracy and resolving power of Q-Exactive Orbitrap mass spectrometer.

5.3 A new and improved pipeline for LC-MS based metabolomics with a focus on primary metabolism

Any untargeted metabolomics experiment consists of three major steps: 1) extraction of metabolites, 2) chromatographic separation and mass spectrometric measurement, and 3) data analysis. The complexity of the metabolome in plant tissue, the availability of a wide array of equipment and unreliable data analysis platforms (Jorge et al., 2016) calls for the optimization of protocols specific to the experimental design at hand. Given that the primary metabolites are our main focus, we evaluated a few extraction procedures to determine the best condition suitable for mass spectrometric measurements and showed that methanol:water (4:1, v/v) extraction yields best results.

To address the second issue of chromatographic separation, we demonstrated an optimal gradient for HILIC partitioning with reliable separation of amino acids and with shorter run times compared to GC or other LC based methods. The HILIC method employs a zwitterionic bonded silica column and enables separation of hydrophilic compounds without prior derivatization and is compatible with mass spectrometric detection (Langrock et al., 2006). Here, we achieved better separation and shorter run time as demonstrated with amino acids (Figure 4.5) by taking advantage of the HILIC water partitioning mode between 10-30% A in the mobile phase. A combination of packages in R, that include IPO (Libiseller et al., 2015), XCMS (Smith et al., 2006) and McMillan correction (McMillan et al., 2016) were used for data analysis and optimal feature list detection. It is however important to remember that compound identification is still a major bottleneck in metabolomics studies, and hence a simultaneous data dependent or targeted acquisition of MS/MS fragmentation is necessary for success in metabolomics experiments. The user also requires a thorough understanding of the “seven golden rules

of heuristic filtering of molecular formulas obtained by accurate mass spectrometry," that were reported by Kind and Fiehn (2007) for efficient compound identification. Successful adoption of this pipeline has been demonstrated in at least one published study, which is focused on identifying changes in primary metabolism using untargeted metabolomics (Chen et al., 2017).

5.4 Use of isotopic label incorporation followed by HRMS for novel pathway identification

The use of isotopic label incorporation by treatment with ^{13}C or ^{15}N containing compounds followed by tracking the isotopic signature has been previously demonstrated to be extremely useful in the discovery of novel pathways and the quantification of metabolic flux. However, several limitations exist in the methods previously adopted, including the use of targeted metabolite analysis (which requires knowledge of the metabolites involved) via GC or LC followed by mass spectrometry, use of only a single labelled compound or resorting to laborious techniques, such as NMR (Allen et al., 2009; Giavalisco et al., 2011; Kikuchi and Hirayama, 2007). To address this issue, at least to a certain extent, we resorted to HILIC chromatography (described above) followed by high resolution mass spectrometry (HRMS), taking advantage of the mass accuracy along with the resolving power of Q-Exactive Orbitrap mass spectrometer, and demonstrated a successful incorporation of dual labels ($^{13}C_5$, $^{15}N_2$) supplied to wild-type *Arabidopsis*, in the form of glutamine followed by identification of compounds in the primary metabolic pathways described in Chapters 2 and 3 (section 4.2.4). Using this method, isotopic signatures can be extracted from an untargeted metabolomics experiment with an in-house python script and hence unknown pathways can be successfully identified.

5.5 Future prospects in understanding N metabolism and improving NUE in plants

One of the goals of near future is to use the metabolomics and labelling strategies developed for deciphering the role of GAT1_2.1 in C/N partitioning as well as its physiological consequence (i.e. the shoot branching phenotype). In the post-genomic era, metabolomics has emerged into a powerful tool for elucidating complex phenotypes.

Establishing pathways and validation of genes involved in these pathways holds great potential for improving our understanding of nitrogen metabolism and signaling in general. Considering the continuously inclining statistics for nitrogen fertilizer use/misuse and the requirement for increased Nitrogen Use Efficiency (NUE) to achieve sustainable yields and reduce fertilizer consumption, a thorough understanding of N metabolism is of utmost importance. The future holds the possibility of using modern technologies such as gene editing to target genes involved in N metabolism for achieving said NUE and hence sustainable crop production.

5.6 Literature cited

Allen, D.K., Libourel, I.G.L., and Shachar-Hill, Y. (2009). Metabolic flux analysis in plants: coping with complexity. Plant, Cell & Environment 32:1241-1257.

Boex-Fontvieille, E.R., Gauthier, P.P., Gilard, F., Hodges, M., and Tcherkez, G.G. (2013). A new anaplerotic respiratory pathway involving lysine biosynthesis in isocitrate dehydrogenase-deficient Arabidopsis mutants. New Phytologist 199:673-682.

Chen, C., Li, C., Wang, Y., Renaud, J., Tian, G., Kambhampati, S., Saatian, B., Nguyen, V., Hannoufa, A., Marsolais, F., et al. (2017). Cytosolic acetyl-CoA promotes histone acetylation predominantly at H3K27 in Arabidopsis. Nature Plants.

Coschigano, K.T., Melo-Oliveira, R., Lim, J., and Coruzzi, G.M. (1998). *Arabidopsis* mutants and distinct Fd-GOGAT genes: implications for photorespiration and primary nitrogen assimilation. The Plant cell 10:741.

Curthoys, N.P., and Watford, M. (1995). Regulation of glutaminase activity and glutamine metabolism. Annual Review of Nutrition 15:133-159.

Erickson, J.W., and Cerione, R.A. (2010). Glutaminase: a hot spot for regulation of cancer cell metabolism? Oncotarget 1:734-740.

Fait, A., Fromm, H., Walter, D., Galili, G., and Fernie, A.R. (2008). Highway or byway: the metabolic role of the GABA shunt in plants. Trends in Plant Science 13:14-19.

Gaufichon, L., Masclaux-Daubresse, C., Tcherkez, G., Reisdorf-Cren, M., Sakakibara, Y., Hase, T., ClÉMent, G., Avice, J.-C., Grandjean, O., Marmagne, A., et al. (2013). *Arabidopsis thaliana* ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. Plant, Cell & Environment 36:328-342.

Giavalisco, P., Li, Y., Matthes, A., Eckhardt, A., Hubberten, H.M., Hesse, H., Segu, S., Hummel, J., Kohl, K., and Willmitzer, L. (2011). Elemental formula annotation of polar and lipophilic metabolites using (13) C, (15) N and (34) S isotope labelling, in combination with high-resolution mass spectrometry. The Plant journal : for cell and molecular biology 68:364-376.

Gutiérrez, R.A., Stokes, T.L., Thum, K., Xu, X., Obertello, M., Katari, M.S., Tanurdzic, M., Dean, A., Nero, D.C., McClung, C.R., et al. (2008). Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. Proc Natl Acad Sci U S A 105:4939-4944.

Jorge, T.F., Rodrigues, J.A., Caldana, C., Schmidt, R., van Dongen, J.T., Thomas-Oates, J., and António, C. (2016). Mass spectrometry-based plant metabolomics: Metabolite responses to abiotic stress. Mass Spectrometry Reviews 35:620-649.

Kikuchi, J., and Hirayama, T. (2007). Practical aspects of uniform stable isotope labeling of higher plants for heteronuclear NMR-based metabolomics. In: Metabolomics: Methods and Protocols--Weckwerth, W., ed. Totowa, NJ: Humana Press. 273-286.

Kind, T., and Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. BMC Bioinformatics 8:105.

Lea, P.J., Sodek, L., Parry, M.A.J., Shewry, P.R., and Halford, N.G. (2007). Asparagine in plants. Annals of Applied Biology 150:1-26.

Li, J.Y., Fu, Y.L., Pike, S.M., Bao, J., Tian, W., Zhang, Y., Chen, C.Z., Li, H.M., Huang, J., Li, L.G., et al. (2010). The *Arabidopsis* nitrate transporter NRT1.8 functions in nitrate removal from the xylem sap and mediates cadmium tolerance. Plant Cell 22:1633-1646.

Libiseller, G., Dvorzak, M., Kleb, U., Gander, E., Eisenberg, T., Madeo, F., Neumann, S., Trausinger, G., Sinner, F., Pieber, T., et al. (2015). IPO: a tool for automated optimization of XCMS parameters. BMC Bioinformatics 16:118.

Lima, J.D., and Sodek, L. (2003). N-stress alters aspartate and asparagine levels of xylem sap in soybean. Plant Science 165:649-656.

McMillan, A., Renaud, J.B., Gloor, G.B., Reid, G., and Sumarah, M.W. (2016). Post-acquisition filtering of salt cluster artefacts for LC-MS based human metabolomic studies. Journal of Cheminformatics 8:44.

Oliveira, I.C., and Coruzzi, G.M. (1999). Carbon and amino acids reciprocally modulate the expression of glutamine synthetase in *Arabidopsis*. Plant physiology 121:301.

Roitinger, E., Hofer, M., Köcher, T., Pichler, P., Novatchkova, M., Yang, J., Schlögelhofer, P., and Mechtler, K. (2015). Quantitative phosphoproteomics of the Ataxia Telangiectasia-Mutated (ATM) and Ataxia Telangiectasia-Mutated and Rad3-related (ATR) dependent DNA damage response in *Arabidopsis thaliana*. Molecular & Cellular Proteomics 14:556-571.

Sakakibara, H., Takei, K., and Hirose, N. (2006). Interactions between nitrogen and cytokinin in the regulation of metabolism and development. Trends in Plant Science 11:440-448.

Smith, C., Want, E., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching and identification. Analytical Chemistry 78.

Stitt, M., Müller, C., Matt, P., Gibon, Y., Carillo, P., Morcuende, R., Scheible, W.R., and Krapp, A. (2002). Steps towards an integrated view of nitrogen metabolism. Journal of Experimental Botany 53:959-970.

Tsai, F.Y., and Coruzzi, G.M. (1990). Dark-induced and organ-specific expression of two asparagine synthetase genes in *Pisum sativum*. The EMBO Journal 9:323-332.

Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B. (2009). Understanding the warburg effect: the metabolic requirements of cell proliferation. Science (New York, N.Y.) 324:1029-1033.

Zhu, H., and Kranz, R.G. (2012). A nitrogen-regulated glutamine amidotransferase (gat1_2.1) represses shoot branching in *Arabidopsis*. Plant physiology 160:1770-1780.

APPENDICES

Appendix A: Differentially expressed gene list FDR < 0.001

	TAIR ID	Gene Symbol	Gene description
Down-regulated	AT1G08440	AT1G08440	aluminum activated malate transporter family protein
	AT1G24880	LpxC2	UDP-3-O-acyl N-acetylglycosamine deacetylase family protein
	AT2G26150	HSFA2	heat shock transcription factor A2
	AT4G21680	NRT1.8	NITRATE TRANSPORTER 1.8
	AT4G22485	AT4G22485	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT3G10040	AT3G10040	sequence-specific DNA binding transcription factor
	AT3G45060	NRT2.6	high affinity nitrate transporter 2.6
	AT3G46130	MYB48	myb domain protein 48
	AT3G28945	AT3G28945	transposable_element_gene
	AT2G43890	AT2G43890	Pectin lyase-like superfamily protein
	AT5G24660	LSU2	response to low sulfur 2
	AT5G65207	AT5G65207	hypothetical protein
	AT1G32450	NRT1.5	nitrate transporter 1.5
Up-regulated	AT3G44260	CAF1a	Polynucleotidyl transferase
	AT3G59930	AT3G59930	defensin-like protein
	AT4G17500	ERF-1	ethylene responsive element binding factor 1
	AT2G24580	AT2G24580	FAD-dependent oxidoreductase family protein
	AT5G66985	AT5G66985	hypothetical protein
	AT1G05700	AT1G05700	Leucine-rich repeat transmembrane protein kinase protein
	AT3G63110	IPT3	isopentenyltransferase 3
	AT1G23760	JP630	BURP domain-containing protein

	AT3G54040	AT3G54040	PAR1 protein
	AT1G27730	STZ	salt tolerance zinc finger
	AT2G30040	MAPKKK14	mitogen-activated protein kinase kinase kinase 14
	AT4G24570	DIC2	dicarboxylate carrier 2
	AT2G02680	AT2G02680	Cysteine/Histidine-rich C1 domain family protein
	AT2G24762	GDU4	glutamine dumper 4
	AT1G60960	IRT3	iron regulated transporter 3
	AT3G53730	AT3G53730	Histone superfamily protein
	AT3G16330	AT3G16330	Avr9/Cf-9 rapidly elicited protein
	AT5G01810	CIPK15	CBL-interacting protein kinase 15
	AT4G29780	AT4G29780	nuclease
	AT5G01830	SAUR21	ARM repeat superfamily protein
	AT5G40730	AGP24	arabinogalactan protein 24
	AT5G25940	AT5G25940	early nodulin-like protein
	AT5G01040	LAC8	laccase 8
	AT4G22470	AT4G22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	AT2G44370	AT2G44370	Cysteine/Histidine-rich C1 domain family protein
	AT4G32480	AT4G32480	sugar phosphate exchanger
	AT3G12700	NANA	Eukaryotic aspartyl protease family protein
	AT5G14330	AT5G14330	transmembrane protein
	AT2G07709	AT2G07709	pseudogene
	AT1G80840	WRKY40	WRKY DNA-binding protein 40
	AT2G39530	AT2G39530	Uncharacterized protein family (UPF0497)
	AT1G19380	AT1G19380	sugar
	AT2G24570	WRKY17	WRKY DNA-binding protein 17
	AT1G79850	RPS17	ribosomal protein S17

	AT1G30760	AT1G30760	FAD-binding Berberine family protein
	AT3G07390	AIR12	auxin-induced in root cultures-like protein
	AT2G28720	AT2G28720	Histone superfamily protein
	AT5G04310	AT5G04310	Pectin lyase-like superfamily protein
	AT1G25560	TEM1	AP2/B3 transcription factor family protein
	AT1G49500	AT1G49500	transcription initiation factor TFIID subunit 1b-like protein
	AT1G10970	ZIP4	zinc transporter
	AT1G12090	ELP	extensin-like protein
	AT1G22550	AT1G22550	Major facilitator superfamily protein
	AT4G35480	RHA3B	RING-H2 finger A3B
	AT1G10960	FD1	ferredoxin 1
	AT5G09440	EXL4	EXORDIUM like 4
	AT5G48430	AT5G48430	Eukaryotic aspartyl protease family protein
	AT2G40330	PYL6	PYR1-like 6
	AT3G26470	AT3G26470	Powdery mildew resistance protein
	AT3G16570	RALF23	rapid alkalinization factor 23
	AT1G02360	AT1G02360	Chitinase family protein
	AT2G19060	AT2G19060	SGNH hydrolase-type esterase superfamily protein
	AT1G25400	AT1G25400	transmembrane protein
	AT3G13610	AT3G13610	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
	AT4G11210	AT4G11210	Disease resistance-responsive (dirigent-like protein) family protein
	AT3G04880	DRT102	DNA-damage-repair/toleration protein (DRT102)
	AT4G39940	AKN2	APS-kinase 2
	AT4G22070	WRKY31	WRKY DNA-binding protein 31
	AT1G80180	AT1G80180	hypothetical protein

	AT1G68880	bZIP	basic leucine-zipper 8
	AT2G38530	LTP2	lipid transfer protein 2
	AT3G15210	ERF4	ethylene responsive element binding factor 4
	AT3G54420	EP3	homolog of carrot EP3-3 chitinase
	AT5G62865	AT5G62865	hypothetical protein
	AT4G15800	RALFL33	ralf-like 33
	AT1G80450	AT1G80450	VQ motif-containing protein
	AT5G66200	ARO2	armadillo repeat only 2
	AT2G37750	AT2G37750	hypothetical protein
	AT1G71880	SUC1	sucrose-proton symporter 1
	AT3G09220	LAC7	laccase 7
	AT5G07110	PRA1.B6	prenylated RAB acceptor 1.B6
	AT4G12480	EARLI1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT3G01290	HIR2	SPFH/Band 7/PHB domain-containing membrane-associated protein family
	AT2G26370	AT2G26370	MD-2-related lipid recognition domain-containing protein
	AT2G42360	AT2G42360	RING/U-box superfamily protein
	AT5G65690	PCK2	phosphoenolpyruvate carboxykinase 2
	AT1G15670	AT1G15670	Galactose oxidase/kelch repeat superfamily protein
	AT2G23620	MES1	methyl esterase 1
	AT4G12470	AZI1	azelaic acid induced 1
	AT5G07030	AT5G07030	Eukaryotic aspartyl protease family protein
	AT5G05440	PYL5	Polyketide cyclase/dehydrase and lipid transport superfamily protein
	AT2G30620	AT2G30620	winged-helix DNA-binding transcription factor family protein
	AT3G16240	DELTA-TIP	delta tonoplast integral protein

	AT2G48080	AT2G48080	oxidoreductase
	AT5G22920	AT5G22920	CHY-type/CTCHY-type/RING-type Zinc finger protein
	AT1G14900	HMGA	high mobility group A
	AT5G43520	AT5G43520	Cysteine/Histidine-rich C1 domain family protein
	AT4G31550	WRKY11	WRKY DNA-binding protein 11
	AT5G41080	GDPD2	PLC-like phosphodiesterases superfamily protein
	AT1G78830	AT1G78830	Curculin-like (mannose-binding) lectin family protein
	AT2G38310	PYL4	PYR1-like 4
	AT3G45960	EXLA3	expansin-like A3
	AT4G25810	XTR6	xyloglucan endotransglycosylase 6
	AT4G28940	AT4G28940	Phosphorylase superfamily protein
	AT5G55050	AT5G55050	GDSL-like Lipase/Acylhydrolase superfamily protein
	AT5G57560	TCH4	Xyloglucan endotransglucosylase/hydrolase family protein
	AT3G63380	AT3G63380	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
	AT1G14960	AT1G14960	Polyketide cyclase/dehydrase and lipid transport superfamily protein
	AT4G38470	STY46	ACT-like protein tyrosine kinase family protein
	AT2G48130	AT2G48130	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT1G80240	DGR1	choice-of-anchor C domain protein
	AT4G17340	TIP2;2	tonoplast intrinsic protein 2;2
	AT4G01250	WRKY22	WRKY family transcription factor
	AT3G45840	AT3G45840	Cysteine/Histidine-rich C1 domain family protein
	AT3G46280	AT3G46280	kinase-like protein
	AT4G08950	EXO	Phosphate-responsive 1 family protein
	AT2G30930	AT2G30930	hypothetical protein

	AT1G67110	CYP735A2	cytochrome P450
	AT5G01050	AT5G01050	Laccase/Diphenol oxidase family protein
	AT5G42510	AT5G42510	Disease resistance-responsive (dirigent-like protein) family protein
	AT3G52420	OEP7	outer envelope membrane protein 7
	AT5G38710	AT5G38710	Methylenetetrahydrofolate reductase family protein
	AT3G50350	AT3G50350	membrane insertase
	AT1G62510	AT1G62510	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT4G30270	XTH24	xyloglucan endotransglucosylase/hydrolase 24
	AT1G73260	KTI1	kunitz trypsin inhibitor 1
	AT5G12340	AT5G12340	DUF4228 domain protein
	AT5G19120	AT5G19120	Eukaryotic aspartyl protease family protein
	AT4G38340	AT4G38340	Plant regulator RWP-RK family protein
	AT3G12500	HCHIB	basic chitinase
	AT1G62980	EXPA18	expansin A18
	AT5G18840	AT5G18840	Major facilitator superfamily protein
	AT1G06760	AT1G06760	winged-helix DNA-binding transcription factor family protein
	AT4G35770	SEN1	Rhodanese/Cell cycle control phosphatase superfamily protein
	AT2G28860	CYP710A4	cytochrome P450
	ATMG00020	RRN26	rRNA
	AT2G44380	AT2G44380	Cysteine/Histidine-rich C1 domain family protein
	AT5G58650	PSY1	plant peptide containing sulfated tyrosine 1
	AT3G27270	AT3G27270	TRAM
	AT1G53540	AT1G53540	HSP20-like chaperones superfamily protein
	ATMG00660	ORF149	hypothetical protein
	AT5G20250	DIN10	Raffinose synthase family protein

	AT5G44130	FLA13	FASCICLIN-like arabinogalactan protein 13 precursor
	ATMG00030	ORF107A	hypothetical protein
	AT4G06477	AT4G06477	transposable_element_gene
	AT5G64110	AT5G64110	Peroxidase superfamily protein
	AT4G10500	AT4G10500	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
	AT3G30775	ERD5	Methylenetetrahydrofolate reductase family protein
	AT4G26320	AGP13	arabinogalactan protein 13
	AT2G20670	AT2G20670	sugar phosphate exchanger
	AT5G64120	AT5G64120	Peroxidase superfamily protein
	AT5G57625	AT5G57625	CAP (Cysteine-rich secretory proteins
	AT1G35140	PHI-1	Phosphate-responsive 1 family protein
	AT5G50760	AT5G50760	SAUR-like auxin-responsive protein family
	AT2G19800	MIOX2	myo-inositol oxygenase 2
	ATCG00020	PSBA	photosystem II reaction center protein A
	AT3G21351	AT3G21351	transmembrane protein
	AT2G39510	UMAMIT14	nodulin MtN21 /EamA-like transporter family protein
	AT3G41768	AT3G41768	rRNA
	AT5G22890	AT5G22890	C2H2 and C2HC zinc fingers superfamily protein
	AT4G25470	CBF2	C-repeat/DRE binding factor 2
	AT1G75030	TLP-3	thaumatin-like protein 3
	AT1G15040	GAT1_2.1	Class I glutamine amidotransferase-like superfamily protein
	AT3G27690	LHCB2.3	photosystem II light harvesting complex protein 2.3
	AT2G21040	AT2G21040	Calcium-dependent lipid-binding (CaLB domain) family protein
	AT2G01021	AT2G01021	hypothetical protein

	AT2G31141	AT2G31141	hypothetical protein
	AT1G18280	AT1G18280	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT1G08115	U1-5	U1
	ATMG01380	RRN5	rRNA
	AT1G15405	NA	NA
	AT5G44120	CRA1	RmlC-like cupins superfamily protein
	AT1G48130	PER1	1-cysteine peroxiredoxin 1
	AT4G28520	CRU3	cruciferin 3

Appendix B: Differentially expressed gene list FDR < 0.01

	TAIR ID	Gene Symbol	Gene description
Down-regulated	AT1G08440	AT1G08440	aluminum activated malate transporter family protein
	AT1G24880	LpxC2	UDP-3-O-acyl N-acetylglycosamine deacetylase family protein
	AT2G26150	HSFA2	heat shock transcription factor A2
	AT4G21680	NRT1.8	NITRATE TRANSPORTER 1.8
	AT4G22485	AT4G22485	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT3G10040	AT3G10040	sequence-specific DNA binding transcription factor
	AT3G45060	NRT2.6	high affinity nitrate transporter 2.6
	AT4G22505	AT4G22505	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT4G00910	AT4G00910	aluminum activated malate transporter family protein
	AT3G46130	MYB48	myb domain protein 48
	AT3G28945	AT3G28945	transposable_element_gene
	AT2G43890	AT2G43890	Pectin lyase-like superfamily protein
	AT5G24660	LSU2	response to low sulfur 2
	AT5G65207	AT5G65207	hypothetical protein
	AT2G45860	AT2G45860	hypothetical protein
	AT4G17670	AT4G17670	senescence-associated family protein (DUF581)
	AT5G58770	cPT4	Undecaprenyl pyrophosphate synthetase family protein
	AT2G33480	NAC041	NAC domain containing protein 41
	AT1G32450	NRT1.5	nitrate transporter 1.5
	AT5G60520	AT5G60520	Late embryogenesis abundant (LEA) protein-like protein
	AT3G25930	AT3G25930	Adenine nucleotide alpha hydrolases-like superfamily protein
	AT4G39190	AT4G39190	nucleolar-like protein

	AT2G16660	AT2G16660	Major facilitator superfamily protein
	AT1G77580	AT1G77580	filament-like protein (DUF869)
	AT1G17180	GSTU25	glutathione S-transferase TAU 25
	AT5G27950	AT5G27950	P-loop containing nucleoside triphosphate hydrolases superfamily protein
	AT1G78340	GSTU22	glutathione S-transferase TAU 22
	AT4G10450	AT4G10450	Ribosomal protein L6 family
	AT1G60470	GolS4	galactinol synthase 4
	AT2G30750	CYP71A12	cytochrome P450 family 71 polypeptide
Up-regulated	AT3G44260	CAF1a	Polynucleotidyl transferase
	AT3G19030	AT3G19030	transcription initiation factor TFIID subunit 1b-like protein
	AT3G49940	LBD38	LOB domain-containing protein 38
	AT3G59930	AT3G59930	defensin-like protein
	AT2G35270	GIK	Putative AT-hook DNA-binding family protein
	AT1G27020	AT1G27020	plant/protein
	AT2G21020	AT2G21020	pseudogene
	AT1G18140	LAC1	laccase 1
	AT4G17500	ERF-1	ethylene responsive element binding factor 1
	AT3G20110	CYP705A20	cytochrome P450
	AT2G24580	AT2G24580	FAD-dependent oxidoreductase family protein
	AT5G66985	AT5G66985	hypothetical protein
	AT1G05700	AT1G05700	Leucine-rich repeat transmembrane protein kinase protein
	AT5G10430	AGP4	arabinogalactan protein 4
	AT4G31730	GDU1	glutamine dumper 1
	AT4G16370	OPT3	oligopeptide transporter
	AT3G13310	AT3G13310	Chaperone DnaJ-domain superfamily protein

	AT3G63110	IPT3	isopentenyltransferase 3
	AT5G49700	AT5G49700	Putative AT-hook DNA-binding family protein
	AT4G14040	SBP2	selenium-binding protein 2
	AT1G20390	AT1G20390	transposable_element_gene
	AT2G48140	EDA4	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT1G74010	AT1G74010	Calcium-dependent phosphotriesterase superfamily protein
	AT1G23760	JP630	BURP domain-containing protein
	AT2G01190	PDE331	Octicosapeptide/Phox/Bem1p family protein
	AT3G12750	ZIP1	zinc transporter 1 precursor
	AT2G47840	Tic20-II	Uncharacterized conserved protein ycf60
	AT1G79870	AT1G79870	D-isomer specific 2-hydroxyacid dehydrogenase family protein
	AT1G13260	RAV1	related to ABI3/VP1 1
	AT5G16370	AAE5	acyl activating enzyme 5
	AT1G80820	CCR2	cinnamoyl coa reductase
	AT5G42100	BG_PPAP	beta-1
	AT4G25760	GDU2	glutamine dumper 2
	AT3G54040	AT3G54040	PAR1 protein
	AT1G27730	STZ	salt tolerance zinc finger
	AT2G30040	MAPKKK14	mitogen-activated protein kinase kinase kinase 14
	AT2G34080	AT2G34080	Cysteine proteinases superfamily protein
	AT4G24570	DIC2	dicarboxylate carrier 2
	AT5G64310	AGP1	arabinogalactan protein 1
	AT1G10480	ZFP5	zinc finger protein 5
	AT2G38110	GPAT6	glycerol-3-phosphate acyltransferase 6
	AT5G47450	TIP2;3	tonoplast intrinsic protein 2;3

	AT4G18340	AT4G18340	Glycosyl hydrolase superfamily protein
	AT5G15960	KIN1	stress-responsive protein (KIN1) / stress-induced protein (KIN1)
	AT2G02680	AT2G02680	Cysteine/Histidine-rich C1 domain family protein
	AT3G51340	AT3G51340	Eukaryotic aspartyl protease family protein
	AT2G24762	GDU4	glutamine dumper 4
	AT5G63160	BT1	BTB and TAZ domain protein 1
	AT3G04720	PR4	pathogenesis-related 4
	AT1G60960	IRT3	iron regulated transporter 3
	AT3G53730	AT3G53730	Histone superfamily protein
	AT5G20650	COPT5	copper transporter 5
	AT2G37180	RD28	Aquaporin-like superfamily protein
	AT3G16330	AT3G16330	Avr9/Cf-9 rapidly elicited protein
	AT2G47130	SDR3	NAD(P)-binding Rossmann-fold superfamily protein
	AT1G14890	AT1G14890	Plant invertase/pectin methylesterase inhibitor superfamily protein
	AT1G20225	AT1G20225	Thioredoxin superfamily protein
	AT2G38940	PHT1;4	phosphate transporter 1;4
	AT2G48030	AT2G48030	DNAse I-like superfamily protein
	AT4G28990	AT4G28990	RNA-binding protein-like protein
	AT5G01810	CIPK15	CBL-interacting protein kinase 15
	AT2G37980	AT2G37980	O-fucosyltransferase family protein
	AT1G10550	XTH33	xyloglucan:xyloglucosyl transferase 33
	AT4G29780	AT4G29780	nuclease
	AT5G01830	SAUR21	ARM repeat superfamily protein
	AT5G40730	AGP24	arabinogalactan protein 24
	AT4G20390	AT4G20390	Uncharacterized protein family (UPF0497)
	AT5G25940	AT5G25940	early nodulin-like protein

	AT5G01040	LAC8	laccase 8
	AT4G08685	SAH7	Pollen Ole e 1 allergen and extensin family protein
	AT4G22470	AT4G22470	protease inhibitor/seed storage/lipid transfer protein (LTP) family protein
	AT2G34490	CYP710A2	cytochrome P450
	AT2G44370	AT2G44370	Cysteine/Histidine-rich C1 domain family protein
	AT5G24230	AT5G24230	Lipase class 3-related protein
	AT2G07717	AT2G07717	pseudogene
	AT1G80380	AT1G80380	P-loop containing nucleoside triphosphate hydrolases superfamily protein
	AT5G63510	GAMMA CAL1	gamma carbonic anhydrase like 1
	AT5G60660	PIP2;4	plasma membrane intrinsic protein 2;4
	AT2G35710	PGSIP7	Nucleotide-diphospho-sugar transferases superfamily protein
	AT2G38860	YLS5	Class I glutamine amidotransferase-like superfamily protein
	AT4G32480	AT4G32480	sugar phosphate exchanger
	AT3G12700	NANA	Eukaryotic aspartyl protease family protein
	AT4G25110	MC2	metacaspase 2
	AT1G08880	H2AXA	Histone superfamily protein
	AT4G30660	AT4G30660	Low temperature and salt responsive protein family
	AT5G14330	AT5G14330	transmembrane protein
	AT2G39710	AT2G39710	Eukaryotic aspartyl protease family protein
	AT4G35490	MRPL11	mitochondrial ribosomal protein L11
	AT2G07709	AT2G07709	pseudogene
	AT3G25717	RTFL16	ROTUNDIFOLIA like 16
	AT3G44380	AT3G44380	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
	AT4G32950	AT4G32950	Protein phosphatase 2C family protein

	AT1G64660	MGL	methionine gamma-lyase
	AT3G03150	AT3G03150	hypothetical protein
	AT3G26760	AT3G26760	NAD(P)-binding Rossmann-fold superfamily protein
	AT2G44790	UCC2	uclacyanin 2
	AT1G13245	RTFL17	ROTUNDIFOLIA like 17
	AT2G40590	AT2G40590	Ribosomal protein S26e family protein
	AT1G80840	WRKY40	WRKY DNA-binding protein 40
	AT2G39530	AT2G39530	Uncharacterized protein family (UPF0497)
	AT5G59680	AT5G59680	Leucine-rich repeat protein kinase family protein
	AT1G19380	AT1G19380	sugar
	AT2G24570	WRKY17	WRKY DNA-binding protein 17
	AT1G79850	RPS17	ribosomal protein S17
	AT1G30760	AT1G30760	FAD-binding Berberine family protein
	AT3G07390	AIR12	auxin-induced in root cultures-like protein
	AT2G28720	AT2G28720	Histone superfamily protein
	AT1G72230	AT1G72230	Cupredoxin superfamily protein
	AT2G01410	AT2G01410	NHL domain-containing protein
	AT5G04310	AT5G04310	Pectin lyase-like superfamily protein
	AT4G21120	AAT1	amino acid transporter 1
	AT1G25560	TEM1	AP2/B3 transcription factor family protein
	AT4G08300	UMAMIT17	nodulin MtN21 /EamA-like transporter family protein
	AT2G21100	AT2G21100	Disease resistance-responsive (dirigent-like protein) family protein
	AT1G49500	AT1G49500	transcription initiation factor TFIID subunit 1b-like protein
	AT5G54640	RAT5	Histone superfamily protein
	AT1G10970	ZIP4	zinc transporter

	AT1G12090	ELP	extensin-like protein
	AT1G22550	AT1G22550	Major facilitator superfamily protein
	AT4G35480	RHA3B	RING-H2 finger A3B
	AT2G20820	AT2G20820	hypothetical protein
	AT3G11370	AT3G11370	Cysteine/Histidine-rich C1 domain family protein
	AT1G22500	ATL15	RING/U-box superfamily protein
	AT5G23220	NIC3	nicotinamidase 3
	AT2G30210	LAC3	laccase 3
	AT1G10960	FD1	ferredoxin 1
	AT5G09440	EXL4	EXORDIUM like 4
	AT2G01080	AT2G01080	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
	AT5G48430	AT5G48430	Eukaryotic aspartyl protease family protein
	AT2G40330	PYL6	PYR1-like 6
	AT3G26470	AT3G26470	Powdery mildew resistance protein
	AT5G10980	AT5G10980	Histone superfamily protein
	AT5G04770	CAT6	cationic amino acid transporter 6
	AT2G40230	AT2G40230	HXXXD-type acyl-transferase family protein
	AT3G16570	RALF23	rapid alkalinization factor 23
	AT5G42500	AT5G42500	Disease resistance-responsive (dirigent-like protein) family protein
	AT1G02360	AT1G02360	Chitinase family protein
	AT3G22160	AT3G22160	VQ motif-containing protein
	AT2G19060	AT2G19060	SGNH hydrolase-type esterase superfamily protein
	AT1G25400	AT1G25400	transmembrane protein
	AT3G13610	AT3G13610	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
	AT4G11210	AT4G11210	Disease resistance-responsive (dirigent-like protein) family protein

	AT3G04880	DRT102	DNA-damage-repair/toleration protein (DRT102)
	AT4G39940	AKN2	APS-kinase 2
	AT1G68620	AT1G68620	alpha/beta-Hydrolases superfamily protein
	AT4G22070	WRKY31	WRKY DNA-binding protein 31
	AT1G80180	AT1G80180	hypothetical protein
	AT4G18510	CLE2	CLAVATA3/ESR-related 2
	AT1G68880	bZIP	basic leucine-zipper 8
	AT2G38530	LTP2	lipid transfer protein 2
	AT3G15210	ERF4	ethylene responsive element binding factor 4
	AT3G54420	EP3	homolog of carrot EP3-3 chitinase
	AT2G35750	AT2G35750	transmembrane protein
	AT5G62865	AT5G62865	hypothetical protein
	AT5G17820	AT5G17820	Peroxidase superfamily protein
	AT4G15800	RALFL33	ralf-like 33
	AT2G41380	AT2G41380	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
	AT4G15150	AT4G15150	glycine-rich protein
	AT1G80450	AT1G80450	VQ motif-containing protein
	AT2G34500	CYP710A1	cytochrome P450
	AT5G22355	AT5G22355	Cysteine/Histidine-rich C1 domain family protein
	AT5G66200	ARO2	armadillo repeat only 2
	AT2G37750	AT2G37750	hypothetical protein
	AT1G71880	SUC1	sucrose-proton symporter 1
	AT4G28840	AT4G28840	mediator of RNA polymerase II transcription subunit
	AT3G09220	LAC7	laccase 7
	AT1G23205	AT1G23205	Plant invertase/pectin methylesterase inhibitor superfamily protein
	AT5G07110	PRA1.B6	prenylated RAB acceptor 1.B6

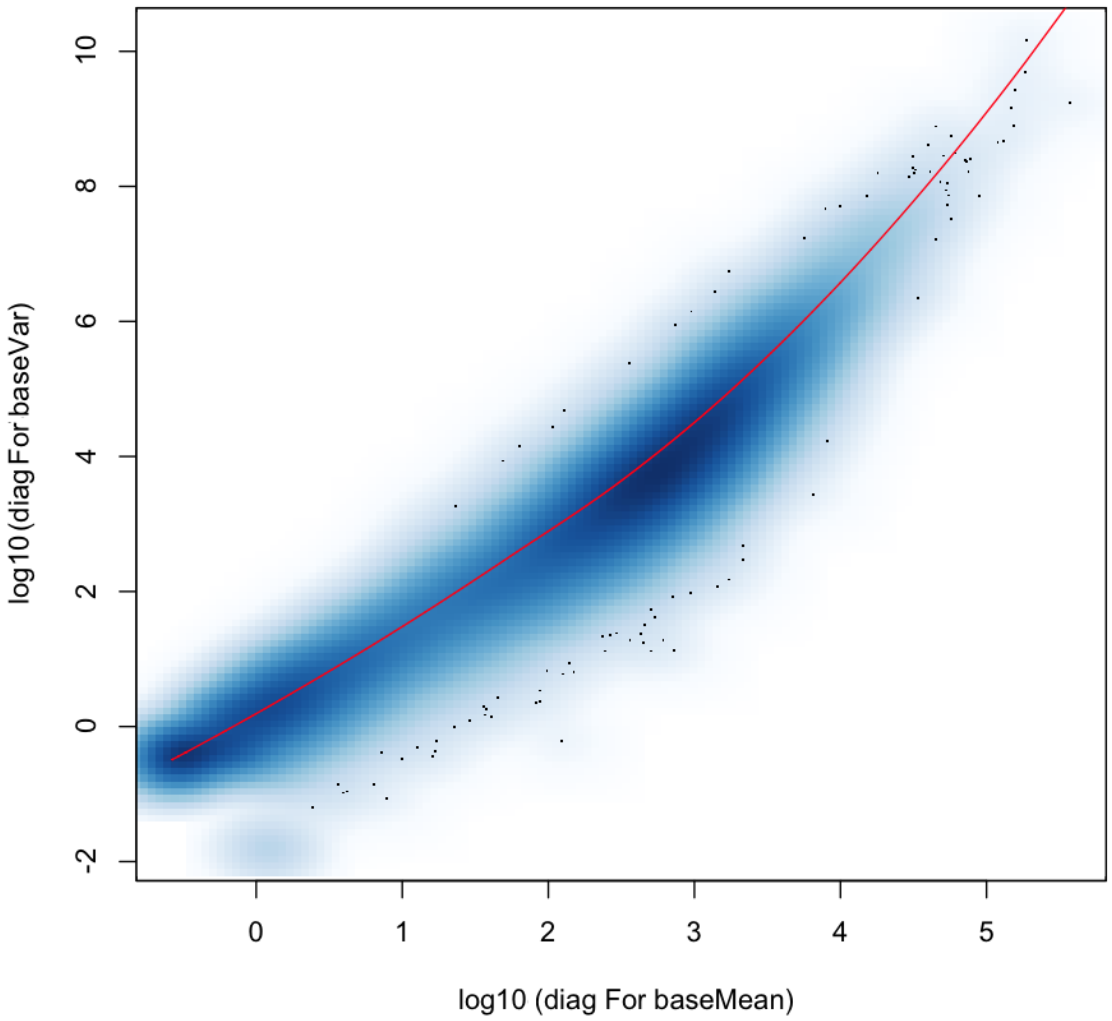
	AT4G12480	EARLI1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT4G29740	CKX4	cytokinin oxidase 4
	AT3G01290	HIR2	SPFH/Band 7/PHB domain-containing membrane-associated protein family
	AT2G26370	AT2G26370	MD-2-related lipid recognition domain-containing protein
	AT3G04530	PPCK2	phosphoenolpyruvate carboxylase kinase 2
	AT2G42360	AT2G42360	RING/U-box superfamily protein
	AT5G65690	PCK2	phosphoenolpyruvate carboxykinase 2
	AT1G15670	AT1G15670	Galactose oxidase/kelch repeat superfamily protein
	AT2G23620	MES1	methyl esterase 1
	AT4G12470	AZI1	azelaic acid induced 1
	AT5G07030	AT5G07030	Eukaryotic aspartyl protease family protein
	AT5G05440	PYL5	Polyketide cyclase/dehydrase and lipid transport superfamily protein
	AT1G54740	AT1G54740	FANTASTIC four-like protein (DUF3049)
	AT5G64100	AT5G64100	Peroxidase superfamily protein
	AT2G30620	AT2G30620	winged-helix DNA-binding transcription factor family protein
	AT1G63530	AT1G63530	hypothetical protein
	AT3G16240	DELTA-TIP	delta tonoplast integral protein
	AT2G48080	AT2G48080	oxidoreductase
	AT5G22920	AT5G22920	CHY-type/CTCHY-type/RING-type Zinc finger protein
	AT1G14900	HMGA	high mobility group A
	AT5G43520	AT5G43520	Cysteine/Histidine-rich C1 domain family protein
	AT1G20030	AT1G20030	Pathogenesis-related thaumatin superfamily protein
	AT4G31550	WRKY11	WRKY DNA-binding protein 11
	AT5G41080	GDPD2	PLC-like phosphodiesterases superfamily protein

	AT5G18850	AT5G18850	Low-density receptor-like protein
	AT1G78830	AT1G78830	Curculin-like (mannose-binding) lectin family protein
	AT2G38310	PYL4	PYR1-like 4
	AT3G45960	EXLA3	expansin-like A3
	AT2G35980	YLS9	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
	AT4G25810	XTR6	xyloglucan endotransglycosylase 6
	AT4G28940	AT4G28940	Phosphorylase superfamily protein
	ATCG00480	PB	ATP synthase subunit beta
	AT5G55050	AT5G55050	GDSL-like Lipase/Acylhydrolase superfamily protein
	AT5G57560	TCH4	Xyloglucan endotransglucosylase/hydrolase family protein
	AT3G63380	AT3G63380	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
	AT1G14960	AT1G14960	Polyketide cyclase/dehydrase and lipid transport superfamily protein
	AT2G32660	RLP22	receptor like protein 22
	AT4G38470	STY46	ACT-like protein tyrosine kinase family protein
	AT4G24180	TLP1	THAUMATIN-LIKE PROTEIN 1
	AT2G48130	AT2G48130	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT1G80240	DGR1	choice-of-anchor C domain protein
	AT4G17340	TIP2;2	tonoplast intrinsic protein 2;2
	AT4G01250	WRKY22	WRKY family transcription factor
	AT3G45840	AT3G45840	Cysteine/Histidine-rich C1 domain family protein
	AT3G46280	AT3G46280	kinase-like protein
	AT3G46230	HSP17.4	heat shock protein 17.4
	AT4G08950	EXO	Phosphate-responsive 1 family protein
	AT4G28040	UMAMIT33	nodulin MtN21 /EamA-like transporter family

			protein
	AT2G30930	AT2G30930	hypothetical protein
	AT1G67110	CYP735A2	cytochrome P450
	AT3G57520	SIP2	seed imbibition 2
	AT5G01050	AT5G01050	Laccase/Diphenol oxidase family protein
	AT1G08630	THA1	threonine aldolase 1
	AT5G42510	AT5G42510	Disease resistance-responsive (dirigent-like protein) family protein
	AT3G52420	OEP7	outer envelope membrane protein 7
	AT5G38710	AT5G38710	Methylenetetrahydrofolate reductase family protein
	AT3G50350	AT3G50350	membrane insertase
	AT1G62510	AT1G62510	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT4G30270	XTH24	xyloglucan endotransglucosylase/hydrolase 24
	AT1G73260	KTI1	kunitz trypsin inhibitor 1
	AT5G12340	AT5G12340	DUF4228 domain protein
	AT5G19120	AT5G19120	Eukaryotic aspartyl protease family protein
	AT1G16390	42646	organic cation/carnitine transporter 3
	AT4G38340	AT4G38340	Plant regulator RWP-RK family protein
	AT3G12500	HCHIB	basic chitinase
	AT1G62980	EXPA18	expansin A18
	AT5G18840	AT5G18840	Major facilitator superfamily protein
	AT5G47560	TDT	tonoplast dicarboxylate transporter
	AT1G06760	AT1G06760	winged-helix DNA-binding transcription factor family protein
	AT4G35770	SEN1	Rhodanese/Cell cycle control phosphatase superfamily protein
	AT5G23950	AT5G23950	Calcium-dependent lipid-binding (CaLB domain) family protein
	AT2G28860	CYP710A4	cytochrome P450

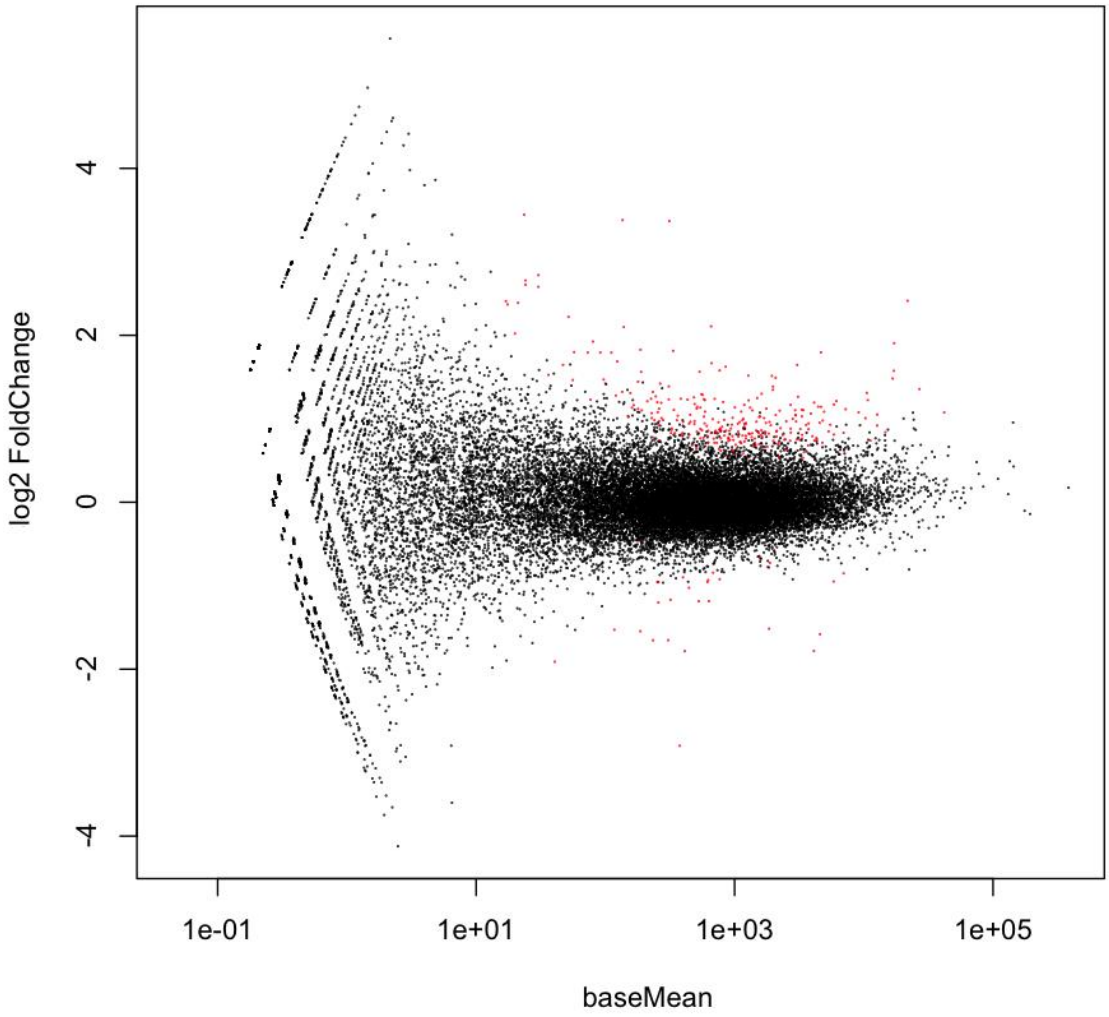
	ATMG00020	RRN26	rRNA
	AT2G44380	AT2G44380	Cysteine/Histidine-rich C1 domain family protein
	AT5G58650	PSY1	plant peptide containing sulfated tyrosine 1
	AT3G27270	AT3G27270	TRAM
	AT4G33020	ZIP9	ZIP metal ion transporter family
	AT1G53540	AT1G53540	HSP20-like chaperones superfamily protein
	AT1G17345	AT1G17345	SAUR-like auxin-responsive protein family
	ATMG00660	ORF149	hypothetical protein
	AT5G20250	DIN10	Raffinose synthase family protein
	AT5G44130	FLA13	FASCICLIN-like arabinogalactan protein 13 precursor
	ATMG00030	ORF107A	hypothetical protein
	AT4G06477	AT4G06477	transposable_element_gene
	AT3G21370	BGLU19	beta glucosidase 19
	AT5G64110	AT5G64110	Peroxidase superfamily protein
	AT4G10500	AT4G10500	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
	AT3G30775	ERD5	Methylenetetrahydrofolate reductase family protein
	AT4G26320	AGP13	arabinogalactan protein 13
	AT2G20670	AT2G20670	sugar phosphate exchanger
	ATCG00470	ATPE	ATP synthase epsilon chain
	AT5G64120	AT5G64120	Peroxidase superfamily protein
	AT5G57625	AT5G57625	CAP (Cysteine-rich secretory proteins
	AT1G35140	PHI-1	Phosphate-responsive 1 family protein
	AT4G25790	AT4G25790	CAP (Cysteine-rich secretory proteins
	AT4G33730	AT4G33730	CAP (Cysteine-rich secretory proteins
	AT5G50760	AT5G50760	SAUR-like auxin-responsive protein family
	AT2G19800	MIOX2	myo-inositol oxygenase 2

	ATCG00020	PSBA	photosystem II reaction center protein A
	AT3G21351	AT3G21351	transmembrane protein
	AT2G39510	UMAMIT14	nodulin MtN21 /EamA-like transporter family protein
	AT3G41768	AT3G41768	rRNA
	AT5G22890	AT5G22890	C2H2 and C2HC zinc fingers superfamily protein
	AT4G25470	CBF2	C-repeat/DRE binding factor 2
	AT1G75030	TLP-3	thaumatin-like protein 3
	AT1G15040	GAT1_2.1	Class I glutamine amidotransferase-like superfamily protein
	AT3G27690	LHCB2.3	photosystem II light harvesting complex protein 2.3
	AT5G54075	U3D	U3D
	AT2G21040	AT2G21040	Calcium-dependent lipid-binding (CaLB domain) family protein
	AT5G53740	AT5G53740	hypothetical protein
	AT2G01021	AT2G01021	hypothetical protein
	AT2G31141	AT2G31141	hypothetical protein
	AT1G18280	AT1G18280	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
	AT1G08115	U1-5	U1
	ATMG01380	RRN5	rRNA
	AT1G15405	NA	NA
	AT5G44120	CRA1	RmlC-like cupins superfamily protein
	AT1G48130	PER1	1-cysteine peroxiredoxin 1
	AT4G28520	CRU3	cruciferin 3



Appendix C: Diagnostic fit of variance function

The fitted variance (red line) follows the gene variance estimates closely. Due to the use of four replicates the per gene variance levels show a fairly tight spread.



Appendix D: M-A plot of fold change vs base means for each gene

Significantly differentially expressed genes (FDR < 0.01) are highlighted in red. Considerably more genes are upregulated than downregulated. Considerable noise can be seen with genes possessing a mean of less than 10 reads mapped.

Appendix E: Copyright permission to include figure previously published

SPRINGER LICENSE
TERMS AND CONDITIONS

Oct 24, 2017

This Agreement between Mr. Shrikaar Kambhampati ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4215310891991
License date	Oct 24, 2017
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Advances in Asparagine Metabolism
Licensed Content Author	Shrikaar Kambhampati, Ebenezer Ajewole, Frédéric Marsolais
Licensed Content Date	Jan 1, 2017
Type of Use	Thesis/Dissertation
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	1
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Original figure numbers	Figure 1
Title of your thesis / dissertation	Role of Asparagine as a nitrogen signal and characterization of a nitrogen responsive class I glutamine amidotransferase, GAT1_2.1 in Arabidopsis thaliana
Expected completion date	Dec 2017
Estimated size(pages)	150

Appendix F: Copyright permission to include text excerpts previously published

SPRINGER LICENSE
TERMS AND CONDITIONS

Oct 24, 2017

This Agreement between Mr. Shrikaar Kambhampati ("You") and Springer ("Springer") consists of your license details and the terms and conditions provided by Springer and Copyright Clearance Center.

License Number	4215311191316
License date	Oct 24, 2017
Licensed Content Publisher	Springer
Licensed Content Publication	Springer eBook
Licensed Content Title	Advances in Asparagine Metabolism
Licensed Content Author	Shrikaar Kambhampati, Ebenezer Ajewole, Frédéric Marsolais
Licensed Content Date	Jan 1, 2017
Type of Use	Thesis/Dissertation
Portion	Excerpts
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	Role of Asparagine as a nitrogen signal and characterization of a nitrogen responsive class I glutamine amidotransferase, GAT1_2.1 in Arabidopsis thaliana
Expected completion date	Dec 2017
Estimated size(pages)	150

Curriculum Vitae

Shrikaar Kambhampati

Education:

2014-present Ph.D. Biology (Cell and Molecular Biology),
Western University,
London, ON, Canada

2011-2013 MSc. (Environmental and Life Sciences)
Trent University,
Peterborough, ON, Canada

2007-2011 B.Sc. (Hons) Agriculture
Acharya N G Ranga Agricultural University
Bapatla, Andhra Pradesh, India

Awards and Honors:

- Biology department undergraduate teaching award for the year, 2015-16.
- Great Ideas for Teaching (GIFT) Award by the Teaching Support Center at Western University for the year, 2015-16.
- Great Ideas for Teaching (GIFT) Award runner-up, 2016-17.
- Invited speaker at the Winter Teaching Conference (2016) held at Western University where I spoke about innovative pedagogical methods.
- Guest speaker at GradCast radio (SOGS podcasts) in 2016.

Selected conference presentations:

- **Kambhampati, S**., Pandurangan, S., Renaud, J.B., Austin, R.S., Sumarah, M.W., Marsolais, F. (2017) Combined transcriptomic and metabolomic approaches provide new insights into C/N partitioning in roots of Arabidopsis thaliana. Plants from Sea to Sky (CSPB annual conference), Vancouver, BC, Canada. [Poster]
- **Kambhampati, S**., Pandurangan, S., Renaud, J.B., Austin, R.S., Sumarah, M.W., Marsolais, F. (2016) Combined transcriptomic and metabolomic approaches provide new insights into C/N partitioning in roots of Arabidopsis thaliana. Plant Biotech, 2016. Kingston, ON, Canada. [Oral]
- **Kambhampati, S**., Marsolais, F. (2015) Role of a class I glutamine amidotransferase (GAT1_2.1) in maintaining C/N balance in roots of Arabidopsis thaliana. Botany, 2015. Edmonton, AB, Canada. [Oral]
- **Kambhampati, S**., Coates, K.E., Zhang, Z., Emery, R.J.N. (2013) Role of cytokinin oxidase/dehydrogenase (CKX) in seed development and the analysis of

their gene family members in Soybean (Glycine max). Plant Biology, 2013, American Society of Plant Biologists. Providence, RI, USA. [Poster]

- **Kambhampati, S**., Coates, K.E., Zhang, Z., Emery, R.J.N. (2013) Role of cytokinin oxidase/dehydrogenase (CKX) in seed development and the analysis of their gene family members in Soybean (Glycine max). CSPB, 2013, Canadian Society of Plant Biologists. Quebec city, Quebec, Canada. [Oral]

Refereed publications:

1. **Kambhampati, S.**, Kurepin, L.V., Bruce, K.E., Cober, E.R., Morrison, M.J., Emery, R.J.N. **(2017)** Yield associated traits correlate with cytokinin profiles in developing pods and seeds of field-grown soybean cultivars. **Field Crops Research** 214: 175-184.

2. Chen, C., Li, C., Wang, Y., Renaud J.B., Tian, G., **Kambhampati, S.**, Saatian, B., Nguyen, V., Hannoufa, A., Marsolais, F., Yuan, Z., Yu, K., Austin, R.S., Kohalmi, S.E., Cui, Y. **(2017)** Cytosolic acetyl-CoA promotes histone acetylation predominantly at H3K27 in Arabidopsis. **Nature Plants** 3: 814-824

3. **Kambhampati, S.**, Ajewole, E., Marsolais, F. **(2017)** Advances in Asparagine metabolism. Progress in Botany, Springer International Publishers, Cham, Switzerland. Volume: 79. (DOI: https://doi.org/10.1007/124_2017_13)

In Preparation:

4. **Kambhampati, S.**, Renaud, J.B., Sumarah, M., Marsolais, F. Role of a class I glutamine amidotransferase in C/N partitioning in roots of *Arabidopsis thaliana***.**

5. **Kambhampati, S.**, Pandurangan, S., Renaud, J.B., Austin, R.S., Sumarah, M., Marsolais, F. Transcriptome and metabolomic data from Asparagine treated *Arabidopsis* roots provide new insights into nitrogen metabolism.