Functional genomics through metabolite profiling and gene expression analysis in *Arabidopsis thaliana*

Diego Fernando Cortes Bermudez

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Genetics, Bioinformatics and Computational Biology

Vladimir Shulaev, Chair
Allan W. Dickerman
Reinhard C. Laubenbacher
Pedro J. P. Mendes
Brenda S. J. Winkel

July 25th, 2008
Blacksburg, Virginia

Keywords: Functional Genomics, SABATH Methyltransferases, Plant Signaling Molecules, Metabolomics, Transcriptomics

Abstract

Functional genomics through metabolite profiling and gene expression analysis in *Arabidopsis thaliana*

Diego Fernando Cortes Bermudez

Vladimir Shulaev, Chair

In the post-genomic era, one of the most important goals for the community of plant biologists is to take full advantage of the knowledge generated by the Arabidopsis thaliana genome project, and to employ state-of-the-art functional genomics techniques to assign function to each gene. This will be achieved through a complete understanding of what all cellular components do, and how they interact with one another to produce a phenotype.

Among the proteins encoded by the Arabidopsis genome are 24 related carboxyl methyltransferases that belong to the SABATH family. Several of the SABATH methyltransferases convert plant hormones, like jasmonic acid, indole-3-acetic acid, salicylic acid, gibberellins, and other plant constituents into methyl esters, thereby regulating the biological activity of these molecules and, consequently, myriad important physiological processes.

Our research aims to decipher the function of proteins belonging to the SABATH family by applying a combination of genomics tools, including genome-wide expression analysis and gas-chromatography coupled with mass spectrometry-based metabolite profiling. Our results, combined with available biochemical information, provide a better understanding of the physiological role of SABATH methyltransferases, further insights into secondary plant metabolism and deeper knowledge of the consequences of modulating the expression of SABATH methyltransferases, both at the genome-wide expression and metabolite levels.

To my mother and father,
to my brother

Acknowledgments

The writing of a dissertation can be a lonely and isolating experience, yet it is obviously not possible without the personal and practical support of numerous people. Faculty, friends, and family members have helped me to complete this dissertation. I would like to express my gratitude to these individuals for their support and assistance.

The faculty of this program have provided me with a tremendous graduate education: they have taught me how to think about transdisciplinary problems; they have provided me with scientific opportunities and economic support; and they have shown me how to approach my work as a scientist. Several individuals deserve special mention for their contributions to this dissertation.

Vladimir Shulaev has been a strong and supportive advisor to me throughout my graduate school career, but he has always given me great freedom to pursue independent work. In reviewing my writings, he offers in-depth comments, but always respects my voice. Throughout my doctoral work he continually stimulated my analytical thinking and encouraged me to develop independent research skills.

I am also very grateful for having an exceptional doctoral committee and wish to thank Allan Dickerman, Reinhard Laubenbacher, Pedro Mendes and Brenda Winkel for their continual support and encouragement. With critiques and simple words they taught me how to think about scientific problems.

A special word of gratitude to everyone that, either in Vladimir's group or in the Virginia Bioinformatics Institute, at large, crossed the path of my life throughout the years. This opportunity to interact with people from different cultural backgrounds, and diverse areas of expertise, has made my graduate experience more flourishing.

A special thanks to Leepika Tuli, with whom I shared an office, the experience of graduate education, and many great conversations. I am deeply indebted to Joel Shuman, who was very generous with his time and knowledge. Joel always took time from his busy schedule to assist me on each step to master mass spectrometry.

I have been especially fortunate to know Mihaela Babiceanu. She has been a true friend. During this years of graduate studies we shared several courses and research projects together, her company made the academics more fun and enjoyable. No one could ask for a better friend in life than she has been to me. I also praise her daughter Laura, for inspiring and amazing me every day. With the unbiased judgment that only the innocence of childhood provide, Laura questions my answers, brings joy and make my days brigher. I want to credit my best friend and fellow countryman, Juan Jairo Ruiz, for his company, endless encouragement and support all this time. All together Mihaela, Laura, and Juan have been my family in this foreign lands.

My enormous debt of gratitude can hardly be repaid to Ruth Irwin, who offered me hospitality, welcoming me so warmly to Blacksburg, and for giving me the respite that chamber music offers.

I want to thank the good friends that I have made in Virginia Tech for their support throughout my time in graduate school. More importantly, they have given me an arena where I can let off steam.

I most want to thank my family for their love, sacrifice, and kind indulgence. This work is also yours.

I also gratefully acknowledge the institutional support that I have received while working on this research. In particular, I thank the National Science Foundation and the Virginia Bioinformatics Institute for supporting me with generous fellowships.

Finally, despite all the assistance provided by my doctoral committee and others, I alone remain responsible for the content of the following, including any errors or omissions which may unwittingly remain.

Contents

List of Figures

List of Tables

List of Abbreviations

BSMT	Benzoic and salicylic acid methyltransferase
DFA	Discriminant function analysis
FA	Farnesoic acid
FAMT	Farnesoic acid methyltransferase
GA-DFA	Genetic algorithms-discriminant function analysis
GAMT1	Gibberellic acid methyltransferase 1
GAMT2	Gibberellic acid methyltransferase 2
GAs	Gibberellins
GC/MS	Gas chromatography coupled to mass spectrometry
GCRMA	GeneChip robust multiarray analysis
GO	Gene ontology
GUS	β-glucuronidase
IAA	Indole-3-acetic acid
IAMT	Indole-3-acetic acid methyltransferase
JA	Jasmonic acid
JMT	Jasmonic acid methyltransferase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC/MS	Liquid chromatography coupled to mass spectrometry
MeBA	Methyl benzoate
MeFA	Methyl farnesoate
MeGA	Methyl gibberellins
MeIAA	Methyl indole-3-acetic acid
MeJA	Methyl jasmonate
MeSA	Methyl salicylate
MT	Methyltransferase
NA	Nicotinic acid
NAMT	Nicotinic acid methyltransferase
RMA	Robust multiarray analysis
RT-PCR	Reverse transcriptase - polymerase chain reaction
SA	Salicylic acid
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine

Declaration

The work in this thesis is based on research carried out at the Biochemical Profiling Group, under the supervision of Vladimir Shulaev at the Virginia Bioinformatics Institute at Virginia Tech, Blacksburg VA. This work is part of a collaborative project with Eran Pichersky at the Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan and Joseph Noel at The Salk Institute for Biological Studies, La Jolla, California.

The Virginia Bioinformatics Institute (VBI) at Virginia Tech has made available its facilities to carry out the proposed research project. The microarray experiments were carried out at the Core Laboratory Facility (CLF), while the metabolite proling was done in Dr. Vladimir Shulaev's laboratory

This project is funded through research grants, MCB0312857 to Vladimir Shulaev, MCB0312466 to Eran Pichersky, and MCB0312449 to Joseph Noel from the National Science Foundation.

No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced otherwise in the text.

Chapter 1

Introduction

"This has been in the works for a long time.
This was something that was always obvious
to do as a next step"

Craig Venter

1.1 Motivation

Until recently, the analysis of biological systems to understand the orchestrated inner mechanisms that produce the final phenotype, depended heavily on monitoring macroscopic physiological characters and few entities at the molecular level, due to the limited analytical techniques available. Under these conditions, understanding gene function was limited to the evaluation of few macroscopic markers and molecular entities selected under a hypothesis-driven fashion. In consequence, any conclusion, knowledge and models derived from this analytical approach relied significantly on the validity of the initial hypothesis. Furthermore, limited to a few and specific number of molecular or macroscopic measurements, the researcher was usually unable to observe and correlate simultaneously occurring phenomena that were not consider under the initial hypothesis.

Advances in understanding biological processes at the molecular level, together with innovations in computer science and robotics, permitted the development of high-throughput technologies for the analysis of the biological systems in a not hypothesis-driven manner. Furthermore, with the comprehensive analysis of the different type of cellular components, there are more opportunities to observe and correlate simultaneously occurring phenomena, generating a more detailed and accurate understanding of biological systems.

Given the components and evolutionary complexity of biological systems, it is not possible to understand them comprehensively on the basis of hypothesis-driven research alone. Likewise, it is not possible to do so solely through "*omic*" studies of genes, proteins, metabolites and their macromolecular complexes. These two research aproaches are complementary and synergistic [1].

Genome-wide transcription profiling with cDNA microarrays or Affymetrix GeneChip has been the most widely used high-throughput analysis platform in the post-genomic era. However, there is a clear understanding that complete knowledge on how biological systems function, can only be achieved with the comprehensive integration of quantitative and qualitative data for all the cellular components. Such integration has the potential to provide insight about the function of unknown genes, the metabolic processes where they participate, as well as the networks among different cellular components. With this motivation, this dissertation address the major challenges of gene function elucidation, based on quantitative high-throughput analysis of genome-wide transcription and metabolic state of a biological system.

1.2 Functional Genomics: From Sequence Information to Gene Function

Large-scale genome sequencing projects have greatly changed the way scientist approach biology. With frequency Genomics is often referred to as a new field that has led to a shift in the way science is performed. Taking advantage of the vast amounts of genome sequence data, the post-genomic era has emerged providing scientist with new tools and approaches to answer biological questions. Nowadays, researchers do not need to approach biological questions exclusively in a hypothesis-driven way, instead they can collect and analyze data in a non-biased and broader fashion [2]. Recent technological advances and the rapid development of novel tools now permit the interrogation of a complete genome all at once and in a single experiment. Currently, significant part of biological research is devoted to add gene function information to genome sequence data, adding value to the nucleotide sequence collections. Knowing the exact sequence and location of all the genes of a given organism is only the first step towards understanding how all the parts of a biological system work together. Functional genomics is a general approach toward understanding how the genes of an organism work together by assigning new functions to unknown genes. Information about the hypothesized function of an unknown gene may be deduced from its sequence, using already known functions of similar genes as the basis for comparison. In addition, the location of given gene in the context of the chromosome may allow speculation with respect to gene function, providing that the function and chromosomal location of genes with similar sequence is known [3].

It is often said that genomics has "revolutionized" plant biology [4]. The availability of genomic techniques has been significantly enhanced in the past years, in order to characterize the structure of genomes and to better understand the relationship between genotype and phenotype. The completed genome sequence of the model plant *Arabidopsis thaliana* has provided plant biologist with a "blueprint" to better understand the inner mechanisms of plant cells and the plant as a whole, expediting the application of novel technologies to functional analysis of plant genetic information. As a long term goal of the plant biology community, the elucidation of the function of all genes in Arabidopsis represent a big step forward to understand how a plant works as a system.

Nowadays, several rapid and multiparallel applications, generally known as *functional genomics,* have become routine approaches for assigning function to new genes. Such methods enable innovative and thorough analysis of the cellular components that are related to gene function. Genomic analysis at three different levels can be combined to provide an understanding of the whole organism: gene transcripts (mRNA), proteins, and low molecular weight intermediates (metabolites) [5]. Consequently, three different approaches, named *transcriptomics, proteomics* and *metabolomics*, have evolved into different strategies, within the functional genomics platform, and are fully integrated into plant biology [6, 7]. However, the role of genes with unknown function cannot be fully determined and comprehended by

use of a single genomics technique. Only integration of information obtained using different genomics tools will facilitate the assignment of function to the large number of plant genes that continue to be undetermined. The integration of the transcriptome and metabolome or detailed targeted metabolite analysis is helpful in identifying the function of unknown genes and determining all gene-to-metabolite correlations in cells. Fortunately, the number of successful studies in identifying novel gene functions by this approach is growing [6, 8–11].

The ultimate goal of large-scale sequencing projects is to assign function to all the genes identified in the genome. In all higher organisms, the majority of genes that have been identified and sequenced have not been functionally characterized [12]. Thereby, the assignment of function has become one of the major goals of functional genomics. On the publication of the Arabidopsis genome sequence, less than 10% of the potential genes had been functionally characterized [13]. Most genes were identified based on the similarity of their nucleotide or amino acid sequence to others available in public databases. Moreover, the function of approximately 30% of the putative Arabidopsis genes could not even be speculated because their sequences were distinct from previously characterized genes [13]. Wheter the cellular components analyzed by any of the mentioned "omics" applications is mRNA, proteins, metabolies, or a combination of them, the genomic approach designed to determined the function of an unknown gene falls into one of two strategies: *forward genetics* and *reverse genetics.* Forward genetics refers to the use of natural or artificially generated mutants with characteristics contrasting to the wild-type plant to clone the gene responsible for the mutant phenotype. In contrast, reverse genetics starts with a specific sequenced gene that is "knocked-out" or "knocked-down" using direct mutagenesis to evaluate the resulting change in phenotype.

With the complete genome sequence of *A. thaliana* available, it has been realized how little is known about the many gene functions in this organism. Even with the relatively large number of known metabolites, the task of determining the genes and enzymes responsible for the formation of these metabolites still remains a difficult task.

In the post-genomic era, one of the most important milestones for the community of plant biologists is to take full advantage of the knowledge generated by the *A. thaliana* genome project and to employ state-of-the-art functional genomics techniques to assign function to each gene. This will be achieved through a complete understanding of what proteins do, and how do proteins interact with other molecules in the cell to produce the final phenotype.

In the 1980's, plant scientists adopted the widespread use of an easily manipulated model called Arabidopsis, and established an international research effort called The Multinational Coordinated Arabidopsis Genome Research Project. As an outcome of this effort, the entire genome sequence of this plant was completed in Decamber 2000. For the first time, the sequence of the estimated 25,000 genes that are necessary for a plant to function as a flowering plant, was known. As a continuation to the Arabidopsis genome sequencing efforts, an important and revolutionary initiative—to exploit the revolution in plant genomics by understanding the function of all genes of a reference species within their cellular, organis-

mal, and evolutionary context by the year 2010 (The 2010 Project)—was proposed by the community of plant biologists [4].

Among the genes yet to be completely understood with regard to their functional role, the *A. thaliana* genome contains 24 related genes encoding enzymes that belong to the SABATH family of methyltransferases (MTs). Experiments reported in the literature suggest that different SABATH methyltransferases convert several important hormones and other plant constituents into their methyl esters, thereby exerting important effects on the biological activity of these molecules and, consequently, on myriad important physiological processes.

Different "*omics*" approaches evolved within the functional genomics platform, are fully integrated into plant biology to help assigning function to unknown genes. However, applying a single genomic approach cannot decipher the function of an unknown gene. A combination of genomic approaches together with biochemical analysis has proven to be a powerful strategy for assigning function to candidate genes and enzymes, particularly those in specialized plant metabolism.

A research group, led by Eran Pichersky at the University of Michigan, has isolated and cloned all the 24 gene members of the SABATH family in Arabidopsis. In addition, they have obtained T-DNA insertion lines, developed homozygous knockouts, and also overexpressing transgenic plants for most of the 24 genes. These plant materials were made available for further characterization of their phenotypes, and constitute a great resource to undertake a functional genomics enterprise elucidating the function of the AtSABATH genes. Current knowledge on AtSABATH methyltransferases indicates that this family of proteins are involved in important processes in the plant secondary (specialized) metabolism, through methylation of important signaling molecules. I strongly believe that the correlative analysis of extensive metabolic profiling and gene expression profiling data, together with the already available and ongoing biochemical analysis, might bring a complete understanding to assign function to the SABATH genes.

It has been shown that SABATH methyltransferases convert several important hormones and other plant metabolites into their methyl esters, thus these MTs play an important role on the biological activity of these molecules and consequently on many important physiological processes. Members of the SABATH family in Arabidopsis exert important roles in different metabolic processes mediated by jasmonic acid (JA), salicylic acid (SA), indole-3-acetic acid (IAA), gibberellins (GAs), and possibly other signaling molecules.

It is generally known among plant biologists and physiologists the importance of this specialized group of metabolites and their crucial role in processes like defense response mechanisms, systemic acquired resistance and other stress related responses, as well as development and growth. Despite the important role of this molecules in different processes in the plant, the molecular basis and mechanisms of action have not been completely established. It is intriguing to study the consequences of the particular actions of these enzymes, as well as to establish the role of the members of this family of proteins in Arabidopsis metabolism.

The current knowledge about the biosynthesis of plant hormones and other metabolites has been acquired mostly by individual researchers concentrating on one or a few reactions at the time. The advent of large scale automated sequencing, methods for analysis of gene expression profile to interrogate gene function, bioinformatics, improvements in chemical analysis instrumentation and the ability to purify proteins and analyze their function faster than before, now allow for a functional genomics approach to this problem. Not only will it soon be possible to identify the function of all enzymes in biochemical pathways in selected plant species, but also we will be able to find important regulatory elements, such as transcription factors and signaling pathway members. This will enable us to better understand, and be able to control, the regulation and flux in specific pathways for the production of target compounds which are found to be important for human health and/or for industrial application [14–17].

1.3 Research Objectives

The main objective of this research was to apply functional genomics to decipher the roles of the methyltransferases belonging to the SABATH family by a combination of methods that involve gene expression profiling and metabolite profiling. We aimed to gain a better understanding of specialized metabolites in Arabidopsis, further insights into specialized metabolism and deeper knowledge of the consequences of modulating the expression of SABATH methyltransferases. In order to achieve the main objective, the following specific aims were pursued:

Specific Aim 1. To determine the effect of normal and aberrant (knockout and overexpressing) expression levels of SABATH genes on the genome-wide transcription profile of Arabidopsis. The following specific issues were addressed:

1. Identification of gene expression profiles associated with single perturbations of specific SABATH genes.
2. Identification of genes and pathways differentially expressed as a result of perturbation in the expression levels of AtSABATH genes.

Specific Aim 2. To establish the effect of normal and aberrant (knockout and overexpressing) expression levels of Arabidopsis SABATH genes on metabolic profiles. The following specific issues were addressed:

1. Understanding the metabolic consequences of modulating the expression of AtSABATH genes.

2. Characterization of metabolic fingerprints specific to different AtSABATH gene perturbations.

3. Identification of metabolic networks affected by specific AtSABATH gene perturbations.

Specific Aim 3. Parallel and complementary analysis of the transcriptome and metabolome of wild-type and mutants for gibberellic acid methyltransferase 1 (GAMT1) and gibberellic acid methyltransferase 1 (GAMT2). The following specific issues were addressed:

1. Identification of genes and pathways differentially expressed as a result of perturbation in the expression levels of GAMT1 and GAMT2.

2. Identification of metabolic networks and target metabolites affected by perturbation in the expression levels of GAMT1 and GAMT2.

3. Explore the physiological roles of GA methylation in Arabidopsis.

1.4 Description of the Thesis

This dissertation focuses on functional genomics through a combination of metabolite profiling, and gene expression analysis in *A. thaliana.* This dissertation is organized in 6 chapters.

Chapter 1: Describes the motivation for the thesis and the importance of functional genomics based on a combination of high-throughput analyses to identify novel gene functions. The specific aims pursued in this research are also stated.

Chapter 2: Provides an overview of the current knowledge on the biology of different proteins members of the SABATH family of methyltransferases in Arabidopsis and across different species. The possible implications of methylation and demethylation of signaling molecules is also discussed.

Chapter 3: Presents results obtained from analysing the effects that modulating the expression of different AtSABATH genes have on the genome-wide expression profile of Arabidopsis. The information on differentially expressed genes was systematically analyzed to elucidate the biological consequences of changes in the gene expression profile.

Chapter 4: Results obtained from metabolite profiling and metabolite fingerprinting analysis of different Arabidopsis plant tissues from wild type, and mutants with perturbations in the gene expression for several AtSABATH are presented. The results obtained from multivariate statistical analysis are presented and discussed in a biological context.

Chapter 5: This chapter describes the results of a parallel and complementary analysis of genome-wide expression and metabolite profiling data to understand the gene function of gibberellic acid methyltransferases GAMT1 (At4g26420) and GAMT2 (At5g56300). The results indicate that GA methylation is a mechanism to regulate the availability of active gibberelin in seeds and the downstream events that are triggered by active gibberellins in seeds. Physiological interpretation and hypotheses about the role of GA methylation in Arabidopsis seed development and germination are discussed.

Chapter 2

Regulating the Biological Activity of Signaling Molecules in Plants with Carboxyl Methyltransferases[1]

[1]This chapter is based on published material, used with permission of Elseiver:
Y. Yang, M. Varbanova, J. R. Ross, G. Wang, D. Cortes, E. Fridman, V. Shulaev, J. P. Noel, E. Pichersky. Methylation and demethylation of plant signaling molecules. J. T. Romeo, R. K. Ibrahim, L. Varin, V. De Luca, redaktorzy, *Recent Advances in Phytochemistry*, Volumen 40. Elsevier Science Ltd, Oxford, 2006.

"Aromatic plants bestow no spicy fragrance while they grow; but crush'd or trodden to the ground, diffuse their balmy sweets around."

Oliver Goldsmith

2.1 Introduction

The transmission of chemical signals from one organ to another, or between tissues or neighboring cells plays an important role in many developmental and physiological processes in plants. Some of the molecules carrying out this role in plants include ethylene, auxins, gibberellins, jasmonates, salicylic acid, cytokinins and brassinosteroids. These chemical signals are designated as hormones or signaling molecules.

In response to a stimulus, signaling molecules are synthesized at the source, transmitted, and finally perceived at the target site in order to trigger a response. After eliciting a response, the signal molecule must be switched off, or the physiological process and response will continue to be perceived and propagated after the need to respond to the initial stimulus expires. This can be achieved by either negative feed-back, to reduce *de novo* synthesis of the hormone, inactivation through conjugation, or degradation to different moieties. Activation of the hormone pathway often down-regulates hormone metabolism, which leads to the inactivation of the hormone signals.

Historically, the evidence for the presence of signalling compounds has typically come from genetic experiments, through the analysis of mutans that were defective in the initiation of an specific process. In other cases, the exogenous application of an array of compounds has revealed effects on physiological or developmental processes.

Despite many exciting discoveries on the mechanisms involved in plant signalling, for many of the compounds we do not yet have a complete understanding of the details of the pathways involved in their synthesis and catabolism. In some cases, the exact chemical structure of the actual mobile compound is not clear.

In this chapter we review the current knowledge on the SABATH family of methyltransferases and the possible implications for the methylation of signalling molecules.

2.2 Methylation in Plant Specialized Metabolism

Among the large number of estimated metabolites in plants, a wide variety of volatile ester compounds have been found to constitute components of complex cocktails emitted to serve in both reproductive and vegetative processes [18,19]. More than 1.000 low-molecular-weight organic compounds have been identified in plants to date [19], most of them components of floral scents. The flower volatiles function as pollinator attractants but may also be involved in defense mechanism, such as airborne signals to attract natural predators of herbivores, to directly repel herbivores or to activate defense-related genes in neighboring plants or healthy tissue in infected plants [20, 21]. Methylbenzoate, methylsalicylate, benzylbenzoate and benzylacetate are found in floral scents of *Clarkia breweri* and *Arabidopsis thaliana*. Together with cinnamic esters, terpenoids, fatty acid derivatives, C5-branched compounds, various nitrogen compounds and several sulfur containing compounds, benzenoid esters form the core physiologically active volatiles in plants [19]. Certain volatiles such as methylsalicylate and methyl jasmonate are also thought to be involved in cell-to-cell signaling [21].

Methylation is a common enzymatic modification in plant secondary metabolism. Almost all classes of plant metabolites are known to be methylated, most commonly on carbon, nitrogen, sulphur and oxygen atoms. The most widely used methyl donor for enzymatic methyl transfer is the co-factor S-adenosyl-L-methionine (SAM) [22]. A large number of enzymes that perform SAM-dependent methylation reactions have been reported in plants. Depending on the target group of compounds they methylate and based on their sequence similarity, these plant small molecule methyltransferases are classified in different groups. Type 1 MTs, with specific activity for oxygen atoms, *O*-methyltransferases (OMTs) methylate hydroxyl moieties of phenylpropanoid-based compounds. Type 2 MTs, currently understood to be specific for coenzyme-A derivatized phenylpropanoid compounds, are found in all lignin-producing plants. The third family of plant MTs consists mostly of enzymes that convert carboxylic acid to methyl ester derivatives, but it also include some alkaloid N-methyltransferases (NMT) involved in caffeine biosynthesis [22, 23].

In studies of the biosynthesis of floral volatiles in Eran Pichersky's laboratory, at the University of Michigan, a new group of methyltransferases with SAM-dependent activity was discovered. These proteins are capable of methylate various compounds at the carboxyl moieties to produce volatile methyl esters that serve to attract pollinators [18, 23, 24]. Methylsalicylate (MeSA) and methyl jasmonate (MeJA), two such esters, are the methylation products of the plant hormones SA and JA, respectively [23–25]. To distinguish this new group type type 3 MTs from those previously reported, this new class was designated the SABATH family, based on three of the first five (**SA**MT Salicylic Acid MT from *Clarkia breweri*; **BA**MT Benzoic Acid MT from snapdragon flowers; **Th**eobromine synthase from coffee [24, 26, 27]) genes isolated and characterized. In Arabidopsis, a group of 24 methyltransferases enzymes that share high sequence similarity to each other, but not significant similarity to the methyltransferases previously characterized in plants, represents the SABATH family (Figure 2.2) [14,23]. The first enzyme isolated and characterized from this family was a salicylic acid carboxyl

methyltransferase (SAMT) from flowers of *Clarkia brewi*, the enzyme responsible for the synthesis of methyl salicylate (MeSA) from salicylic acid (SA) [24]. The three-dimensional crystal structure of the *Clarkia brewi* SAMT in complex with the substrate SA and the demethylated product S-adenosyl-L-homocysteine was determined by Joseph Noel's group (Figure 2.4) [28]. The protein structural information, together with sequence analysis of databases and repositories of gene information revealed that *Clarkia brewi* SAMT defines a new family of S-adenosyl-L-methionine (SAM)-dependent methyltransferases [14, 23].

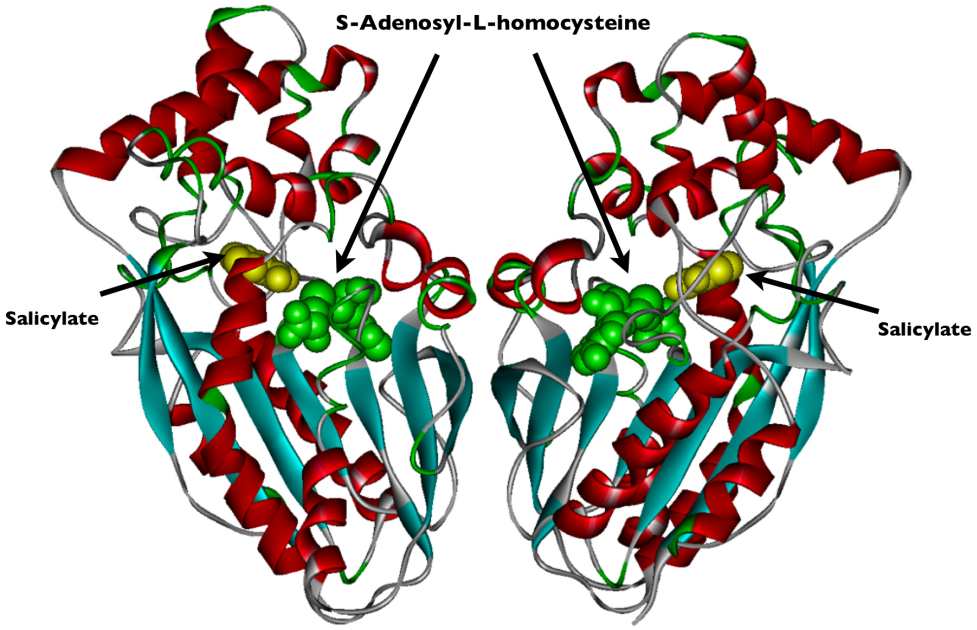


Figure 2.1: Ribbon diagram of the dimeric arrangement of the *Clarkia brewi* SAMT polypeptide backbone.
Highlighted is the binding pocked of SAMT are the methylation products of methyl salicylate and S-adenosyl-L-homocysteine.

Additionally, other methyltransferases have been identified, including a benzoic acid carboxyl methyltransferase (BAMT) from snapdragon flowers, known to catalyze the formation of methyl benzoate (MeBA). Moreover the N-methyltransferases, caffeine synthase (TCS1) from tea and theobromine synthase (CaMXMT) from coffee were characterized and shown to be related to SAMT and BAMT [29, 30].

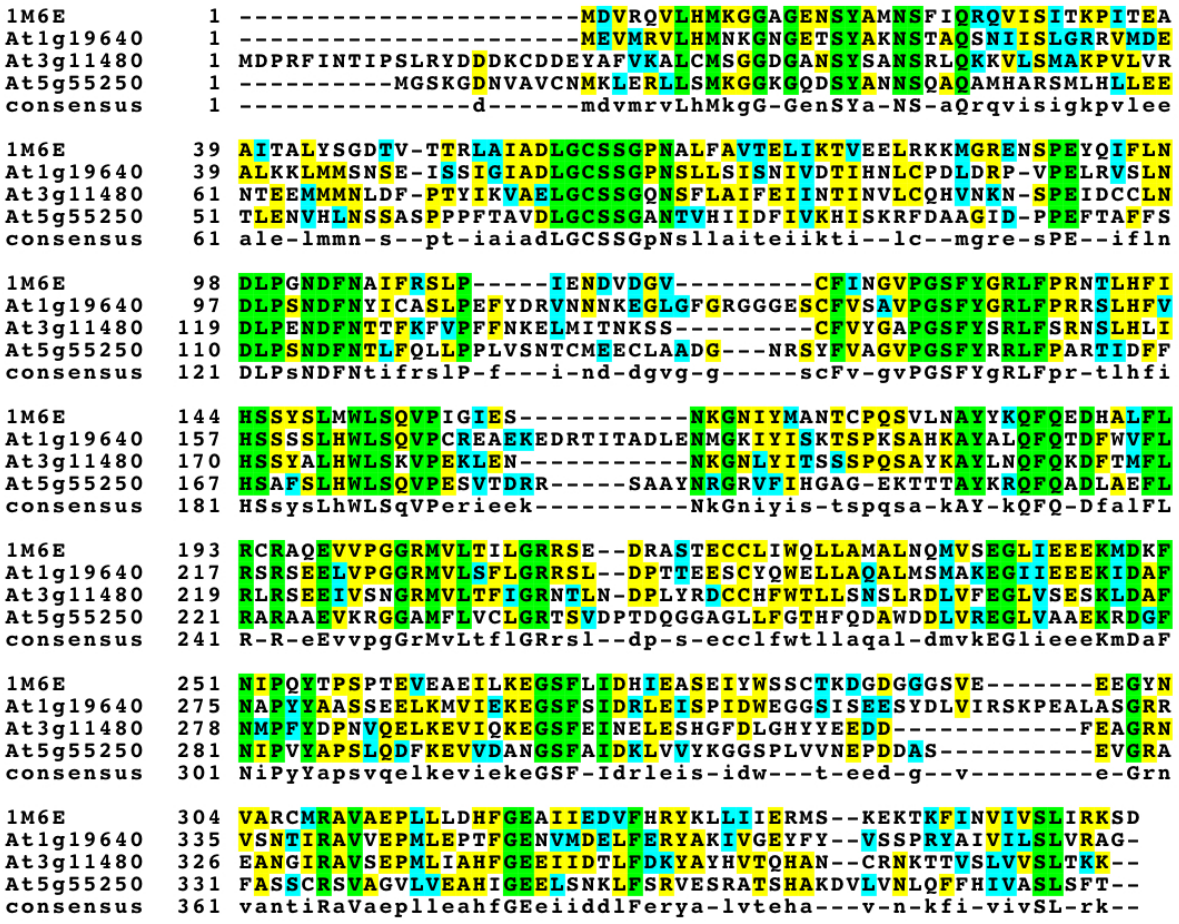


Figure 2.2: Partial sequence alignment of *Clarkia brewi* SAMT (1M6E) and three representative members of the AtSABATH methyltransferases.
At1g19640 (Jasmonic acid MT), At3g11480 (Benzoic acid and salicylic acid MT), and At5g55250 (Indole-3-acetic acid MT). The multiple sequence alignment was obtained with Clustal W and highlighed with Boxshade. Completely conserved residues are shaded in green, identical residues in yellow, and residues with similar chemical properties in cyan.

S-adenosyl-L-methionine-dependent methyltransferases [E.C. 2.1.1-] catalyze the transfer of a methyl group of SAM to an acceptor molecule, with the simultaneous formation of the corresponding methylated derivative and S-adenosyl-L-homocysteine as products (Figure 2.3). Methyltransferases are essential enzymes that play a variety of roles in maintenance of biological activities in all organisms. In general, the enzymatic methylation of natural plant products inactivates the reactivity of their hydroxyl and/or carboxyl group and, thus, alters their solubility and intracellular compartmentalization. In plants, MTs are involved in the biosynthesis of a wide array of natural products related to the plant's survival, including phenylpropanoids, lignins, flavonoids, alkaloids, and phytoalexins [31].

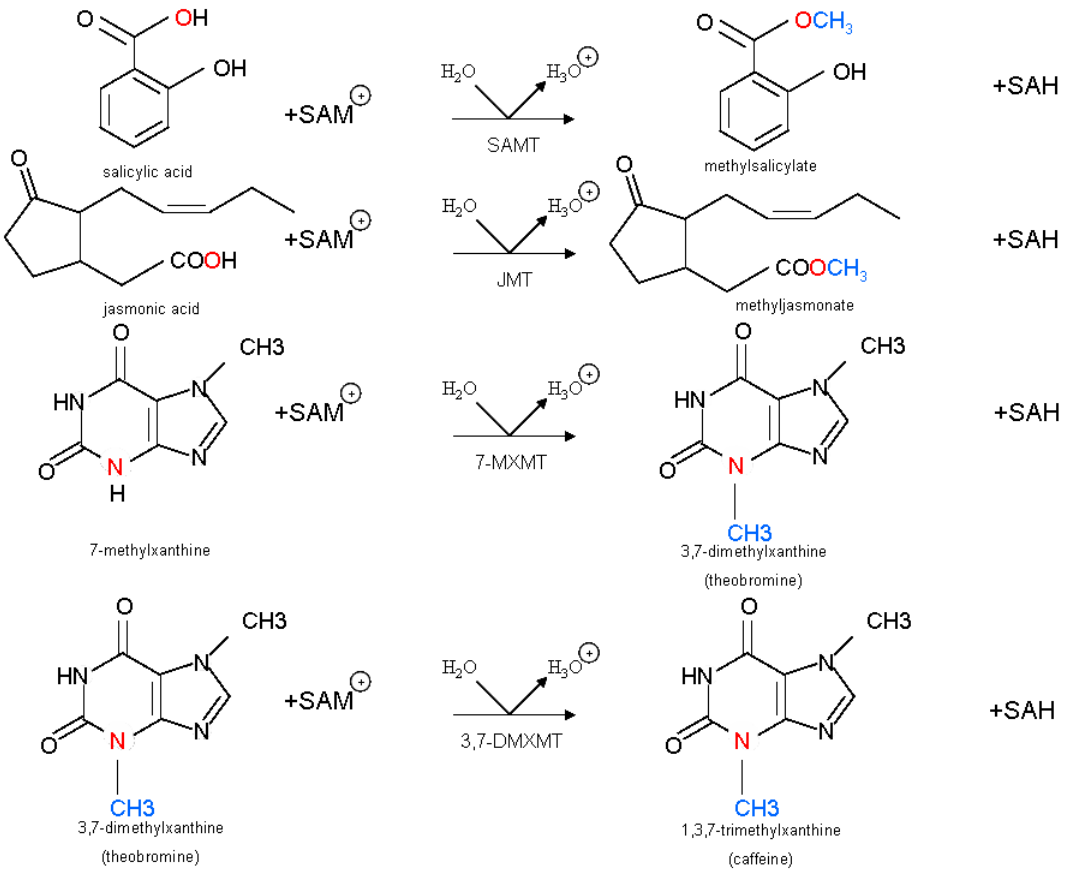


Figure 2.3: Reactions catalyzed by representative members of the SABATH family of plant methyltransferases. Adapted from Zubieta *et al.* (2003)

It has been reported that SABATH MTs convert several important hormones and other plant constituents into their methyl esters, thus these MTs play an important role on the biological activity of these molecules and consequently on many important physiological processes. Members of the SABATH family in Arabidopsis play important roles in different metabolic processes mediated by jasmonic acid (JA), salicylic acid (SA), indole-3-acetic acid (IAA), gibberellins (GAs), and possibly other signaling and hormones molecules, including abscisic acid (ABA) and cytokinins.

The importance of this specialized group of metabolites and their crucial role in defense response mechanisms, systemic acquired resistance and other stress related responses, as well as their role in development and growth, are well documented by plant biologists and physiologists. Despite the important role of this molecules in different processes in the plant, the molecular basis and mechanisms of action have not been completely established. It is interesting to study the consequences of the particular actions of these enzymes, as well as to establish the roles of the members of this family of proteins in Arabidopsis metabolism.

The recently identified family of SABATH methyltransferases and the available evidence showing that some of these methyltransferases are responsible for the synthesis of known plant metabolites with a significant importance in physiological processes, hava open the door to unravel new biochemical pathways or complete the knowledge available for others.

2.3 *Arabidopsis thaliana* SABATH Methyltransferases and their Putative Substrates

2.3.1 *At1g19640* Encodes an Enzyme that Methylates Jasmonic Acid

An enzyme catalyzing the methylation of jasmonic acid (JA) to form methyl jasmonate (MeJA), was identified by Seo *et al.* [25], the first AtSABATH methyltransferase to be characterized and designated as jasmonic acid methyltransferase (JMT). JMT was isolated as orthologous to *NTR1,* a nectary-specific gene isolated from *Brassica campestris.* However, both NTR1 and JMT exhibit protein structural features similar to those of SAMT a S-adenosyl-L-methionine-dependent salicylic acid carboxyl methyltransferase previously characterized from the flowers of *Clarkia breweri* [24]. Based on the homology to SAMT, Seo *et al.* assayed JMT for activity with several common plant acids in the presence of S-adenosyl-methionine, and showed that JA reacted with JMT to produce its methyl ester, MeJA. Expression of the gene encoding JMT is induced in different organs, at different developmental stages, and in response to various stimuli. JMT expression has been detected in most parts of mature plants, particularly in rosettes, cauline leaves and flowers. JMT expression is also induced by wounding, alamethicin, and hervibory [23, 25]. Treating Arabidopsis plants with MeJA also induce JMT expression locally and systemically, evidencing the role of MeJA as an intracellular signal transducer or airborne signal mediating intra and inter-plant communications. Transgenic plants that constitutively express the JMT gene show elevated levels of MeJA without altering JA content [25]. Moreover, according to the microarray data repository for Arabidopsis, Genevestigator (`www.genevestigator.com`) [32], JMT is significantly up-regulated in response to several stimuli including treatment with MeJA or SA, infection with *P. syringae,* and more interestingly in response to treatment with chemical substances that act as inhibitors of other hormones, such as auxin inhibitors, gibberellin biosynthesis inhibitors, brassinosteroids and ethylene inhibitors. This observation may indicate that crosstalk among the biosynthetic pathways of different signaling molecules may be orchestrated by JA/MeJA in ways were not anticipated.

However, analysis of a null allele obtained through T-DNA insertional mutagenesis, has not uncovered any obvious physiological phenotype of this mutation *in planta* [33]. Similarly to ethylene and active gibberellins, which control the expression of genes in their biosynthesis pathways, MeJA positively induces the expression of genes for JA biosynthesis. In turn,

over-expression of JMT caused the constitutive expression of jasmonate-responsive genes including lipoxygenase 2 (LOX2) and allene oxide synthase (AOS), the first two enzymes in the jasmonate biosynthesis pathway. Furthermore, genes coding defensins, resistant-related proteins and oxidative stress-related genes are also expressed, and plants are more resistant than wild-type to attack by the fungus *Botrytis cinerea* [34] and the bacterium *Pseudomonas syringae* pv. *tomato* [35, 36].

Because of the role of MeJA in defense response mechanisms, plants have been engineered to accumulate high levels of MeJA from endogenous JA, which increase their capacity to response to biotic and abiotic stress, circumventing problems associated with uptake of exogenous sprays, and ensuring the presence of appropriate isomers of MeJA. JMT overexpressing plants are more resistant against herbivores, pathogens and competitors, compared to the wild-type. However, the overproduction of MeJA by constitutive expression of JMT negatively impact total seed production and seed germination rates in Arabidopsis *[35].*

To the best of our knowledge no MeJA esterases have been isolated in Arabidopsis. However, MeJA esterases have been isolated and characterized in tomato and tobacco, which hydrolyzes MeJA to JA [37], a further indication that plants have specific enzymes to hydrolyze methylated hormones. Recently it has been shown that the tobacco protein SABP2 (Salicylic acid binding protein-2) hydrolyzes methyl salicylate to salicylic acid [38, 39]. Furthermore, there is a group 20 homologs of SABP2 in the genome of *Arabidopsis thaliana*, named AtMES (for methyl esterases), and among them several candidate methyl indole-3-acetic acid (MeIAA) esterases that could hydrolyze MeIAA. [40]. If MeJA esterases are present in Arabidopsis, they will soon be discovered and characterized.

2.3.2 *At3g11480* Encodes an Enzyme that Methylates Salicylic and Benzoic Acids

Methylsalicylate (MeSA) and methylbenzoate (MeBA) are constituents of scents emitted from flowers to attract pollinators and also from leaves to function as animals and insect toxins [14]. MeSA and MeBA are synthesized *in planta* through a reaction catalyzed by methyltransferases, which transfer a methyl group from S-adenosyl-L-methionine to the carboxyl group of salicylic acid (SA) or benzoic acid (BA), respectively. SA and BA are very similar compounds, differing only in the presence of an ortho hydroxyl on the benzyl ring of SA. In contrast to other species, whic produce enzymes with either SA methyltransferase or BA methyltransferase activity [24, 41–43], Arabidopsis contains an expressed SABATH gene encoding an enzyme that can methylate both SA and BA with similar efficiency, and it was therefore designated as BSMT [14].

Similar to JMT, BSMT expression is induced by wounding, hervibory and alamethicin, a channel-forming peptide from the fungus *Tricoderma viridae* that mimics the effect of pathogen attack, suggesting that the function of these genes are induced by stresses that

trigger MeSA emission. Furthermore, in plants grown under standard conditions and profiling different tissues, BSMT transcripts detected by RT-PCR were abundant in sepals of flowers and, in lower levels, in leaf trichomes and hydathodes [14]. In response to treatment with MeJA, wounding, attack of herbivorous insects, and uprooting, BSMT transcripts are accumulated in leaves [14]. This expression pattern is similar to that reported in Genevestigator [32]. According to this repository of Arabidopsis microarray data, BSMT is significantly expressed in inflorescence tissues, particularly in sepals under standard growth conditions. Additionally, BSMT expression is induced in adult rosette leaves treated with either *Pseudomonas syringae* pv. *tomato* or *Botrytis cinerea,* and in experiments where plants are treated with abscisic acid, MeJA or exposed to high light intensity for more than 3 hours.

Several studies have provided evidence that JA negatively regulates the expression of SA-responsive genes in Arabidopsis [44–46]. The antagonistic cross-talk between SA and JA signaling appears to be complex, and these pathways communicate at multiple regulatory points [47]. BSMT1 was induced by MeJA or wounding, as mentioned before, but not by SA [14]. More recently Koo *et al. [47],* found that JA induces BSMT, the activation of BSMT leads to an interruption in the accumulation of SA, which may contribute to an antagonistic effect on SA signaling pathways by depleting the SA pool in plants, suggesting that the BSMT gene may be a crucial switch that mantains a balance between JA and SA signaling pathways.

Is has been suggested that MeSA emission in Arabidopsis plays a role in attracting predators of herbivorous insect pests [48], but the spiked induction of BSMT expression in response to abiotic stress, and the lack of detectable emissions of MeSA or MeBA, suggest that these highly toxic compounds might play a role in direct defense [14]. However, T-DNA insertional null mutants for BSMT have not shown any obvious morphological or physiological phenotype under conditions tested thus far.

2.3.3 *At5g55250* Encodes an Enzyme that Methylates Indole-3-Acetic Acid

The major naturally-occuring auxin is indole-3-acetic acid (IAA). This growth regulating substance is important for many developmental processes during the life cycle of plants. IAA is physiologically active in the form of the free acid, but can also be found in conjugated forms in plant tissues, as IAA-sugar and IAA-amino acid, that are not active but can be converted reversibly to free IAA [40]. In Arabidopsis the SABATH gene *At5g552500* encodes an IAA methyltransferase (IAMT), that coverts IAA into IAA methyl ester (MeIAA) *in vitro,* with a kinetic specific constant ($Kcat/Km$) higher for IAA than for salicylic acid or other tested compounds [28].

Qin *et al.* [49], have found that *IAMT* expression pattern is developmentally regulated in leaves and that over-expression of *IAMT* results in a upward-curling leaf phenotype and

agraviotropic growth in both hypocotyls and roots, suggesting that IAA methylation has a significant effect on auxin homeostasis and plant development. Experiments with *GUS* reporter gene fusions indicated that *IAMT* is expressed ubiquitously in rosette leaves in young seedlings, but after the eighth true leaves emerged the expression of this gene gradually fades away from the center of the leaves and is detected primarily towards the edges of the leaves as leaf development progresses. Apart from leaves, IAMT-*GUS* expression was also detected at high levels in flowers and siliques. According data in the Genevestigator database [32], the expression in siliques is significantly higher than in flowers and rossette leaves. Also based on these data, IAMT is down-regulated in adult leaves infected with *P. syringae,* but up-regulated in plants treated with brassinosteroids or IAA.

Additionally, a T-DNA insertional mutant for *IAMT* (SALK-072125) showed diminished levels of MeIAA without obvious phenotype [33]. A possible explenation for this observation is that other MTs genes may be able to compensate for the loss of function of IAMT, due to the lack of specific catalytic activity of some of the characterized enzymes and because there are still additional, as yet uncharacterized carboxyl MTs in Arabidopsis [23, 28].

Recently the crystal structure of IAMT was determined and refined to 2.75A resolution [50]. The overall tertiary and quaternary structures closely resemble the two-domain bilobed monomer and the dimeric arrangement, respectively, previously observed for the related salicylic acid carboxyl methyltransferase from *Clarkia breweri* (CbSAMT) [50].

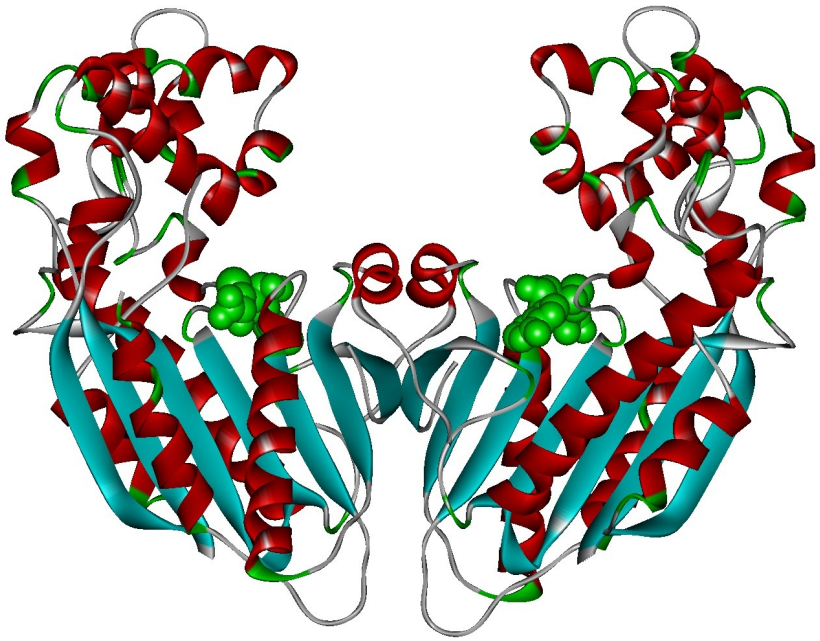


Figure 2.4: Ribbon diagram of the dimeric arrangement of the *Arabidopsis thaliana* IAMT polypeptide backbone.
Highlighted is the binding pocked of IAMT is the methylation product S-adenosyl-L-homocysteine.

2.3.4 *At3g44860* Encodes an Enzyme that Methylates Farnesoic Acid *in vitro*

Tha SABATH gene *At3g44860* encodes an enzyme with high catalytic activity *in vitro* towards farnesoic acid (FA). Although methyl farnesoate (MeFA), an insect juvenile hormone precursor, has not yet been detected in Arabidopsis, the presence of a farnesoic acid carboxyl methyltransferase (FAMT) may indicate that Arabidopsis is capable of producing MeFA, as a presumed defense mechanism against insect herbivory like some other plants [51]. Gene expression pattern studies indicated that transcript levels for this gene increased upon induction of the plant defense response. Based on RT-PCR, the highest levels of *FAMT* transcripts was detected in leaves, compared to roots, stems, flowers and siliques. In Arabidopsis leaves, induction of *FAMT* expression was higher when treated with the fungal elicitor alamethicin than with SA, MeJA or physical wounding [51]. The gene expression pattern found by Yang *et al.* [51], resemble those found based on microarray expression data from Genevestigator, where *FAMT* transcripts were higher in adult leaves than in any other tissue. Furthermore, accoring Genevestigator, *FAMT* gene expression is up-regulated when Arabidopsis plants are subject to stress conditions like wounding, drought, high salt and infection with *P. syrngae* pv. *tomato*, also when plants are treated with MeJA or ethylene inhibitors. MeFA and FA have not been detected in Arabidopsis in spite of the high catalytic activity of FAMT towards FA; perhaps FAMT methylates other yet unknown FA-like compounds [33, 51].

2.3.5 *At5g04370* Encodes an Enzyme that Methylates Nicotinic Acid *in vitro*

Among the 24 members of the SABATH family in Arabidopsis, one enzyme was identified showing high specific activity with nicotinic acid (NA), an important intermediate compound in the salvage bio synthetic pathway of NAD^+, where NA is found in the amide form. NAD^+ is an essential compound in all organisms, both as a coenzyme for oxidoreductase and as a source of ADP-ribosyl groups used in various reactions. The carboxyl methylation of NA, by S-Adenosyl-L-methionine: nicotinic acid methyltransferase (NAMT) to give methyl nicotinate, precludes it from forming the amide. NAMT expression was detected in flowers, roots, and siliques and it was induced by NaCl, ABA, trigonelline, quinolinic acid (a NA analog) and nicotinic acid [33]. According to Genevestigator, *NAMT* is expressed at high levels in the hypocotyl and radicles of seedling, also in roots and seeds of adult plants. NAMT expression is also induced by several stimuli including drought, high salt, osmotic stress and treatmen with abscisic acid. The physiological consequences of carboxyl methylation of NA in Arabidopsis are still unknown [33].

2.3.6 *At4g26420* and *At5g56300* Encode Enzymes that Methylate Gibberellins

Gibberellins are signaling molecules known to play key regulatory roles in a variety of growth and development processes in plants such as germination, leaf expansion, seed development and germination, stem elongation, bolting, flowering and fruit set [52]

Two genes belonging to the SABATH methyltransferase family, At4g26420 and At5g56300, respectively designated *GAMT1* and *GAMT2*, encode enzymes that catalyze the formation of the methyl esters of gibberellins (GAs) *in vitro.* Sequence analysis indicated that GAMT1 and GAMT2 are more similar to each other than to any other member of the SABATH methyltransferase family in Arabidopsis, having 58% identity. GAMT1 showed MT activity with several GAs, such as GA9, GA20, GA4, GA3 and GA1, from high to low catalytical activity, while GAMT2 was optimally active with GA4. In Arabidopsis, tobacco and petunia, constitutively expressing GAMT1 or GAMT2 under the control of 35S-promoter, resulted in plants that showed typical GA-deficiency phenotypes, such as dwarfism, reduced fertility. Proportionally to the severity of the GA deficiency phenotype the levels of precursor and bioactive GAs were lower in GAMT mutants compared to the wild-type. Plants over-expressing GAMT2 contained lower levels of bioactive GA4 than wild-type plants did, but, in general, the depletion of endogenous GAs in GAMT2-overexpressing lines was not as severe as those in GAMT1-overexpressing lines [53].

In Arabidopsis, GAMT1 overexpressing plants are dark green dwarf plants with shortened pedicels and petioles and are sterile. Supplementation with bioactive GAs partially restore the fertility of this plants. GAMT2-overexpressing plants in addition to the dwarf habit, exhibited a bushy phenotype with extensive branching and auxiliary floral meristem. GAMT1, but not GAMT2-overexpressing plants can recover normal fertility levels when they are sprayed with GA_4 [33, 53].

Gene expression studies with RT-PCR indicated that *GAMT1* and *GAMT2* are expressed predominantly in siliques, with very low transcript levels in other tissues, with the exception of germinating seeds, where a weak but detectable expression was observed. *GAMT1* and *GAMT2* transcripts levels in siliques at different stages of development showed that expression of both genes began at early stages of silique development, peaked in the second half of this process, and then began to decrease after the start of desiccation [53]. Data from Genevestigator [32], showed the same expression profile for GAMT2, with expression predominately in developing seeds. Because of a mistake in the annotation of GAMT1 in TAIR, and the fact that based on this erroneous annotation all 11 oligonucleotides chosen by Affimetrix to represent GAMT1 on the ATH1 microarray are derived from sequences downstream of GAMT1 that are not part of the GAMT1 gene, there is no relevant information on the expression of GAMT1 on these websites. The corrected and updated annotation of GAMT1 (At4g26420) was made available from TAIR and the National Center for Biotechnology Information (NCBI) on the Arabidopsis genome release (TAIR7) January, 2007.

2.4 SABATH Methyltransferases in other species

S-adenosyl-L-methionine dependent carboxyl methyltransferases have been identified in several plant species other than Arabidopsis [24, 43, 54, 55]. Flower-specific benzenoid carboxyl MTs have been characterised in *Stephanotis floribunda* and *Nicotiana suaveolens* [54]. Pott *et al.* [43], has shown that *S. floribunda* SAMT and *N. suaveolens* BSMT, are responsible for the differential emission of MeSA and MeBA observed in the floral scents of both species. This two enzymes are examples of specificity and promiscuity in the specialized metabolism, both *S. floribunda* SAMT and *N. suaveolens* BSMT are able to methylate a range of other benzenoid-related compounds, an in the case of *S. floribunda* SAMT, also several cinnamic acid derivatives. This observation seems to be consistent with the structure of the active site of these two enzymes compared to the *C. breweri* SAMT [54], supporting the hypothesis that less specific or recently evolving enzymes may carry out the same chemical reaction with several different substrates, adapting to changes in substrate availability [56, 57].

More recently, an enzyme capable of methylating both BA and SA was identified in *Oryza sativa*, and therefore termed *OsBSMT1 [47].* Arabidopsis plants over-expressing *OsBSMT1* accumulate higher levels of MeSA and MeBA, and when challenged with bacterial or fugal pathogens, they fail to accumulate SA and SA-glucoside, becoming more susceptible to disease compared to wild-type plants. However, neighboring wild-type plants free of bacterial or pathogen infection, were able to trigger defense response mechanisms induced by the airborne signal MeSA [47].

To identify the complete SABATH gene family from the fully sequenced rice genome (International Rice Genome Sequencing Project, 2005), the protein sequence of *C. breweri* SAMT was used to search the rice genome sequence database using the BLASTP algorithm iteratively. A total of 41 sequences encoding proteins bearing significant similarity to known SABATH proteins were identified in the rice genome, with conspicuous expression patterns in different rice tissues including roots, stems, leaves, panicles and germinating seeds, but detailed expression analysis of OsSABTH genes is still needed to understand the biological roles of these genes [50]. Among these OsSABATH genes, in addition to OsBSMT1 [47] mentioned above, another enzyme was recently characterized as OsIAMT1 based on its high sequence similarity to AtIAMT1 [50].

Zhao *et al.* [58], recently isolated and characterized the gene PtIAMT1 from poplar (*Popolus trichocarpa)*, based on sequence similarity to Arabidopsis IAMT. The presence of PtIAMT transcripts in leaves, roots and stems suggests involvement of this gene in development of these organs in poplar. Negre *et al.* [41], isolated and characterized a cDNA that encodes a SAMT from *antirrhium majus,* which according to induction experiments with salicylic and jamonic acid, may serve a physiological role in defense by producing methyl salicylate. Several enzymes with FAMT activity have been isolated and characterized in non-plant species [59, 60], however neither of these enzymes have sequence similarity to each other or to Arabidopsis FAMT, suggesting that FA methylation arouse independently in multiple

evolutionary events [61]. Kapteyn *et al.,* using a biochemical genomics based approach, identified the enzyme responsible for formation of methyl cinnamate in sweet basil [62]. This enzyme, cinnamic acid carboxymethyl transferase (CACMT), like all in the SABATH family uses SAM as methyl group donor. A cDNA encoding CACMT was found by identifying an EST with sequence homology to known carboxyl methyltransferases such as AtBSMT, AtSAMT, and AtJMT. CACMT was only expressed in basil lines that produce methyl cinnamate.

In addition to *O*-methyltransferases, the SABATH family includes a number of enzymes with N-methyltransferase activity, which targets the nitrogen atom in several small molecules involved in caffeine biosynthesis [30]. Caffeine (1,3,7-trimethylxanthine) is an alkaloid secondary metabolite produced by certain plant species, with coffee (*Coffea arabica*) and tea (*Camellia sinica*) being the best known. Caffeine is synthesized from xanthosine after three nitrogen-directed methylation steps, catalyzed by xanthosine MT (XMT), 7-methylxanthine MT (theobromine synthase), and 3,7-methylxanthine MT (caffeine synthase), respectively [29]. Based on sequence analysis, N-methyltransferases involved in caffeine biosynthesis [27, 30, 63, 64], are more closely related to SABATH carboxylmethyltransferases than to others plant N-methyltransferases, that are involved in the biosynthesis of nicotine and coclaurine [65]. Thus, the carboxylate *O*-methyltransferases constitute a much larger structural family than originally anticipated, containing not only carboxyl group specific *O*-methyltransferases but also some alkaloid N-methyltransferases.

2.5 Possible Roles For Methylation of Signaling Molecules

Considering the chemical activities of carboxyl MTs, we can see that by converting bioactive acid signaling molecules to their methyl esters these enzymes can control both the functional activity and the cellular location of reaction products. Many signaling molecules are methylated as a result of the catalytic activity of carboxyl MTs, and like other conjugates, the methyl esters of these compounds are inactive [37, 53]. Unlike other carboxylic acid conjugates, methyl esters are volatile compounds, airborne molecules that can diffuse throughout plant tissues without the help of active transport. After reaching the target tissue of action or even neighboring plants, methyl esters can be hydrolyze back to the bioactive acid form by carboxyl esterases [39]. Despite that several carboxyl esterases for the methyl ester resulting from the catalytic activity of SABATH MTs are still unknown, it is possible to speculate that they exist and are antagonist to carboxyl MTs. With MTs or carboxyl esterases ideally placed, the activity of signaling molecules can be regulated through selective hydrolysis. Specific examples of this molecules are SA, JA and IAA with their respective methyl ester conjugate forms.

Many reports indicate that MeSA is produced in and emitted from vegetative parts of plants upon onset of systemic acquired resistance and during herbivory in general [21, 48]. While in some cases the MeSA may help attract parasitic wasps that prey on the caterpillars inflicting

tissue damage [66], in some other cases it may act as a signal to trigger defense response mechanism systemically in other tissues or neighboring plants, as it occurred in tobacco leaves during tobacco mosaic virus (TMV) infection [21]. SA is an essential component of the signal transduction pathway in SAR [20,67]. In tobacco an enzyme originally designated as SA binding protein (SABP2), posses methylsalicylate esterase activity, capable of de-esterifying MeSA back to SA. When this SABP2 expression is suppressed, tobacco plants infected with TMV cannot initiate an effective SAR [39]. Likely, MeSA itself is not capable of eliciting a response in the target cell, and it must to be converted back to SA to be unset defense response, this is supported by the fact that SABP2-silenced plants fail to develop syatemic acquired resistanse and have suppressed local defense responses [38]. Despite of MeSA not being bioactive, it is more non-polar than SA and capable of diffusing through membranes and transported by air [38, 39].

Methyl jasmonate, JA and its amino acid conjugates, collectively referred as jasmonates, are important cellular regulators mediating diverse developmental processes, including root growth, pollen production and plant resistance to insects and pathogens [67]. MeJA is one of the JA metabolites proposed to play an important role in inter and intra-plant signaling [25, 36, 67, 68]. When plants are exposed to volatile MeJA, they quickly elicit a series of JA mediated defense responses, like MeSA, this methyl ester is much more non-polar that its free acid and more readily penetrates into plant tissue [14, 67, 69]. Moreover, MeJA treatment is the most common means of eliciting herbivory resistance in plants, because it was generally assumed that once MeJA enters the cell, it is de-esterified back to the bioactive JA. However, only recently have several enzymes showing methyljasmonate esterase activity been characterized [37, 70]. We have examined the overexpressing line constructed by Seo *et al.* [25] by DNA microarray analysis and have found that a member of the Arabidopsis esterase family, encoded by *At5g58310* is induced 4-fold in the transgenic plants. A cDNA of *At5g58310* was expressed in *E. coli* and the protein was shown to have methyl esterase activity with MeJA and MeIAA but no with MeSA. The augment in disease resistance of plants overexpressing *JMT* observed by Seo *et al.* [25] may in consequence be due to the high levels of MeJA being mobilized and converted back to JA to execute its regulatory function in the target cell.

Pathways that regulate IAA activity in plants are complicated and poorly understood [71]. Based on the recently identified IAA methyltransferases in different species, IAMT activity appears to be highly conserved among not so recently divergent plants, adding a new issue of complexity to this network. Despite the high sequence similarity, methylation of IAA is likely to have multiple biological consequences. AtIAMT, OsIAMT1 and PtIAMT1 all exhibited evident expression in multiple tissues [49, 50, 58], implying that IAMT activity may be involved in a wide range of biological processes. Although the mode of action of IAA is not well understood, we can speculate that by regulating the homeostasis of IAA, IAMTs are involved generally in initiation and elongation of different tissues during plant development, acquiring lineage-specific roles in different plant species. In addition to the free acid form, IAA exists in the cell as conjugates, which can be hydrolyzed back to IAA

and are active in auxin bioassays. Like IAA, the IAA-ester and IAA-amide conjugates are taken up and effluxed via auxin polar transport. However, MeIAA is essentially non-polar with no auxin activities per se, providing a distinct way to regulate IAA activity. MeIAA like other IAA conjugates may be hydrolyzed into IAA by a esterase [49]. Li *et al.* [72], recently found that MeIAA unlike other IAA conjugates, entered plant cells by diffusion and produced phenotypes related with elevated IAA, indicating that MeIAA is probably converted to IAA by a methyl esterase after diffusion, which represent a rapid and effective way to modulate local IAA concentrations and gradients. In fact, several esterases recently identified belonging to the AtMES esterase family in Arabidopsis, exhibit high activity with MeIAA, indicating that manifestations of MeIAA *in vivo* activity are due to the action of free IAA that is generated from MeIAA upon hydrolysis by one or more plant esterases [40].

It is generally known that GAs are synthesized during seed development and during germination [73], but little is known about the distribution and specific roles of GAs in Arabidopsis developing siliques and seeds [74, 75]. Methylesters of GAs have been shown to be inactive in several angiosperms species [76, 77]. The observation that transgenic Arabidopsis overexpressing GAMT1 and GAMT2 have lower levels of GAs, showing phenotypes characteristics of GA-deficiency, which can be partially reverted by the external application of GAs, suggest that MeGAs are inactive in Arabidopsis as well [53]. It is possible that methylation of GAs is a step in the irreversible deactivation of GAs, perhaps tagging them for further catabolism by epoxidation, oxidation, or glycosylation, reactions known to occur in other parts of the plant or other plant species [78, 79]. Methylated GAs are more hydrophobic than their free acid forms, therefore methylation may be required for the hormone to diffuse throught membranes and reach the cellular compartment of degradation. Furthermore, the expression peak of GAMT1 and GAMT2 at the later stages of silique development when the embryos are fully developed, nearing the end of fast cell division and elongation, a process controlled by GAs, suggest that GAMTs exert inactivation of GAs *in planta* [53]. Different from MeSA, MeJA and MeIAA, where carboxyl methylation is known to be a reversible process due to the catalytic activity of esterases which hydrolyze them back to the free acid form, there is not available evidence that a similar phenomena occurs in MeGAs [53]. Hence, all available information suggest that methylation of GAs in Arabidopsis siliques and seeds in an irreversible process to inactivate gibberellins.

There is no clear understanding of what the role of methylating FA may be in Arabidopsis plants. In insects FA is methylated on its carboxyl group and the resulting MeFA is converted to Juvenile hormone III (JH III) [59]. MeFA and JH III have been isolated form the Malaysian plant *Cyperus ira,* and it was proposed that the high content of JH III plays a role in the plant defense against insects thought its ability to interfere with insect development [80]. However, in Arabidopsis attempts to measure insect juvenile hormone-like compounds such as FA or MeFA have not been successful. If these compounds are produced, they may be present at very low levels and their production may be subject to temporal and spacial regulation [61]. In spite of the large number of chemical compounds tested for *in vitro* activity with AtFAMT and the high catalytic activity with FA, we should not rule out the possibility that AtFAMT

may be capable of methylating phytochemical structures similar to FA.

Carboxyl methylation and ester hydrolysis of small molecules dramatically alter the bioactivity of both xenobiotics and natural products in plants and can be mediated rapidly. Examples of this have been presented above when considering the roles of SABATH MTs and their counteracting esterases in regulating phytohormone bioactivation, turnover and signaling. There is also a clear role for SABATH MTs in modulating the transport of biologically inactive forms of signaling molecules to exert their function in target cells. With the recent identification of SABATH methyltransferases in rice and other plants, there are more possibilities to explore whether these enzymes methylate other signaling molecules and if plants have the corresponding carboxyl esterases to reverse such reactions. No doubts that over the coming years many more SABATH MTs will be added to the list and new classification systems will evolve based on their diversification in function. With the wide range of enzymatic activities observed for different members of the SABATH family of methyltransferases in Arabidopsis and other plant species, we speculate that they are derived from a super-family of proteins that are able to assume multiple functions based on the plasticity of their structure and their enzymatic activity.

Chapter 3

Modulating the Expression of SABATH Methyltransferases: Consequences on the Genome-wide Expression Profile of *Arabidopsis thaliana*

"So you do the experiment, and you stare at it and you say, Now, does it mean anything? What is the world really going to look like? The messenger experiment had that quality"

Walter Gilbert

3.1 Introduction

Living cells interact with their environment and respond to external signals. They adapt their metabolic processes in a highly regulated fashion. mRNA is the first step towards protein synthesis and, for many genes, the amount of produced proteins is thereby mainly regulated by the amount of the corresponding mRNA.

High-throughput analysis of differential gene expression with microarrays is a powerful tool for discovering novel genes or for gaining additional information about certain biological processes on a genomic scale. Transcriptional profiling using microarrays has developed into the most prominent tool for functional genomics and has convincingly demonstrated how information from raw sequence data can be converted into a broad understanding of gene function [81].

The aim of microarray technology is the measurement of the levels of thousands of mRNA molecules at once, thus providing insight into the transcriptional state of a cell (transcriptome). Two types of microarrays are commonly used for expression profiling in plants: oligonucleotide and DNA spotted microarrays. Microarray analysis is an important tool in the deciphering of biotic and abiotic stress-responses in plants [82,83]. Moreover, microarrays can also be used to analyze changes in the transcriptome of transgenic plants [84, 85]. For a specific cell status, mRNA levels corresponding to all transcribed genes are measured in parallel. To this end, mRNA prepared from cell samples is reversibly transcribed to cDNA by enzymes with simultaneous incorporation of marker molecules. The marker molecules later allow for a quantitative analysis of the cDNA molecules. Marker molecules based on radioactive or fluorescent labeling, are commonly used. The labeled cDNA molecules are then applied to the microarray [83, 86], where they hibridize to the target molecules fixed onthe array matrix, and based on the amount of probe hybridized to a target we can estimate the expression level of a particular gene. Monitoring the transcriptome allows to assign putative function to unknown genes. Considering the vast number of scientific reports as indicator, transcriptional profiling with microarrays appears to be the most powerful and versatile technology for functional genomics [87].

Despite the use of DNA microarrays analysis as a high-throughput method to accelerate the characterization of gene function and decipher complex gene expression patterns, there are no reports of genome-wide transcription profiling related to modification in the expression

levels of Arabidopsis SABATH genes. Here we report a transcriptomics study based on oligonucleotide DNA chip microarrays to understand the transcriptional profile of Arabidopsis under normal and aberrant levels of expression for several SABATH MTs genes.

3.2 Literature Review

3.2.1 Microarray Technology - Affymetrix GeneChip

Developed by Schena *et al.* [88], microarray technology was initially based on DNA spotted arrays, where mRNA relative expression levels are measured by co-hybridization of cDNAs derived from two different mRNA preparations (e.g., control and treatment), labelled with two fluorescent dyes. Microarray have also been developed using photo-lithographic synthesis *in situ* of gene-specific oligonucleotides, allowing of up to 300,000 oligos/cm^2 (Affymetrix, Santa Clara, CA).

The Affymetrix GeneChip™Arabidopsis ATH1 Genome Array, was used in this research to study the function of several SABATH methyltransferases. This GeneChip designed by Affymetrix in collaboration with TIGR, contains more than 22,500 probe sets representing approximately 24,000 genes based on information from the international Arabidopsis sequencing project that was formally completed in December, 2000. In Affymetrix GeneChip up to 20 short oligonucleotides (probe set) are used to probe each gene, and probe sets targeting the same gene, but different portions of the gene's coding region, are distributed in various locations of the chip [89]. Arabidopsis ATH1 Genome Array contains probe sets of 11 short oligonucleotides (25-mer), with more than 22,500 probe sets representing 24,000 gene sequences. Each target hybridization oligo (PM: perfect match) is associated with a "negative control" oligonucleotide (MM: miss match) to evaluate the target hybridization specificity. MM probes are created by changing the middle (13th) base of the PM probe with the intention of measuring non-specific binding. A strong hybridization signal on the MM probe indicates the presence of cross-hybridizing targets. However, there are reports that MM probes also hybridized the specific target gene, making their capacity to serve as non-specific hybridization controls, questionable [89].

Affymetrix GeneChips offer several advantages over DNA spotted arrays [90]. However, hybridization to GeneChips generates absolute intensity values describing the mRNA expression level in the one target, requiring at least two GeneChips have to be used for a simple experiment (control vs treatment) [89].

3.2.2 Methods for Microarray Data Analysis

The main issue in a microarray experiment is data analysis and the subsequent extraction of biological knowledge. The limited number of possible experiment replications and large number of variables, makes data analysis in microarrays experiments and other "*omic*" techniques, complicated. Unfortunately, is not possible to identify a single piece of software that is accepted by the scientific community as the gold standard in microarray data analysis [91]. However, there is a consensus that, independently of the computational tools, a microarray data analysis can not be performed without quality control, data processing (normalization), differential expression detection, and extraction of biological knowledge [90,91]. In the research described on this thesis, quality control, normalization and differential expression detection was performed with Bioconductor (`www.bioconductor.org`), a microarray analysis suite based on R package (`cran.r-prject.org`), which is an integrated suite of software tools for data analysis and graphical display.

3.2.2.1 Quality Control

Quality assessment should be carried out during the initial analysis of any Affymetrix GeneChip data set. To evaluate the quality of the GeneChips before proceeding to downstream procedures, we used `affyPLM` package from Bioconductor which provides a number of useful tools based on probe-level modeling procedures. A probe level model (PLM) is a model that is fit to probe-intensity data. More specifically, it allows the analysis of the data at the probe level and chip level, on a probeset by probeset basis, arranging the probe-intensity data for a probeset so that the rows are probes and the columns are chips [92, 93].

Several quality diagnostics tools are available in `affyPLM`. We used Normalized Unscaled Standard Errors (NUSE) method. The standard error estimates obtained for each gene on each array are taken and standardized across arrays so that the median standard error for that gene is 1 across all arrays. This process accounts for differences in variability between genes. The array with elevated SE relative to other arrays is typically of lower quality. Box plots of these values, can be used to compare arrays (`http://bmbolstad.com/`).

3.2.2.2 Data Pre-processing

Affymetrix GeneChips generates data at the probe level. The raw fluorescence signal has to be summarized or transformed into gene level signal and normalized for experimental errors. The probe set summary methods currently used widely by the scientific community are mainly model-based methods. These methods model probe set summaries using the information derived from a multi-array experiment [91].

Robust Multi-Array Normalization

The Robust Multi-array Average (RMA) method for computing an expression value begins by computing background-corrected PM intensities for each PM cell on every GeneChip. The exact details of the background correction method are unpublished, however a description can be found in Irizarry *et al* [93]. Background-corrected and log-transformed PM intensities are normalized using the quantile normalization method. In the quantile normalization, the highest background-corrected and log-transformed PM intensities are determined. Then, these values are averaged, and the original individual values are replaced with the average. This process is then repeated with the second highest intensity on each GeneChip, the third highest, etc.

Following quantile normalization, an additive linear model is fit to the normalized data to obtained a expression measure for each probe on each GeneChip. The particular model for a probe set can be written as

$$Y_{\mathrm{ij}} = m_{\mathrm{i}} + a_{\mathrm{j}} + e_{\mathrm{ij}}$$

where Y_{ij} is the normalized PM probe value corresponding to the i^{th} GeneChip and the j^{th} probe within the probe set, m_i is the log-scaled expression for the probe set in the i^{th} GeneChip, a_j is the probe affinity effect for the j^{th} probe within the probe set, and e_{ij} denotes the random error term. Ultimately, the GeneChip-specific log-scaled expression values (m_i), estimated by Turkey's median polish are reported as the RMA measures of expression for this probe set or gene [92, 93].

GeneChip Robust Multi-Array Normalization

RMA uses a global background adjustment based on a PM-only estimation procedure and does not take into account non-specific binding in probe set background calculation. Instead, GeneChip Robust Multi-Array Normalization (GCRMA) converts background adjusted probe intensities to expression measures using the same normalization and summarising methods as RMA, but using probe sequence information [94].

GCRMA uses probe sequence information to estimate probe affinity for non-specific binding. The sequence information is summarized in a more complex way than the simple GC content. Instead, the base types (A,T,G or C) at each position (1-25) along the probe determine the affinity of each probe. The parameters of the position-specific base contributions to the probe affinity is estimated in an non-specific binding experiment in which only non-specif binding but no gene-specific hybridization is expected. GCRMA performs better for lower expressed genes and RMA for high ones [94]. Because in most hybridization experiments few genes fall in the high-expressed category here we used GCRMA for data pre-processing.

3.2.2.3 Detection of Differentially Expression

Extraction of biological information is the most important goal of microarray data analysis, and requires appropriate statistical methods. In the early days of microarrays, a simple fold change cutoff was used to determine differential expression. Today, this practice is generally unacceptable.

The main issue in differential expression analysis is the extremely high number of observations for each sample and the limited number of replicates. Modifications to traditional analysis of variance (ANOVA) and t-test have been made in order to adjust for multiple measures and small sample sizes [95, 96]. To identify differentially expressed genes we used the `limma` (Linear Models for Microarray Data) [97] package, part of the Bioconductor project. Based on `limma`, the design of any microarray experiment can be represented in terms of a linear model for each gene. The approach is outlined in Smyth *et al.* 2004 [97]. The approach requires two matrices to be specified. The first is the design matrix which indicates in effect which RNA samples have been applied to each array. The second is the contrast matrix which specifies which comparisons to make between the RNA samples. To test if the differences in expression values for each gene between two groups (knockout vs wild-type, or overexpress vs wild-type) is significant we used the moderated t-test which is based on empirical Bayesian statistics and is more powerful than the standard unpaired t-test when the number of replicates is low [97].

For microarray experiments with more than two groups (wild-type, knockout, overexpress, double knockout), where we were interested in identifying genes that were differentially expressed in one or more groups relative to the others we used one-way ANOVA. This allows to determine whether the variability observed in expression value for a particular gene across the different groups is due to the experimental differences or is solely due to random sources [89].

Using, both linear models and ANOVA, differentially expressed genes are selected based on hypothesis testing and probability statistics. The probability that a gene is differentially expressed is denoted by the p-value. The smaller the p-value, the less likely that the observed differentially expressed gene have occurred by chance. However, due to the large number of observations on a GeneChip (more than 24,000 ATH1 Affymetrix Genome array), traditional threshold of p=*0.05* will identify 5% (1,200 genes) by chance alone. This is a problem for microarray data analysis, which is common to other "omic" approaches where the number of false discoveries can increase significantly when the numbers of test is large.

Several methods have been developed to adjust the probability statistic for multiple measures and small sample sizes. False Discovery Rate (FDR) [98] is a widely accepted approach to control the number of false positives resulting from multiple testing. This procedure makes use of ordered p-values $P_{(1)} \leq \ldots \leq P_{(\mathrm{m})}$, with corresponding null hypotheses $H_{(1)}, \ldots, H_{(\mathrm{m})}$. For a desire FDR level q, the ordered p-value P_{i} is compared to the critical value $q \cdot i/m$. Let k be the maximum i for which $P_{(i)} \leq q \cdot i/m$. Then reject all $H_{(1)}, \ldots, H_{(\mathrm{k})}$ if such a k

exist. Benjamini and Hochberg [98] show that when the test statistics are independent, this procedure controls the FDR at the level q.

3.2.2.4 Biological Knowledge Extraction

Extracting clear and coherent hypotheses from transcriptomics data remains an important challenge. The main difficulty in the analysis consist not in the identification of differentially expressed genes but in their interpretation. The problem is more complicated when the pathway of interest involves moderate effects that are not captured by the genes at the top of the list of differentially expressed genes. In the current study, rather than looking at individual gene functions we focused on the discovery of biological pathways to provide an overview of the underlying processes that are affected as a result of modulating the expression of SABATH methyltransferases.

Certain biological categories within a set of differentially expressed genes derived from a gene expression profiling experiment, are prevalent even when the profiling is done using different transcript selection methods [99]. Therefore, even though not all the genes in a particular pathway might be found in the differentially expressed genes list, they all represent subsets of transcripts associated to a specific biological pathway or functional category [99].

The Gene Ontology (GO), this is perhaps the most commonly used database for functional annotation of transcription profiling. Because of the parent-child relationship based on which the GO is structured, a particular transcript is often associated with multiple categories and a large number of GO categories might appear significant, producing redundant information than can be difficult to interpret [91]. In addition to the GO, transcription profiles can be explored with biological knowledge databases like the Kyoto Encyclopedia of Genes and Genomes (KEGG), to map differentially expressed transcripts in specific biological knowledge domains.

Two of the most used statistics to evaluate association between functional pathways and differential expression are the one-tailed Fisher exact test (FET) [99–101] and Gene Set Enrichment Analysis (GSEA) [102, 103]. FET is a statistical significance test used in the analysis of categorical data where the sample size is small. The test is used to examine the significance of association between two variables in a 2X2 contingency table [99–101]. GSEA on the other hand, evaluates microarray data at the level of "gene sets". With the gene sets defined based on prior biological knowledge, e.g. published information about biochemical pathways or co-expression in previous experiment [102, 103]. To our knowledge all tools available for pathway analysis in Arabidopsis are based on FET rather than GSEA.

3.3 Methods

3.3.1 Plant Material and Growth Conditions

Wild-type *A. thaliana* L. ecotype Col-0 was used on this study for all experiments. Plants were grown in soil in growth chambers under long day conditions (16 h light at $150uE/8$ h dark cycle) at $22^{o}C \pm 1^{o}C$. Seeds for all knockout and overexpressing lines for the AtSABATH were provided by Dr. Eran Pichesrky from the Department of Molecular, Cellular and Developmental Biology at the University of Michigan, Ann Arbor, as part of a collaborative NSF2010 project. Knockout mutant lines with Col-0 background were obtained from Arabidopsis Biological Resource Center [104], Syngenta Arabidopsis Insertion Library or SAIL (formerly GARLIC) Collection [105] and GABI-Kat [106] and confirmed at Dr. Pichersky laboratory. Overexpressing transgenic plants were obtained by *Agrobacterium tumefasis*-mediated floral dip transformation [107] of wild-type plants (Col-0) with cDNA of each gene in sense orientation, driven by the cauliflower mosaic virus 35S promoter.

3.3.2 RNA Extraction

For RNA extraction we selected the tissues and developmental stages were each particular gene was highly express. Information about the gene expression profiles of AtSABATH methyltransferases was obtained by querying the Genevestigator `www.genevestigator.com` database, which is a repository of high quality microarray data for Arabidopsis and other model organisms. Information about the expression profile of several AtSABATH methyltransferases was obtained from the literature [14, 23, 33, 51]. Consequently, RNA was extracted from flowers for IAMT and BSMT mutants, and leaves for JMT mutants. RNA was also extracted from each one of the tissues mentioned above from the wild-type.

Total RNA from frozen tissue, was extracted with RNeasy Midi kit from Qiagen™. Following the extraction, DNase treatment was performed following standard procedures with Qiagen™ columns. Quality and concentration of the extracted RNA was estimated with 2100 Bioanalyzer from Agilent.

3.3.3 Affymetrix™ Microarray Analysis

Three independent biological replicates were analyzed for each genotype. Labeling and hybridization of RNA were conducted using standard Affymetrix protocols by the Core Laboratory Facilities at the Virginia Bioinformatics Institute. Briefly, ATH1 Arabidopsis GeneChips (Affymetrix, Santa Clara, CA) were used for measuring changes in gene expression levels. Total RNA was converted into cDNA, which was in turn used to synthesize biotinylated cRNA. The cRNA was fragmented into smaller pieces and then was hybridized

to the GeneChips. After hybridization, the chips were automatically washed and stained with streptavidin phycoerythrin using a fluids station. The chips were scanned by the GeneArray scanner by measuring light emitted at 570 nm when excited with 488-nm wavelength light. Data from the chips were compiled using MicroArray Suite 5.0 software.

3.3.4 Microarray Data Processing

Data from the GeneChip experiments were analyzed using the R packages Limma [97] and AffyImGUI [108] available through the Bioconductor Project [109]. The raw signal intensities of the individual probe pairs, from the Affymetrix CEL data files generated for each GeneChip were loaded into the software indicating the source of RNA hybridized on each chip and the replicate arrays. Diagnostic tools to access the quality of the date were used, such as histograms of the raw intensities, RNA degradation plots and MA plots of the perfect match probes. Quality assessment was followed by background correction, normalization and summarizing of the probe-level data into probe-set expression values. These three steps were accomplished with GCRMA algorithm [94]. The normalized probe-set expression values were then used to analyze the differential expression in terms of linear models. For each comparison a number of differential statistics is provided, however we decided to used the FDR [98] adjusted p-value $\leqq 0.05$, as the cut-off to consider a gene differentially expressed.

3.3.5 Pathway Analysis

To further understand the biological role of the genes found differentially expressed based on statistical testing, we performed a pathway analysis using the classification of the Kyoto Encyclopedia of Genes and Genomes (KEGG) biochemical pathways [110] and the Gene Ontology classification (GO) provided by the Gene Ontology Consortium through The Arabidopsis Information Resource (TAIR). Functional annotation and biological category enrichment were performed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) (`http://david.abcc.ncifcrf.gov/`) [101]. DAVID uses the list of differentially expressed gene IDs as input, to estimate whether a pathway or process is significantly enriched with differentially expressed genes when modulating the expression of SABATH MTs. To maximize the amount of useful information we carried out the pathway analysis using three sets of gene IDs: all differentially expressed genes, up-regulated genes and down-regulated genes.

If the genes in a particular GO categories or KEGG pathways are those significantly represented in the list of genes that are differentially expressed for each particular SABATH mutant versus wild-type comparison, that GO category or KEGG pathway is considered significant. Additionally, GO categories and KEGG pathways were separated in up-regulated and down-regulated, if they were enriched with up-regulated or down-regulated genes, respectively. Only the GO categories statistically significant with a p-value $\leqq 0.001$, and KEGG

pathways statistically significant with a p-value $\leqq$0.05 are reported here.

3.4 Results and Discussion

3.4.1 Gene Expression Profiling of JMT Overexpressing Mutant

To examine the effect of gene expression modulation of JMT on the genome-wide transcription profile of *A. thaliana,* we examined the JMT overexpressing line constructed by Seo *et al.* [25], using DNA microarray. RNA was isolated from leaves of 30-day old plants which have fully developed rosettes. In the table 3.1 are shown the KEGG pathways that were significantly up-regulated and down-regulated in JMT overexpressing plants. GO categories enriched with significantly differentially expressed genes and their corresponding p-values are listed in Tables A.2 and A.1.

Table 3.1: KEGG Pathways significantly enriched in Arabidopsis plants overexpressing JMT. Green and red color of p-values indicate whether the pathway is down or up-regulated, respectively.

KEGG Category	***p*-value**
ath00941:Flavonoid biosynthesis	0.00519
ath00040:Pentose and glucuronate interconversions	0.02246
ath00710:Carbon fixation	0.02810
ath00450:Selenoamino acid metabolism	0.03267
ath00500:Starch and sucrose metabolism	0.03367
ath00480:Glutathione metabolism	0.04203
ath00030:Pentose phosphate pathway	0.01643
ath00564:Glycerophospholipid metabolism	0.04070

α-Linolenic Acid Metabolism and the Jasmonate Biosynthesis Pathway

Jasmonates modulate various physiological events such as resistance to pathogens and insects, fruit ripening, maturation of pollen, root growth and senescence [25]. Jasmonates are synthesized from linolenic acid by oxygenation with lipoxygenase (LOX), then converted to 12-oxo-phytodienoic acid by allene oxide synthase (AOS) and allene oxide cyclase (AOC). Jasmonic acid (JA) is synthesized from 12-oxo-phytodienoic acid through reduction by 12-oxo-phytodienoic acid reductase (OPR) and three steps of β-oxidation. JMT methylates JA using S-adenosyl-methionine as methyl donor to produce methyl jasmonate (MeJA). JA and MeJA are accumulated when plants are exposed to different biotic and abiotic stresses [67].

Among the statistically significant differentially expressed genes we found 44 genes that

change their mRNA expression levels in response to jasmonic acid stimulus (Table:3.4). Moreover, the overexpression of JMT, the enzyme responsible for the methylation of JA, induced changes in the expression of the all the genes coding the enzymes up-stream in the JA biosynthesis pathway (LOX2, LOX5, AOS, AOC1, AOC2, OPR1, OPR2, ACX3, ACX4, KAT2, and MFP2) (Table 3.2 and Figure 3.1). Sasaki *et al.* [34] found similar results when wild-type Arabidopsis plants were treated with $30\mu M$ MeJA. This might suggest that MeJA is a key compound in the jasmonate-signalling pathway, capable of controlling its own production by a positive feed back.

S-adenosyl-L-methionine (SAM) is the methyl donor for SABATH MTs. In addition to the genes involved in the jasmonates biosynthesis we found changes in the expression of genes in the methionine biosynthesis pathway, with directionality towards reducing the amount of SAM available for methylation (Fig:3.2). Interestingly, methionine adenosyltransferase (EC 2.1.5.6), the enzyme catalyzing the production of SAM from L-methionine was down regulated, while adenosyl-methionine decarboxylase (EC 4.1.1.50) and aminopropyltransferase (EC 2.5.1.16), enzymes responsible for the transformation of SAM into other subproducts were up regulated (Table 3.3), suggesting that reducing the availability of the methyl donor could reduce the activity of JMT. Furthermore, in another branch of the methionine metabolism, the genes encoding 1-aminocyclopropane-1-carboxylate synthase (EC 4.1.1.14), and aminocyclopropane-carboxylate oxidase (EC 1.14.17.4), were down-regulated, these enzymes convert SAM into ethylene. We speculate that high methylation of JA regulates the ethylene biosynthesis by depleting SAM, but further studies need to be performed to test this hypothesis.

Interestingly, among the different GO categories enriched with differentially expressed genes (Tables A.2 and A.1), we observed that JMT overexpression induced changes in genes that respond to the stimulus of other phytohormones like, salicylic acid, ethylene, abscisic acid and auxins, suggesting that the intricate mechanisms of cross-talk among of the signal transduction pathways of phytohormones is very complicated and yet poorly understood.

Table 3.2: Differentially expressed genes in the jasmonate biosynthesis pathway. The locus identifier is highlighed in green or red for down and up-regulated genes, respectively.

Locus Identifier	Annotation
AT3G51840	ACX4 (ACYL-COA OXIDASE 4); oxidoreductase
AT4G29010	AIM1 (ABNORMAL INFLORESCENCE MERISTEM); enoyl-CoA hydratase
AT2G33150	KAT2/PED1 (PEROXISOME DEFECTIVE 1); acetyl-CoA C-acyltransferase
AT1G19640	JMT (JASMONIC ACID CARBOXYL METHYLTRANSFERASE)
AT3G06860	MFP2 (MULTIFUNCTIONAL PROTEIN); enoyl-CoA hydratase
AT5G42650	AOS (ALLENE OXIDE SYNTHASE); hydro-lyase/ oxygen binding
AT3G22400	LOX5; lipoxygenase
AT1G76690	OPR2 (12-oxophytodienoate reductase 2)
AT1G76680	OPR1 (12-oxophytodienoate reductase 1)
AT3G06690	ACX3 (ACYL-COA OXIDASE 3); acyl-CoA oxidase
AT3G45140	LOX2 (LIPOXYGENASE 2)
AT3G25760	AOC1 (ALLENE OXIDE CYCLASE 1)
AT3G25770	AOC2 (ALLENE OXIDE CYCLASE 2)

Table 3.3: Differentially expressed genes in the methionine biosynthesis pathway. The locus identifier is highlighed in green or red for down and up-regulated genes, respectively.

Locus Identifier	Annotation
AT1G05010	EFE (ETHYLENE FORMING ENZYME)
AT3G01120	MTO1 (METHIONINE OVER-ACCUMULATION 1)
AT3G02470	SAMDC (S-ADENOSYL-METHIONINE DECARBOXYLASE)
AT5G15950	Adenosyl-methionine decarboxylase family protein
AT2G36880	MAT3 (METHIONINE ADENOSYLTRANSFERASE 3)
AT3G17390	MTO3 (S-adenosyl-methionine synthase 3)
AT1G12010	1-aminocyclopropane-1-carboxylate oxidase, putative / ACC oxidase, putative
AT4G37770	ACS8 (1-Amino-cyclopropane-1-carboxylate synthase 8)
AT1G62380	ACO2 (ACC OXIDASE 2)
AT5G18930	BUD2/SAMDC4 (BUSHY AND DWARF 2); adenosyl-methionine decarboxylase
AT1G70310	SPDS2 (SPERMIDINE SYNTHASE 2)
AT1G23820	SPDS1 (SPERMIDINE SYNTHASE 1)

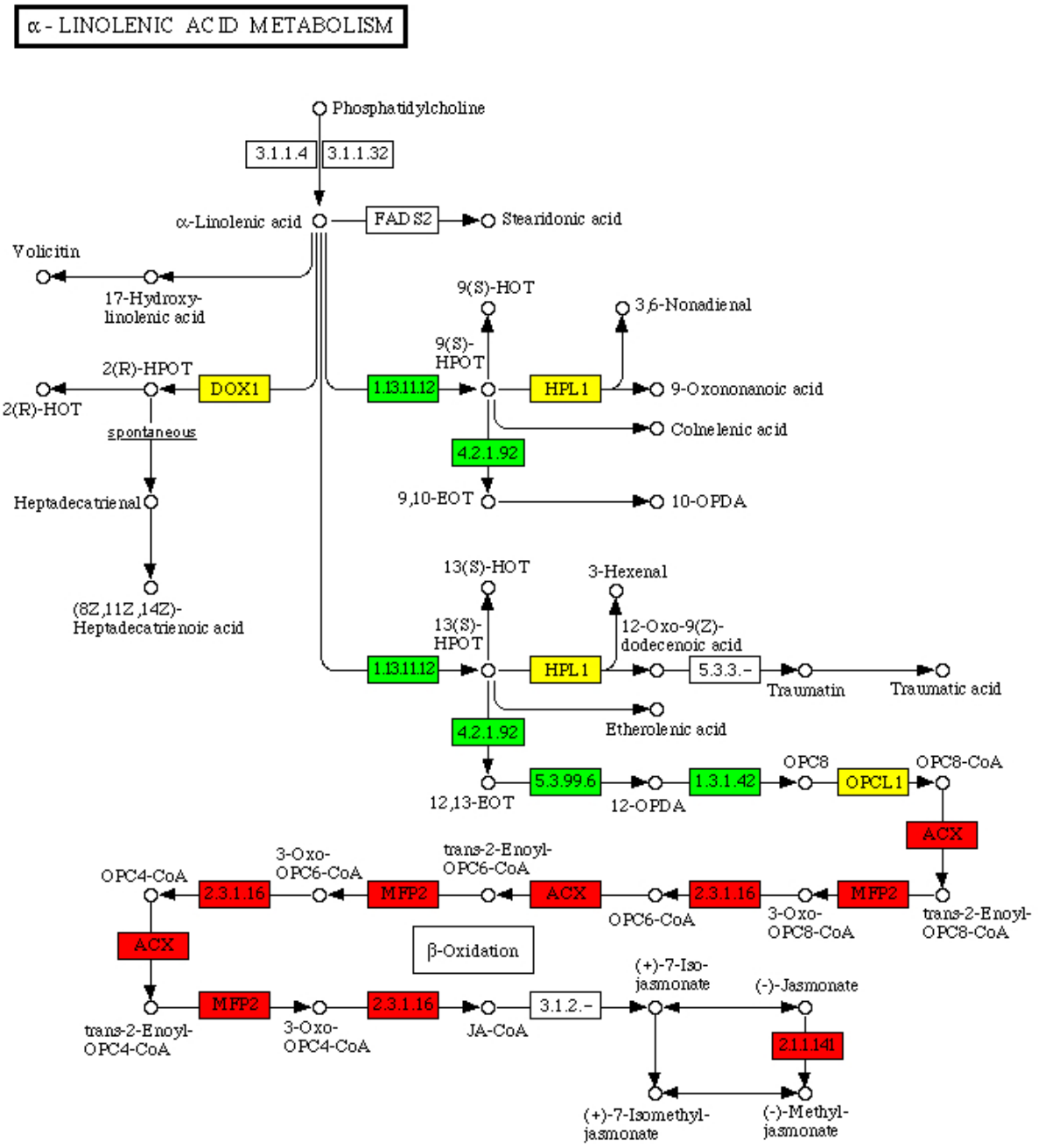


Figure 3.1: α-Linolenic Acid metabolism and the jasmonate biosynthesis pathway. Colored boxes indicate genes identified in the *A. thaliana* genome. Green and red indicate down-regulated and up-regulated genes, respectively, in response to JMT overexpression. Yellow indicates the genes that did not change significantly.

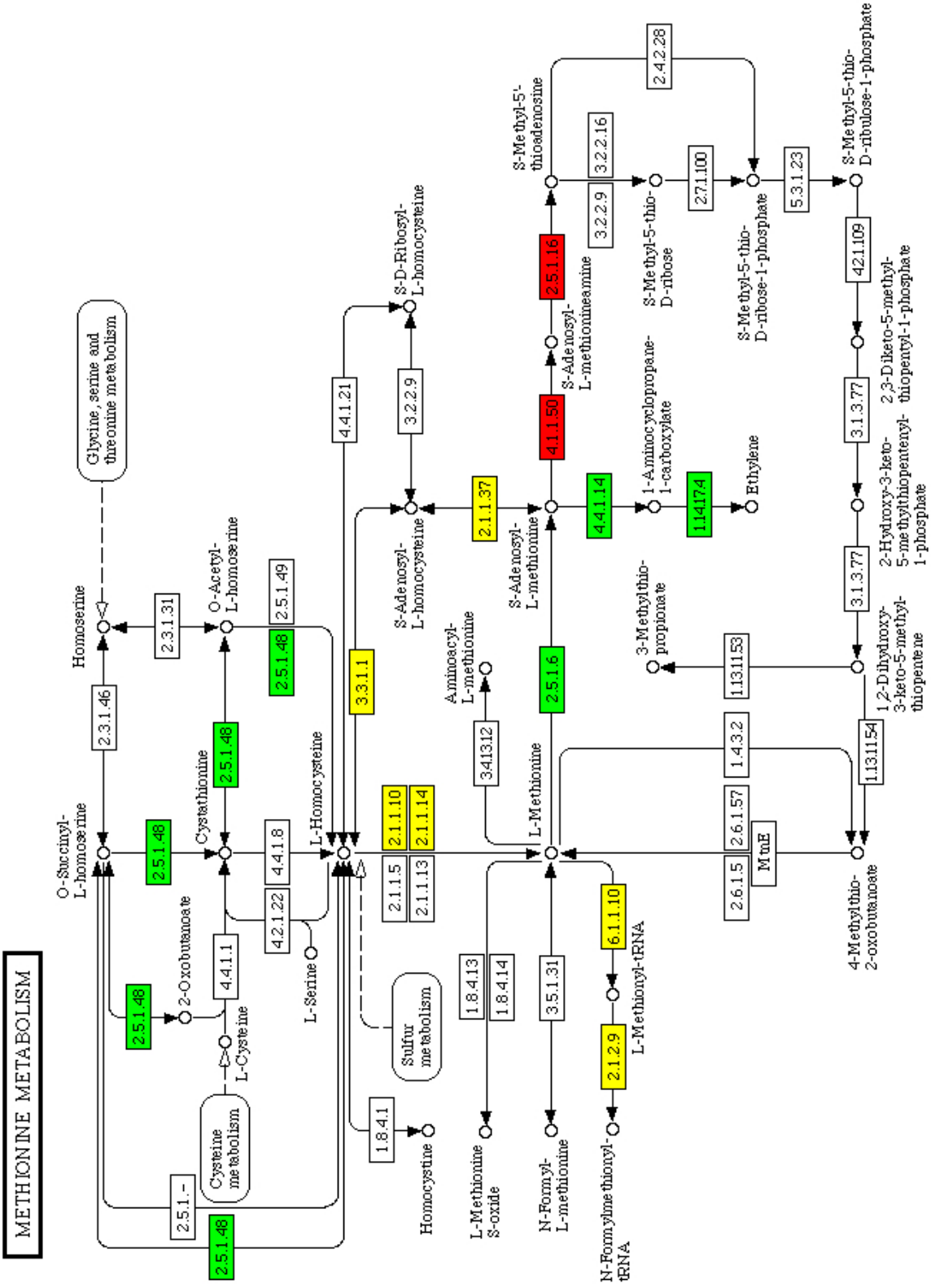


Figure 3.2: Methionine metabolism affected by JMT overexpression. Colored boxes indicate genes identified in the *A. thaliana* genome. Green and red indicate down-regulated and up-regulated genes, respectively, in response to JMT overexpression. Yellow indicates the genes that did not change significantly.

Table 3.4: Jasmonate stimulus-responsive genes regulated by JMT overexpression.

Locus Identifier	Annotation
AT2G34810	FAD-binding domain-containing protein
AT2G16720	MYB7 (myb domain protein 7); DNA binding / transcription factor
AT5G67300	ATMYB44/ATMYBR1/MYBR1 (MYB DOMAIN PROTEIN 44)
AT2G28900	OEP16 (OUTER ENVELOPE PROTEIN 16)
AT3G16350	myb family transcription factor
AT4G34710	ADC2 (ARGININE DECARBOXYLASE 2)
AT5G40440	ATMKK3 (MITOGEN-ACTIVATED KINASE KINASE 3); MAP kinase kinase
AT1G32640	ATMYC2 (JASMONATE INSENSITIVE 1); transcription factor
AT3G09940	ATMDAR3/MDHAR (MONODEHYDROASCORBATE REDUCTASE)
AT1G17520	DNA-binding protein, putative
AT5G62470	MYB96 (myb domain protein 96); transcription factor
AT2G39940	COI1 (CORONATINE INSENSITIVE 1); ubiquitin-protein ligase
AT5G07690	MYB29 (myb domain protein 29); transcription factor
AT4G01280	myb family transcription factor
AT4G26850	VTC2 (VITAMIN C DEFECTIVE 2)
AT4G34990	AtMYB32 (myb domain protein 32); transcription factor
AT4G02570	ATCUL1 (CULLIN 1)
AT5G27380	GSH2/GSHB (GLUTATHIONE SYNTHETASE 2); glutathione synthase
AT1G70000	DNA-binding family protein
AT1G76930	ATEXT4 (EXTENSIN 4)
AT1G22640	MYB3 (myb domain protein 3); transcription factor
AT5G52660	myb family transcription factor
AT3G09600	myb family transcription factor
AT5G60890	ATMYB34/ATR1/MYB34 (ALTERED TRYPTOPHAN REGULATION, MYB DOMAIN PROTEIN 34)
AT1G49950	ATTRB1/TRB1 (TELOMERE REPEAT BINDING FACTOR 1)
AT4G37760	SQE3 (SQUALENE EPOXIDASE 3); oxidoreductase
AT5G13930	ATCHS/CHS/TT4 (CHALCONE SYNTHASE); naringenin-chalcone synthase
AT3G50060	MYB77; DNA binding / transcription factor
AT1G71030	ATMYBL2 (Arabidopsis myb-like 2); DNA binding / transcription factor
AT3G11820	SYP121 (syntaxin 121); SNAP receptor
AT5G37500	GORK (GATED OUTWARDLY-RECTIFYING K+ CHANNEL)
AT2G46830	CCA1 (CIRCADIAN CLOCK ASSOCIATED 1); transcription factor
AT3G46590	TRFL1 (TRF-LIKE 1); DNA binding
AT4G09460	ATMYB6 (MYB DOMAIN PROTEIN 6); DNA binding / transcription factor
AT2G46370	JAR1 (JASMONATE RESISTANT 1)
Continued ...	

Locus Identifier	Annotation
AT1G07640	OBP2 (OBF BINDING PROTEIN 2); DNA binding / transcription factor
AT2G39770	CYT1 (CYTOKINESIS DEFECTIVE 1); nucleotidyltransferase
AT1G74430	MYB95 (myb domain protein 95); DNA binding / transcription factor
AT3G16470	JR1 (Jacalin lectin family protein)
AT2G24850	TAT3 (TYROSINE AMINOTRANSFERASE 3); transaminase
AT1G18710	AtMYB47 (myb domain protein 47); DNA binding / transcription factor
AT1G57560	AtMYB50 (myb domain protein 50); DNA binding / transcription factor
AT3G45140	LOX2 (LIPOXYGENASE 2)

Flavonoid Biosynthesis

Jasmonates induce the production of specialized metabolites in many plant species [111]. Previous macroarray experiments have identified that JA induces some genes for enzymes in secondary metabolism [34, 69]. We found that the flavonoid biosynthesis pathway was significantly represented among the down-regulated genes (Table 3.6 and Figure 3.3), and approximately 20% of genes identified in this experiment encoded enzymes involved in the production of specialized metabolites or compounds such as amino acids that can serve as their precursors, suggesting that one of the principal effects of JA is to control specialized metabolism. Application of MeJA has been reported to induce accumulation of anthocyanins in the Arabidopsis leaves [112, 113], and several genes involved in the production of anthocyanins were up-regulated, including chalcone synthase, anthocyanin synthase, and leucoanthocianidin dyoxygenase [114]. Here, we found that the overexpression of JMT, and a possible increase in the endogenous levels of MeJA, represses the expression of genes in the flavonoid biosynthesis pathway (Fig:3.3). This, perhaps indicate that JA and not MeJA promote accumulation of anthocyanins. We speculate that MeJA applied to the plant is converted back into JA by the action of a methyesterase, and JA in turn triggers the accumulation of anthocyanins by enhancing the expression of genes in the flavonoid biosynthesis pathway. When JMT is overexpressed, most, if not all, of the JA produced by the plant is being converted into MeJA, and the amount of free JA to stimulate the flavonoid biosynthesis is limited or below the minimum threshold.

A number of plant species produce flavonoid compounds as protective agents against environmental stresses, or in response to exogenous MeJA [112, 113, 115, 116]. Rapid and transient induction of flavonoid gene expression also characterizes plant stress responses [114]. We therefore presume that anthocyanins contribute to the characterised JA-mediated responses against biotic and abiotic stress.

Table 3.6: Flavonoid biosynthesis genes down-regulated by JMT overexpression.

Locus Identifier	Annotation
AT5G17050	UDP-glucoronosyl/UDP-glucosyl transferase family protein
AT5G42800	DFR (DIHYDROFLAVONOL 4-REDUCTASE); dihydrokaempferol 4-reductase
AT5G08640	FLS (FLAVONOL SYNTHASE)
AT3G55120	A11/CFI/TT5 (TRANSPARENT TESTA 5); chalcone isomerase
AT3G51240	F3H (TRANSPARENT TESTA 6); naringenin 3-dioxygenase
AT5G54160	ATOMT1 (O-METHYLTRANSFERASE 1)
AT1G30530	UDP-glucoronosyl/UDP-glucosyl transferase family protein
AT5G13930	ATCHS/CHS/TT4 (CHALCONE SYNTHASE); naringenin-chalcone synthase
AT5G07990	TT7 (TRANSPARENT TESTA 7); flavonoid 3'-monooxygenase
AT4G22870	[AT4G22870, leucoanthocyanidin dioxygenase, putative
AT4G22880	[AT4G22880, LDOX (TANNIN DEFICIENT SEED 4)]

Carbon Fixation, Starch and sucrose Metabolism

We found that carbohydrate and energy metabolism are significantly represented among the genes regulated by JMT overexpression. The starch and sucrose metabolism, pentose and glucoronate interconvertions, and carbon fixation pathways, are significantly down-regulated by JMT overexpression. While the pentose phosphate pathway is up-regulated. High levels of JA have been reported in tissues actively importing carbon and nitrogen [117]. Specially, those where large amounts of carbon and nitrogen accumulate as polymers in the form of starch, fructans, and proteins [117, 118].

A role for jasmonic acid in energy metabolism and protein storage in plants was suggested, in part, because jasmonate levels are high in developing tissues. In six-week-old soybean seedlings, JA levels are higher in young growing leaves that are importing carbon and nitrogen than in older fully expanded leaves [118]. Also, another reason to suggest that jasmonate plays an important role in energy metabolism and protein storage during plant development derives from the discovery that genes encoding vegetative storage proteins [117] are regulated by jasmonate. It has also been reported that application of JA to leaves decreases expression of nuclear and chloroplast genes involved in photosynthesis [119]. Based on the tight relationship between photosynthesis and carbon fixation, JA has been proposed to act as inhibitor of premature accumulation of the photosynthetic apparatus proteins in meristematic cells of the leaf base while stimulating accumulation of carbon and nitrogen reserves needed for later cell development [117]. We found that a possible reduction in free JA, due to the overexpression of JMT, repressed the expression of genes involved in energy metabolism and carbohydrate storage (Figures 3.4 and 3.5). In our study, the down-regulation of the carbon fixation, and the starch and sucrose metabolism, support this hypothesis. The limited

amounts of free JA due to high methylation rate to produce MeJA, has a negative effect on the fixation of carbon and accumulation of energy in the forms of starch or sugars.

Glutathione Biosynthesis

Another significantly down-regulated metabolic pathway in JMT overexpressing plants was the glutathione biosynthesis. Glutathione (GSH) plays a pivotal role in protecting plants from environmental stresses, mainly oxidative stress. We found that a reduction of free JA, due to the overexpression of JMT, regulates the expression of genes involved in gluthatione metabolism (Figure 3.6). Xiang and Oliver [120], have shown that jasmonic acid treatment increased mRNA levels of GSH biosynthesis genes. Even thought this study was looking at glutathione accumulation in response to heavy metals, they tested several chemical compounds associated with stress to see if it is possible to elicit the same response, and found that JA, unlike other compounds tested, increased the mRNA levels of all three genes involved in GSH biosynthesis. Xiang and Oliver's [120] findings, together with our observation of down regulation in gluthatione biosynthetic genes as result of low availability of free JA, provide evidence supporting the role of JA in defense against oxidative stress through increased expression of GSH metabolic genes.

It has been reported that, in spite of the high transcripts levels of GSH biosynthetic genes induced by JA, it does not result in GSH accumulation [120]. When MeJA was applied to the Arabidopsis plants it did not alter the glutathione content in unstressed plants [120]. This is not surprising because GSH homeostasis is tightly controlled. Perhaps, under the induction condition, there is no demand for increasing GSH level because oxidative stress was not present. The exact mechanism on how JA stimulates GSH biosynthesis is unknown. We speculate that JA does enhance the capacity for GSH synthesis enhancing responsiveness when an oxidative stress is imposed upon the plant.

Selenoamino acid metabolism / Methionine biosynthesis

We found that selenoamino acid metabolism is down regulated when JMT is overexpressed. In higher plants, Se is mainly taken up from soil in the form of selenate or selenite [121]. There are no reports of a direct relationship between regulation of selenoamino acid metabolism and JA. We believe that the up-regulation observed in the selenoamino acid metabolism is due to the presence of common enzymes with the methionine biosynthesis pathway, which produces S-adenosyl-L-methionine, the methyl donor for JA methylation. Cystathionine γ-synthase (EC 2.1.5.48), S-adenosyl-methionine synthase (EC 2.5.1.6) and methionine adenosyltransferase (EC 2.5.1.6), participate in both pathways and are down-regulated. This caused the Se-methylation inhibition by a reduction in the SAM level due to the repression of methionine adenosyltransferase activity. Therefore the up-regulation of the selenoamino acid metabolism is indirect consequence of the depletion of SAM due to the high activity of JMT.

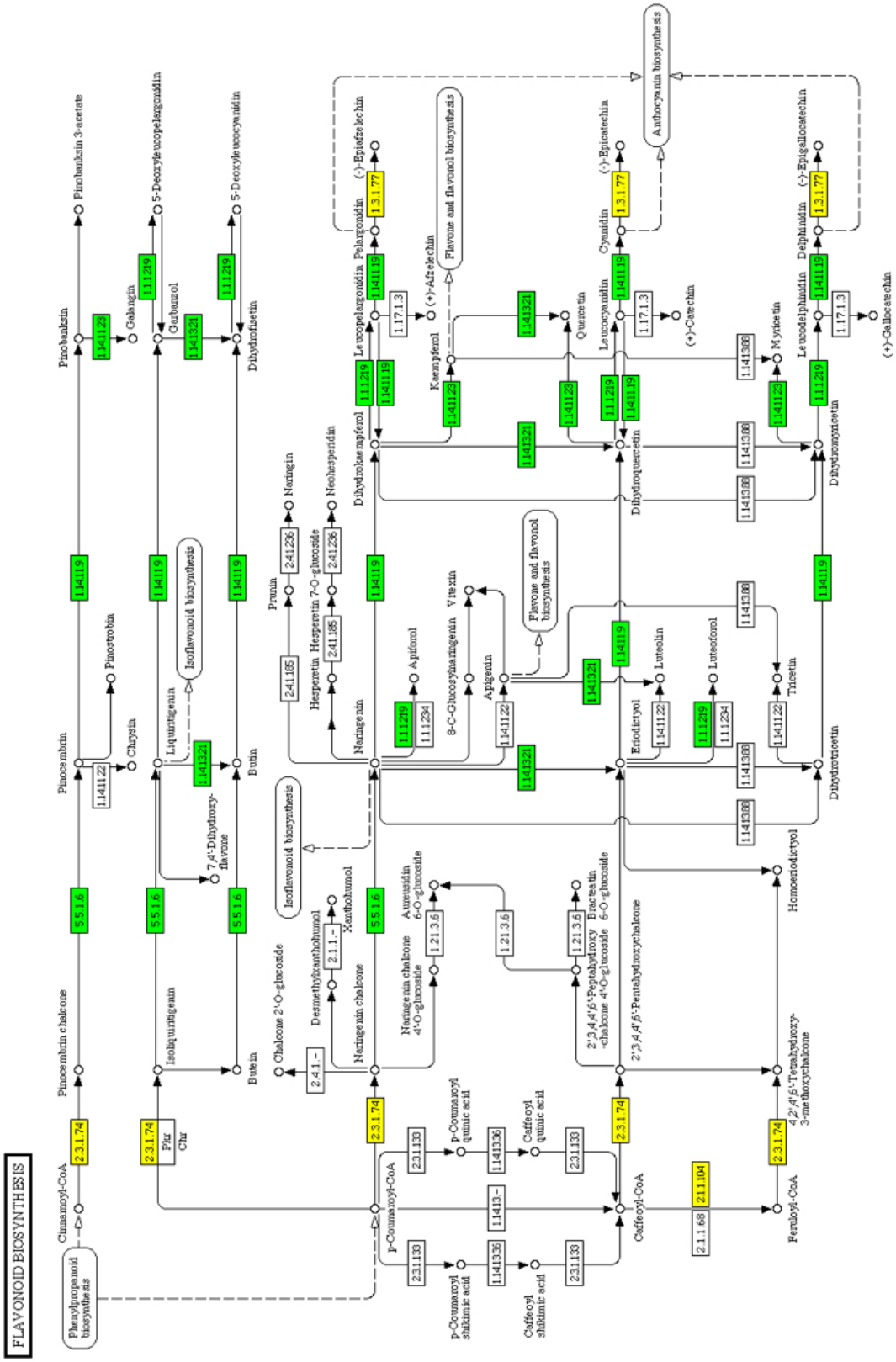


Figure 3.3: Flavonoid biosynthesis pathway from KEGG.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by JMT overexpression. Yellow indicates the genes that did not change significantly.

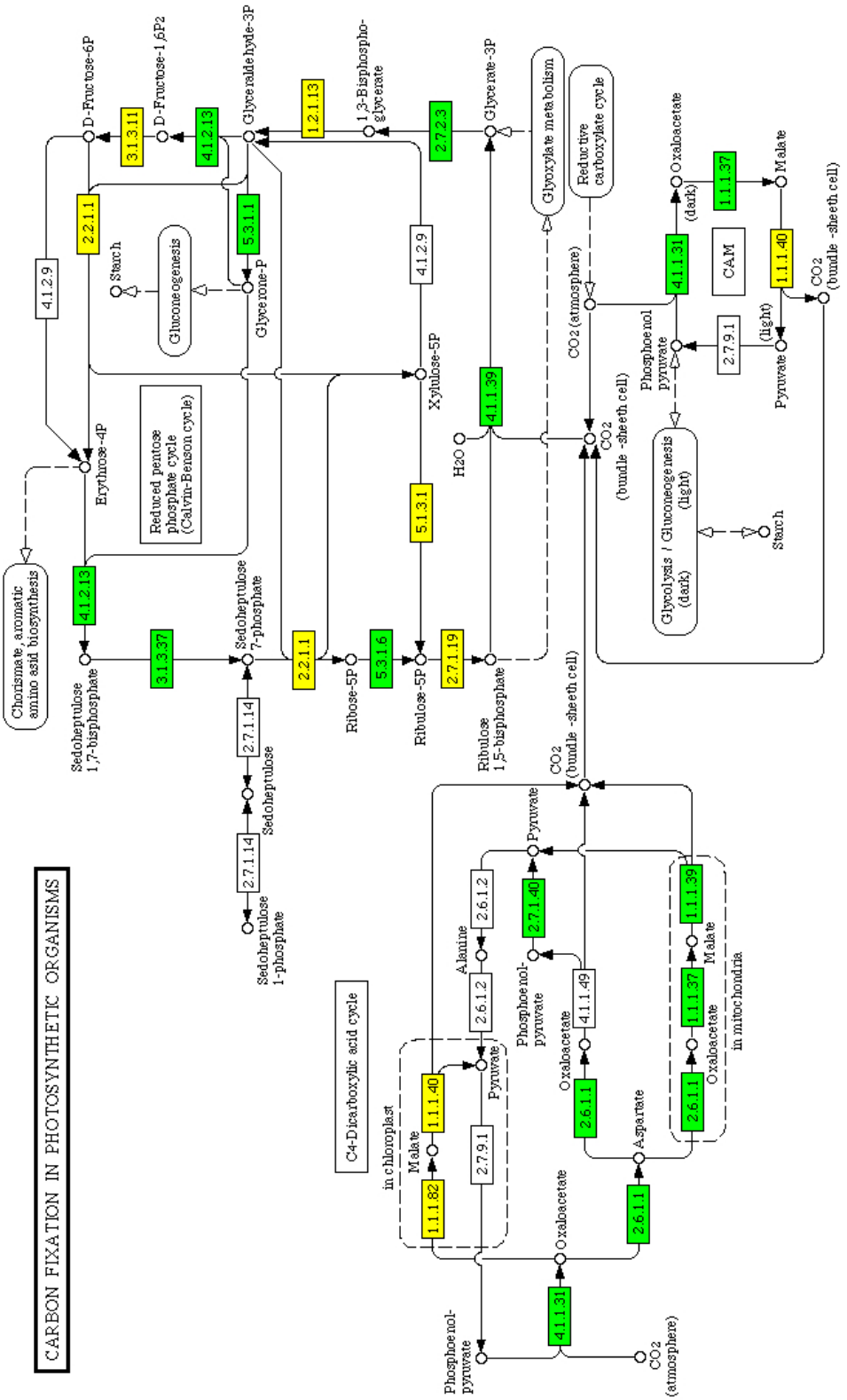


Figure 3.4: Carbon fixation pathway from KEGG.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by JMT overexpression. Yellow indicates the genes that did not change significantly.

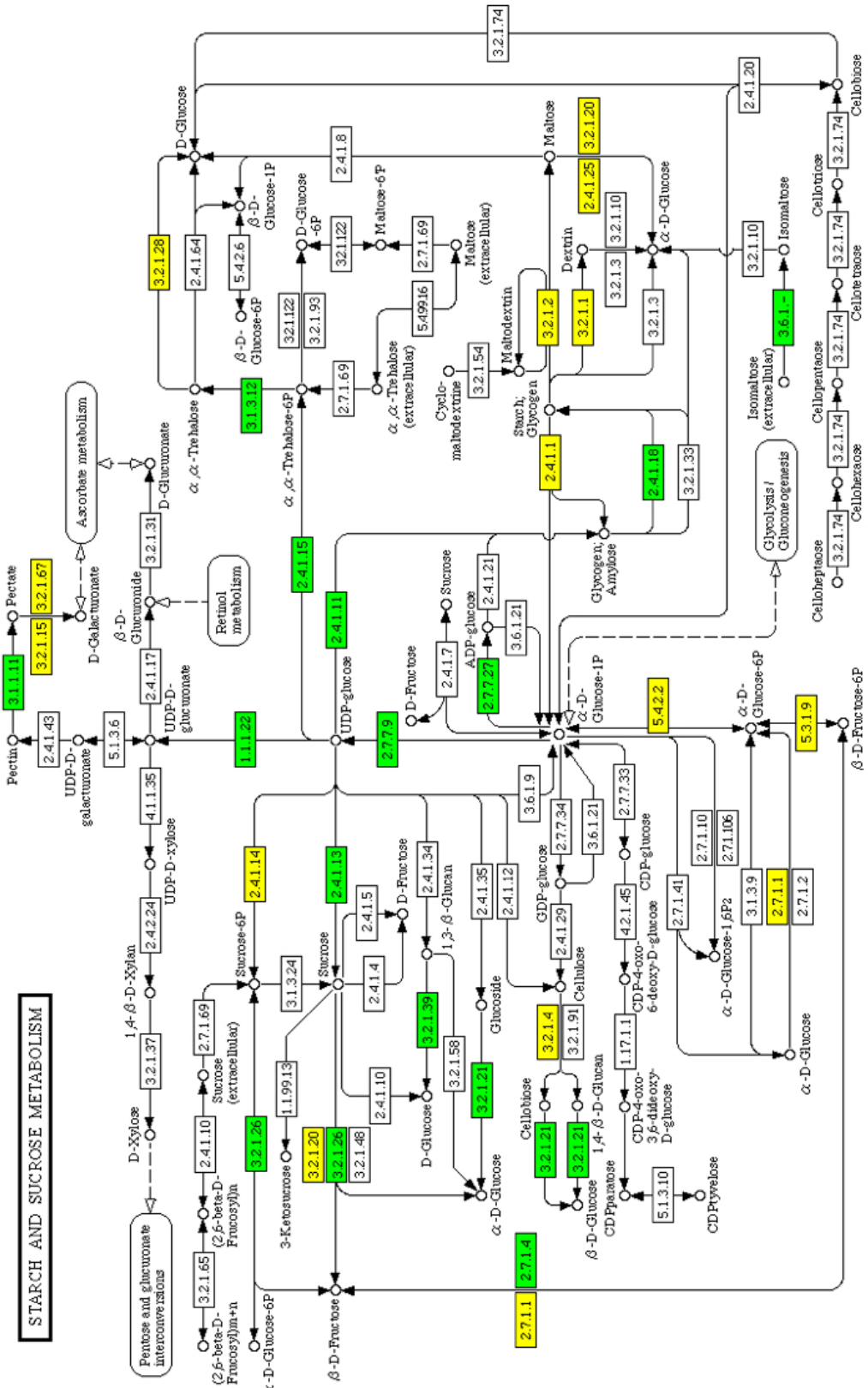


Figure 3.5: Starch and sucrose metabolism pathway from KEGG.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by JMT overexpression. Yellow indicates the genes that did not change significantly.

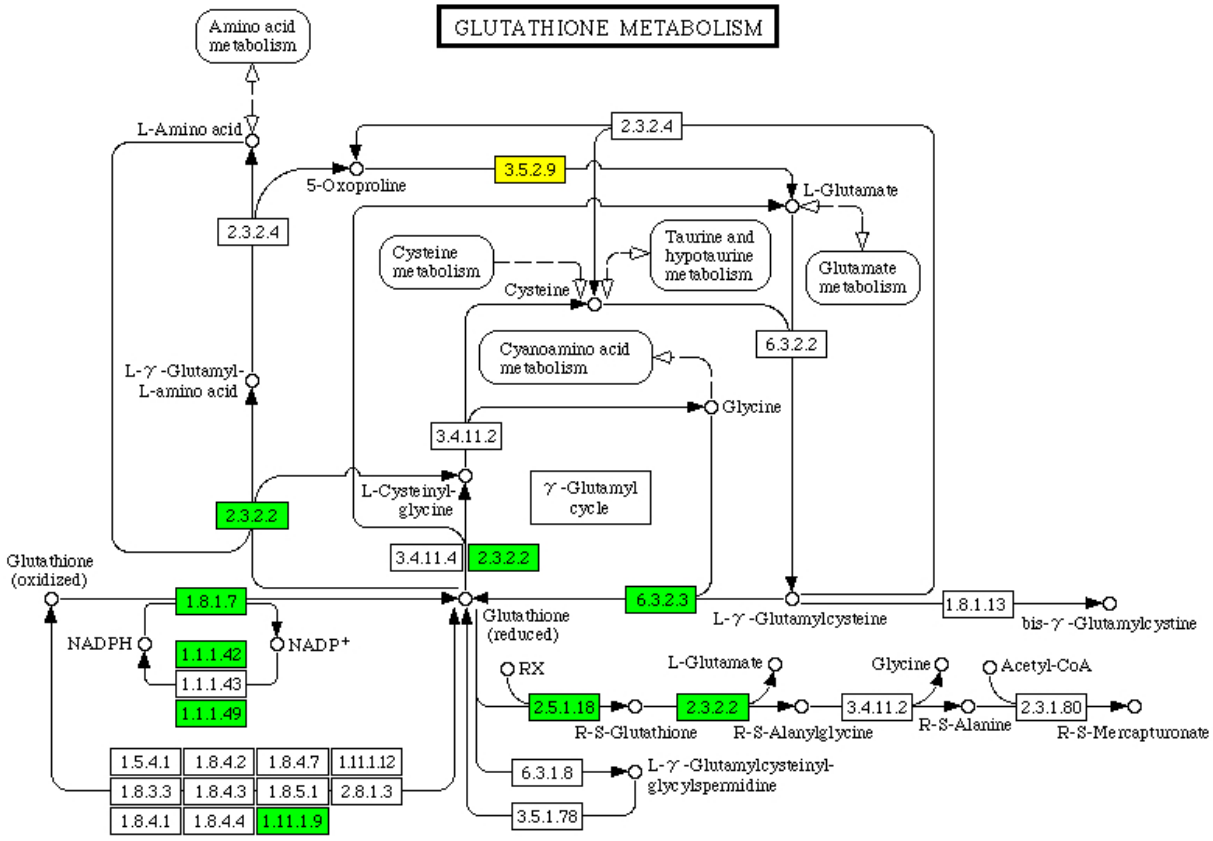


Figure 3.6: Glutathione metabolism pathway from KEGG.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by JMT overexpression. Yellow indicates the genes that did not change significantly.

3.4.2 Gene Expression Profiling of IAMT Overexpressing Mutant

Several overexpressing and mutant transgenic lines are available to study the function of indole-3-acetic acid methyltransferase (IAMT) in Arabidopsis [49]. To examine the effect of regulating the activity of IAMT on the genome-wide transcription profile of *A. thaliana,* we examined the IAMT overexpressing line constructed by Qin *et al.* [49], using DNA microarray. In addition to the overexpression line, a T-DNA insertion mutant of IAMT (SALK-072125), was identified from the Salk T-DNA collection [104], but no obvious phenotype was observed for this T-DNA line. The T-DNA insertion occurs at the end of the third of the four exons of this gene, causing an insertion of nine nucleotides at the site of integration and interrupting the C-terminal end of the protein. Despite the T-DNA insertion, a truncated transcript of IAMT was detected, which would encode a truncated protein with a deletion of 92 amino acids at the C terminus [49]. It is possible that this altered protein still has IAMT activity. Instead of a knockout this mutant should be consider a knock-down. Alternatively, it is possible that other methyltransferase genes may be able to compensate for the loss of function of IAMT, because there are additional, still uncharacterized carboxyl methyltransferase genes in the Arabidopsis genome. Therefore, for the purpose of understanding the effects of modulating the expression of IAMT in Arabidopsis we did not include the IAMT mutant, and limited the analysis to the 35S-IAMT overexpressing line.

Qin *et* al, constructed transgenic plants that express β-glucoronidase (GUS) reporter gene under the control of IAMT promoter, and found an ubiquitous GUS expression in several tissues, including seedlings, rosette leaves and flowers, with a severe curly leaf phenotype [49]. For our study, RNA was isolated from flowers of 35 days old plants, where approximately 30% of the flowers to be produced have opened, and siliques have not matured.

Table 3.7 list the KEGG pathways that were enriched with genes that change significantly when IAMT is overexpressed. Additional information about GO categories enriched with differentially expressed genes is listed in Tables A.4 and A.3.

Table 3.7: KEGG Pathways significantly enriched in Arabidopsis plants overexpressing IAMT.
Green and red p-values indicate whether the pathway is down-regulated or up-regulated, respectively.

KEGG Category	***p*-value**
ath00196:Photosynthesis - antenna proteins	4.0E-08
ath00195:Photosynthesis	3.1E-07
ath00907:Carotenoid biosynthesis - Organism group-specific	0.04684
ath00906:Carotenoid biosynthesis - General	0.04684
ath03010:Ribosome	8.75E-11
ath03050:Proteasome	8.02E-06
ath00592:alpha-Linolenic acid metabolism	0.00094

Auxin is involved in many aspects of plant growth and development. Many molecular components and processes regulated by auxins have been identified and their biological function is at least partly understood [122]. As shown in Table 3.7, several key processes in the plant seem to be regulated by auxins. Recently, it was found that methyl indole-3-acetic acid (MeIAA) is an inactive form of IAA [72] and that several esterases belonging to the *AtMES* Esterase are capable of hydrolyze and activate MeIAA *in vitro* [40]. We speculate that overexpression of IAMT might result in high levels of free IAA in two ways. First, by increasing the demand of IAA as substrate, IAMT stimulates *de novo* synthesis of free IAA. Second, IAMT methylates IAA at higher rates producing more MeIAA, which in turn is converted back to IAA by MeIAA esterases. Although, we did not find any of the known MeIAA esterases up-regulated, they may still have effect at their minimum expression levels. Without further prove, the regulation of the pathways shown in table 3.7 is for now attributed solely to the changes in the transcription levels of IAMT.

Auxin Responsive Genes

Auxin enhances the transcription of several classes of hormone responseive genes, such as the Aux/IAA , Gretchen Hagen3 (GH3), and SMALL AUXIN UP RNA (SAUR) gene family members [123]. Among the GO categories significantly represented in the set of differentially expressed genes, we found the response to auxin stimulus category (GO:0009733). SAUR-AC1 (At4g38850), GH3.5/WES1 indole-3-acetic acid amido synthetase (At4g27260), DFL1/GH3.6 indole-3-acetic acid synthetase (At5g54510), and 10 Aux/IAA inducible transcription factors were up-regulated (Table 3.8). The Aux/IAA genes, which are rapidly up-regulated in response to auxins such as IAA, encode the unstable IAA proteins believe to be transcriptional repressors [124]. The Arabidopsis genome has 29 genes that encode putative Aux/IAA proteins [13]. The Aux/IAA proteins regulate auxin-mediated gene expression by interacting with members of the AUXIN RESPONSE FACTOR protein family. Auxin induces the expression of many, but not all, Aux/IAA gene family members. Recent biochemical and genetic studies indicate that the Aux/IAA proteins are the postulated repressors of auxin-regulated transcriptional activation. The molecular genetic and biochemical data indicate that rapid turnover of Aux/IAA proteins is required for normal auxin response [123].

The high number of up-regulated auxin responsive genes, including these groups of transcription factors that regulate auxin-mediated gene expression, support the hypothesis stated above, indicating that IAMT overexpression might result in high IAA levels in the plant.

Ribosome, Protein Synthesis and IAA

Cell division, growth, maturation and differentiation are all processes associated with protein synthesis and auxin regulation [122]. Ribosomes function as an assembly line for proteins. We observed that genes encoding ribosomal proteins were significantly down-regulated by

IAMT overexpression (Figure 3.7). A tremendous amount of evidence had accumulated until the late 1960s to support the claim that RNA and protein synthesis were required for auxin to induce growth in a wide variety of tissues [125]. The effect of IAA on RNA and protein synthesis was establish very early in etiolated sub-apical sections of *Pisum sativum* [125]. There is substantial evidence that continued RNA and protein syntheses are required for auxin to induce cell elongation in excised and intact non-growing tissues [126, 127]. Work on soybean hypocotyl tissue showed that auxin treatment alters the pattern of protein synthesis in both elongating and basal sections by promoting the synthesis of specific polypeptides, or by repressing the synthesis of certain polypeptides. Changes in the patterns of protein synthesis induced by auxin are also shown to be tissue specific, because different tissues synthesize largely different polypeptides in response to auxin [126]. Additionally, applications of auxin to the tips of intact aged pea *Pisum sativum* epicotyls resulted in an increase in the content of polyribosomes and poly(A) and in the capacity of isolated polysomes to support protein synthesis in vitro [127].

While the presence of a significant large number of auxin responsive genes, described in the previous section, might be a result of high IAA levels, the down regulation of the ribosome and protein synthesis suggest that the IAA levels in IAMT overexpressing plants is low, consequently reducing the plant growth rate, consistent with the observed limited leaf development phenotype. These two observations might seem contradictories. A possible explanation is that the overexpression of IAMT may result in a high turnover of IAA, allowing for the occurance of both observations. IAMT overexpression results in high rate synthesis of IAA, which activate early auxin-reposive genes, in turn IAA is inactivated by methylation, and ribosomes and protein synthesis which are activated late in the auxin-response cascade remain inactive. Further studies would need to be performed to test our hypothesis.

Ubiquitin-Protesome pathways and Auxins Response

In eukaryotic organisms, ubiquitin conjugation to target proteins for subsequent degradation by the multisubunit proteosome plays an important role in diverse cellular processes ranging from cell-cycle regulation to signal transduction [122]. Several auxin resistant-mutans have been isolated and analysed [122]. Sequence analysis of the auxin-resistant mutant (*axr1*) indicates that the AXR1 gene encodes a protein that participates in activation of the ubiquitin-related protein RUB1, which prepares proteins for ubiquitin-mediated degradation [128]. Additionally, the auxin transport inhibitor resistant mutant *(tir1)* is defective in synthesis of a protein having a conserved amino acid domain similar to that found in ubiquitin protein ligase [129]. The characterization of the auxin resistant mutants *axr1* and *tir1* in Arabidopsis, led to the discovery that the ubiquitin–proteasome pathway is involved in auxin response [130]. Based on this information it has been proposed that IAA role in cell expansion and growth is mediated by the ubiquitin-proteosome pathway [130].

Substantial evidence strongly suggest that each state of development is associated with a

unique set of proteins, some of which must be degraded before the subsequent stage can be initiated [130, 131]. The critical proteins of one stage are proposed to repress by feedback the transcription of genes necessary for the next stage [130,131]. Ubiquitin-mediated protein degradation regulates the growth and development of eukaryotic organisms. Proteins that are destined to be destroyed are tagged with a poly-ubiquitin chain by a cascade of reactions involving three enzymes, known as the ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2) and ubiquitin protein ligase (E3). Ubiquitinated proteins are recognized and degraded by the 26S proteasome, a multiprotein complex comprising a 20S core unit and two 19S regulatory particles [131].

Strong evidence show that AUX/IAA proteins generally have a high turnover rate. The half-life of these proteins in wild-type Arabidopsis seedlings ranges from 10 min to 80 min, depending on the protein. However, this short half-life can be extended several fold by treatment with proteasome inhibitors, indicating that the degradation of the AUX/IAA proteins is associated with the proteasome pathway [132, 133]. Our observation of genes encoding proteosome assembly proteins being down regulated, might further indicate that the overall levels of IAA are changing due to the modulation of IAMT (Fig 3.8).

Is Jasmonic Acid Biosythesis regulated by IAA ?

As mentioned above, jasmonates are synthesized through the α-linolenic acid biosynthetic pathway. We found that α-linolenic acid metabolism was significantly down-regulated as a result of IAMT overexpression (Figure 3.9). In a review article Saniewski *et al.* [134] describe the interaction between jasmonates and auxin in regulation of some physiological processes in plant growth and development. Saniewski *et al.* [134] showed that some jasmonate-induced processes are reduced by auxins and some auxin stimulated physiological processes are inhibited by jasmonates. Substantial evidence presented by experimental data indicates that auxins prevent or reduce some physiological phenomena induced by jasmonates [134]. Since both jasmonates and auxin are present at high concentrations in the same young structures of tissues, it is probable that auxin reduces some processes induced by jasmonates and vice versa. At the gene expression level auxin was shown to alter the level of specific mRNAs, inhibiting accumulation of genes encoding proteins in the jasmonate pathway, like lipoxygenase (LOX A) [135]. On the contrary, like we observed when JMT is overexpressed, jasmonates have been reported to increase an accumulation of LOX A mRNA, and other jasmonate biosynthetic enzymes, which are inducible by wounding [135, 136]. In wounded tissue, a decrease in endogenous levels of auxin combined with wound-induced increase in jasmonates as internal factors may contribute to wound-induced expression of genes encoding jasmonate biosynthetic proteins [137].

Photosynthesis and IAA

All the processes in which auxin regulation is required demand energy, therefore the indirect regulation of photosynthesis by auxins is not unexpected. Application of IAA to leaves has been shown to increase the rate of photosynthetic CO_2, uptake in a number of different plant species [138, 139]. In the early 1930's it was establish in *Chlorella vulgaris* that auxin does not affect respiration but it markedly affects the rate of photosynthesis. The effect is indirect through the changes in the chlorophyll content and the cell size. In the young cultures the rate of photosynthesis is proportional to the chlorophyll concentration, while in the old cultures it is proportional to both the chlorophyll concentration and the extent of the cell surface [140]. This increase in photosynthetic rate promoted by IAA could result from a number of different interactions of the hormone within the plant cell. Thus photosynthesis could be stimulated as a result of increased CO_2 diffusion, increased utilization of photosynthetic products or by a direct effect of IAA on chloroplasts [141].

Several studies have shown the effect of auxins on photorespiration in different plant systems and different tissues within the same plant. Treatment of developing primary leaves of of *Sinapis alba* with IAA at low concentrations caused a decrease of dry weight, soluble reducing sugars, soluble protein, chlorophylls, carotenoids and cytochromes. It also caused a lower ratio of protein to chlorophyll, a lower ratio of chlorophyll A to chlorophyll B and a higher ratio of chlorophyll per cytochrome F [141].

Up-regulation of different components of the photorespiration machinery observed as consequence of overexpression of IAMT (Figure 3.10), might be indirect result of transient high levels of IAA before it is methylated or perhaps the diffusion of MeIAA through the chloroplast membrane and its conversion back to IAA to exert local effect. Despite the fact that the effect of IAA in photorespiration has been study extensively, there are no reports correlating IAA and photosynthesis in Arabidopsis.

Carotenoids and Abscisic Acid Biosynthesis

Interestingly, in addition to the photosynthetic pathway, we found that the carotenoids biosynthesis is significantly up-regulated as well (Figure 3.11). We can consider two possible reasons to explain the effect of IAA on carotenoids biosynthesis. First, the quantitative and qualitative composition of chlorophylls and carotenoids pigments is directly connected with photosynthesis. Second, the synthesis of β-carotene is an indirect abscisic acid (ABA) biosynthesis pathway, consequently carotenoids biosynthesis up-regulation might be another cross-talk mechanisms between IAA and ABA.

Due to the role of auxins on tissue growth and development it is frequently used as hormone supplement in tissue culture experiments, consequently most of the reports about its effects on plant pigments concentration, comes from tissue culture studies. It has been documented that IAA strongly stimulated the increase of chlorophylls A and B and carotenoids (espe-

cially β-carotene and lutein + zeaxanthin) [142, 143]. In *Vigna sinensis* Savi, water stress decreased all photosynthetic pigments (chlorophyll A, chlorophyll B and carotenoids), but IAA solutions supplied to the plant at different concentration increased photosynthetic pigments [142]. In banana (*Musa* spp. AAA) leaves, under micro-propagation, the levels of IAA in normal leaves were significantly higher than those found in green and yellow sectors of variegated leaves [143].

The carotenoid biosynthetic pathway is well known in plants and several reactions are affected by mutation of carotenoid biosinthetic genes in Arabidopsis and maize. However, the mechanism of action of IAA on plant pigments is not well understood. It is possible that IAA regulates photosynthesis to satisfy the energy demand for growth and development, indirectly increasing the demand for photosynthetic pigments including chlorophylls and carotenoids. We mentioned earlier the regulation of ABA biosynthesis pathway by IAA as a cross-talk as an interaction of the signaling pathways of these two hormones. Similarly to IAA, ABA is known to regulate myriad aspects of plant growth and development. Many physiological studies have suggested that these two hormones functionally interact in roots as regulators of growth, development and tropism [144]. Additionally, it has been proposed that in Arabidopsis and probably all higher plants, ABA and IAA signalling is mediated by both redundant and independent mechanisms, some of which also appear to affect responses to other signal molecules, including each others' [144]. Connection between ABA and auxin responses are suggested by the properties of the *axr2-1* and *hy11* mutations. For example, the *axr2-1* mutant is resistant to ABA and increase the stability of the IAA response protein IAA7, believe to be a transcriptional repressor [124]. However, direct effect of auxin responsive proteins over the carotenoids metabolism on the carotenoids biosynthesis has not been considered. We speculate that one or several of the auxin responsive proteins up-regulated when IAMT is overexpressed, and thought to be transcription factors, might be involved in regulating transcription of genes responsible for the synthesis of ABA. The coordinated regulation of ABA/IAA concentrations is what control the processes of cell growth and development, rather than each one of these signaling molecules independently.

Table 3.8: Auxin Responsive Genes (GO:0009733) significantly up-regulated in Arabidopsis plants overexpressing IAMT.

Locus Identifier	**Annotation**	**Locus Identifier**	**Annotation**
AT3G60690	auxin-responsive family protein	AT1G04240	SHY2 (SHORT HYPOCOTYL 2); transcription factor
AT1G74840	myb family transcription factor	AT2G22670	IAA8 (indole-acetic acid-induced protein 8); transcription factor
AT1G29490	auxin-responsive family protein	AT4G38860	auxin-responsive protein, putative
AT4G14550	IAA14 (SOLITARY ROOT); transcription factor	AT4G36110	auxin-responsive protein, putative
AT1G49010	myb family transcription factor	AT1G29500	auxin-responsive protein, putative
AT3G23050	IAA7 (AUXIN RESISTANT 2); transcription factor	AT2G21210	auxin-responsive protein, putative
AT4G02890	UBQ14 (ubiquitin 14); protein binding	AT4G27260	GH3.5/WES1; indole-3-acetic acid amido synthetase
AT2G46690	auxin-responsive family protein	AT4G10100	CNX7/SIR5; catalytic
AT4G38840	auxin-responsive protein, putative	AT5G42190	ASK2 (ARABIDOPSIS SKP1-LIKE 2); ubiquitin-protein ligase
AT3G03840	auxin-responsive protein, putative	AT1G17520	DNA-binding protein, putative
AT4G38620	MYB4 (myb domain protein 4); transcription factor	AT2G36910	ATPGP1 (ARABIDOPSIS THALIANA P GLYCOPROTEIN1); calmodulin binding
AT1G19830	auxin-responsive protein, putative	AT2G33310	IAA13 (indole-acetic acid-induced protein 13); transcription factor
AT4G01280	myb family transcription factor	AT4G02570	ATCUL1 (CULLIN 1)
AT1G75580	auxin-responsive protein, putative	AT4G36800	RCE1 (RUB1 CONJUGATING ENZYME 1); small protein conjugating enzyme
AT1G70000	DNA-binding family protein	AT3G61830	ARF18 (AUXIN RESPONSE FACTOR 18); transcription factor
AT1G04550	IAA12 (AUXIN-INDUCED PROTEIN 12); transcription factor	AT2G21220	auxin-responsive protein, putative
AT5G15310	AtMIXTA/AtMYB16 (myb domain protein 16); DNA binding / transcription factor	AT5G19140	auxin/aluminum-responsive protein, putative
AT3G09600	myb family transcription factor	AT1G22640	MYB3 (myb domain protein 3); DNA binding / transcription factor
AT5G56290	PEX5 (PEROXIN 5); peroxisome matrix targeting signal-1 binding	AT4G34760	auxin-responsive family protein
Continued ...			

Locus Identifier	Annotation	Locus Identifier	Annotation
AT1G70940	PIN3 (PIN-FORMED 3); auxin:hydrogen symporter/ transporter	AT5G54510	DFL1/GH3.6 (DWARF IN LIGHT 1); indole-3-acetic acid amido synthetase
AT5G47370	HAT2; transcription factor	AT4G14560	IAA1 (INDOLE-3-ACETIC ACID INDUCIBLE); transcription factor
AT4G05050	UBQ11 (UBIQUITIN 11); protein binding	AT4G36740	ATHB40/HB-5 (ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 40); DNA binding / transcription factor
AT3G23030	IAA2 (indole-acetic acid-induced protein 2); transcription factor	AT1G29430 AT5G27780	[AT1G29430, auxin-responsive family protein];[AT5G27780, auxin-responsive family protein]
AT4G38850	SAUR_AC1 (SMALL AUXIN UP RNA 1 FROM ARABIDOPSIS THALIANA ECOTYPE COLUMBIA)	AT2G28350	ARF10 (AUXIN RESPONSE FACTOR 10); miRNA binding / transcription factor
AT1G29440	auxin-responsive family protein	AT5G03240	UBQ3 (POLYUBIQUITIN 3); protein binding
AT1G29460	auxin-responsive protein, putative	AT1G74660	MIF1 (MINI ZINC FINGER 1); DNA binding / transcription factor
AT5G20620	UBQ4 (ubiquitin 4); protein binding	AT2G46830	CCA1 (CIRCADIAN CLOCK ASSOCIATED 1); transcription factor
AT2G04550	IBR5 (INDOLE-3-BUTYRIC ACID RESPONSE 5); protein tyrosine/serine/threonine phosphatase	AT1G75590	auxin-responsive family protein
AT3G62980	TIR1 (TRANSPORT INHIBITOR RESPONSE 1); ubiquitin-protein ligase	AT4G28640	IAA11 (indole-acetic acid-induced protein 11); transcription factor
AT5G43700	ATAUX2-11 (indole-acetic acid-induced protein 4); transcription factor	AT4G09460	ATMYB6 (MYB DOMAIN PROTEIN 6); DNA binding / transcription factor
AT3G17600	IAA31 (indole-acetic acid-induced protein 31); transcription factor	AT1G77690	amino acid permease, putative
AT4G18710	BIN2 (BRASSINOSTEROID-INSENSITIVE 2); kinase	AT3G04730	IAA16 (indole-acetic acid-induced protein 16); transcription factor
AT2G01420	PIN4 (PIN-FORMED 4); auxin:hydrogen symporter/ transporter	AT2G42620	MAX2 (MORE AXILLARY BRANCHES 2); ubiquitin-protein ligase
AT5G66700	HB53 (homeobox-8); DNA binding / transcription factor	AT2G21200	auxin-responsive protein, putative
AT5G18060	auxin-responsive protein, putative		

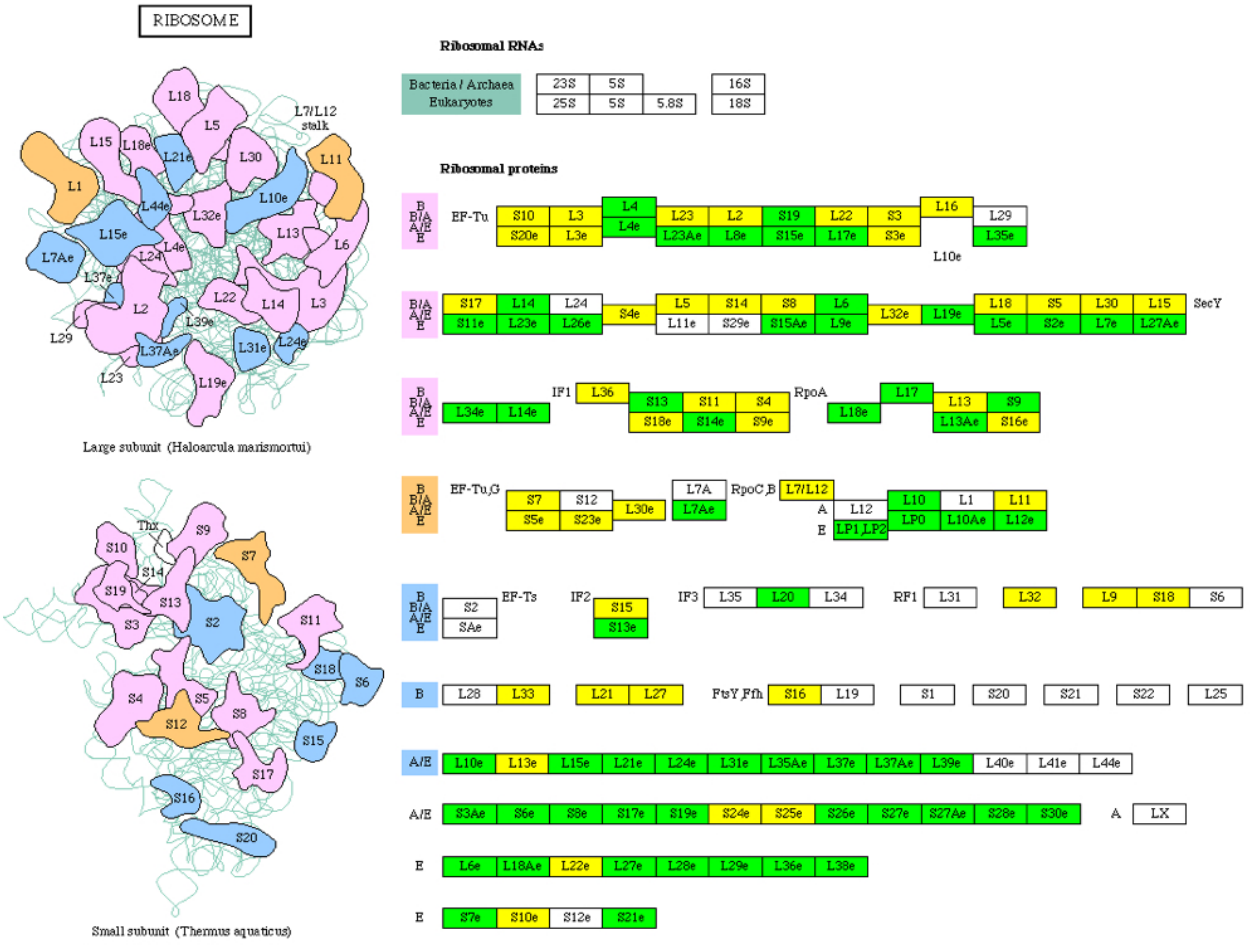
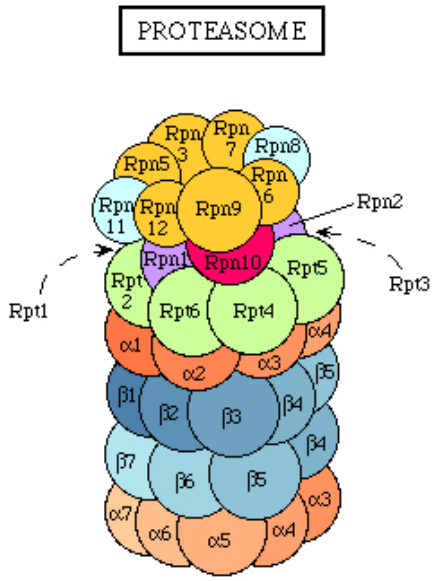


Figure 3.7: Ribosome proteins from KEGG down-regulated by IAMT overexpression. Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by IAMT overexpression. Yellow indicates the genes that did not change significantly.



Rpn1	Rpn2	Rpn3	Rpn4	Rpn5	Rpn6	
Rpn7	Rpn8	Rpn9	Rpn10	Rpn11	Rpn12	
Rpt1	Rpt2	Rpt3	Rpt4	Rpt5	Rpt6	
α1	α2	α3	α4	α5	α6	α7
β1	β2	β3	β4	β5	β6	β7

Figure 3.8: Proteosome proteins from KEGG down-regulated by IAMT overexpression. Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by IAMT overexpression. Yellow indicates the genes that did not change significantly.

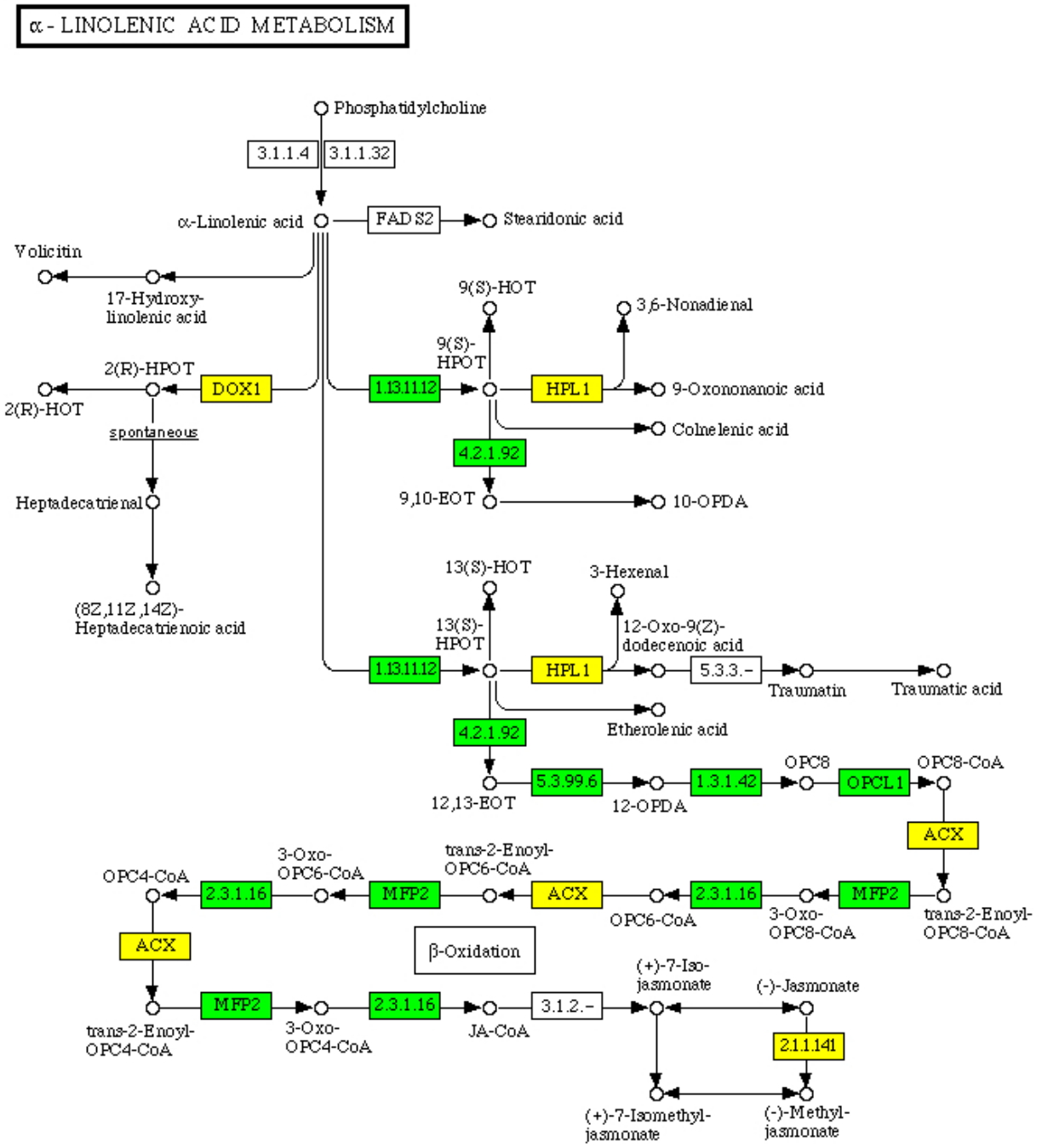


Figure 3.9: α-Linoleic Acid Metabolism from KEGG is down-regulated by IAMT overexpression.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by IAMT overexpression. Yellow indicates the genes that did not change significantly.

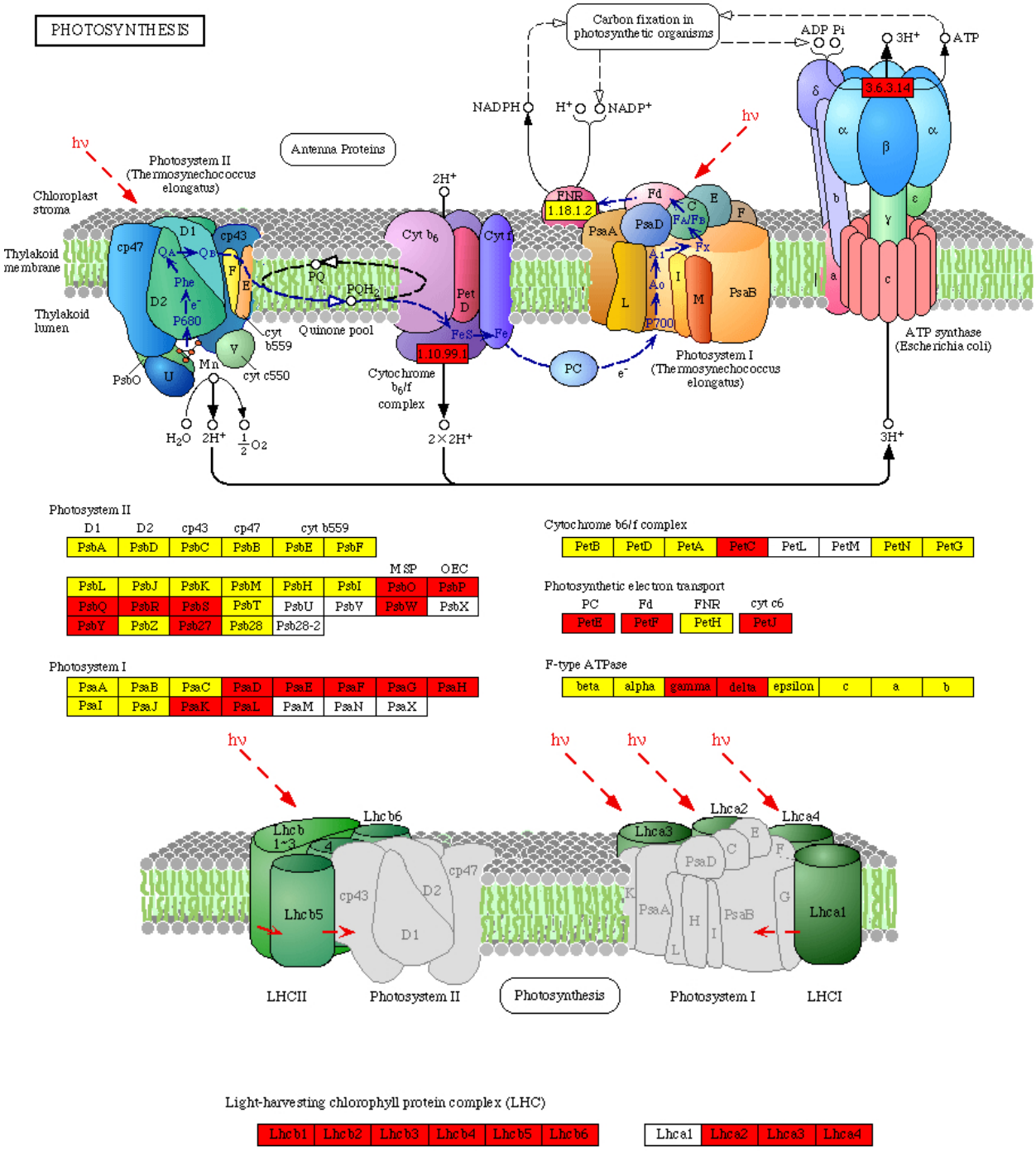


Figure 3.10: Photosynthesis from KEGG up-regulated by IAMT overexpression. Colored boxes indicate the genes identified in the Arabidopsis genome. Red indicates the genes that were significantly induced by IAMT overexpression. Yellow indicates the genes that did not change significantly.

3.4.3 Gene Expression Profiling of BSMT Knockout

BSMT is responsible for the methylation of salicylic and benzoic acids [14]. To study the role of BSMT in Arabidopsis plants we used two methods of gene expression modulation: overexpression and T-DNA insertion mutants. The transgenic overexpressor for BSMT was obtained at Eran Pichersky laboratory at the University of Michigan, by genetic transformation of *A. thaliana* with a genetic construct based on 35S promoter and BSMT cDNA. The T-DNA insertion mutant of BSMT (GARLIC-776-B10), was identified from the Syngenta Arabidopsis Insertion Library or SAIL (formerly GARLIC) Collection [105]. Under normal conditions, the T-DNA insertion mutant *bsmt*, showed not obvious morphological or physiological phenotype, and no differences were observed relative to the wilt-type plants. To examine the gene expression profile of BSMT overexpressing, knockout and wild-type lines, RNA was extracted from flowers collected from 30-day old plants before siliques full development and maturation. Surprisingly, only two genes were found significantly different between overexpressor and knockout lines. Additionally, BSMT overexpressor or knockout lines compared to the wild-type, are enriched on genes that up-regulated or down-regulate the same pathways. Analysis of the expression of the BSMT gene only, indicated that it is significantly down-regulated in the T-DNA knockout, but is not significantly different in the overexpressor compared to the wild-type. These observations clearly suggest a co-supression effect in BSMT overexpressing as a consequence of the high transcription levels.

Co-supression by overexpression is a phenomenon reported before in plants [145–147]. Overexpression, where a gene is expressed from a high-level constitutive or tissue-specific promoter in transgenic plants, can produce either plants that accumulate high levels of transcription factor or knock-out plants through inactivation of the transgene and/or the endogenous gene by cosuppression [145–147]. Overexpression and inactivation of the transgene do not necessarily occur in all plant cells and the extent of overexpression or cosuppression is difficult to assess. Therefore for the purpose of analysing the effect of modulating the expression of BSMT in Arabidopsis we focus only on the T-DNA knockout.

In table 3.10 is listed the KEGG pathways that were enriched with genes that change significantly when BSMT is knocked out. Additional information about GO categories enriched with differentially expressed genes is listed in Tables A.6 and A.5.

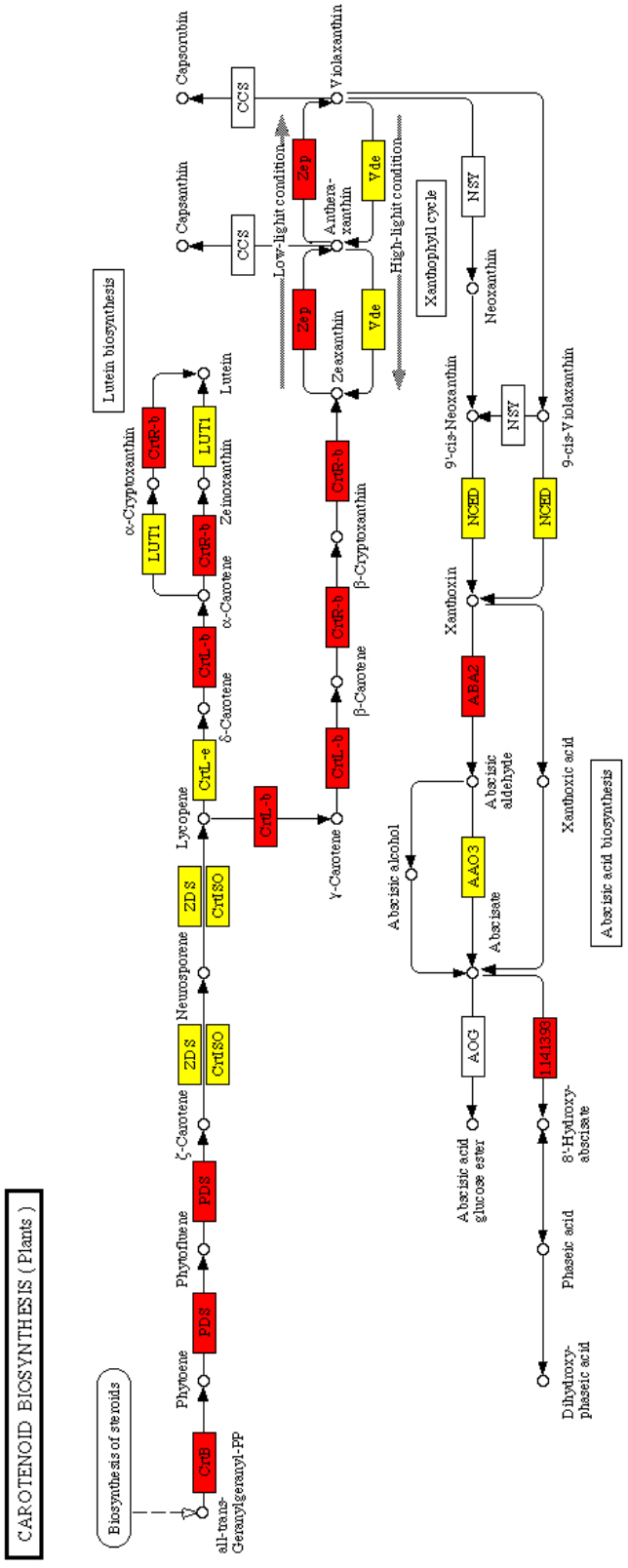


Figure 3.11: Carotenoids and Abscisic Acid Biosynthesis from KEGG is up-regulated by IAMT overexpression. Colored boxes indicate the genes identified in the Arabidopsis genome. Red indicates the genes that were significantly induced by IAMT overexpression. Yellow indicates the genes that did not change significantly.

Table 3.10: KEGG Pathways significantly enriched in BSMT T-DNA knockout Arabidopsis plants. Green and red p-values indicate whether the pathway is down-regulated or up-regulated.

KEGG Category	p-value
ath00592:alpha-Linolenic acid metabolism	4E-07
ath00196:Photosynthesis - antenna proteins	1E-04
ath00562:Inositol phosphate metabolism	0.0015
ath04070:Phosphatidylinositol signaling system	0.023
ath03010:Ribosome	2E-05
ath00010:Glycolysis / Gluconeogenesis	7E-04
ath00710:Carbon fixation	6E-04
ath00030:Pentose phosphate pathway	0.0039
ath00640:Propanoate metabolism	0.026
ath00051:Fructose and mannose metabolism	0.040

SA acid JA signaling pathways cross-talk

The interactions between SA and JA signaling mechanisms appear to be complex, and there is evidence for both positive and negative interactions between theses pathways. They appear to exert mutually antagonistic effects [47]. The inhibitory effect of SA on JA signaling is well demonstrated in tomato and Arabidopsis [148, 149]. Among the genes down-regulated by BSMT knockout we found a significant enrichment in the α-Linolenic metabolism, which is responsible for the synthesis of JA (Fig 3.12). This suggest SA accumulation in plant as direct consequence of BSMT inactivity, results in a negative feed-back that reduce the biosynthesis of JA.

SA is a negative regulator of JA-dependent defense responses in plants [149]. SA accumulation by pathogen infection triggers NPR1, a protein-protein interaction mediator, and WRKY70, a transcription factor. NPR1 represses the JA biosynthetic gene LOX2 and represses the JA biosynthesis and responses [47]. On the contrary, Koo *et al.* [47], found that JA induce BSMT, which may contribute to an antagonistic effect on SA signaling pathways by depleting SA in plants.

Looking for differentially expressed genes that mediate in the SA and JA cross-talk we found WRKY40 and WRKY54, both transcription factors. The up-regulated WRKY54 is an homolog of WRKY70, and prevents excessive accumulation of SA. Like WRKY70, WRKY54 play dual roles as negative regulators of SA biosynthesis and positive regulators of SA-Mediated Gene Expression and Resistance [150]. On the other hand, the down-regulated WRKY40 is a negative regulator of defense, it binds other WRKY transcription factors to block the gene expression of genes that can only be trigger by SA [150].

We believe that JA biosynthesis and WRKY40 down-regulation, together with WRKY54

high expression, suggest that BSMT in addition to producing the airborne signal MeSA, it helps reducing the accumulation of SA which can be toxic to the plant at high levels.

Phosphatidylinositol signaling system

The phosphatidylinositol (PI) metabolic pathway is considered critical in plant responses to many environmental factors. Previous studies have indicated the involvement of multiple PI-related gene families during cellular responses [151]. Phosphoinositides have been implicated in a vast range of cellular functions, including signal transduction, vesicle trafficking, endocytosis and cytoskeletal rearrangement, by which growth factors, hormones and neurotransmitters can exert their physiological effects.

Lin *et at* [151], studied in Arabidopsis the involvement of the phosphatidylinositol signaling pathway in multiple plant responses to hormone and abiotic treatments. They found that many phosphatidylinositol pathway-related genes were differentially expressed following the treatment with different plant hormones, including SA [151]. More recently, using Arabidopsis suspension cells as a model, Krinke *et al.* showed that SA addition induced a rapid and early (within few minutes) decrease in a pool of phosphatidylinositol [152]. This decrease paralleled with an increase in phosphatidylinositol 4-phosphate and phosphatidylinositol 4,5-biphosphate. Both phosphatidylinositol 4-phosphate and phosphatidylinositol 4,5-biphosphate are mostly found in the endoplasmic reticulum, which led to speculations that SA mediates vesicle trafficking during stress response. However, the physiological role of regulating the phosphatidylinositol signaling pathway following hormone treatment remains to be elucidated.

Light harvesting apparatus and SA Levels

The photosynthesis antenna proteins were significantly down-regulated when BSMT expression was knockout (Figure 3.13). Several Arabidopsis mutants (*cpr1-1, cpr5-1, cpr6-1* and *dnd1-1),* which accumulate high endogenous levels of SA have been assessed on their adjustment of photosynthetic processes to short-term acclimation to high-light [153]. The operating efficiency of the photosystem II and chlorophyll A fluorescence analysis indicated that both low and high levels of endogenous SA, affect significantly the photosynthetic electron transport required for light harvesting [153]. Mateo *et al.,* concluded that controlled levels of SA are required for optimal functioning of the photosynthetic apparatus. This observations further support our hypothesis that one of BSMT roles is to maintain the homeostasis of SA endogenous levels by its conversion into MeSA.

Carbohydrate Metabolism and SA

Several pathways related with the carbohydrate metabolism were significantly up-regulated when BSMT expression was knockout (Table 3.10). Glycolysis/gluconeogenesis, carbon fixation, pentose phosphate pathway, propanoate metabolism, and fructose and mannose metabolism are all biochemical processes responsible for the formation, breakdown and interconversion of carbohydrates to provide short term energy to the cell or reserves for the long term.

The effects of SA applied exogenously and the accumulation of endogenous SA have been studied extensively in different plant systems [20, 21, 154]. In addition to the activation of gene expression and synthesis of defense proteins, it has been reported that accumulation of SA induce depolarization of the plasma membrane in plant cells [155, 156]. Raskin *et al. [154],* proposes that SA triggers proton signaling via cytoplasm acidification, however a complete dissipation of the proton gradient in the cell causes cell death. To compensate, energy consumption is required to restore both proton and potassium gradients (due to the activation of ATPases, which ensures H^+ release and K^+ accumulation). Studies about the effect of SA accumulation in wheat roots, indicates that the activation of the glycolysis and Krebs cycle due to high SA levels represents a stage of total enhancement of energy exchange aimed at the prevention of excessive acidification of the cytoplasm [156]. This suggest that in addition to satisfying the energy demand that comes with the activation of defense mechanisms by SA signaling, energy consumption reduces the cytoplasm acidification resulting from SA accumulation.

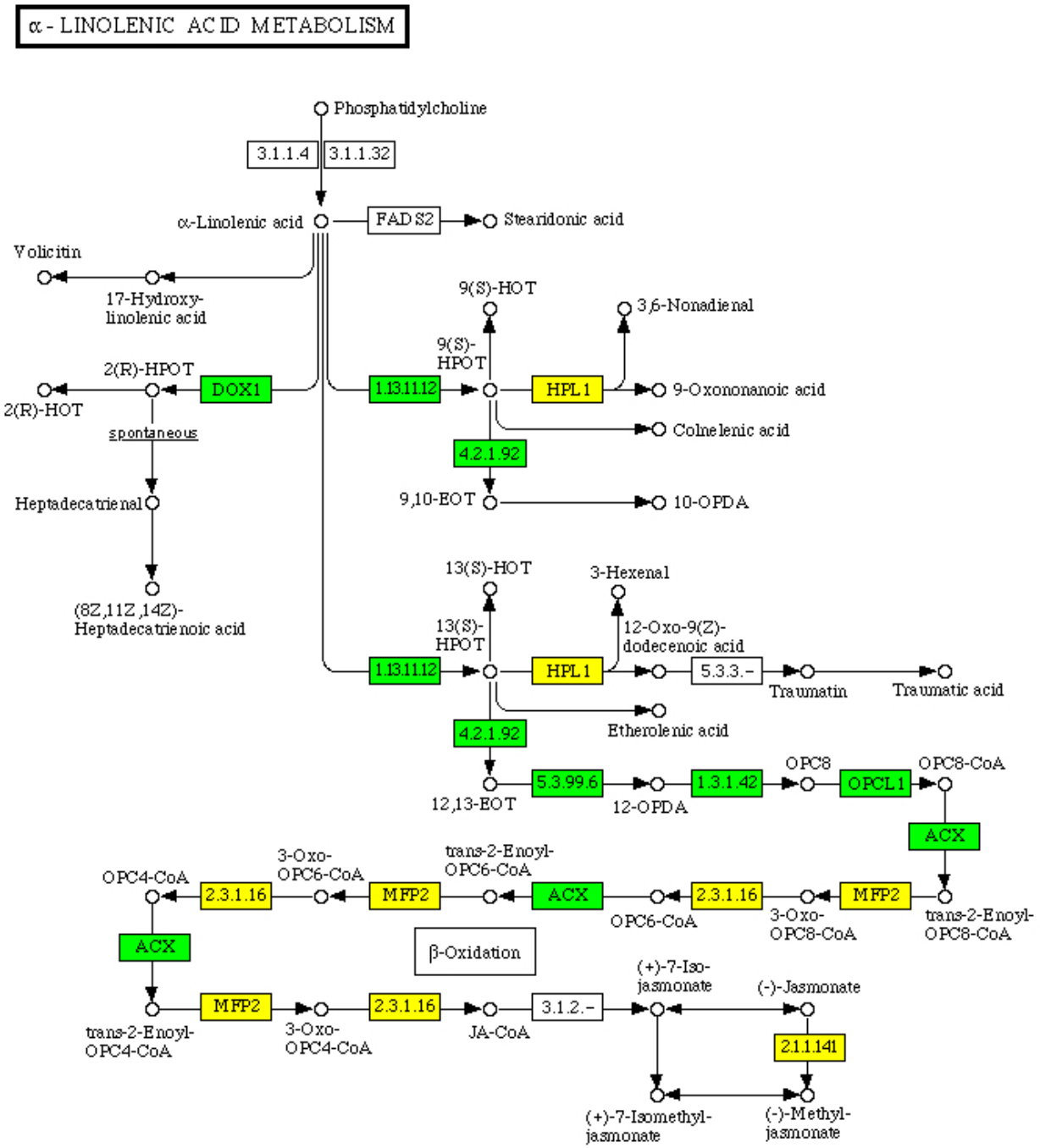


Figure 3.12: α-Linolenic acid metabolism and jasmonic acid biosynthesis from KEGG down-regulated in BSMT knockout.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by BSMT overexpression.

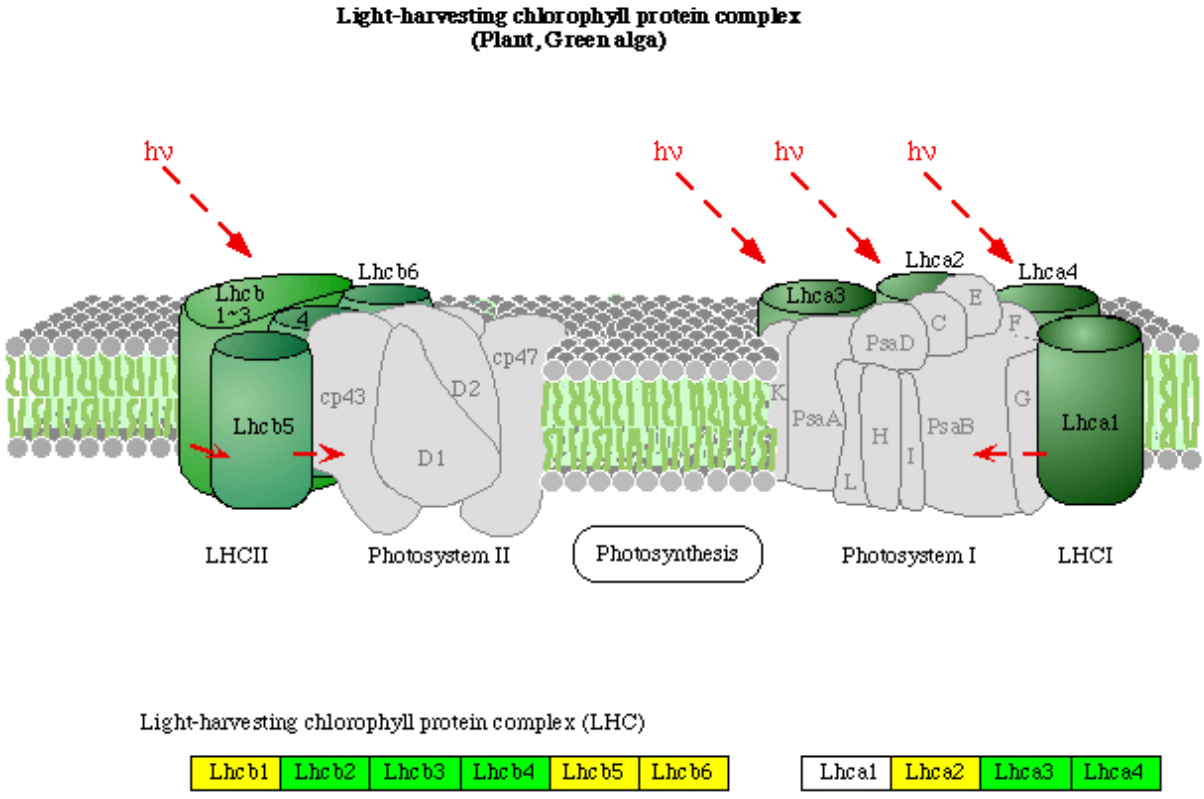


Figure 3.13: Photosynthesis antenna from KEGG is down-regulated in BSMT knockout. Colored boxes indicate the genes identified in the Arabidopsis genome. Green indicates the genes that were significantly repressed by BSMT overexpression.

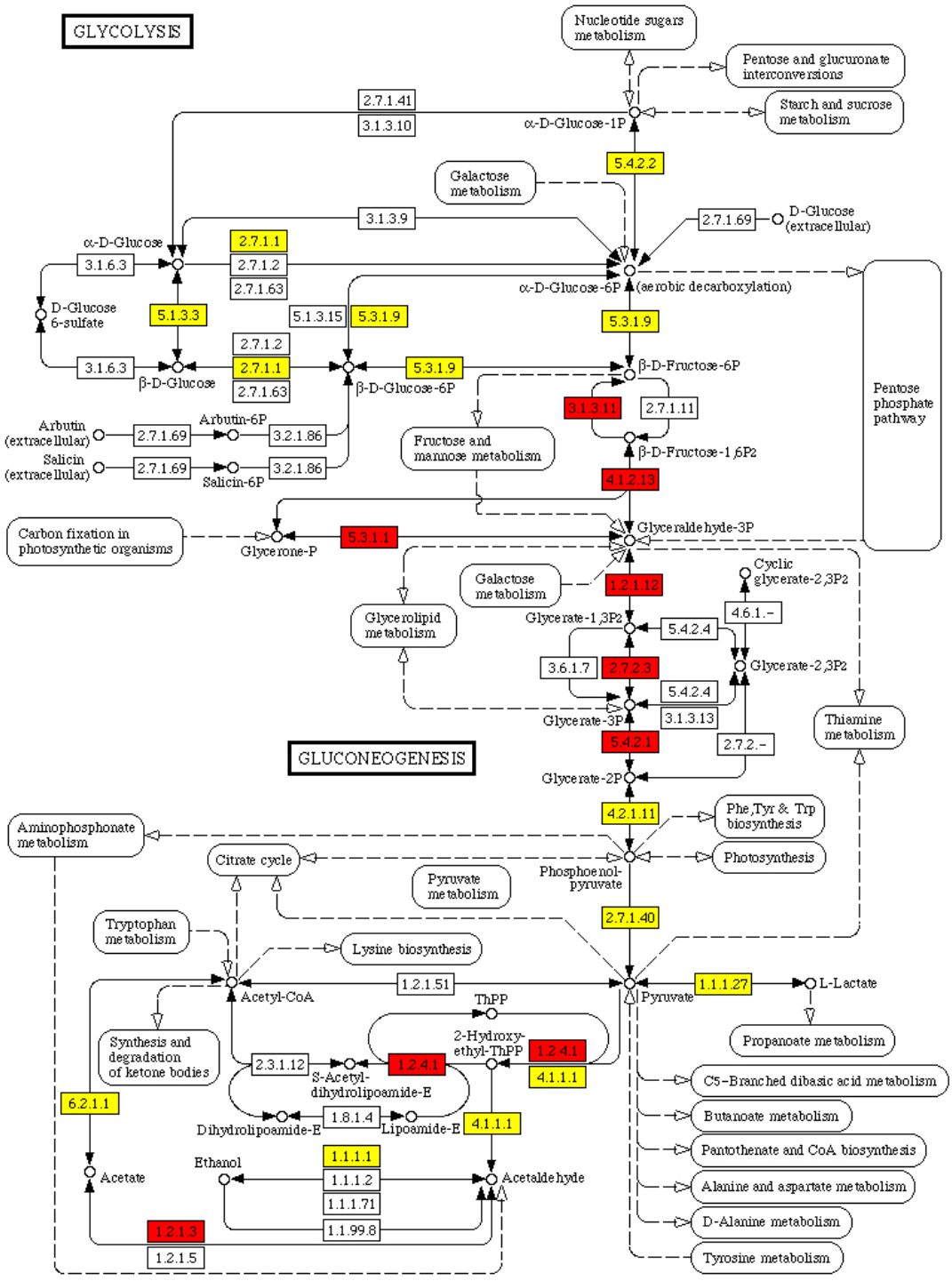


Figure 3.14: Glycolysis/gluconeogenesis pathway from KEGG up-regulated in BSMT knock-out.
Colored boxes indicate the genes identified in the Arabidopsis genome. Red indicates the genes that were significantly repressed by BSMT overexpression.

3.5 Conclusions

In the past few decades, molecular genetic approaches have been widely used to study gene functions in plants. Arabidopsis mutants with altered gene expression level are generated, and the effect of this modification on plant growth, development and metabolism are determined. Certainly, this approach has been very successful studying in plant metabolism, particularly genes involved in signaling pathways responsible for growth, development and stress response. The combination of reverse genetics and biochemical analysis has made possible the identification of chemical substrates for several SABATH methyltransferases, however gene function goes beyond the specific catalytic reaction in which a gene/protein is involved in. Gene fucntion knowledge is also about understanding the biological processes affected, the specific phenotype of the knockout or lost-of-function allele, in broarder scope understanding the consequences of that catalytic activity or reaction in the plant.

We found at the gene expression level, that converting signalling molecules to their methylesters can control their functional activity by regulating the homeostasis of their bioactive forms. Without intentional biotic or abiotic challenge, Arabidopsis plants growth under normal conditions, but with altered expression levels of SABATH methyl transferases, regulated genes involved in physiological processes known to be affected by high or low levels of bioactive signalling molecules.

The results presented in this chapter indicate that microarray analysis is a powerful tool for identifying genes involve in the molecular mechanisms, that may function in response to the regulation of signaling molecules by methylation. Mutations with no obvious phenotye like JMT overexpressor and BSMT loss-of-function mutant would likely be missed by traditional genetic screens. Despite the fact that important genes and pathways were regulated in response to abnormal levels of JMT and BSMT, the lack of obvious phenotype suggest that functional redundancy among the genes regulated by JMT and BSMT result in lack of phenotype observed for theses mutants.

The pathway analysis of the microarray data was based on gene enrichment according to current knowledge of the Arabidopsis metabolism. Further analysis of all the genes identified in this study may expand our knowledge of genes involve in responses modulated by signaling molecules, like JA, SA, and IAA, and may aid understanding the regulation of these molecules by methylation.

Chapter 4

Metabolite Fingerprinting of *Arabidopsis thaliana* with altered gene expression levels of SABATH Methyltransferases

"Nature goes her own way and all that to us seems an exception is really according to order"

Johann Wolfgang von Goethe

4.1 Introduction

The chemical phenotype of an organism is the measurement of all metabolites, as well as their fluctuations, in response to environmental changes. The comprehensive study of the chemical phenotype is referred to as *metabolomics*. Metabolomics seeks for a truly unbiased quantitative and qualitative analysis of all biochemical intermediates in a sample. Metabolites are defined as intermediate or final products of the cellular metabolism; the entire set of cellular metabolites found in an organism as a whole, is known as the *metabolome*. According to Sumner *et al.* [5], the ultimately goal of metabolomics involves and unbiased detection of the entire metabolome . However, it is not as simple as that, metabolomics is a challenging area of research considering the complexity of the metabolome. Whereas DNA and RNA are polymers of four nucleotides and proteins are made of the combination of 22 amino acids, there are estimates that the number of metabolites across living organisms are on the order of hundred of thousands [57, 157], made of a wide range of chemical components with significant variations in their chemical and physical properties. But, since metabolites have been proposed to offer a direct link between genotype and phenotype [158], and responsible for performing important biological functions, the importance of their study does not need to be mention to a long extend. Metabolite levels are the result of changes in growth and development, as well as environmental factors affecting the enzyme activity. As genes switch on and off, this usually result on changes in the proteins, and finally the metabolites levels change accordingly.

Three major approaches are commonly used in metabolomics studies. *Target analysis* is constrained to the measurement of one or a very few known target compounds. Such targets are usually quantified in an absolute manner using calibration curves and/or stable isotope labeled internal standards. Because of the diversity in their chemical properties, most analytical methods are design to target specific compounds [5, 12, 159, 160]. In the case of *metabolite profiling*, it attempts to measure a set of metabolites shared among different pathways or all metabolites of a specific pathway. However, metabolite profiling can be untargeted, attempting to detect all the metabolites present in a sample, to get insight knowledge when the metabolites affected by a particular condition under study are unknown or to acquire information about the overall metabolism glancingly [161, 162]. On the other hand, *metabolic fingerprinting* is different from the other two approaches in that it does not aim to physically separate individual metabolites nor attempt to identify all the metabolites in a sample; it rather considers the profile as a whole [163–165]. In metabolic fingerprinting samples are compared by multivariate statistical analysis to identify spectral regions

that discriminate samples based on their biological characteristics. Regardless of the approach taken, metabolomic analyses thereby aim to provide a comprehensive insight into the metabolic state of a system.

To investigate the effects of modulating the expression of several SABATH MTs at the metabolic level, we used gas-chromatography coupled to mass-spectrometry (GC/MS) to analyse the metabolite profiles of Arabidopsis plants where the expression of SABATH MTs differed from the wild-type. Using a metabolic fingerprinting approach to study the metabolic effects of single gene perturbations, we intend to identify the major consequences that changes in the expression levels of several SABATH MTs have on the overall metabolite composition. Thereby, here we present our results on the metabolite fingerprinting of wild-type, knockout and overexpressing lines, for several SABATH MTs, particularly for plant tissues where the genes are shown to be highly express.

4.2 Literature Review

4.2.1 Metabolite profiling in functional genomics

Plant functional genomics based on the analysis of gene products such as mRNA and proteins from transgenic and mutant plants, has contributed significantly to the understanding of gene functions otherwise unknown. However, these methods alone do not provide information about how changes on gene or protein expression translate into changes in biological function. Due to the complex regulatory mechanism at all levels in the cell, changes in gene expression, not always result in change in function or phenotype. Metabolomics can be used to explain the biochemical function of annotated genes. It can also be used to define phenotypes and to bridge the genotype-to-phenotype gap [158]. Moreover, metabolomics contributes significantly to the elucidation of gene function, because it offers the unbiased ability to differentiate phenotypes based on metabolites levels that may or may not produce a visible phenotype [166, 167]. For instance, in cases where mutations or expression of transgenes lead to measurable phenotypic changes, metabolite profiling can be used to determine the biochemical cause or consequences of the observed phenotype. Under the metabolite profiling concept the profiling concept, in order to define precisely the biochemical function of plant metabolism, the analysis should be unbiased and sensitive. Such analyses complement existing functional genomics approaches, such transcriptomics and proteomics, while offering a direct link between a gene sequence and the function of the metabolic network in plants [12]. Metabolomics is even more powerful when performed at large scale and integrated with corresponding data on transcriptomics and proteomics, few examples of this combined approach applied to plant functional genomics have been reported [8–10, 84, 168, 169]. Furthermore, metabolite profiling can elucidate links and relationships that occur primarily through regulation at the metabolic level [158, 170].

4.2.2 Challenges of Metabolomics

It has been stated that the ultimate goal of metabolomics is the complete and unbiased detection and measurement of all metabolites. The Identification of molecular substrates for proteins is a challenging task. The plant metabolome is massive, diverse and complex [5, 57]. It is constituted by a large number of compounds that differ significantly in structure, molecular weight and polarity. There have been estimates that the plant kingdom produces between 90,000 and 200,000 different metabolites [167]. With the improvement in analytical techniques the number of metabolites increases as well. In plants multiple metabolite subclasses include, but are no limited to: amino acids, flavonoids, lipids, fatty acids and organic acids [5]. Currently, no single analytical technique allows simultaneous analysis of all plant metabolites, because they separation modes lack sufficient resolution and selectivity to resolve thousands of compounds with a wide-range of physicochemical properties. Separating metabolites into several large groups and analysis of each individual class separately has become the technique of choice to reduce the chemical complexity of biological matrices [171, 172]. However, there is not clear definition between sub-metabolomes in which to built an analytical strategy, because a single analytical property does not offer enough selectivity to separate a single class from the entire metabolome. Consequently, different analytical techniques are being used to target metabolites that share similar chemical, physical and structural properties [165].

Another challenge is the dynamic range of metabolite concentrations present in the cell [165]. While metabolites that are in high abundance could be detected, very low abundance metabolites will never be found. How to cover the entire spectrum of metabolite concentrations and accurately detect molecules over a huge dynamic range. Additionally, many of the compounds measured in biological systems still represent unknowns. These analytical challenges have forced scientist to focus on measuring small subsets of metabolites and to refine their techniques to target specific compounds. Analytical developments are on the way to do untargeted metabolite profiling in more efficient way than how is being done nowadays [165].

4.2.3 Analytical technology for metabolite profiling

All commonly used techniques characterized the samples according to more than one physical parameter: in a first step metabolites are separated by chromatography according to their chemical properties (for example hydrophobicity or polarity). Depending on the substance class of metabolites under interest, gas chromatography (GC), capillary electrophoresis (CE), or liquid chromatography (LC), are among the most frequently used techniques. After the chromatographical separation of the samples, mass spectrometry (MS) is applied for metabolite detection. The parallel use of GC/MS and LC/MS comprehensively covers the accurate identification and quantification of a large fraction of all available metabolites [165, 167, 173].

In plant metabolite profiling many different analytical platforms have been used, includ-

ing GC/MS, LC/MS, CE/MS, HPLC-PDA (high-performance liquid chromatography with photodiode-array detection) and NMR (nuclear magnetic resonance). However, GC/MS or LC/MS are the tools of choice for generating high-throughput data for the identification and quantitation of small-molecular-weigh metabolites, and with the combination of them is possible to cover a wide range of chemical species [165].

4.2.3.1 Gas Chromatography-Mass Spectrometry

The earliest GC analyses of biological molecules were performed in the 1960's. Vanden-Heuvel *et al* [174]. were the first practitioner in steroid analysis, and it was the combination of GC with mass spectrometry that opened the door to the analysis of complex mixtures of metabolites. GC/MS provides a robust system with excellent separation capacities and high thought-put possibilities, and is therefore the most commonly used analytical technique for routine analyses in the field of plant metabolomics [175,176]. The separation of the analytes in gas chromatography is dependent on analyte interactions with the stationary face and the boiling point [177]. Since only volatile and heat-stable compounds can be analyzed by GC/MS, plant extract must be derivatise before GC analysis. Although, there are different different strategies for derivatising compound prior GC/MS analysis, the standard procedure in plant metabolomics is first use methoxyamine (CH_3-O-NH_2) in pyridine to stabilize carbonyl moieties in the metabolites and them functional groups such -OH, -COOH, -SH or -NH, are converted into TMS-ethers, TMS-esters, TMS-sulfides or TMS-amides, respectively, using trimethylsilyl (TMS) reagents, usually BSTFA or MSTFA. After derivatization a portion of the sample is introduced into the inlet of the GC instrument. The inlet temperature is often higher than 250°C, at which many metabolites are evaporated. Two different injection methods are most widely used; splitless and split. In splitless injections the whole sample is introduced onto the high resolution capillary column. In slip injections, only a portion of the volatilized sample goes into the column. Splitless injections is preferable to split injections for trace analyses. The length of the capillary column varies between 10 to 60 metres. The polarity of the column can also be varied by changing the stationary phases [177].

Metabolites eluding from GC are ionized by Electron-Impact (EI) or Chemical Ionization (CI). For metabolic analysis EI is the most commonly used technique. In EI vaporized metabolites are impacted by a beam of electrons with sufficient energy to fragment and ionize the molecule. The electrons are emitted from a heated filament and accelerated typically to 70eV. The number of fragment ions that is produced of each metabolite is a function of the electron impact energy. The source is designed so that when the ions are formed they are pushed out from the source and into the mass analyzer. EI results in molecular ion fragmentation, which is later used for structural interpretation of the metabolites and identification. In comparison with EI, CI is a much softer ionization technique, in which the ions are allowed to collide with reagent gas (often proton-rich like methane) to form abundant ad duct ions that contain the intact molecular species. This is advantageous for determining the molecular weight of metabolites [177]. To identify compounds, commercially

available databases of molecular ion fragmentation patterns of molecules, such as National Institute for Standard and Technology NIST (`http://www.nist.gov/srd/nist1.htm`), can be used, which contains approximately 200000 electron impact spectra. Unfortunately, many of the spectra included in NIST are of synthetic compounds, the number of derivatized plant metabolites is limited and their retention index information is incomplete. In addition to commercial libraries, in-house standard libraries have also been compiled containing spectra and corresponding retention indices, however this is a very laborious task to be completed in each laboratory that performs GC/MS analysis. The Max Planck Institute in Golm, Germany, has made its mass spectra library, which includes spectra of standard compounds, publicly available (`http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html`). Although this library is a help for many laboratories, the number of identified peaks in Arabidopsis extracts analysed by GC/MS is only 30-40%. Further information is needed to increase the number of identifications.

4.2.3.2 Liquid Chromatography-Mass Spectrometry

Another common method for metabolomics analysis is LC/MS. Instead of gas, the mobile face of the chromatography is liquid. Similarly like in GC/MS, once the metabolites are separated by LC they are detected by mass spectrometry. Unlike GC/MS, in LC/MS analyses the metabolites do not have to be volatile or possible to derivatise. After extraction and, if necessary, sample purification, the metabolites are separated by LC according to the differences in chemical properties of the metabolites present. Since there is not need for derivatise compounds prior to the MS analysis when using LC/MS, fewer sample preparation steps are required in LC/MS than in GC/MS analyses.

Several types of LC have been developed and serve different purposes. Reverse-phase liquid chromatography (RP-HPLC) methods are widely used and recognized and used in the field of metabolomics for the separation of non-polar compounds such as flavonoids, lipids, or phenolics. The separation of the extract depends on how the metabolites interact with the alkyl bonded spherical silica stationary phases. Many biologically important compounds do not separate easily on reversed-phase packing material. For example, ATP and GTP are very difficult to retain on a C_{18} column, due to their high polarity, but for such substances HILIC-type columns can be used instead. Hydrophilic interaction liquid chromatography (HILIC) is an alternative method of separation of highly polar compounds [178, 179]. C_{18} monolithic silica capillary columns have been used in plant metabolomics to improve chromatographic resolution [179]. Recently, a newly introduced very high-resolution "ultra performance LC" (UPLC) based on sub 2μm packing materials and high-pressure solvent delivery, in combination with MS, have been developed to improve the separation efficiency of metabolites [180, 181]. Both UPLC and monolithic columns improve peak separation, resulting in the ability to detect more peaks. However, despite the improvements provided by the introduction of of UPLC systems and monolithic columns, the optimal packing materials for separating specific classes of compounds differ widely. Consequently, the LC separation

of whole metabolomes is currently impossible using a single chromatographic system [179].

Following LC separation, the liquid mobile face carrying the compounds of interest, is introduced into the MS and the metabolites are ionized by appropriated techniques, such as electrospray ionisation (ESI) or atmospheric pressure chemical ionisation (APCI). In LC-ESI-MS the LC effluent is transported through a capillary with a high voltage (2-5 kV). This leads to an electric field gradient forming on the water surface. The polarity of the voltage (positive or negative) chosen depends on the analyte. By the action of a nebulising gas, from the capillary tip a "Taylor cone" is generated, containing charged liquid droplets. The droplets fly in atmosphere pressure towards the entrance of the mass analyzer. in ESI ionisation takes place at atmospheric pressure and mainly $[M+H]^+$ or $[M+H]^-$ ions are formed. Since the sensitivity of ESI-based methods towards different compounds varies, depending on whether positive or negative ions are detected, each sample should be analysed twice using both positive and negative modes, alternatively using positive/negative switching mode. On the other hand, APCI a form of chemical ionization for LC/MS, takes place at atmospheric pressure, and is usually used for substances that can not be analysed by ESI [171]. In comparison with electron impact in GC/MS both ESI and APCI fragmentation of molecular ions during ionization are much softer, and thus yield less information for mass interpretation. However, by running tandem MS (MS/MS) to provide fragmentation information [5, 179] and/or measuring accurate masses [181] structural information can be generated for identification. Mass spectra libraries for LC/MS analysis are less common than GC/MS libraries, partly because tandem mass spectra vary depending on the instrument and techniques used.

4.2.4 Data mining of metabolomics data

Due to the large amounts of data arising from metabolomics analysis, like other functional genomics approaches, a number of tools are available for computational needs, and others are under development [182–184]. Metabolomics data analysis has unique bioinformatics requirements, in addition to other common to microarray and proteomics data [172].

It is difficult to create a general work-flow that covers all the possible paths that could betaken in a metabolomic data analysis approach, because it depends in the type of questions that the scientist is looking to answer. In Figure 4.1 a generalized data mining strategy for data mining in metabolomics is outlined.

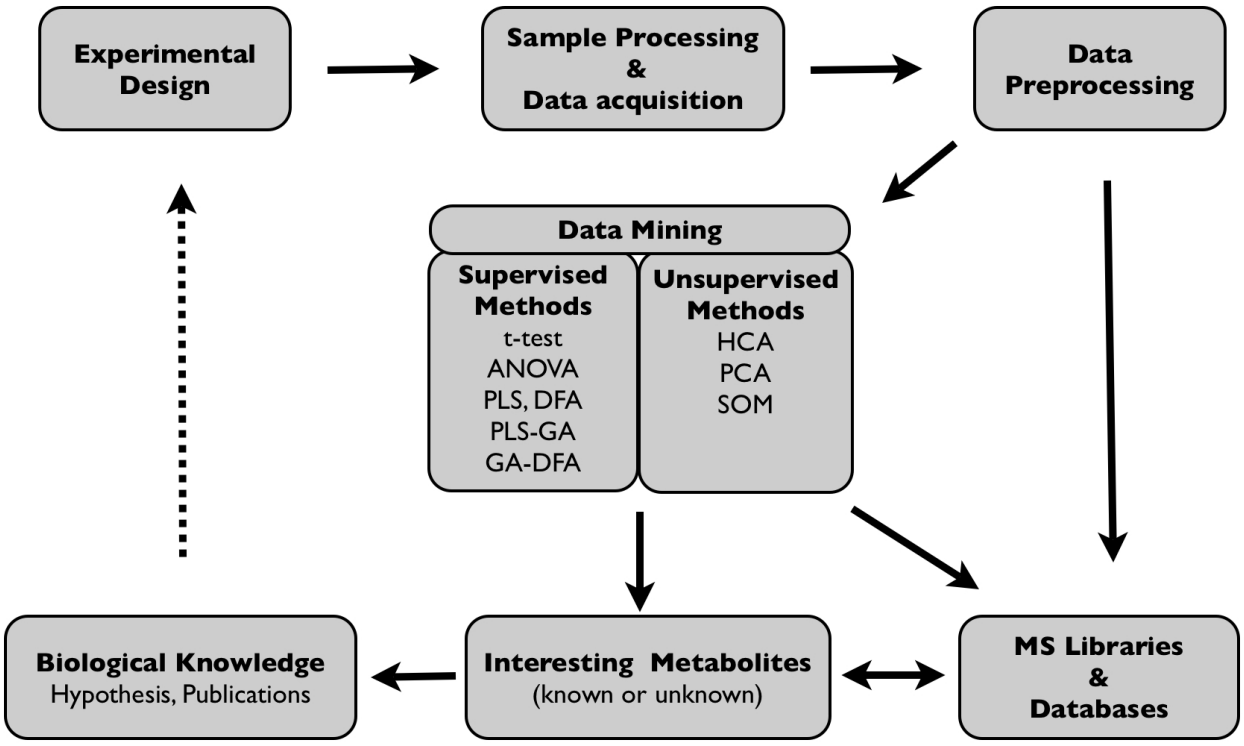


Figure 4.1: Schematic representation of a generalized approach to metabolomics data mining

Data analysis of metabolite data, at this moment in time, is still considered a developing area. Metabolomics data analysis have unique data processing requirements, but once the raw data has been processed other tools common to microarray, proteomics, and multivariate data in general, can be applied [171, 185].

The process usually starts with a particular experimental design, trying to answer a specific question. Samples are collected, extracted and the data is collected by means of one or several metabolite profiling platforms described above. The purpose of the data pre-processing is to extract from the chromatograms, the quantitative information of the metabolites, to generate variables that can be compared to each other between samples using data mining tools. This process represent a challenge for data acquired as a result of mass spectrometry coupled to a chromatography separation, because of its multidimensionality. This type of data has four dimensions: m/z (mass per charge), intensity, retention time and the different observations (samples and replicates), which need to be reduced to discreet variables per sample in order to make the comparable across a particular sample set. One option is to reduce dimensionality to m/z and intensity by summation over the time domain. In this case the output is similar to that obtained with direct in metabolic fingerprinting procedures like nuclear magnetic resonance (NMR) or fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), without the matrix co-suppression interference of this methods. the other option is to identify each one of the individual components (metabolites) through

deconvolution of the total ion chromatogram. Then the metabolites (known or unknown) become the variables, which must be align across all samples, creating comparable variables.

Before subjecting the data to further analysis it is pre-process depending on the biological information to be obtained, and also on the data analysis method chosen since different data analysis methods focus on different aspects of the data and are based on different assumptions. Centering, scaling and transformation are among the procedures used for metabolomics data and are discussed by van den Berg *et al.* [186].

With data extraction and pre-processing we have produces a data set containing the variables (metabolites or m/z) in a format suitable for further analysis. For an experimental design where samples are classified in two classes "A" and "B", it may be sufficient to compare and rank the list of metabolites according to p-values calculated by performing a t-test to determine if the means are equal or distinct. However, experimental designs are often more complex than two classes, *e.g.* time series, different developmental stages, different treatments or mutants. In experiments like these, a wide range of statistical tools, and machine-learning algorithms, generally called multivariate statistic methods, can be applied to metabolomics data. Multivariate analysis can be divided into supervised and non-supervised classification methods. Non-supervised algorithms like hierarchical clustering (HCA), principal components analysis (PCA) and self organizing maps (SOM), attempt to to find patterns that. explain the variance in a data set. the grouping obtained is derived from the data itself, rather than from information provided by the analyst. These methods are very sensitive to outliers and experimental error. On the other hand, supervised methods such as ANOVA, partial least squares (PLS) and discriminant function analysis (DFA), look at the differences among the biological groups based on classification of the samples in the data matrix. A more powerful approach to select discriminant variables, is the combination of supervised methods with evolutionary algorithms, such as genetic algorithm (GA) or genetic programing. Examples of this are PLS-GA [187], GA-DFA [188].

In addition to multivariate analysis and statistical methods, a number of databases, data mining and visualization tools are publicly available on-line. These include, for instance, pathway viewers and metabolic pathway databases such as KEGG (`www.genome.jp/kegg`), MetaCyc (`metacyc.org`), AraCyc (`www.arabidopsis.org/tools/aracyc`), KaPPA-View (`kpv.kazusa.or.jp/kappa-view`), and MapMan (`gabi.rzpd.de/projects/MapMan`); combined analysis of different omics data with PRIME (`prime.psc.riken.jp`); analysis of genome-wide mRNA, proteins and metabolite data with MetNet (`metnet.vrac.iastate.edu`); and a data model for plant metabolomics ArMet (`www.armet.org`).

4.2.4.1 Discriminant Function Analysis

We mentioned in the previous section that a powerful approach to select discriminant variables, is the combination of supervised methods with evolutionary algorithms, such as genetic algorithm (GA) [187, 188]. Pedro Mendes and the Biochemical Networks Modeling Group

at the Virginia Bioinformatics Institute, has written the software OMETER, which carries a number of multivariate statistical analysis and machine learning algorithms, applied to functional genomics and systems biology data (`http://mendes.vbi.vt.edu/tiki-index.php?page=ometer`). Among the different tool provided by OMETER there are two implementations of discriminant analysis with genetic algorithm variable selection. Both of them used as input a data file containing the information of the variables measured for each sample, and the classification of the samples into groups based on their biological characteristics or treatment. This information is used to minimize within-group variance and maximize between-group variance [189]. In the first algorithm (GA-DFA), similar to the one discribed by Jarvis and Goodacre [188], GA selects a predetermined (fixed) number of variables from the full data set, to formulate a robust model that give an indication of those variables as the most important to discriminate among the different classes of samples being analyzed. In the second algorithm (GA-DFA2), the initial number of variables is arbitrary, but this number will increase or decrease during the analysis execution, in an attempts to find the optimal number of variables for the classification. With either algorithm a limited number of variables from the full data set is selected and a model is build to distinguish between the diffreent type of samples, expressed as a single linear function by DFA [189]. Because of the stochastic nature of the GA-DFA algorithms, it ussually returns a different combination of variables as discriminant every time it is run. Therefore, the application of GA-DFA methods is best done by iterated run of the algorithm, followed by statistical analysis of the results to select the variables more frequently selected in the classification. For the purpose of our research, we used GA-DFA2 to find discriminant variables when metabolite data for multiple classes, such overexpress, knockout and wild-type samples were compared.

4.2.4.2 Data pre-processing and analysis with XCMS

Many software packages for metabolite analysis have become commercially available recently, especially for biomarker detection. Approximately 20 available alignment programs for GC–MS, LC–MS, nuclear magnetic resonance (NMR), and MS are known [190, 191]. In this study, we used the freely available tool XCMS [192](various forms (X) of chromatography mass spectrometry) for chromatogram alignment, data pre-processing and basic statistics analysis. The software XCMS is implemented with the freely available R statistical language, allowing selection of important peaks and/or components, and uses a t-test and the resulting p-values to rank important components. This program previously has been used successfully for several metabolomics-based studies, initially for the analysis of LC/MS data, but more recently for GC/MS data as well [192–198].

The XCMS package reads full scan data in a variety of formats including mzXML, mzDATA and NetCDF files. With XCMS the pre-processing and analysis is performed by several subsequent independent modules [192]. With the data files classified according to their biological nature, the first step is to identify the peaks per m/z per sample. This is to identify the maximum intensity values for each individual mass unit (± 0.1 m/z) in the chromatographic

time domain. The peak for each m/z are selected using a filtering function and a signal-to-noise ratio cut-off. After peak identification and filtering, the peaks representing the same components across samples are placed into groups and this information is used to correct correlated drifts in retention time from sample to sample. Aligned peaks with the adjusted retention time are the used to reread the raw data files to identify and fill missing peaks. After pre-processing the data can be exported a number of statistical analysis can be done. However, for the task of identifying differentially regulated metabolites, XCMS provides a simple t-test to identify metabolites whose intensities are significantly from sample class to sample class. In our analysis we used the list of significant m/z and their retention time generated by XCMS, to identify individual metabolites after peak deconvolution of the original raw data files with AMDIS (Automated Mass Spectral Deconvolution and Identification System) from the USA National Institute for Standards and Technology (NIST).

4.3 Methods for Metabolite Profiling

4.3.1 Plant Material and Growth Conditions

Plant material and growth conditions are the same as described in section 3.3. Similarly to the gene expression profiling, for metabolite profiling the analysis was performed in three different tissues, where each particular gene is highly expressed. The metabolite analysis was performed in wild-type, overexpressing lines and knockout mutants for JMT, IAMT, BSMT, GAMT1 and GAMT2. For JMT analysis, leaves were collected from 25-30 days old plants, which have fully developed rosettes but have not started blooming yet. For IAMT and BSMT analysis, flowers were collected from 35 days old plants, where approximately 30% of the flowers to be produced have opened, and siliques have not developed. Finally, for GAMTs analysis siliques half-way through maturation or fully matured, were collected from plants 45-60 days old.

4.3.2 Metabolite Extraction

Extraction and derivatization were performed according to Roessner *et al.* (2000) [199] and Fiehn *et al.* (2000) [160]. The different plant tissues, leaves, flowers and siliques were harvested, weighted, frozen in liquid nitrogen, and stored at –80°C. A minimum of 5 to 10 biological replicates were processed for each genotype-tissue. Additionally, five blanks consisting of extraction buffer with internal standards were also analyzed to account for potential carryover, intrument performance, and derivatization efficiency. Ten milligrams of freeze dried tissue were grounded using 3.2 mm chrome-steel beads in a Retsh mixer mill, followed by extraction with 500 μL of 50:50 methanol:water with 0.5 mg/mL solution of ribitol, added prior the extraction. After the extraction buffer was added mixture was

stirred on Retsch mill for 1 minute at 30 reps/sec, sonicated for 1 min in ultrasonic bath and incubated on dry ice for 5 min. Three cycles of sonication and dry ice incubation were carried out before spinning out the extract 15 min at 4^{o}C at 13,000 rpm. One hundred and fifty μL of the clear supernatant was transferred into limited volume HPLC vials and analyzed by LC-MS.

For GC/MS, the remaining 350 μL were transferred back to the original tubes with the sample pellet, 500 μL more of extraction buffer were added and the mixture was vortex briefly. The samples were incubated at 70^{o}C for 15 min. The extract was partitioned into polar and non-polar fractions using liquid-liquid extraction with 400 μL of chloroform containing docosanol internal standard (40mg/mL). After vortex the samples were centrifuged 5 min at 4^{o}C at 13,000 rpm. The upper polar phase, approximately 750 μL were transferred to 1.1mL chromacol vials and dried in a Centrivap benchtop centrifugal concentrator (Labconco, Kansas City, MO) overnight. For derivatization 80 μL of 20 mg/mL methoxyamine hydrochloride, were added to the dried polar phase, followed by 90 min incubation at 30°C. The mixture was trimethylsilylated with 80μL of MSTFA+1%TMSC for 30 min at 37°C to prepare volatile and thermally stable derivatives. Samples were incubated at room temperature for 2 h before injection into GC/MS.

4.3.3 GC-MS Analysis

For GC/MS analysis, sample volumes of 1 μL were injected at a split ratio of 25:1 into a Trace DSQ GC/MS system (Thermo Finnigan, Austin, TX) equipped with Combi-Pal autosampler (Leap Technologies, Carrboro, NC). Tuning was done using tris(perfluorobutyl)amine (CF43) as a reference gas. Chromatography was performed using a 30m x 0.25mm x 250-μm Alltech AT-5MS column (Alltech Associates, Deerfield, IL). Injection temperature was 230°C, the interface was kept at 250°C, and the ion source was kept at 200°C. Oven temperature program was 2 min at 70°C, followed by a 5°C/min ramp to 310°C, 1 min at 310°C, and a final 5 min at 70°C before the next injection. Carrier gas was helium at a constant flow of 1 mL/min. Ion source filament energy was 70eV. Mass spectra were recorded at two scans per second over a range of 50 to 650 m/z.

4.3.4 Statistical Analysis of MS Data

4.3.4.1 Discriminant Function Analysis

For the selection of variables discriminant among the mutants for each gene and the corresponding wild-type, we have been employing a novel approach to metabolic fingerprinting as recently described [189], with modifications. Rather than attempting to identify all the molecules obtained from metabolic profiling, we focus only on those metabolites that demonstrate to be determinant among sample groups. We applied the mutivariate statisti-

cal method called discriminant function analysis (DFA) coupled with genetic algorithm (GA) variable selection (GA-DFA), discribed in the previous section, to process data generated through GC/MS. The classification software used in the analysis was OMETER v. 0.60 written by Pedro Mendes and available at `http://mendes.vbi.vt.edu/tiki-index.php?page=ometer`.

Before GA-DFA analysis the data was preprocessed as follow: (1) for each sample, only data collected within 7-50 min were used to eliminate unretained compounds, specially those used during the sample derivatization; (2) following metodologies described by Allen *et al. [163],* each GC/MS array was reduced into a single MS vector by summing the ion counts of a given m/z ratio over the total scan time, then each MS vector, containing m/z ratios and their total intensity, was simplified to unit m/z ratio (ion counts of fractional m/z ratios were added to the nearest integer m/z). As a result, after this initial data reduction, an MS chromatogram with m/z range 50 to 650 will be reduced to a single vector having 600 values. Reduction of the data to a single MS vector was performed with Xcalibur software (v 2.0, ThermoFinnigan); finally (3), every MS vector was normalized to the total ion count for that vector, so that different spectra can be compared quantitatively. The data obtained upto this point in the analysis is equivalent to that obtained by infusion, without the co-suppression problem. All the MS vectors in a data set were exported and formatted before analysis into a single matrix of N objects $\times$ V variables, where N is the number of MS vectors or samples and V the m/z ratios in the mass range. This matrix is then used as input file for OMETER, stating the class to which each sample belongs according to its nature (wild-type, overexpress or knockout). GA-DFA2 was applied with 5000 iterations with a initial number of discriminant variables set to 9. The results were then compiled in tables of frequency, indicating how many times each m/z ratio was selected by GA-DFA. The top selected m/z ratios are ones most likely to discriminate between classes. To validate this multivariate analysis approach, we carried out DFA analysis with selected m/z ratios that were most frequently selected during the 5000 iterations. The top selected m/z ratios and a subset of m/z ratios that were selected together with the top m/z ratios, were used to identify cadidate metabolites using the original spectro-chromatograms and a custom made mass spectra library, built in Dr. Vladimir Shulaev's laboratory using comertially available compounds.

4.3.4.2 Pre-processing and Analysis with XCMS

XCMS was used for automated and unbiased peak finding, using the MS NetCDF file format. At the Shulaev laboratory, Xcalibur software (v 2.0, ThermoFinnigan) is used for general data handling. Converting the Xcalibur (*.raw) files to NetCDF (*.cdf) format was done with the included Xconvert program. The CDF files were subsequently processed in batches (*e.g.* wild-type, overexpress and knockout for each JMT) by XCMS using default settings except for the following XCMS processing parameters: "`minfrac`" the minimum fraction of samples necessary in at least one of the sample groups for it to be a valid group was set

to 0.5; "`minsamp`" the minimum number of samples necessary in at least one of the sample groups for it to be a valid group was set to 4; "`bw`" the bandwidth of gaussian smoothing kernel was set to 30 s. The XCMS software identified a specific m/z at a specific retention time, which were subsequently aligned with respect to the retention time across the analyzed samples. The processed m/z ions are reported in the format M256T987, which represent an ion with m/z 256 at the retention time of 987 s. For each sample, an intensity value is associated with the respective m/z ion. Also associated with each m/z ion is the retention time interval and median value in which that specific m/z identified and the number of times a certain m/z was present among the samples within each class.

4.4 Results and Discussion

4.4.1 Metabolic Fingerprinting and GA-DFA

Genetic algorithms have been previously applied for the identification of discriminatory variables withing multivariate data [164, 189, 200]. In a mass spectra data set, genetic algorithm combined with discriminant fuction analysis (GA-DFA), can be used to search for m/z ratios that differentiate between *a priori* known classes, that is, it looks for m/z ratios present in all the samples belonging to one class but that are absent in the others, or are significantly different in intensity [189, 200]. GA-DFA can be run mutiple times on a single data set producing different discriminant models on each iteration using a subset of variables from the available data. Then the variables used in each model can be combined and the frequency of selection for each variable calculated [164].

In this study, to identify specific compounds that are significantly altered by changes in the expression of JMT, IAMT, and BSMT, we performed gas chromatography-mass spectrometry (GC/MS) analysis of polar compounds of the different mutants and wild-type, followed by GA-DFA to identify discriminant m/z ratios among the mass spectra of the wild-type and the mutants for each gene. For this analysis we used the same batch of tissue used for microarray analysis, and the sample extracts were prepared from the tissues used for microarray analysis described in section 3.3. Leaves extracts were used for JMT wild-type and mutants, while flowers extracts were used for IAMT and BSMT wild-type and mutants.

Simple analysis of GC/MS data files for leaves extracts of wild-type plants using AMDIS software, revealed the presence of 219 different components, with 155 matching targets in the mass spectra library. Similarly, the in the GC/MS profiles of flowers 826 compounds were found by deconvolution with AMDIS, and 444 matched targets in the mass spectra library. Even though the numbers of possible metabolites present in the extracts may be overestimated by the sensitivity of the deconvolution algorithm, it gives an approximated idea of their complexity. To identify the most discriminant m/z ratios, and therefore their corresponding metabolites, we ran GA-DFA for 5000 iterations. Separately for the wild-type

and mutant lines for each gene, GA-DFA was performed on three classes of samples at the time (wild-type, overexpress and knockout). The selection frequency of discriminant m/z ratios is summarized in Figures: 4.2, 4.4, and 4.6, for JMT, IAMT and BSMT, respectively. Using the top m/z ratios selected by GA-DFA after 500 iterations, a DF model was generated for each comparison, which resulted in a clear separation among wild-type, and the mutants (Figures 4.3, 4.5, and 4.7). Despite the fact that most of the top m/z ratios were selected by GA-DFA, less that 10% of the total number of times the algorithm was ran, they still contain enough information to discriminate the different samples classes. The DF models generated with a subset of the top m/z ratios have a classification accuracy of 100%, 86% and 87% for JMT, IAMT and BSMT sample sets, respectively (Figures 4.3, 4.5, and 4.7). For IAMT and BSMT, we belive, the accuracy of the classification obtained is expected according with the conditions of the experiment. As mentioned in section 3.4.2, the T-DNA insertion mutant *iamt* (SALK-072125), is not a complete knockout because it produces a truncated transcript for IAMT. Similarly, BSMT overexpressing showed significant similarities to the knockout, suggesting that the phenomenon of co-suppression by overexpression may have occured. An improved classification model can be obtained if only IAMT overexpressing and *bsmt* knockout are compared to the wild-type.

Metabolic fingerprinting is the rapid classification of samples according to their biological origin or nature [12], where is not necessary or feasible to determine the levels of metabolites individually, it rather develop a rapid and efficient approach to identify the specific metabolic signatures of each sample class, allowing the sample classification and pattern recognition among large data sets [201]. Overall, metabolite profiling combined with GA-DFA was succesful in distinguishing wild-type from mutant samples, and identifying the speficific m/z ratios characteristics of the gene expression variation for JMT, IAMT and BSMT. Although the biochemical significance of the m/z ratios, and the metabolites they belong to, requires further investigation, GA-DFA has made it possible to reduce the number of candidate metabolites that are changing when the expression of JMT, IAMT and BSMT is modulated. The next logical step will be to used the top discriminant m/z ratios selected by GA-DFA to identify candidate metabolites that the phenotype observed as a consequence of modulating the gene expression of the SABATH MTs studied.

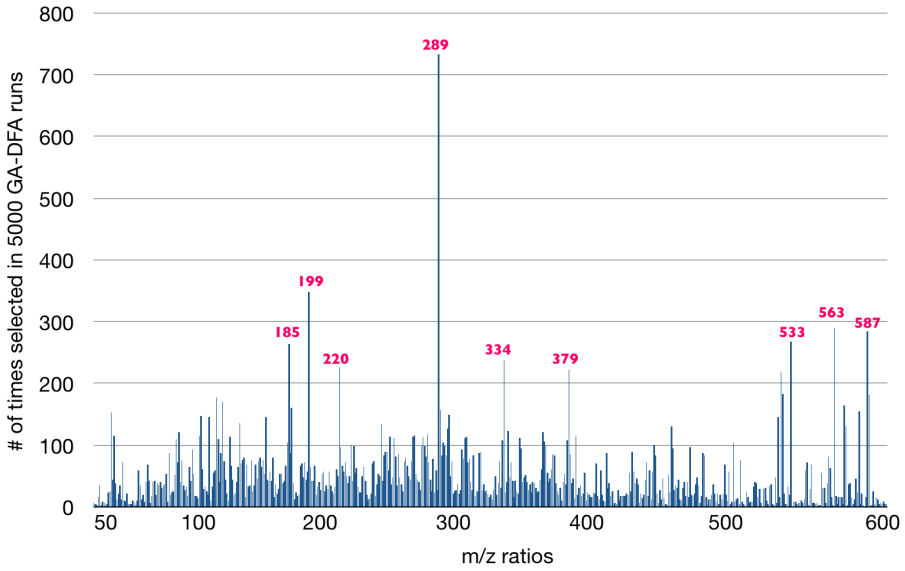


Figure 4.2: Histogram of m/z ratios chosen by GA-DFA on mass spectra data of leaves extracts for JMT overexpress, knockout and wild-type.
GA-DFA was applied to three classes of samples at the time: wild-type, JMT overexpress, and *jmt* mutant. Represented in the histogram are the frequency of selection for each m/z ratio to distinguish among the three classes in 5000 GA-DFA runs. The top m/z ratios selected as the most discriminant were: 185, 199, 220, 289, 334, 379, 553, 563, and 587.

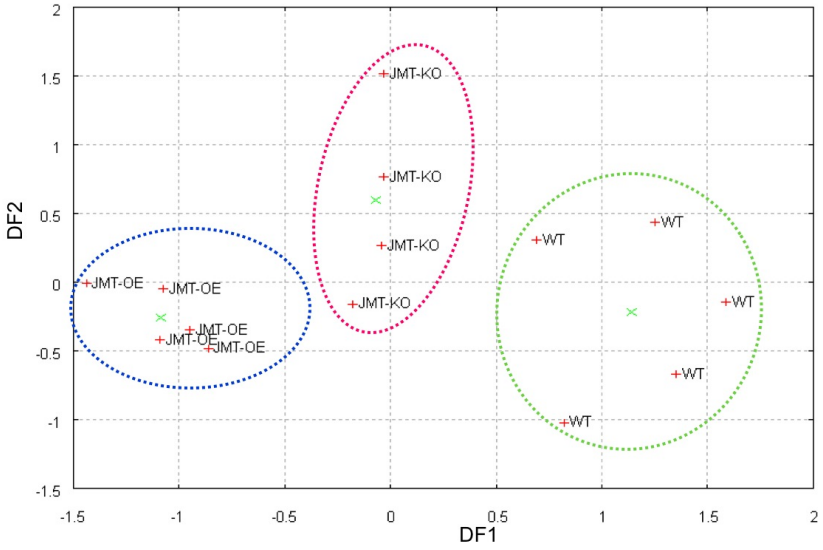


Figure 4.3: DFA with a subset of selected m/z ratios chosen by GA-DFA on mass spectra data for wild-type, JMT overexpress and *jmt* mutant. Three of the top m/z ratios (199, 289, 563) for DFA analysis.

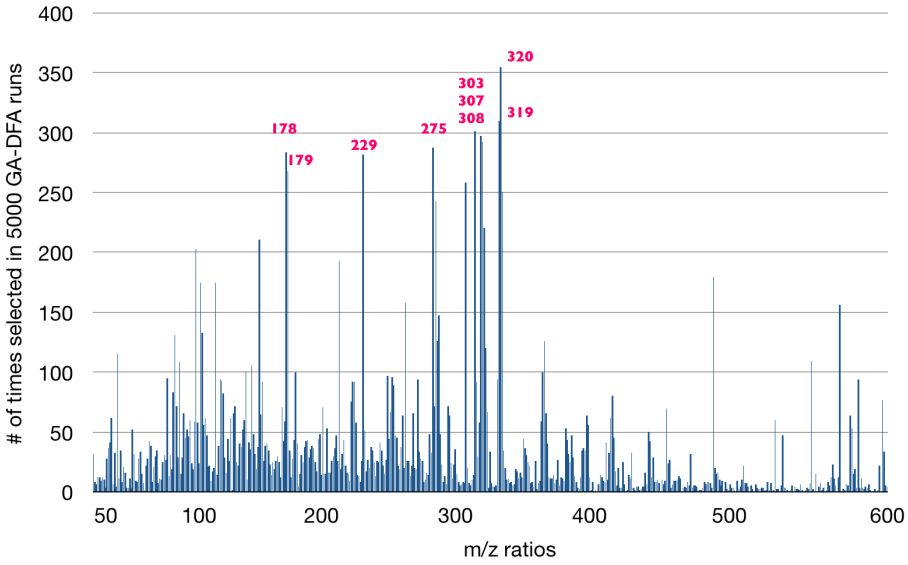


Figure 4.4: Histogram of m/z ratios chosen by GA-DFA on mass spectra data of flower extracts for IAMT overexpress, knockout and wild-type.
GA-DFA was applied to three classes of samples at the time: wild-type, IAMT overexpress, and *iamt* mutant. Represented in the histogram are the frequency of selection for each m/z ratio to distinguish among the three classes in 5000 GA-DFA runs. The top m/z ratios selected as the most discriminant were: 178, 179, 229, 275, 303, 307, 308, 319, and 320

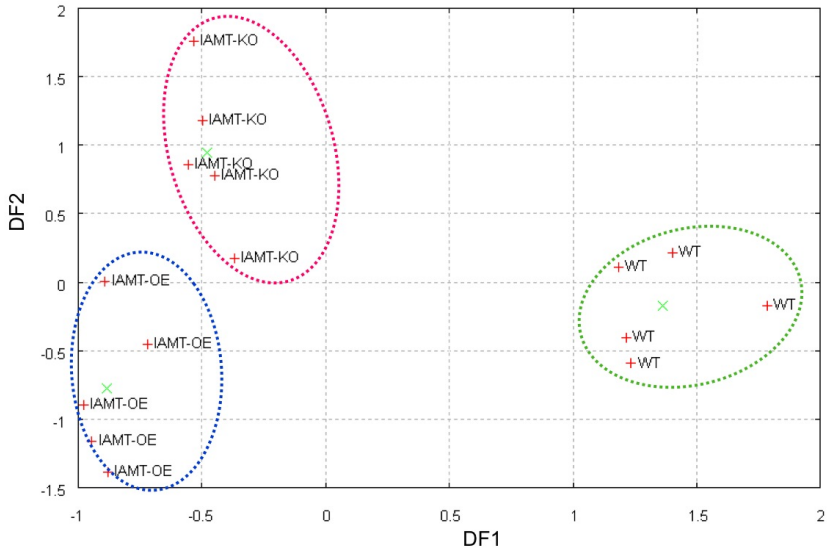


Figure 4.5: DFA with a subset of selected m/z ratios chosen by GA-DFA on mass spectra data for wild-type, IAMT overexpress and *iamt* mutant. The top five m/z ratios (178, 179, 303, 319, 320) selected by GA were used for DFA analysis.

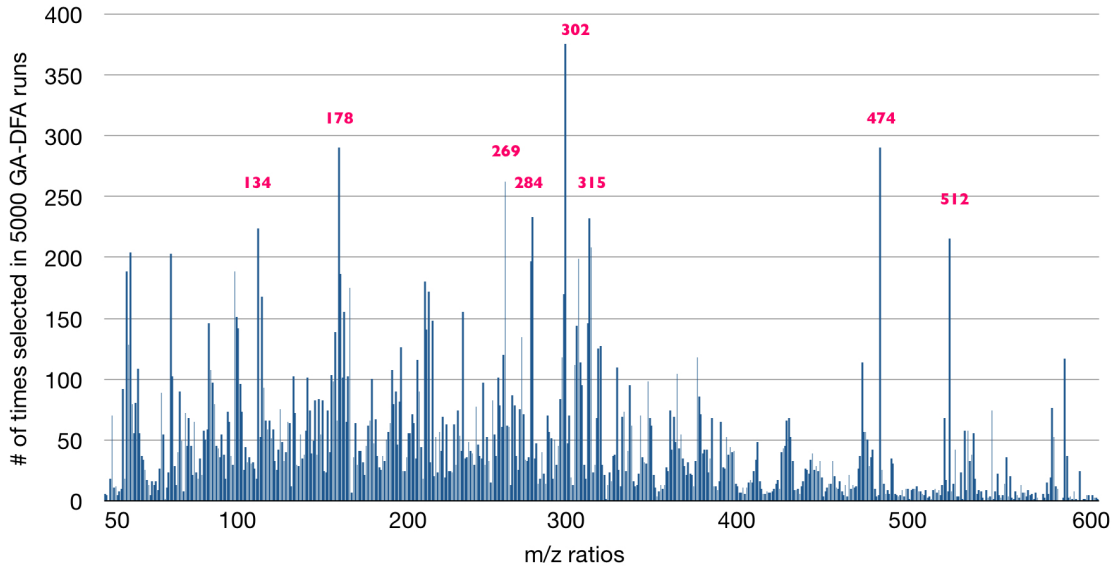


Figure 4.6: Histogram of m/z ratios chosen by GA-DFA on mass spectra data of leaves extracts for BSMT overexpress, knockout and wild-type.
GA-DFA was applied to three classes of samples at the time: wild-type, BSMT overexpress, and *bsmt* mutant. Represented in the histogram are the frequency of selection for each m/z ratio to distinguish among the three classes in 5000 GA-DFA runs. The top m/z ratios selected as the most discriminant were: 134, 178, 269, 284, 302, 315, 474, 1nd 512.

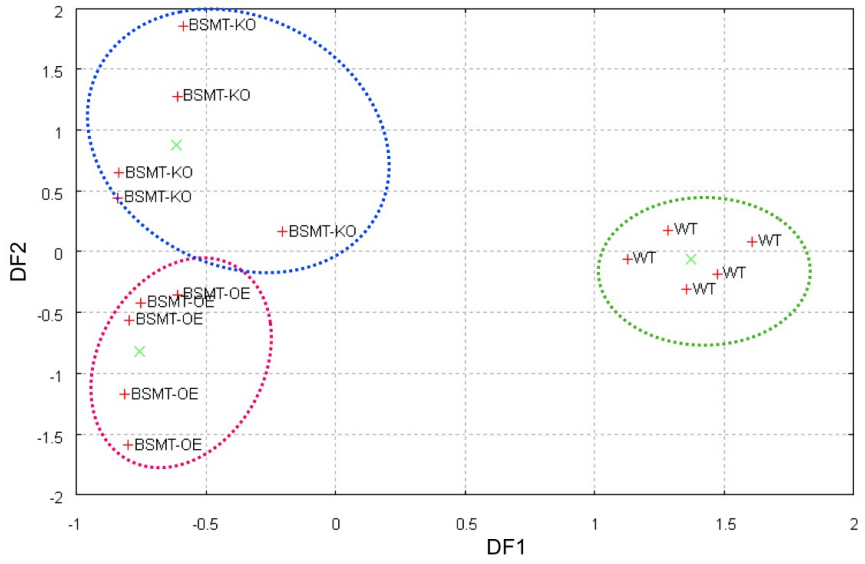


Figure 4.7: DFA with a subset of selected m/z ratios chosen by GA-DFA on mass spectra data for wild-type, BSMT overexpress and *bsmt* mutant. The top five m/z ratios (178, 269, 284, 302, 474) selected by GA were used for DFA analysis.

4.4.2 GC/MS Preprocessing and Analysis with XCMS

The process of converting analytical data to biological knowledge requires the convertion of raw analytical data (m/z vs retention time vs ion intensity) to a data matrix that is aligned in both the m/z and retention time dimentions [193]. In addition to the multivariate statistics analysis of the GC/MS data with GA-DFA, we performed an alternative data analysis, which combined preprocessing the mass spectra data with XCMS, followed by ANOVA. One of the major advantages of GA-DFA, is that it allows to identify the most dicriminant m/z ratios in a complex multiclass mass spectra dataset, reducing significantly the number of putative metabolites responsible for the differences among the original samples. However, further analysis needs to be done to identify the metabolites related with the discriminant m/z ratios. This identification process can be easier for LC/MS data, specially if it is obtained through an accurate mass spectrometry intrument. However, in the case of GC/MS-EI (electron impact) data a particular m/z ratio, may have originated from multiple molecules. Therefore, complementary to the GA-DFA analysis we used XCMS preprocessing combined with ANOVA. As a result is possible to stablish the statistical significance of each m/z ratio, and its retention time, which allows the identification of the metabolite.

We used XCMS on the same sample set for JMT wild-type, knockout and overexpress, used for GA-DFA analysis in the previous section. The total number of features found by XCMS after peak grouping and retention time correction in the JMT sample set was 1320. Depending on the performance of the peak picking algorithm these features either can represent m/z ratios that were found to differ from background noise or can be associated with real components or metabolites. This number is significantly different compared to the correct deconvolution of peaks with AMDIS software, with 219 individual component dentified. In XCMS a component is assigned to an m/z ratio in a particular time window, while in AMDIS the deconvolution process takes in consideration multiple m/z ratios at the time to identify a particular component.

In XCMS the aplication of statistics is limited to a paired t-test. In our case with more than two classes to be compared (wild-type, overexpress, and knockout) the data was exported from XCMS and a simple one-way ANOVA was performed, which allowed the identification of 221 features statistically significant with a p-value $\leqq 0.05$. It is important to mention that there is not linear correspondance between the number of features identified by XCMS and the number of significant metabolites. In GC/MS-EI multiple m/z ratios can be obtained for a single metabolite, due to the electron impact fragmentation. Consequently, multiple significant m/z ratios in a window of time may corresponde to a single metabolite. In Table 4.1, is listed a representative subset of the 221 GC/MS-XCMS components that were found significant when JMT mutants and wild-type. The XCMS analysis assigns names to each one of the features in the format M147T1161, as shown in Table 4.1, where M147 is the m/z ratio and T1161 is the retention time in seconds where that particular m/z ratio is present. With the convined information for several or all the m/z ratios significant at the same retention time, we identified the corresponding metabolites using the devonvolution

and mass spectra library search software AMDIS.

We compared the significant compunds identified by XCMS-ANOVA with those identified with GA-DFA. While 7 components were identified with GA-DFA, 14 were identified with XCMS-ANOVA, including those identified by GA-DFA (Table 4.2).

XCMS preprocessing was also applied to the GC/MS data sets for IAMT, BSMT and GAMTs. The results obtained for GAMT1 and GAMT2 XCMS-ANOVA are presented in Chapter 5. However, for IAMT and BSMT the quality of the deconvoluted data was not good, like the one obtained on the data set for JMT or GAMTs. There are two possible reasons to explain why this may have occured. First, the GC/MS data for both IAMT and BSMT was obtained through splitless injection. In this mode all the analyte sample vaporized in the injector goes onto the column for chromatography, this used when the sample contains very small amounts of analyte. However, because all the analyte mass in a 1 μL injection goes on column, it may result in column overloading and significant drifts in retention time, which impose a challenge to peak grouping and retention time alignment by XCMS. This lead us to the second reason, that the XCMS parameters need to be optimized for this complex data sets. It has been reported that the operation of XCMS software is complex, but flexible, with many parameters to be tuned by the user [193]. In previous studies it has been shown that a great deal of time need to be expended optimizing several parameters that can significantly affect the quality of the peak alignment, retention time correction, and ultimatelly the reproducibility of the results [193, 194, 196]. This could be considered a significant advantage of GA-DFA over analyses based on XCMS.

Table 4.1: Results of statistical analysis of GC/MS data by XCMS and ANOVA, showing a subset of the most significant components that can be used for discrimination of JMT wild-type and mutants, and the putative compounds they derived from according to mass spectra library search.

GC/MS-XCMS Component Name	Component ID	RT	F value	p value
M122T618	Unknown	10.30	5.52	0.022
M159T701	Unknown	11.69	4.87	0.031
M187T874	GSSG-2TMS	14.56	5.70	0.020
M100T1011	Ethanolamine-TMS	16.86	9.46	0.004
M147T1161	Fumaric Acid	19.35	5.84	0.019
M373T1258	Unknown	20.97	4.87	0.031
M177T1478	Threonic Acid	24.63	16.42	0.003
M185T1647	Xylose	27.45	6.58	0.013
M395T1703	Unknown	28.38	4.64	0.035
M311T1889	Fructose	31.49	9.44	0.004
M180T1917	Galactose	31.95	10.65	0.003
M525T2039	Gluconic Acid	33.98	16.91	0.000
M380T2138	Myo-inositol	35.63	8.87	0.005
M347T2686	D-Sucrose	44.77	19.80	0.000

Table 4.2: Metabolites detected by GC/MS in polar extracts of Arabidopsis siliques of wild-type, JMT overexpressing,and *jmt* knockout.

Derivative	RT[a]	Wild-type		JMT overexpressing			*jmt* knockout		
		Resp ratio[b]	SE[c]	Fold Change	Resp ratio	SE	Fold Change	Resp ratio	SE
Unknown_RT10.31	10.31	0.023	0.006	1.244	0.029	0.005	0.853	0.020	0.002
Unknown_RT11.69	11.69	0.010	0.002	**2.391**	0.024	0.005	1.359	0.014	0.004
GSSG-TMS2	14.58	1.647	0.453	1.016	1.673	0.281	1.020	1.680	0.254
Ethanolamine	16.86	0.100	0.049	0.676	0.067	0.009	0.428	0.043	0.007
Fumaric Acid*	19.39	23.127	8.561	1.297	30.006	9.224	1.235	28.558	2.309
Unknown-RT20.97	20.97	0.001	0.001	0.298	0.000	0.000	**2.195**	0.003	0.001
Threonic Acid	24.63	0.010	0.003	**2.243**	0.022	0.007	**1.667**	0.016	0.002
Xylose*	27.43	0.193	0.024	1.407	0.272	0.071	1.073	0.207	0.023
Unknown-RT28.38	44.77	0.005	0.000	0.876	0.005	0.001	1.001	0.005	0.000
Keto-Gluconic Acid*	29.4	0.006	0.003	0.093	0.001	0.000	0.094	0.001	0.000
Citric-Acid*	30.3	1.154	0.404	0.584	0.674	0.180	**1.698**	1.959	0.511
Fructose-MeOX-5TMS*	31.43	0.545	0.217	**2.471**	1.346	0.413	**2.623**	1.428	0.196
Galactose-MeOX*	31.95	2.898	1.045	**1.837**	5.322	2.227	1.704	4.937	0.744
Gluconic Acid	33.93	0.030	0.016	**14.958**	0.454	0.180	**13.043**	0.396	0.075
Myo-Inositol*	35.65	0.001	0.000	**2.152**	0.003	0.001	**1.779**	0.003	0.000
Unknown-RT39.74	39.74	0.070	0.027	**3.091**	0.218	0.084	**2.716**	0.191	0.020
D-Sucrose	44.77	0.813	0.210	0.026	0.021	0.010	0.123	0.100	0.034

Among over 219 components found present in Arabidopsis leaves polar extracts by GC/MS analysis, these metabolites were identified using the top m/z ratios selected by GA-DFA and the most discriminant components identified with XC-MS-Anova. GC/MS analysis of polar extracts from Arabidopsis leaves of wild-type, JMT overexpressing, and *jmt* knockout. Polar extracts were derivatized and analyzed as described in section 4.3. All samples were methoximated and trimethylsilylated. Compounds with a fold change increase of 1.5 or more are indicated in bold.[a]RT for retention time in minutes. [b]Response ratios are peak areas compared to the internal standard ribitol/adonitol. [c]SE, standard error, n=5. *Components identified with GA-DFA only. Values of 0.000 are <0.0001

4.5 Conclusions

A common apporach in functional genomics is to design experiments on gene knockouts or controlled overexpressors to study the phenotype related to a particular gene function. Similar to other "omics" approaches, metabolomics can generate large amounts of data from these genetic manipulations, however identifying which metabolites are of key importance to the gene function is essential to generate new knowledge.

Our approach to apply GC/MS profiling in combination with GA-DFA enable us to classify forward genetic mutants according to their metabolic fingerprints, even in cases where there are not obvious phenotypical differences, like JMT overexpressor and knockout, or BSMT overexpressor and knockout. The application of GA-DFA identified a subset of key m/z ratios within the mass spectra, for the discrimination of wild-type and mutants for each gene. Using this information the number of metabolites that change in response to modulating the gene expression of JMT, IAMT, and BSMT is reduced to only those metabolites which mass spectrum contains the dicriminant m/z ratios selected by GA-DFA. The logical extension of this research is to identify the metabolites where the discriminant m/z ratios are present in the original GC/MS raw data files.

Processing of raw data using XCMS-ANOVA approach resulted in time aligned ion features, defined as unique m/z ratios at a unique retention time, which allowed the identification of the significan metabolites. XCMS, is a useful tool for data analysis if the software parameter are carefully optimized and the quality of the aligment results.

A combined analysis with GA-DFA and XCMS-ANOVA, allowed the identification of a set of known and unknown metabolites that change significantly when the expression of JMT is modulated. This is the first step to understand the consequences of JMT expression modulating at the metabolic level.

Chapter 5

Regulating Biological Activity in Arabidopsis Seeds with Gibberelic Acid Methyltransferases[1]

[1]Part of the information contained in this chapter is based on published material, used with permission of the American Society of Plant Physiologists, and The Plant Cell Journal :
M. Varbanova, S. Yamaguchi, Y. Yang, K. McKelvey, A. Hanada, R. Borochov, F. Yu, Y. Jikumaru, J. Ross, D. Cortes, C. Je Ma, J. P. Noel, L. Mander, V. Shulaev, Y. Kamiya, S. Rodermel, D. Weiss, E. Pichersky. Methylation of gibberellins by arabidopsis GAMT1 and GAMT2. *The Plant Cell*, 19(1):37–45, 2007.

"Science moves with the spirit of an adventure characterized both by youthful arrogance and by the belief that the truth, once found, would be simple as well as pretty"

James D. Watson

5.1 Introduction

Gibberellins (GAs) form a large group of tetracyclic diterpenoid carboxylic acids. Biologically active GAs, function as natural regulators and control diverse aspects of plant growth and development, including seed germination, stem elongation, leaf expansion, and flower and seed development. Among more than a hundred GAs identified from plants (`http://www.plant-hormones.info/gibberellin_nomenclature.htm`), only a small number of them, such as GA_1 and GA_4, are thought to function as hormones [202]. In general, concentrations of active GAs in plants are highest in areas of rapid cell elongation, such as stem, shoots, and developing seeds [53]. GA-deficient mutants have a dwarf phenotype due to a reduction in the internode lenght, and often are sterile or with high rate of seed abortion. On the contrary, plants that produce excessive amounds of GAs are taller, with a longer hypocotyl, lighter-green leaves, increased stem elongation, earlier flowering, and decreased seed dormancy when compared to wild-type plants [203].

The GA metabolic pathway in plants has been studied for a long time, and a number of genes encoding the metabolic enzymes have been identified. More recently, the availability of genomics tools in model plant species has accelerated the identification of additional genes involved in the GA metabolic pathway [202]. We recently reported that two members of the SABATH methyltransferase gene family in the Arabidopsis genome, GAMT1 and GAMT2, encode enzymes that use S-adenosine-L-methionine (SAM) as a methyl donor to methylate the carboxyl group of GAs, resulting in the methyl esters of GAs (MeGAs). Both genes are expressed most highly in the siliques during seed development, and their mutation results in higher concentrations of bioactive GAs in mature seeds [53]. Until GAMT1 and GAMT2 were characterized, the only known deactivation mechanisms to regulate bioactive GAs levels were 2β-hydroxilation catalyzed by GA 2-oxydases, and epoxidation of the 16,17-double bond of non 13-hydroxylated GAs, catalyzed by a cytochrome P450 monoxigenase [202]. We proposed that methylation of GAs may constitute an additional mechanism for modulating cellular GA concentrations in addition to the oxidative pathways previously described. To further understand the role of GA methylation *in planta* we characterized the effects that modulating the expression of GAMT1 and GAMT2 might have on gene expression and metabolism inArabidopsis.

Transcriptomics and metabolomics approaches are sources of rich information about the biological adaptation of cellular responses, with each approach independently providing a

different and complementary view. Therefore, we employed the complementary functional analyses at the transcriptome and metabolome levels to provide insights into the mechanisms that underline the molecular events associated with the phenotype of GAMT1 and GAMT2 mutants, allowing to understanding the physiological role of GA methylation in develoing siliques and seeds.

5.2 Literature Review

Gibberellins were first discovered in 1912 by a Japanese scientist studing a rice disease characterized by excessive stem elongation named "bakenae" (foolish seedling) [204]. The symptoms of the disease were found to be linked to compounds released from the fungus *Gibberella fujikuroi*. The compound was isolated in 1935 and named after the fungus as "Gibberellin". Futher experiments showed that a variety of GA-like compounds are also present in plants and play key roles in plant growth and development (`http://www.plant-hormones.info/gibberellinhistory.htm`).

The term gibberellins describes a group of tetracyclic diterpenoid compunds that share the *ent*-gibberellane skeleton (Fig 5.1), which is divided into two sub-groups depending on the presence of a carbon at position 19 and 20. Most C_{19}-GAs are biosynthetic intermediates of biologically active C_{20}-GAs. Rather than complicated chemical nomenclature, GAs have been given trivial names based on the order of discovery ($GA_{1,2,3...,n}$).

The wide range of physiological processes regulated by GAs has been defined by the effect of mutations disrupting the GA biosynthesis pathway and by the effect of exogenously applied GAs. Plant physiological responses to GAs can be divided in two major categories. First, GAs promote cell expansion and, perhaps, cell division that lead to elongation and development of tissues. Second, GAs regulates the mobilization of nutritional reserves and, thus, modulates the course of some physiological processes [202, 205].

5.2.1 The Gibberelin Biosynthesis Pathway

GAs are biosynthesized from geranylgenaryl diphosphate (GGDP), a common C_{20} precursor for diterpenoids. In the early, plastid localized portion of the pathway, GGDP is converted to tetracyclic hydrocarbon *ent*-kaurene, through a series of cyclization reactions catalyzed by the terpene synthases *ent*-copalyl diphosphate synthase (CPS) and *ent*-kaurene synthase (KS). *ent*-kaurene is then converted to GA_{12} by two P450s, the enzymes *ent*-kaurene oxidase (KO or CYP701A), located in the outer membrane of the plastid, and *ent*-kaurenoic acid oxidase (KAO or CYP88A), present in the endoplasmic reticulum (Fig 5.1).

The final reactions of the GA biosynthesis pathway in Arabidopsis branch from GA_{12} into two parallel pathways that form either GA_1 or GA_4. Through oxidations on C_{20} and C_3

by GA 20-oxidase (GA20ox) and GA 3-oxidase (GA3ox), respectively, GA_{12} is converted to GA_4, a bioactive form. GA20ox is a multifunctional enzyme catalyzing a series of oxidations from GA_{12} and GA_{53} to GA_9 and GA_{20}. This results in the convertion of GA intermediates from C_{20}-GAs to C_{19}-GAs by loss of a methyl group (Fig 5.1). On the other branch, GA_{12} is substrate for GA13ox for the production of GA_{53}, which is a precursor for GA_1 in the 13-hydroxylated pathway. GA3ox, which participates in both braches of the pathway, performs only a single oxidation at C_3 to form bioactive GA_1 and GA_4 from GA_9 and GA_{20} (Fig 5.1).

5.2.2 Gibberellin catabolism and deactivation

Deactivation is important for effective regulation of the concentrations of bioactive hormones in plants. The definition of "active" versus "inactive" GAs is based on primarily on physiologica assays, but active GAs also share some chemical properties. Bioactive GAs have been identified by their ability to restore growth and devlopment in GA-deficinet mutants after exogenous application. Naturally occuring bioactive GAs are also characterized by a hydroxyl group on C_3, generally added at the final step of biosynthesis (Fig 5.1).

GAs are metabolically deactivated in several different ways. The primary known deactivation reaction is 2β-hydroxylation catalyzed by a class of deoxigenases designated as 2 oxidase enzymes (GA 2-oxidases or GA2oxs). GA2oxs convert the presursors GA_9 and GA_{20}, as well as bioactive GA_1 and GA_4, to inactive GA_{51}, GA_{29}, GA_{34}, and GA_8, respectively (Fig 5.1). Recent work on a recessive tall rice mutant (*eui*), revealed a new GA deactivation mechanism. EUI is a P450 that epoxidizes the 16,17-double bond of non13- hydroxylated GAs, including GA_4, GA_9 , and GA_{12} [79]. More recent work has shown that GAMTs play a role of deactivating GAs in seeds by methylation (Fig 5.2). GAMT1 and GAMT2 utilize a variety of GAs, including bioactive GAs and their precursors, as substrates *in vitro*, and produce the corresponding methyl esters [33, 53].

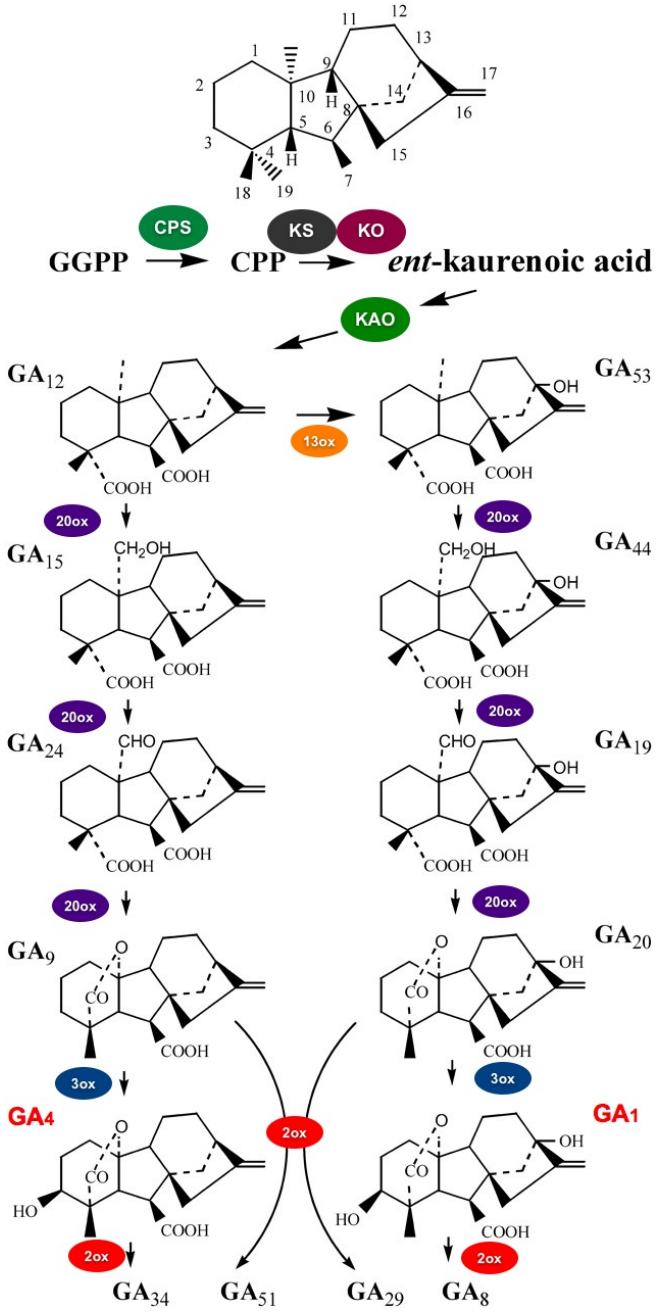


Figure 5.1: The gibberelic acid biosynthetic pathway.
The pathway for biosynthesis of some active GAs, and their catabolic fates, are shown. GA_{34}, GA_{51}, GA_{29}, and GA_8 are catabolic products of GA degradation. Bioactive GA_4 and GA_1, are highlighted in red. The basic structure of *ent*-gibberellane with the numbering system of the carbons is shown at the top. GGPP, geranylgeranyl diphosphate; CPP, ent-copalyl diphosphate; CPS, ent-copalyl diphosphate synthase; KS, ent-kaurene synthase; KO, ent-kaurene oxidase; KAO, ent-kaurenoic acid oxidase; 2ox, GA 2-oxidase; 3ox, GA 3-oxidase; 13ox, GA 13-oxidase; 20ox, GA 20-oxidase. Modified from Varbanova *et al.* Plant Cell 2007;19:32-45.

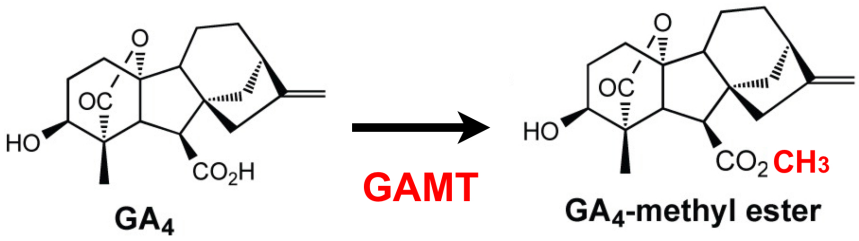


Figure 5.2: Gibberelin A4 deactivated by GAMT through methylation.

5.2.3 Regulation of GA Biosynthesis

Because of the variety of physiological processes that GAs influences, and because of the complex biosynthetic pathway, GAs synthesis is regulated carefully to mantain normal plant development and appropiated response to envirotmental changes. Ligh and GAs, are required for the promotion of germination [206]. GA3oxs in Arabidopsis are induced during the germination by a red light pulse, and the induction is reversed by far red light, suggesting that the active form of the phytochrome may be required for GA3oxs induction [207].

Auxin and GAs play, to some extend, similar roles promoting apical dominance, stem elongation, cell enlargement and division. The control of their biosynthesis and signaling pathways is expected due to the overlapping functions. Strong evidence suggest that auxin regulates GA biosynthesis. In pea, IAA application acts by promoting the down-regualtion of GA2ox, and up-regulation of GA20ox and GA3ox [208]. This suggest that the interaction between GA biosynthesis and auxins may function through different genes.

Additionally, GA biosynthesis pathway shows evidence of self-regulation. Xenobiotic bioactive GAs, repressed GA20ox and GA3ox genes, while GA deficiency results in their up-regulation. In contrast, the expression of GA deactivation genes, GA2ox2 and GA2ox2, is up-regulated upon GA treament [202]. No feedback regulation has been reported on up-stream biosynthetic genes in Arabidopsis, including CPS, KS and KO. It is likely that feedback may operate on selected downstream genes to fine-tune GA levels.

5.3 Methods

The procedures outlined in the section 3.3, were followed to study the gene expression profile of mutants deficient in the expression of GA methyltransferases, with the following specifics:

5.3.1 Plant Material

Wild-type *A. thaliana* L. ecotype Col-0 was used on this study for all experiments. Plants were grown in soil in growth chambers under long day conditions (16 h light at $150uE/8$ h dark cycle) at $22^{o}C \pm 1^{o}C$. Seeds for all knockout and overexpressing lines for the AtSABATH were provided by Dr. Eran Pichesrky (Department of Molecular, Cellular and Developmental Biology at the University of Michigan, Ann Arbor), as part of a collaborative NSF2010 project. Four SALK knockout mutant lines of GAMT1 (SALK_088960 and SALK-047730) and GAMT2 (SALK-109505 and SALK-143728) with Col-0 background were obtained from the Arabidopsis Biological Resource Center [104], and confirmed by the Pichersky laboratory. Additionally, a double mutant line SALK-088960 x SALK-109505, was obtained by the Pichersky laboratory. Overexpressing transgenic plants were obtained for GAMT1 and GAMT2 genes by *Agrobacterium tumefaciens*-mediated floral dip transformation [107] of wild-type plants (Col-0) with cDNA of each gene in sense orientation, driven by the cauliflower mosaic virus 35S promoter.

5.3.2 Gene expression profiling

The procedures outlined in the section 3.3, were followed to study the gene expression profile of mutants deficient in the expression of GA methyltransferases, with the following specifics:

5.3.2.1 RT-PCR and qRT-PCR

For qRT-PCR analysis, total RNA was isolated from different organs of Arabidopsis and from siliques at different stages of development as defined by Bowman [209] using the RNA procedure described by Maes and Messens [210]. The RNA was subjected to DNase treatment using the DNA-free kit (Ambion), and first-strand cDNA was synthesized by AMV reverse transcriptase with poly(T)18 primer in parallel with a negative control reaction in which no AMV reverse transcriptase was added. All samples were brought to 100 μL in volume, after which 1 μL was used in a 20 μL of qPCR reaction containing 20 mM Tris, pH 8.4, 50 mM KCl, 3 mM $MgCl_2$, 5% DMSO, 250 mM deoxynucleotide triphosphate, 200 nM primer, 0.5 units of Taq polymerase B (Promega), 1X CYBR-Green I dye (Molecular Probes), and 10 nM fluorescein (Bio-Rad). GAMT1 and GAMT2 gene-specific primers were designed as follows: GAMT1 forward 5'-TGTTGTTTATGCTGATGGGTGGTC- 3' and GAMT1 reverse 5'-

CGCAATCTCTTCGGTGGTTCTAA- 3'; GAMT2 forward 5'-CGTCCTTCAGGCTCAAG TAGTC- 3' and GAMT2 reverse 5'-CCCTATCTTGAAACCACCACAACGGTC- 3'. Amplification of the ubiquitin gene ubq10 (At4g05320) using the forward primer 5'-AGGAGTCCAC ACTTCACTTGGTC-3' and the reverse primer 5'-GGTGTCAGAGCTCTCTACCTCCA-3' was used as internal control. The qPCR was performed on an iCycler thermocycler (Bio-Rad), with the following conditions: 95°C for 3 min, 50 cycles of 95°Cfor 15 s, 60°C for 30 s, and 72°C for 30 s, followed by a melting cycle of 55 to 95°C with an increasing gradient of 0.5°C, and a 10-s pause at each temperature. All reactions were performed in duplicate. Data processing was done by iCycler real-time detection system software (version 3.0). cDNA of RNA from siliques at half-way through development was serially diluted to generate the standard curve. Analyses of the melting curves were performed to ensure amplification of one specific gene product. At least three sets of independent experiments were performed to calculate a mean value and standard deviation. The cycle threshold (Ct) values generated by the iCycler software for each cDNA sample were used only when they met the following parameters: significant difference in Ct numbers between cDNA sample and the negative control was observed and the melting curve of the PCR product generated by the melting cycles was indicative of specific amplification.

5.3.2.2 RNA Extraction

Based on GAMT1 and GAMT2 promoter-GUS reporter expreiments performed by Varbanova *et al.* and information about the gene expression characteristics of these genes obtained from Genevestigator (`www.genevestigator.com`), total RNA for microarray analysis was extracted from siliques half-way through maturation collected from 45 - 60 days old plants. Instead of the RNA extraction protocol described for RT-PCR and qRT-PCR, to characterized the expression characteristics of GAMT1 and GAMT2, described in the previous section, Ruuska *et al.* [211] protocol for RNA extraction from Arabidopsis seeds was followed to produce high quality RNA free of the carbohydrates and polyphenols, usually found in siliques tissues, which reduced the hybridization efficiency in trial experiments. Following the extraction, DNase treatment was performed following standard procedures Qiagen™ columns. Quality and concentration of the extracted RNA was estimated with 2100 Bioanalyzer from Agilent.

5.3.2.3 Semi-quantitative Reverse Transcription-Polymerase Chain Reaction

Semi-quantitative RT-PCR was performed for selected genes significantly regulated in *gamt1-gamt2* double knockout. cDNA synthesis and specific gene amplification were carried out using Ready-to-Go RT-PCR beads from Amersham. Two micrograms of RNA was reverse transcribed into cDNA in a 25-µL reaction with poly(dT)18 priming using a M-MLV reverse transcriptase (Promega Corporation), and following the manufacture instructions. Following first strand cDNA synthesis the reverse transcriptase was inactivated at 65°C for 10 minutes

and the cDNA diluted with water in a final volume 100 µL. The PCR volume was 25, using PuReTaq PCR Ready-To-Go Beads (GE Healthcare) containing 100 ng of each primer, 2.5 units of PuReTaq DNA polymerase, 10 mM Tris-HCl, 50 mM KCl, 1.5 mM MgCl2, 200 µM of each dNTP, BSA. 2 mM and 2.5 µL of cDNA. A GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA) was used with an initial denaturation step of 94°C for 5 min, followed by 25 cycles of 94°C for 30 s, 54°C for 30 s, 72°C for 60 s, and a final elongation step of 72°C for 10 min. RT-PCR analysis was performed in duplicate using RNA from two different sets of plants. Specific forward and reverse primers for AtSABATH MTs and selected genes found differentially express in GAMT1/GAMT2 double knockout are listed on the appendix (Table A.9). Amplification of the Actin2 (*At3g18780*) and β-tubulin (*At5g12250*) were used as controls.

5.3.3 Metabolite Profiling

The procedures outlined in section 4.3 were followed to study the effect of modulating the gene expression of GAMT1 and GAMT2 on the metabolite profile Arabidopsis, with the following specific procedures for GA measurements carried out at Shinjiro Yamaguchi's laboratory at the RIKEN Plant Science Center in Japan.

5.3.3.1 GA Measurements

Seeds from Arabidopsis Col-0 of the wild type, *gamt1*, *gamt2*, and *gamt1-gamt2* mutants, and overexpressing lines were sown in a Petri dish with MS medium supplemented with 10% sugar. After 2 weeks at 23°C, the seedlings were transplanted to soil and the plants were grown in a growth chamber at 23°C. GA measurements for knockout mutants were conducted by liquid chromatography–selected reaction monitoring using 2H-labeled GAs as internal standards. We used a liquid chromatography tandem mass spectrometry system consisting of a quadrupole/time-offlight tandemmass spectrometer (Q-Tof Premier;Waters) and an Acquity Ultra Performance liquid chromatograph (Waters) equipped with a reverse-phase column (Acquity UPLC BEH-C18; Waters). Approximately 200 mg (dry weight) of lyophilized plant materials were used for each measurement.

Lyophilized plant material (200 mg dry weight) was ground in 10 mL of acetone, and [17,17-$^2H_2$]GAs (500 pg each) were added as internal standards. This mixture was incubated for 12 h at 4°C and then centrifuged at 3000 g for 20 min at 4°C. The supernatant was concentrated to dryness, and dissolved in 2 mL of aqueous acetonitrile (1:1, v/v). The solution was partitioned against an equal volume of n-hexane, and the n-hexane phase was discarded. After the removal of acetonitrile by evaporation in vacuo, the pH was adjusted to 8.0 with 500 mM potassium phosphate buffer. The sample was loaded onto a polyvinylpyrrolidone (PVP, 500 mg; Tokyo Kasei, Japan) cartridge, and eluted with 5 mL of 100 mM potassium phosphate buffer. The pH of this eluate was adjusted to 3.0 with HCl, and loaded onto a

reverse-phase cartridge (Oasis HLB, 60 mg; Waters). After washing with 3 mL of water containing 2% formic acid, GAs were eluted with 6 mL of 80 % acetonitrile containing 1% formic acid. The eluate was dried, dissolved in methanol, and then loaded onto an ion-exchange column (Bond Elut DEA, 500 mg; Varian). GAs were eluted with 6 mL of methanol containing 0.5 % acetic acid after washing with 6 mL of methanol. The eluate was dried, dissolved in CHCl3:ethylacetate=1:1 (v/v) containing 1% acetic acid, and then loaded onto a SepPak silica cartridge (100 mg; Waters). GAs were eluted with 3 mL of chloroform: ethylacetate=1: 1 (v/v) containing 1% acetic acid. The eluted GA-containing fraction was concentrated to dryness, dissolved in 20 μL of water, and then subjected to LC-MS/MS analysis. The LC-MS/MS system consisted of a quadrupole/ time-of-flight tandem mass spectrometer (Q-Tof Premier; Waters) and an Acquity Ultra Performance LC (Waters) equipped with a reverse-phase column (Acquity UPLC BEH-C18; 2.1 x 50 mm, 1.7 μm particle size; Waters). LC separations were performed at a flow rate of 200 μL/min using the following program with solvent A (water) and solvent B (acetonitrile containing 0.05% [v/v]): a linear gradient of B from 3% to 65% over 20 min followed by an isocratic elution with 98% of B for 5 min. The water and acetonitrile used were of LC-MS grade (Kanto Chemical, Japan). GAs that were quantified in this study are listed in the appendix. The collision energy and the sampling cone voltage for individual GAs, was optimized to obtain sensitive detections of specific fragment ions that allow us to confirm the identity of GAs during quantifications. The levels of GAs were determined using a calibration curve ($R2 > 0.997$), which was obtained on each occasion by injecting a series of standard solutions (10 μL) that contained a fixed concentration of [$^2H_2$]GAs (50 pg/μL) and varying concentrations of unlabeled GAs (0.5-100 pg/μL). We used a software tool (MassLynx v. 4.1, Waters) to calculate GA concentrations from the LC/MS/MS data.

5.3.4 Stress Assays

In collabortaion with Dr. Ron Mittler at the Department of Biochemistry, University of Nevada, we performed several stress assays to evaluate the hypothesis that GAs accumulation in the lack-of-function GAMTs mutants have a different stress response than wild-type. For the analysis of stress-tolerance, seeds of wild type, *gamt1* and *gamt2* single mutants, and *gamt1-2* double mutant were surface-sterilized with bleach and placed in rows on 1% agar plates (0.5 × MS medium), containing different concentrations of NaCl [212, 213]. Plates were maintained vertically in a growth chamber (21–22°C, constant light, 100 μmol m−2 s−1) and root length, root growth and % germination were scored at different times after seed plating [213]. Four- or five-day-old seedlings grown on 0.5 × MS agar plates were also subjected to heat (38°C) or cold stress (10°C) for different times, allowed to recover for 24 h, and analyzed [213]. All stress experiments were performed with 3–5 technical replications, each containing 15–30 seeds per line, and repeated at least three times. Statistical analysis was performed as described in [213].

5.4 Results and Discussion

5.4.1 GAMT1 and GAMT1 predominantly expressed in siliques methylate gibberellins

GAMT1 and GAMT2, which we have shown to methylate the carboxyl group of various GAs to form the corresponding MeGA esters, belong to the plant SABATH family of methyltransferases. The Arabidopsis genome has 24 genes belonging to this family [14, 23], and among the proteins encoded by these 24 genes, GAMT1 and GAMT2 are the most similar proteins to each other. The level of divergence between GAMT1 and GAMT2 proteins (58% identity) indicates that the duplication that gave rise to the genes encoding them (At4g26420 and At5g56300) is not very recent.

GAMT1 and GAMT2 are both expressed primarily in developing siliques, with peak transcript levels toward the end of this process. Plants expressing a β-glucuronidase (GUS) reporter gene under the control of the GAMT1 promoter (GAMT1-GUS) and stained for GUS activity showed staining in the developing seeds and in the anthers (Fig 5.3) [53]. GUS staining of transgenic plants with a GAMT2 promoter–GUS fusion gene indicated expression in the anthers but, in contrast with the qRT-PCR results (Fig 5.4, no staining in the developing seeds. qRT-PCR results concerning the expression of GAMT2 are consistent with data obtained from microarray experiments and available online (`www.genevestigator.ethz.ch`) [32]. Because of the mistake in the annotation of GAMT1 in TAIR, and the fact that based on this erroneous annotation all 11 oligonucleotides chosen by Affymetrix to represent GAMT1 on the ATH1 microarray are derived from sequences downstream of GAMT1 that are not part of the GAMT1 gene, there is no relevant information on the expression of GAMT1 on these websites. Nonetheless, qRT-PCR measurements of GAMT1 transcripts showed that this gene has a very similar pattern of expression to GAMT2 throughout the developmental process of the siliques, with little expression seen elsewhere.

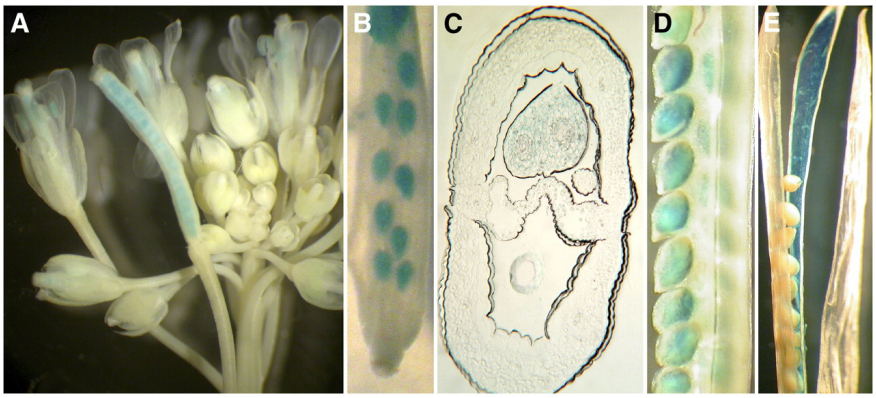


Figure 5.3: Staining for GUS Activity with Plants Transgenic for the GAMT1 Promoter–GUS Transgene.
(A) Inflorescence with flowers and siliques at different stages of development.
(B) A young silique showing staining in the developing seeds.
(C) Cross section of a developing seed.
(D) Maturing seeds.
(E) A mature silique showing staining in the septum and funiculus.
Varbanova *et al*. Plant Cell 2007;19:32-45.

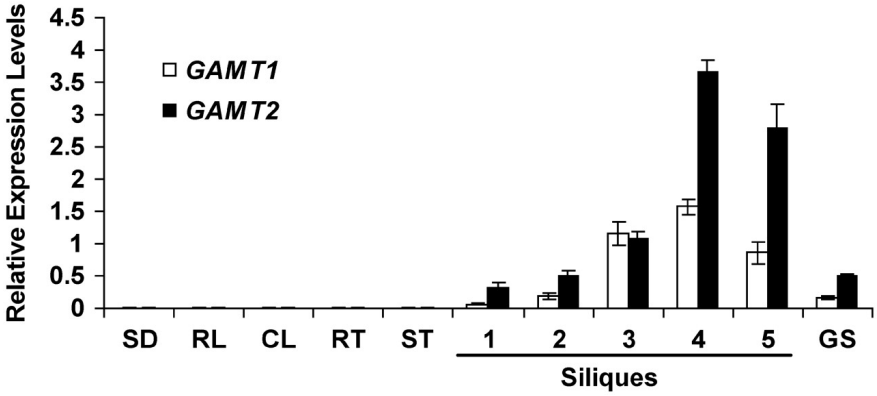


Figure 5.4: Detection of GAMT1 and GAMT2 Transcripts in Different Arabidopsis Organs and Stages of Seed Development by qRT-PCR.
Relative levels of GAMT1 and GAMT2 transcripts (compared with levels of transcripts of the control, the ubi10 gene) are presented as mean values ± SD, calculated from three sets of independent experiments. Silique samples are as follows: 1, flower buds and stage 1 siliques; 2, stages 2 and 3; 3, stages 4 and 5; 4, stages 6, 7, and 8; 5, stages 9 and 10. The 10 stages of the siliques were defined according to Bowman (1994), and they span the development process from the smallest silique (stage 1) to fully mature and dry siliques (stage 10). SD, 2-week-old seedlings; RL, rosette leaf; CL, cauline leaf; RT, root; ST, stem; GS, germinating seeds. Varbanova *et al*. Plant Cell 2007;19:32-45.

We obtained Arabidopsis plants that expressed GAMT1 and GAMT2 under the control of the 35S cauliflower mosaic virus promoter. The majority of independently obtained transgenic Arabidopsis lines (>10) expressing GAMT1 either under the control of the 35S promoter exhibited a similar phenotype that included small, dark-green rosette leaves (Fig 5.6) that grew exceedingly slow, and they developed into dwarfed, bushy plants. Some never bloomed, and those that did bloom started doing so after 3 months (whereas wild-type plants under these conditions began blooming in 6 weeks). Their flowers were small and mostly sterile. Spraying plants with GA4 recovered some fertility and caused partial amelioration of the dwarf phenotype of the whole plant. Spraying with MeGA4 did not ameliorate the dwarf phenotype. Expression of GAMT2 in transgenic Arabidopsis plants under the control of the 35S promoter did not produce a dwarf phenotype at the early stage of plant development when grown on Murashige and Skoog (MS) medium, and seedlings were indistinguishable from wild-type seedlings. However, after transfer to soil on the third week after germination, the majority of transgenic plants (>10 independent lines) showed a different growth pattern from the wild type, growing slowly and exhibiting a higher number of branches compared with wild-type plants, giving them a semidwarf, bushy look (Fig 5.7). While some of these plants had reduced fertility, the majority had normal seed set.

We obtained Arabidopsis Columbia (Col) plants that contained T-DNA insertions in GAMT1 and GAMT2. *gamt1-1* (SALK-047730) had an insertion in exon 3, and *gamt1-2* (SALK-088960) had an insertion in exon 2 that included a small deletion of the coding region. *gamt2-1* (SALK-143728) had an insertion in exon 2, and *gamt2-2* (SALK-109505) had an insertion intron 2 that caused a partial deletion and rearrangement of exon 3 . Lines homozygous for either mutation contained no functional transcript. In addition, we obtained the homozygous double mutant line by crossing *gamt1-2* and *gamt2-2* mutants. None of the single mutant or double mutant plants showed visible phenotypes, such as altered germination rate, growth habit, or seed set, under normal growth conditions. These verified null mutants grew normally and had no visible morphological differences with wild-type plants.

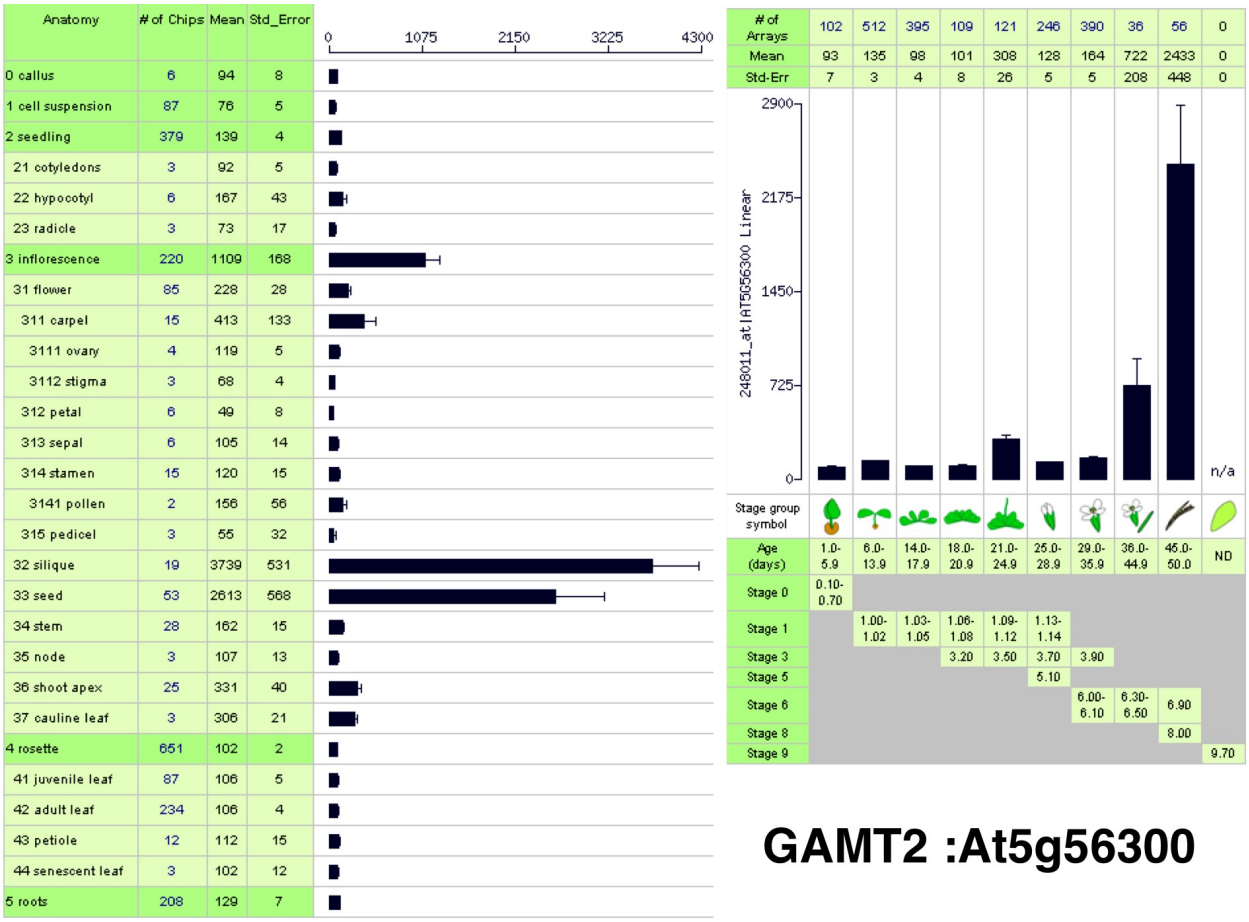


Figure 5.5: Gene expression characteristics of GAMT2 based on data from the microarray repository Genevestigator.

Relative levels GAMT2 transcripts are presented as mean values ± SD or normalized intensity values, calculated from over 1000 independent experiments microarray expeiments on wild-type Col0, in a wide range of plant tissues and organs. The chart on the left shows the gene expression of GAMT2 relative to the plant anatomy. On the left the gene expression is relative to the developmental stage of the Arabidopsis plant according to the nomenclature described by Boyes *et al* (2001).

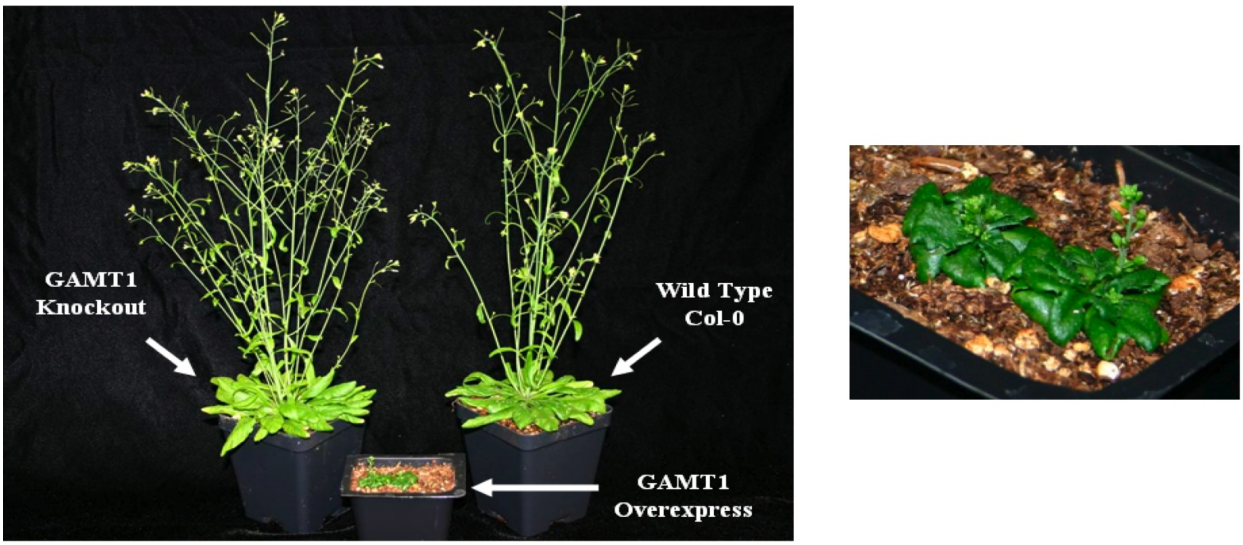


Figure 5.6: Phenotypical effect of modulating the expression of GAMT1 in Arabidopsis Col-0.
Eight-week-old *gamt1* (SALK-088960) mutant plant (left), next to a same-age plant overexpressing GAMT1 (middle) due to 35S-promoter, followed by a wild-type plant (right). On the rightmost a close-up picture of GAMT1 overexpressing line.

Figure 5.7: Phenotypical effect of modulating the expression of GAMT2 in Arabidopsis Col-0.
Eight-week-old *gamt2* (SALK-143728) mutant plant (left), next to a same-age plant overexpressing GAMT2 (middle) due to 35S-promoter, followed by a wild-type plant (right).

5.4.2 Genome-wide Expression Profiling of *gamt1* and *gamt2* mutants.

The expression characteristics of GAMT1 and GAMT2 suggest that the enzymes they encode play a specific role during seed development. To examine the moleculat mechanisms effected by the methylation of GAs in during the development of Arabidopsis seeds, the homozygous single and double mutant seeds were subject to microarray analysis. Shortly after our article our article reporting the discovery of enzymes that deactivate gibberellins through methylation was published, Xing *et al. [214],* reported similar findings for GAMT2. The characterized the gene expression profile of an overexpressing line for GAMT2, but did not consider a GAMT2 knockout. Here we report on the genome expression profiles of GAMT1 single knockout (*gamt1* SALK-088960*),* and GAMT2 single knockout (*gamt2* SALK-143728*).* Furthermore, considering the possibility of function redundancy based on their sequence similarity and promiscuos enzyme activity *in vitro*, the double mutant (*gamt1-gamt2)* was included in the microarray analysis. Principal component analysis (PCA) and hierarchical clustering analysis based on the normalized gene expression data (Fig 5.8), suggest that *gamt2* mutant has a higher effect on gene expression than *gamt1* mutant compared to the wild-type, and that the lack of function for both genes significanly affect the gene expression profile in developing seeds, more than any one of them alone (Table 5.1).

In the table 5.1 is listed the different pathways that were significantly represented among the genes that were regulated in each one of the mutants. Additional information about GO categories enriched with differentially expressed genes is listed in Tables A.8 and A.7.

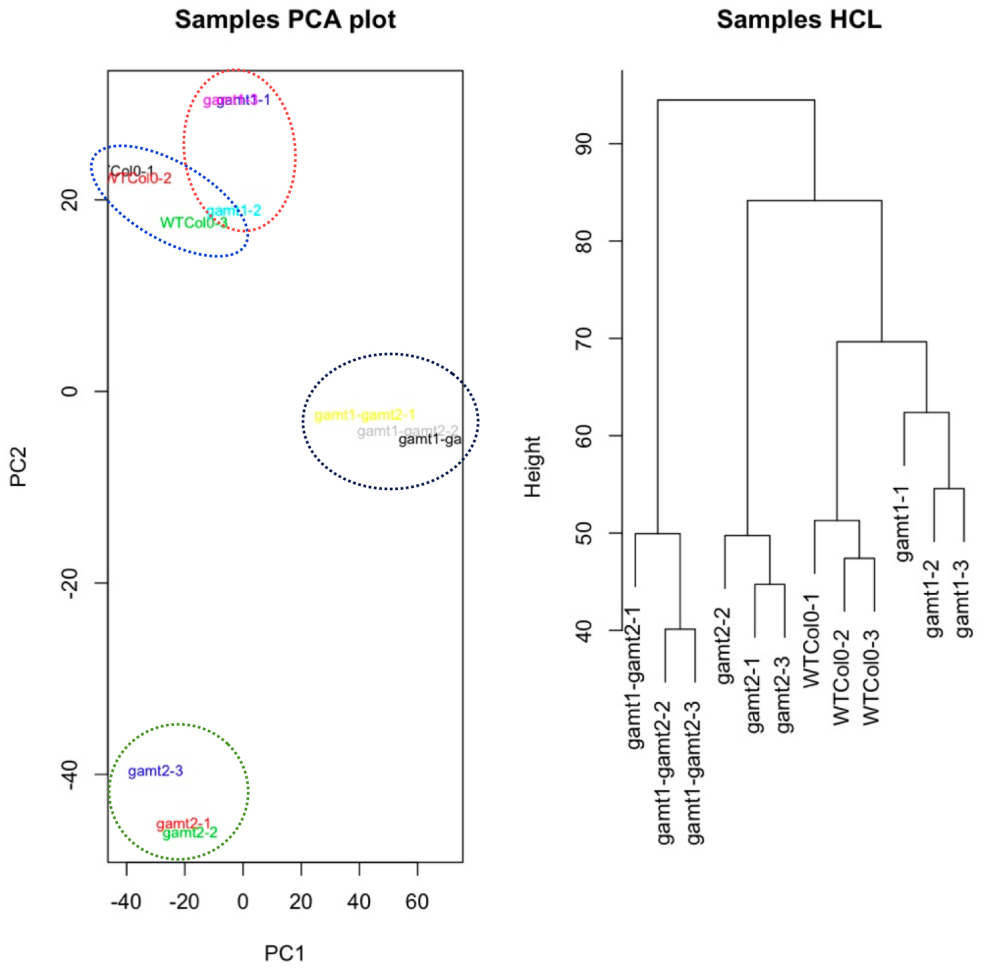


Figure 5.8: Principal component analysis (PCA) and hierarchical clustering analysis (HCA) of wild-type, and *gamt* single and double mutants based on the normalized gene expression microarray data.
Three array replicates for wild-type, *gamt1, gamt2* and gamt1-2, highlighted in blue, red, black and green, respectively.

Table 5.1: KEGG Pathways significantly regulated in *gamt1* and *gamt2,* single mutants, and *gamt1-gamt2* double mutant Arabidopsis plants.
Green and red *p*-values indicate whether the pathway is down-regulated or up-regulated, respectively. In addition to the specific KEGG pathway, we also list the KEGG category in the hierarchy where they are classified into.

KEGG Pathway	***gamt1-2* *p*-value**	***gamt2* *p*-value**	***gamt1* *p*-value**	**Category**
ath00450:Selenoamino acid metabolism ath00271 Methionine metabolism	0.043	>0.05	>0.05	Amino Acid Metabolism
ath00941:Flavonoid biosynthesis	0.025	>0.05	>0.05	Biosynthesis of Secondary Metabolites
ath00196:Photosynthesis - antenna proteins	9.8E-04	>0.05	>0.05	Energy Metabolism
ath00195:Photosynthesis	0.005	>0.05	>0.05	Energy Metabolism
ath00710:Carbon fixation	0.022	>0.05	>0.05	Energy Metabolism
ath00500:Starch and sucrose metabolism	0.002	>0.05	>0.05	Energy Metabolism
ath00901:Indole and ipecac alkaloid biosynthesis	0.011	>0.05	>0.05	Biosynthesis of Secondary Metabolites
ath00940:Phenylpropanoid biosynthesis	0.015	>0.05	>0.05	Biosynthesis of Secondary Metabolites
ath00620:Pyruvate metabolism	0.017	>0.05	>0.05	Carbohydrate Metabolism
ath00910:Nitrogen metabolism	0.003	>0.05	>0.05	Energy Metabolism
ath04120:Ubiquitin mediated proteolysis	0.002	>0.05	>0.05	Folding, Sorting and Degradation
ath00361:gamma-Hexachlorocyclohexane degradation	0.034	>0.05	>0.05	Xenobiotics Metabolism
ath00626:Naphthalene and anthracene degradation	0.038	>0.05	>0.05	Xenobiotics Metabolism
ath00643:Styrene degradation	0.041	>0.05	>0.05	Xenobiotics Metabolism
ath00290:Valine, leucine and isoleucine biosynthesis	0.012	0.046	>0.05	Amino Acid Metabolism
ath00920:Sulfur metabolism	0.036	0.037	>0.05	Energy Metabolism
ath00260:Glycine, serine and threonine metabolism	>0.05	0.037	>0.05	Amino Acid Metabolism
ath00251:Glutamate metabolism	>0.05	0.034	>0.05	Amino Acid Metabolism
ath00272:Cysteine metabolism	>0.05	0.040	>0.05	Amino Acid Metabolism
ath03050:Proteasome	>0.05	0.019	>0.05	Folding, Sorting and Degradation
ath03010:Ribosome	0.016	9E-07	>0.05	Protein Biosynthesis

Gibberellin biosynthesis genes are regulated in *gamt1-gamt2*

The level of bioactive GAs in plants are maintained via feedback and feedforward regulation of GA metabolism [202]. Convertion of GAs into MeGAs in plants may lead to changes in the availavility of bioactive GAs and thereafter expression changes of GA biosynthetic genes and downstream GA-responsive genes. Analysis of regulated gees in *gamt1-gamt2* showed that the expression of some important genes involved in the GA biosynthesis, such as KAO2 (*ent*-kaurenoic acid hydroxylase 2), AtGA2ox2 (gibberellin 2-oxidase 2), and AtGA2ox1 (gibberellin 2-oxidase 2), were up-regulated, while AtGA20ox3 (gibberellin 20-oxidase 3), AtGA20ox2 (gibberellin 20-oxidase 2), and AtGA3ox (gibberellin 3-β-dioxygenase GA4), were down-regulated (Fig 5.9, Table 5.2).

Self-regulation of the gibberellins biosynthestic pathway is well documented [215]. We suggest that GA20ox activity in Arabidopsis seeds is down-regulated due to the high amounts of bioactive gibberellins, like GA_1 and GA_4. It has been proposed that GA20ox is a primary target for feedback regulation. Support for feedback regulation of GA20ox activity was provided by work with GA-biosynthesis mutants. For example, the level of mRNA for the stem-specific GA20ox is higher in *ga5* (GA-deficient mutant), is higher than in wild-type. Treatment of *ga5*, and to a lesser extend, wild-type with GA_4, caused a reduction of GA20ox transcrip levels. Strong down-regulation of GA20ox transcrip levels has also been observed in pea and rice. On the contrary, low endogenous GA levels after treatment with GA biosynthesis inhibitors, resulted in increased GA20ox mRNA levels [215].

Like GA20ox, GA3ox is also subject to feedback regulation. Gain-of-funtion mutations in the repressors or loss-of-function mutations in the positive components of GA response pathway results often in higher levels of bioactive GAs and up-regulation of both GA20ox and GA3ox gene expression [216]. To the best of our knowledge, there is not experimental data about the regulation of GA2ox on GA biosynthetic mutants. However, it has been predicted that changes in GA2ox mRNA levels would be opposite to those of GA20ox and GA3ox genes because GA2ox are feedforward regulated by the amount of bioactive GAs [217]. As mentioned abovewe found that in Arabidopsis GA2ox2 and GA2ox1 are regulated in an opposite fashion compared to GA20ox2 and GA3ox, when the levels of bioactive GAs increase in *gamt1-gamt2* double mutant, due to the lack of GAMT methylating activity. These further indicate that the methylation of GAs by GAMT1 and GAMT2 leads to inactivation of GAs.

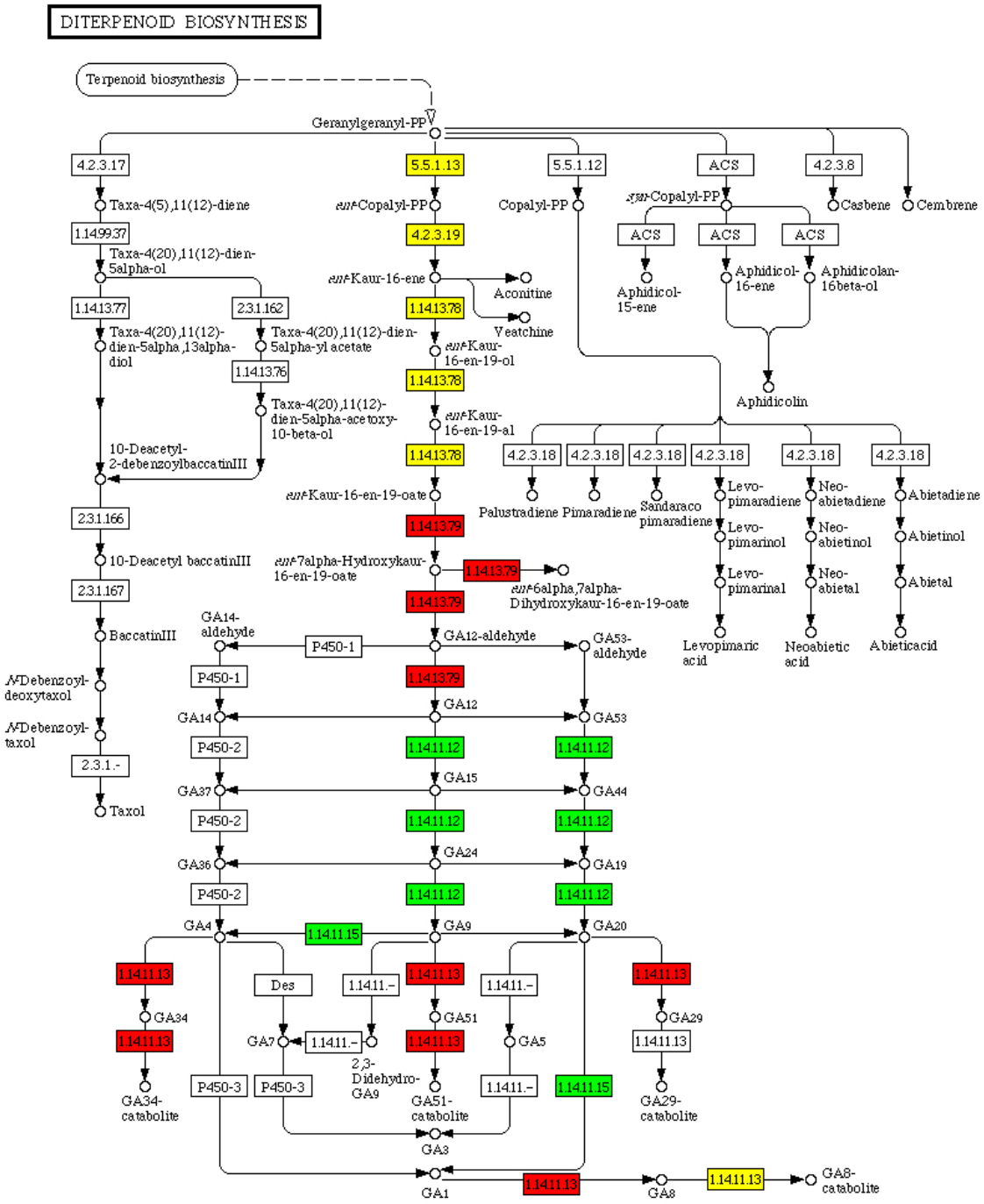


Figure 5.9: Regulation of genes in the diterpenoid biosynthesis pathway for *gamt1-gamt2* double mutant.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green and red indicate the genes that were significantly repressed or induced, respectively, in *gamt1 gamt2* double mutant. Yellow indicates genes that did not change significantly.

Table 5.2: Differentially expressed genes in the gibberellins biosynthesis pathway in *gamt1-gamt2* double mutant Arabidopsis plants.
Locus identifier in green and red indicate whether the gene is up or down-regulated, respectively.

Locus Identyfier	Annotation
AT2G32440	KAO2 (ENT-KAURENOIC ACID HYDROXYLASE 2)
AT1G30040	AtGA2ox2; gibberellin 2-beta-dioxygenase
AT1G78440	AtGA2ox1 (GIBBERELLIN 2-OXIDASE 1)
AT5G07200	AtGA20ox3 YAP169 (GIBBERELLIN 20 OXIDASE 3)
AT5G51810	AtGA20ox2 (GIBBERELLIN 20 OXIDASE 2);
AT1G15550	AtGA3ox GA4 (GA REQUIRING 4); gibberellin 3-beta-dioxygenase

Table 5.3: Differentially expressed genes in the methione biosynthesis pathway in *gamt1-gamt2* double mutant Arabidopsis plants.
Locus identifier in green and red indicate whether the gene is up or down-regulated, respectively.

Locus Identyfier	Annotation
AT3G01120	MTO1 (METHIONINE OVERACCUMULATION 1)
AT3G22740	HMT3 (Homocysteine S-methyltransferase 3)
AT5G15950	AT5G15950, adenosylmethionine decarboxylase family protein
AT1G12010	1-aminocyclopropane-1-carboxylate oxidase, putative
AT2G19590	ACO1 (ACC OXIDASE 1); 1-aminocyclopropane-1-carboxylate oxidase
AT2G36880	MAT3 (METHIONINE ADENOSYLTRANSFERASE 3)
AT3G23810	SAHH2 (S-ADENOSYL-L-HOMOCYSTEINE (SAH) HYDROLASE 2)
AT5G49160	MET1 (DECREASED METHYLATION 2DNA)

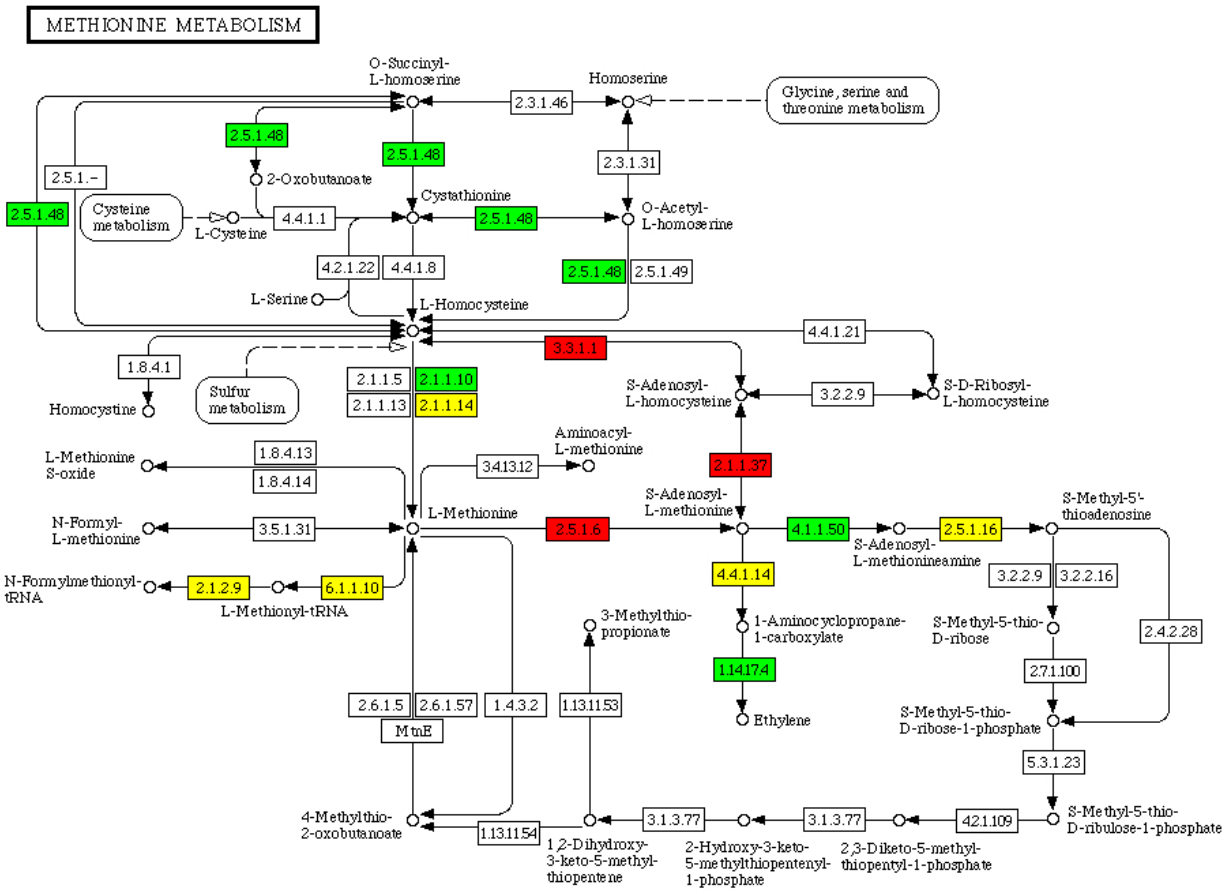


Figure 5.10: Regulation of genes in the methionine metabolism pathway for *gamt1-gamt2* double mutant.
Colored boxes indicate the genes identified in the Arabidopsis genome. Green and red indicate the genes that were significantly repressed or induced, respectively, in *gamt1 gamt2* double mutant.

Methionine biosynthesis and GAs methylation

Similar to our results described in section 3.4.1, where genes in the methinone biosynthesis pathway were regulated in response to JMT overexpression, we found that the same genes were regulated in *gamt1-gamt2* mutant, but the changes were in the opossite direction. Genes down-regulated by overexpression of JMT, are down-regulated in the loss of function mutant *gamt1-gamt2*. S-adenosyl-L-methionine, the methyl donor for reactions catalyzed by SABATH MTs, is synthesized in the methionine biosynthesis pathway.

The genes encoding enzymes upstream of S-adenosyl-L-methionine, like MAT3 (At2g36880 methionine adenosyltransferase 3), and SAHH2 (At3g23810 S-adenosyl-l-homocysteine (SAH)

hydrolase 2), were significantly up-regulated. While genes downstream S-adenosyl-L-methionine are down regulated (Fig 5.10, Table 5.3). Suggesting that in develping seeds, genes in the methinone biosynthesis pathway are regulated by bioactive GAs by a feedback loop mechanism. When GAs have achived their role in seed development they are inactivated by methylation. Methionine biosyntheic genes responde to high levels of GAs by incresing the pool of S-adenosyl-L-methionine, which is in turn use by GAMTs to inactivate GAs. Once GAs are inactivated by methylation, the mRNA levels of S-adenosyl-L-methionine biosynthetic genes is reduced.

Energy metabolism regulated by gibberellin biosynthesis in seeds

Plants have the capacity to accumulate large amounts of carbon and nitrogen in specific cells and tissues and to mobilize these materials for use in other parts of the plant. This capacity is used during seed formation when nutrients are moved from the vegetative plant to developing seeds, and during seed germination when carbon and nitrogen are mobilized for seedling development.

Several pathways related with energy metabolism were significantly up-regulated in *gamt1-gamt2* double mutant (Table 5.1, Fig 5.12, 5.11). This might result from the stimulus of bioactive GAs at high levels. The effect of gibberellins in enhancing the growth rate in plants by stimulating photosynthesis and the energy metabolism has been target of research since the early 1960's [218]. A major effect of gibberellins on growning seedlings is often an increased dry weight. This observation has frecuently led to suggestions that this particular gibberellin effect may be related to an increased photosynthetic carbon-fixation, but the mechanisms by which the photosynthestic rate was increased have not been elucidated completely. Recently, the molecular mechanisms underlying the effect of GAs on photosynthetic carbon-fixation were studied at the moleculat level [219]. Yuan *et al.,* stablish that application of GA3 to broad bean and soybean leaves increased the photosynthesic rate by more than 20% in intact leaves and protoplasts. Transgenic tobacco plants with either Arabidopsis GA20ox or GA2ox under the control of 35S cauliflower mosaic virus promoter, have high and low levels of bioactive GAs, respectivelly. In accordance with the effect of GAs on plant growth, biomass accumulation and photosynthetic rate was increased or decreased in AtGA20ox and AtGA2ox overexpressing plants, respectively [220].

As mentioned above, the GA biosynthetic genes, GA20ox and GA2ox were significantly up and down regulated due to the self-regulation mechanisms of GA biosynthesis in response to high levels of bioactive GAs. Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) and RuBPCase (ribulose-1,5-bisphosphate carboxylase), key enzymes controling carbon fixation in plants were significantly up-regulated in response to the incresed levels of bioactive GAs in *gamt1-gamt2* mutant. It is believe that the quantity of activated RuBPCase is closely related to the rate of photosynthetic carbon assimilation. Additionally, sereval other genes encoding proteins of the light-harvestion complex, photosynthetic electron transport, photosystem

I and II, all of them involved directly in photosynthetic carbon fixation processes, were significantly up-regulated. Based on our experimental result it is suggested that the up-regulation of energy metabolism mediated by photosynthetic carbon fixation, is due to the high levels of bioactive GAs, which may increase the photosynthetic rate.

Flavonoid biosynthesis pathway and GAs

Flavonoid biosynthesis is regulated by a complex interaction between internal and external stimuli such temperature, light, carbohydrates, water stress and hormones [221]. The role of plant hormones on anthocyanin accumulation is intriguing, but the literature on this subject is still controversial [221]. We found that ten genes in the flavonoid biosynthesis pathway are significantly up-regulated in siliques of the loss-of-funtion *gamt1-gamt2* mutant, possibly due to the accumulation of bioactive GAs (Table 5.1, Fig 5.13). The influence of GAs on anthocyanin accumulation is not fully understood. While other hormones, like ethylene appears to have a univocally possitive effect on anthocyanin accumulation, the results for GAs are contradictory. GAs appear to have positive role in petunia flowers, while playing an inhibitory role in other plant tissues [221]. To our knowledge, the only report connecting anthocyanin accumulation and GAs showed that low phosphate-grown Arabidopsis have low levels of bioactive GAs which contributes to anthocyanin accumulation [222]. In our case is possible that the up-regulation of flavonoid biosynthetic genes is not a direct consequence of high levels of bioactive GAs, but instead a sugar-mediated process. We found that several pathways related to energy metabolism were significantly up-regulated. Particularly, the starch and sucrose metabolism was up-regulated, which may be related with a sugar-dependent gibberellin-induction of the expression of flavonoid and anthocyanin biosynthetic genes [221, 223]. Additionally, seeds have energy reserves in the form of starch, which is broken-down to simple sugars, a process accelerated by GAs through the induction of hydrolytic enzymes [75].

In petunia (*Petunia hybrida*) corolas, the induction of anthocyanin synthesis and anthocyanin biosynthetic gene expression is a process that requires both sugar and GAs, none of them alone can induce this processes [221]. Moalem-Beno *et al.* [223], showed that only metabolic sugars, fructose, trehalose, sucrose galactose and glucose, promoted GAs-induced anthocyanin accumulation. Sugars are essential as a general source of carbohydrates for carbom metabolism, upon which the induction of pigmentation is dependent [221]. Sugars may play a similar role in Arabidopsis seeds. High levels of bioactive GAs increase the transcription of genes encoding hydrolytic enzymes, like α-amylase, consequently inducing starch degradation and high levels of sugars, like sucrose, which represent a rich pool of carbon readily available for anthocyanin and flavonoids biosynthesis. Furthermore, studies in Arabidopsis suggest the existence of a crosstalk between the sucrose and hormone signalling pathways in the regulation of the anthocyanin biosynthetic pathway [221]. We speculate that in Arabidopsis seeds, flavonoid biosynthesis up-regulation is a sugar-mediated process, induced by gibberellins.

Starch Hydrolysis and Gibberelin Biosynthesis

Genes in the starch and sucrose metabolism were significantly up-regulated in siliques of *gamt1-gamt2* double mutant (Table 5.1, Fig 5.14). It is well established, that during germination of cereals seeds, gibberellins are synthesized in the embryo after imbibition, and induce hydrolitic enzymes such as α-amylase, which are involved in the degradation of starch to less complex carbohydrates, like sucrose and glucose. However, only recently the role of GA-induced starch degradation was investigated in Arabidopsis seed development [75]. Kim *et al.,* showed that genes encoding GA biosynthetic enzymes and α-amylases expressed almost simultaneusly around starch granules in the outer integument, preceding the disapperance of those granules. GA20ox2, GA20ox3, and GA3ox4, spacially overlapped with the α-amylase gene, *amy3*. Additionally, α-amylase transcrips were not observed in a GA-deficient mutant where GA3ox4 was knockout. Further indication that starch degradation is a GA-regulated process. As a possible physiological role for starch degradation induced by GAs in the outter integument, it has been proposed that timely supply of the starch hydrozylate to the surrounding tissues of the seed surface is essential for the structural formation of the seed coat, which protects the embryo from mechanical damage and pathogen attack [75]. Nevertheless, the relationship between the degradation of starch granules and development of the seed coat needs to be examine further.

Interestingly, the expression peak of GAMT1 and GAMT2 is in the later stages of seed development, when the seed has matured and the seed coat is fully developed. This suggest that once the developmental processes regulated by GAs have been completed, including the seed coat formation through starch degradation, the GA-reponse in turned down by irreversible inactivation of bioactive GAs by methylation. In the *gamt1-gamt2* mutant this inactivation mechanism is not available, therefore the starch degradation results in the accumulation of sucrose, glucose and other sugars, which might trigger the flavonoid biosytetic genes as discussed in the previous section.

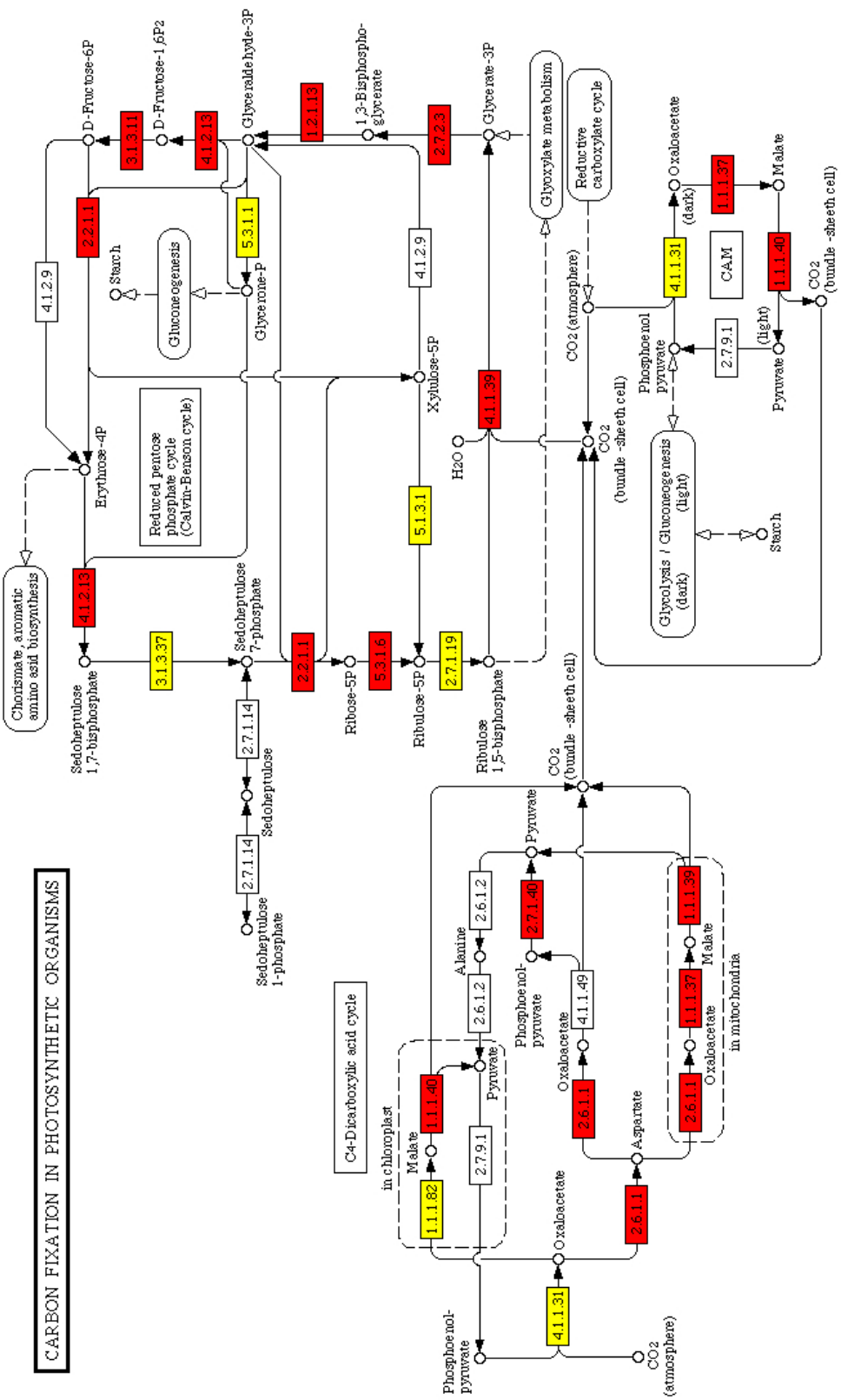


Figure 5.11: Carbon fixation pathway from KEGG showing up-regulated genes in *gamt1-gamt2* double mutant. Colored boxes indicate the genes identified in the Arabidopsis genome. Red indicate the genes that were significantly induced in *gamt1-gamt2* double mutant. Yellow indicates genes that did not change significantly.

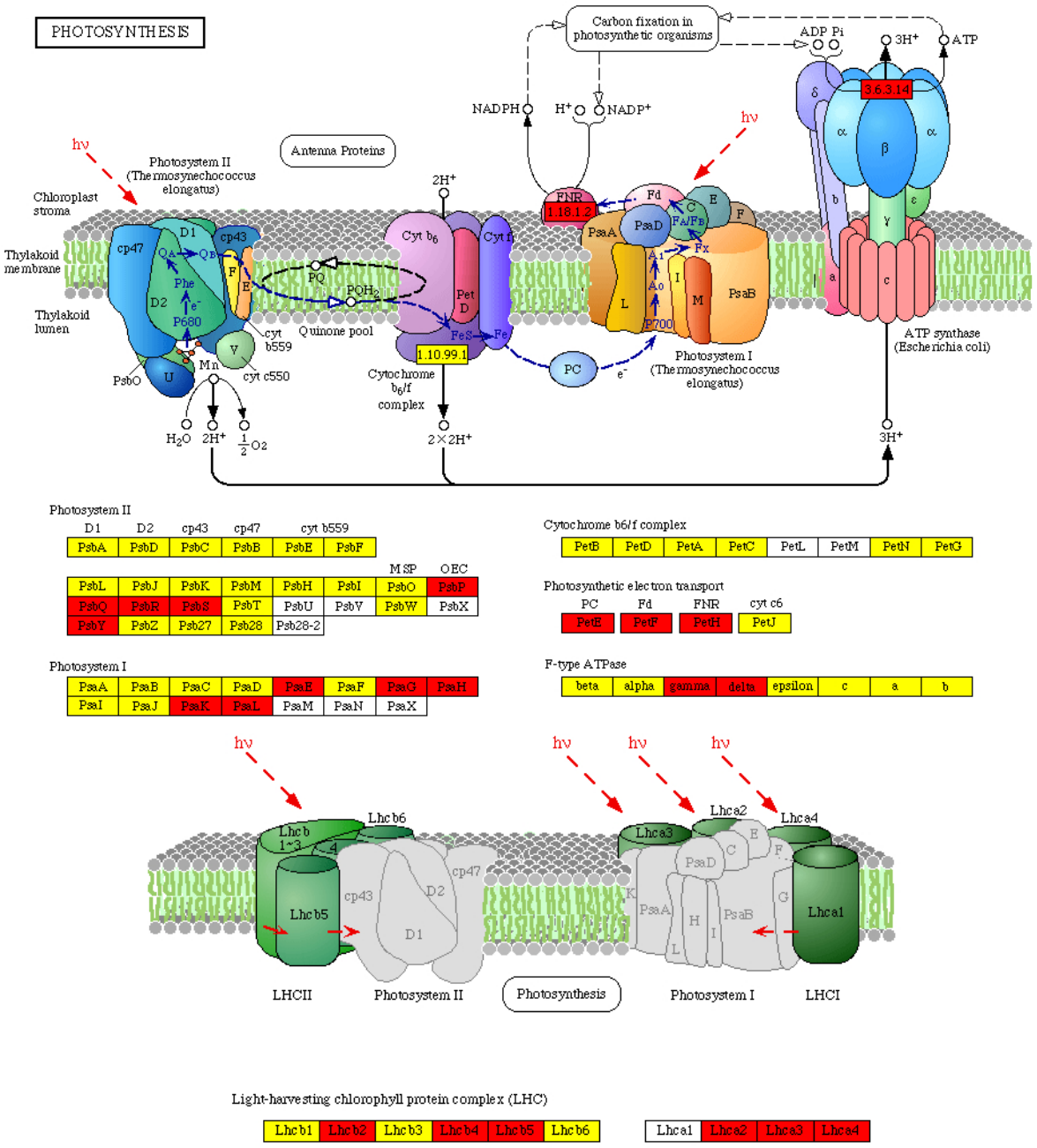


Figure 5.12: Photosynthesis antenna complex from KEGG showing up-regulated genes in *gamt1-gamt2* double mutant.

Colored boxes indicate the genes identified in the Arabidopsis genome. Red indicate the genes that were significantly induced in *gamt1-gamt2* double mutant. Yellow indicates genes that did not change significantly.

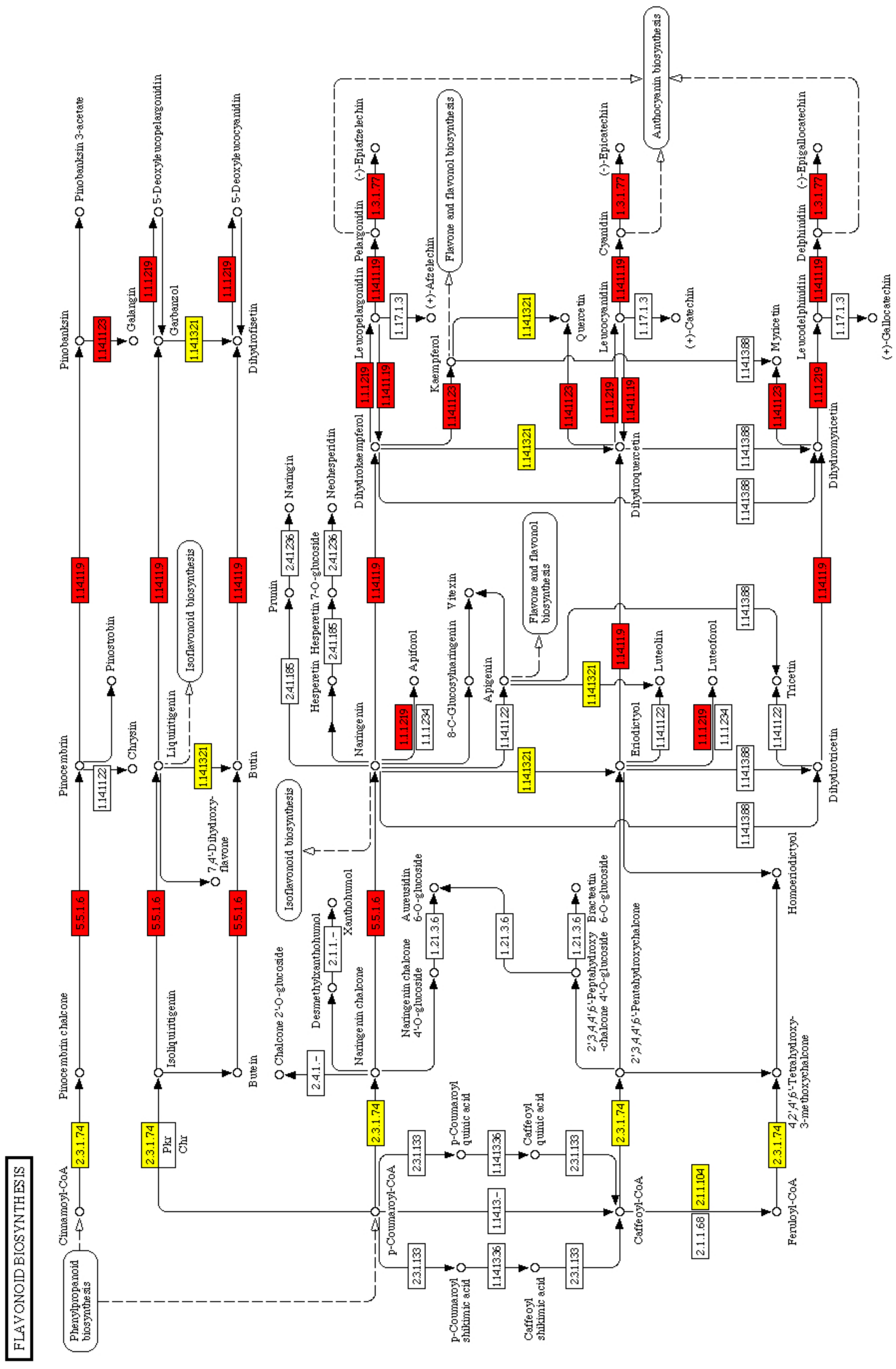


Figure 5.13: Flavonoid biosynthesis pathway from KEGG showing up-regulated genes in *gamt1-gamt2* double mutant. Colored boxes indicate the genes identified in the *Arabidopsis* genome. Red indicate the genes that were significantly induced in *gamt1-gamt2* double mutant. Yellow indicates genes that did not change significantly.

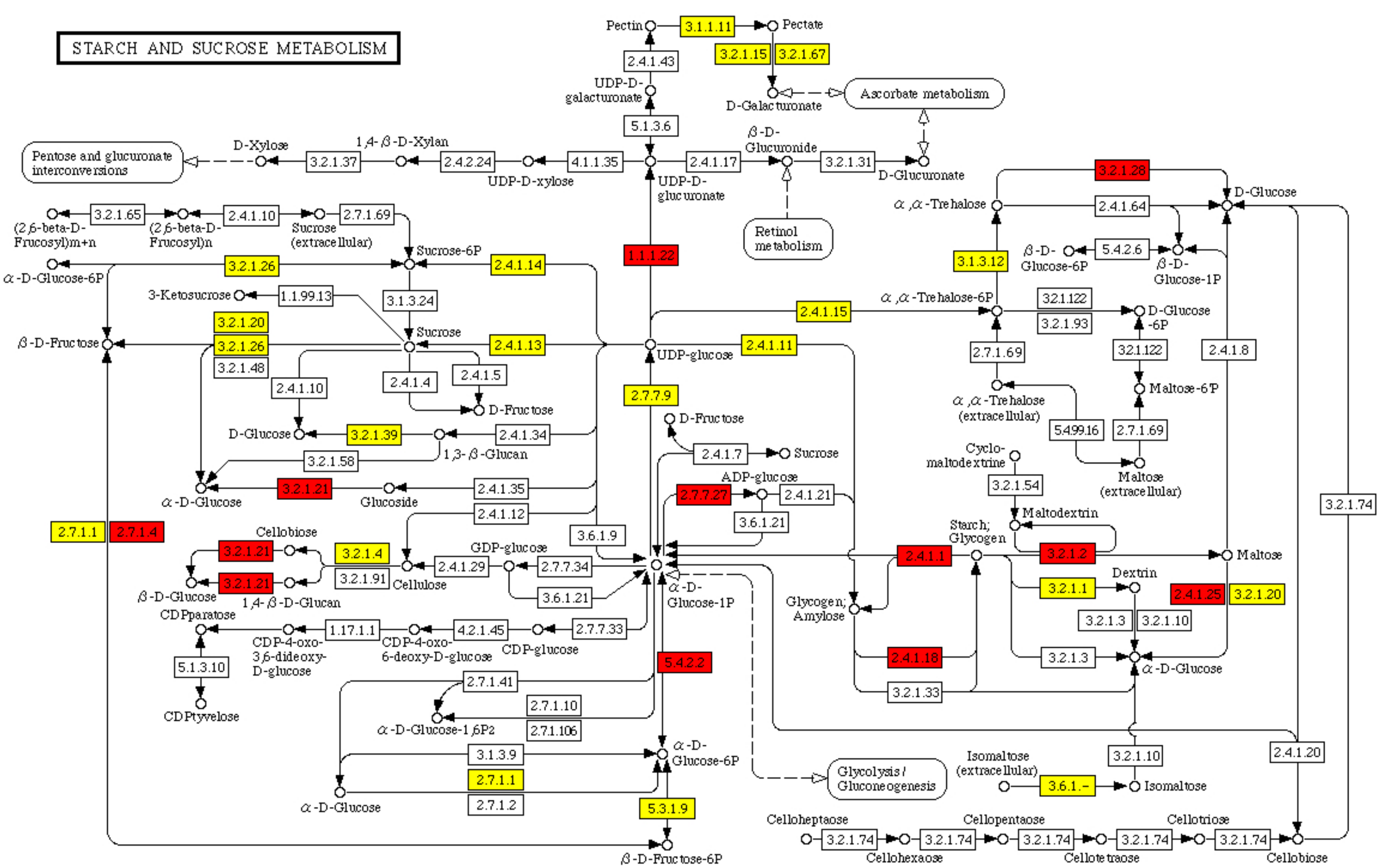


Figure 5.14: Starch and sucrose metabolism from KEGG showing up-regulated genes in *gamt1-gamt2* double mutant. Colored boxes indicate the genes identified in the *Arabidopsis* genome. Red indicate the genes that were significantly induced in *gamt1-gamt2* double mutant. Yellow indicates genes that did not change significantly.

5.4.3 Semiquantitative RT-PCR analysis of GA-responsive genes

In addition to the gene expression information obtained through microarray analysis, a selected set of 17 genes differentially expressed in *gamt1-gamt2* double mutant were selected for further analysis by semiquantitative RT-PCR. Based on the current annotation of this genes they are expressed in response to GAs stimulus. As expected GAMT1 and GAMT2 transcripts were not detect on *gamt1-gamt2* mutant. Based on the RT-PCR results, significant differences in transcrips abundance were not observed for all the genes, as expected. GA2ox1, GA2ox2, GASA1 and GASA2 were up-regulated in *gamt1-gamt2* according to the microarray analysis, and accordingly, the transcripts levels detected by RT-PCR were higher in *gamt1-gamt2* than in wild-type. For the other genes is difficult to conclude whether the differences are significant. To further validate the magnitude of regulation determined by microarray, quantitative RT-PCR should be carried out and RT-PCR to verify the validity of selected regulated genes.

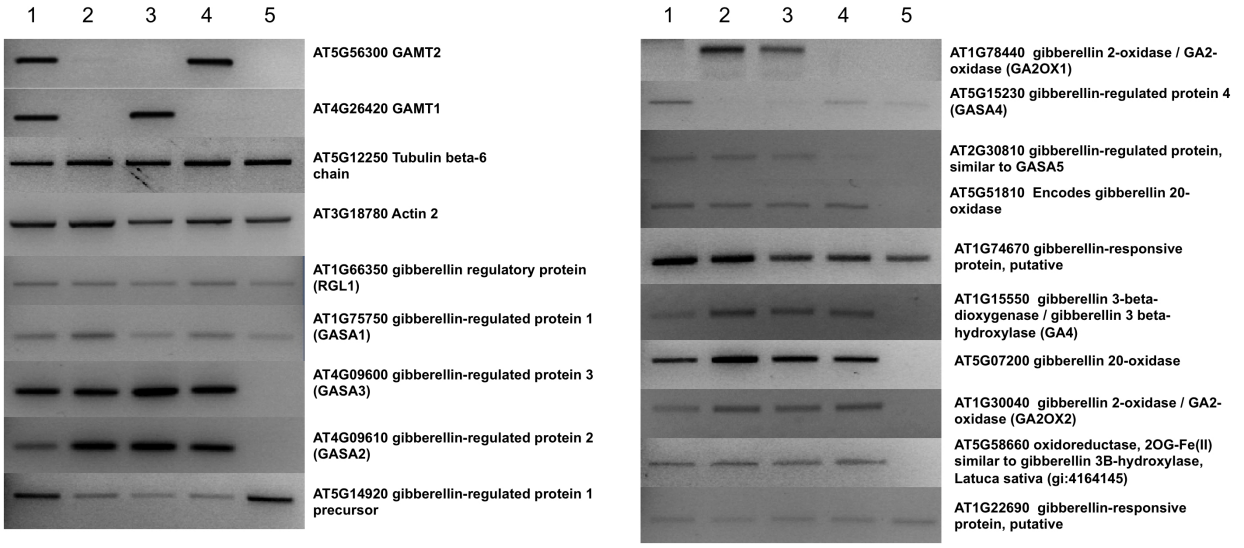


Figure 5.15: RT-PCR analysis of transcript levels of individual gibberellin related genes. The expression of 17 gibberellin responsive genes was examined by RT-PCR. Lines 1-4 correspond to cDNA obtained from developing siliques of (1) wild-type Col-0, (2) *gamt1-gamt2* double knockout, (3) *gamt2* and (4) *gamt1*, while line 5 correspond to cDNA obtained from matures wild-type leaves. RT-PCR with primers for β-tubulin and actin-2 was used to judge equality of the concentration of the cDNA templates used for each reaction.

5.4.4 Metabolic profiling of *gamt1* and *gamt2* mutants

Siliques of Mutant Plants Have Higher Levels of GAs than Wild-Type Plants

We measured the content of some GAs in the siliques of homozygous *gamt1-2* (SALK-088960), *gamt2-2* (SALK-109505), and *gamt1-2 x gamt2-2* plants using siliques of developmental stages 5, 6, and 7, where GAMT1 and GAMT2 expression levels are close to or are at their highest levels. The levels of bioactive GA4 and GA1 in siliques were significantly higher in the *gamt1-gamt2* double mutant than those in wild-type plants, with the level of GA_1 especially elevated in the double mutant (~10-fold) compared with that in the wild type (Figure 5.16). GA_4 and GA_1 levels were slightly more abundant in *gamt1* and *gamt2* single mutants than in wild-type plants as well (Figure 5.16). The effects of *gamt* single and/or double knockout were observed on the levels of the deactivated forms GA34 and GA8 and on some precursor GAs in the *gamt1* plants. These results support the conclusion that GAs are endogenous substrates for GAMT1 and GAMT2 in wild-type siliques.

An even stronger indication that GAs are the endogenous substrates of GAMT1 and GAMT2 is the observation that in *gamt1*, *gamt2*, and *gamt1-2* mutant lines, the levels of several GAs in the siliques, and in particular, the levels of the active GA_1 and GA_4, are higher than in the siliques of wild-type plants. Here, a general correlation is seen between increases in levels of individual GAs and the relative activity levels of GAMT1 and GAMT2 with these compounds. For example, GAMT1 has strong *in vitro* activity with GA_{20}, and *gamt1* shows increased levels of GA_{20} in siliques, while GAMT2 has weak activity with GA_{20} *in vitro,* and the *gamt2* mutant shows little or no differences in GA_{20} levels in siliques. In some cases, when both enzymes appear to have similar activity with a given GA, for example GA_1, there is a small increase in the level of this GA in either *gamt1* or *gamt2*, while the increased levels are higher in the *gamt1-2* double mutant.

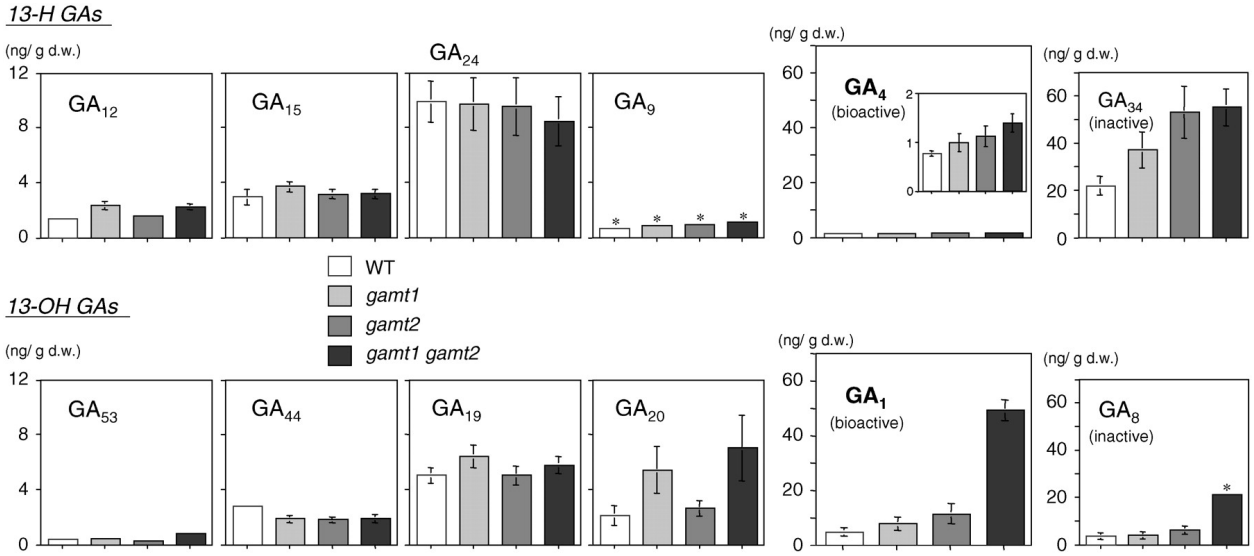


Figure 5.16: GA Levels in Developing Siliques of *gamt* Single and Double Mutants. Shown are endogenous levels of GAs in developing siliques of the wild type, *gamt1-1* and *gamt2-2* single mutants, and the *gamt1-2 gamt2-2* double mutant plants. Results of respective GAs are displayed according to their order in the biosynthesis pathway (see Figure 5.1). Results of GA_4 are also indicated in the inset with a different y axis scale to clarify the differences among the genotypes. Note that the y axis scale for precursor GAs (GA_{12}, GA_{15}, GA_{24}, GA_9, GA_{53}, GA_{44}, GA_{19}, and GA_{20}) is different from that for bioactive (GA_4 and GA_1) and deactivated (GA_{34} and GA_8) forms. GA measurements were conducted three times using independently prepared plant extracts. Means with SE (which are displayed only when they are >0.2) from triplicates are shown. Asterisks indicate that these GAs could be measured reliably only once due to comigration of impurities on the liquid chromatograph. d.w., dry weight. Varbanova *et al.* Plant Cell 2007;19:32-45.

Metabolic Fingerprints and Discriminant Analysis

To examine the most discriminant components associated with the metabolite fingerprint of Arabidopsis wild-type, *gamt1, gamt2,* and *gamt1-2,* we performed gas chromatography-mass spectrometry (GC/MS) analysis of polar compounds extracted from siliques of the different mutants and wild-type and applied GA-DFA analysis on the resulting data. For this analysis we used the same batch of tissue that was used for microarray analysis. Visual inspection of the chomatograms generated by GC/MS, clearly show that there are differences between the wild-type and the mutants (Figure 5.19) In general the relative abundance of several peaks is changing when the expression of GAMT1 and GAMT2 is repressed in the knockout mutants. However, with approximately 150 different components detected in polar extracts of Arabidopsis siliques, it would be necessary to devonvolute each one of them and create a processing method where they are accounted for. Instead of performing this tedious task, we used GA-DFA analysis to distinguish between the mass spectra of wild-type and the mutants profiles, to identify those m/z ratios that are discriminant (significantly different in abundance) among the different sample classes. The selection frequency of discriminant m/z ratios is summarized in Figure 5.17. Using the top m/z ratios to formulate a DFA model resulted in a clear separation among wild-type, single mutants in one cluster, and double mutant in another cluster (Figure 5.18). We were able to observed a discrimination between the wild-type and the mutants with the top variables selected by GA-DFA. However, these m/z ratios are not able to separate between *gamt1* and *gamt2* in the DF model generated. This observation is consistent with the microarray results were a small number of differentially expressed genes was obtained when *gamt* and *gamt2* were compared. These results provide evidence supporting the hypothesis that GAMT1 and GAMT2 may be redundant in their catalytic activity, and consequently, the lack of one protein is compensated by the other. Figure 5.18, represents the discrimination among the wild-type and mutants based on a DF model formulated with a subset of 4 top m/z ratios (62, 138, 332 and 256), indicating that these m/z ratios alone contained enough information to differentiate among wild-type and mutants. The DF model generated with these m/z ratios have a classification accuracy of 70%, mostly due to the misclassification of *gamt1* and *gamt2* mutants (Figure 5.18). We believe that the classification obtained are reasonable considering the similarities previously observed for *gamt1* and *gamt2*, as it was stated earlier that the misclassification may be due to the significant similarities between these two mutants. An improved classification model could be obtained if *gamt1* and *gamt2* are assigned to the same sample class.

The top m/z ratios selected by GA-DFA as the most discriminant were used to identify the specific metabolites in the GC/MS chromatograms. Consequently, rather than developing a data processing method for all the components present, we focus on few components where the discriminant m/z ratios are present. A list of the metabolites identified with the discriminant m/z ratios, and their relative amount changes are listed in Table 5.4.

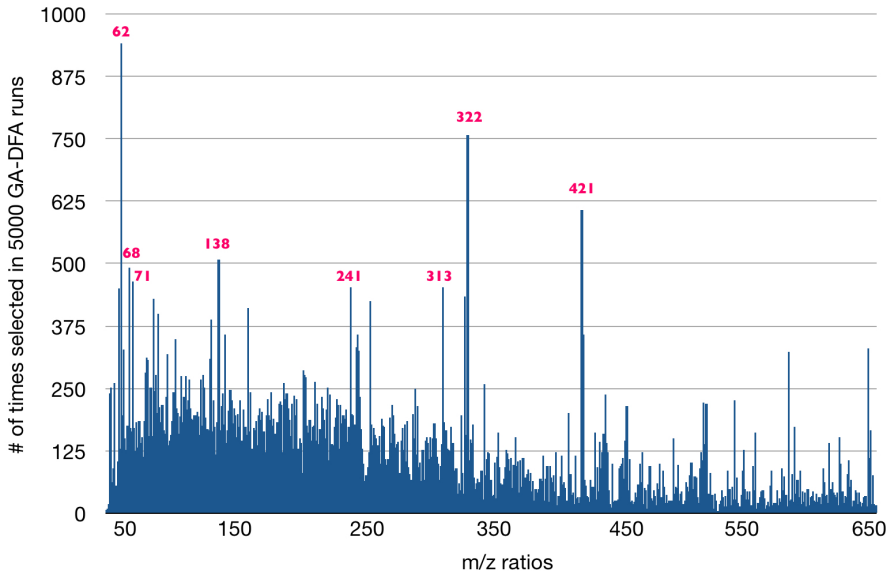


Figure 5.17: Frequency of *m/z* ratios chosen by GA-DFA on mass spectra data of siliques extracts for *gamt* mutants.

GA-DFA was applied to four classes of samples at the time: wild-type, *gamt1, gamt2*, and *gamt1*-2 double mutant. Represented in the histogram are the frequency of selection for each *m/z* ratio to distinguish among the four classes in 5000 GA-DFA runs. The top *m/z* ratios selected as the most discriminant were: 62, 68, 71, 241, 313, 332 and 421.

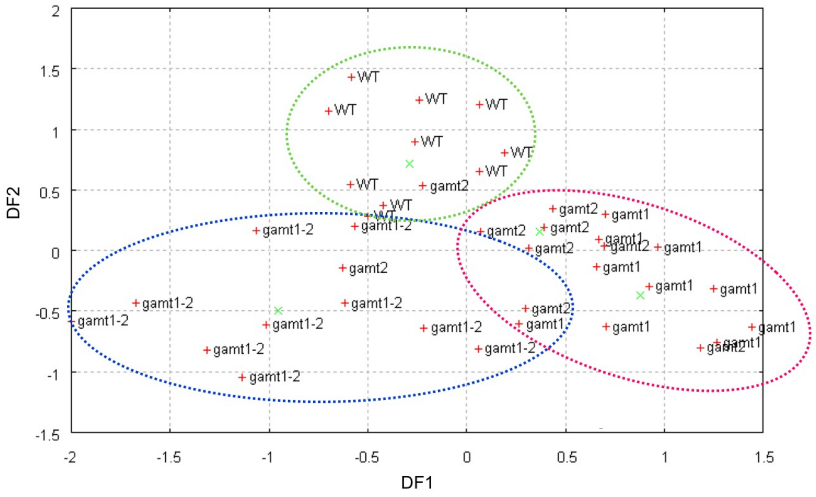


Figure 5.18: DFA with a subset of selected *m/z* ratios chosen by GA-DFA on mass spectra data for wild-type, *gamt1*, *gamt2* and *gamt1-2.*

DFA was applied to four classes of samples at the time: wild-type, *gamt1, gamt2*, and *gamt1*-2 double mutant. A subset of *m/z* ratios that were most often selected together with the top *m/z* ratio were used (62, 138, 332 and 256).

Energy Metabolism, Sugars, and Gibberellins

Several compounds like fumaric acid, malic acid, succinic acid, citric acid and oxalic acid, all of them intermediate molecules in the citric acid cycle were identified using a combination of top m/z ratios selected by GA-DFA. Similarly, several sugars like uridine-5-diphosphoglucose, α-D-glucose-1,5-diphosphate, fructose, glucose, maltose, sucrose and trehalose, were identified. Additionally, an increase in sugar composition was observed in all the mutants compared to the wild-type (Table 5.4).

Based on our initial observation on the microarray analysis results, that several pathways related with energy metabolism were significantly regulated in *gamt1-2* mutant, specially starch and sucrose metabolism and carbon fixation, we speculated that GAMTs regulate the levels of bioactive GAs that promote starch hydrolysis in the outer integument of seeds. It has been proposed that timely supply of starch hydrolyze to the tissues in the seed surface are essential for the structural formation of the seed coat [75]. Our results indicate that the elevated levels of bioactive GAs in seeds due to the lack of GAMT1 and GAMT2 activity results in an increased energy metabolism and elevated starch hydrolysis, which results in changes in the sugars composition and intermediate metabolites in the citric acid cycle.

Expression of some α-amylase genes is induced by GAs and repressed by abscisic acid [205]. Alpha-amylases act in the breakdown of starch to generate soluble sugars, during the formation of the seed coat in Arabidopsis and during the germination of starch-bearing seeds such rice, barley and wheat [75, 224]. In addition to playing a central role in energy metabolism, soluble sugars such glucose and sucrose help regulate many developmental and physiological processes [225–227]. A role for sugars in seed germination is suggested based on the observation that in the presence of abscisic acid, exogenous soluble sugars allow wild-type seeds to germinate [228]. Furthermore, strong evidence support the role of sugars regulating the expression of a significant number of plant genes, including photosynthesis [225, 227]. However the pathways by which plants respond to sugars as signalling molecules remain poorly understood. Characterization of this pathways is complicated by the fact that they form part of a complex regulatory network that include phytohormones and environmental-response pathways [224, 229].

Considering the important role of sugars, the large number differentially genes expressed in *gamt1-2,* and the different pathways regulated; one possible mechanism of action to consider for GAs, is that they elicit these changes indirectly by changing the sugar concentration or flux through the activation of starch hydrolases, like α-amylase. Consequently, rather than being processes regulated by GAs directly all the changes in gene expression observed in the double mutant of the multiple orchestrated mechanisms regulated by sugars. Additional studies are, however, required to further examine this hypothesis.

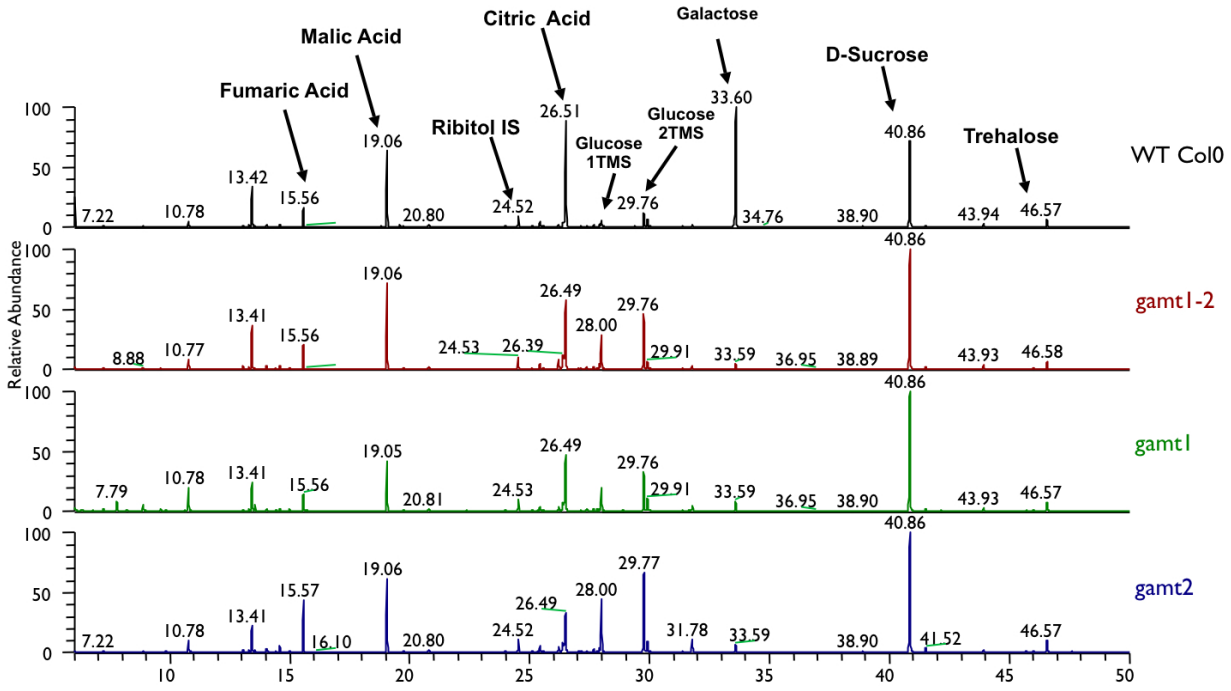


Figure 5.19: GC profiles of polar extracts obtained from siliques of wild-type, *gamt1*, *gamt2* and *gamt1-2*.

Specific compounds identified by both GA-DFA and XC/MS-Anova analysis approaches are shown with their respective retention time.

Table 5.4: Metabolites detected by GC/MS in polar extracts of Arabidopsis siliques of wild-type, *gamt1, gamt2,* and *gamt1-2.*
Among over 150 components found present in Arabidopsis siliques polar extracts by GC/MS analysis, these metabolites were identified using the top m/z ratios selected by GA-DFA.

Derivative	Wild-type		*gamt1-2*			*gamt1*			*gamt2*		
	Resp ratio[a]	SE[b]	Fold Change	Resp ratio	SE	Fold Change	Resp ratio	SE	Fold Change	Resp ratio	SE
L-Alanine	0.550	0.060	1.169	0.643	0.040	**1.797**	0.989	0.177	0.967	0.532	0.047
GSSG-TMS2	0.849	0.067	1.269	1.077	0.077	0.984	0.835	0.160	1.076	0.913	0.195
Ethanolamine	0.680	0.033	0.837	0.569	0.035	1.040	0.707	0.047	1.026	0.698	0.052
L-Serine-TMS2	0.484	0.045	1.696	0.821	0.070	0.807	0.391	0.074	0.570	0.276	0.038
Uridine-5-diphosphoglucose	6.167	0.777	1.194	7.361	0.519	0.900	5.550	1.205	0.817	5.039	0.509
L-Threonine-TMS2	0.437	0.027	1.411	0.616	0.057	1.029	0.450	0.061	0.879	0.384	0.052
Succinic Acid	1.078	0.053	**1.495**	1.611	0.128	**1.629**	1.756	0.244	1.007	1.086	0.107
Fumaric Acid	2.941	0.252	1.358	3.993	0.226	**2.206**	6.487	0.490	1.426	4.194	0.398
DL-Malic Acid	10.080	0.829	1.158	11.675	0.468	0.970	9.782	0.911	0.718	7.238	0.406
DL-Pyroglutamic Acid	0.711	0.247	0.786	0.558	0.044	0.757	0.538	0.090	0.797	0.567	0.094
Alpha-D-Glucose-1,5-diphosphate	1.546	0.127	0.881	1.362	0.064	0.873	1.350	0.093	0.809	1.251	0.086
Fructose-5TMS	0.988	0.122	**4.055**	4.006	0.552	1.733	1.712	0.299	1.448	1.431	0.172
Citric Acid	6.775	1.023	0.397	2.689	0.909	0.498	3.376	0.558	0.543	3.677	0.258
Glucose-RT28.01	1.750	0.145	**7.551**	13.210	2.637	**3.236**	5.661	1.197	**1.749**	3.060	0.456
Oxalic Acid	0.009	0.001	**7.204**	0.063	0.013	0.911	0.008	0.003	0.541	0.005	0.001
Glucose-RT29.76	2.193	0.217	**8.076**	17.714	3.359	**3.477**	7.628	1.773	**1.752**	3.844	0.617
D-Maltose	0.117	0.010	1.207	0.141	0.010	1.403	0.164	0.017	**1.757**	0.205	0.013
Myo-Inositol	0.264	0.036	**3.573**	0.942	0.186	**2.982**	0.786	0.117	**2.205**	0.581	0.040
Galactose-RT33.60	34.273	2.644	0.033	1.126	0.101	0.041	1.421	0.096	0.064	2.180	0.254
D-Sucrose	5.203	0.624	1.154	6.005	0.916	**1.887**	9.816	0.807	**1.871**	9.736	0.392
Trehalose	0.439	0.064	1.286	0.564	0.054	**1.638**	0.718	0.040	**1.797**	0.788	0.055

GC/MS analysis of polar extracts from Arabidopsis siliques of wild-type, *gamt1, gamt2* and *gamt1-2* mutants. Polar extracts were derivatized and analyzed as described in section 4.3. All samples were methoximated and trimethylsilylated. Compounds with a fold change increase of 1.5 or more are indicated in bold. [a]Response ratios are peak areas compared to the internal standard ribitol/adonitol. [b]SE, standard error, n=10.

Table 5.5: Results of statistical analysis of GC/MS data by XCMS-ANOVA showing a subset of the most significant components that can be used for discrimination of wild-type and GAMT mutants, and the putative compounds they derived from according to mass spectra library search

GC/MS-XC/MS Component Name	Component ID	RT	F value	p value
M150T362	Ethylene glycol-TMS	6.03	42.02	0.000
M86T537	Unknown-RT8.95	8.95	15.97	0.007
M156T619	Unknown-RT10.31	10.31	12.78	0.028
M201T934	Succinic Acid	14.60	13.84	0.017
M128T934	Fumaric Acid	15.56	14.92	0.011
M234T1143	Malic Acid	19.05	17.54	0.004
M441T1584	Fructose-5TMS	26.39	15.58	0.008
M274T1590	Citric Acid-4TMS	26.49	15.68	0.008
M261T1661	Erythritol-4TMS	27.69	15.48	0.008
M194T1680	Glucose-RT28.01	28.01	22.55	0.001
M374T1743	L-Ascorbic Acid	29.05	24.98	0.000
M347T1787	Glucose-RT29.78	29.78	21.99	0.001
M220T1792	Maltose	29.86	15.54	0.008
M266T1907	Myo-Inositol-6TMS	31.79	21.02	0.001
M318T2016	Galactose-RT33.60	33.60	75.34	0.000
M159T2451	D-Sucrose	40.85	32.79	0.000
M331T2795	Trehalose	46.58	18.49	0.002

5.4.5 Differential Response of *gamt1-2* mutant to Stress Conditions

To test the relative contribution of GAMTs to abiotic stress tolerance in Arabidopsis, we subjected seedlings of *gamt* single and double mutants with lost-of-function of GAMT1 and GAMT2, to different abiotic stresses, and scored the seedlings for root growth and lenght, parameters that reflect the overall health and stress tolerance of the plants [213]. As shown in Fig. 5.20, significant differences were found in the tolerance of *gamt1-2* to cold and heat stress compared to the response of the wild-type and single mutants. Additionally, *gamt1-2* was significantly sensitive to salt concentrations above 100mM. Interestingly, these observations are in agreement with the GO categories enrichment found for genes differentially expressed in *gamt1-2*. While genes in the GO category response to abiotic stimulus (GO:09628), response to cold (GO:09409), response to temperature stimulus (GO:09266) were significantly up regulated, the GO category response to salt stress (GO:09651) was significantly down-regulated.

Our the results obtained with microarray gene expression analysis, suggested that there is a function redundancy in the methylation activity of GAMT1 and GAMT2 *in planta*. The significantly different response of *gamt1-2* to cold, heat and salinity, and the absence of significant differences between wild-type and the single knockouts, in terms of stress response, support the hypothesis that even though GAMT1 and GAMT2 may have specific methylation activity *in vitro*, they can partially compensate for the lack-of-function of each other.

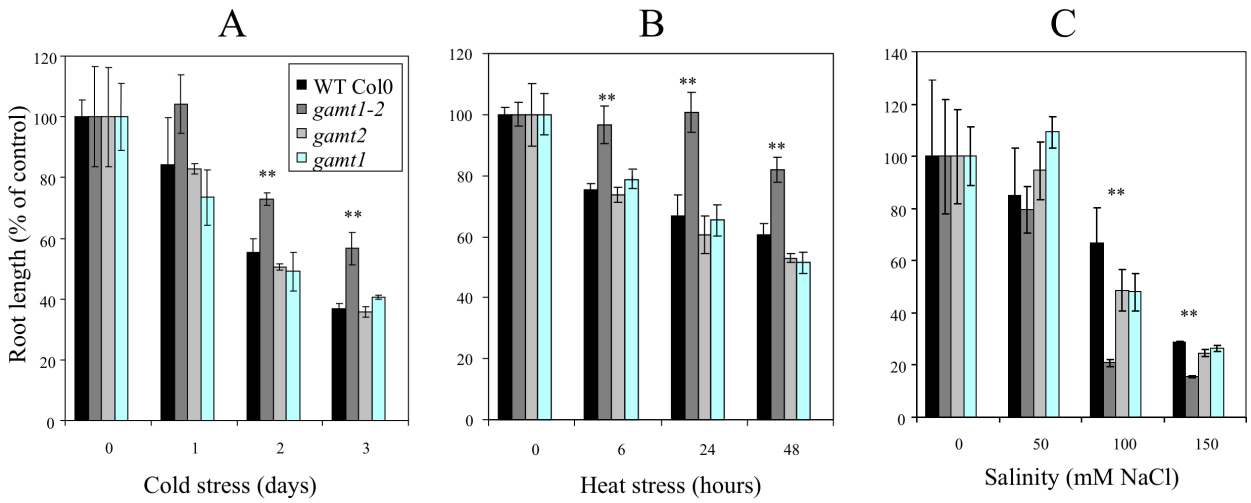


Figure 5.20: Enhanced tolerance of *gamt1-2* double mutant plants to cold and heat stress conditions, and increased sensitivity to salinity.
Effects of abiotic stresses imposed by cold (A), heat (38ºC; B), and salinity (C) on root elongation in 5-d-old seedlings of the different mutant lines and wild-type control were measured with seedlings grown on agar plates. **, P = 0.01 in *t*-test.

Upon exposure to stresses, plants reduce their growth rate by changing hormonal production. Evidence has emerged that the GA metabolism pathway is altered in response to abiotic stresses. A relationship between gibberellin levels and plant stress protection has been tested using near-isogenic lines of a normal and dwarf barley (*Hordeum vulgare* L.) [230]. Their results suggest that reduced GA levels or sensitivity to GA, with a concomitant reduction in height are important for induction of stress tolerance [231]. In Arabidopsis, DDF1 and AP2 transcription factors strongly induced by high-salinity stress, causes a reduction in GA4 content and dwarfism. DFF1 is closely related to the dehydration responsive element–binding proteins (DREBs) involved in stress responses [232]. In addition, transgenic plants overexpression DDF1, as well as GA-deficient *ga1-3* mutant exhibit a higher survival rate under salinity conditions, whereas exogenous GA treatment reduces survival rate [232]. It has been proposed that the salt-inducible DDF1 gene is involved in growth responses under high salinity conditions in part through altering GA levels [202]. In a recent review of the gibberellin metabolism and its regulation, Yamagushi [202] mentioned that in Arabidopsis GA2ox7, which encodes a GA2ox that specifically deactivates C_{20}-GAs (bioactive GAs), is a target of DDF1, suggesting that salt decreases bioactive GAs levels through elevated deactivation. This could be further tested measuring the level of bioactive GAs in *gamt1-2* and monitoring the expression of GA2ox7 and DDF1 in the mutant under salinity stress conditions.

In Arabidopsis, the evidence available about the role of gibberellins in abiotic stress response is limited to salinity. To the best of our knowledge, this is the first report connecting alterations in the GA biosynthesis and abiotic stress responses other than salt.

5.4.6 Gibberellin Methylation Role in Arabidopsis Seed Development

Gibberellins are synthesised during seed development and germination [73], but the distribution and specific role of GA during siliques and seeds development in Arabidopsis is poorly understood [74, 75]. A significant number of enzymes involve in the biosynthesis and catabolism of GAs identified [202, 215, 216], but the role of many of these enzymes in plant development remain to be fully characterized. Gibberellin 3-oxidase (GA3ox) catalyzes the final step in the synthesis of bioactive GAs. A recent study examined the patterns of four different GA3oxs, and found that in developing siliques, GA3ox1 is mainly expressed in the replums, funiculi, and the silique receptacles, whereas the other GA3ox genes are only expressed in developing seeds [233]. According to Hu *et al.* [233] active GAs appear to be transported from the seed endosperm to the surrounding maternal tissues where they promote growth. Homologs of GA 2-oxidases capable of deactivating GAs, expressed in Arabidopsis developing seeds have been identified, but their physiological role in GA deactivation is yet to be fully established [217, 234]. Methyl ester of GA_4 and GA_3, have been proven to be inactive in Arabidopsis and Petunia, respectively [76, 77], although GA_3 and GA_3 appear to be bioactive molecules during the formation of the antheridium in ferns [235].

The observation that transgenic Arabidopsis plants overexpressing GAMT1 or GAMT2 have lower levels of GAs, show varying levels of dwarfism, delayed flowering, and sterility and that these phenotypes can be partially reversed by the external application of GAs suggest that MeGAs are indeed inactive in Arabidopsis as well [53]. Furthermore, MeGA4 applied to the GAMT1-overexpressing lines have no effect in the dwarf phenotype [53]. Therefore, methylation is part of the irreversible deactivation of GAs. Methylated GAs are more hydrophobic, allowing modified GAs to diffuse through membranes to reach their site of degradation. The expression of GAMT1 and GAMT2 in the siliques peaks at the later stages of silique development, when the embryos in the siliques are nearing the end of their fast cell division and elongation, a process that is controlled by GAs, further suggesting a role for GAMT1 and GAMT2 in inactivation of GAs.

Methylation is reversible process, in fact several esterases for different signalling molecules have been recently characterized [39, 40, 70]. While it is possible that methylation of GAs, like GA glycosylation, is in some cases a reversible process in plants, the results of ancymidol inhibition experiments indicated that the wild-type Arabidopsis seeds are unable to recover the GAs that had been lost due to the activities of GAMT1 and GAMT2 [53]. Since both GAMT1 and GAMT2 are expressed mainly in the developing siliques (including seeds), the enzymes they encode have partially overlapping substrate specificities, and double mutants show an additive effect of GAMT1 and GAMT2 expression on levels of some GAs, it can be concluded that the two genes have partially overlapping functions in silique and seed development.

5.5 Conclusions

We have reported that two members of the SABATH methyltransferase gene family in Arabidopsis, GAMT1 (At4g26420) and GAMT2 (At5g56300), encode enzymes that methylate the carboxyl group of gibberellic acids (GAs), resulting in the methyl ester of GAs (MeGAs) [53]. Both genes are expressed mostly in siliques during seed development. It is generally known that GAs are synthesized during seed development and germination. However, the specific physiological role of GA methylation in siliques, seeds or other plant tissues has not been examined. To the best of our knowledge, this is the first study exploring the consequences of GAs methylation in Arabidopsis.

Transcriptomics and metabolomics approaches are sources of rich information in terms of the biological adaptation of cellular responses, with each strategy independently providing to some extend a different, but complementary view of the changes in the inner cellular processes. Thereby, complementary gene expression analyses, at the transcriptome and metabolome levels, have provided insights into the mechanisms that underlie the molecular events associated with the phenotype of GAMT1 and GAMT2 mutants. In this study we used transcriptional and metabolite profiling to explain the phenotypic responses in GAMTs mutants, and to understand the physiological role of GA methylation in siliques and seeds.

Our observations indicate that GAMT1 and GAMT2 are responsible for the methylation of bioactive GAs in Arabidopsis plants, resulting the inactivation of these signaling molecules. According to the observation that *gamt1-2* double mutant showed additive effect in the accumulation of bioactive GAs, and induced more significant changes on the gene expression profile than *gamt1* or *gamt2* alone, we believe that GAMT1 and GAMT2 have overlapping function. As direct consequence of the lack of GA methylation activity in *gamt1-2*, the levels of bioactive GAs in siliques increased, compared to the wild-type. In Arabidopsis seeds specifically, this increased levels of bioactive GAs results in the activation of feedback and feedforward mechanisms in the gibberellin biosynthetic pathway to regulate *de novo* synthesis of bioactive GAs.

Several metabolic pathways and genes are significantly regulated in response to the high levels of bioactive GAs, particularly those related with energy metabolism, and biosynthesis of secondary metabolites. The starch hydrolysis, a well documented process, known to be regulated by GAs in developing and germinating seeds [75], was significantly up-regulated. As a consequence of the elevated expression of starch hydrolases, the levels of sugars like glucose and sucrose, increased significantly in *gamt1-2* compared to the wild-type. The flavonoid biosynthesis and other specialized biosynthetic processes were activated by gibberellins through a regulation process that we believe may be sugar-mediated.

As reported previously, our results indicate that high levels of bioactive GAs in Arabidopsis contribute to higher sensitivity to salinity stress. Additionally, for the first time reported, we found that high GAs levels, contribute to increase the tolerance to cold and heat stress. However, the specific molecular mechanism about how GAs may regulated such processes need further studies.

Overall, the results obtained through the gene expression and metabolite profiling indicates that bioactive GAs, and the biological processes they are involved in seeds, are regulated by the catalytic activity of GAMT1 and GAMT2 which inactivates GAs through irreversible methylation.

Bibliography

[1] Weinstein JN: **'Omic' and hypothesis-driven research in the molecular pharmacology of cancer.** *Curr Opin Pharmacol* 2002, **2**(4):361–365.

[2] Appleford NEJ, Evans DJ, Lenton JR, Gaskin P, Croker SJ, Devos KM, Phillips AL, Hedden P: **Function and transcript analysis of gibberellin-biosynthetic enzymes in wheat.** *Planta* 2006, **223**(3):568–82.

[3] Chen F, Ro DK, Petri J, Gershenzon J, Bohlmann J, Pichersky E, Tholl D: **Characterization of a root-specific Arabidopsis terpene synthase responsible for the formation of the volatile monoterpene 1,8-cineole.** *Plant Physiol* 2004, **135**(4):1956–66.

[4] Chory J, Ecker JR, Briggs S, Caboche M, Coruzzi GM, Cook D, Dangl J, Grant S, Guerinot ML, Henikoff S, Martienssen R, Okada K, Raikhel NV, Somerville CR, Weigel D: **National Science Foundation-Sponsored Workshop Report: "The 2010 Project"Functional Genomics and the Virtual Plant. A Blueprint for Understanding How Plants Are Built and How to Improve Them**. *Plant Physiology* 2000, **123**(2):423–426.

[5] Sumner LW, Mendes P, Dixon RA: **Plant metabolomics: large-scale phytochemistry in the functional genomics era.** *Phytochemistry* 2003, **62**(6):817–36.

[6] Goossens A, Hakkinen ST, Laakso I, Seppanen-Laakso T, Biondi S, De Sutter V, Lammertyn F, Nuutila AM, Soderlund H, Zabeau M, Inze D, Oksman-Caldentey KM: **A functional genomics approach toward the understanding of secondary metabolism in plant cells.** *Proc Natl Acad Sci U S A* 2003, **100**(14):8595–600.

[7] Hedden P, Holdsworth MJ, Simpson GG, Edwards KJ, Lenton JR: **New frontiers in plant development: from genes to phenotype.** *Trends Plant Sci* 2002, **7**(9):381–3.

[8] Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuhara M, Arita M, Fujiwara T, Saito K: **Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in Arabidopsis thaliana.** *Proc Natl Acad Sci U S A* 2004, **101**(27):10205–10.

[9] Hirai MY, Klein M, Fujikawa Y, Yano M, Goodenowe DB, Yamazaki Y, Kanaya S, Nakamura Y, Kitayama M, Suzuki H, Sakurai N, Shibata D, Tokuhisa J, Reichelt M, Gershenzon J, Papenbrock J, Saito K: **Elucidation of gene-to-gene and metabolite-to-gene networks in arabidopsis by integration of metabolomics and transcriptomics.** *J Biol Chem* 2005, **280**(27):25590–5.

[10] Phelps TJ, Palumbo AV, Beliaev AS: **Metabolomics and microarrays for improved understanding of phenotypic characteristics controlled by both genomics and environmental constraints.** *Curr Opin Biotechnol* 2002, **13**:20–4.

[11] Tokimatsu T, Sakurai N, Suzuki H, Ohta H, Nishitani K, Koyama T, Umezawa T, Misawa N, Saito K, Shibata D: **KaPPA-View. A Web-Based Analysis Tool for Integration of Transcript and Metabolite Data on Plant Metabolic Pathway Maps**. *Plant Physiology* 2005, **138**(3):1289–1300.

[12] Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L: **Metabolite profiling for plant functional genomics.** *Nat Biotechnol* 2000, **18**(11):1157–61.

[13] Initiative TAG: **Analysis of the genome sequence of the flowering plant Arabidopsis thaliana**. *Nature* 2000, **408**(6814):796–815.

[14] Chen F, D'Auria JC, Tholl D, Ross JR, Gershenzon J, Noel JP, Pichersky E: **An Arabidopsis thaliana gene for methylsalicylate biosynthesis, identified by a biochemical genomics approach, has a role in defense.** *Plant J* 2003, **36**(5):577–88.

[15] Fridman E, Pichersky E: **Metabolomics, genomics, proteomics, and the identification of enzymes and their substrates and products.** *Curr Opin Plant Biol* 2005, **8**(3):242–8.

[16] Gomase VS, Changbhale SS, Patil SA, Kale KV: **Metabolomics.** *Curr Drug Metab* 2008, **9**:89–98.

[17] Zhu M, Zhao S: **Candidate gene identification approach: progress and challenges.** *Int J Biol Sci* 2007, **3**(7):420–427.

[18] Pichersky E, Gershenzon J: **The formation and function of plant volatiles: perfumes for pollinator attraction and defense.** *Curr Opin Plant Biol* 2002, **5**(3):237–43.

[19] Dudareva N, Pichersky E, Gershenzon J: **Biochemistry of plant volatiles.** *Plant Physiol* 2004, **135**(4):1893–902.

[20] Ribnicky DM, Shulaev V, Raskin I: **Intermediates of salicylic acid biosynthesis in tobacco**. *Plant Physiol* 1998, **118**(2):565–72.

[21] Shulaev V, Leon J, Raskin I: **Is Salicylic Acid a Translocated Signal of Systemic Acquired Resistance in Tobacco?** *Plant Cell* 1995, **7**(10):1691–1701.

[22] Noel JP, Dixon RA, Pichersky E, Zubieta C, Ferrer JL, Romeo JT: *Structural, functional, and evolutionary basis for methylation of plant small molecules*, Elsevier, *Volume 37* 2003 chap. 2, :37–58.

[23] D'Auria JC, Chen F, Pichersky E: **The SABATH family of MTs in Arabidopsis thaliana and other plant species**. In *Recent Advances in Phytochemistry, Volume 37*. Edited by Romeo JT, Ibrahim RK, Varin L, De Luca V, Elsevier Science Ltd, Oxford 2003:253–283.

[24] Ross JR, Nam KH, D'Auria JC, Pichersky E: **S-Adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, an enzyme involved in floral scent production and plant defense, represents a new class of plant methyltransferases.** *Arch Biochem Biophys* 1999, **367**:9–16.

[25] Seo HS, Song JT, Cheong JJ, Lee YH, Lee YW, Hwang I, Lee JS, Choi YD: **Jasmonic acid carboxyl methyltransferase: a key enzyme for jasmonate-regulated plant responses.** *Proc Natl Acad Sci U S A* 2001, **98**(8):4788–93.

[26] Murfitt LM, Kolosova N, Mann CJ, Dudareva N: **Purification and characterization of S-adenosyl-L-methionine:benzoic acid carboxyl methyltransferase, the enzyme responsible for biosynthesis of the volatile ester methyl benzoate in flowers of Antirrhinum majus.** *Arch Biochem Biophys* 2000, **382**:145–51.

[27] Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H: **Caffeine synthase gene from tea leaves.** *Nature* 2000, **406**(6799):956–7.

[28] Zubieta C, Ross JR, Koscheski P, Yang Y, Pichersky E, Noel JP: **Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family.** *Plant Cell* 2003, **15**(8):1704–16.

[29] Kato M, Mizuno K, Fujimura T, Iwama M, Irie M, Crozier A, Ashihara H: **Purification and characterization of caffeine synthase from tea leaves.** *Plant Physiol* 1999, **120**(2):579–86.

[30] Ogawa M, Herai Y, Koizumi N, Kusano T, Sano H: **7-Methylxanthine methyltransferase of coffee plants. Gene isolation and enzymatic properties.** *J Biol Chem* 2001, **276**(11):8213–8.

[31] Ibrahim RK, Muzac I: **The methyltransferase gene superfamily: a tree with multiple branches**. In *Recent Advances in Phytochemistry, Volume 34*. Edited by Romeo JT, Ibrahim RK, Varin L, De Luca V, Elsevier Science Ltd, Oxford 2000:349–384.

[32] Zimmermann P, Hirsch-Hoffmann M, Hennig L, Gruissem W: **GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox.** *Plant Physiol* 2004, **136**:2621–2632.

[33] Yang Y, Varbanova M, Ross JR, Wang G, Cortes D, Fridman E, Shulaev V, Noel JP, Pichersky E: **Methylation and Demethylation of Plant Signaling Molecules**. In *Recent Advances in Phytochemistry, Volume 40*. Edited by Romeo JT, Ibrahim RK, Varin L, De Luca V, Elsevier Science Ltd, Oxford 2006.

[34] Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Amagai M, Kuwata C, Tsugane T, Masuda T, Shimada H, Takamiya K, Ohta H, Tabata S: **Monitoring of methyl jasmonate-responsive genes in Arabidopsis by cDNA macroarray: self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways.** *DNA Res* 2001, **8**(4):153–61.

[35] Cipollini D: **Consequences of the overproduction of methyl jasmonate on seed production, tolerance to defoliation and competitive effect and response of Arabidopsis thaliana**. *New Phytologist* 2007, **173**:146–153.

[36] Ament K, Kant MR, Sabelis MW, Haring MA, Schuurink RC: **Jasmonic acid is a key regulator of spider mite-induced volatile terpenoid and methyl salicylate emission in tomato.** *Plant Physiol* 2004, **135**(4):2025–2037.

[37] Wu J, Wang L, Baldwin IT: **Methyl jasmonate-elicited herbivore resistance: does MeJA function as a signal without being hydrolyzed to JA?** *Planta* 2008, **227**(5):1161–1168.

[38] Kumar D, Klessig DF: **High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity.** *Proc Natl Acad Sci U S A* 2003, **100**(26):16101–16106.

[39] Forouhar F, Yang Y, Kumar D, Chen Y, Fridman E, Park SW, Chiang Y, Acton TB, Montelione GT, Pichersky E, Klessig DF, Tong L: **Structural and biochemical studies identify tobacco SABP2 as a methyl salicylate esterase and implicate it in plant innate immunity.** *Proc Natl Acad Sci U S A* 2005, **102**(5):1773–8.

[40] Yang Y, Xu R, Ma C, Vlot A, Klessig D, Pichersky E: **Inactive Methyl Indole-3-Acetic Acid Ester Can Be Hydrolyzed and Activated by Several Esterases Belonging to the AtMES Esterase Family of Arabidopsis thaliana.** *Plant Physiol* 2008.

[41] Negre F, Kolosova N, Knoll J, Kish CM, Dudareva N: **Novel S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, an enzyme responsible for biosynthesis of methyl salicylate and methyl benzoate, is not involved in floral scent production in snapdragon flowers.** *Arch Biochem Biophys* 2002, **406**(2):261–70.

[42] Negre F, Kish CM, Boatright J, Underwood B, Shibuya K, Wagner C, Clark DG, Dudareva N: **Regulation of methylbenzoate emission after pollination in snapdragon and petunia flowers.** *Plant Cell* 2003, **15**(12):2992–3006.

[43] Pott MB, Effmert U, Piechulla B: **Transcriptional and post-translational regulation of S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase (SAMT) during Stephanotis floribunda flower development.** *J Plant Physiol* 2003, **160**(6):635–43.

[44] Brodersen P, Petersen M, Bjorn Nielsen H, Zhu S, Newman MA, Shokat KM, Rietz S, Parker J, Mundy J: **Arabidopsis MAP kinase 4 regulates salicylic acid- and jasmonic acid/ethylene-dependent responses via EDS1 and PAD4.** *Plant J* 2006, **47**(4):532–546.

[45] Petersen M, Brodersen P, Naested H, Andreasson E, Lindhart U, Johansen B, Nielsen HB, Lacy M, Austin MJ, Parker JE, Sharma SB, Klessig DF, Martienssen R, Mattsson O, Jensen AB, Mundy J: **Arabidopsis map kinase 4 negatively regulates systemic acquired resistance.** *Cell* 2000, **103**(7):1111–1120.

[46] Kachroo P, Shanklin J, Shah J, Whittle EJ, Klessig DF: **A fatty acid desaturase modulates the activation of defense signaling pathways in plants.** *Proc Natl Acad Sci U S A* 2001, **98**(16):9448–9453.

[47] Koo Y, Kim M, Kim E, Song J, Jung C, Moon J, Kim J, Seo H, Song S, Kim J, Lee J, Cheong J, Choi Y: **Overexpression of salicylic acid carboxyl methyltransferase reduces salicylic acid-mediated pathogen resistance in Arabidopsis thaliana.** *Plant Mol Biol* 2007.

[48] Van Poecke RM, Posthumus MA, Dicke M: **Herbivore-induced volatile production by Arabidopsis thaliana leads to attraction of the parasitoid Cotesia rubecula: chemical, behavioral, and gene-expression analysis.** *J Chem Ecol* 2001, **27**(10):1911–1928.

[49] Qin G, Gu H, Zhao Y, Ma Z, Shi G, Yang Y, Pichersky E, Chen H, Liu M, Chen Z, Qu LJ: **An indole-3-acetic acid carboxyl methyltransferase regulates Arabidopsis leaf development.** *Plant Cell* 2005, **17**(10):2693–704.

[50] Zhao N, Ferrer JL, Ross J, Guan J, Yang Y, Pichersky E, Noel JP, Chen F: **Structural, biochemical, and phylogenetic analyses suggest that indole-3-acetic acid methyltransferase is an evolutionarily ancient member of the SABATH family.** *Plant Physiol* 2008, **146**(2):455–467.

[51] Yang Y, Yuan JS, Ross J, Noel JP, Pichersky E, Chen F: **An Arabidopsis thaliana methyltransferase capable of methylating farnesoic acid.** *Arch Biochem Biophys* 2006, **448**(1-2):123–32.

[52] Pimenta Lange MJ, Lange T: **Gibberellin Biosynthesis and the Regulation of Plant Development**. *Plant Biology* 2006, (3):281–290.

[53] Varbanova M, Yamaguchi S, Yang Y, McKelvey K, Hanada A, Borochov R, Yu F, Jikumaru Y, Ross J, Cortes D, Je Ma C, Noel JP, Mander L, Shulaev V, Kamiya Y, Rodermel S, Weiss D, Pichersky E: **Methylation of Gibberellins by Arabidopsis GAMT1 and GAMT2**. *The Plant Cell* 2007, **19**:37–45.

[54] Pott MB, Hippauf F, Saschenbrecker S, Chen F, Ross J, Kiefer I, Slusarenko A, Noel JP, Pichersky E, Effmert U, Piechulla B: **Biochemical and Structural Characterization of Benzenoid Carboxyl Methyltransferases Involved in Floral Scent Production in Stephanotis floribunda and Nicotiana suaveolens**. *Plant Physiology* 2004, **135**(4):1946–1955.

[55] Xu R, Song F, Zheng Z: **OsBISAMT1, a gene encoding S-adenosyl-L-methionine: salicylic acid carboxyl methyltransferase, is differentially expressed in rice defense responses.** *Mol Biol Rep* 2006, **33**(3):223–31.

[56] Pichersky E, Gang DR: **Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective.** *Trends Plant Sci* 2000, **5**(10):439–45.

[57] Schwab W: **Metabolome diversity: too few genes, too many metabolites?; Plant Metabolomics**. *Phytochemistry* 2003, **62**(6):837–849.

[58] Zhao N, Guan J, Lin H, Chen F: **Molecular cloning and biochemical characterization of indole-3-acetic acid methyltransferase from poplar.** *Phytochemistry* 2007, **68**(11):1537–1544.

[59] Shinoda T, Itoyama K: **Juvenile hormone acid methyltransferase: a key regulatory enzyme for insect metamorphosis.** *Proc Natl Acad Sci U S A* 2003, **100**(21):11986–11991.

[60] Holford KC, Edwards KA, Bendena WG, Tobe SS, Wang Z, Borst DW: **Purification and characterization of a mandibular organ protein from the American lobster, Homarus americanus: a putative farnesoic acid O-methyltransferase.** *Insect Biochem Mol Biol* 2004, **34**(8):785–798.

[61] Yang Y, Yuan JS, Ross J, Noel JP, Pichersky E, Chen F: **An Arabidopsis thaliana methyltransferase capable of methylating farnesoic acid.** *Arch Biochem Biophys* 2006, **448**(1-2):123–132.

[62] Kapteyn J, Qualley AV, Xie Z, Fridman E, Dudareva N, Gang DR: **Evolution of Cinnamate/p-coumarate carboxyl methyltransferases and their role in the biosynthesis of methylcinnamate.** *Plant Cell* 2007, **19**(10):3212–3229.

[63] Mizuno K, Okuda A, Kato M, Yoneyama N, Tanaka H, Ashihara H, Fujimura T: **Isolation of a new dual-functional caffeine synthase gene encoding an enzyme for the conversion of 7-methylxanthine to caffeine from coffee (Coffea arabica L.).** *FEBS Lett* 2003, **534**(1-3):75–81.

[64] Mizuno K, Kato M, Irino F, Yoneyama N, Fujimura T, Ashihara H: **The first committed step reaction of caffeine biosynthesis: 7-methylxanthosine synthase is closely homologous to caffeine synthases in coffee (Coffea arabica L.).** *FEBS Lett* 2003, **547**(1-3):56–60.

[65] Misako K, Kouichi M: **Caffeine synthase and related methyltransferases in plants.** *Front Biosci* 2004, **9**(1093-4715 (Electronic)):1833–42.

[66] Turlings TC, Loughrin JH, McCall PJ, Rose US, Lewis WJ, Tumlinson JH: **How caterpillar-damaged plants protect themselves by attracting parasitic wasps.** *Proc Natl Acad Sci U S A* 1995, **92**(10):4169–4174.

[67] Cheong JJ, Choi YD: **Methyl jasmonate as a vital substance in plants.** *Trends Genet* 2003, **19**(7):409–13.

[68] Gray WM: **Hormonal regulation of plant growth and development.** *PLoS Biol* 2004, **2**(9):E311.

[69] Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Masuda T, Shimada H, Takamiya K, Tabata S, Ohta H: **Genome-wide expression-monitoring of jasmonate-responsive genes of Arabidopsis using cDNA arrays.** *Biochem Soc Trans* 2000, **28**(6):863–4.

[70] Stuhlfelder C, Mueller MJ, Warzecha H: **Cloning and expression of a tomato cDNA encoding a methyl jasmonate cleaving esterase.** *Eur J Biochem* 2004, **271**(14):2976–2983.

[71] Leyser O: **Molecular genetics of auxin signaling.** *Annu Rev Plant Biol* 2002, **53**:377–398.

[72] Li L, Hou X, Tsuge T, Ding M, Aoyama T, Oka A, Gu H, Zhao Y, Qu LJ: **The possible action mechanisms of indole-3-acetic acid methyl ester in Arabidopsis.** *Plant Cell Rep* 2008, **27**(3):575–584.

[73] Ogawa M, Hanada A, Yamauchi Y, Kuwahara A, Kamiya Y, Yamaguchi S: **Gibberellin biosynthesis and response during Arabidopsis seed germination.** *Plant Cell* 2003, **15**(7):1591–604.

[74] Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G: **AtGA3ox2, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis.** *Plant Physiol* 2004, **136**(3):3660–9.

[75] Kim YC, Nakajima M, Nakayama A, Yamaguchi I: **Contribution of gibberellins to the formation of Arabidopsis seed coat through starch degradation.** *Plant Cell Physiol* 2005, **46**(8):1317–25.

[76] Weiss D, Van Der Luit A, Knegt E, Vermeer E, Mol J, Kooter J: **Identification of Endogenous Gibberellins in Petunia Flowers (Induction of Anthocyanin Biosynthetic Gene Expression and the Antagonistic Effect of Abscisic Acid).** *Plant Physiol* 1995, **107**(3):695–702.

[77] Cowling RJ, Kamiya Y, Seto H, Harberd NP: **Gibberellin dose-response regulation of GA4 gene transcript levels in Arabidopsis.** *Plant Physiol* 1998, **117**(4):1195–1203.

[78] Zhang Y, Zhu Y, Peng Y, Yan D, Li Q, Wang J, Wang L, He Z: **Gibberellin homeostasis and plant height control by EUI and a role for gibberellin in root gravity responses in rice.** *Cell Res* 2008, **18**(3):412–421.

[79] Zhu Y, Nomura T, Xu Y, Zhang Y, Peng Y, Mao B, Hanada A, Zhou H, Wang R, Li P, Zhu X, Mander LN, Kamiya Y, Yamaguchi S, He Z: **ELONGATED UPPERMOST INTERNODE encodes a cytochrome P450 monooxygenase that epoxidizes gibberellins in a novel deactivation reaction in rice.** *Plant Cell* 2006, **18**(2):442–456.

[80] Bede JC, Teal PE, Goodman WG, Tobe SS: **Biosynthetic pathway of insect juvenile hormone III in cell suspension cultures of the sedge Cyperus iria.** *Plant Physiol* 2001, **127**(2):584–593.

[81] Holtorf H, Guitton MC, Reski R: **Plant functional genomics.** *Naturwissenschaften* 2002, **89**(6):235–49.

[82] Gadjev I, Vanderauwera S, Gechev TS, Laloi C, Minkov IN, Shulaev V, Apel K, Inze D, Mittler R, Van Breusegem F: **Transcriptomic footprints disclose specificity of reactive oxygen species signaling in Arabidopsis.** *Plant Physiol* 2006, **141**(2):436–45.

[83] Walsh B, Henderson D: **Microarrays and beyond: What potential do current and future genomics tools have for breeders?** *Journal of Animal Science* 2004, **82**(13_suppl):E292–299.

[84] Hirai MY, Saito K: **Post-genomics approaches for the elucidation of plant adaptive mechanisms to sulphur deficiency.** *J Exp Bot* 2004, **55**(404):1871–9.

[85] Yang GX, Jan A, Shen SH, Yazaki J, Ishikawa M, Shimatani Z, Kishimoto N, Kikuchi S, Matsumoto H, Komatsu S: **Microarray analysis of brassinosteroids- and gibberellin-regulated gene expression in rice seedlings**. *Molecular Genetics and Genomics* 2004, **V271**(4):468–478.

[86] Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J: **Microarray Analysis of Developing Arabidopsis Seeds**. *Plant Physiology* 2000, **124**(4):1570–1581.

[87] Borevitz JO, Ecker JR: **PLANT GENOMICS: The Third Wave**. *Annual Review of Genomics and Human Genetics* 2004, **5**:443–477.

[88] Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270**(5235):467–470.

[89] Stekel D: *Microarrays Bioinformatics*. Cambridge 2003.

[90] Clarke JD, Zhu T: **Microarray analysis of the transcriptome as a stepping stone towards understanding biological systems: practical considerations and perspectives**. *The Plant Journal* 2006, **45**(4):630–650.

[91] Cordero F, Botta M, Calogero RA: **Microarray data analysis and mining approaches.** *Brief Funct Genomic Proteomic* 2007, **6**(4):265–281.

[92] Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP: **Summaries of Affymetrix GeneChip probe level data.** *Nucleic Acids Res* 2003, **31**(4):e15.

[93] Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data**. *Biostatistics* 2003, **4**(2):249–264.

[94] Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F: **A model-based background adjustment for oligonucleotide expression arrays.** *Journal of the American Statistical Association* 2004, **99**(468).

[95] Kerr MK, Martin M, Churchill GA: **Analysis of variance for gene expression microarray data.** *J Comput Biol* 2000, **7**(6):819–837.

[96] Kerr MK, Churchill GA: **Statistical design and the analysis of gene expression microarray data.** *Genet Res* 2001, **77**(2):123–128.

[97] Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3**:Article3.

[98] Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing**. *Journal of the Royal Statistical Society. Series B (Methodological)* 1995, **57**:289–300.

[99] Hosack DA, Dennis GJ, Sherman BT, Lane HC, Lempicki RA: **Identifying biological themes within lists of genes with EASE.** *Genome Biol* 2003, **4**(10):R70.

[100] Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S, Bussey KJ, Riss J, Barrett JC, Weinstein JN: **GoMiner: a resource for biological interpretation of genomic and proteomic data.** *Genome Biol* 2003, **4**(4):R28.

[101] Huang D, Sherman B, Tan Q, Collins J, Alvord WG, Roayaei J, Stephens R, Baseler M, Lane HC, Lempicki R: **The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists**. *Genome Biology* 2007, **8**(9).

[102] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**(43):15545–15550.

[103] Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP: **GSEA-P: a desktop application for Gene Set Enrichment Analysis.** *Bioinformatics* 2007, **23**(23):3251–3253.

[104] Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadrinab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseeuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR: **Genome-wide insertional mutagenesis of Arabidopsis thaliana.** *Science* 2003, **301**(5633):653–657.

[105] Sessions A, Burke E, Presting G, Aux G, McElver J, Patton D, Dietrich B, Ho P, Bacwaden J, Ko C, Clarke JD, Cotton D, Bullis D, Snell J, Miguel T, Hutchison D, Kimmerly B, Mitzel T, Katagiri F, Glazebrook J, Law M, Goff SA: **A high-throughput Arabidopsis reverse genetics system.** *Plant Cell* 2002, **14**(12):2985–2994.

[106] Li Y, Rosso MG, Strizhov N, Viehoever P, Weisshaar B: **GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in Arabidopsis thaliana.** *Bioinformatics* 2003, **19**(11):1441–1442.

[107] Clough SJ, Bent AF: **Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana.** *Plant J* 1998, **16**(6):735–743.

[108] Wettenhall JM, Simpson KM, Satterley K, Smyth GK: **affylmGUI: a graphical user interface for linear modeling of single channel microarray data.** *Bioinformatics* 2006, **22**(7):897–899.

[109] Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**(10):R80.

[110] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M: **KEGG: Kyoto Encyclopedia of Genes and Genomes.** *Nucleic Acids Res* 1999, **27**:29–34.

[111] Gundlach H, Muller MJ, Kutchan TM, Zenk MH: **Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures.** *Proc Natl Acad Sci U S A* 1992, **89**(6):2389–2393.

[112] Feys B, Benedetti C, Penfold C, Turner J: **Arabidopsis Mutants Selected for Resistance to the Phytotoxin Coronatine Are Male Sterile, Insensitive to Methyl Jasmonate, and Resistant to a Bacterial Pathogen.** *Plant Cell* 1994, **6**(5):751–759.

[113] Ton J, De Vos M, Robben C, Buchala A, Metraux JP, Van Loon LC, Pieterse CMJ: **Characterization of Arabidopsis enhanced disease susceptibility mutants that are affected in systemically induced resistance.** *Plant J* 2002, **29**:11–21.

[114] Devoto A, Ellis C, Magusin A, Chang HS, Chilcott C, Zhu T, Turner JG: **Expression profiling reveals COI1 to be a key regulator of genes involved in wound- and methyl jasmonate-induced secondary metabolism, defence, and hormone interactions.** *Plant Mol Biol* 2005, **58**(4):497–513.

[115] Tamari G, Borochov A, Atzorn R, Weiss D: **Methyl jasmonate induces pigmentation and flavonoid gene expression in petunia corollas: A possible role in wound response**. *Physiologia Plantarum* 1995, **94**:45–50.

[116] Winkel-Shirley B: **Biosynthesis of flavonoids and effects of stress.** *Curr Opin Plant Biol* 2002, **5**(3):218–223.

[117] Creelman R, Mullet J: **BIOSYNTHESIS AND ACTION OF JASMONATES IN PLANTS.** *Annu Rev Plant Physiol Plant Mol Biol* 1997, **48**:355–381.

[118] Creelman RA, Mullet JE: **Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress.** *Proc Natl Acad Sci U S A* 1995, **92**(10):4114–4119.

[119] Voros K, Feussner I, Kuhn H, Lee J, Graner A, Lobler M, Parthier B, Wasternack C: **Characterization of a methyljasmonate-inducible lipoxygenase from barley (Hordeum vulgare cv. Salome) leaves.** *Eur J Biochem* 1998, **251**(1-2):36–44.

[120] Xiang C, Oliver DJ: **Glutathione metabolic genes coordinately respond to heavy metals and jasmonic acid in Arabidopsis.** *Plant Cell* 1998, **10**(9):1539–1550.

[121] Lyi SM, Heller LI, Rutzke M, Welch RM, Kochian LV, Li L: **Molecular and biochemical characterization of the selenocysteine Se-methyltransferase gene and Se-methylselenocysteine synthesis in broccoli.** *Plant Physiol* 2005, **138**:409–420.

[122] Benjamins R, Scheres B: **Auxin: the looping star in plant development.** *Annu Rev Plant Biol* 2008, **59**:443–465.

[123] Overvoorde PJ, Okushima Y, Alonso JM, Chan A, Chang C, Ecker JR, Hughes B, Liu A, Onodera C, Quach H, Smith A, Yu G, Theologis A: **Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana.** *Plant Cell* 2005, **17**(12):3282–3300.

[124] Fedoroff NV: **Cross-talk in abscisic acid signaling.** *Sci STKE* 2002, **2002**(140):RE10.

[125] Trewavas AJ: **Effect of IAA on RNA and protein synthesis. Effects of 3-indolylacetic acid on the metabolism of ribonucleic acid and protein in etiolated subapical sections of Pisum sativum.** *Arch Biochem Biophys* 1968, **123**(2):324–335.

[126] Zurfluh L, Guilfoyle T: **Auxin-induced changes in the patterns of protein synthesis in soybean hypocotyl.** *Proc Natl Acad Sci U S A* 1980, **77**:357–361.

[127] Schuster A, Davies E: **Ribonucleic Acid and Protein Metabolism in Pea Epicotyls : III. Response to Auxin in Aged Tissue.** *Plant Physiol* 1983, **73**(3):822–827.

[128] Lincoln C, Britton JH, Estelle M: **Growth and development of the axr1 mutants of Arabidopsis.** *Plant Cell* 1990, **2**(11):1071–1080.

[129] Ruegger M, Dewey E, Gray WM, Hobbie L, Turner J, Estelle M: **The TIR1 protein of Arabidopsis functions in auxin response and is related to human SKP2 and yeast grr1p.** *Genes Dev* 1998, **12**(2):198–207.

[130] Dharmasiri N, Estelle M: **Auxin signaling and regulated protein degradation.** *Trends Plant Sci* 2004, **9**(6):302–308.

[131] Dharmasiri S, Estelle M: **The role of regulated protein degradation in auxin response.** *Plant Mol Biol* 2002, **49**(3-4):401–409.

[132] Ramos JA, Zenser N, Leyser O, Callis J: **Rapid degradation of auxin/indoleacetic acid proteins requires conserved amino acids of domain II and is proteasome dependent.** *Plant Cell* 2001, **13**(10):2349–2360.

[133] Zenser N, Dreher KA, Edwards SR, Callis J: **Acceleration of Aux/IAA proteolysis is specific for auxin and independent of AXR1.** *Plant J* 2003, **35**(3):285–294.

[134] Saniewski M, Ueda J, Miyamoto K: **Relationships between jasmonates and auxin in regulation of some physiological processes in higher plants**. *Acta Physiologiae Plantarum* 2002, **24**(2):211–220.

[135] Theologis A, Huynh TV, Davis RW: **Rapid induction of specific mRNAs by auxin in pea epicotyl tissue.** *J Mol Biol* 1985, **183**:53–68.

[136] Bell E, Mullet JE: **Lipoxygenase gene expression is modulated in plants by water deficit, wounding, and methyl jasmonate.** *Mol Gen Genet* 1991, **230**(3):456–462.

[137] DeWald D, Sadka A, Mullet J: **Sucrose Modulation of Soybean Vsp Gene Expression Is Inhibited by Auxin.** *Plant Physiol* 1994, **104**(2):439–444.

[138] Turner WB, Bidwell RG: **Rates of Photosynthesis in Attached and Detached Bean Leaves, and the Effect of Spraying with Indoleacetic Acid Solution.** *Plant Physiol* 1965, **40**(3):446–451.

[139] Bidwell R, Turner W: **Effect of Growth Regulators on CO(2) Assimilation in Leaves, and its Correlation with the Bud Break Response in Photosynthesis.** *Plant Physiol* 1966, **41**(2):267–270.

[140] Yin H: **Effect of auxin on Chlorella vulgaris and studies on the movement of leaves**. *PhD thesis*, California Institute of Technology 1937.

[141] Zerbe R, Wild A: **The effect of indole-3-acetic-acid on the photosynthetic apparatus of Sinapis alba**. *Photosynthesis Research* 1980, **1**(2):71–81.

[142] Khalil S, Mandurah HM: **Combined Effect of Soil Water Availability and Growth Substances on Some Aspects of Chemical Composition of Cowpea Plants**. *Journal of Agronomy and Crop Science* 1989, **162**(2):81–92.

[143] Zaffari GR, Peres LEP, Kerbauy GB: **Endogenous Levels of Cytokinins, Indoleacetic Acid, Abscisic Acid, and Pigments in Variegated Somaclones of Micropropagated Banana Leaves**. *Journal of Plant Growth Regulation* 1998, **17**(2):59–61.

[144] Rock CD, Sun X: **Crosstalk between ABA and auxin signaling pathways in roots of Arabidopsis thaliana (L.) Heynh.** *Planta* 2005, **222**:98–106.

[145] Angenent GC, Franken J, Busscher M, Weiss D, van Tunen AJ: **Co-suppression of the petunia homeotic gene fbp2 affects the identity of the generative meristem.** *Plant J* 1994, **5**:33–44.

[146] Samach A, Kohalmi SE, Motte P, Datla R, Haughn GW: **Divergence of function and regulation of class B floral organ identity genes.** *Plant Cell* 1997, **9**(4):559–570.

[147] Schwechheimer C, Zourelidou M, Bevan M: **PLANT TRANSCRIPTION FACTOR STUDIES.** *Annu Rev Plant Physiol Plant Mol Biol* 1998, **49**:127–150.

[148] Pena-Cortés H, Albrecht T, Prat S, Weiler EW, Willmitzer L: **Aspirin prevents wound-induced gene expression in tomato leaves by blocking jasmonic acid biosynthesis**. *Planta* 1993, **191**:123–128.

[149] Gupta V, Willits MG, Glazebrook J: **Arabidopsis thaliana EDS4 contributes to salicylic acid (SA)-dependent expression of defense responses: evidence for inhibition of jasmonic acid signaling by SA.** *Mol Plant Microbe Interact* 2000, **13**(5):503–511.

[150] Wang D, Amornsiripanitch N, Dong X: **A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants.** *PLoS Pathog* 2006, **2**(11):e123.

[151] Lin WH, Ye R, Ma H, Xu ZH, Xue HW: **DNA chip-based expression profile analysis indicates involvement of the phosphatidylinositol signaling pathway in multiple plant responses to hormone and abiotic treatments**. *Cell Res* 2004, **14**:34–45.

[152] Krinke O, Ruelland E, Valentova O, Vergnolle C, Renou JP, Taconnat L, Flemr M, Burketova L, Zachowski A: **Phosphatidylinositol 4-kinase activation is an early response to salicylic acid in Arabidopsis suspension cells.** *Plant Physiol* 2007, **144**(3):1347–1359.

[153] Mateo A, Funck D, Muhlenbock P, Kular B, Mullineaux PM, Karpinski S: **Controlled levels of salicylic acid are required for optimal photosynthesis and redox homeostasis**. *Journal of Experimental Botany* 2006, **57**(8):1795–1807.

[154] Raskin I: **Role of Salicylic Acid in Plants**. *Annual Review of Plant Physiology and Plant Molecular Biology* 1992, **43**:439–463.

[155] Khurana J, Cleland C: **Role of Salicylic Acid and Benzoic Acid in Flowering of a Photoperiod-Insensitive Strain, Lemna paucicostata LP6.** *Plant Physiol* 1992, **100**(3):1541–1546.

[156] Gordon LK, Minibayeva FV, Ogorodnikova TI, Rakhmatullina DF, Tzentzevitzky AN, Kolesnikov OP, Maksyutin DA, Valitova JN: **Salicylic acid induces dissipation of the proton gradient on the plant cell plasma membrane.** *Dokl Biol Sci* 2002, **387**:581–583.

[157] Kell DB: **Metabolomics and systems biology: making sense of the soup.** *Curr Opin Microbiol* 2004, **7**(3):296–307.

[158] Fiehn O: **Metabolomics–the link between genotypes and phenotypes.** *Plant Mol Biol* 2002, **48**(1-2):155–71.

[159] Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH, Trethewey RN, Lange BM, Wurtele ES, Sumner LW: **Potential of metabolomics as a functional genomics tool.** *Trends Plant Sci* 2004, **9**(9):418–25.

[160] Fiehn O, Kopka J, Trethewey RN, Willmitzer L: **Identification of uncommon plant metabolites based on calculation of elemental compositions using gas chromatography and quadrupole mass spectrometry.** *Anal Chem* 2000, **72**(15):3573–80.

[161] Weckwerth W, Fiehn O: **Can we discover novel pathways using metabolomic analysis?** *Curr Opin Biotechnol* 2002, **13**(2):156–60.

[162] Weckwerth W: **Metabolomics in systems biology.** *Annu Rev Plant Biol* 2003, **54**(1543-5008 (Print)):669–89.

[163] Allen J, Davey HM, Broadhurst D, Heald JK, Rowland JJ, Oliver SG, Kell DB: **High-throughput classification of yeast mutants for functional genomics using metabolic footprinting.** *Nat Biotechnol* 2003, **21**(6):692–696.

[164] Johnson HE, Broadhurst D, Goodacre R, Smith AR: **Metabolic fingerprinting of salt-stressed tomatoes**. *Phytochemistry* 2003, **62**(6):919–928.

[165] Dettmer K, Aronov PA, Hammock BD: **Mass spectrometry-based metabolomics.** *Mass Spectrom Rev* 2007, **26**:51–78.

[166] Dixon RA, Gang DR, Charlton AJ, Fiehn O, Kuiper HA, Reynolds TL, Tjeerdema RS, Jeffery EH, German JB, Ridley WP, Seiber JN: **Applications of Metabolomics in Agriculture.** *J Agric Food Chem* 2006, **54**(24):8984–8994.

[167] Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie A: **Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems.** *Plant Cell* 2001, **13**:11–29.

[168] Carrari F, Baxter C, Usadel B, Urbanczyk-Wochniak E, Zanor MI, Nunes-Nesi A, Nikiforova V, Centero D, Ratzka A, Pauly M, Sweetlove LJ, Fernie AR: **Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior.** *Plant Physiol* 2006, **142**(4):1380–96.

[169] Nikiforova VJ, Daub CO, Hesse H, Willmitzer L, Hoefgen R: **Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response**. *Journal of Experimental Botany* 2005, **56**(417):1887–1896.

[170] Fiehn O: **Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks**. *Comparative and Functional Genomics* 2001, **2**(3):155–168.

[171] Moritz T, Johansson AI: *Metabolomics, Metabonomics and Metabolite Profiling*, RSC Biomolecular Sciences 2007 chap. Plant Metabolomics.

[172] Shulaev V: **Metabolomics technology and bioinformatics.** *Brief Bioinform* 2006, **7**(2):128–39.

[173] Fiehn O: **Metabolite profiling in Arabidopsis.** *Methods Mol Biol* 2006, **323**(1064-3745 (Print)):439–47.

[174] Vandenheuvel FA, Court AS: **Refference high-efficiency nonpolar packed columns for the gas-liquid chromatography of nanogram amounts of steroids. II. Identification of steroids.** *J Chromatogr* 1969, **39**:1–16.

[175] Hall RD: **Plant metabolomics: from holistic hope, to hype, to hot topic.** *New Phytol* 2006, **169**(3):453–68.

[176] Hall R, Beale M, Fiehn O, Hardy N, Sumner L, Bino R: **Plant metabolomics: the missing link in functional genomics strategies.** *Plant Cell* 2002, **14**(7):1437–40.

[177] Gullberg J: **Metabolomics: a tool for studying plant biology**. *PhD thesis*, Swedish University of Agricultural Sciences 2005.

[178] Tolstikov VV, Fiehn O: **Analysis of highly polar compounds of plant origin: combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry.** *Anal Biochem* 2002, **301**(2):298–307.

[179] Tolstikov VV, Fiehn O, Tanaka N: **Application of liquid chromatography-mass spectrometry analysis in metabolomics: reversed-phase monolithic capillary chromatography and hydrophilic chromatography coupled to electrospray ionization-mass spectrometry.** *Methods Mol Biol* 2007, **358**:141–155.

[180] Wilson ID, Nicholson JK, Castro-Perez J, Granger JH, Johnson KA, Smith BW, Plumb RS: **High resolution "ultra performance" liquid chromatography coupled to oa-TOF mass spectrometry as a tool for differential metabolic pathway profiling in functional genomic studies.** *J Proteome Res* 2005, **4**(2):591–598.

[181] Wilson ID, Plumb R, Granger J, Major H, Williams R, Lenz EM: **HPLC-MS-based methods for the study of metabonomics.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, **817**:67–76.

[182] Bhalla R, Narasimhan K, Swarup S: **Metabolomics and its role in understanding cellular responses in plants.** *Plant Cell Rep* 2005, **24**(10):562–71.

[183] Mendes P: **Emerging bioinformatics for the metabolome.** *Brief Bioinform* 2002, **3**(2):134–45.

[184] Mendes P: **Metabolomics and the challenges ahead.** *Brief Bioinform* 2006, **7**(2):127.

[185] Younghoon K, Inho P, Doheon L: **Integrated Data Mining Strategy for Effective Metabolomic Data Analysis**. In *Optimization and Systems Biology: The First International Symposium on Optimization and Systems Biology*, Lecture Notes in Operations Research 7. Edited by Xiang-Sun Zhang LYW Luonan Chen, Wang Y 2007.

[186] van den Berg RA, Hoefsloot HCJ, Westerhuis JA, Smilde AK, van der Werf MJ: **Centering, scaling, and transformations: improving the biological information content of metabolomics data.** *BMC Genomics* 2006, **7**(1471-2164 (Electronic)):142.

[187] Shi H, Paolucci U, Vigneau-Callahan KE, Milbury PE, Matson WR, Kristal BS: **Development of biomarkers based on diet-dependent metabolic serotypes: practical issues in development of expert system-based classification models in metabolomic studies.** *OMICS* 2004, **8**(3):197–208.

[188] Jarvis RM, Goodacre R: **Genetic algorithm optimization for pre-processing and variable selection of spectroscopic data**. *Bioinformatics* 2005, **21**(7):860–868.

[189] Henriques IDS, Aga DS, Mendes P, O'Connor SK, Love NG: **Metabolic footprinting: a new approach to identify physiological changes in complex microbial communities upon exposure to toxic chemicals.** *Environ Sci Technol* 2007, **41**(11):3945–3951.

[190] Katajamaa M, Oresic M: **Processing methods for differential analysis of LC/MS profile data.** *BMC Bioinformatics* 2005, **6**:179.

[191] Katajamaa M, Miettinen J, Oresic M: **MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data**. *Bioinformatics* 2006, **22**(5):634–636.

[192] Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G: **XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification.** *Anal Chem* 2006, **78**(3):779–787.

[193] Dunn W, Broadhurst D, Brown M, Baker P, Redman C, Kenny L, Kell D: **Metabolic profiling of serum using Ultra Performance Liquid Chromatography and the LTQ-Orbitrap mass spectrometry system.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2008.

[194] Kind T, Tolstikov V, Fiehn O, Weiss RH: **A comprehensive urinary metabolomic approach for identifying kidney cancerr**. *Analytical Biochemistry* 2007, **363**(2):185–195.

[195] Margaria T, Kubczak C, Steffen B: **Bio-jETI: a service integration, design, and provisioning platform for orchestrated bioinformatics processes.** *BMC Bioinformatics* 2008, **9 Suppl 4**:S12.

[196] Nordstrom A, O'Maille G, Qin C, Siuzdak G: **Nonlinear data alignment for UPLC-MS and HPLC-MS based metabolomics: quantitative analysis of endogenous and exogenous metabolites in human serum.** *Anal Chem* 2006, **78**(10):3289–95.

[197] Nordstrom A, Want E, Northen T, Lehtio J, Siuzdak G: **Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics.** *Anal Chem* 2008, **80**(2):421–429.

[198] Wikoff WR, Gangoiti JA, Barshop BA, Siuzdak G: **Metabolomics identifies perturbations in human disorders of propionate metabolism.** *Clin Chem* 2007, **53**(12):2169–2176.

[199] Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L: **Technical advance: simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry.** *Plant J* 2000, **23**:131–142.

[200] Ellis DI, Broadhurst D, Kell DB, Rowland JJ, Goodacre R: **Rapid and quantitative detection of the microbial spoilage of meat by fourier transform infrared spectroscopy and machine learning.** *Appl Environ Microbiol* 2002, **68**(6):2822–2828.

[201] Villas-Boas SG, Rasmussen S, Lane GA: **Metabolomics or metabolite profiles?** *Trends Biotechnol* 2005, **23**(8):385–6.

[202] Yamaguchi S: **Gibberellin Metabolism and its Regulation.** *Annu Rev Plant Biol* 2008.

[203] Huang S, Raman AS, Ream JE, Fujiwara H, Cerny RE, Brown SM: **Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis.** *Plant Physiol* 1998, **118**(3):773–81.

[204] Kurosawa E: **Experimental studies on the nature of the substance secreted by the "bakanae" fungus**. *Nat Hist Soc Formosa* 1926, **16**:213–227.

[205] Bethke PC, Jones RL: **Gibberellin signaling.** *Curr Opin Plant Biol* 1998, **1**(5):440–6.

[206] Derkx MPM, Vermeer E, Karssen CM: **Gibberellins in seeds of Arabidopsis thaliana: biological activities, identification and effects of light and chilling on endogenous levels**. *Plant Growth Regulation* 1994, **V15**(3):223–234.

[207] Yamaguchi S, Sun Tp, Kawaide H, Kamiya Y: **The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberellin biosynthesis.** *Plant Physiol* 1998, **116**(4):1271–8.

[208] Gazzarrini S, McCourt P: **Cross-talk in plant hormone signalling: what Arabidopsis mutants are telling us.** *Ann Bot (Lond)* 2003, **91**(6):605–12.

[209] Bowman J: *Arabidopsis: An Atlas of Morphology and Development*. Springer-Verlag 1994.

[210] Maes M, Messens E: **Phenol as grinding material in RNA preparations.** *Nucleic Acids Res* 1992, **20**(16):4374.

[211] Ruuska SA, Ohlrogge JB: **Protocol for small-scale RNA isolation and transcriptional profiling of developing Arabidopsis seeds.** *Biotechniques* 2001, **31**(4):752, 754, 756–8.

[212] Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R: **When defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress.** *Plant Physiol* 2004, **134**(4):1683–96.

[213] Davletova S, Schlauch K, Coutu J, Mittler R: **The Zinc-Finger Protein Zat12 Plays a Central Role in Reactive Oxygen and Abiotic Stress Signaling in Arabidopsis**. *Plant Physiology* 2005, **139**(2):847–856.

[214] Xing S, Qin G, Shi Y, Ma Z, Chen Z, Gu H, Qu LJ: **GAMT2 Encodes a Methyltransferase of Gibberellic Acid That is Involved in Seed Maturation and Germination in Arabidopsis**. *Journal of Integrative Plant Biology* 2007, **49**(3):368–381.

[215] Hedden P, Kamiya Y: **GIBBERELLIN BIOSYNTHESIS: Enzymes, Genes and Their Regulation**. *Annual Review of Plant Physiology and Plant Molecular Biology* 1997, **48**:431–460.

[216] Olszewski N, Sun TP, Gubler F: **Gibberellin signaling: biosynthesis, catabolism, and response pathways.** *Plant Cell* 2002, **14 Suppl**(1040-4651 (Print)):S61–80.

[217] Thomas SG, Phillips AL, Hedden P: **Molecular cloning and functional expression of gibberellin 2- oxidases, multifunctional enzymes involved in gibberellin deactivation.** *Proc Natl Acad Sci U S A* 1999, **96**(8):4698–703.

[218] Haber AH, Carrier WL, Enochs NJ: **Gibberellin Action on Growth of Seedlings in the Absence of Photosynthesis**. *Nature* 1961, **190**(4780):1034–1035, [http://dx.doi.org/10.1038/1901034a0].

[219] Yuan L, Xu DQ: **Stimulation effect of gibberellic acid short-term treatment on leaf photosynthesis related to the increase in Rubisco content in broad bean and soybean**. *Photosynthesis Research* 2001, **68**:39–47.

[220] Biemelt S, Tschiersch H, Sonnewald U: **Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants.** *Plant Physiol* 2004, **135**:254–265.

[221] Loreti E, Povero G, Novi G, Solfanelli C, Alpi A, Perata P: **Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis.** *New Phytol* 2008.

[222] Jiang C, Gao X, Liao L, Harberd NP, Fu X: **Phosphate Starvation Root Architecture and Anthocyanin Accumulation Responses Are Modulated by the Gibberellin-DELLA Signaling Pathway in Arabidopsis**. *Plant Physiology* 2007, **145**(4):1460–1470.

[223] Moalem-Beno D, Tamari G, Leitner-Dagan Y, Borochov A, Weiss D: **Sugar-Dependent Gibberellin-Induced Chalcone Synthase Gene Expression in Petunia Corollas.** *Plant Physiol* 1997, **113**(2):419–424.

[224] Gibson SI: **Sugar and phytohormone response pathways: navigating a signalling network.** *J Exp Bot* 2004, **55**(395):253–64.

[225] Koch K: **CARBOHYDRATE-MODULATED GENE EXPRESSION IN PLANTS.** *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:509–540.

[226] Pego JV, Kortstee AJ, Huijser C, Smeekens SC: **Photosynthesis, sugars and the regulation of gene expression.** *J Exp Bot* 2000, **51 Spec No**:407–416.

[227] Smeekens S: **Sugar regulation of gene expression in plants.** *Curr Opin Plant Biol* 1998, **1**(3):230–234.

[228] Finkelstein RR, Lynch TJ: **Abscisic acid inhibition of radicle emergence but not seedling growth is suppressed by sugars.** *Plant Physiol* 2000, **122**(4):1179–1186.

[229] Gibson SI: **Control of plant development and gene expression by sugar signaling**. *Current Opinion in Plant Biology* 2005, **8**:93–102.

[230] Vettakkorumakankav N, Falk D, Saxena P, Fletcher R: **A crucial role for gibberellins in stress protection of plants**. *Plant Biol.* 1999, **40**.

[231] Sarkar S, Perras MR, Falk DE, Zhang R, Pharis RP, Austin Fletcher R: **Relationship between gibberellins, height, and stress tolerance in barley (Hordeum vulgare L.) seedlings**. *Plant Growth Regulation* 2004, **42**(2):125–135.

[232] Magome H, Yamaguchi S, Hanada A, Kamiya Y, Oda K: **dwarf and delayed-flowering 1, a novel Arabidopsis mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor.** *Plant J* 2004, **37**(5):720–9.

[233] Hu J, Mitchum MG, Barnaby N, Ayele BT, Ogawa M, Nam E, Lai WC, Hanada A, Alonso JM, Ecker JR, Swain SM, Yamaguchi S, Kamiya Y, Sun TP: **Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis.** *Plant Cell* 2008, **20**(2):320–336.

[234] Schomburg FM, Bizzell CM, Lee DJ, Zeevaart JAD, Amasino RM: **Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants.** *Plant Cell* 2003, **15**:151–63.

[235] Yamauchi T, Oyama N, Yamane H, Murofushi N, Schraudolf H, Pour M, Furber M, Mander LN: **Identification of Antheridiogens in Lygodium circinnatum and Lygodium flexuosum**. *Plant Physiology* 1996, **111**(3):741–745.

Appendix A

Supplementary Tables

Table A.1: Biological function Gene Ontology categories significantly up-regulated in JMT overexpressing Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0009628 response to abiotic stimulus	1.1E-29	GO:0009909 regulation of flower development	1.3E-06
GO:0032501 multicellular organismal process	2.6E-19	GO:0048519 negative regulation of biological process	1.3E-06
GO:0006950 response to stress	2.0E-15	GO:0007623 circadian rhythm	1.9E-06
GO:0048518 positive regulation of biological process	5.2E-15	GO:0006350 transcription	2.2E-06
GO:0048856 anatomical structure development	1.3E-14	GO:0009755 hormone-mediated signaling	2.5E-06
GO:0009266 response to temperature stimulus	1.9E-13	GO:0045449 regulation of transcription	3.7E-06
GO:0022414 reproductive process	2.5E-13	GO:0010468 regulation of gene expression	4.6E-06
GO:0009737 response to abscisic acid stimulus	3.7E-13	GO:0009893 positive regulation of metabolic process	1.0E-05
GO:0009409 response to cold	1.0E-12	GO:0009658 chloroplast organization and biogenesis	1.9E-05
GO:0048608 reproductive structure development	1.8E-12	GO:0005983 starch catabolic process	2.0E-05
GO:0009791 post-embryonic development	4.4E-12	GO:0031325 positive regulation of cellular metabolic process	2.7E-05
GO:0050793 regulation of developmental process	8.9E-12	GO:0009908 flower development	3.6E-05
GO:0009416 response to light stimulus	1.1E-11	GO:0010286 heat acclimation	4.3E-05
GO:0009719 response to endogenous stimulus	3.9E-11	GO:0007154 cell communication	6.7E-05
GO:0009314 response to radiation	3.9E-11	GO:0006333 chromatin assembly or disassembly	8.6E-05
GO:0006970 response to osmotic stress	4.4E-11	GO:0009269 response to desiccation	8.9E-05
GO:0050789 regulation of biological process	1.6E-10	GO:0065003 macromolecular complex assembly	9.1E-05
GO:0009651 response to salt stress	2.6E-10	GO:0006334 nucleosome assembly	1.1E-04
GO:0009414 response to water deprivation	4.0E-10	GO:0009738 abscisic acid mediated signaling	1.5E-04
GO:0009415 response to water	1.0E-09	GO:0006323 DNA packaging	1.8E-04
GO:0009725 response to hormone stimulus	1.2E-09	GO:0051094 positive regulation of developmental process	2.0E-04
GO:0009657 plastid organization and biogenesis	2.3E-09	GO:0006996 organelle organization and biogenesis	2.1E-04
Continued ...			

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0048522 positive regulation of cellular process	2.3E-09	GO:0046686 response to cadmium ion	2.3E-04
GO:0050794 regulation of cellular process	1.5E-08	GO:0045941 positive regulation of transcription	2.6E-04
GO:0009793 embryonic development ending in seed dormancy	1.1E-08	GO:0051093 negative regulation of developmental process	2.8E-04
GO:0042221 response to chemical stimulus	2.0E-08	GO:0007165 signal transduction	3.9E-04
GO:0019222 regulation of metabolic process	7.1E-08	GO:0009631 cold acclimation	4.4E-04
GO:0048316 seed development	8.5E-08	GO:0009734 auxin mediated signaling pathway	8.1E-04
GO:0031323 regulation of cellular metabolic process	8.7E-08	GO:0009605 response to external stimulus	1.0E-06
GO:0009639 response to red or far red light	2.5E-07	GO:0005982 starch metabolic process	1.1E-06

Table A.2: Biological function Gene Ontology categories significantly down-regulated in JMT overexpressing Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0050896 response to stimulus	8.8E-33	GO:0006955 immune response	4.8E-05
GO:0042221 response to chemical stimulus	8.7E-21	GO:0006979 response to oxidative stress	5.9E-05
GO:0009628 response to abiotic stimulus	2.4E-19	GO:0002376 immune system process	6.4E-05
GO:0051707 response to other organism	9.0E-15	GO:0006970 response to osmotic stress	6.6E-05
GO:0009607 response to biotic stimulus	2.0E-14	GO:0009651 response to salt stress	7.1E-05
GO:0051704 multi-organism process	6.5E-14	GO:0008652 amino acid biosynthetic process	7.5E-05
GO:0006950 response to stress	1.2E-13	GO:0019318 hexose metabolic process	8.3E-05
GO:0009266 response to temperature stimulus	2.6E-12	GO:0009737 response to abscisic acid stimulus	8.3E-05
GO:0009725 response to hormone stimulus	3.0E-11	GO:0007264 small GTPase mediated signal transduction	8.5E-05
GO:0007154 cell communication	3.8E-10	GO:0009812 flavonoid metabolic process	1.1E-04
GO:0007165 signal transduction	1.5E-09	GO:0044248 cellular catabolic process	1.1E-04
GO:0044262 cellular carbohydrate metabolic process	1.5E-09	GO:0005996 monosaccharide metabolic process	1.2E-04
GO:0009719 response to endogenous stimulus	6.8E-09	GO:0009813 flavonoid biosynthetic process	1.2E-04
GO:0009409 response to cold	8.3E-09	GO:0007169 transmembrane receptor protein tyrosine kinase signaling pathway	1.4E-04
GO:0009605 response to external stimulus	9.1E-09	GO:0007167 enzyme linked receptor protein signaling pathway	1.4E-04
GO:0006807 nitrogen compound metabolic process	1.8E-08	GO:0019319 hexose biosynthetic process	1.6E-04
GO:0006952 defense response	9.0E-08	GO:0051179 localization	1.6E-04
GO:0016051 carbohydrate biosynthetic process	1.3E-07	GO:0019748 secondary metabolic process	1.7E-04
GO:0006519 amino acid and derivative metabolic process	3.2E-07	GO:0009753 response to jasmonic acid stimulus	1.8E-04
GO:0009416 response to light stimulus	4.6E-07	GO:0009250 glucan biosynthetic process	1.9E-04
GO:0042440 pigment metabolic process	5.5E-07	GO:0009056 catabolic process	2.0E-04
GO:0007242 intracellular signaling cascade	5.7E-07	GO:0006575 amino acid derivative metabolic process	2.2E-04
GO:0009308 amine metabolic process	6.4E-07	GO:0046364 monosaccharide biosynthetic process	2.2E-04
GO:0005975 carbohydrate metabolic process	1.1E-06	GO:0055086 nucleobase, nucleoside and nucleotide metabolic process	3.0E-04
Continued ...			

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0009314 response to radiation	1.1E-06	GO:0009414 response to water deprivation	3.2E-04
GO:0009611 response to wounding	1.3E-06	GO:0009755 hormone-mediated signaling	3.2E-04
GO:0006790 sulfur metabolic process	1.3E-06	GO:0008361 regulation of cell size	3.2E-04
GO:0044271 nitrogen compound biosynthetic process	2.8E-06	GO:0009415 response to water	3.2E-04
GO:0019752 carboxylic acid metabolic process	3.1E-06	GO:0009620 response to fungus	3.8E-04
GO:0006082 organic acid metabolic process	3.2E-06	GO:0040007 growth	4.7E-04
GO:0007166 cell surface receptor linked signal transduction	5.4E-06	GO:0006816 calcium ion transport	5.5E-04
GO:0006073 glucan metabolic process	6.5E-06	GO:0030244 cellulose biosynthetic process	9.8E-04
GO:0009739 response to gibberellin stimulus	7.0E-06	GO:0009624 response to nematode	9.9E-04
GO:0009408 response to heat	1.0E-05	GO:0005976 polysaccharide metabolic process	3.8E-05
GO:0006520 amino acid metabolic process	1.6E-05	GO:0009309 amine biosynthetic process	6.2E-06
GO:0046148 pigment biosynthetic process	2.7E-05	GO:0045087 innate immune response	3.4E-05
GO:0044264 cellular polysaccharide metabolic process	4.3E-05		

Table A.3: Biological function Gene Ontology categories significantly up-regulated in IAMT overexpressing Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0009628 response to abiotic stimulus	1.29E-26	GO:0009751 response to salicylic acid stimulus	3.37E-05
GO:0009725 response to hormone stimulus	1.25E-24	GO:0009648 photoperiodism	4.08E-05
GO:0042221 response to chemical stimulus	1.11E-21	GO:0009605 response to external stimulus	4.50E-05
GO:0015979 photosynthesis	5.99E-21	GO:0007165 signal transduction	4.86E-05
GO:0009719 response to endogenous stimulus	1.03E-19	GO:0048573 photoperiodism, flowering	4.95E-05
GO:0009416 response to light stimulus	2.48E-18	GO:0010118 stomatal movement	6.12E-05
GO:0009314 response to radiation	1.26E-17	GO:0019685 photosynthesis, dark reaction	6.54E-05
GO:0009314 response to radiation	1.26E-17	GO:0042440 pigment metabolic process	1.08E-04
GO:0009733 response to auxin stimulus	3.25E-17	GO:0043467 regulation of generation of precursor metabolites and energy	1.43E-04
GO:0006970 response to osmotic stress	1.19E-12	GO:0051656 establishment of organelle localization	2.08E-04
GO:0009651 response to salt stress	1.38E-11	GO:0050793 regulation of developmental process	1.64E-04
GO:0006950 response to stress	9.72E-10	GO:0009734 auxin mediated signaling pathway	2.91E-04
GO:0009414 response to water deprivation	2.62E-09	GO:0009791 post-embryonic development	2.97E-04
GO:0046686 response to cadmium ion	2.69E-09	GO:0009637 response to blue light	3.43E-04
GO:0009737 response to abscisic acid stimulus	4.38E-09	GO:0009644 response to high light intensity	3.55E-04
GO:0009765 photosynthesis, light harvesting	4.94E-09	GO:0009638 phototropism	5.00E-04
GO:0009415 response to water	1.61E-08	GO:0019253 reductive pentose-phosphate cycle	5.00E-04
GO:0010038 response to metal ion	5.16E-08	GO:0051640 organelle localization	5.84E-04
GO:0009753 response to jasmonic acid stimulus	9.99E-08	GO:0006949 syncytium formation	6.45E-04
GO:0009755 hormone-mediated signaling	1.03E-07	GO:0044270 nitrogen compound catabolic process	6.48E-04
GO:0009723 response to ethylene stimulus	4.33E-07	GO:0019748 secondary metabolic process	6.63E-04
GO:0010035 response to inorganic substance	4.64E-07	GO:0009767 photosynthetic electron transport	7.22E-04
GO:0007242 intracellular signaling cascade	4.98E-07	GO:0010228 vegetative to reproductive phase transition	7.22E-04
Continued			

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0009739 response to gibberellin stimulus	5.90E-07	GO:0010119 regulation of stomatal movement	7.22E-04
GO:0009639 response to red or far red light	1.48E-06	GO:0006512 ubiquitin cycle	7.52E-04
GO:0018298 protein-chromophore linkage	2.58E-06	GO:0006766 vitamin metabolic process	9.62E-04
GO:0032501 multicellular organismal process	2.75E-06	GO:0007154 cell communication	3.33E-05
GO:0009642 response to light intensity	7.48E-06	GO:0032502 developmental process	2.88E-05
GO:0048856 anatomical structure development	1.69E-05		

Table A.4: Biological function Gene Ontology categories significantly down-regulated in IAMT overexpressing Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0006996 organelle organization and biogenesis	6.57E-17	GO:0005975 carbohydrate metabolic process	5.57E-05
GO:0044249 cellular biosynthetic process	4.03E-13	GO:0006457 protein folding	6.31E-05
GO:0022613 ribonucleoprotein complex biogenesis and assembly	2.08E-11	GO:0009793 embryonic development ending in seed dormancy	8.37E-05
GO:0006950 response to stress	3.08E-11	GO:0051707 response to other organism	8.55E-05
GO:0042254 ribosome biogenesis and assembly	3.17E-11	GO:0009695 jasmonic acid biosynthetic process	9.79E-05
GO:0009266 response to temperature stimulus	4.66E-10	GO:0016053 organic acid biosynthetic process	1.43E-04
GO:0019752 carboxylic acid metabolic process	7.13E-10	GO:0046394 carboxylic acid biosynthetic process	1.43E-04
GO:0006082 organic acid metabolic process	7.68E-10	GO:0016052 carbohydrate catabolic process	
GO:0009628 response to abiotic stimulus	2.74E-09	GO:0044275 cellular carbohydrate catabolic process	1.51E-04
GO:0016043 cellular component organization and biogenesis	3.59E-09	GO:0006519 amino acid and derivative metabolic process	1.62E-04
GO:0005982 starch metabolic process	6.33E-09	GO:0006325 establishment and/or maintenance of chromatin architecture	1.72E-04
GO:0009409 response to cold	7.33E-09	GO:0006323 DNA packaging	1.72E-04
GO:0009056 catabolic process	1.30E-08	GO:0048316 seed development	1.82E-04
GO:0009611 response to wounding	1.11E-07	GO:0019318 hexose metabolic process	1.95E-04
GO:0048856 anatomical structure development	1.26E-07	GO:0044264 cellular polysaccharide metabolic process	2.11E-04
GO:0000003 reproduction	1.47E-07	GO:0043648 dicarboxylic acid metabolic process	2.27E-04
GO:0044271 nitrogen compound biosynthetic process	1.89E-07	GO:0006412 translation	2.75E-04
GO:0009309 amine biosynthetic process	6.11E-07	GO:0006006 glucose metabolic process	2.82E-04
GO:0006631 fatty acid metabolic process	9.29E-07	GO:0009791 post-embryonic development	3.36E-04
GO:0006073 glucan metabolic process	1.34E-06	GO:0000910 cytokinesis	3.51E-04
GO:0031497 chromatin assembly	1.35E-06	GO:0045333 cellular respiration	3.92E-04
GO:0044265 cellular macromolecule catabolic process	1.46E-06	GO:0040007 growth	3.99E-04
GO:0009605 response to external stimulus	1.97E-06	GO:0046164 alcohol catabolic process	4.50E-04
GO:0008652 amino acid biosynthetic process	2.60E-06	GO:0022607 cellular component assembly	5.53E-04
GO:0009607 response to biotic stimulus	3.99E-06	GO:0009168 purine ribonucleoside monophosphate biosynthetic process	5.65E-04
Continued ...			

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0005983 starch catabolic process	3.99E-06	GO:0007166 cell surface receptor linked signal transduction	5.97E-04
GO:0031408 oxylipin biosynthetic process	5.18E-06	GO:0005976 polysaccharide metabolic process	6.18E-04
GO:0006633 fatty acid biosynthetic process	5.91E-06	GO:0006007 glucose catabolic process	6.52E-04
GO:0006807 nitrogen compound metabolic process	9.74E-06	GO:0030163 protein catabolic process	6.97E-04
GO:0009790 embryonic development	1.27E-05	GO:0055086 nucleobase, nucleoside and nucleotide metabolic process	6.97E-04
GO:0006334 nucleosome assembly	1.37E-05	GO:0006629 lipid metabolic process	6.98E-04
GO:0042221 response to chemical stimulus	1.40E-05	GO:0046365 monosaccharide catabolic process	7.13E-04
GO:0032787 monocarboxylic acid metabolic process	1.68E-05	GO:0019320 hexose catabolic process	7.13E-04
GO:0009861 jasmonic acid and ethylene-dependent systemic resistance	1.85E-05	GO:0044255 cellular lipid metabolic process	7.86E-04
GO:0005996 monosaccharide metabolic process	3.84E-05	GO:0009060 aerobic respiration	8.81E-04
GO:0009308 amine metabolic process	4.48E-05	GO:0051276 chromosome organization and biogenesis	9.95E-04

Table A.5: Biological function Gene Ontology categories significantly up-regulated in BSMT knockout Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0044249 cellular biosynthetic process	1.11E-10	GO:0043094 metabolic compound salvage	1.18E-04
GO:0009058 biosynthetic process	1.26E-10	GO:0009059 macromolecule biosynthetic process	1.90E-04
GO:0051186 cofactor metabolic process	3.22E-08	GO:0009057 macromolecule catabolic process	2.40E-04
GO:0006732 coenzyme metabolic process	1.51E-07	GO:0055086 nucleobase, nucleoside and nucleotide metabolic process	2.83E-04
GO:0019253 reductive pentose-phosphate cycle	9.14E-07	GO:0009409 response to cold	2.89E-04
GO:0009628 response to abiotic stimulus	1.46E-06	GO:0042254 ribosome biogenesis and assembly	3.09E-04
GO:0006066 alcohol metabolic process	2.59E-06	GO:0016052 carbohydrate catabolic process	3.12E-04
GO:0019685 photosynthesis, dark reaction	7.36E-06	GO:0044275 cellular carbohydrate catabolic process	3.12E-04
GO:0005996 monosaccharide metabolic process	3.84E-05	GO:0009790 embryonic development	3.53E-04
GO:0015977 carbon utilization by fixation of carbon dioxide	5.32E-05	GO:0009853 photorespiration	4.18E-04
GO:0015979 photosynthesis	5.39E-05	GO:0022613 ribonucleoprotein complex biogenesis and assembly	5.36E-04
GO:0006007 glucose catabolic process	6.11E-05	GO:0006412 translation	8.25E-04
GO:0051188 cofactor biosynthetic process	6.41E-05	GO:0050793 regulation of developmental process	8.80E-04
GO:0019320 hexose catabolic process	6.53E-05	GO:0009206 purine ribonucleoside triphosphate biosynthetic process	8.86E-04
GO:0046365 monosaccharide catabolic process	6.53E-05	GO:0009201 ribonucleoside triphosphate biosynthetic process	8.86E-04
GO:0009416 response to light stimulus	7.83E-05	GO:0009145 purine nucleoside triphosphate biosynthetic process	8.86E-04
GO:0046164 alcohol catabolic process	8.49E-05	GO:0006096 glycolysis	9.37E-04
GO:0019318 hexose metabolic process	9.12E-05	GO:0006739 NADP metabolic process	9.58E-04
GO:0044265 cellular macromolecule catabolic process	1.04E-04	GO:0006098 pentose-phosphate shunt	9.58E-04
GO:0006006 glucose metabolic process	1.09E-04	GO:0009314 response to radiation	1.16E-04

Table A.6: Biological function Gene Ontology categories significantly down-regulated in BSMT knockout Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0006950 response to stress	3.22E-18	GO:0000302 response to reactive oxygen species	1.24E-05
GO:0050896 response to stimulus	1.92E-15	GO:0009651 response to salt stress	1.73E-05
GO:0009611 response to wounding	7.81E-14	GO:0045087 innate immune response	5.94E-05
GO:0042221 response to chemical stimulus	3.96E-12	GO:0009607 response to biotic stimulus	7.69E-05
GO:0009415 response to water	1.17E-11	GO:0009719 response to endogenous stimulus	1.49E-04
GO:0009605 response to external stimulus	6.29E-11	GO:0006955 immune response	2.04E-04
GO:0009414 response to water deprivation	7.72E-11	GO:0005984 disaccharide metabolic process	2.08E-04
GO:0051704 multi-organism process	3.53E-10	GO:0002376 immune system process	2.34E-04
GO:0009861 jasmonic acid and ethylene-dependent systemic resistance	2.03E-09	GO:0009737 response to abscisic acid stimulus	2.97E-04
GO:0009628 response to abiotic stimulus	4.30E-08	GO:0019748 secondary metabolic process	3.95E-04
GO:0009695 jasmonic acid biosynthetic process	7.69E-08	GO:0007154 cell communication	4.74E-04
GO:0009694 jasmonic acid metabolic process	1.28E-07	GO:0042542 response to hydrogen peroxide	5.14E-04
GO:0051707 response to other organism	1.53E-07	GO:0007169 transmembrane receptor protein tyrosine kinase signaling pathway	7.90E-04
GO:0006979 response to oxidative stress	2.39E-07	GO:0007167 enzyme linked receptor protein signaling pathway	7.90E-04
GO:0031408 oxylipin biosynthetic process	5.64E-07	GO:0010035 response to inorganic substance	8.43E-04
GO:0031407 oxylipin metabolic process	7.52E-07	GO:0009725 response to hormone stimulus	8.79E-04
GO:0006970 response to osmotic stress	4.83E-06	GO:0009814 defense response, incompatible interaction	5.83E-06

Table A.7: Biological function Gene Ontology categories significantly up-regulated i*n gamt1-gamt2* double mutant Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0009628 response to abiotic stimulus	2.9E-16	GO:0009719 response to endogenous stimulus	2.1E-05
GO:0005982 starch metabolic process	4.4E-13	GO:0044264 polysaccharide metabolic process	2.4E-05
GO:0015979 photosynthesis	1.1E-11	GO:0009251 glucan catabolic process	2.5E-05
GO:0009409 response to cold	2.4E-11	GO:0009612 response to mechanical stimulus	3.0E-05
GO:0009266 response to temperature stimulus	3.5E-09	GO:0009607 response to biotic stimulus	3.9E-05
GO:0044262 carbohydrate metabolic process	6.9E-09	GO:0044248 cellular catabolic process	4.5E-05
GO:0006950 response to stress	1.6E-08	GO:0009765 photosynthesis, light harvesting	5.5E-05
GO:0042221 response to chemical stimulus	4.7E-08	GO:0009733 response to auxin stimulus	7.1E-05
GO:0009725 response to hormone stimulus	3.0E-07	GO:0006970 response to osmotic stress	9.1E-05
GO:0006073 glucan metabolic process	3.2E-07	GO:0008361 regulation of cell size	1.0E-04
GO:0005983 starch catabolic process	5.6E-07	GO:0016051 carbohydrate biosynthetic process	2.4E-04
GO:0015995 chlorophyll biosynthetic process	3.5E-06	GO:0006334 nucleosome assembly	2.6E-04
GO:0006996 organelle organization and biogenesis	3.6E-06	GO:0009314 response to radiation	3.1E-04
GO:0016052 carbohydrate catabolic process	4.0E-06	GO:0032989 cellular structure morphogenesis	3.2E-04
GO:0019685 photosynthesis, dark reaction	5.1E-06	GO:0009624 response to nematode	3.7E-04
GO:0009415 response to water	5.6E-06	GO:0008643 carbohydrate transport	3.8E-04
GO:0051707 response to other organism	6.3E-06	GO:0044265 macromolecule catabolic process	3.9E-04
GO:0009414 response to water deprivation	1.1E-05	GO:0019321 pentose metabolic process	6.5E-04
GO:0009411 response to UV	1.4E-05	GO:0009737 response to abscisic acid stimulus	7.0E-04
GO:0051258 protein polymerization	1.9E-05	GO:0019253 reductive pentose-phosphate cycle	8.1E-04
GO:0019684 photosynthesis, light reaction	1.9E-05	GO:0046907 intracellular transport	3.3E-04

Table A.8: Biological function Gene Ontology categories significantly down-regulated in *gamt1-gamt2* double mutant Arabidopsis plants. GO categories with a p-value $\leqq 0.001$ are reported.

Gene Ontology Category	***p*-value**	**Gene Ontology Category**	***p*-value**
GO:0042221 response to chemical stimulus	3.3E-11	GO:0000096 sulfur amino acid metabolic process	6.4E-05
GO:0019748 secondary metabolic process	3.7E-09	GO:0019344 cysteine biosynthetic process	1.5E-04
GO:0009628 response to abiotic stimulus	1.0E-08	GO:0055082 cellular chemical homeostasis	1.9E-04
GO:0007275 multicellular organismal development	1.5E-08	GO:0006873 cellular ion homeostasis	1.9E-04
GO:0032502 developmental process	2.4E-08	GO:0006534 cysteine metabolic process	2.0E-04
GO:0048856 anatomical structure development	5.4E-07	GO:0019725 cellular homeostasis	2.3E-04
GO:0048608 reproductive structure development	1.4E-06	GO:0010035 response to inorganic substance	3.1E-04
GO:0003006 reproductive developmental process	1.4E-06	GO:0006807 nitrogen compound metabolic process	3.4E-04
GO:0006790 sulfur metabolic process	1.4E-06	GO:0006979 response to oxidative stress	3.7E-04
GO:0022414 reproductive process	2.6E-06	GO:0006519 amino acid metabolic process	4.3E-04
GO:0019915 sequestering of lipid	5.2E-06	GO:0044271 nitrogen compound biosynthetic process	5.1E-04
GO:0009651 response to salt stress	8.9E-06	GO:0009790 embryonic development	5.7E-04
GO:0006082 organic acid metabolic process	1.1E-05	GO:0032787 monocarboxylic acid metabolic process	6.2E-04
GO:0044248 cellular catabolic process	1.7E-05	GO:0009636 response to toxin	9.3E-04
GO:0019752 carboxylic acid metabolic process	1.9E-05	GO:0006869 lipid transport	7.8E-04
GO:0016137 glycoside metabolic process	2.0E-05	GO:0044272 sulfur compound biosynthetic process	8.2E-04
GO:0006970 response to osmotic stress	2.2E-05	GO:0010038 response to metal ion	8.2E-04
GO:0044255 lipid metabolic process	2.3E-05	GO:0009793 embryonic development ending in seed dormancy	7.3E-04
GO:0009725 response to hormone stimulus	5.7E-05		
GO:0048316 seed development	6.2E-05		

Table A.9: Selected gibberellin responsive genes differentially expressed in *gamt1-gamt2* mutant, used for semi-quantitative RT-PCR analysis.

Locus Identifier	**Gene Name**	**Forward Primer**	**Reverse Primer**
AT5g56300	GAMT2	GCCGGAGTTTGAGGCTTTCTTCTG	TTAAACTCGGATGGCGGAAAAC
AT4g26420	GAMT1	TGGAAACGGTGGTTGAGGTGT	CCGGCTTGAGCATAGTTAGC
AT5G12250	β-tubulin	ACCACTCCTAGCTTTGGTGATCTG	AGGTTCACTGCGAGCTTCCTCA
AT3G18780	Actin2	TCCCTCAGCACATTCCAGCAGAT	AACGATTCCTGGACCTGCCTCATC
AT1G66350	RGL1 transcription factor	GCTCTGTGGTGGTTTTTGGAT	GGATGAGCTAAGAGGCGATG
AT1G75750	GASA1	TGATAAACAACCACTCAAAGGTAA	TGCATGAGTTTCGAACATGG
AT4G09600	GASA3	GCATGCTTTTTGAGACATGG	GGGTTGCTAAACTTACAGCAAAA
AT4G09610	GASA2	TGCATGAGTTTCGAACATGG	TTTTGTCCTCCAAAACATAGCA
AT5G14920	GA-regulated protein	AACACTTCCATCGCCATCTC	TCCACCACGTGTCTTCATGT
AT5G15230	GASA4	AACCAAAAACTCCTCTCTCAGTACA	CATTCATAGGGCCACTCTCAT
AT2G30810	GA-regulated protein	TGATGAAGCTCATAGTTGTCTTTG	TTGAAAAACAAGGATTACAAACATAG
AT1G78440	GA2ox1	CAGATGGGCTAGGGATCAAA	AAAAGATGGGAGCGATGTTG
AT5G51810	GA20ox2	TGACATTCCAGAGCTCAACG	TCGCGCTCTCTCTATTCACA
AT1G74670	GA-regulated protein	CAAGAGCTAGTCATGGCCAAA	TAACCATCACGACAACAAAAGC
AT1G15550	GA3ox GA4	CTTCGGGCTACCTGTCCAAC	TGTGGAAGAGGTCACCAACG
AT5G07200	GA20ox3	CTTAATCAGCACTCGCACCA	CCCCAAGACTCATTCCAAGA
AT1G30040	GA2ox2	GTCGTCAACCTAGCCGATCC	TCACCATCTTCTCCGCCTCT
AT5G58660	oxidoreductase, 2OG-Fe(II)	GGAGATATGATGCAGGCCAT	ATGGCCTGCATCATATCTCC
AT1G22690	GA-regulated protein	TGAATGTGGTGGCTTTTGTTAC	TAAGCTCACACATCACACATGC

Appendix B

Supplementary Figures

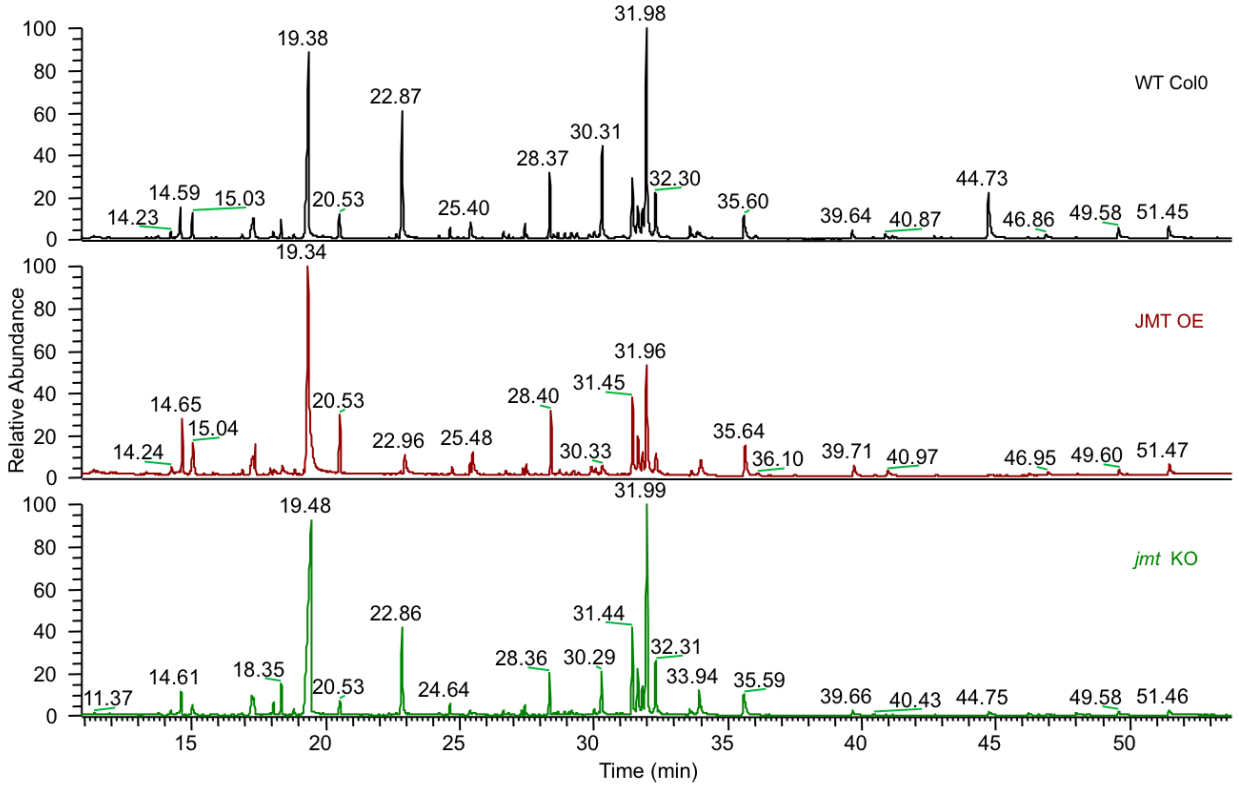


Figure B.1: GC profiles of polar extracts obtained from leaves of wild-type, JMT overexpressing, and *jmt* knockout.

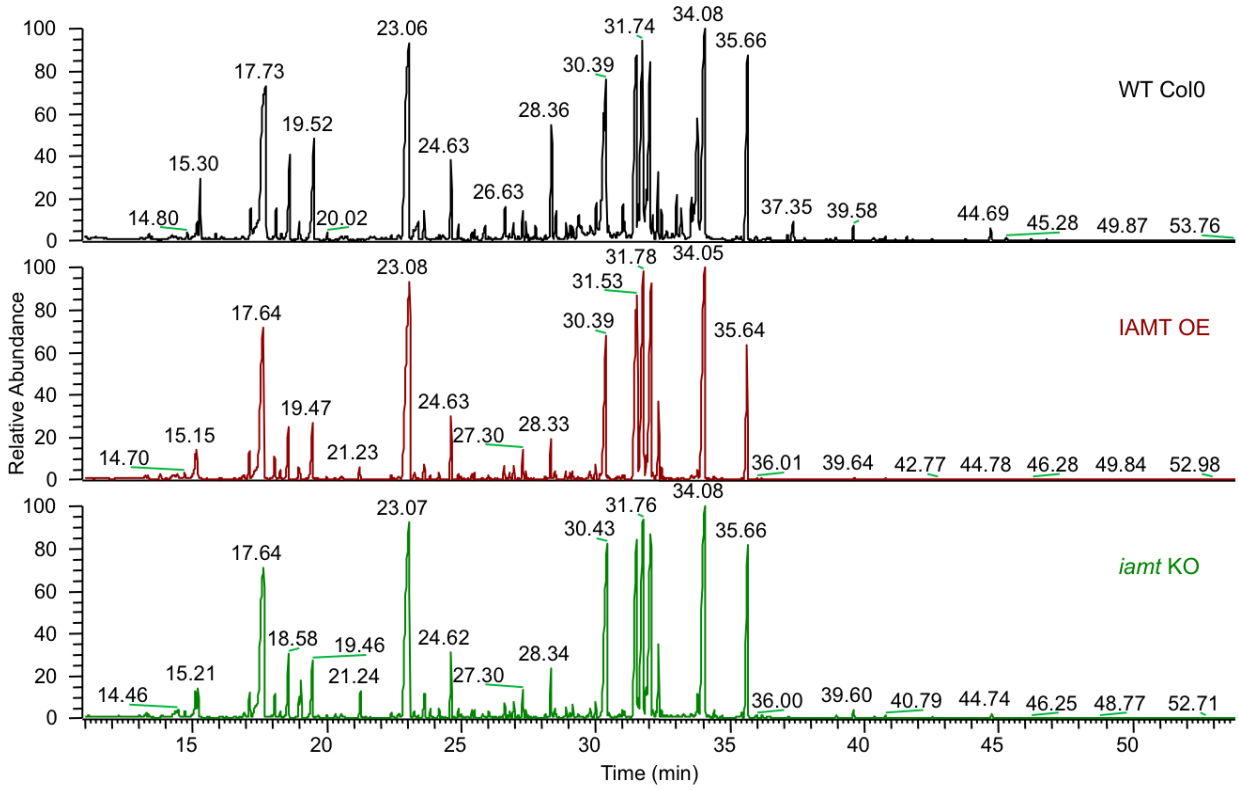


Figure B.2: GC profiles of polar extracts obtained from flowers of wild-type, IAMT overexpressing, and *iamt* knockout.

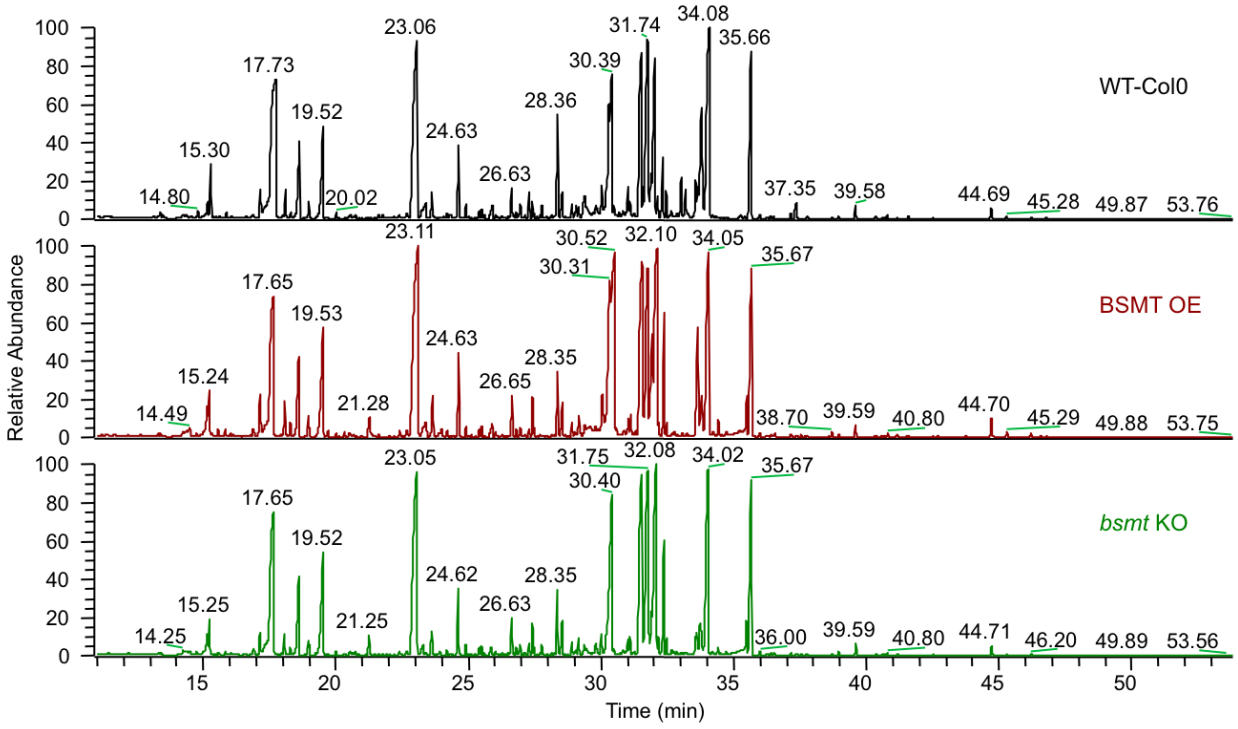


Figure B.3: GC profiles of polar extracts obtained from leaves of wild-type, BSMT overexpressing, and *bsmt* knockout.