MECHANISMS OF CROSS-TALK IN SECONDARY METABOLISM PATHWAYS IN *ARABIDOPSIS* REVEALED BY INTEGRATIVE FUNCTIONAL GENOMICS

AMIT RAI
(BACHELOR OF ENGINEERING IN BIOTECHNOLOGY)

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

DECEMBER 2012

DECLARATION

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

_____

Amit Rai

Date:

ACKNOWLEDGEMENT

This thesis is a reflection of hard work, dedication, patience and sacrifices of the past four years. I am indebted to many people for their guidance and support throughout this period without which this thesis would have been a mere dream.

First and foremost I am deeply grateful to my supervisor Professor Sanjay Swarup for showing his confidence in me and giving me an opportunity to work in his laboratory. He has always been patient and encouraging in times of new ideas and difficulties and for many motivational discussions. I appreciate all his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating.

It is my pleasure to thank Associate Professor Yu Hao and Professor Prakash Kumar for their insightful comments, valuable advice and discussions throughout my project. I was very fortunate as they were always there to help whenever I needed some suggestions. My work has being greatly benefited from suggestions and kind encouragement from them.

I would like to extend my gratitude to all my lab members who were always willing to help and give their best suggestions, particularly Wei Ling, Chui Ching, Gourvendu Saxena and Sheela Reuben for commenting and giving their best suggestion regarding my work as well as thesis, Umashankar Shivshankar for his invaluable help with all kind of bioinformatics analysis.

It is my pleasure to thank Ms Liew Chye Fong for her excellent support. I very much appreciate her enthusiasm and willingness to help with all lab-related technical issues.

My lunch time was made enjoyable in large part due to Martin and Petra that became a part of my life. I would like to thank them and I learned a lot from them about life, research, how to tackle new problems and how to develop techniques to solve them.

I gratefully acknowledge the funding sources that made my Ph.D. work possible. My work was supported by Research Scholarship funded by National University of Singapore. I am deeply appreciative for the administrative support of the office staff from the Department of Biological Sciences and special thanks to Ms. Reena Devi and Ms. Priscilla Li for being so helpful in graduate administration matters.

I would like to thank my family, specially my Dad for all their love and encouragement. They have always provided unwavering love, motivation and encouragement. Thanks for believing in me.

More importantly, I would like to thank my loving, supportive, encouraging, and patient wife Megha for her immense love. She was always there cheering me up and stood by me through the good as well as bad times. This would have not been reality without her emotional support.

Last but not least I would like to thank God for giving me the opportunity, strength and courage to fulfil my dream.

PUBLICATIONS

- Wen Cai Zhang, Ng Shyh-Chang, He Yang, **Amit Rai**, Shivshankar Umashankar, Siming Ma, Boon Seng Soh, Li Li Sun, Bee Choo Tai, Min En Nga, Kishore Kumar Bhakoo, Senthil Raja Jayapal, Massimo Nichane, Qiang Yu, Dokeu A. Ahmed, Christie Tan, Wong Poo Sing, John Tam, Agasthian Thirugananam, Monireh Soroush Noghabi, Yin Huei Pang, Haw Siang Ang, Wayne Mitchell, Paul Robson, Philipp Kaldis, Ross Andrew Soo, **Sanjay Swarup**, Elaine Hsuen Lim, Bing Lim Glycine Decarboxylase Activity Drives Non-Small Cell Lung Cancer Tumor-Initiating Cells and Tumorigenesis**.** Cell, 148: 259-272, 2012.

- Sheela Reuben**, Amit Rai**, Bhinu VS, Rodrigues Amrith, Yeoh Hock Hin, **Sanjay Swarup**. A Bacterial Quercetin Oxidoreductase QuoA-Mediated Perturbation in Phenylpropanoid Pathway Reveals Novel Properties of Metabolic Network in Arabidopsis. (ready for submission in December, 2012).

- **Amit Rai**, Shivshankar Umashankar, Narasimhan,K, **Sanjay Swarup**. Analysis of Arabidopsis Gene Expression and Gene Network based on Promoter Signature Revealed Regulome of Glycosylation (Manuscript in preparation).

BOOK CHAPTERS

- **Amit Rai**, Shivshankar Umashankar, **Sanjay Swarup**. Plant metabolomics and its integration with other functional genomics. Legume Genomics: Methods and Protocols. Human Press, USA (In preparation, to be submitted in 2013).

CONFERENCES

- **Amit Rai**, Sheela Reuben, **Swarup S**. Study of PhpA mediated metabolic perturbation in phenylpropanoid metabolic pathways. 14th Biological Sciences Graduate Congress, Bangkok. Chulalongkorn University: Bangkok, Thailand, 2009.Oral presentation

TABLE OF CONTENTS

CHAPTER 4: Results and discussion

Summary

Cellular metabolic pathways and their interconnected branches constitute metabolic networks. Knowledge of such networks provides close understanding of the biochemical phenotype of cells. Fluxes of metabolites through metabolic networks, therefore, represent, integrated responses of cells to extrinsic and intrinsic signals. Plant secondary metabolites, such as phenylpropanoids play a critical role in plant defence against biotic and abiotic stresses. Their biosynthetic pathways are highly interconnected to form a densely interconnected network. To understand the properties of the phenylpropanoid metabolic network and its association with other metabolic networks, we perturbed the flavonoid sub-network using genetic tools and studied the perturbational responses at both metabolite and gene expression levels.

Previously, our laboratory had characterized a bacterial oxidoreductase coding gene, *QuoA*, which converts the flavonoid, naringenin to quercetin. *QuoA* gene was previously isolated and expressed ectopically in transgenic *Arabidopsis* to effectively reverse the native biosynthesis of quercetin to naringenin. Here, three *QuoA* expressing transgeniclines were chosen from a screen of 16 independent transformed lines on the basis of transgene expression levels and enzymatic activity for further characterization. Lignins and anthocyanins had increased levels in QuoA lines, while metabolites from phenolamide pathway had reduced levels. These increased levels were corroborated with phenotype of increased stem stiffness and hypocotyl coloration. This is the first report showing accumulation of both lignins and anthocyanins at metabolite levels and phenotypes. Results here suggest that phenolamide pathway plays a major role in channelling of metabolic intermediates to the lignin pathway. Five branch points in phenylpropanoid metabolic network had

increased levels of first product with a concomitant decreased expression of the biosynthetic gene encoding the branch points enzyme. This indicated feedback controls at both biochemical and genetic levels at these branch points within the phenylpropanoid network.

Non-targeted metabolomics studies of QuoA transgenic lines revealed perturbational effects in seemingly distant pathways of isoprenoid metabolic network as compared to the point of perturbation in the flavonoid pathway. As these results cannot be explained solely by changes in metabolic channelling, we tested the hypothesis whether these distant pathways could be co-ordinately regulated by a common mechanism. *Arabidopsis* mutant lines of the transcriptional factor TTG1 singly or in combination with selected flavonoid *tt* mutant lines had a metabolic pheno-copy of QuoA lines, including the perturbation in isoprenoid metabolic network. This raised a question whether TTG1 could be a common transcription factor that is regulating both phenylpropanoid and isoprenoid networks?

To identify direct effects of TTG1 perturbation, we identified metabolic pathways responding to increased activity of TTG1 upon induction by dexamethasone (dex) in transgenic lines. *TTG1* over-expression resulted in accumulation of isoprenoids besides the expected increase in anthocyanins. As TTG1 complexes with TT8 to regulate anthocyanin genes, we also tested the effects of latter regulator. Transgenic lines with TT8 induction behaved somewhat similarbut with less pronounced effect than that in TTG1 induction lines. Further, TT8 loss-of-function and induced over-expression lines showed that TT8 regulates glycosylation of metabolites other than previously shown flavonoids. Glycosylation of several nucleotides and flavonoids

were up-regulated in dex-induced TT8 transgenic lines. Microarray analysis of TT8 loss-of-function and dex-induced TT8 transgenic lines showed several carbohydrate-active genes to be directly affected by TT8. Quantitative real-time PCR validated these results for selected genes.

After identifying the metabolic pathways directly affected by TTG1 and TT8, we next identified the genes directly affected by both these transcriptional factors. For this purpose, we first performed microarray profiling of RNA from inducible over-expression transgenic lines of TTG1 and TT8, respectively which were created in this study. Eleven genes from isoprenoid metabolic network were up-regulated in response to induced over-expression of TTG1 and eight were down-regulated in TTG1 loss-of-function mutant lines. TT8 induction led to up-regulation of five genes from isoprenoid metabolic network, but, the effect and number of genes affected were less pronounced than in TTG1 over-expression lines. Interestingly, several stress-responsive genes such as *PCC1*, *UGT75B1*, and *PDR12* were up-regulated by TTG1. These genes are involved in pathogen response mechanism. Next, we established that these changes in RNA levels of selected genes were due to transcript changes induced by TTG1 and TT8 induction by applying the translation inhibitor cycloheximide (chx) in combination with the TTG1 and TT8 inducing chemical, dex. Quantitative real-time PCR of 32 genes selected from the above treatments confirmed these results. Lastly, using chromatin immunoprecipitation assay, we were able to show that TTG1 direct binds to promoters of 10 of these 32 selected genes, thus established novel targets of TTG1.

Overall, this study has demonstrated the utility of a bacterial enzyme to reveal novel properties of secondary metabolism network, which can be used for metabolic

engineering of lignin pathways. We have also identified a common regulatory mechanism for stress-response pathways and distant secondary metabolic networks. Therefore, a longstanding question of how secondary network triggered during stress response seems to have been partially answered. Further studies involving this mechanism will be useful in better understanding of how plants co-ordinate their large-scale responses to various biotic and abiotic stresses.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µ- micron

µg- microgram

µl- microliter

4CL- 4-coumarate ligase

4PB-4-propynyloxybenzoic acid

AACT- acetoacetyl-CoA thiolase

ABA - abscissic acid

ABA1- zeaxanthin epoxidase

ALA- 5-aminolaevulinic acid

ANS - anthocyanidin synthase

Å- Angstrom

AT1G31910- phosphomevalonate kinase

AT2G26930- 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase

ATD2- glutamine 5-phosphoribosylpyrophosphate amidotransferase

ATP Adenine triphosphate

BAP1- 3-hydroxy-3-methylglutaryl coenzyme A synthase

bHLH- Basic helix- loop- helix

BLAST- The Basic Local Alignment Search Tool

BR6OX 1- brassinosteroid -6-oxidase

C- Centigrade

C3H- Coumarate 3-hydroxylase

C4H- Cinnamic acid 4-hydroxylase

CAS1- Cycloartenol synthase 1

CCR- cinnamoyl-CoA reductase

cDNA- complementary DNA

CDP-ME- 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol

CDP-ME2P- 2-Phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol

CHI - chalcone isomerase

CHS - chalcone synthase

CS- chlorophyll synthase

CHX- cycloheximide

CLA1- 1-deoxy-D-xylulose-5-phosphate synthase

CMK- 4-(cytidine 5’-diphospho)-2-C-methyl-D-erythritol kinase

CoA- Coenzyme A

CPD (CBB3), cytochrome P450 monooxygenase

DFR - dihydroflavonol reductase

DHNAPT- 1,4-dihydroxy-2-naphtoate phytyl transferase

DIG- Digoxigenin

DMAPP- dimethylallyl diphosphate

DNA- Deoxyribonucleic acid

DS- diterpenoid synthase

DXR- 1-deoxy-D-xylulose-5-phosphate reductoisomerase

DXS- 1-deoxy-D-xylulose-5-phosphate synthase

ESI-MS- Electrospray Ionisation Mass Spectrometry

F3’5’H - Flavonoid 3’, 5’ hydroxylase

F3’H- Flavonoid 3’monooxygenase

F3H- Flavanone-3- dehydroxylase

FAS - Ferrous Ammonium Sulphate

FLS - Flavonol synthase

FPS- farnesyl diphosphate synthase

FRET - Fluorescence energy resonance transfer

FTICR-MS -Fourier transform ion cyclotron resonance mass spectrometry

g- gram

GA-3P- D-glyceraldehyde 3-phosphate

GBA- 3-geranyl-4-hydroxybenzoate

GCEN- Gene co-expression network

GC-MS- Gas chromatography Mass Spectrometry

GcpE- 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase

GGPS- geranylgeranyl diphosphate synthase

GGPS1/GGPS6- geranylgeranyl pyrophosphate synthase

GGR- geranyl geranyl reductase

GO- Gene Ontology

GPP- Geranyl diphosphate

GPS- geranyl diphosphate synthase

GRN-Gene Regulatory Network

h- hour

HBPT- 4-hydroxybenzoate polyprenyltransferase

HCA - Hierarchical cluster analysis

HCHL- Enoyl-CoA hydratase

HCl- Hydrochloric acid

HCT - Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase

HDR- 1-hydroxy-2-methyl-2-butenyl 4-diphosphate reductase

HDS-1-hydroxy-2-methyl-2-butenyl 4-diphosphate synthase

HMG1/HMG2- 3-hydroxy-3-methylglutaryl coenzyme A reductase

HMGR- 3-hydroxy-3-methylglutaryl-CoA reductase

HMGS- 3-hydroxy-3-methylglutaryl-CoA synthase

HPLC- High Performance Liquid Chromatography

hp-light-hyper responsive high pigment

HPT- homogentisate phytyl transferase

HST- homogentisate solanesyl transferase

IAA- Indole acetic acid

IPP- isopentenyl pyrophosphate

IPP1/IPP2- isopentenyl diphosphate isomerase

IPT- isopentenyl transferase

IspF- 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

IspH- 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase

KEGG- Kyoto Encyclopedia of Genes and Genomes

LB- Luria Bertani

LDOX- leucoanthocyanidin dioxygenase

LUT1- Lutein-deficient 1

m/z- mass/charge

MCS- 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase

MCT- 2-C-methyl-D-erythritol 4- phosphate cytidyltransferase

MEP- 2-C-methyl-D-erythritol 4- phosphate

mg- milligram

min-minute

MK- mevalonate kinase

ml- milliliter

mm- millimeter

mM- millimolar

M- Molar

MN- Metabolic Network

MPDC- diphosphomevalonate Decarboxylase

mRNA- messenger RNA

MS agar – Murashige and Skoog agar

MS- Mass spectrometry

mTPS- monoterpenoid synthase

MS/MS- Tandem Mass Spectrometry (Fragmentation Mass Spectrometry)

MVA- mevalonate

MVAP- mevalonate-5-phosphate

MVAPP- mevalonate-5-diphosphate

MVD1- mevalonate diphosphate decarboxylase

MVK- mevalonate kinase

NAD- Nicotinamide adenine dinucleotide

NADME- NAD-dependent malic enzyme

nm- nanometer

NP- Natural products

OMT- Flavonol 3'-o-methyltransferase

OPPP - oxidative pentose phosphate pathway

OPS- oligoprenyl diphosphate synthase

P. putida –*Pseudomonas putida*

PAL- L-phenylalanine ammonia-lyase

PCA- Principal Component Analysis

Pca- protocatechuic acid

Pcd- protocatechuate 3, 4-dioxygenase

PGPR - Plant Growth Promoting Rhizobacteria

phpA- Phenylpropanoid catabolism gene A

PIP- Piperonylic acid

PMK- phosphomevalonate kinase

PMSF- Phenylmethylsulphonyl fluoride

pp- Phenylpropanoid

PPI- Protein protein interaction network

PPS- polyprenyl diphosphate synthase

PSY- phytoene synthase

qRT-PCR- Real-time Quantitative Polymerase Chain Reaction

RNA- Ribonucleic acid

rpm- rotation per minute

RT PCR- Reverse Transcriptase Polymerase Chain Reaction

s- seconds

S. cerevisiae- Saccharomyces cerevisiae

SA- Salicylic acid

SPE- Solid phase extraction

SPS- solanesyl diphosphate synthase

SQS- squalene synthase

SQS1- Squalene synthase 1

SS- sesquiterpenoid synthase

tAPX- thylakoid-bound ascorbate peroxidase

TK- Transketolase

tt- transparent testa

UDP- Uridine diphosphate

UPGMA- Unweighted Average distance

UV- Ultraviolet

w.r.t- with respect to

Wt Col –Wild- type *Arabidopsis* Colombia

Wt Ler- Wild- type *Arabidopsis* Landsberg erecta

Wt-Wild type

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Metabolites are the intermediates and end products of metabolic pathways. They are the closest link to biochemical phenotype of cells. Many of these metabolites additionally serve as (i) second messengers within cellular signalling system, (ii) ligands of DNA, mRNA and proteins, thereby altering their structures and functions or (iii) modes of inter-cellular or inter-organisimal communication. Hence, system level understanding of the response behaviour of a biological system in a particular environment or state is increasingly relying on knowledge of the metabolic content together with the knowledge of genome, gene and protein expression. While gene expression and proteome analyses by themselves do not provide a holistic understanding of cellular function, metabolic profiling can give an instantaneous snapshot of the biochemistry of that cell. Systems level approaches to interrogate genomic and proteomic state of an organism have been developed for longer time, while those for metabolite analyses have gained interest more recently. The study of metabolic state in a biological system is termed as metabolomics, which involves non-targeted measurements of changes in the relative concentrations of metabolites in response to perturbations. Metabolomics studies of a biological system in a particular state lead to description of the metabolites and pathways, which are critical to the adaptation and survival of the organism in that state. Among different biological systems, plants are particularly useful for metabolomics studies. They produce more

than 200,000 metabolites with large diversity, a lot of which are beneficial to mankind as food, medicines and industrial raw materials.

Metabolic pathways are defined as series of chemical reactions occurring within a cell with each step leading to production of a metabolite. Metabolic network is the complete set of interconnected metabolic pathways and physical processes that determine the biochemical phenotype of a cell, which includes metabolism and its regulation. Metabolic network is characterized by its plasticity, robustness, modularity and interconnectivity with several other metabolic network modules in a species. Regulatory and metabolic functions of biological systems are mediated by networks of interacting biochemical components. Although most specific functions involve a limited set of components, several studies have shown that connectivity among proteins or metabolites extends far beyond the limits of individual function.. Understanding of metabolic networks and its properties plays an important role in designing and engineering novel metabolic networks with improved metabolic processes that can increase product yield and enhance cellular properties. Studies on microbial metabolic network have played an exceptional role in improving *S. cerevisiae* strains for industrial applications such as baking, brewing, and wine making industries. Golden rice is another example of metabolic engineering application which leads to vitamin A biosynthesis, thus improving nutritional value of this important cereal grain. Although, metabolic engineering approaches have been extremely successful in microbial systems, it has not achieved expected success in engineering transgenic plants with desired features. Microbial metabolic systems are relatively simple compared to plant metabolic networks, which get further complicated by subcellular compartmentalization of various pathways and interaction

of different networks. Efforts to modify the biosynthesis of a variety of plant primary and secondary metabolites have been confounded by the very metabolic flexibility that researchers are trying to exploit.

Metabolic networks in general and plant metabolic networks in particular are highly complex. Components of metabolic network i.e. genome, gene expression, protein levels and metabolites all change in a synchronized manner in response to genetic or environmental perturbation. A number of studies have shown that perturbations in a particular metabolic network are not only confined to that metabolic network, but are extended to other networks as well. Such links between metabolic networks are referred to as crosstalk. Thus, it is very difficult to understand properties of metabolic network by a single systems level or 'omics' approach. Holistic approach to understand metabolic network is required to achieve desired success in plant metabolic engineering. Application of integrative omics approaches to study metabolic pathways and networks are being increasingly adopted in recent times. For example, gene-to-metabolite networks regulating sulphur and nitrogen nutrition and secondary metabolism in *Arabidopsis* have been investigated (Hirai et al., 2004). Transcriptome and metabolome analysis revealed that specific responses are induced when deficiency of sulphur or nitrogen is observed in several metabolic pathways. Glucosinolates, important bioactive secondary product, are found in the family *Brassicaceae.* Using integrative omics approach that combined transcriptome and metabolomics data, this poorly described metabolic pathway has been well-characterized.

Metabolic networks are connected to other pathways or network through several hub metabolites such as NADH or ATP. Interconnectivity in networks and existence of such hubs tends to affect several pathways at the same time in response to changes in specific biological system. Although such an interconnected design could in principle lead to the simultaneous flow of substrates in numerous directions, in practice metabolic fluxes pass through specific pathways. Moreover, the flux is readily optimized to maximize metabolic efficiency under different conditions. This suggests connectivity of network is not random but is specific and is under tight regulation. While effects of metabolic channelling on networks are better understood, recent genome-wide expression studies have revealed the prominent role of transcription in regulating metabolic flow in response to specific perturbations (Almaas et al., 2006). However, it remains unclear as to how and to what extent the modulation of enzyme expression determines functional metabolic units and how this affects the global properties of the metabolic network.

Perturbing is one of the approaches used to understand properties of metabolic network. Perturbations can be caused in uncover relationships that exist between different parts of the networks. While physiological or environmental perturbations can be used to study such relationships, they can simultaneously affect different points in multiple pathways. Genetic approaches, on the other hand, have been used to cause perturbations of specific steps in targeted metabolic pathways. These perturbations can be induced via loss-of-function approaches such as by gene deletions or silencing. In comparison to these approaches, expression of plant or bacterial enzymes in transgenic plants provides a means for perturbing metabolic pathways by gain-of-function. This approach has been effective in introducing new

functions to increase the accumulation of desired compounds or to reroute metabolites. Expression of bacterial enoyl-CoA hydratase/lyase, for example, results in rerouting in the phenylpropanoid pathway and leads to depletion of phenolics and accumulation of vanillic acid (Mayer et al., 2001). During perturbation of the isoprenoid pathway by the expression of a bacterial gene encoding 1-deoxy-D-xylulose-5-phosphate synthase in transgenic potato tubers, metabolic channelling of trans-zeatin riboside and carotenoids increases twofold and levels of phytoene increase six- to sevenfold (Morris et al., 2006).

Phenylpropanoids are secondary metabolites derived from phenylalanine. In plants, they play a critical role in signalling and defence against abiotic or biotic stresses. Phenylpropanoid-based polymers in plants, such as lignin, suberin, or condensed tannins, contribute to the physical stability and robustness towards environmental damages from drought or wounding. Metabolic intermediates from the upper phenylpropanoid biosynthesis pathway serve as starting point for multiple pathways connected by branches to form flavonoids, phytoalexins, anthocyanins, proanthocyanins, lignins and phenolamides. Together, all these pathways constitute the phenylpropanoid metabolic network. A plethora of phenylpropanoid biosynthetic mutants have facilitated the identification and functional characterization of genes and enzymes of the phenylpropanoid metabolic network. Among these, several enzymes of the flavonoid pathway, such as chalcone synthase (CHS), chalcone isomerase (CHI) and dihydroflavonol 4-reductase (DFR) form a complex in order to regulate the partitioning of intermediates among competing pathways and determine the intracellular deposition of end-products (Burbulis et al., 1999). These enzyme complexes are

present in the compartments, and play an important role in channelling of intermediates, limiting their diffusion into the surrounding milieu, which determines metabolic network properties under different conditions. Hence, keeping such complexes intact, while perturbing metabolic pathways in order to estimate network behaviour, would be an ideal approach. Given the highly complex nature of phenylpropanoid metabolic network and the importance of its constituent metabolites, it makes an ideal target to decipher the relationships among the pathways of this network. Phenylpropanoid pathways contribute to all aspects of plant responses to biotic and abiotic stimuli, which suggest possibility of its participation to a larger network with other secondary metabolic pathways coordinating towards all stresses.

We previously reported a *Pseudomonas putida* strain PML2 capable of utilizing quercetin as a sole carbon source for growth and elucidated a catabolic pathway for quercetin degradation in this strain using a comparative metabolomics approach (Pillai and Swarup, 2002). The first step in this pathway is the conversion of quercetin to naringenin involving dehydroxylation at two sites. The corresponding gene has been designated as Quercetin oxidoreductaseA (*QuoA)*, which is the new name of *PhpA* as reported by Pilli and Swarup (2002). Since, phenylpropanoid biosynthesis in plants involve formation of quercetin from naringenin, we envisaged that *QuoA* expression in plants will provide us with a genetic tool to "reverse" this biosynthetic step using gain-of-function approach and enable us to monitor the perturbational effects in metabolic networks. Thus, *QuoA* transgenic lines, and transparent testa mutant lines provides us a tool to understand network properties of phenylpropanoid metabolic network properties. Several studies reported gene-to-metabolites

relationship or regulation of genes in phenylpropanoid metabolic network, hence, providing a sound knowledge base in this network (Vogt, 2010). Previous studies in our lab on *QuoA* transgenic lines revealed several interesting properties of phenylpropanoid metabolic network (Reuben et al., 2012 (manuscript submitted)). This study is an extension of our previous work on understanding of metabolic perturbational response to gain-of-function *QuoA* transgenic lines. The overall aim of this study is to understand properties of the phenylpropanoid metabolic network, and global network response to its perturbation.

1.2 Objectives

Objective 1 To understand network properties of phenylpropanoid metabolic network using systems biology approach

To understand metabolic network properties, gain- and loss-of-function transgenic *Arabidopsis* lines were generated. These lines were used to obtain their metabolic and expression profiles at different developmental stages. We selected three *QuoA* transgenic lines, and compared with F3H loss-of-function mutant lines. We selected three independent *QuoA* transgenic lines based on high, moderate and low levels of gene expression and *QuoA* enzyme activity. Gain-of-function and loss-of-function transparent testa (*tt*) mutant lines were compared to understand network properties of phenylpropanoid metabolic network. Using multi-omics approaches, we investigated global effects in response to the above metabolic perturbations. This enabled us to understand effect of perturbation in a linear manner, and helped to understand local effect of perturbation in phenylpropanoid metabolic network.

Objective 2 To understand role of F3H and TTG1 loss-of-function in perturbational response of *Arabidopsis*.

Perturbational response of QuoA gain-of-function and *F3H* loss-of-function transgenic lines showed several metabolic pathways other than phenylpropanoid metabolic network were perturbed. Since point of perturbation was at phenylpropanoid metabolic network, we hypothesized that observed perturbational responses were due to QuoA activity and specific mutations in endogenous pathway would produce its phenocopy. To test this hypothesis, we compared metabolic profile of ten loss-of-function lines of phenylpropanoid metabolic network. Of these, *ttg1* and *tt6* mutant lines emerged as sharing similar profiles as that of *QuoA* transgenic lines.

Objective 3 To identify direct and indirect targets of TTG1 and its interacting partner TT8

Metabolic profiling of *ttg1* mutant and its double mutant with *tt3*, *tt4*, *tt5* and *tt6* mutant lines suggested that TTG1 could regulate metabolic pathways other than phenylpropanoid metabolic network. To study this, inducible over-expression and inducible silencing transgenic lines of TTG1 and its known interacting partner, TT8 were created. RNA and metabolic profiling were used to identify possible targets of TTG1 and TTG1-TT8 complex, which were verified using RT-PCR. ChIP experiment was performed to verify in-vivo binding of TTG1 to its target genes.

Objective 4 To study genetic regulation of glycosylation of secondary metabolites

Based on previous reports about TT8 regulating flavonoid glycosylation and this study that it regulates multiple metabolic pathways, we studied the role of TT8 in regulating glycosylation at a systems level. Microarray and metabolic profiling of *TT8*

loss-of-function and inducible over-expression line was performed to identify regulome and targets of glycosylation of secondary metabolites.

CHAPTER 2

LITERATURE REVIEW

2.1 Biology of plant metabolites

2.1.1 Metabolites and Metabolism

Metabolites are the products of enzyme-catalyzed reactions that occur naturally within cells. They are synthesized for both essential functions, such as growth and development by "primary metabolites", and for specialized functions, such as environmental response or defence against pathogens by "secondary metabolites". Since primary metabolites are directly involved in growth and development, they are present in all organisms including plants. But, absence of secondary metabolites can result in immediate as well as long-term impairment of the organism's survivability, fecundity, or aesthetics. Primary metabolites (e.g. amino acids, sugars, sugar phosphates, and organic acids) as well as secondary metabolites (e.g. phenylpropanoids) are components of the complex networks of biochemical pathways in plant metabolism. However, this classification is now considered obsolete because many secondary metabolites have essential functions and contribute towards plant growth and development (Hartmann, 2007). For instance, isoprenoid end products participate in a wide range of physiological processes acting in plants both synergistically, such as chlorophyll and carotenoids during photosynthesis and antagonistically, such as gibberellic acid and abscisic acid during seed germination. Metabolic profile of a biological system represents its overall environmental, developmental or growth condition which is specific to that state (Messerli et al., 2007). Specificity of metabolic profile or a metabolite to a tissue, developmental stage

or environmental condition emphasize important role that metabolites play in the response of the system to that state and its closest association with its phenotype (Messerli et al., 2007; Park et al., 2012). *Metabolism* is the set of chemical reactions that result in metabolite synthesis and plays an important role to sustain life. These processes allow organisms to grow and reproduce, maintain their structures, and respond to their environments. Metabolism allows optimized trade-offs among several requirements, such as defence response and adaptability via tolerance to various perturbations and evolvability in metabolite networks.

The diversity of metabolites found in plants is by far greater than in most other organisms. Metabolites in the plant kingdom are extremely diverse chemically, with estimates indicating as many as 200,000 different types of chemical substances (Riedelsheimer et al., 2012). The structural diversity of secondary metabolites is an important part of the nature of plants as a rich source of useful phytochemicals for humans. Different conjugation processes, such as glycosylation, acylation from a few parent molecules results in huge metabolic diversity of secondary metabolites. Comparison of the metabolite compositions of plant tissues investigated for specific secondary metabolites in *Arabidopsis* (Brown et al., 2003) and for primary metabolites in *Lotus japonicus* (Desbrosses et al., 2005) revealed that plants have evolved metabolic systems for producing a variety of metabolites in a tissue-dependent manner to improve plant fitness. Genome sequencing has uncovered the genetic background of the diversity of plant genomes, which encode large families of metabolism-related genes such as cytochrome P450s and glycosyltransferases (D'Auria and Gershenzon, 2005; Yonekura-Sakakibara and Saito, 2009).

2.1.2 Conjugation of metabolites

Conjugation in a biological system is a result of specialized enzymatic reaction, leading to modification of a metabolite by linking it to a molecule or a chemical group. Conjugation of parent metabolites changes its property, a strategy that biological systems use to regulate metabolic network by regulating the availability of metabolites. Conjugation reactions, such as the formation of sulphates, glycosylated forms, glucuronides and amino acid derivatives of biologically active compounds, are classically considered as detoxification processes. There are eight types of biochemical reactions that are involved as the process of conjugation namely glucuronidation, glycosylation, sulfation, methylation, acetylation, cyanide detoxication, glutathione conjugation and amino acid conjugation.

Conjugation not only leads to production of a large number of bio-active metabolites from a small number of precursor metabolites, this also works for detoxification of metabolites and regulation of biological system. In order to combat the phytotoxicity of xenobiotics, plants have developed versatile detoxification systems. Modification of xenobiotics through covalent linkage to glutathione is one of the important detoxification mechanisms. The resulting glutathione conjugates are exported from cytosol to vacuole by ATP-dependent tonoplast transporter (Coleman et al., 1997). This reaction is catalyzed by glutathione transferases (Dixon et al., 2010). The flavonoid skeletons are modified by glycosylation, malonylation, methylation, hydroxylation, acylation, prenylation, or polymerization, leading to the diversity of end-products (Winkel-Shirley, 2001). These substitutions have important effects on flavonoid function, availability, mobility, and degradation. Several metabolites, such as flavonoids are toxic in nature and thus are stored in conjugated form. Conjugation

of flavonoids, mainly glycosylation leads to conversion of flavonoids into less toxic molecules and in this form, flavonoids are stored. Deglycosylation takes place to make flavonoids available when it is required to the system. Transport of flavonoids takes place by conjugation of glutathione with flavonoids in the cytoplasm, followed by ATP-driven transport via glutathione S-transferase pumps (Mueller et al., 2000; Goodman et al., 2004). Glycosylation gives rise to nearly 4,000 flavonoid structures from nearly 30 common skeleton molecules (Heim et al., 2002). It changes the basic chemical properties of lipophilic acceptor metabolites thus affecting their subcellular mobility, stability, compartmentalization as well as bioactivity. Highly complex families of carbohydrate-active proteins, such as transporters and enzymes, are implicated in these processes. Recently, in a comprehensive study, a rhamnosyltransferase for flavonols was described in *Arabidopsis* and its gene placed in a pathway with other glycosyltransferases using metabolomics and transcriptomics approach (Yonekura-Sakakibara et al., 2007). However, role of other carbohydrate-active proteins such as transporters and hydrolases, aspects of coordinated regulation, use of mutants for pathway validation of metabolic pathways were not addressed. There are very few reports on regulators of glycosylation of metabolites.

Conjugations of metabolites are also used as a regulatory mechanism to control the availability and activity of bioactive compounds. Several reports on plant hormones have shown conjugation as an important tool to regulate biological processes. Cytokinins level depends on the balance between de novo synthesis, the import and export rate, interconversion of distinct forms, transient inactivation by conjugation (mainly glucosylation), and catabolic reactions resulting in a complete loss of biological activity (Sakakibara, 2006). Auxin conjugates are generally considered as

temporary reservoirs of inactive growth hormone Indole-3-acetic acid (IAA), releasing the free active hormone upon hydrolysis (Fluck et al., 2000; Jakubowska and Kowalczyk, 2005). Increased amount of auxin leads to its increased conjugation, and thus plants survive against high auxin level (Staswick et al., 2005; Jutta Ludwig-Muller, 2011). Experimental evidence connects auxin conjugates to the development of a body plan, vasculature, and embryo development (Sztein et al., 1999, 2000; Cooke et al., 2002). Work from Sztein et al. (1999, 2000) has revealed evidence for auxin conjugate formation in mosses (e.g. *Funaria, Polytrichum, Sphagnum*) and hornworts (*Phaeoceros*), whereas in liverworts (e.g. *Marchantia, Pallavicinia, Sphaerocarpus*) the auxin conjugate levels/synthesis of conjugates was very low and occurred at a slow rate. *Zea mays* regulates its growth by glycosylation of IAA.

2.1.3 Metabolome and Metabolomics

Metabolome refers to total metabolites present in a particular cell, tissue, organ or organism (Fiehn et al., 2001; Goodacre et al., 2004). It comprises the complete set of metabolites, the non-genetically encoded substrate, intermediates, and products of metabolic pathways, associated to a cell. *Metabolomics* is metabolome analysis, which seeks to identify and quantify the entire collection of intracellular and extracellular metabolites (Oliver et al., 1998; Tweeddale et al., 1998; Fiehn et al., 2001). The general aim of metabolomics is to identify, measure and interpret the complex time related concentration, activity and flux of metabolites in cells and tissues (Weckwerth et al., 2004; Saito and Matsuda, 2010).

Metabolomics involves a non-targeted analysis of metabolic profile of an organism in response to several factors such as abiotic stress or genetic manipulation, which then

works towards narrowing the gap in understanding gene functions in one form of functional genomics. One advantage of metabolomics as an approach for functional genomics is that it can be used on any biological system without knowing about their metabolome or genome (Macel et al., 2010). Metabolites are chemical compounds and thus applying principles of basic chemistry and analytical tools such as mass spectrometry enables one to study effect of genetic or environmental perturbation without knowing its genetic sequences. Finally, in contrast to transcripts or protein identification, metabolites are not organism-specific and thus same analytical protocol equally works to prokaryotes, fungi, plants and animal kingdom (Goodacre et al., 2007). The early biochemical changes that occur in any diseased condition can also be mapped by the metabolomics approach, which can be helpful in developing predictive biomarker (Kaddurah-Daouk et al., 2008).

The study of metabolomics has gained attention in the past few years. One of the earliest attempts of metabolomics was by Roger Williams in the late 1940s where he used paper chromatography to identify specific metabolic patterns in urine and saliva suggesting that all individuals have a particular metabolic profile of their biological fluids. This field has come a long distance by then and is evolving rapidly. While in 1999, three articles containing keywords metabolomics were published, the number increased to 203 articles in 2004 and now, it has further increased to 4,683. Whenever a perturbation is introduced in the living system, it responds by changing its behaviour on genetic level by altering gene expression as well as metabolic level by changing composition or abundance of metabolites. Evaluation of interaction between the level of metabolites, RNA and protein in response to applied perturbation result in generation of "omics" data. Metabolomics focuses on detection, identification and

quantification of metabolites present and hence has become an important tool to facilitate functional genomics. Metabolic profiling can be performed by targeted or non-targeted manner. Non-targeted metabolomics enables us to identify the broad metabolite profiles of samples and to find novel metabolites that can be used as biomarkers (Glinski and Weckwerth., 2006). However, many difficulties exist in achieving this goal, e.g. it is technically difficult and extremely time-consuming to merge all the data obtained in different formats by different instruments, and to identify the unknown metabolites (Werner et al., 2008). Targeted metabolomics is, in general, used to test a specific biological hypothesis by detecting and quantify analytes from biological system. Since this approach involves a small number of metabolites to capture, it has high sensitivity with high accuracy. Knowledge of metabolites and its chemical nature is prerequisite for this approach. Several metabolomics studies use combination of targeted and non-targeted approaches to identify novel metabolites.

2.1.4 Metabolic profiling and metabolomics technology platforms

Unlike transcripts and proteins, the identification of metabolites cannot be deduced from genomic information. Thus, the identification and quantification of metabolites must rely on sophisticated and highly sensitive instrumentation based methods such as Mass Spectrometry (MS), Nuclear Magnetic Resonance spectroscopy (NMR), and laser-induced fluorescence detection. Each of these instruments has its own advantages and limitations and selection of a particular technology depends on the nature of research questions. These instruments differ in sensitivity, selectivity, and speed apart from the principles these are based upon.

NMR spectroscopy is one of the most used techniques in structural analysis because of its highly selective and non-destructive approach (Quinones et al., 2009). It is generally accepted as the gold standard in metabolite structural elucidation, but it suffers from relatively low sensitivity (Dunn et al. 2005), which restricts metabolic coverage that can be achieved from it, and hence is not very useful in non-targeted approach of metabolomics. Due to high sensitivity, resolution and high throughput approach for identification and quantification of metabolite, MS is the most widely applied technology in both targeted and non-targeted metabolomics.

Mass spectrometry is an analytical technique that measures the mass-to-charge ratio of charged particles. Equipped with research and advancement of technology, modern MS provides highly specific chemical information directly related to the chemical structure such as accurate mass, isotope distribution patterns for elemental formula determination, and characteristic daughter ions by fragmentation for structural elucidation or identification via spectral matching to authentic compound data. High sensitivity of MS allows detection and measurement of primary and secondary metabolites even at picomole to femtomole levels, thus allows identifying even those metabolites which are present at very low levels (Bedair et al., 2008). Structural elucidation of metabolites is achieved by its fragmentation to daughter ions by tandem MS/MS technique and its comparison with fragmentation pattern of target or putative metabolite. Although, with development of several tools and software, one can predict putative structure and chemical formula of an unknown metabolite using fragmentation profile of a metabolite, NMR is still preferred for confirming chemical structure of an unknown metabolite.

Recent advancement in MS offers an array of technologies that differ in operational principles and performance. These variations include ionization technique, mass analyzer technology, resolving power, and mass accuracy. Since MS is based on ionization of metabolites and its detection against an electric field, ionization technology plays an important role in terms of kind of metabolites identified. The ionization techniques in metabolomics include electron ionization (EI), electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), chemical ionization (CI), Matrix-assisted laser desorption/ionization (MALDI), and, more recently, desorption ESI (DESI) and extractive ESI (EESI) have also been used. Each ionization technology has different range and nature of metabolites that can be ionized and selection of one is based on type of metabolites under investigation (Nordstrom et al., 2008). ESI is most used ionization technology for small molecules (primary and secondary metabolites) and covers wide range including polar and non-polar metabolites. Ionization of metabolites followed by separation of ionized metabolites based on their mass to charge ratio and its detection. Mass analyzers with different resolving powers have also been used in metabolomics. These include mass analyzers with low mass resolution such as ion-trap both linear (Q-trap) and three-dimensional (quadrupoles) or ultrahigh and high resolution MS such as Fourier transform ion cyclotron resonance MS (FT-ICR-MS), Orbitrap MS, and multipass TOF-MS. Each of these mass analyzers has its own advantages and limitations. Selection of a specific MS platform for metabolomics depends on the goal of the metabolomics projects, throughput, and instrumental costs. For instance, ion-trap MS with low mass accuracy are still used in targeted approach such as metabolite quantification (Multiple reaction monitoring (MRM) or Single ion monitoring approach (SIM)). Mass analyzers such as FT-ICR or orbitrap offers mass accuracy up-to 1ppm level. Such high mass

accuracy with high sensitivity makes FT-ICR-MS or orbitrap MS a valuable tool owing to its ability to analyze carbohydrates, amino acids, lipids and fatty acids as well as proteins and polysaccharides simultaneously (Harrigan et al., 2004).

MS-based metabolic profiling can be done either by directly injecting complex metabolite matrix in MS or separating metabolites by chromatography before injecting into MS. Direct injection approach produces less complex spectra but does provide a high throughput screening tool that is often the only practical choice for large sample numbers. Although, metabolites can be identified by directly injecting metabolites in MS, direct MS methods are very susceptible to ion suppression or enhancement. In addition, direct MS data interpretations are very challenging, as unique metabolite ions are difficult to distinguish from adduct and product ions. On the other hand, chromatography coupled to MS offers an excellent solution to complex mixture analyses and has been extensively used in metabolomics (Abian et al., 1999). Chromatographic separation of metabolites prior to MS analyses has several advantages such as it reduces matrix effects and ionization suppression, it separates isomers, provides additional and orthogonal data (i.e. retention time/factor/ index) valuable for metabolite annotation, and finally allows for more accurate quantification of individual metabolites. Gas chromatography (GC), liquid chromatography (LC), and capillary electrophoresis (CE) are three predominant chromatographic techniques that have been incorporated in MS-based metabolomics. Additionally, multidimensional separation techniques such as two-dimensional GC and LC have further enabled the separation of even more complex biological mixtures but these approaches are more applicable in proteomics.

GC-MS based approaches are suitable for the analyses of both volatile and non-volatile compounds following derivatization. Since GC based chromatography requires metabolites to be in gas phase, metabolites are required to be thermo-labile. However, relatively few compounds meet this requirement in their native state (e.g. short chain alcohols, acid, esters, and hydrocarbons (Broeckling et al., 2005; Begley et al., 2009). Plants are the source of a large number of volatiles metabolites that contribute to vegetable and fruit aromas and plant defence responses. This has made GC-MS as a popular choice for plant metabolomics (Hall et al., 2002; Bleeker et al., 2009; Mayer et al., 2008). However, the limitation as to the size and types of metabolites that can be analyzed and extensive preparation and derivatization required is a big concern.

In comparison with GC-MS, LC-MS uses mild heating and is suitable for both thermo-labile and thermo-stable metabolites. It thus gives better coverage and is most suitable for non-targeted metabolic profiling. More recently, there has been development of fast and more efficient ultra-HPLC (UHPLC), which utilizes higher column pressures (12,000–15,000 p.s.i. compared with 6000 p.s.i. for HPLC) and less than 2µm size of column packing particle. This has substantially increased chromatographic resolution, decreased the analysis time and peak capacity compared with HPLC. UHPLC has been shown to increase number of detectable compounds by 20% compared with HPLC or LC method (Nordstrom et al., 2008). Since changing HPLC or UHPLC ends systems and their running conditions changes orthogonal data of metabolites, the LC-MS results are not that reproducible across different labs and does not have extensive database as is the case in GC-MS. Hence, LC-MS based metabolomics uses high mass accuracy combined with MS/MS fragmentation for

identification of metabolites. For this reason, rather than retention time, mass spectrometers that provide high resolution combined with higher mass accuracy (HRAMS) are essential for metabolomics analysis.

2.2 Conceptual approaches to study metabolic pathways and network

Metabolic pathways consist of series of biochemical reactions that are connected by their intermediates, leading to production of specific metabolites. Metabolic pathways in an organism serve a specific purpose and are tightly regulated against any perturbation. Glycolysis, which serves to produce energy molecules in cells, was the first metabolic pathway discovered (Bordel et al., 2010). The metabolism encompasses all biochemical reactions that are catalyzed by enzymes in a cell. Taken together, these reactions form interconnected metabolic pathway that shape a dynamical circuitry referred to as metabolic network (Schilling et al., 2002; Reed et al., 2003). These networks are highly connected chains of metabolic pathways that encompass heterogeneous biological entities including DNA, mRNA, proteins and metabolites. Metabolic networks are a reflection of various biological processes that take place in the cells such as metabolism, transcription, translation and post-translational modifications. Thus, to interrogate a metabolic network, integrative omics approach needs to be adopted. The combination of several high-throughput technologies, such as high mass accuracy mass spectrometry, microarray, protein array, next generation DNA/RNA sequencing and others have generated huge dataset capturing capacity at the level of genes, proteins, metabolites, and the associated interactions among them (Zhang et al., 2010, Dietmair et al., 2012). These studies have resulted in the generation of large data sets that have served as the foundation for the reconstruction of metabolic regulatory, signalling and protein-protein interaction

networks (Klein et al., 2012). An important goal of building metabolic network is to integrate detailed biochemical information with biological intuition to produce testable predictions and thus apply it in metabolic engineering. Hence, the knowledge here in metabolic network is evolving with new information being added and hence predictability of outcome from a perturbation in a network is becoming more accurate (Zhao et al., 2006, Mithani et al., 2009, Handorf et al., 2005). Since functional states of metabolic networks that correspond to phenotypic functions are accessible, structural properties of networks can be accurately compared for their functional states.

Application of modelling approach such as constraint-based modelling (CBM) and relying on simple physical-chemical constraints and gene to metabolite relationships were used to create large-scale microbial metabolic networks. These have proven to be highly successful in predicting metabolic phenotypes in metabolic engineering and many other applications (Feist and Palsson, 2008; Henry et al., 2010). Construction of metabolic network models for multicellular eukaryotes is significantly more challenging than that for bacteria, because of the larger size of the networks, the subcellular compartmentalization of metabolic processes, and the considerable variation in tissue-specific metabolic activity (Shira Mintz-Oron, 2011). Plant metabolic networks are even more complex compared to other organismal models due to its metabolome size and high-level complexity because of extensive secondary metabolism. Subcellular compartmentalization of metabolic pathways, which share common substrate in plants further increases complexity of metabolic network. Several efforts on plants serving as factories for biomass in energy generation and

nutraceuticals are going on. However, much more has to be achieved in plant system as compared to what has been achieved in microbial metabolic engineering.

Currently, plant metabolic engineering mostly involves trial-and-error approaches (Shira Mintz-Oron, 2011), but studies are increasingly showing that perturbations due to genetic or environmental factors are not subjected to single but multiple pathways (Kooke and Keurentjes, 2011). Systems biology approach and integration of multi-omics data will help to build pathways-to-pathways and metabolic network-to-network relationships. Knowledge of metabolic networks, their interaction and regulation will improve our understanding of biological processes that take place in plants. This would be instrumental for developing strategies for metabolic engineering, eventually leading to synthetic biology applications (See section 2.6.4).

2.3 Properties of metabolic networks

2.3.1 Connectivity of metabolic networks

Metabolic networks in a biological system are characterized by high connectivity with metabolites serving as common substrate to a number of different processes. Traditionally, metabolism has been divided into discrete pathways. However, it has become increasingly clear that metabolism operates as a highly integrated network (Sweetlove et al., 2008). Metabolites undergo several processes, such as conjugation to produce various bio-active metabolites or in regulatory process in a biological system, thus increasing interconnectivity between metabolic networks. Presence of isoenzymes further increases complexity of metabolic network as network possesses alternative pathways for the same functions. In most of the metabolic networks, about half of the metabolites can be only converted to a very limited number (usually less

than 10) of metabolites (Hong et al., 2003). It was shown that the probability of a path existing between randomly chosen substrate and product pair is less than 20% which indicated that a metabolic network is far from a fully connected network. Metabolic networks characteristically have some key metabolites such as NADH or ATP, which are involved in a number of reactions and are referred to as hub metabolites (Pfeiffer et al., 2005). Group transfer reactions are involved in production of hub metabolites and these hubs participate in the transfer of different biochemical groups. These hub metabolites increase connectivity between metabolic networks. Several studies on metabolic pathways regulation in Arabidopsis has shown multiple pathways affected in response to perturbational effect. Studies on salicylic acid (SA) and jasmonic acid (JA) crosstalk have shown mutual antagonistic interaction between SA and JA dependent signalling, while in a few cases, synergistic interactions have been described as well (Schenk et al., 2000; Van Wees et al., 2000; Mur et al., 2006). Such cross-talk of signalling process can occur through connectivity of metabolic network. Comprehensive metabolic shift of reproductive transitions was explained by observed feedback mechanisms to optimally channel resources, suggesting strong interconnectivity of metabolic network in *Artemisia annua* (Arsenault et al., 2010).

2.3.2 Topology of gene expression networks

Topology of a metabolic network is a schematic description of the arrangement of a network, including its nodes and connecting lines. The nature of the links between nodes determines the topology of a network. It reflects dynamic nature of metabolic network and evolution of metabolic process (Wagner et al., 2001). Although many biological networks are described as being scale free, i.e. does not depend upon length scale at which the network is observed, networks with hierarchical structures that

allow hubs and module structures within the network currently seem to be best suited for capturing most of the features of biological networks (Barabasi et al., 2004). The global topological properties and local structural characteristics, network motifs, network comparison approach and query, detection of functional modules, function prediction from network analysis, inferring molecular networks from biological data as well as representative databases and software tools has been reviewed with respect to molecular network (Zhang et al., 2007). The structure and properties of metabolic network especially the topology have been described in detail in a book by Palsson (2006).

2.3.3 Plasticity of metabolic networks

Plasticity of a system or a network may be defined as its ability to tolerate alterations in its components or structure and it cannot be directly measured (Dalmolin et al., 2011). Every metabolic network has a fundamental property of being robust that allows them to adapt to the perturbations introduced by environment or genes. Metabolic networks exhibit robustness even under genetic perturbation caused by elimination of enzyme-coding genes. Their structure evolves rapidly through mutations which results in elimination of such genes and addition of new enzyme-coding genes through horizontal gene transfer. Single mutation introduced in two different genotypes having similar phenotype yielded different novel phenotype (Rodrigues et al. 2009). This implicated the application of evolutionary plasticity and robustness of metabolic networks in the emergence of new metabolic abilities.

Plasticity of metabolic network is also related to importance of particular pathways for the biological system. Adaptability of metabolic networks to perform the functions

by a number of alternative pathways increases the survivability of the organisms under diverse environmental conditions. In general, primary metabolic pathways are more robust to the changes, as they are important and directly related to growth and survival of the system and hence are under tight regulatory control (Hart et al., 2011; Stitt et al., 2010). Secondary metabolic networks, on the other hand, do not have tight regulatory systems. This comparison underlines role of evolutionary processes to the plasticity of metabolic network. One component of this robustness is redundancy mediated by isoenzymes performing the same catalytic reaction (Skinner and Cooper 1974; Primerano and Burns 1983), multiple pathways or routes that generate a particular metabolite (Ramos et al. 2008), and systems that maintain homeostasis of the cellular environment (Takahashi et al.,2002).

An in vivo genetic analysis demonstrated that thiamine synthesis can be uncoupled from the early purine biosynthetic pathway in *Salmonella enterica* (Bazurto et al., 2011). The study provided evidence that the five steps of the common HMP/purine biosynthetic pathway can be bypassed in the synthesis of 5-aminoimidazole ribonucleotide. Robustness of a regulatory network in any organism decreases when they switch to stable environment (Soyer et al., 2010). Computer simulation depicting metabolic network evolution was performed under stable and fluctuating conditions. The study showed that those networks that evolved in fluctuating environment were able to tolerate the effect of single gene mutation in specific environment more efficiently than the networks evolving in stable environment. When the evolution of those networks was examined in stable environment, the robustness was found to be completely lost. These studies thus suggested evolution as major player towards plasticity of a metabolic network in an organism. Investigations of functional

plasticity of reconstructed metabolic networks in *Helicobacter pylori, Escherichia coli, and Saccharomyces cerevisiae* by computational method of flux-based analysis revealed a set of metabolic reactions forming a connected metabolic core and carrying non-zero fluxes under all growth conditions and had highly correlated flux variations (Almaas et al., 2005). It was further shown that the large numbers of species-specific reactions of cellular metabolism which are indispensable for the growth of microorganisms were organized around that evolutionary active metabolic core. The effect of available oxygen on central metabolism with different concentration of available oxygen to *Arabidopsis* suspension culture has been studied, which revealed stability of the central metabolic network and its capacity to respond to physiological perturbations with the minimum of rearrangement of metabolic processes. Although steady-state metabolic flux analysis (MFA) using [1- ^{13}C]Glc showed differences in the metabolite profile of organic acids, amino acids, and carbohydrates at different O_2 levels, balance between biosynthesis and respiration within the tricarboxylic acid cycle (TCA) was unchanged, suggesting stability of central metabolic network (Williams et al., 2008). Due to high plasticity of metabolic network, knocking out a gene in an organism often causes no phenotypic affect. One possible explanation could be existence of duplicate genes due to which loss of one is compensated by its homologs. The *Arabidopsis* genes, *TT8, GL2* and *EGL3* have been shown to be able to replace and complement each other. Thus, loss-of-function of any one of these transcriptional factors does not bring distinct changes in metabolic network (Zhang et al., 2003, Morohashi et al., 2007). Another explanation for this could be the existence of alternative pathways. Study on functional compensation of primary and secondary metabolites by duplicate genes in *Arabidopsis* concluded existence of alternative

pathways and high connectivity of metabolic network as main cause of plasticity of metabolic network (Hanada et al., 2011).

2.3.4 Modularity in metabolic network

The concept of modularity stands for independent and self-contained property of a system. Biological systems often display modularity, in the sense that they can be decomposed into nearly independent subsystems, which can perform specific functions. Studies have suggested that modular structure of a network emerges in response to change in environmental conditions, since a small part of metabolic network is needed to establish homeostasis against the applied changes. This property of the metabolic system is also supported by the fact that different biological systems are unique only in small part of metabolic system whileshares great similarity. It is suggested that, during evolution, only a subset of metabolic network changed in biological system. Parter et al., (2007) have supported this theory by their study on the modularity of networks. The authors observed that the metabolic networks of organisms in variable environments were significantly more modular than networks of organisms that evolved under more constant conditions. Metabolites that participate in only a few reactions but connect different modules are more conserved than hubs whose links are mostly within a single module. Some other studies suggested evolution of modularity of metabolic network was in response to need for biological system to accommodate different niches (Kashtan et al., 2007; Parter et al., 2007; Kreimer et al., 2008).

Comparison of metabolic network in hundreds of bacterial species showed modularity to be moderately concordant with organismal phylogeny along the tree of life

(Kreimer et al., 2008). This study also showed network size as a key determinant of metabolic network modularity. Endosymbiotic organisms, which have smaller network, had lower modularity scores than non-symbiotic organism. The modularity values of pathogens and commensals were as low as or lower than those of endosymbionts. Analysis of *Homo sapiens* metabolic network revealed a defined topology based on functional and evolutionary modularity (Zhao et al., 2007). The analysis of the metabolic network showed that it had core modules that performed basic metabolic functions and peripheral modules that were highly modular and performed specialized functions. Isoprenoid metabolic network in *Arabidopsis* presents an excellent example of network modularity where different branches such as carotenoids, brassinosteroids metabolic pathways exist independently despite being connected at branch points. Isoprenoid metabolic network has two alternative routes for synthesis of metabolites due to subcellular compartmentalization of metabolic pathways (Vranova et al., 2012). In *Arabidopsis,* both mevalonate and non-mevalonate pathways operate independently with mevalonate pathway being associated with synthesis of sterols, sesquiterpenes and the side chain of ubiquinone while non-mevalonate pathway for synthesis of isoprenes, carotenoids, side chains of chlorophyll and plastoquinone. But, studies have established interaction between them (Wille et al., 2004).

2.4 Experimental approaches to study metabolic network

Metabolic network of biochemical pathways are studied by introducing perturbations through environmental stress, chemical inhibitions or mutations targeted to specific steps in the pathways under study. In order to overcome the introduced changes, the metabolic networks self-organize and re-accommodate to

allow the organism to perform stablely and reliably (Moran et al., 2005). This process of reorganization and adaptability is captured using metabolic profiling and flux measurement, expressional profiling as well as biochemical analysis of enzymes related to the pathways. The integration of these studies establishes metabolic network. Among several approaches to introduce perturbation in metabolic pathways, genetic approach is considered as a more directed approach towards studies of a particular pathway. Genetic perturbation can be achieved by either gain- of-function or loss-of-function approach. In gain-of-function approach, a foreign gene or endogenous gene related to the pathways is introduced in the target pathways in such a way that the expression of the gene uses metabolite from the target pathways as its substrate. Overexpression lines are an example of gain-of-function approach. In loss-of-function approach, mutation or silencing of endogenous gene from metabolic pathway is created through transposon mutagenesis, chemical inhibitors or using RNAi or amiRNA approaches. Metabolic perturbation can also be achieved by applying biotic such as pathogens or abiotic stress such as UV radiation or pathogenic stress, which causes metabolic pathways to respond to overcome this environmental stress. Gain-of-function approach together with loss-of-function approach has been used to characterize *PDR9* gene in *Arabidopsis* (Ito et al., 2006). In this study, gain-of-function mutant exhibited increased tolerance to 2,4-dichlorophenoxyacetic (2,4-D) acid while loss-of-function showing hypersensitivity to it suggesting role of PDR9 transporter in effluxing 2,4-D out of plant cells without having any effect on endogeneous auxin transport. Loss-of-function using chemical inhibitors lovastatin (inhibitor of mevalonate pathways) and formidomycin (inhibitor of mevalonate-independent pathways) showed cross-talk between cytosolic and plastidic pathways of

isoprenoid biosynthesis in *Arabidopsis thaliana* (Laule et al, 2003). Gain-of-function approach used for functional study of MYB-like transcription factor *PAP1* gene. The changes in the metabolic profiles caused by *PAP1* over-expression lines showed increased accumulation of anthocyanins. This was combined with gene expression studies which revealed expression of genes known to be involved in anthocyanin production was up-regulated and thus other up-regulated genes could be tentatively assigned a role in the formation of anthocyanins, such as members of the glycosyltransferase, acyltransferase and glutathione S-transferase families.

Expression of bacterial gene in plant system have also being used as an important tool to introduce perturbation in metabolic network and to genetically engineer the plant for enhanced growth, resistance and for over-production of desired products. It also helps in identifying regulatory mechanism and interaction among primary and secondary metabolism. In plants, arogenate is used as precursor to synthesize phenylalanine. To identify any alternative pathway for the production of phenylalanine, a bacterial bi-functional *PheA* (chorismate mutase/prephenate dehydratase) gene involved in conversion of chorismate via prephenate into phenylpyruvate has been expressed in *Arabidopsis thaliana* (Tzin et al., 2009). Transgenic lines expressing PheA transgene was reported to have increased level of phenylalanine suggesting the conversion of phenylpyruvate into phenylalanine. Also, the PheA-expressing lines were seen to be more sensitive towards 5-methyl-Trp, a tryptophan biosynthesis inhibitor. This suggests that chorismate is the common precursor for the synthesis of phenylalanine and tryptophan. Other primary metabolites and transcriptome profiling was minimally affected by the increase in phenylalanine level but the secondary metabolites derived from

aromatic amino acids were altered. This suggests that minor regulatory cross-interaction exists between the aromatic amino acid biosynthesis network and the *Arabidopsis* transcriptome and primary metabolism and regulatory cross-interactions between the flux of aromatic amino acid biosynthesis from chorismate and their further metabolism into various secondary metabolites. Perturbation of isoprenoid pathway in potato tubers via expression of a bacterial gene encoding 1-deoxy-D-xylulose5 5-phosphate synthase (DXS) has been reported to cause two-fold increase the level of carotenoids, six to seven fold increase in phytoene and increase in level of trans-zeatin (Morris et al., 2006). A bacterial gene, enoyl-CoA hydratase, converting 4-coumaroyl-CoA, caffeoyl-CoA and feruloyl-CoA to their respective hydroxybenzaldehydes, when expressed in plants to perturb phenylpropanoid pathway has been shown to cause phenotypic abnormalities, depletion of phenolics and over-accumulation of several novel metabolites including the glucosides and glucose esters of 4-hydroxybenzoic acid and vanillic acid and the glucosides of 4-hydroxybenzyl alcohol and vanillyl alcohol (Mayer et al., 2001). Since, gain-of-function approach using bacterial gene to perturb metabolic network keeps enzyme complexes intact while acting on its specific substrate, bacterial genes are used to study metabolic network properties in a targeted manner.

One of the most efficient tools to explore functionality of genes is to control gene expression during different developmental stages of plant (Masclaux et al., 2004). Therefore, inducible system is considered to be the best method for the functional study of gene as it allows spatial as well temporal control of transgene expression (Geng and Mackey, 2011). Also, inducible system causes local changes in level of gene expression thereby causing no change in development of the entire plant

(Borghi, 2010). While constitutive promoters have been routinely used for the expression of transgene, it proves unsuitable for those genes whose suppression or over-expression pose a detrimental effect or lethality to the host plant (Zuo et al., 2006) and when the effect of expression of the target gene is to be seen at a particular developmental stage (Roslan et al., 2001), inducible system is preferred over it. There are several inducible systems used including ethanol-inducible expression with the AlcR/AlcA system, dexamethasone-inducible expression with GR fusions, GVG/UAS or pOp/LhGR systems, β-estradiol-inducible expression with the XVE/OlexA system, heat shock-inducible system and Tet (tetracycline)-regulated expression systems (Borghi, 2010).

Since the introduction of inducible system as a tool for functional analysis of gene, it has been associated with several discoveries. The estradiol-inducible XVE/ pER8 expression vector has also been used for the creation of COS library which provided controlled expression of cDNA so that their overexpression may not prove lethal. This COS system was screened multiple times aiming against a particular stress response. Several genes were identified which when overexpressed could confer the property like enhancement in the germination rate or increased survival rate of seedlings in saline environment by screening the seedlings for salt tolerance (Papdi et al., 2008). Glucocorticoid receptor based inducible system has been developed in rice (Ouwerkerk et al., 2001). GVG inducible system was used to study the function of two putative transcription factors namely LjNDX and LjCPP1 in *L. japonicus*. In this study, it was shown that growth disturbances were seen due to GVG system. The activation of chimeric GVG transcription factor affected shoot internode elongation and root pericycle cell division (Andersen et al., 2003). Glycolate oxidase (GLO)

antisense gene driven by an estradiol-inducible promoter has been used to study the functionality of glycolate oxidase by controlled suppression of GLO. The study suggested GLO to be a typical photorespiratory enzyme. It was also shown that possibly through feedback inhibition on rubisco activase exerts a strong regulation over photosynthesis and when GLO is supressed in rice, glyoxylate cycle may be partially activated in order to compensate for the photorespiratory glyoxylate (Xu et al., 2009). Functional estradiol inducible system along with microarray has been used to identify *AGL-24* induced genes, including Suppressor of Overexpression of Co1 (SOC1), a floral pathway integrator. The study suggested a positive feedback loop during floral transition which integrates flowering signals (Liu et al., 2008). Transgenic rice plants have been created in which cDNA of *DREB1A* and *DREB1B* were introduced under the control of the stress-inducible *rd29* promoter to overexpress the DREB protein. It was shown that the transgenic plants showed increased tolerance to drought and salinity because of overexpression of DREB protein regulated by inducible *rd29* promoter (Datta et al., 2012).

2.5 Regulation of metabolic network

The ability to regulate metabolism is a fundamental process in living systems. Understanding of metabolic network and its regulation enables a better understanding of the biochemical pathways that control development and is the key to the success of metabolic engineering. All life forms regulate their metabolism in order to adapt to the environment (Gutteridge et al., 2007). Biological system recruits a number of strategies for the regulation of metabolic networks in response to perturbations such as genetic/transcriptional regulation (Yeang et al., 2006), enzyme inhibition/activation by small molecules (Reichard, 2002), mRNA attenuation, ribo-switches (Winkler,

2005) and cellular compartmentalization (Baillie et al., 2005). Since regulation of metabolic pathways occurs in coordinated manner with all cellular components participating, integrative omics approach provides holistic view of regulatory process.

Integrative 'omics' approach was used to model lysine biosynthesis pathway in yeast by integrating data from gene expression, growth phenotypes, and proteomics data (Gat-Viks et al., 2004). The model enabled to identify novel regulatory mechanisms in the pathway which were then validated experimentally. Several regulatory processes occur in synchronized way to orchestrate homeostasis in biological system (Venkatesh et al., 2004). *Escherichia coli trp* regulon demonstrates control by genetic regulation, mRNA attenuation and enzyme inhibition, where tryptophan inhibits the enzymes as well as production of those enzymes that are required for its synthesis (Simão et al., 2005). Using systems biology approach on the metabolic network from genome-scale metabolic reconstruction, it was possible to identify common transcriptional response in the metabolic network pattern and hence successfully identify reporter metabolites (metabolites around which the most significant transcriptional changes occur) and a set of connected genes with significant and co-ordinated response to genetic or environmental perturbations. It was observed that the cells respond to perturbation by changing the expression pattern of genes involved in that perturbed metabolic pathway, which results in perturbation of other pathways due to high connectivity of metabolic network (Patil et al., 2005). Several instances has been reported where the enzyme gene expression was controlled through feedback mechanism by small metabolites (Yeang et al., 2006).

Regulation of metabolites and metabolic pathways in plants are particularly interesting, as the plant's sessile nature prevents its escape from unfavourable environments and hence they have evolved a highly dynamic and flexible metabolism to cope with changing conditions. Several factors play important role in determining metabolic state of plants. Untargeted metabolic profiling of *Arabidopsis* at different developmental stages, different tissues and different temperature conditions (cold or heat) showed metabolic profile specific to each condition (Giakountis and Coupland, 2008; Espinoza et al., 2010). Since each developmental stage or tissue or temperature state brings specific requirements to be met by plants to survive and grow, each was represented by a specific metabolic profile. Synthesis of several primary metabolites in plants depends upon light; this brings specificity in regulation of metabolic pathways in day and night period. Starch, the major storage molecule in the plant kingdom, is linearly degraded during the night for the continuation of growth in dark conditions. Each day at dawn, 95% of the total starch reserve in form of starch granules is broken down to support metabolism and normal growth (Smith and Stitt, 2007). Interestingly, a sudden, significant decrease in the length of the light period is followed by immediate modifications in the rate of starch breakdown such that there is a constant supply of starch during the unexpected, prolonged night (Graf et al., 2010), suggesting that plants hold a timing mechanism, anticipated on the previous dawn, to optimize plant growth and productivity (Graf and Smith, 2011). This suggests two separate regulatory systems, one for synthesis and storage of starch and the second for starch degradation, which acts opposite to each other and activation of each depends on diurnal cycles. The transformation of a dormant seed into a photosynthetic active seedling requires conversion of storage compounds into essential building blocks and energy carriers. Imbibition leads to increased hexose

sugars level, followed by a large change in gene transcription (Fait et al., 2006; Howell et al., 2009). Most primary metabolites acquire a stable metabolic state which is more or less enhanced for 24 h after exposure to light and inductive temperature indicating synthesis of primary metabolites to increase or maintain high levels of storage molecules (Allen et al., 2010). Thereafter, a major metabolic switch takes place in which many metabolites are consumed to form building blocks and energy carriers. For instance, sucrose was found to decrease throughout the developmental period, and many transcripts were highly correlated with sucrose levels, suggesting metabolic regulation of transcript abundance. The causality in these cases is difficult to prove, and it is likely that there is a very complex feedback regulation between metabolic status, gene transcription, and the environment to maximize germination vigour (Kooke and Keurentjes., 2011). Distinct phases in a germinating seed can be characterized by transient metabolic profiles (Fait et al., 2006) which indicate highly regulated system that switches from one type of regulation to other depending upon the requirement of that state.

2.6 Application of metabolic network

Understanding of metabolic network has several applications. It allows us to analyse problems ranging from single pathway to multi-species studies. Some of the applications of metabolic network are explained below.

2.6.1 Knowledge- gap filling-discovery of missing pathways

The generation of metabolic networks has been greatly facilitated by the development of high throughput ’omics’ approaches and data integration from different platforms. Studies on metabolic flux analysis using steady-state metabolic flux analysis (MFA)

approach have revealed the dynamic behavior of metabolic network (Gomes et al., 2012). Despite this, systematic characterization of gene functions at a genomic scale remains challenging. For instance, up to 13% of gene functions in the intensively studied bacterium *E.coli* remain currently unknown (Peregrín-Alvarez et al., 2009; Vieira et al., 2011). Many of these genes have possible roles in the metabolic network. Improvement of metabolic network has an important role in identifying the missing pathways and enzymes in any system. Subgraph extraction technique, which can be applied to gene expression, protein levels, operons as well as phylogenetic profiles, has been developed to predict pathways from biological networks. In the evaluation process, the method was examined for 7 sub-network extraction approaches on 71 known metabolic pathways from *Saccharomyces cerevisiae* (Faust et al., 2010). The metabolic network was obtained from MetaCyc (Caspi et al., 2012). The best results were found when random walk-based reduction of the graph was combined with a shortest paths based algorithm. It was shown to have an accuracy of ~77%. Information from aberrant pathways and networks are now being successfully applied in the identification of groups of biological entities (e.g. gene, protein) termed 'functional modules'. These perform biological tasks (e.g. biological processes) which the constituent parts could not perform if they were dissociated (Wang et al. 2012).

2.6.2 Contextualization of high throughput data in metabolic networks

With the omics approach, which generates huge amount of data, challenges lie in how to organize, sort, interrelate, and contextualize these datasets. It motivated the development of top-down approach wherein statistical analysis of high throughput data inferred some biochemical network structure and function. But, it has a drawback as it provides computational and technological hurdle (Stark et al., 2003). Metabolic

network combines all the reactions, which relate the cellular compounds to each other. The reconstruction of metabolic network is now considered as powerful technique for contextualization of high throughput data. Metabolic reconstruction is highly useful in context of gene expression data. The mapping of high throughput data on the metabolic data helps to gain a deeper insight of the known metabolic pathways (Oberhardt et al., 2009). Metabolic network reconstructions enable complex relationships amongst molecular components to be represented formally in a biologically relevant manner while respecting physical constraints. Recently, in a comprehensive study, a rhamnosyltransferase for flavonols was described in *Arabidopsis* and its gene placed in a pathway with other glycosyltransferases using metabolomics and transcriptomics approach (Yonekura-Sakakibara et al., 2007). In-silico models derived from such reconstructions can then be queried or interrogated through mathematical simulations (Bordbar et al., 2011).

2.6.3 Study of diseases through metabolic systems biology approach

Previously, it was thought that the relationship between diseases and changes in a single or a few genes is the sole reason for diseases. However, currently there is a lot of evidence which infers that complex diseases results from multiple genetic aberrations in biological pathways and networks (Meyerson et al., 2010). While most of these studies have been carried out in mammalian models and humans, plants suffers from a number of complex diseases, whose study can benefit from biomedical work. Biological pathways represent series of actions among molecules that lead to a certain product or a specific change in a cell. A major challenge in the diagnosis and

treatment of such diseases is to identify characteristic aberrations in the biological pathways and molecular network activities and elucidate their relationship to the disease (Wang et al., 2011). Therefore, the study of metabolic network and integrating the gene expression with networks of protein–protein interactions and transcription factor binding revealed critical insights into cellular behavior and hence will make the study of diseases easier (Moxley et al., 2009). Any aberration in metabolic flux or topology of metabolic network can thus be measured and simulated to give the idea about the gene involved in pathology and effective drugs then can be made via the retrieved information. Systems biology approaches have been used in *Arabidopsis* to identify two Myb transcriptional factors associated with regulation of glucosinolate biosynthesis in response to sulphur and nitrogen deficiency (Hirai et al., 2007) and in barley to identify metabolites conferring resistance to Fusarium head blight (Bollina et al., 2010).

2.6.4 Metabolic engineering and synthetic biology

Existing metabolic networks can be altered to optimize specific cellular functions which might be either pre-existing or newly designed. This type of 'engineering' is made at the molecular level by the introduction of mutations, deletions, insertions and up or down-regulations of enzymes of interest. Because efficient designs are often non-intuitive, predicting the right adjustments of metabolic fluxes necessary to reach the defined objective necessitates computational strategies based on constraint-methods.

Metabolic engineering refers to the use of molecular biological techniques to alter the metabolic pathway to improve cellular properties by designing and implementing

rational genetic modifications. Metabolic engineering generally focuses on the production of high levels of metabolites of medical or industrial interest. It is accomplished by introduction of mutations, deletions, insertions and up or down regulations of enzyme of interest. Because efficient designs are often non-intuitive, predicting the right adjustments of metabolic fluxes necessary to reach the defined objective necessitates computational strategies based on constraint-methods. A lot of studies have shown that changing expression of a single gene may not be sufficient to bring the desired changes, since regulation of metabolic network occurs at specific junctions (Hoefnage et al., 2002). Thus, knowledge of metabolic network behaviour and regulatory system becomes essential to enable us to achieve desired changes in the biological system. With the help of the knowledge obtained from metabolic pathway, the flux is directed towards those compounds without causing any damage to the cellular environment. They are so far mostly performed in bacteria or yeast. With the use of metabolic engineering, flower colour can be specifically modified without changing the other characteristics. Novel flower colour varieties such as brick red petunias and violet carnations have been successfully made by expression of heterologous flavonoid gene (Tanaka et al., 1998). Human steroid hormones such as hydroxycortisone can also be synthesized in yeast. A future field of application will likely be the area of synthetic biology, which has gained increased interest with the recent completion and incorporation of the first synthetic bacterial genome. However, the complexity of metabolic networks, compounded by multiple layers of transcriptional, protein, and substrate-level regulation of metabolic enzymes, renders predictable metabolic engineering extremely difficult, and often causes unwanted consequences or sub-optimal outcomes when local network maps or intuitive knowledge are the basis of engineering decisions (Kim et al., 2008; Nevoigt, 2008).

Synthetic biology focuses to assemble well-known nucleic acids and proteins into biological 'devices' to design a novel cellular function whose behaviour can be predicted (Haynes et al., 2009). The first synthetic gene circuit were engineered in bacteria, as the components of bacterial cell are well-defined and thus offers an attractive system to operate these gene circuits. But, more recently, synthetic gene circuits are also being established in mammalian, yeast and plant cells (Drubin et al., 2007). Synthetic biology has been used to produce cost effective anti-malarial precursor in yeast and bacteria (Ro et al., 2006; Anthony et al., 2009). Plants are source of numerous bioactive compounds known as natural products (NPs) that have widespread use in pharmaceutical and biotechnological industry. In the last 25 years, these NPs constitute 75% of antibiotic and anticancer drugs approved by FDA (Dudareva et al., 2012). But, the extraction of these NPs from plants is not favoured because of low yield as well as complicated down-stream process. Chemical processes also suffer with low-yield along with use of toxic catalyst and extreme reaction conditions (Koffas et al., 2008). Thus synthetic biology approach involving microbial metabolic engineering is emerging as a promising approach to synthesize these products. *E.coli* and *S.cervisiae* have been engineered with plant secondary metabolic network and has been reported to synthesize diverse range of natural products including fatty acids, terpenoids, flavonoids, polyketides and alkaloids (Xu et al., 2012). Complete understanding of plant's metabolic network at systems level will further help us to employ synthetic biology approach and engineer the microorganisms in order to produce several beneficial compounds thus, overshadowing the limitations imposed by their extraction from plants and chemical process.

2.7 Phenylpropanoid metabolic network

Starting from the amino acids phenylalanine and tyrosine, the phenylpropanoid pathway produces the majority of phenolic compounds found in nature. It has been estimated that approximately 20% of the carbon fixed by photosynthesis is directed into the phenylpropanoid pathway (Vogt, 2010). More than 100,000 phenylpropanoid compounds have been identified from various species, making them one of the largest natural product families in the world. Phenylpropanoid pathway is one of the important secondary metabolic pathways and includes important class of compounds like anthocyanins, flavonoids, proanthocyanidins, catechins, sinapoyl esters, hydroxycinnamates, phenolamides and lignins (Fig 2.1). The magnificent diversity of phenylpropanoids, in principle, is the result of efficient modification and amplification of a very limited set of core structures, derived from the shikimate pathway (Herrmann, 1995). Many of these compounds such as quercetin, kaempferol, cyanidin and pelargonidin are essential for normal plant growth and development. Phenylpropanoid-based polymers, like lignin, suberin, or condensed tannins, contribute to the stability and robustness towards mechanical or environmental damage, like drought or wounding. They also play important roles in defense against pathogens and plant–microbe/animal interactions. Phenylpropanoid pathway networks can be seen as complex biological regulatory systems that have evolved in vascular plants during their successful transition to land, and which are ultimately essential for their growth, development and survival**.** Furthermore, when considered from both evolutionary and functional perspectives, such networks have afforded vascular plants with competitive advantages for successful land plant adaptation (Lewis and Davin,

1994, 1999; Lewis et al., 1999), which include mechanisms for obtaining and transporting water and nutrients (Lewis and Yamamoto, 1990); maintenance of a high water potential which facilitates active metabolism in desiccating environments (Bernards et al., 1995; Bernards and Lewis, 1998); minimizing the effects of temperature, humidity and (UV) light variations; withstanding and modulating forces of compression which act upon plant structures during growth and development (Lewis et al., 1999), and formation of specialized structures which permit maintenance of pollen grain viability (Wiermann and Gubatz, 1992).

Phenylpropanoid pathways contribute to all aspects of plant responses to biotic and abiotic stimuli, which suggest the possibility of its participation in a larger network with other secondary metabolic pathways . A plethora of mutants for genes from phenylpropanoid metabolic pathways have facilitated the identification and functional characterization of genes and enzymes (Alonso et al., 2003; Ossowski et al., 2008). Several flavonoid metabolic enzymes were shown to form complexes in order to regulate the partitioning of intermediates among competing pathways and determine the intracellular deposition of end products (Burbulis et al., 1999). Chalcone synthase (CHS), chalcone isomerase (CHI) and dihydroflavonol 4-reductase (DFR) from the flavonoid biosynthesis pathway were shown to form a complex in an orientation dependent manner. In the same study, interactions were shown between CHS, CHI, and flavonol-3-hydroxylase in lysates from *Arabidopsis* seedlings.

Transparent testa (tt) mutant lines are mutations in genes encoding phenylpropanoid pathway enzyme (i.e *F3H*, Flavanone-3-hydroxylase) or

regulators (i.e. *TTG1*), which are not able to produce anthocyanin and thus their seeds are yellowish in colour (and hence called transparent testa (Table 2.1). Several studies have been done to characterize the effect of mutation in phenylpropanoid pathways in *tt* mutants (Peer et al., 2001). But these studies were directed towards observing changes at flavonoid level changes and their localization or regulation of transcriptional factors. There are no studies done till now upon interaction of phenylpropanoid with other secondary metabolic network. As phenylpropanoid plays an important role in several plant processes, it is highly likely that these pathways link with other metabolic networks to achieve multiple effects. Given their central role and diversity of metabolites produced from its several branches, phenylpropanoid pathways are ideal for deciphering the nature and properties of their metabolic network. Applications of systems biology approach will enable us to improve the understanding of phenylpropanoid metabolic network.

Table 2.1 List of transparent testa mutant lines used in this study

Lines	Enzyme lesion	Locus	References
tt1	**WIP1 domain transcriptional factor (Zinc finger protein)**	**AT1G34790**	**Sagasser et al., 2002**
tt2	**MYB domain protein 123**	**AT5G35550**	**Nesi et al., 2001**
tt3	**Dihydroflavonol 4-reductase**	**AT5G42800**	**Shirley et al., 1995**
tt4	**Chalcone synthase**	**AT5G13930**	**Koornneef, 1990**
tt5	**Chalcone isomerase**	**AT3G55120**	**Shirley et al., 1995**
tt6	**Flavanone 3-hydroxylase**	**AT3G51240**	**Pelletier and Shirley, 1996**
tt7	**Flavonoid 3' monooxygenase**	**AT5G07990**	**Schoenbohm et al., 2000**
tt8	**bHlH domain protein**	**AT4G09820**	**Nesi et al., 2000**
tt9	**Unknown**	**not assigned**	
tt10	**Laccase-15**	**AT5G48100**	**Pourcel et al., 2005**
ttg1	**WD40 repeat protein**	**AT5G24520**	**Walker et al., 1999**

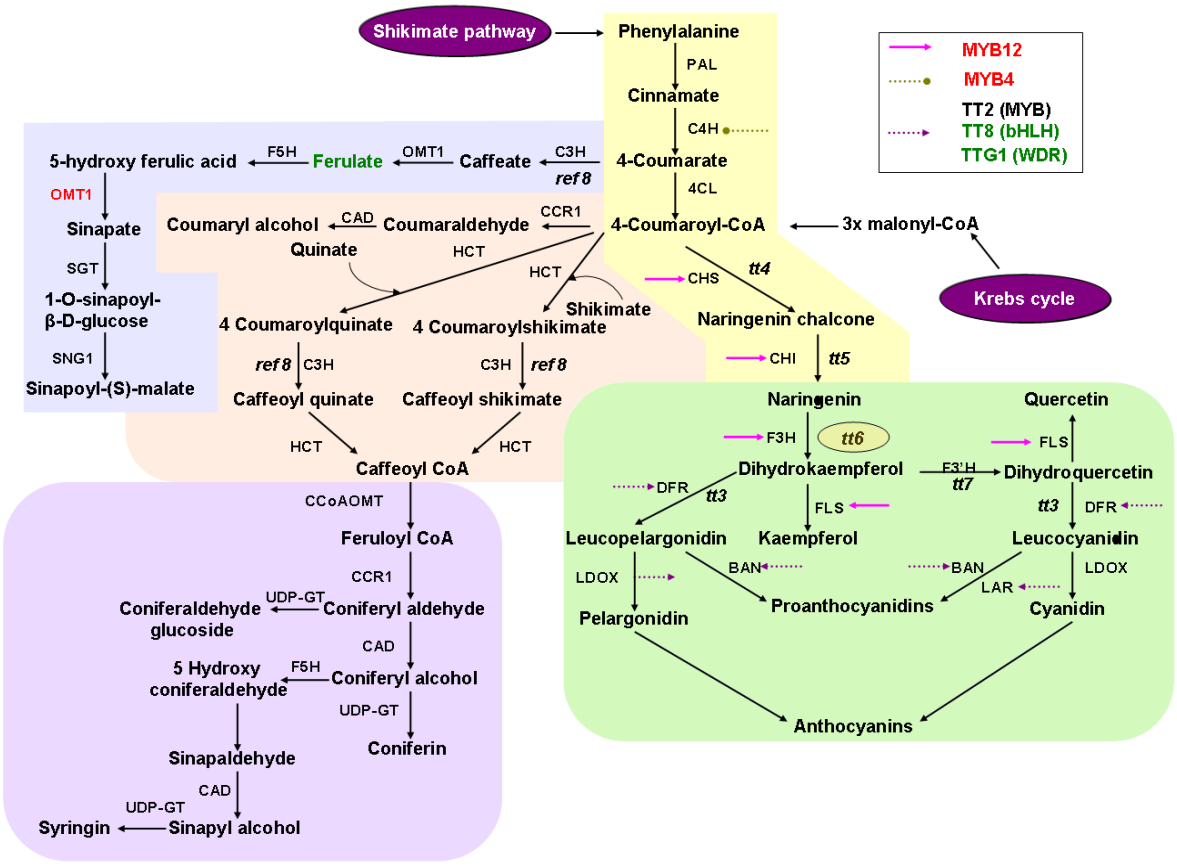


Fig 2.1 Phenylpropanoid metabolic network map in *Arabidopsis* (Sheela et al., 2012) *PAL*: Phenylalanine ammonia lyase, *C4H*: Cinnamate 4-hydroxylase, *4CL*: 4-coumaroyl-coenzyme A ligase, *CHS*: Chalcone synthase, *CHI*: Chalcone-flavanone isomerase; *F3H*: Flavanone 3-hydroxylase, *F3'H*: Flavonoid 3'monooxygenase *DFR*: Dihydroflavonol 4-reductase, *FLS*: Flavonol synthase, *LAR*: Leucoanthocyanidin reductase, *LDOX*: Leucoanthocyanidin dioxygenase, *BAN*: Banyuls, *ANR*: Anthocyanidin reductase, *ANS*: Anthocyanidin synthase, *FGT*: Flavonoid glycosyl transferases, *C3H*: Coumarate-3-hydroxylase, *F5H*: Ferulate-5-hydroxylase, CCR1: Cinnamoyl-CoA reductase 1, *OMT1*: Flavonol 3'-o-methyltransferase, *SGT*: Sinapate-1-glucosyltransferase, *SNG1*: Sinapoylglucose:malate sinapoyltransferase, *HCT*: Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase, *CAD*: Cinnamyl alcohol dehydrogenase, *CCoAOMT*: Caffeoyl CoA o-methyltransferase, *UDP-GT*: UDP glucose transferase, *SHT*: Spermidine hydroxycinnamoyl transferase, bHLH: basic Helix-Loop-Helix, WDR: WD-repeat protein with WRKY domains, TT: transparent testa, *TTG1*: Transparent testa glabra 1.

CHAPTER 3

MATERIALS AND METHODS

3.1 Plant materials and growth conditions

Arabidopsis seeds were obtained from the *Arabidopsis* stock centre (*Arabidopsis* Biological Resource Centre, USA). Seeds were stratified for 4 days at 4^0C and then were sown to pots containing pre-mixed potting mixture. These pots were grown under standard room condition (22°C, long days: 16 h light/8 h dark cycle) with relative humidity of 60-70 and illumination 70wm^{-2}.

3.2 Growing seedlings on MS media

Arabidopsis seeds were surface sterilized using 30% Clorox with 7 min incubation. This was followed by 6 times washing with autoclaved water. Seeds were then placed on full strength MS medium with 5% sucrose and 5% Phytoagar® for 4 days before placing them in growth chamber at 22°C for 16 h light/8 h dark cycle. Seedlings were harvested 6 days after transferring the plates in light.

3.3 Chemicals and reagents

HPLC grade acetonitrile and methanol (J. T. Baker, NJ, USA) were used for all the experiments. Analytical reagent (AR) grade hydrochloric acid; glacial acetic acid and sodium hydroxide were from Amersham Pharmacia Biotech (Uppsala, Sweden). Quercetin, spermidine, kaempferol and naringenin were obtained from Sigma Chemical Co. (St. Louis, MO, USA). Quercetin-3-glucoside, quercetin-3,4-diglucoside and kaempferol-7-glucoside were obtained from Extrasynthese™ (United

States). Flavonoid stock solutions were prepared in 80% methanol (Pillai and Swarup, 2002).

3.4 Plasmid construction

To construct p2x35s:tt8-GR and p2x35s:ttg1-GR, full length cDNA for tt8 and ttg1 were amplified with introducing cla1 and xho1 restriction sites at 5' and 3' ends respectively. GR vector from Prof Yu Hao was used to create this construct. Amplified DNA was digested by Cla1 and Xho1, and cloned into GR vector.

To create pER:ami_TT8 and pER:ami_TTG1, web microRNA designer was used to create artificial microRNA specific to its target *TT8* and *TTG1*. Designed primers were used to amplify genes using artificial miRNA vectors pRS300, and were digested using Xho1 and Spe1 from 5' to 3' end respectively. Digested PCR product was cloned into pER8 vector.

3.5 Anthocyanin extraction

For anthocyanin extraction, six-day-old seedlings were used because it is the developmental stage when anthocyanin accumulation has been shown to be high (Pourcel et al., 2010). 0.1g seedlings ground in 0.2ml acidified methanol (1% v/v concentrated HCl) and stored at 4ºC overnight to facilitate the total extraction of anthocyanins. Three biological replicates each consisting of 40-50 seedlings (0.1g) were analyzed. The suspension was centrifuged at 14,000 rpm for 20 min. Absorbance was read at 530nm using Beckman Spectrophotometer DU 640 B (Beckman Coulter, Inc. CA, USA). Relative anthocyanin units were calculated as $A=A_{530}-1/4xA_{650}$ per gram fresh weight (Pourcel et al., 2010).

3.6 Stem stiffness test

Five-week-old *Arabidopsis* plants were used to test tensile strength. Inflorescence stems were measured for its cross sectional area, length and thickness. Stems were then placed from both ends to two fix ends of tensile strength tester (5900 series mechanical testing systems from Instron). Displacement speed for the two ends were fixed to 0.5mm/min. Maximum load and extension at breaking point was measured and modulus calculated as stress over strain in MPa using instruments and software provided by the manufacturer.

3.7 Lignin staining

Five-week-old stem tissues (near to the junction between stem and root) of WT (Col and Ler), pp10 and *tt6 Arabidopsis* lines were hand-sectioned and immediately immersed for 2 min in 10% phloroglucinol (Sigma-Aldrich Chemical Co., USA) dissolved in 100% ethanol. Following this, the sections were incubated in concentrated HCl for one minute. Sections were rinsed and images were taken at 80x magnification under a dissection microscope with bright field illumination.

3.8 Plant transformation

3.8.1 Preparation of *Agrobacterium tumefaciens* GV3101 competent cells

Agrobacterium tumefaciens GV3101 cells were inoculated in 100ml LB media with gentamycine, riphampicin and tetracycline antibiotics, incubated at 28°C with shaking till it grows to OD600~0.6. The culture was placed on ice and then centrifuged at 3500 rpm for 15 min at 4°C. Thepellet obtained was washed 5 times with ice-cold double autoclaved H_2O by resuspension and centrifugation. A final 5ml cold double

autoclaved H_2O was added to suspend all cell pellet and aliquots of 200µl/tube were flash frozen in liquid for future use.

3.8.2 Electroporation to competent cell

Prepared competent cells tubes were added with 200 ng plasmid and incubated on ice for 1 hour. Cells were transferred to 1mm Gene pulser cuvettes (Bio Rad, USA), and electroporated using MicroPulser Electropotor (Bio Rad, USA). *Agrobacterium* cells were recovered in 1ml of LB broth for 2 hours at 28°C with shaking. Cells were collected and plated on LB agar plate supplemented with 10µg/ml tetracycline, 50µg/ml rifampicin, 10µg/ml gentamycin and kanamycin 100µg/ml for GR constructs or spectinomycin for pER constructs. The plates were incubated at 28°C for 2 days. Selected colonies were tested for insert and PCR confirmed colonies were used for floral dip.

3.8.3 Method of transformation of *Arabidopsis thaliana*

Plasmids were extracted from *Agrobacterium* colonies obtained from selection platesand were sequenced to confirm the presence of construct. Positive *Agrobacterium* colonies were inoculated into liquid LB medium supplemented with antibiotics and incubated at 28 °C with shaking until OD600 reached 0.6. The *Agrobacterium* cells pelleted by centrifugation at 8000rpm for 5 minuwere resuspended in the same volume of medium containing 5% sucrose and 0.01% Silvet L-77® (LEHLE Seeds, USA), incubated at room temperature for 1 h. Flower buds of *Arabidopsis* wild type Col-0 to be transformed were dipped into the *Agrobacterium* culture for around 2 min, after which plants were covered and placed in dark for 16 h

to improve transformation efficiency. Plants were then placed back in growth chambers for normal growth and seeds were harvested.

3.8.4 Screening of transgenic plants

Eight-day-old seedlings were screened by spraying 0.2% Basta® solution twice at one week interval with first spray on transgenic plants. Surviving T1 plants were genotyped to confirm the successful transformation.

3.9 Genotyping

One rosette leaf of ten-day-old putative transgenic plants was cut and ground in 200µl extraction buffer (0.2M tris-HCl, 0.4M LiCl, 25mM EDTA and 1% SDS) in Eppendrof tubes until the solution turned light green. The homogenized sample was centrifuged at 14000g in a mini-centrifuge for 5 min at room temperature. 150-µl of supernatant was mixed with 150µl isopropanol in a freshtube by inverting several times. After centrifuging at 14000g for 8 min, the pellet was washed with 70% ethanol and dried in a vacuum. Pellets were dissolved in 30µl nuclease-free water and used for genotyping by PCR.

3.10 Gene expression analysis

Total RNA was extracted from 0.1 g plants using Omega® plant RNA extraction kit (Omega BIO-TEK®, USA) using the manufacturer's protocol. 1µg total RNA was used to convert into cDNA using RevertAid® first strand cDNA synthesis kit (Thermo Scientific, USA) using manufacturers' instructions. Single stranded cDNA were diluted to a concentration of 50ng/µl and were used as template for quantitative real time PCR analysis. Quantitative real-time PCR was performed in triplicates using

an ABI 7500 real time PCR system (Applied Biosystems, USA) using Maxima SYBR Green qPCR mix (Thermo Scientific, USA). Delta delta Ct method was used to calculate relative expression level using TUB2 as endogenous control.

3.11 Primer and Oligonucleotides

Primers were designed using the NCBI primer designing tool and synthesized by Sigma Aldrich (Singapore). Lyophilized primers were dissolved in the appropriate volume of TE buffer to reach a final concentration of 100µM. For working solutions, primers were further diluted 1:10 with ddH2O. All primers used for this work are listed in Table 3.1.

Table 3.1 List of all primers used in this study

Primer Name	Squence (5'-3')	Annotation
ttg1_(i)miR-s	GACTGATACGCCTTTGTTAAGACTCTCTCTTTTGTATTCC	amiRNA primers for ttg1 silencing
ttg1_(ii)miR-a	GAGTCTTAACAAAGGCGTATCAGTCAAAGAGAATCAATGA	
ttg1_(iii)miR-s*	GAGTATTAACAAAGGGGTATCAGTCACAGGTCGTGATATG	
ttg1_(iv)miR-a*	GACTGATACCCCTTTGTTAATACTCTACATATATATTCCT	
tt8_(i)miR-s	GATTAACGTATGGGTTGACTCAATCTCTCTTTTGTATTCC	amiRNA primers for tt8 silencing
tt8_(ii)miR-a	GATTGAGTCAACCCATACGTTAATCAAAGAGAATCAATGA	
tt8_(iii)miR-s*	GATTAAGTCAACCCAAACGTTATTCACAGGTCGTGATATG	
tt8_(iv)miR-a*	GAATAACGTTTGGGTTGACTTAATCTACATATATATTCCT	
tt6_(i)miR-s	GATAATACGTTACAATCTCTCTGTCTCTCTTTTGTATTCC	amiRNA primers for tt6 silencing
tt6_(ii)miR-a	GACAGAGAGATTGTAACGTATTATCAAAGAGAATCAATGA	
tt6_(iii)miR-s*	GACAAAGAGATTGTATCGTATTTTCACAGGTCGTGATATG	
tt6_(iv)miR-a*	GAAAATACGATACAATCTCTTTGTCTACATATATATTCCT	
primer a	CTGCAAGGCGATTAAGTTGGGTAAC	primers for Prs300 (miR319A Pbsk
primer b	GCGGATAACAATTTCACACAGGAAACAG	
tt2RT-PCRf	TTGATGGTTTGGACTGTG	Quantitative RT PCR primers
tt2RT-PCRr	GCAATATAGGCTCAACAAGT	Quantitative RT PCR primers
ttg2RT-PCRf	CCCCACAACTTTCTAAGCAAACA	Quantitative RT PCR primers
ttg2RT-PCRr	TGCTTAGGAAGTTGTGAGTGAAG	Quantitative RT PCR primers
myb118RT-PCRf	ATGTGCATCACAAACCATCG	Quantitative RT PCR primers
myb118RT-PCRr	AGTTGGTGGGGTGGTATAGG	Quantitative RT PCR primers
myb5RT-PCRf	TCTCATCCTCCGTCTTCACC	Quantitative RT PCR primers

myb5RT-PCRr	TTGCATCAAGAGGCTTGTGG	Quantitative RT PCR primers
DFR-RT-PCRf	TGGTGGTCGGTTCCATTCAT	Quantitative RT PCR primers
DFR-RT-PCRr	GAGAGAGCGCGGTGATAAGG	Quantitative RT PCR primers
BAN-RT-PCRf	ACATTTGCTGTCTTACAACACAAGT	Quantitative RT PCR primers
BAN-RT-PCRr	CGAAAGCCTTCATTGATAAGTTTTTGCG	Quantitative RT PCR primers
p ER-f	ACCCTTCCTCTATATAAGGAAGTTC	
p ER-r	TCGGTTTCGACAACGTTCGTCAA	Vector pER8 primers
35Spromoter f	GACCCTTCCTCTATATAAGGAAGTTC	
PGP2r	CGACGGCCAGTGAATTGTAATACG	Vector HY109 primers
tt8:gr-f	AACTCGAGATGGATGAATCAAGTATTATTC	
tt8:gr-r	ATTATCGATTAGATTAGTATCATGTATTATGAC	primers for tt8:gr construct
ttg1:gr-f	AAACTCGAGATGGATAATTCAGCTCC	
ttg1:gr-r	ATATCGATAACTCTAAGGAGCTGCATTTTG	primers for ttg1:gr construct
GL2-RT-PCRf	AAGCTCGTCGGCATGAGT	Quantitative RT PCR primers
GL2-RT-PCRr	TTCTCTCGATTTCACTGTCTGG	Quantitative RT PCR primers
TT10-RT-PCRf	CAATGCATTGGCATGGTGTAGAG	Quantitative RT PCR primers
TT10-RT-PCRr	CTCACATCCCTCTTCCACCAC	Quantitative RT PCR primers
HMG2-RT-PCRf	TACTCGGTGTGAAAGGATCAAACA	Quantitative RT PCR primers
HMG2-RT-PCRr	CCGAACCAGCCACTATTCTTG	Quantitative RT PCR primers
ISPD-RT-PCRf	GCGGTTGGAGCAGCTGTACTTG	Quantitative RT PCR primers
ISPD-RT-PCRr	GGTGTCTGCATTTCCCATAGGGTTT	Quantitative RT PCR primers
GPPS-RT-PCRf	ACGTCGCAAGTCTACTGCACGATGA	Quantitative RT PCR primers
GPPS-RT-PCRr	AAGCCCGGGAGAGCAAGAAGTCTC	Quantitative RT PCR primers
BR6OX1-RT-PCRf	TGGCCAATCTTTGGCGAA	Quantitative RT PCR primers
BR6OX1-RT-PCRr	TCCCGTATCGGAGTCTTTGGT	Quantitative RT PCR primers
CBB2-RT-PCRf	CCCTCCGCTACCGCTTAAAAGTC	Quantitative RT PCR primers
CBB2-RT-PCRr	AACGCGAGTGATTCGAGAAGAGAAC	Quantitative RT PCR primers
GGR-RT-PCRf	ACCGCCTGTGCCCTAGAAATGGT	Quantitative RT PCR primers
GGR-RT-PCRr	GCCAGAGCCGTAGACAGTGTGGTTA	Quantitative RT PCR primers
PR5-RT-PCRf	AAT GTC AAG CTG GGG A	Quantitative RT PCR primers
PR5-RT-PCRr	AGG TGC TCG TTT CGT C	Quantitative RT PCR primers
PR1-RT-PCRf	TCC GAT AGC GGC AAA G	Quantitative RT PCR primers
PR1-RT-PCRr	CCC ACT GGC ATA GGA T	Quantitative RT PCR primers
LURP1-RT-PCRf	CAACTTCCATTCCGCAACAATTCG	Quantitative RT PCR primers
LURP1-RT-PCRr	GAAGGGGTCAAAGACATGATTCTCG	Quantitative RT PCR primers
DOGT1-RT-PCRf	CCGGGGTTGAACAGCCTATGAAA	Quantitative RT PCR primers
DOGT1-RT-PCRr	TGGGTTCTGCCAGTTCCATTATGTC	Quantitative RT PCR primers
PCC1-RT-PCRf	GCGTCAACCCCGAAGCCATAAT	Quantitative RT PCR primers
PCC1-RT-PCRr	GGTTTGGGCAACGACTTCTGTCTC	Quantitative RT PCR primers
PDR12-RT-PCRf	GTTTCTTGAGTTTCCAGAGGAGTTTC	Quantitative RT PCR primers
PDR12-RT-PCRr	CCAAGCGAGTCCTAGTATGAGAAGAAAC	Quantitative RT PCR primers
SAG13-RT-PCRf	GAACTCAAGATGGAGTCTTGG	Quantitative RT PCR primers
SAG13-RT-PCRr	ATTGTTGACGAGGATGTTGAG	Quantitative RT PCR primers
CYP76C2-RT-PCRf	CCGTACCCCCACTAACTCGATACG	Quantitative RT PCR primers

CYP76C2-RT-PCRr	GAGTCGCCGACAGTTTTCTCAAAAG	Quantitative RT PCR primers
WAK1-RT-PCRf	GCA CAT GAC TTC TTT TCA CGA	Quantitative RT PCR primers
WAK1-RT-PCRr	GCG GTA ACC AGA TTG ACA CTT	Quantitative RT PCR primers
SUS2-RT-PCRf	TGCCATGAATAATGCCGATTTC	Quantitative RT PCR primers
SUS2-RT-PCRr	TTGCCCAACATTGTTCTTGCTT	Quantitative RT PCR primers
SUS4-RT-PCRf	AAGGAATCGTTCGCAAATGG	Quantitative RT PCR primers
SUS4-RT-PCRr	TTTCAGCGGCAACATCCTC	Quantitative RT PCR primers
ATSUC7-RT-PCRf	TTGGATGGGTCGTGAAGTGTATGGT	Quantitative RT PCR primers
ATSUC7-RT-PCRr	TTGCCAAACACACGGCGAGAATAAT	Quantitative RT PCR primers
UGT85A1-RT-PCRf	ATCGAGAAGGGCTTATGTCCGCTAA	Quantitative RT PCR primers
UGT85A1-RT-PCRr	CGCGGAGGGCGAAACTAATCATAAC	Quantitative RT PCR primers
NIT4-RT-PCRf	CGCCACGCTAGATAAGGCAGAGA	Quantitative RT PCR primers
NIT4-RT-PCRr	TGGCCATTAACGCTAATCGTTCCAC	Quantitative RT PCR primers
WAK3-RT-PCRf	CTGCACGAGGCCCGAATACAAAC	Quantitative RT PCR primers
WAK3-RT-PCRr	TGTCGTCGGAGCTTGGTATACTTCC	Quantitative RT PCR primers
CSLE1-RT-PCRf	CTCTCGCGAGCTGCACTTATGAGG	Quantitative RT PCR primers
CSLE1-RT-PCRr	TTCCGGGTTCAGGTAGGCTGATTTC	Quantitative RT PCR primers
UGT75B1-RT-PCRf	CTTCCACAACTCCATGATCGCAAAC	Quantitative RT PCR primers
UGT75B1-RT-PCRr	TGAGATTCACCGACCTTTTCTGACG	Quantitative RT PCR primers
BAN-P1-f	TCTTAGGTGAAGACAAGTTGGTT	ChIP qRT PCR primers
BAN-P1-r	TGAGTTTATCGTCTTGAGACTTCTA	ChIP qRT PCR primers
BAN-P2-f	TGCAACTCATTTTAAGATCGGCA	ChIP qRT PCR primers
BAN-P2-r	ACCAACTTGTCTTCACCTAAGAA	ChIP qRT PCR primers
BAN-P3-f	GCCTTCTTTTGTTTTCGATTTAGGA	ChIP qRT PCR primers
BAN-P3-r	CGAGTCAATTCAACATAACTTGCT	ChIP qRT PCR primers
BAN-P4-f	GCAATGGAGTTTGGTTTCGGT	ChIP qRT PCR primers
BAN-P4-r	GGCATTGGCAATAGCATTCCG	ChIP qRT PCR primers
BAN-P5-f	GTGTTTGCCGGTGGTTCTTC	ChIP qRT PCR primers
BAN-P5-r	TCCAACTCCTTAGCTTGCAACA	ChIP qRT PCR primers
TT10-P1-f	ACGGACCAAGAAGATGATGCT	ChIP qRT PCR primers
TT10-P1-r	GAAGTTCATCAAAGAATAGAGAGGC	ChIP qRT PCR primers
TT10-P2-f	TGCTTGGCCTAACAAAACTCA	ChIP qRT PCR primers
TT10-P2-r	AGCATCATCTTCTTGGTCCGT	ChIP qRT PCR primers
TT10-P3-f	ACATGTAGCTTTGGCGTGGA	ChIP qRT PCR primers
TT10-P3-r	TGTTAGGCCAAGCATTACCTT	ChIP qRT PCR primers
TT10-P4-f	AGAAATCGACATTTCACCAGTCA	ChIP qRT PCR primers
TT10-P4-r	ACGAAAGAGATGTTTCGCTAAC	ChIP qRT PCR primers
TT10-P5-f	GACATGAAGAGGTTTGACA	ChIP qRT PCR primers
TT10-P5-r	CTGGTGAACTTCAAAGAATC	ChIP qRT PCR primers
GGR-P1-f	TCGATGTTCTTCTTCTGTTCTTGT	ChIP qRT PCR primers
GGR-P1-r	TGGGTTCAGCCAAACTCTCT	ChIP qRT PCR primers
GGR-P2-f	GAAAGAGGGTTGCGAAGGGA	ChIP qRT PCR primers
GGR-P2-r	ACAGCTTTTCATCTGTAAGAGAACC	ChIP qRT PCR primers

GGR-P3-f	AAAGCGGCATCGTGAGGTAA	ChIP qRT PCR primers
GGR-P3-r	TCCCTTCGCAACCCTCTTTC	ChIP qRT PCR primers
GGR-P4-f	CTTGTGGGGCATTGGACAGA	ChIP qRT PCR primers
GGR-P4-r	TTACCTCACGATGCCGCTTT	ChIP qRT PCR primers
GGR-P5-f	CTCTCCAATGCCCGTGTTCT	ChIP qRT PCR primers
GGR-P5-r	AGGACGGCCTCATTTTCCTG	ChIP qRT PCR primers
CBB2-P1-f	ACAGCTGTGACCACTCTTCC	ChIP qRT PCR primers
CBB2-P1-r	TGTGAGAGAGAAAAGTGTGG	ChIP qRT PCR primers
CBB2-P2-f	TGAGATTAAGGATTTTGTGATGCC	ChIP qRT PCR primers
CBB2-P2-r	AAGGGAAGAGTGGTCACAGC	ChIP qRT PCR primers
CBB2-P3-f	AGGTGGTTGGGGGTAAAATGT	ChIP qRT PCR primers
CBB2-P3-r	TGGCATCACAAAATCCTTAATCT	ChIP qRT PCR primers
CBB2-P4-f	TCACCGATCAAGATCCAACCA	ChIP qRT PCR primers
CBB2-P4-r	ACCCCCAACCACCTATCTCT	ChIP qRT PCR primers
CBB2-P5-f	TCACATTGAATGCTCCAAATC	ChIP qRT PCR primers
CBB2-P5-r	TGCACTCACTGTCACCCAAT	ChIP qRT PCR primers
PCC1-P1-f	CGCTAAGGTAGGTCTGTCCG	ChIP qRT PCR primers
PCC1-P1-r	GGAGCTGAGGAGTGAGGATG	ChIP qRT PCR primers
PCC1-P2-f	TCATGTACGTGACGAAAGCCA	ChIP qRT PCR primers
PCC1-P2-r	CGGACAGACCTACCTTAGCG	ChIP qRT PCR primers
PCC1-P3-f	CATGGTTCAAAGGCAGTTGGT	ChIP qRT PCR primers
PCC1-P3-r	CACGTACATGACGTCGTAAAACA	ChIP qRT PCR primers
PCC1-P4-f	CCATGATTCCCATTTGCTTCGT	ChIP qRT PCR primers
PCC1-P4-r	TGGCTTTGTTCCTTATGTTGGTG	ChIP qRT PCR primers
PCC1-P5-f	AGTTGTATCCAGGGTTGCGG	ChIP qRT PCR primers
PCC1-P5-r	GGATCTCTTTCAACCAGTGAAGC	ChIP qRT PCR primers
PDR12-P1-f	GTGGAAGAAAGATTCAGGAAG	ChIP qRT PCR primers
PDR12-P1-r	GATTCCTTTCCTGAGACGATC	ChIP qRT PCR primers
PDR12-P2-f	CAGAGGAGTCACAGCTGACG	ChIP qRT PCR primers
PDR12-P2-r	TTGGGTTTGTGTGCTCCGTA	ChIP qRT PCR primers
PDR12-P3-f	TGGCAAAGACAACAACGA	ChIP qRT PCR primers
PDR12-P3-r	CGTCAGCTGTGACTCCTCTG	ChIP qRT PCR primers
PDR12-P4-f	TCGTTTGTTTTCTTCTCCTCACG	ChIP qRT PCR primers
PDR12-P4-r	TGCAACCACAAAACTTTAATTTG	ChIP qRT PCR primers
PDR12-P5-f	TGAAGATAAAAAGTCCAAGGCTGA	ChIP qRT PCR primers
PDR12-P5-r	ACGTGAGGAGAAGAAAACAAACG	ChIP qRT PCR primers
BR6OX1-P1-f	GTTCCGCTCTCCTTCGATGGA	ChIP qRT PCR primers
BR6OX1-P1-r	TGGTTTCTCATGAAGTTGG	ChIP qRT PCR primers
BR6OX1-P2-f	GCTAGGAGACAACAAAAGTGA	ChIP qRT PCR primers
BR6OX1-P2-r	AGTGCAGCATGCTCAAGAGA	ChIP qRT PCR primers
BR6OX1-P3-f	CGAATCCCGTATATGATCGCGT	ChIP qRT PCR primers
BR6OX1-P3-r	AGCATGATGGTGTTCAAAAGAGT	ChIP qRT PCR primers
BR6OX1-P4-f	TGTTGTAAACGCTTAAAGTGC	ChIP qRT PCR primers
BR6OX1-P4-r	ACGGGATTCGAATTTGGTGT	ChIP qRT PCR primers

BR6OX1-P5-f	GGTTTACGCGGTTGCCTAGG	ChIP qRT PCR primers
BR6OX1-P5-r	CAATATAGGAAACTTAAGC	ChIP qRT PCR primers
GPPS-P1-f	CACTGAACTTGTTGAGTTGG	ChIP qRT PCR primers
GPPS-P1-r	GGCAACAAGCGAAAATGGGT	ChIP qRT PCR primers
GPPS-P2-f	GGGTTACGTTTGCAGAACAACT	ChIP qRT PCR primers
GPPS-P2-r	ACAACTGAAGCAATGTCCAACA	ChIP qRT PCR primers
GPPS-P3-f	TCATGGGTCCTATTTGGATGC	ChIP qRT PCR primers
GPPS-P3-r	TGTTCTGCAAACGTAACCCT	ChIP qRT PCR primers
GPPS-P4-f	TGTGCTAATTATCTCGGTTTTTGT	ChIP qRT PCR primers
GPPS-P4-r	CACTAGTTTTGGCTGATTACTCCA	ChIP qRT PCR primers
GPPS-P5-f	CTCTCGCCTCTCATCGGTTC	ChIP qRT PCR primers
GPPS-P5-r	TCCCCACAACACTCAATTATC	ChIP qRT PCR primers
HMG2-P1-f	ACGATGATGACCCAACAACCT	ChIP qRT PCR primers
HMG2-P1-r	TGGTTATGAGATTATTTTGGGCT	ChIP qRT PCR primers
HMG2-P2-f	TCTTCACTGATCCATGTTGGA	ChIP qRT PCR primers
HMG2-P2-r	TGTCTCCTCCGTCACAATCT	ChIP qRT PCR primers
HMG2-P3-f	TGTGGCCAAATTTACCCAAGA	ChIP qRT PCR primers
HMG2-P3-r	TGAGGCGTTTATTTTGGGTCT	ChIP qRT PCR primers
HMG2-P4-f	GCAAAAGGATCGGGAAAC	ChIP qRT PCR primers
HMG2-P4-r	GCTAATTATTTAGTGATG	ChIP qRT PCR primers
HMG2-P5-f	ACTGCTGAATTCCAAAAGCCA	ChIP qRT PCR primers
HMG2-P5-r	GTTTCCCGATCCTTTTGCGG	ChIP qRT PCR primers
ISPD-P1-f	TCCCAACGGTTTCGCCA	ChIP qRT PCR primers
ISPD-P1-r	AAGAAAGCTCCGGTTCCT	ChIP qRT PCR primers
ISPD-P2-f	AGAAGGCACCAACACCGATT	ChIP qRT PCR primers
ISPD-P2-r	CGAAACCGTTGGGAACAACC	ChIP qRT PCR primers
ISPD-P3-f	CGTGACATTCTAAGAACTTTGC	ChIP qRT PCR primers
ISPD-P3-r	GCATCAGAAACAACCCGCTC	ChIP qRT PCR primers
ISPD-P4-f	TGCAAGGAATGGATTCATCTCAG	ChIP qRT PCR primers
ISPD-P4-r	ATGGTTAGAATCAAGTTTTGGGATT	ChIP qRT PCR primers
ISPD-P5-f	ACCTCAAGGAATGGATTCATCTCAA	ChIP qRT PCR primers
ISPD-P5-r	TTTGAAGTCAAGTTTTGTGGAAGA	ChIP qRT PCR primers

3.12 ChIP Assay

3.12.1 Tissue fixation

Fresh plant materials from two developmental stages, six-day-old seedlings and five to seven-day-old siliques were collected into MC buffer (10mM potassium phosphate, pH 7.0, 50mM NaCl and 0.1 M sucrose) supplemented with 0.01% succinamide glutarate and were infiltrated by applying vacuum for 45 min at 4ºC. Thereafter, MC

buffer was replaced by fresh MC buffer containing 1% formaldehyde and re-infiltrated by applying vacuum for another 45 min at 4ºC. Following this, the reaction was quenched by incubating with 0.15M glycine for 20 min at 4ºC. Fixed tissues were washed by cold MC buffer for 20 min and this process was repeated for three times. Fixed tissues were flash frozen in liquid nitrogen and stored at -80ºC until used for ChIP experiments as described below.

3.12.2 Nuclear protein-DNA extraction

Fixed tissues from the above treatment were ground into powder with mortar and pestle using liquid nitrogen. Freshly prepared chilled M1 buffer (10mM sodium phosphate buffer, pH 7.0, 0.1M NaCl, 10mM beta-mercaptoethanol, 1M hexylene glycol and half tablet of Roche protease inhibitor cocktail (Catalog no 05892791001) was added to the powder resulting in thick slurry. This slurry was transferred to an Eppendorf tube and centrifuged at 14,000rpm for 3 min at 4ºC. Supernatant was discarded and pellet was resuspended to 2ml prechilled M2 buffer (10mM sodium phosphate buffer, 10mM beta-mercaptoethanol, 1M hexylene glycol, 10mM MgCl2,0.5% Triton X-100 and one tablet of Roche protease inhibitor cocktail) followed by centrifugation at 14,000g at 4ºC for 3 min. This process of washing pellet with M2 was repeated for multiple times till pellet turned to light green. Pellet was then resuspended into pre-chilled M3 buffer (10mM sodium phosphate buffer, 10mM beta-mercaptoethanol and half tablet of Roche protease inhibitor cocktail) and centrifuged at 14,000 rpm for 3 mint at 4ºC. The supernatant was discarded, and the pellet was resuspended in sonication buffer.

3.12.3 Sonication of nucleus

Resuspended nuclei pellet in 0.5ml sonication buffer (10mM sodium phosphate buffer, pH 7.0, 0.1M NaCl, 10mM EDTA pH 8.0, 0.5% sarkosyl, one tablet of Roche protease inhibitor cocktail (Catalog no-05892791001)) was placed in 1.5ml Eppendrof tube. The sample tube was kept in an iced water bath and samples were sheared with a probe sonicator for a 15 min continuous pulse at 10% amplitude output (BioRad®, USA). The sonicated sample was centrifuged at maximum speed for 5 min at 4°C. Supernatants were collected in a new tube, and pellet were resuspended in 0.25ml and centrifuged at maximum speed for 5 min at 4°C, supernatant was collected and added to last collected supernatant. 75µl of supernatant was saved in another tube, which was used to test sonication effects, while rest of the samples were subjected to immunoprecipitation.

3.12.4 Chromatin immunoprecipitation (ChIP)

200µl of 25% v/v agarose beads (Sigma Aldrich, USA) were equilibrated with 1ml of IP buffer (50Mm HEPES pH 7.5, 150mM KCl, 5mM $MgCl_2$, 10µM $ZnSO_4$, 1% Triton X-100 and 0.05% SDS), centrifuged at 2500g for one minute. 1ml of supernatant was removed. The final volume was maintained at 200µl. 800µl of solubilised chromatin solution was pre-incubated with 100µl agarose beads and equal volume of IP buffer on an orbital shaker for one hour at 4°C. This was followed by centrifugation at 3000g for 5 min at 4°C and supernatant was collected. Supernatant was incubated with 3µl of anti GR antibody (Milipore®, USA) at 4°C on an orbital shaker for two hours. After adding 100µl of agarose beads, the chromatin-antibody solution was further incubated overnight at 4°C on an orbital shaker. After overnight binding, the solution mix was centrifuged at 2500rpm for 3 min to remove

supernatant, and pellets were washed six times collecting the pellets and re-suspending in IP buffer before re-suspending the beads at 2500rpm for 3 min. To elute the chromatin antibody complex, the beads were incubated with 200µl elution buffer (50mM Tris pH 8.0, 1% SDS and 10mM EDTA) at 65°C for 15 min on a shaker and the suspension was centrifuged at 14000g for 2 min at RT. Supernatant was collected as immunoprecipitated chromatin eluates. Additional elution buffer (100µl each time) was added to repeat the incubation step to improve recovery of chromatin complex. Out of total 500µl combined eluates, 40µl was used for Western blot analysis, while the rest was used for DNA recovery.

3.12.5 Dex, chx and β-estradiol induction of gene expression

Dexamethasone-inducible TTG1 and TT8 over-expression transgenic lines were created in Col-0 background to identify direct and indirect targets of the transcriptional factors tt8 and ttg1. Induction of GR in the transgenic *Arabidopsis* plants was performed according to Morohashi et al. (2007)**,** with numerous modifications. To induce over-expression in the young seedlings, seeds were grown on MS media and 5-day-old seedlings were transferred to MS plate with 50µM dexamethasone prepared in ethanol and incubated for 24 h before RNA extraction. To induce over-expression in siliques, 50µM 0.015% v/v of Silwet L-77® (LEHLE Seeds, USA) was sprayed over 40-day-old transgenic lines till the plants were completely wet. Spray was repeated three times with 8 hours interval between each spray before 5-7 days-old siliques were harvested for RNA extraction. To study reciprocal effect of silencing of *TTG1* and *TT8,* estradiol-inducible XVE system was generated for amiRNA targeted to *TT8* and *TTG1*. In the inducible silencing experiment, 10µM β-estradiol was applied to the 5-day-old seedlings for 24 h before

harvesting tissues for RNA extraction and expression analysis. Cycloheximide treatment was performed similar to dex treatment with final concentration as 100µM used to apply on transgenic lines (Baudry et al., 2006).

3.13 Microarray hybridization and gene expression analysis

3.13.1 cDNA synthesis and sample labelling

Total RNA extracted from siliques or seedlings were tested for their integrity using Agilent 2100 bioanalyzer and 6000 RNA nano chip from Agilent. Samples that had RIN (RNA integration number) value above 8 were used for double stranded cDNA synthesis according to application note for Nimblegen microarray kit (Roche,USA). Total 10µg of RNA were used for cDNA synthesis using Superscript2® (Invitrogen, USA) following the manufacturer's protocol (Invitrogen). After conversion of RNA to cDNA; reaction mix was then purified using GeneJET PCR purification kit (Thermo Scientific, USA) to remove enzymes and exchange buffers from the reaction mix. Converted cDNA were then tested for its quality by Bioanalyzer™® run. Samples with cDNA size ranging from 200-800 bp with mean at 450 bp were selected to perform labelling and microarray hybridization. 1µg of total cDNA were used for Cy3 labelling using NimbleGen one-color cDNA labelling kit (Catalog Number: 06370411001) following manufacturer's protocol. Since Cy3 loses its sensitivity under light, whole experiment was performed with samples protected from light. Cy3 labelled cDNA samples were quantified and 4µg of labelled DNA were transferred to a new 1.5ml eppendrof tube. Contents of the tube were dried under vacuum and preserved in -80^0C protected from light before using them for microarray chip hybridization.

3.13.2 Hybridization and washing

NimbleGen 12x135K *Arabidopsis* (TAIR 9.0) arrays were used for gene expression analysis. Dried Cy3 labelled sample were re-suspended in 3.3µl of sample tracking control provided by NimbleGen microarray kit. Each sample was spiked with a unique sample tracking control. Samples were vortexed briefly and the contents were collected back by centrifugation. NimbleGen hybridization kit was used and following the manufacture's instruction, hybridization master mix was prepared. Prepared samples were added with 8.7µl of hybridization mix, and vortex for 15 s followed by collecting them back using centrifugation. Reaction mix were then incubated at 95ºC for 5 min protected from light. Sample tubes were then placed at 42ºC till microarray chip were ready for sample loading. For gene expression study, NimbleGen 12x135K arrays were used. NimbleGen mixer was mounted on microarray slides using precision mixer alignment tool following manufacturer's instructions. 8µl of prepared samples were loaded on individual port on microarray slides and ports were sealed using mixer seal material provided by NimbleGen. MAUI® Hybe Station (BioMicro® Systems) was kept at 42ºC for 3 h to ensure no temperature fluctuations were present before using them for hybridization. Microarray slides were placed in hybridization chamber and mixer was switched on to allow mixing of loaded samples on microarray slides for 16 h at 42ºC. After hybridization, microarray slides were washed using NimbleGen wash buffer kit and NimbleGen array processing accessory following manufacturer's protocol for washing microarray slides. Slides were dried by spinning in NimbleGen microarray dryer.

3.13.3 One colour array scanning and data extraction

Axon GenePix® 4100A (Molecular Devices, USA) scanner was used to scan microarray slides, which was performed at different power value termed as PM in its software ranging from 300 to 600. Images thus obtained were processed in Deva® software from NimbleGen. Each slide had 12 individual arrays which were all together in the scanned image. Deva software was used to burst the single slide image into individual array, which were then aligned using default parameters. Each slide was then matched with array technology to assign gene name to each spot and to assign a signal intensity associated with expression level of that gene. Default parameters were selected for background correction to the expression profile and RMA analysis was performed. Data thus generated were used to for gene expression analysis using Genespring 12 software. Data were subjected *quantile normalization* and baseline correction was made using mean of samples. Data were filtered by taking 20^{th} to 100^{th} percentile of signal value thus removing low signal gene, followed by filtering data based on coefficient of variance with <20% as cut off. Filtered data were then used for statistical analysis to identify differentially expressed genes in different conditions. Differentially expressed genes were selected based on p value<0.05 and 2-fold changes compared to its control.

3.14 Total RNA extraction from siliqs

50mg of siliques were homogenized using liquid N_2 and transferred to a pre-chilled 2ml tube with 8M LiCl with 2% β-mercaptoethanol and incubated at 4°C for overnight. This mixture was centrifuged at 13,500 rpm at 4°C for 30 min. Supernatant were removed and pallets were added with 75% EtOH kept at -20°C, washed by shaking vigorously and centrifuged at 13,500 rpm at 4°C for 5 min. Supernatant were

removed and pellets were allowed to air dry briefly. Pellets were resuspended in 800µl of solubilisation buffer (100mM NaCl, 10mM Tris/HCl pH 7.6, 0.5% SDS, 25mM EDTA, 2% β-mercaptoethanol (freshly added)) followed by adding 800µl of chloroform, shaken vigorously for 20 s and centrifuged at 4°C, 13,500rpm for 15 min. Aqueous phase was transferred to a new Eppendrof tube and previous step was repeated with phenol, phenol:chloroform:isoamylalcohol (25:24:1), chloroform/ isoamylalcohol (24:1). Last aqueous phase was distributed into 2 tubes containing 600µl of 100% EtOH and 40µl of 3M NaOAc and mixture was incubated at -20°C for 90 min. This was followed by centrifugation at 4°C; 13,500 rpm for 5min, and supernatant were removed. Pellets were washed with 1ml of cold 70% EtOH and centrifuged at 13,500 for 5 min. Supernatants were removed and pellets were air-dried briefly, and total RNA was dissolved in 30µl of DEPC-treated water.

3.15 Metabolomics of *Arabidopsis* lines

3.15.1 Extraction of metabolites

Plant samples were snap frozen by liquid N_2 and homogenized using mortar and pestle. 100mg of homogenized samples were transferred to a 1.5ml Eppendrof tube and 1ml of ice cold 80% methanol was added immediately. This was followed by centrifugation at 4°C, 13,500 rpm for 30 min. Supernatants were transferred to a new Eppendrof tube and centrifuged at maximum speed at 4°C for 10 min. Supernatants were collected and were used to analyze through mass spectrometry.

3.15.2 Ultra high performance liquid chromatography

Agilent 1290 UHPLC system was used for the separation of metabolites through chromatography. Zorbox eclipse plus C18 column with particle size 1.8 µm and

dimensions of 10cm length and 2.1cm diameter was used in-line with UHPLC to perform metabolite separation based their hydrophobicity. Acetonitrile (buffer B) and water (buffer A) were acidified by adding 0.1% formic acid in and used as organic and aqueous buffer, respectively. To achieve separation of metabolites based on their chemical properties, a linear gradient was established across column for 18 min with a flow rate of 300µl/min. The column was subjected to 10% of buffer B for 0.5 min followed by a linear gradient to reach to 100% buffer B at 12^{th} min, was maintained at 100% buffer B till 15^{th} min followed by 10% buffer B at 15.1 min and was maintained till 18^{th} min of run. Column was maintained at 50ºC and 6µl as injection volume. Each run was followed by washing of the column for 5 min using 50% buffer B at 0.4ml/min flow rate.

3.15.3 Electrospray Ionization Mass Spectrometry

Untargeted Mass Profiling was performed by Agilent quadruple time of flight mass spectrometry (Agilent Q-ToF 6540) using ESI probe in positive mode of ionization. UHPLC system with column was in-line with mass spectrometry, and sample run time was 18 min. Drying gas temperature was set at 350ºC with 10L/min (nitrogen) flow rate, sheath gas temperature was set at 400ºC with 12L/min flow rate and capillary voltage was set at 4000V. Data acquisition was performed in centroide mode for MS scan with acquisition rate of 4 spectra/s with mass range set to 50-1200 m/z. Samples were in triplicates and each sample was injected 3 times as technical replicates.

3.15.4 Metabolite data analysis

Metabolic profile data was acquired in ".d" format using Data Acquisition software. Mass Hunter software was used to de-convolute the raw data and data was extracted based on their mass spectral and chromatographic characteristics, a process called as Molecular Feature Extraction (MFE). Any feature (m/z) with a signal value of less than 500cps was considered as noise and deleted from any analysis. Metabolite features with single positive charge with +H, +Na and +K adducts allowed were considered for data extraction and further analysis. Extracted data were then subjected to statistical analysis using Mass Profile Professional software from Agilent. Mass alignment was performed with 2mDa±15ppm window and RT correction was performed with 2mDa±15ppm as parameter. Samples were grouped based on the treatment or mutant groups with allowing only those metabolite features which were present in at-least 75% of a particular group. Aligned metabolic data with p value <0.05 and fold change 2 w.r.t control were selected for interpretation.

CHAPTER 4

RESULTS AND DISCUSSION

Data presented in this Chapter are on gene expression and metabolic state of gain- and loss-of-function transgenic lines and associated phenotypes. First, results are presented from gain-of-function transgenic lines that express a bacterial quercetin oxidoreductase (QuoA) encoding gene, which could induce perturbation in phenylpropanoid metabolic network (Section 4.1). These lines were characterized for their lignin and anthocyanin related phenotypes, metabolites and gene expression levels. Second, in order to ascertain the observed phenotype and metabolic properties of QuoA transgenic lines were not due to QuoA activity on unspecific substrate, metabolic profiling of loss-of-function *tt* mutant lines was performed and compared with QuoA transgenic lines (Section 4.2). Among selected mutant lines, *tt6* and *ttg1* mutant lines were identified to have metabolic phenocopy of *QuoA* transgenic line. Next, metabolic profiling of double mutants of *tt6*, and *ttg1* with *tt3*, *tt4* and *tt5* mutant lines were performed and analysed to identify TTG1 as important for co-ordinated perturbation of phenylpropanoid metabolic network with distant pathways such as brassinosteroids and carotenoids biosynthesis pathways (Section 4.2). In the last part of this study, inducible over-expression lines of *TTG1* were used to further understand the role of transcriptional factor TTG1 and TT8, and to identify their direct targets (Section 4.3). Taken together, combination of non-targeted metabolic profiling and gene expression analysis of genetically perturbed lines was used to understand relationships between the branches of phenylpropanoid pathways and its association with other metabolic networks. All Sections start with a brief introduction

and background, followed by results and discussion. All figures for subsections are included at their end.

4.1 Properties of phenylpropanoid metabolic network revealed by comparing QuoA high, moderate and low expression transgenic lines and loss-of-function tt6 mutant line

We previously reported a *Pseudomonas putida* strain PML2 capable of utilizing quercetin as a sole carbon source for growth and elucidated its catabolic pathway for quercetin degradation using a comparative metabolomics approach (Pillai and Swarup, 2002). The corresponding gene has been designated as *QuoA*. Since, phenylpropanoid biosynthesis in plants involve formation of quercetin from naringenin, we envisaged that *QuoA* expression in plants will provide us with a genetic tool to "reverse" this biosynthetic step using gain-of-function approach (Fig 4.1) and enable us to monitor the perturbational effects in metabolic networks. In previous study we created gain-of-function *QuoA* transgenic lines in *Arabidopsis*. Sixteen independent *QuoA* gain-of-function transgenic lines were characterized using gene expression and metabolite profiling approaches. Non-targeted metabolic profiling of *QuoA* transgenic lines revealed several metabolic channels within the phenylpropanoid pathway. QuoA-mediated perturbation in phenylpropanoid metabolic network resulted in widespread changes in metabolic channelling within the phenylpropanoid network. From screening of sixteen independent QuoA transgenic lines (Reuben, 2008), namely *Q10*, *Q2* and *Q11*, were further characterized based on their high, moderate and low expression and activity of QuoA, respectively. QuoA activity was observed to be consistent with its expression levels in these three transgenic lines in converting quercetin to naringenin. Selected three transgenic lines

based on different levels of QuoA activity were, considered respectively, to have three levels of perturbation of phenylpropanoid metabolic network.

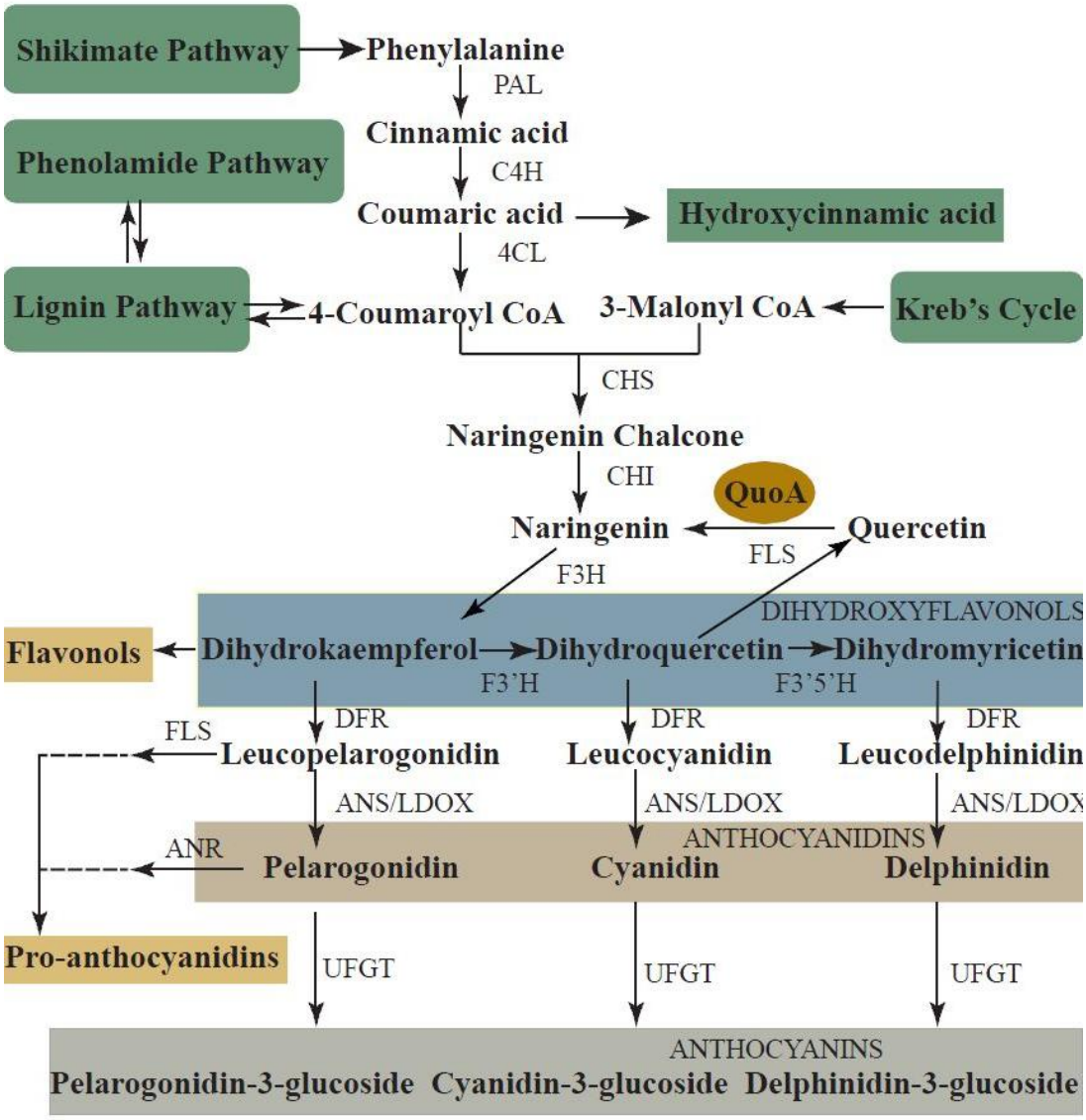


Fig 4.1 Phenylpropanoid metabolic network. A bacterial oxidoreductase QuoA that converts quercetin to naringenin was introduced in *Arabidopsis* to perturb phenylpropanoid network. Enzyme names are mentioned in Fig 2.1. Here, bold arrows refer to direct gene to substrate and product relationship while dotted arrows refer to additional metabolic reactions in between which have not been shown here.

(Adapted from Petroni et al., 2011) Abbreviations for genes have been shown in Fig 2.1.

4.1.1 QuoA expression results in increased coloration of hypocotyls and stem stiffness

In the *Arabidopsis* flavonoid pathway, naringenin is converted to quercetin via the intermediate dihydrokaempferol by flavanone-3-hydroxylase (F3H) encoded by *tt6*. Since this affected step is at the junction of upper and lower phenylpropanoid pathways, we reasoned that the perturbation could possibly have to change in both directions. Lower phenylpropanoid pathway leads to anthocyanin production, which gives a characteristic color to various plant parts. Indeed, *QuoA* transgenic seedlings had distinct purple coloration in their hypocotyls (Fig. 4.2A), which indicated anthocyanin accumulation. Transgenic lines, *Q*10 and *Q*2, showed high and moderate levels of anthocyanin accumulation in the hypocotyl region, respectively, while *Q*11 had similar level of anthocyanins as compared to Col-0. Levels of extracted anthocyanins in *Q*10 were more than double, while *Q*2 had 70% higher anthocyanin than that in Col-0 (Fig. 4.2B). Hence, anthocyanin levels paralleled those of *QuoA* gene expression and enzymatic activity levels, respectively. As expected, the F3H mutant line *tt6* had no observable coloration (Fig. 4.2A).

Having described hypocotyls coloration as an indicator of perturbation in the lower phenylpropanoid pathways, we investigated phenotypic indicators of the upper phenylpropanoid pathway. As F3H is involved in the formation of dihydrokaempferol from naringenin, which leads to formation of quercetin in biosynthetic pathways, we used F3H mutant line *tt6* to compare perturbation with QuoA expression lines. We focused on effects on lignin pathway based on two reasons. In the upper

phenylpropanoid pathway, lignin branch originates very closely from the point of perturbation by QuoA and second, that previous reports (Besseau et al., 2007; Vantrolime et al., 2008; Li et al., 2010) have established close relationship between flavonoid and lignin pathways. While *Arabidopsis* is not a woody plant, its lignification affects stiffness of stem internodes (Zhong et al., 2000, Allen et al., 2009). We, therefore, tested mechanical strength of stems of the three transgenic and four control lines using mechanical tester (Fig 4.3A). Young's modulus of stem internodal segments, which represents stiffness of the stem, was calculated. Tensile strength was significantly affected in the transgenic lines and the trend was in the following order, *Q*10>*Q*2>Col-0>*Q*11>Ler>*tt6*>ref8 (Fig. 4.3B). As expected, *HCT* mutant ref8, which has highly reduced lignin metabolites also had lowest tensile strength (Franke et al., 2002). The tensile strength of the three transgenic lines followed the same order as *QuoA* expression and enzymatic activity. In order to confirm that increase in stiffness is due to lignin accumulation, we stained internodal sections using phloroglucinol-HCl for lignin in the high expression line *Q*10. The lignin staining area is visibly larger in *Q*10 compared to the wild type (Fig. 4.4).

As QuoA acts at the flavonoid pathway step involving F3H (Fig 4.1), we used F3H mutant line, *tt6* to compare its effect with *QuoA* transgenic lines. Metabolic perturbations due to expression of *QuoA* in gain-of-function transgenic lines were expected to be in both upper and lower phenylpropanoid pathways. In comparison to *QuoA* lines which allow metabolic channelling in both upper and lower phenylpropanoid pathways, metabolic mutant *tt6* is expected to block lower pathway, resulting in effects only in upper pathway perturbation. As expected, stiffness of *tt6* stems was significantly lower while lignin staining area was smaller than Ler (Fig. 4.3

and 4.4). In order to identify which branched pathways could account for increased lignin in *QuoA* transgenic lines compared to *tt6*, we next quantified intermediate metabolites from the three linked pathways, namely, phenylpropanoid, lignin and phenolamide pathways.

A

B

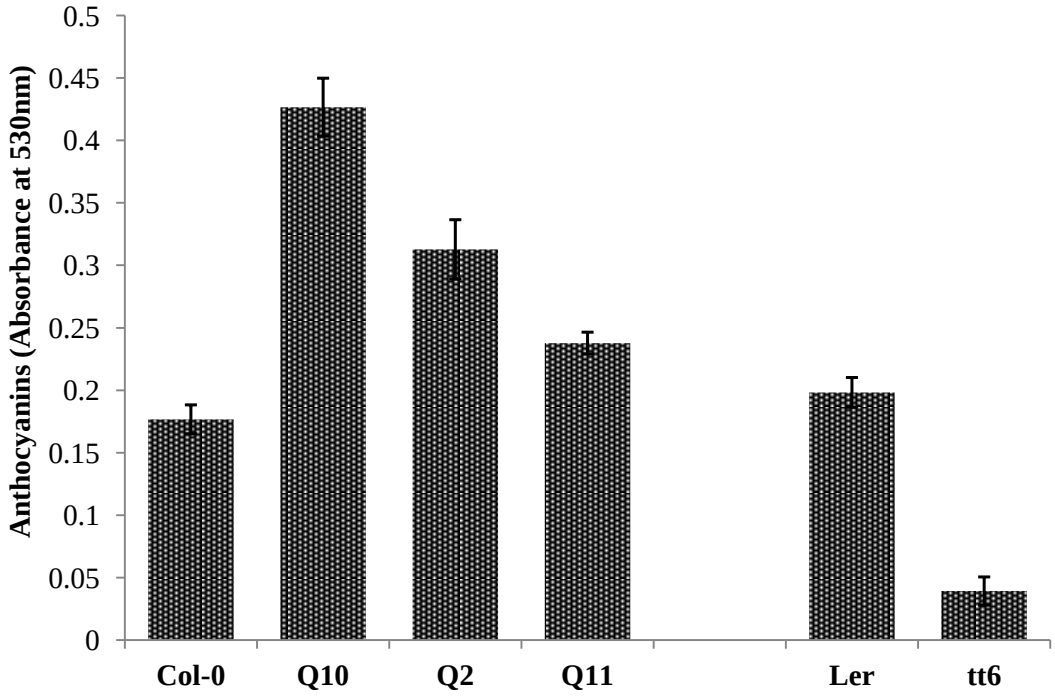
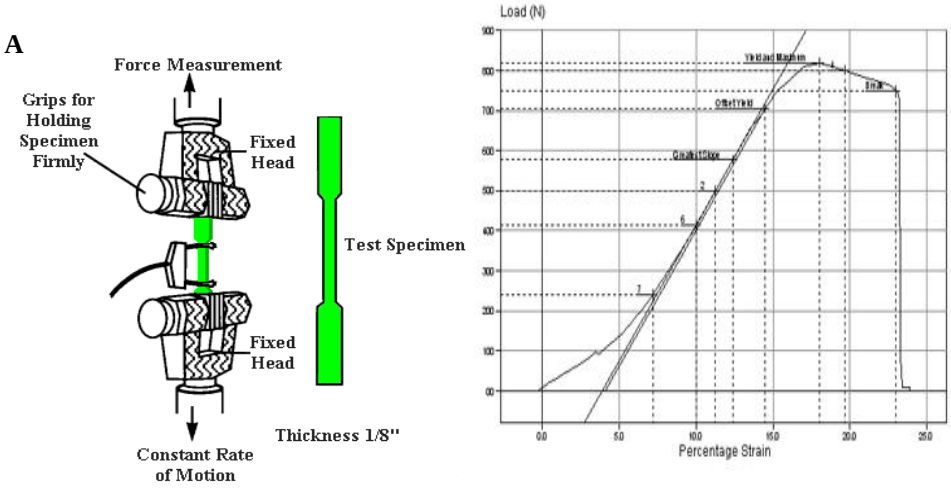


Fig 4.2 Accumulation of anthocyanin levels in QuoA transgenic lines. A, *QuoA* transgenic seedlings showing accumulation of anthocyanins, *tt6* mutant was used as negative control. B, Absorbance values of the anthocyanin extracts at 530nm. Anthocyanins were extracted from 6 day old seedlings using acidified methanol and

absorbance measured at 530nm. Col-0 and *tt6* seedling are shown as controls. All experiments were repeated three times. Standard error of mean is based on three replications ($p<0.01$).



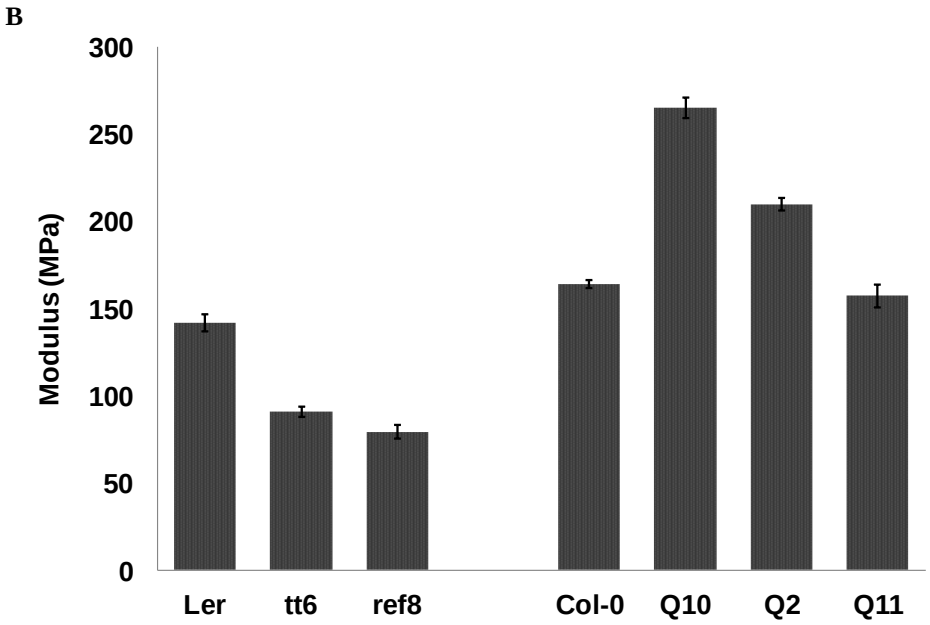


Fig 4.3 Stem stiffness increased in Q*uoA* transgenic transgenic lines. A, Schematic diagram of tensile tester used to measure stem tensile strength. B, Measurement of Young's Modulus of 35 day old stems of transgenic lines. Plants (n=9) were subjected to displacement speed of 0.5mm/min till the stem reached their break point in a tensile tester. Maximum load and extension at breaking point was measured and modulus calculated as stress over strain in MPa. Standard error of mean is based on nine replications ($p<0.0001$).

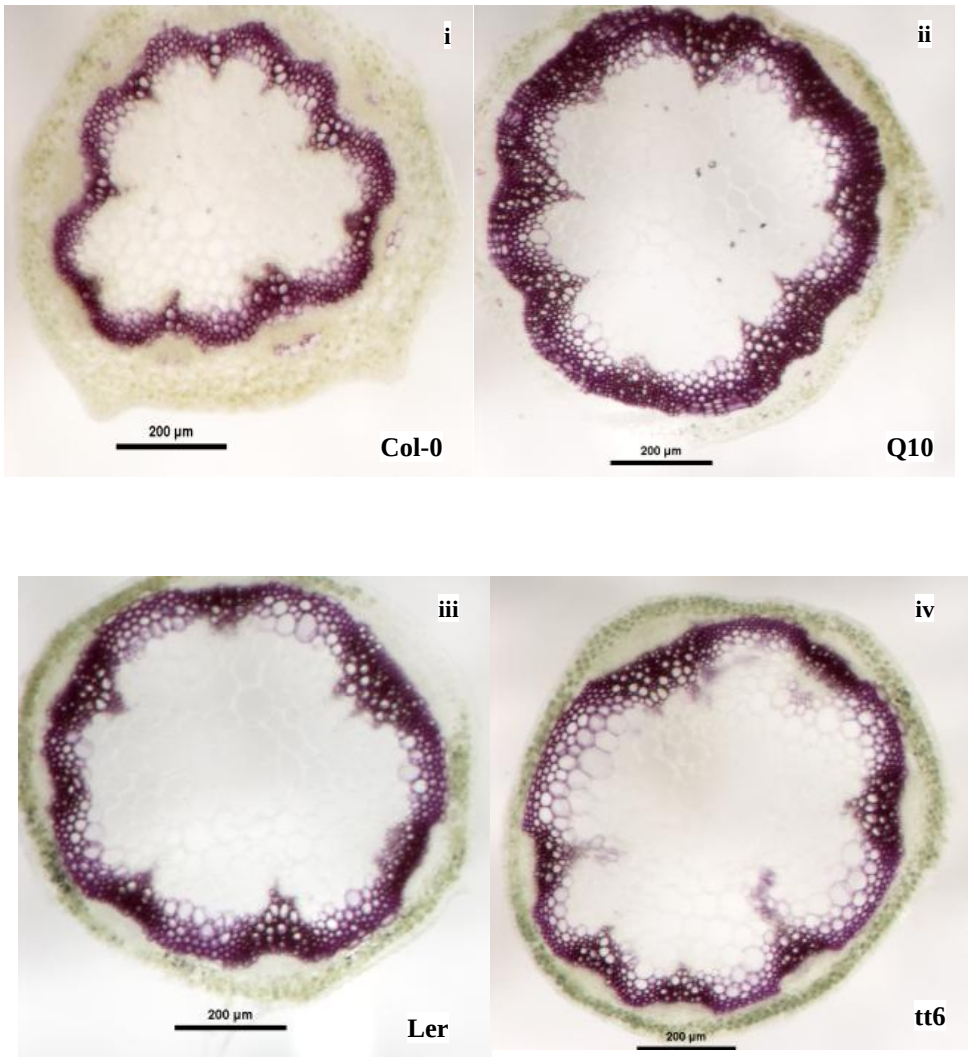


Fig 4.4 Lignin accumulation increased in *QuoA* transgenic lines. Phluoroglucinol staining of lignin in stem cross-sections of wild type Columbia and Landsberg (Col-0; Ler), *QuoA* transgenics (Q10) and *tt6* mutant. Infloresence stems from 5-week-old stems of QuoA transgenic lines were sectioned and stained with phloroglucinol-HCl.

All images were taken at 80x magnification under a dissection microscope with bright field illumination.

4.1.2 *QuoA* perturbation leads to increase in lignin and anthocyanin levels and decrease in levels of phenolamide intermediates

Metabolite levels from phenylpropanoid, phenolamide and lignin pathways were quantified in the three *QuoA* transgenic lines *Q*2, *Q*10 and *Q*11. Steady-state levels of quercetin were reduced by 1.5- to 2-folds and naringenin levels were increased by 4- to 7-folds in transgenic lines compared to wild type. These results can be explained by the *QuoA* activity that contributed towards naringenin pools at the expense of quercetin as substrate. As described in the previous section, increased hypocotyl coloration and anthocyanin accumulation of QuoA transgenic lines shows that metabolites from lower phenylpropanoid pathways increased in all transgenic lines. Hence, the levels of metabolites from both upper and lower pathways from the point of perturbation were significantly affected. Levels of cyanidin and leucopelargonidin increased 4- to 9-folds in the transgenic lines (Fig. 4.5), with highest level of change in *Q*10 compared to other transgenic lines, which is consistent with anthocyanin accumulation estimated in transgenic lines (Fig. 4.2B).

Among the four metabolites from the upper phenylpropanoid pathway, only 4-coumaroyl-CoA showed 1.6-fold significant increase ($p<0.001$) in transgenic lines when compared to Col-0. As the upper phenylpropanoid pathway is connected to lignin biosynthetic pathway via 4-coumaroyl-CoA, we tested perturbational effects in lignins. End-products and many intermediates of lignin metabolic pathways showed increased accumulation. Caffeoyl shikimate, coniferyl aldehyde, ferulate, 5-hydroxy

ferulic acid and end-product coniferin showed 4- to 10-folds accumulations in the transgenic lines with *Q*10 showing highest accumulation compared to other transgenic lines (Fig. 4.6). Highest change was that of coumaroyl alcohol with increase of 2- to 11-folds. We next investigated possible pathways connected to lignins that could have been down-regulated leading to increase in level of lignins. We, therefore, tested intermediates of recently described phenolamide pathway, which is linked to lignins at multiple branch-points (Matsuno et al., 2009). Interestingly, all the metabolites of the phenolamide pathway except tri-coumaroyl spermidine were present at lower levels compared to those in WT (Fig. 4.7). Spermidine levels showed the most decrease by 64-fold in comparison to WT.

As expected, the endogenous mutant *tt6* showed absence of anthocyanins due to block in lower phenylpropanoid pathways (Fig. 4.1 and 4.2). Cyanidin and leucopelargonidin were also not detected in *tt6* mutant line. The *tt6* mutant line had increase in naringenin level by 1.3-fold ($p<0.001$) (Fig. 4.5). Metabolic changes in the lignin pathway were not as high as in the gain-of-function lines, although same trend of increased metabolite levels was observed (Fig. 4.6). Ferulate and coniferin levels, for example, were increased by 3- to 4-folds in *tt6*. Several lignin metabolites such as coumaroyl alcohol, coumaraldehyde and sinapyl alcohol showed 1.3- to 2-fold increase in concentration when compared to Ler. In *tt6*, levels of phenolamide pathway metabolites were not significantly affected (Fig. 4.7).

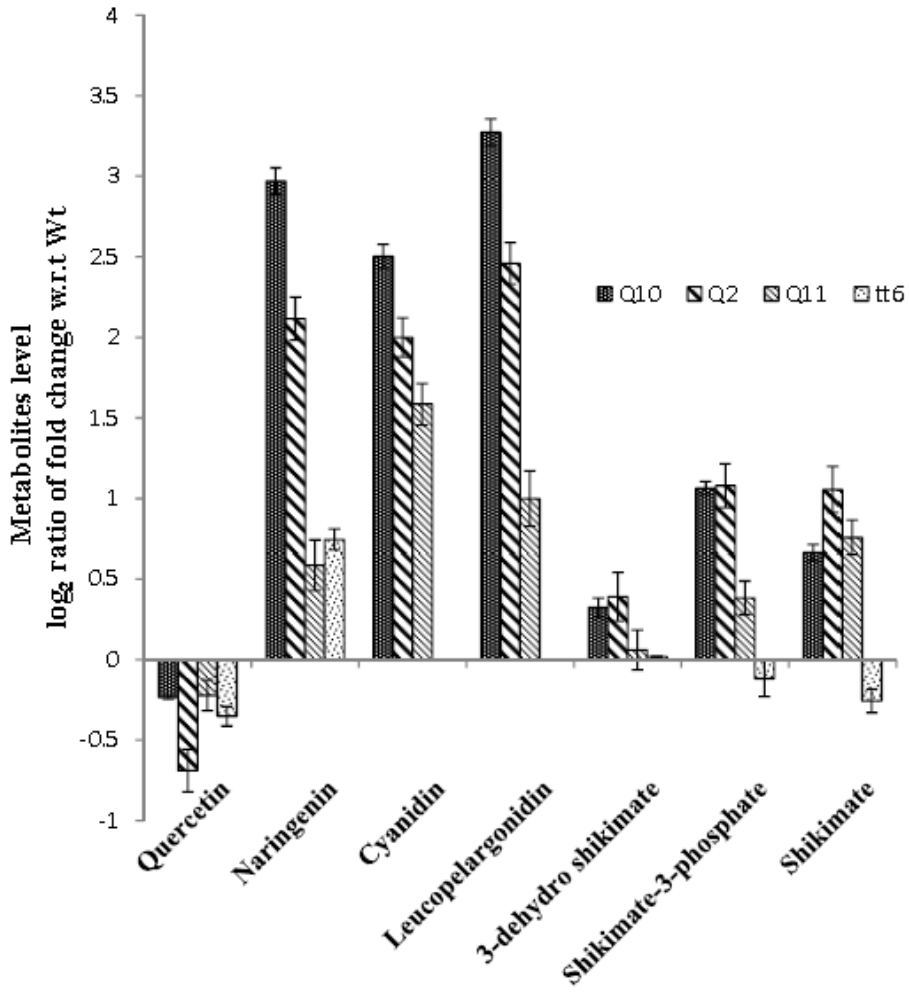


Fig 4.5 Quantification of metabolites from flavonoid metabolic pathways using multiple reaction monitoring (MRM) approach. Metabolites which showed differential levels with respect to control were quantified. *QuoA* transgenic lines were in Col-0 background, while *tt6* was in Ler background. Comparisons of *QuoA* and *tt6* lines were done with their respective wild type lines. Pure standards or Mass bank database (http://www.massbank.jp/index.html?lang=en) was used to select daughter ions for multiple ion monitoring. Standard error of mean is based on three biological replicates and experiments were repeated for three times.

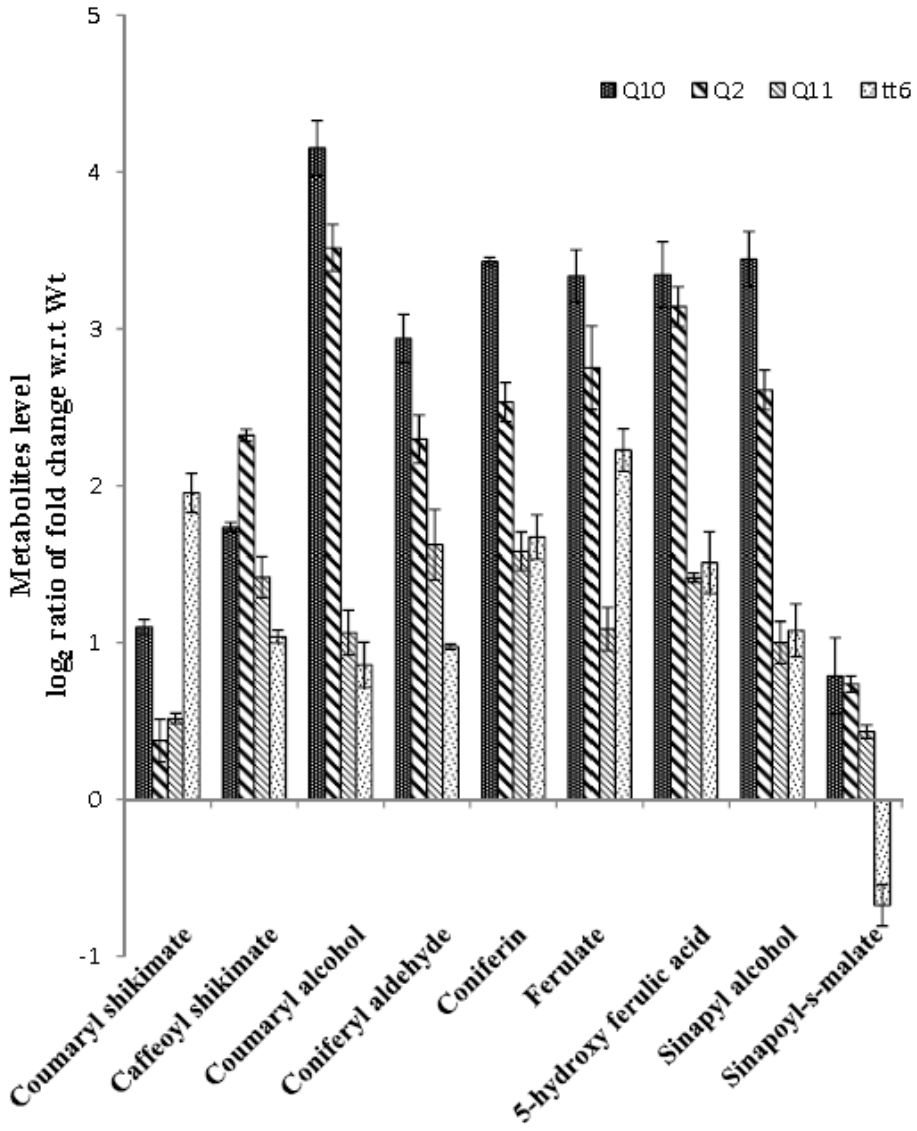


Fig 4.6 Quantification of metabolites from lignin metabolic pathways using multiple reaction monitoring (MRM) approach. Metabolites which showed differential levels with respect to control were quantified. *QuoA* transgenic lines were in Col-0 background while *tt6* was in Ler background. Metabolites from lignin metabolic pathways were all up-regulated in *QuoA* transgenic lines while moderately up-regulated in *tt6* mutant line. Pure standards or Mass bank database (http://www.massbank.jp/index.html?lang=en) was used to select daughter ions for multiple ion monitoring. Standard error of mean is based on three biological replicates and experiments were repeated three times.

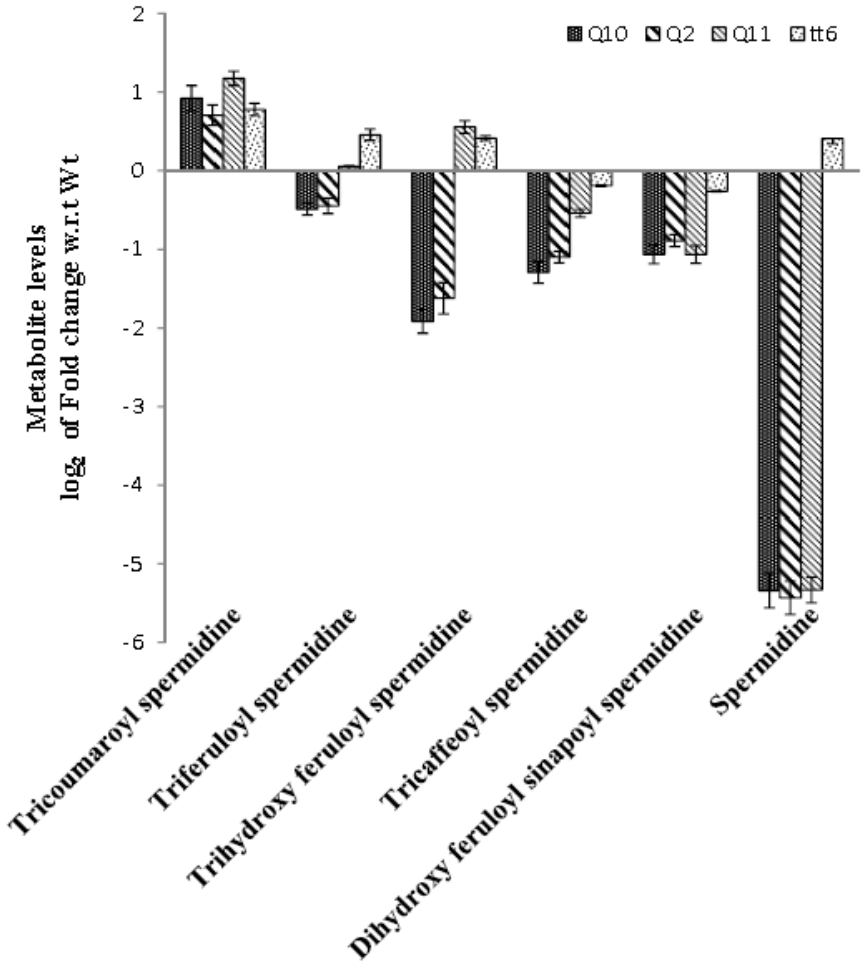
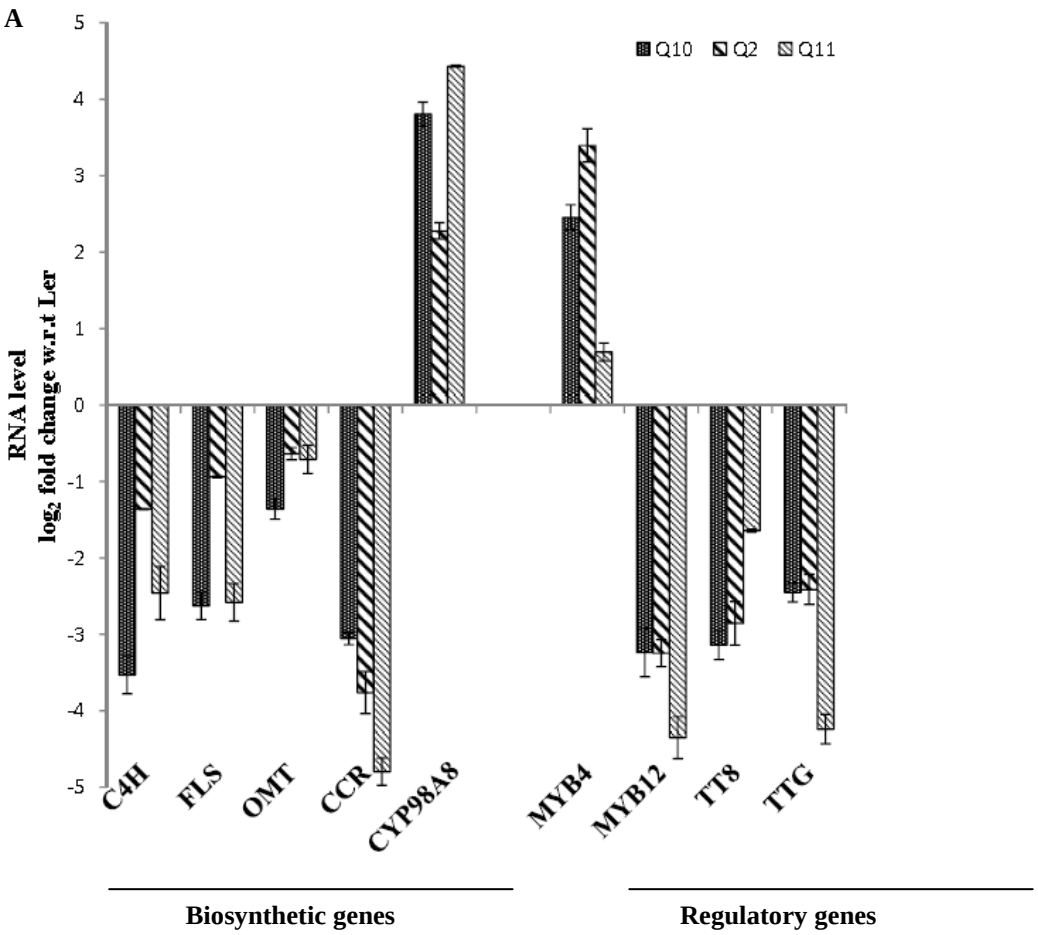


Fig 4.7 Multiple reaction monitoring (MRM) of metabolites from phenolamide metabolic pathways. Differential metabolites from phenolamide metabolic pathways were quantified. *QuoA* transgenic lines were in Col-0 background while *tt6* was in Ler background. Pure standards or Mass bank database (http://www.massbank.jp/index.html?lang=en) was used to select daughter ions for multiple ion monitoring. Standard error of mean is based on three biological replicates and experiments were repeated three times.

4.1.3 Metabolic perturbation by QuoA leads to changes in expression of biosynthetic genes

Since metabolite levels for multiple pathways were co-ordinately affected, we investigated whether some of these resulted from gene expression change or could be explained solely due to changes in metabolic channelling. We systematically compared the RNA levels for (eleven) genes associated with each step in the flavonoid and anthocyanin biosynthetic pathways and for key (six) genes in the lignin pathway. Expression levels of five genes, namely cinnamate-4-hydroxylase (*C4H*) and flavonol synthase (*FLS*) genes in the flavonoid pathway, a cytochrome 450 (*CYP98A8*) from phenolamide pathway, and cinnamoyl-CoA reductase (*CCR*) and flavonol 3'-o-methyltransferase (*OMT*), genes of the lignin pathway, showed differential expression in *QuoA* lines compared to Col-0 levels (Fig. 4.8A). In *tt6* mutant line, *C4H* did not show any expression change but chalcone synthase (*CHS*) showed ~16-fold increase in expression (Fig. 4.8B). There was no change in expression levels of *PAL*, 4-coumarate ligase (*4CL*), chalcone isomerase (*CHI*), *CHS,* dihydroflavonol-4-reductase (*DFR)*, flavanone-3-hydroxylase (*F3H*), flavanoid-3'-monooxygenase (*F'3H*) leucoanthocyanidin dioxygenase (*LDOX*) and banyuls (*BAN*) genes in transgenic plants. In the lignin pathway, *CCR* and *OMT* showed differential expression (Fig. 4.8B). Decrease in RNA levels of flavonol synthase (*FLS*), *CCR1* and *OMT1* showed corresponding changes in the metabolite levels in the transgenics. Significant change of 21-fold difference in expression was observed for *CYP98A8,* a cytochrome P450 involved in phenolamide pathway.


A
Q10
Q2
Q11
RNA level
log2 fold change w.r.t Ler
5
4
3
2
1
0
-1
-2
-3
-4
-5
C4H
FLS
OMT
CCR
CYP98A8
MYB4
MYB12
TT8
TTG
Biosynthetic genes
Regulatory genes

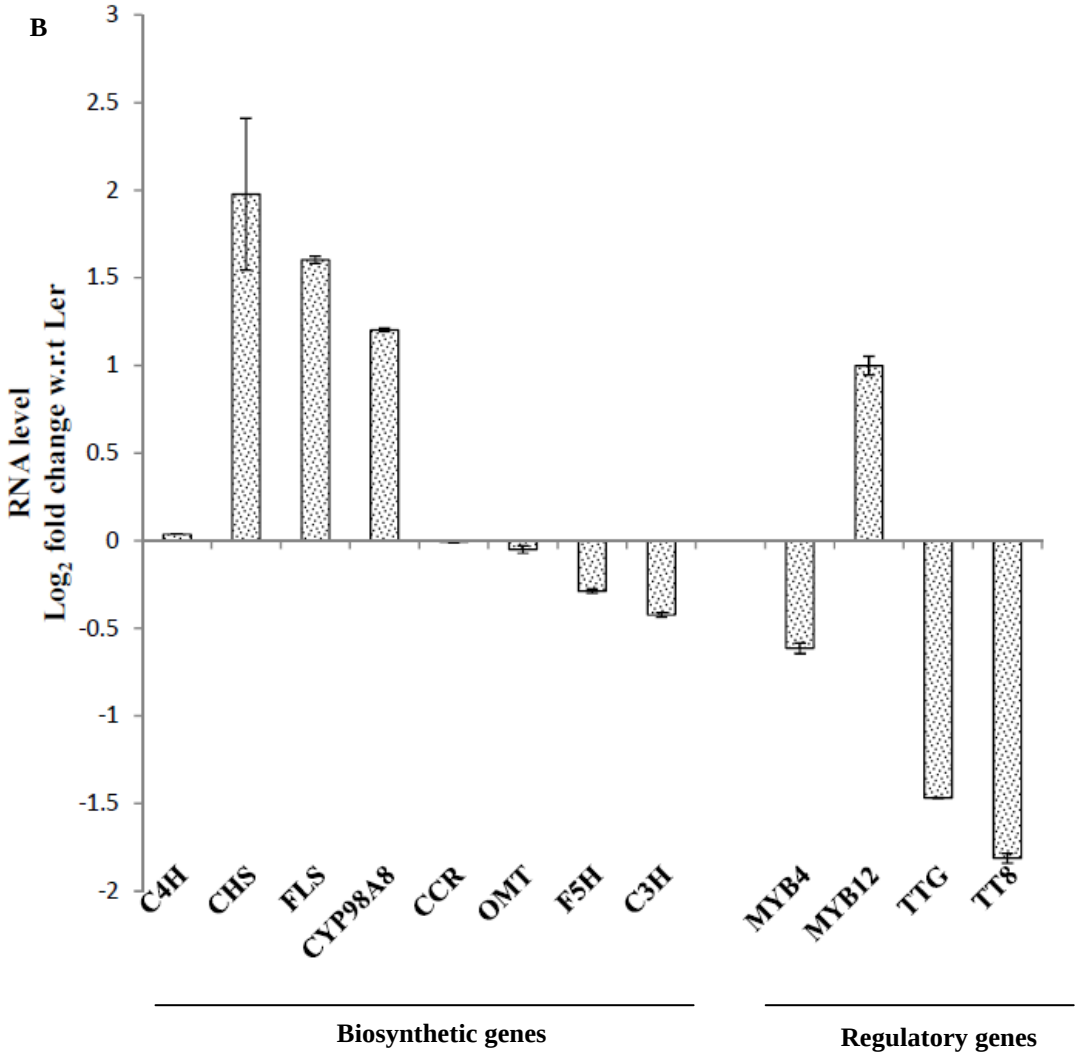


Fig 4.8 RNA levels of selected biosynthetic genes in 5-week-old *QuoA* transgenic lines and *tt6* mutant. A, RNA levels of biosynthetic genes in *QuoA* transgenic plants B, RNA levels of biosynthetic genes in *tt6* mutant. Real-time quantitative PCR was performed and Ct values normalized using tubulin as control. Fold changes were calculated based on ΔΔCt method. *C4H*: cinnamate hydroxylase *FLS*: Flavonol synthase, *CYP98A8*: Phenolamide pathway gene, *CCR*: cinnamoyl-CoA reductase, *OMT*: Flavonol 3'-o-methyltransferase. Standard error of mean is based on three replications and experiments were repeated two times for all genes.

4.1.4 Discussion

The overall nature of the metabolites, enzyme activity and transgene expression data in this study provide support that the perturbational effects in the transgenic lines result from *QuoA* expression and are specific in their nature in response to these perturbations. Three transgenic lines selected with low, medium and high levels of *QuoA* expression resulted in corresponding low, medium and high levels of *QuoA* enzyme activity, hypocotyl coloration due to anthocyanin accumulation and stem stiffness due to lignin accumulation, respectively. The overall *QuoA* perturbational effects observed are unlikely to be due to a general stress response. Several stress-related metabolite markers, such as proline, glycine, betaine or sugar alcohols (Wang et al., 2003), were not affected in *QuoA* lines (data not shown).

Perturbations in the phenylpropanoid network by the expression of bacterial *QuoA* allowed us to gain insights into the relationships between branched pathways. With the exception of F5H (Hang et al., 2009), reduction in the expression of at least 12 lignin biosynthetic genes in its loss-of-function mutants , leads to decrease in lignin content and increase in various flavonoid metabolites (Besseau et al., 2007; Li et al., 2008; Hoffmann et al., 2004; Vantrolime et al., 2008; Li et al., 2010). Lack of an increase in lignins in F3H mutant line *tt6* in this study is also consistent with the previous report. In contrast, there is very little known about the reciprocal effects of direct perturbation of flavonoid pathway on the lignin pathway branch. Simultaneous increase in both anthocyanin and lignins by *QuoA* mediated perturbation clearly establishes the reciprocal effects of flavonoids perturbation on the lignin pathway.

This effect on lignins is also biologically relevant as shown by the increased stem stiffness of *QuoA* transgenic lines.

One possible explanation of the difference in the perturbational effects between the gain-of-function and loss-of-function approaches here is that intact flavonoid enzyme complexes (Burbulis et al., 1999, Saslowsky et al., 2001, Conrado *et al*., 2008) are required for increased channelling of metabolites from flavonoid to lignin pathways. This possibility generates a testable hypothesis that some of the key enzymes of the lignin pathways are part of the flavonoid metabolic complexes. It is quite likely that change in metabolic channelling profile and increase in lignification have resulted from the *QuoA* gain-of-function approach. Phenolamide pathway, on the other hand, seems to be the likely candidate for increased channelling to lignin pathway in response to perturbation in flavonoid pathways. The low, medium and high expression lines of *QuoA* showed a corresponding effect on the overall reduction of the seven intermediates from the phenolamide pathway. The only exception was tri-coumaroyl spermidine, which is associated with the first step in the phenylpropanoid pathways (Fig 4.9). *F3H* mutant *tt6*, which has no increase in lignification, also does not have any perturbational effect on the phenolamide pathway. Lignin up-regulation from *QuoA* perturbation, therefore, seems to occur at the expense of phenolamide pathway intermediates. Our study, therefore, suggests a strong relationship between lignin and phenolamide pathways and opens-up an additional approach to manipulate lignification and stem stiffness characteristics, especially of herbaceous plants.

It is intriguing that *QuoA*-mediated perturbation led to up-regulation of both lignin and anthocyanin pathways, while flavonoid upper pathway intermediates were

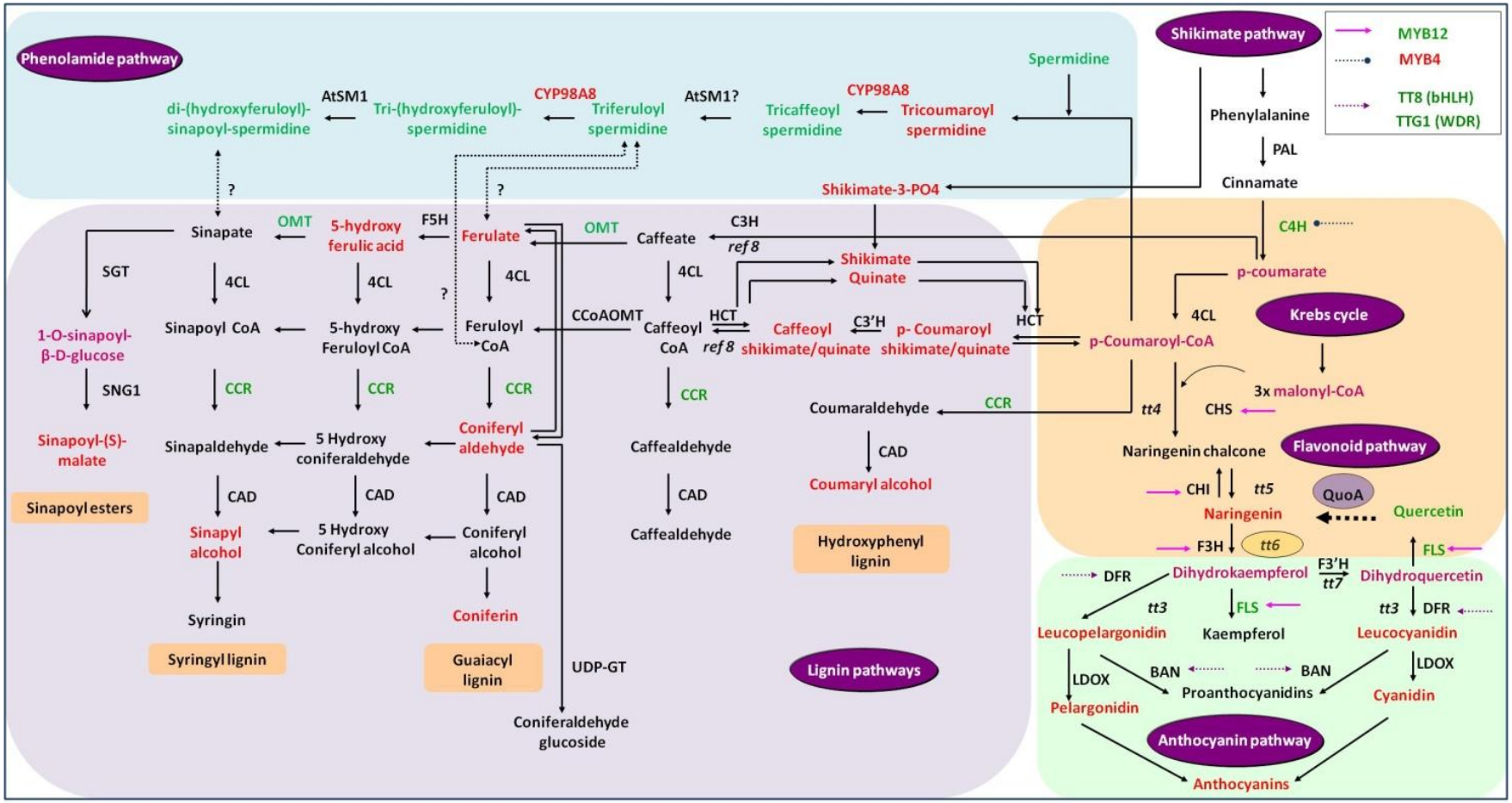
unaffected. So, this raises the question, which pathway contributes towards simultaneous up-regulation of the above pathways? There are no input pathways below the perturbed step in the flavonoid pathway. Hence, increase of the anthocyanins can only be explained by increased channelling from the flavonoid pathways. In the case of lignin pathway, three other input pathways are likely contributors towards its increased channelling. These are upper flavonoid pathways, malonyl-CoA from Krebs cycle and phenolamide pathway. Contributions by the first two pathways cannot explain increased channelling to lignin pathway. In both cases, their respective metabolites levels were unaffected by *QuoA*-induced perturbation. The only exception was tri-coumaroyl spermidine from phenylpropanoid pathway.

Our data strongly suggest that perturbation of metabolite levels is associated with a specific perturbation of gene expression level at the branch points. A complete survey of all genes from these four pathways revealed changes in expression of five genes, all of which are associated with branch points. In each case, the associated products encoded by these genes were also perturbed. The direction of these changes were similar in all cases, while all branch point genes were down-regulated, the products of these reactions were up-regulated.

Several studies have established that the first two steps of flavonoid pathway are under feedback control of metabolites (Peltier et al., 1999; Peer et al., 2001; Morreel et al., 2006; Marinova et al., 2007). More recently, flavonoid glycosylation has been shown to affect PAL expression levels (Yin et al., 2012). However, this metabolite to gene expression regulation seems to be more widespread in the overall phenylpropanoid network. The mechanism of control of the gene-metabolites pair at

the branch points cannot be ascertained. Both possibilities remain open that either there is coordinated control of gene-metabolite pairs via common mechanism or that there are local regulatory controls at each branch point arising from the common primary perturbational effect. It will be interesting to investigate the regulatory mechanisms at such branch-points in the metabolic networks.

Fig 4.9 Summary of perturbational effect metabolite and gene expression perturbation resulting from *QuoA* expression. The metabolites that showed over-accumulation and genes that were over-expressed in *QuoA* transgenic plant Q10 are indicated in red while those that are down-regulated are shown in green respectively. The coloured arrows indicate the target genes of the regulators that showed differential expression in this study. Index for regulators is shown in the inset box. Point of perturbation is boxed in light orange. Only metabolic pathways belonging to *Arabidopsis* are shown (based on TAIR www.arabidopsis.org; Matsuno *et al*., 2009, Umezawa, 2009). *PAL*: Phenylalanine ammonia lyase, *C4H*: Cinnamate 4-hydroxylase, *4CL*: 4-coumaroyl-coenzyme A ligase, *CHS*: Chalcone synthase, *CHI*: Chalcone-flavanone isomerase; *F3H*: Flavanone 3-hydroxylase, *F3'H*: Flavonoid 3'monooxygenase *DFR*: Dihydroflavonol 4-reductase, *FLS*: Flavonol synthase, *LAR*: Leucoanthocyanidin reductase, *LDOX*: Leucoanthocyanidin dioxygenase, *BAN*: Banyuls, *ANR*: Anthocyanidin reductase, *ANS*: Anthocyanidin synthase, *FGT*: Flavonoid glycosyl transferases, *C3H*: Coumarate-3-hydroxylase, *F5H*: Ferulate-5-hydroxylase, CCR1: Cinnamoyl-CoA reductase 1, *OMT1*: Flavonol 3'-o-methyltransferase, *SGT*: Sinapate-1-glucosyltransferase, *SNG1*: Sinapoylglucose:malate sinapoyltransferase, *HCT*: Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase, *CAD*: Cinnamyl alcohol dehydrogenase, *CCoAOMT*: Caffeoyl CoA o-methyltransferase, *UDP-GT*: UDP glucose transferase, *SHT*: Spermidine hydroxycinnamoyl transferase, bHLH: basic Helix-Loop-Helix, WDR: WD-repeat protein with WRKY domains, TT: transparent testa, *TTG1*: Transparent testa glabra 1. The fold changes are represented as >2 Red; < 2 Green; <1.3 Purple. All metabolites/genes represented in black fonts do not show changes.


Phenolamide pathway
Shikimate pathway
MYB12
MYB4
TT8 (bHLH)
TTG1 (WDR)
Spermidine
AtSM1
CYP98A8
AtSM1?
CYP98A8
di-(hydroxyferuloyl)-sinapoyl-spermidine
Tri-(hydroxyferuloyl)-spermidine
Triferuloyl spermidine
Tricaffeoyl spermidine
Tricoumaroyl spermidine
Phenylalanine
PAL
Cinnamate
Shikimate-3-PO4
C4H
p-coumarate
OMT
F5H
OMT
C3H
ref 8
Sinapate
5-hydroxy ferulic acid
Ferulate
Caffeate
Shikimate
Quinate
SGT
4CL
4CL
4CL
4CL
4CL
Krebs cycle
1-O-sinapoyl-β-D-glucose
Sinapoyl CoA
5-hydroxy Feruloyl CoA
Feruloyl CoA
CCoAOMT
Caffeoyl CoA
ref 8
HCT
Caffeoyl shikimate/quinate
C3'H
p- Coumaroyl shikimate/quinate
HCT
p-Coumaroyl-CoA
SNG1
CCR
CCR
CCR
CCR
3x malonyl-CoA
Coumaraldehyde
CCR
tt4
CHS
Sinapoyl-(S)-malate
Sinapaldehyde
5 Hydroxy coniferaldehyde
Coniferyl aldehyde
Caffealdehyde
Naringenin chalcone
Flavonoid pathway
CAD
CAD
CAD
CAD
CAD
Coumaryl alcohol
CHI
tt5
QuoA
Quercetin
Sinapoyl esters
Sinapyl alcohol
5 Hydroxy Coniferyl alcohol
Coniferyl alcohol
Caffealdehyde
Naringenin
F3H
tt6
FLS
Hydroxyphenyl lignin
DFR
Dihydrokaempferol
F3'H
tt7
Dihydroquercetin
tt3
FLS
tt3
DFR
Leucopelargonidin
Kaempferol
Leucocyanidin
Syringin
Coniferin
LDOX
BAN
BAN
LDOX
Syringyl lignin
Guaiacyl lignin
UDP-GT
Lignin pathways
Proanthocyanidins
Pelargonidin
Cyanidin
Anthocyanin pathway
Coniferaldehyde glucoside
Anthocyanins

4.2 Comparative analysis of metabolic profiles of (ten) *tt* mutants with *Q10* revealed *tt6* and *ttg1* mutant lines as its biochemical phenocopy

Preliminary characterization of (sixteen) independent QuoA transgenic lines showed several metabolic pathways to be perturbed in response to perturbation in phenylpropanoid metabolic network (Reuben, 2008). All transgenic lines were shown to have metabolic perturbation in pathways such as brassinosteroids, carotenoids, plant steroids, nucleotides biosynthesis pathways, which are parts of isoprenoid metabolic network. RNA levels of several biosynthetic genes from these pathways were also differentially expressed in *QuoA* transgenic lines. This suggested a possible crosstalk between phenylpropanoid and isoprenoid metabolic networks, which was maximum at young seedling stage and were diluted at later developmental stages of *Arabidopsis*.

Among *QuoA* transgenic lines, maximum effect of phenylpropanoid metabolic perturbation was observed in the isoprenoid network in *Q10* transgenic line, which is also high expression line of *QuoA* transgene. We investigated whether perturbation of these seemingly distant pathways were specific to *QuoA* or were also associated with loss-of-function of different phenylpropanoid genes. For this reason, we selected ten phenylpropanoid loss-of-function mutant lines for metabolic profiling of 6-day-old seedlings.

4.2.1 Loss-of-function mutant lines *tt6* and *ttg1* have similar metabolic profiles as that of *Q10*

Perturbation in different branches of isoprenoid metabolic network in response to perturbation created in phenylpropanoid metabolic network in *QuoA* transgenic lines led us to ask if this is specific to QuoA activity or due to native pathways being co-ordinately regulated with phenylpropanoid metabolic network. To address this question, *tt* mutant lines *tt1, tt2, tt3, tt4, tt5, tt6, tt7, tt8, tt8 (3), tt9, tt10* and *ttg1* were profiled for their metabolites. These *tt* mutants were selected based on loss-of-function of structural genes and transcription factors at different points of lower phenylpropanoid metabolic network. Among those selected transparent testa mutant lines, *tt3, tt4, tt5, tt6, tt7, tt9* and *tt10* are loss-of-function mutant of biosynthetic genes of phenylpropanoid metabolic network while *tt1, tt2, tt8, tt8(3)* and *ttg1* are loss-of-function mutants of transcription factors. Differential metabolites with p-value of 0.05 and fold-change of 2 were then compared across transparent testa differentially expressed metabolites.

Hierarchical clustering of *tt* mutant lines based on metabolic profiling showed formation of three groups (Fig 4.10A). First group (cluster 1) constituted by *tt2, tt8, tt8 (3),* Ler and Ws, second group (cluster 2) with *tt1, tt3, tt7* and *tt10* and third group (cluster 3) with *tt4, tt5, tt6, tt9* and *ttg1* mutants lines with ttg1 and tt6 clustered together (Fig 4.10A). Cluster 1 represents transcription factors mutant lines, *tt8, tt2* and *tt8 (3)* which are loss-of-function mutants of TT2 and TT8. Both TT2 and TT8 have been shown to interact and form protein complex to regulate expression of BAN and DFR, biosynthetic genes from lower phenylpropanoid metabolic network (Baudry et al., 2004). Cluster 2 consisted of *tt1, tt3, tt7* and *tt10,* all of which are associated

with lower phenylpropanoid metabolic network. Except *tt1*, all are loss-of-function mutants of biosynthetic genes. Cluster 3 consisted of *tt4, tt5, tt6, tt9* and *ttg1*, all of which except *ttg1* are loss-of-function mutants at the junction of upper and lower phenylpropanoid metabolic networks. PCA plot based on metabolic profiling of *tt* mutant lines showed similar pattern with *ttg1* and *tt6* clustering together while transcription factors and structural genes formed separate clusters (Fig 4.10B).

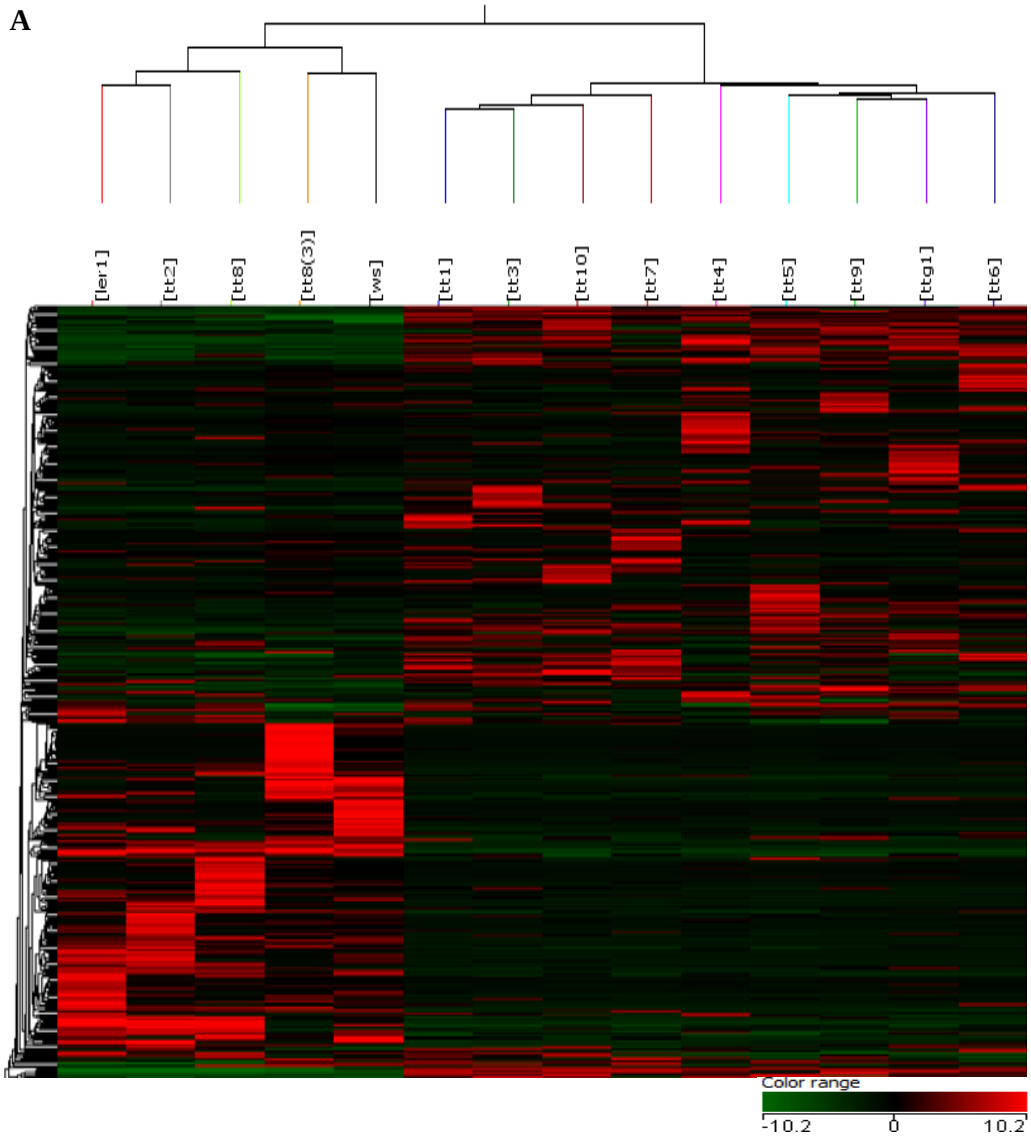
Comparison of differential metabolites common between *Q10* and *tt* mutant lines was done to test if there is any similarity between metabolic profile of *Q10* resulting from QuoA activity and loss-of-function of native genes in Arabidopsis. Differential metabolites between *Q10* and *tt* mutant with p-value 0.05 and fold-change 2 were selected. Fold change values of common metabolites of *tt* mutant lines were subtracted from the fold change values of corresponding metabolites in *Q10*. These values were then plotted as a box plot (Fig 4.11). Transgenic lines with maximum similarity in perturbational pattern with *Q10* were closer to the central line or 0 values, sharing higher similarity in metabolic perturbation behavior, while those transgenic lines with least similarity were far away from central lines. Among all *tt* loss-of-function transgenic lines, two *tt* mutants - *tt6*, loss-of-function of F3H, and *ttg1*, loss-of-function mutant of transcription factor TTG1, were shown to have similar biochemical phenotype as of *Q10*.

As shown earlier, perturbation in phenylpropanoid metabolic network due to loss of *F3H* and due to QuoA activity results in accumulation of same metabolite naringenin (Fig 4.5), which is restricted to upper phenylpropanoid metabolic network in case of *tt6*, while may be channelled to both upper and lower pathways in *Q10*. Hence,

similarity of *tt6* metabolic profile was corroborated with the results of immediate effects of QuoA mediated perturbation in *Q10* transgenic lines. Showing similar metabolic perturbational behavior of ttg1 mutant line as that of *tt6* and *Q10* was a novel finding, since the point of perturbation was different than that of *tt6* and *Q10*. Differential metabolites of *tt6* were mapped on biosynthetic pathways which were mainly from carotenoids, brassinosteroids, steroids, phenylpropanoid and nucleotide synthesis pathway (Fig 4.12), similar to *ttg1* showing same pattern of perturbation in isoprenoid metabolic network.

In order to investigate possible causes of coordinate changes in multiple pathways, we tested RNA level changes of genes involved in phenylpropanoid and then selected pathways. RNA levels of *TTG1, TT8* and *F3H* were tested in *Q10, tt6, tt8* and *ttg1* transgenic lines. RNA levels of *TTG1* and *TT8* were down regulated in all the four transgenic lines (Fig 4.13). With the exception of *tt8* mutant lines, *F3H* was shown to be down regulated in other three transgenic lines. Down regulation of *TT8* and *TTG1* in all the four transgenic lines were consistent with the fact that these two proteins form a complex that co-ordinately regulates biosynthetic genes. However, down regulation of *F3H* in *Q10* and ttg1 transgenic line was not expected since TTG1 does not regulate expression of F3H. Results thus established that perturbation in isoprenoid metabolic network in response to perturbation in phenylpropanoid metabolic network was not specific to QuoA activity, but loss-of-function mutants of native genes, *F3H* and *TTG1* also showed similar perturbational pattern. Since other mutants of phenylpropanoid metabolic network did not show perturbation in isoprenoid network, this suggested existence of unique properties of metabolic network in ttg1 and tt6 loss-of-function mutant.

A


[ler1]
[tt2]
[tt8]
[tt8(3)]
[ws]
[tt1]
[tt3]
[tt10]
[tt7]
[tt4]
[tt5]
[tt9]
[ttg1]
[tt6]
Color range
-10.2
0
10.2

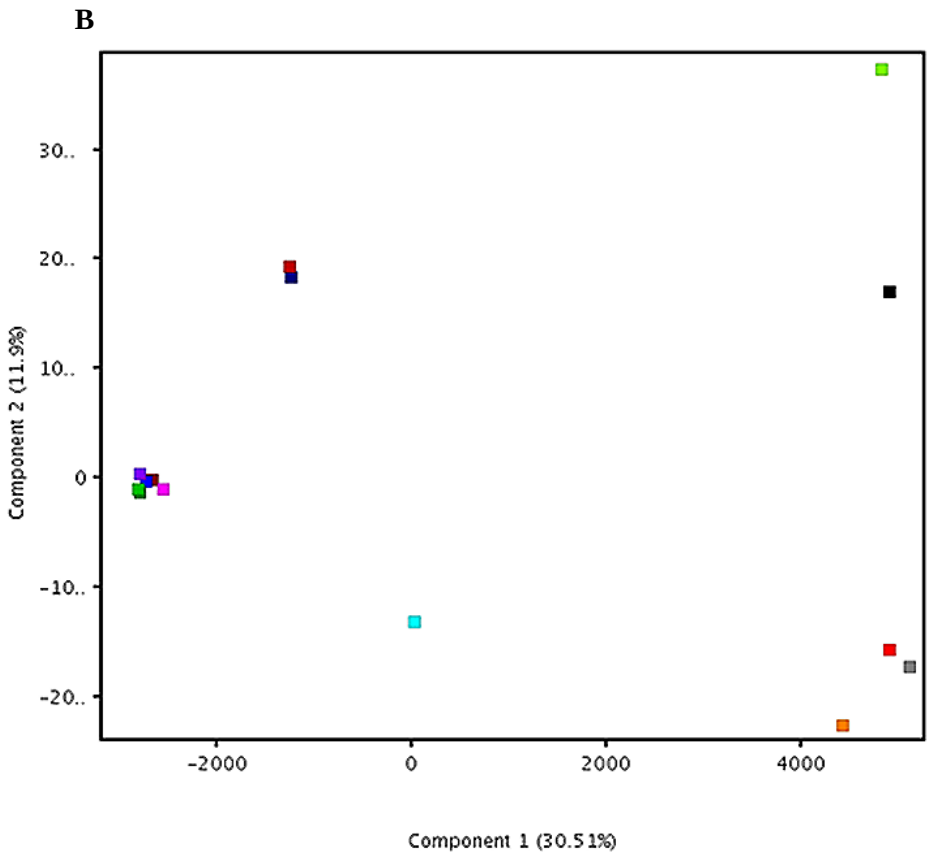


Fig 4.10 Hierarchical clustering (A) and 2D-PCA plot (B) of metabolic profile of *tt* mutant lines. Metabolites were extracted from 6-day-old seedlings and untargeted metabolic profiling of four biological replicate of mutant lines were performed, and each metabolic extract was injected 3 times to serve as technical replicate. Differential metabolites from all *tt* mutant lines were filtered with p-value 0.05 and fold-change 2. Experiments were repeated for two times. Both plots showed *tt6* and *ttg1* grouped together. Colour codes for hierarchical clustering refer to normalized intensity of a particular metabolite in transgenic lines.

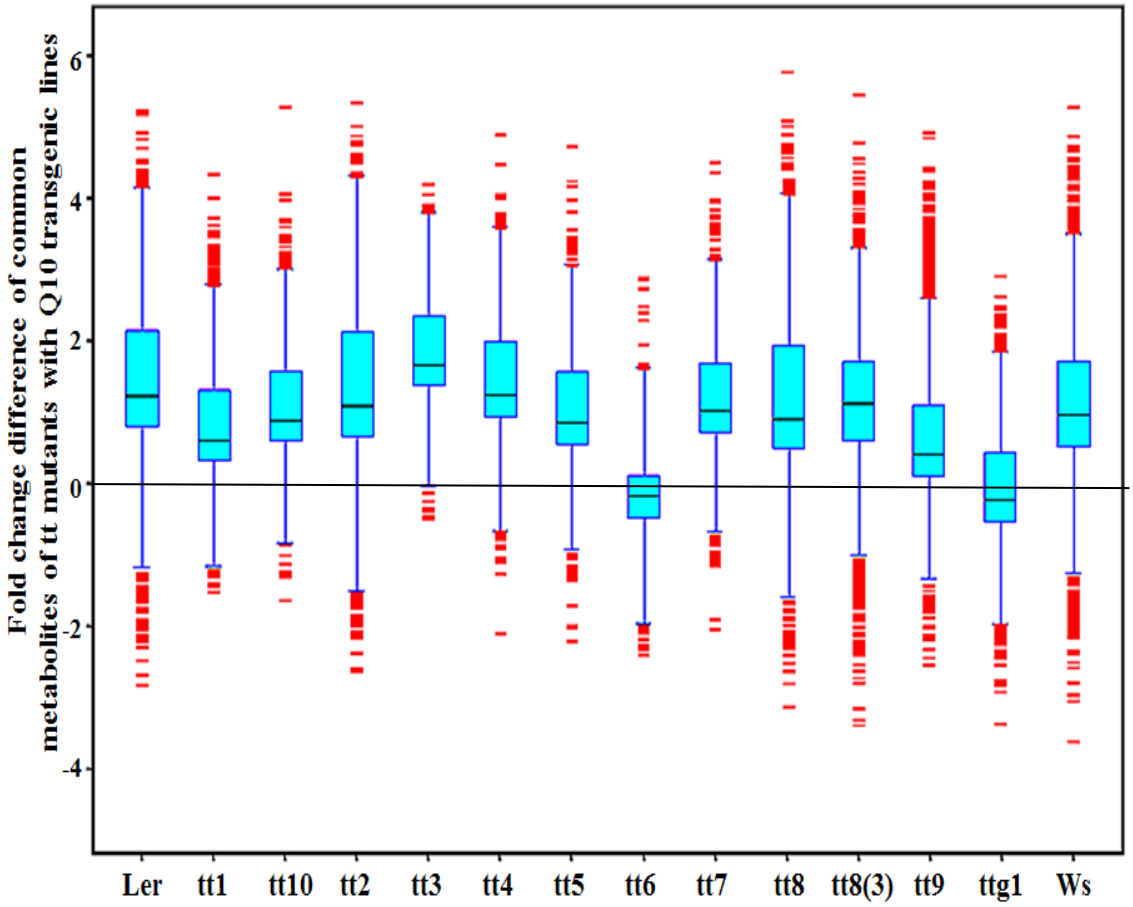


Fig 4.11 Box plot showing similarities between different phenylpropanoid mutants with that of *QuoA* transgenics. The differential metabolites were first determined in *tt* mutants. Those metabolites that were common in the mutants and *Q10* were then selected. The fold change for these metabolites in *tt* mutants w.r.t to *Q10* has been plotted in this figure. This shows the similarity between transgenics and *tt* mutants both in the type of metabolites and in the accumulation levels (based on fold change). The top and bottom blue edges in each rectangular box are the 25th and 75th percentiles of the fold change values, respectively. The distances between the top and bottom edges represent the inter-quartile range for the data. The black horizontal line in the middle of each box is the sample median. If the median is not centred in the box, it shows sample skewness. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length (the adjacent values). Observations beyond the whisker length are marked as outliers (red + sign).

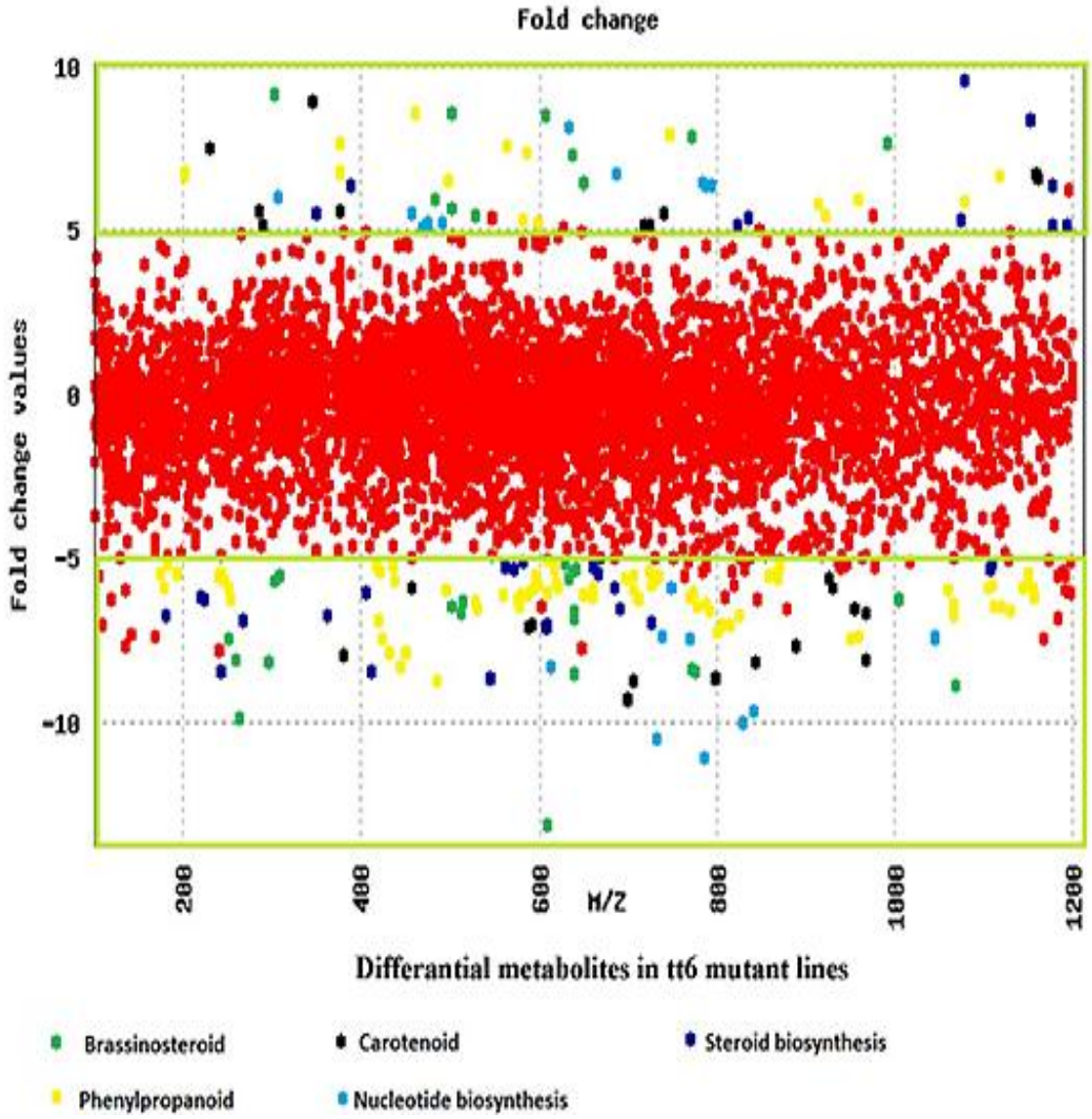


Fig 4.12 Differential metabolites of *tt6* w.r.t Ler were mainly associated to phenylpropanoid and isoprenoid metabolic network. Metabolites with +/- 2 fold changes and p-value 0.05 were used to perform further analysis and mapping on metabolic pathways. As shown in this figure, most of the differential metabolites of fold-change +/- 5 were from phenylpropanoid and different branches of isoprenoid metabolic network

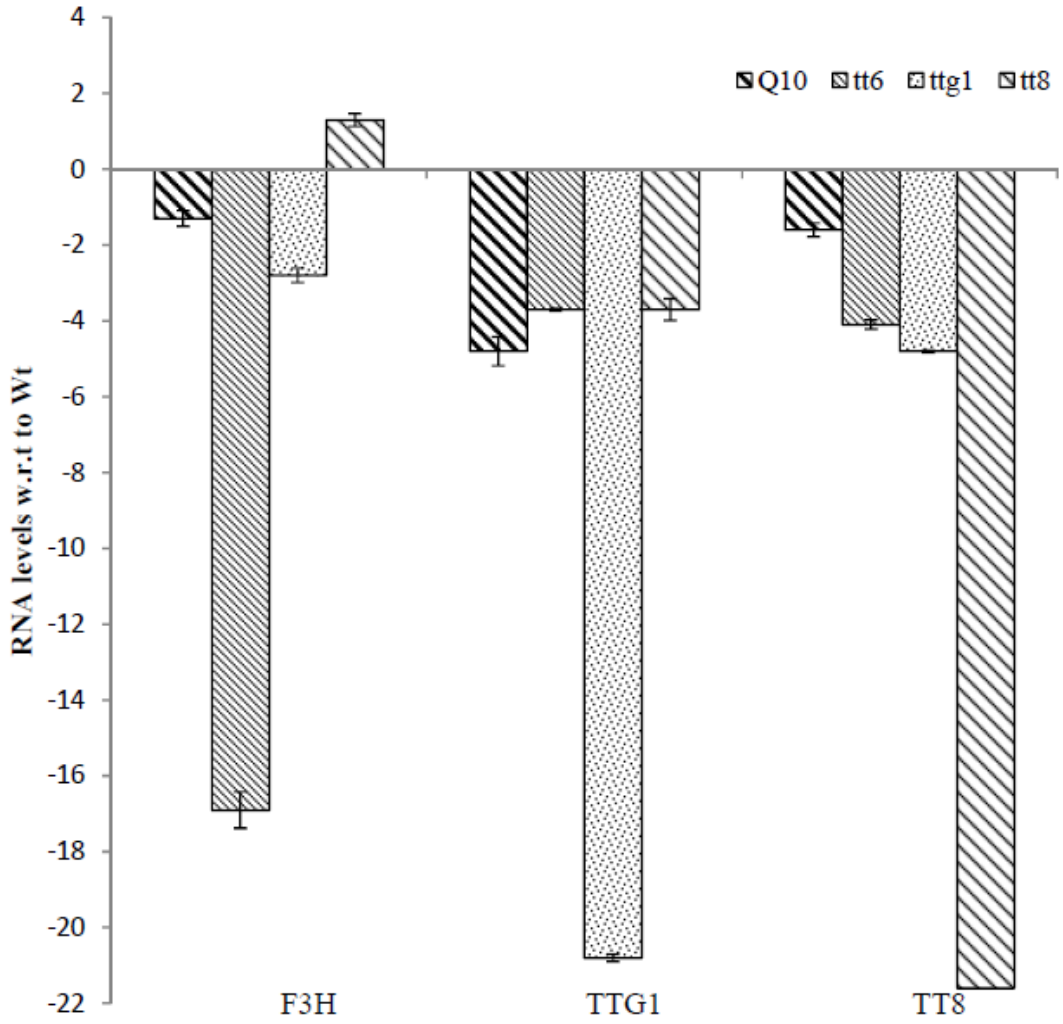


Fig 4.13 RNA levels of *F3H, TTG1* and *TT8* in 6 day old seedlings in Q10, tt6, tt8 and ttg1 transgenic line. RNA levels of F3H, TTG1 and TT8 were measured in *Q10, tt6, ttg1* and *tt8* mutant lines by real-time quantitative PCR. Ct values were normalized using tubulin as control. Fold changes were calculated based on ΔΔCt method. Standard error of mean is based on three biological replications and experiments were repeated two times for all genes.

4.2.2 TTG1 loss-of-function in *tt3, tt4* and *tt5* single mutants caused perturbation in isoprenoid metabolic network

Among selected *tt* mutants, F3H and TTG1 loss-of-function mutant lines showed perturbed isoprenoid metabolic network in response to perturbation in phenylpropanoid metabolic network. Metabolic perturbation was similar to *Q10* transgenic line, which also showed down-regulation of RNA levels of *F3H* and *TTG1*. To address the question if F3H or TTG1 loss-of-function caused perturbation in isoprenoid metabolic network, we performed metabolic profiling of single mutants *tt3*, *tt4*, *tt5* and their double mutants with *ttg1* or *tt6*. We reasoned that since *tt3*, *tt4* and *tt5* did not have perturbation in isoprenoid metabolic network, if their double mutant with *ttg1* or *tt6* introduces perturbation in isoprenoid metabolic network, we may conclude TTG1 or F3H loss of function as being required to bring this change. Metabolites were extracted from 6-day-old seedlings of *tt3, tt4, tt5, tt3-ttg1, tt3-tt6, tt4-ttg1, tt4-tt6, tt5-ttg1, tt5-tt6* and *tt6-ttg1* mutant lines. Non-targeted metabolic profiling was performed. Hierarchical clustering and heat map of single and double mutants based on their metabolic profiling showed formation of two major clusters (Fig 4.14). First cluster consisted of *tt3-ttg1, tt6-ttg1, tt4-ttg1* and *tt5-ttg1*. Second cluster had further four sub-clusters, *tt6* and *ttg1* was clustered together forming first group, Wt, *tt5*, *tt6-tt5*, *tt6-tt3* and *tt6-tt4* being the second group while *tt3* and *tt4* were represented by individual groups. As expected, *ttg1* and *tt6* were clustered together thus strengthening earlier observation of similar metabolic profile. Hierarchical clustering of single and double mutants showed the dominance of TTG1 in determining the metabolic perturbation behavior (Fig 4.14). *ttg1* double mutants with *tt3*, *tt4*, *tt5* and *tt6* formed a single cluster which was further closer to *ttg1* and *tt6*. This was not the case with *tt6*, for instance, *tt6* and *tt5-tt6* were clustered together

suggesting the dominance of *tt5* in determining the metabolic profile pattern. *tt6* double mutants with *tt3* and *tt4* were clustered together along with *tt3* and *tt4*. 2D PCA plot based on metabolic profiling of these loss-of-function single and double mutant lines also showed similar trend as observed by hierarchical clustering (Fig 4.15). 2D PCA plot and hierarchical clustering thus conclusively suggests dominance of *ttg1* in determining the metabolic profile of its double mutants with *tt3*, *tt4*, *tt5* and *tt6*.

In order to gain an insight into the effect of mutation of *ttg1* and *tt6* in *tt3*, *tt4* and *tt5* single mutant lines, the differentially expressed metabolites of *tt3*, *tt4* and *tt5* single mutant were compared with their corresponding double mutants with *ttg1* and *tt6*. As shown in Fig 4.16, effect of *tt6* mutation in *tt3*, *tt4* and *tt5* single mutant was less severe than that of ttg1. For example, differential metabolites between *ttg1-tt3* and *tt3* loss-of-function was 2231 as compared to only 644 differential metabolites between *tt6-tt3* and *tt3* loss-of-function (Fig 4.16). Similarly, other double mutants of *ttg1* showed more differential metabolites while double mutants of tt6 showed less differential metabolic ions. The result clearly suggests role of ttg1 in drastic change in metabolic profile of *tt3*, *tt4* and *tt5* single mutant as compared to tt6. Loss-of-function of TTG1 in *tt3*, *tt4* and *tt5* mutant lines showing more differential metabolites while fewer in case of F3H loss-of-function also suggests that transcriptional factor plays bigger role in regulating phenylpropanoid metabolic network than the metabolic channelling.

We next looked for the pathways that were perturbed by introducing mutation of TTG1 or F3H in *tt3*, *tt4*, *tt5* and *tt6* mutant lines. List of differentially expressed

metabolites in *ttg1* and *tt6* double mutants (with *tt3*, *tt4*, *tt5* and *tt6*) w.r.t single mutants *tt3*, *tt4* and *tt5* were fragmented and MS2 fragmentation pattern of daughter ions were matched to mass-bank (http://www.massbank.jp/index.html?lang=en) database to identified metabolites and those that were confirmed were mapped to respective metabolic pathways. Numbers of differential metabolites from a particular pathway were plotted in heat map to visualize metabolic pathways with number of differential metabolites in double mutant lines w.r.t single mutant lines. Double mutants of *ttg1* showed higher number of differentially expressed metabolites related to isoprenoids and its branch pathways, while double mutants of *tt6* showed fewer metabolites differentially expressed (Fig 4.17 and 4.18). Differential metabolites from ttg1 double mutants w.r.t single mutants were mainly from trans-zeatin biosynthetic pathway, phytate degradation I pathway, mevalonate pathway, cytokinin biosynthesis pathway and geranyl-diphosphate biosynthetic pathway. tt6 double mutants with *tt3*, *tt4* and *tt5* had fewer differential metabolites in isoprenoid metabolic network. Thus, comparing single and double mutant lines, we observed *TTG1* loss-of-function resulting in perturbation of phenylpropanoid and isoprenoid metabolic network. These results suggested TTG1 as main factor associated with co-ordinated metabolic perturbation of phenylpropanoid and isoprenoid metabolic network.

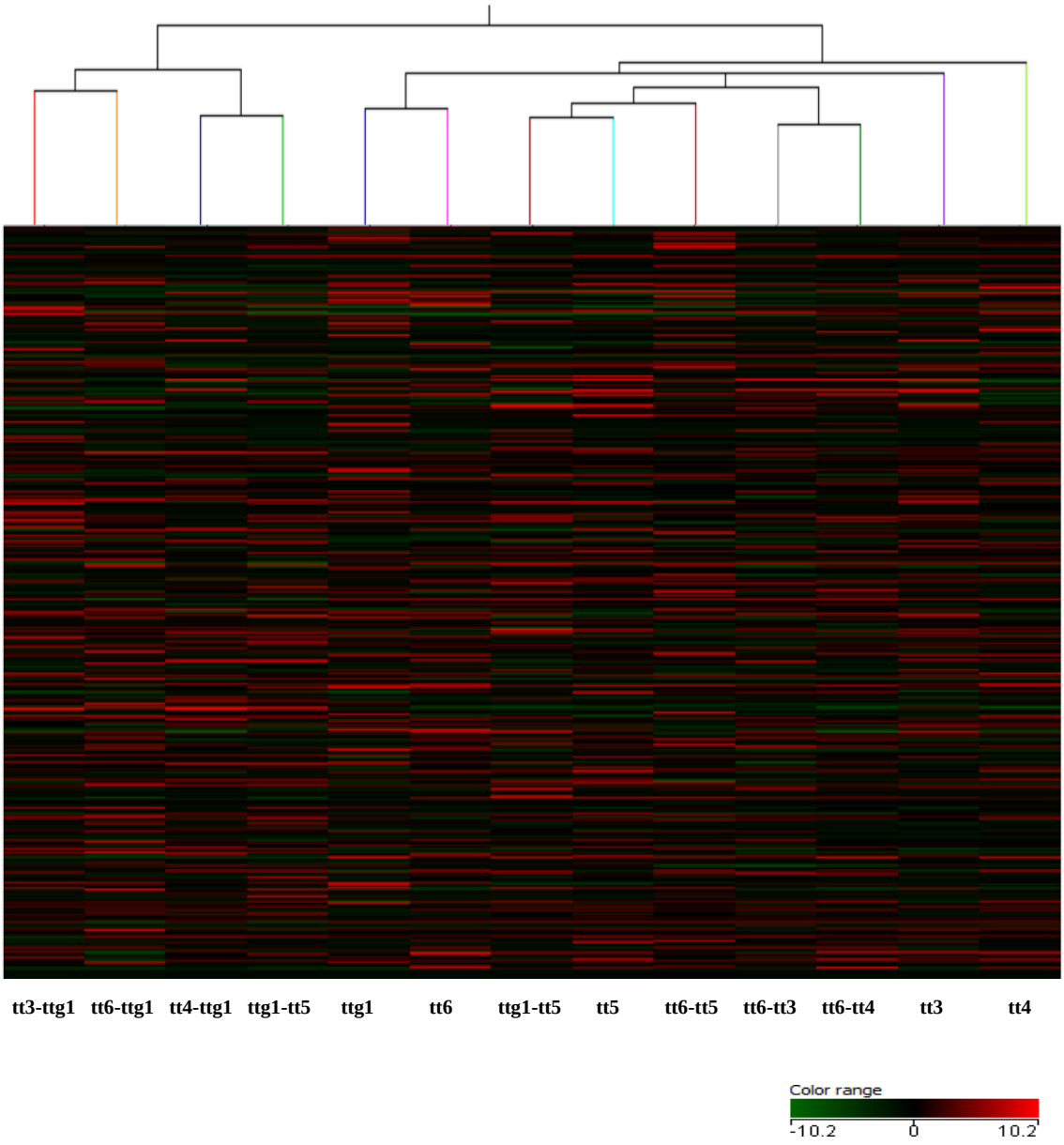


Fig 4.14 Hierarchical clustering of metabolic profile of tt single and double mutants. Hierarchical clustering of differential metabolites w.r.t to Wt was performed based on euclidean distance matrix using ward's linkage criterion. Normalized intensity values of metabolites were used for clustering. Metabolic profiling of 6-day-old seedlings of double mutants were characteristic of either *ttg1* or *tt6* except the case of *tt5*, were metabolic profile of *tt5*, *tt5-ttg1* and *tt5-tt6* were clustered together.

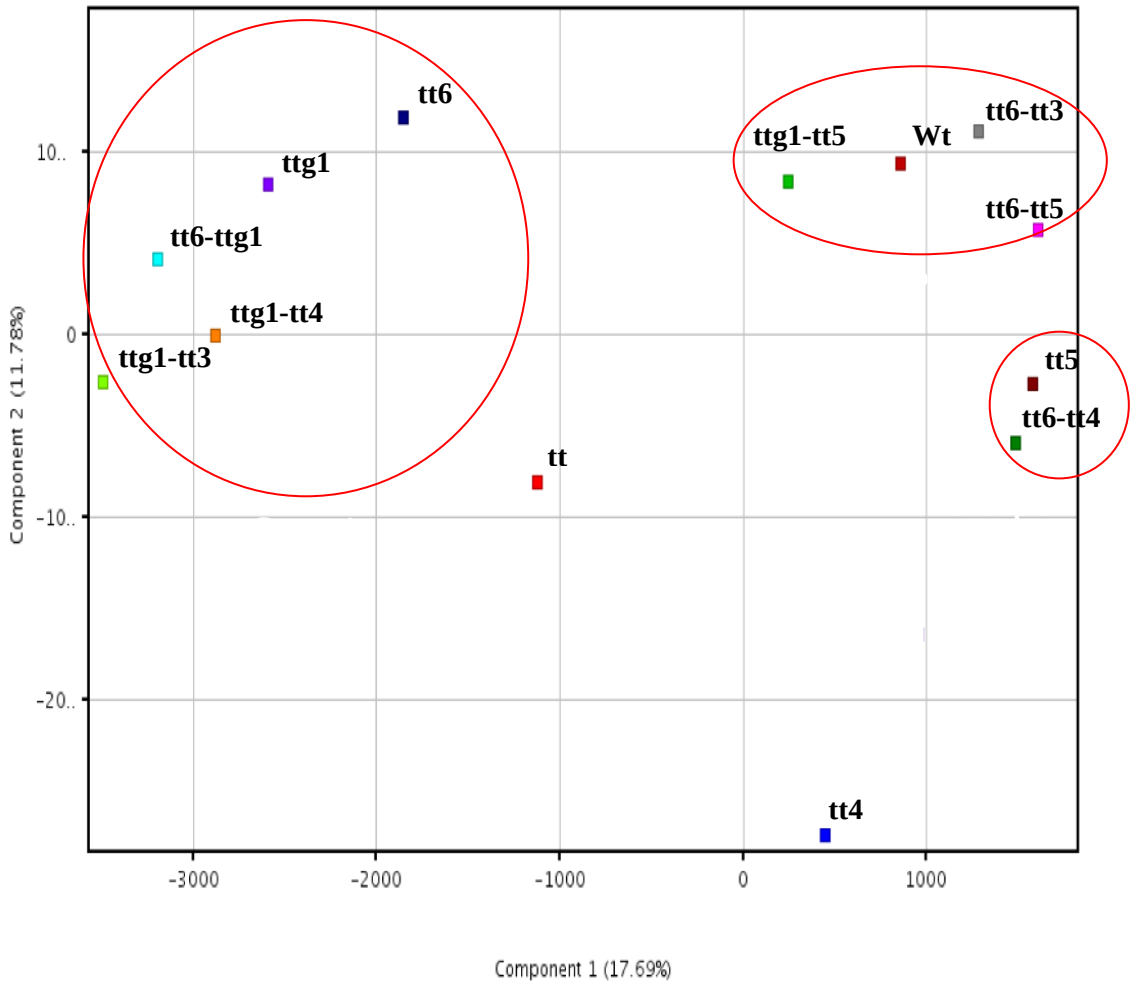


Fig 4.15 2D-PCA plot of metabolic profiling of single and double mutant line. Metabolites from 6 day old seedlings were extracted from single mutant lines *tt3*, *tt4*, *tt5*, *tt6*, *ttg1* and double mutant lines *tt3-tt6*, *tt3-ttg1*, *tt4-tt6*, *tt4-ttg1*, *tt5-tt6*, *tt5-ttg1* and *tt6-ttg1*. Three groups emerged from PCA plot, group1 with *ttg1*, *tt6-ttg1*, *ttg1-tt4* and *ttg1-tt3*, group2 with *ttg1-tt5*, Wt, *tt6-tt5* and *tt6-tt3*, and group3 being *tt6-tt4* and *tt5* while single mutants *tt3*, *tt4* and *tt6* were far apart from these groups with tt6 being closer to group1. It was expected for *tt4* to be separated from these groups since lower phenylpropanoid metabolic pathway is completely blocked in it. Group1 is characterized by *ttg1* single mutant and its double mutants with *tt3*, *tt4* and *tt6*. Group 2 and 3 were mainly *tt6* double mutant with *tt3*, *tt4* and *tt5*. PCA plot thus showed that metabolic profile of double mutant were similar to those characteristic of either *ttg1* or *tt6*. Double mutants of *ttg1* were closer in the group compared to *tt6*.

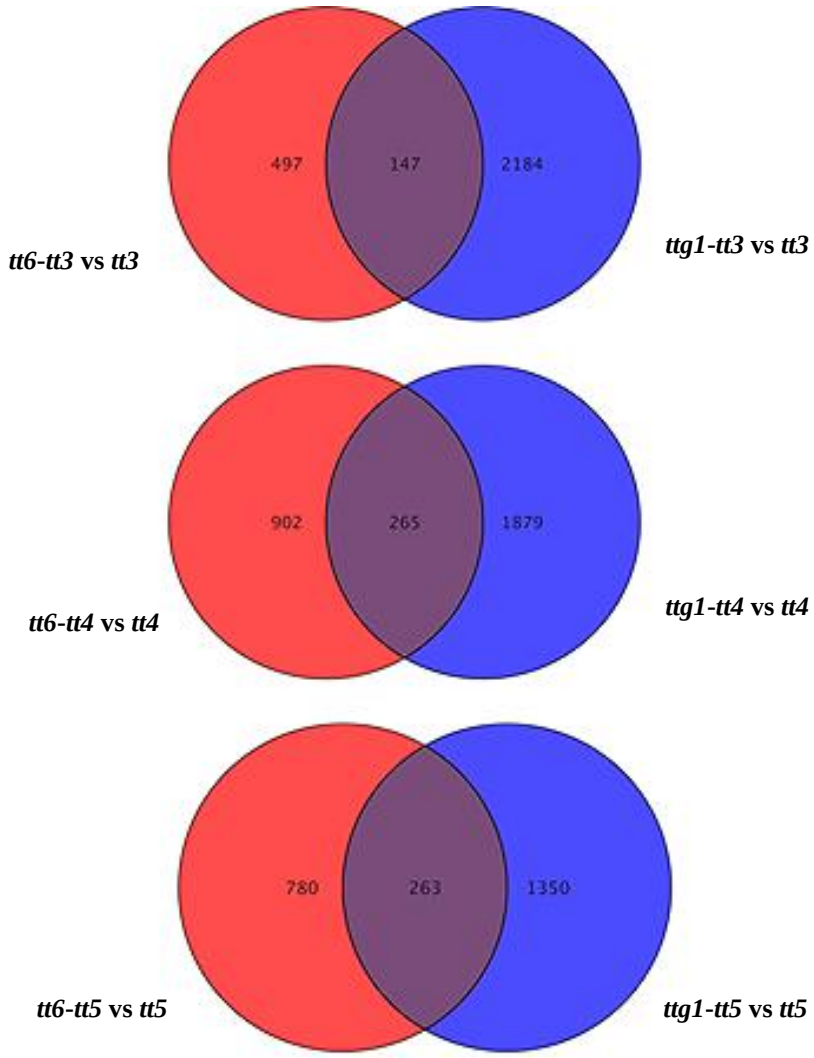


Fig 4.16 Comparison of differential metabolites between single and double mutant lines. Differential metabolites for single and double mutants filtered by p value <0.05 and fold-change >±2 w.r.t Wt were compared to identify effect of *tt6* or *ttg1* mutation on metabolic profile of single mutants *tt3*, *tt4* and *tt5*. As evident from Venn diagram, double mutants with *ttg1* showed more differential metabolites compared with corresponding single mutants.

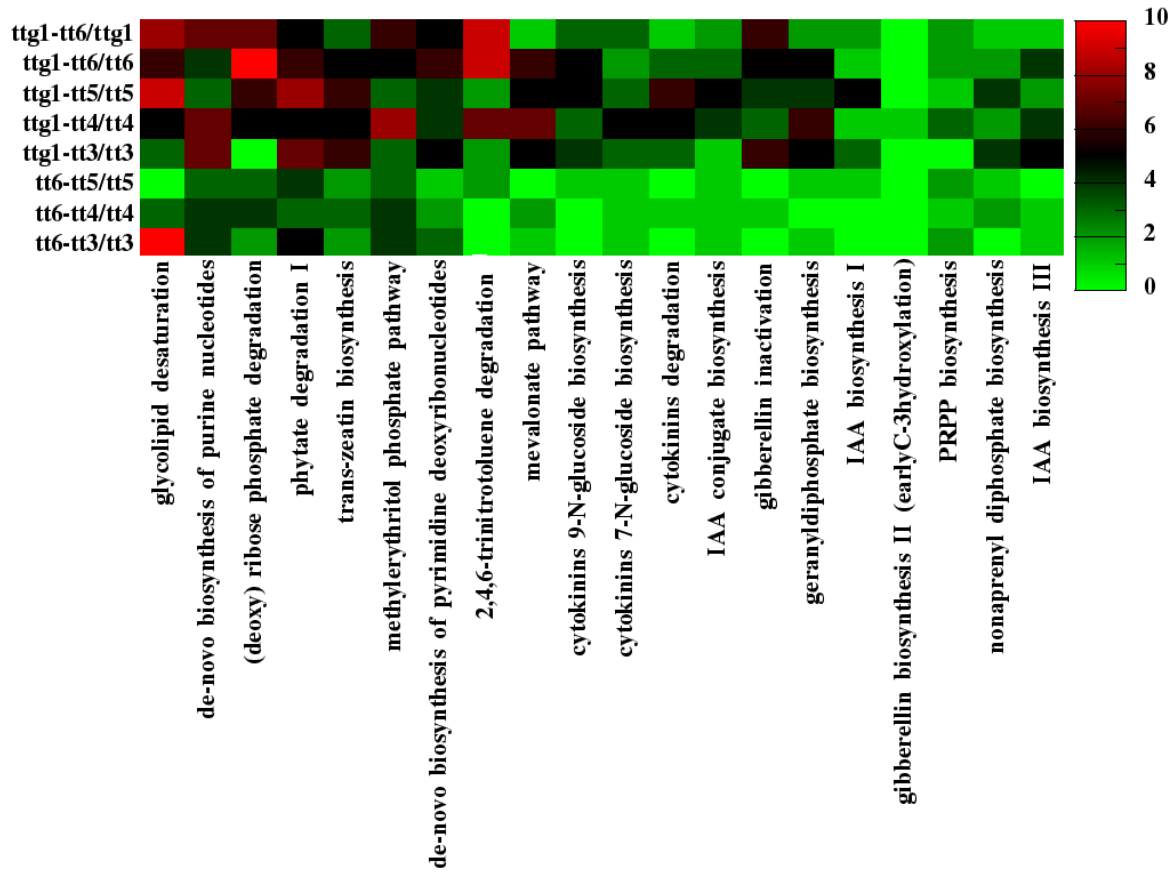


Fig 4.17 Heat map representing number of differential metabolites from a particular metabolic pathway. Differential metabolites for single and double mutants filtered by p value <0.05 and fold-change >±2. Differential metabolites from *ttg1* and *tt6* double mutants w.r.t corresponding single mutants *tt3*, *tt4*, *tt5* and *tt6* were identified by MS2 fragmentation and matching fragmentation pattern with mass-bank database. Identified metabolites were then mapped to metabolic pathways, and number of metabolites differentially expressed from a particular metabolic pathways were visualized here as heat-map.

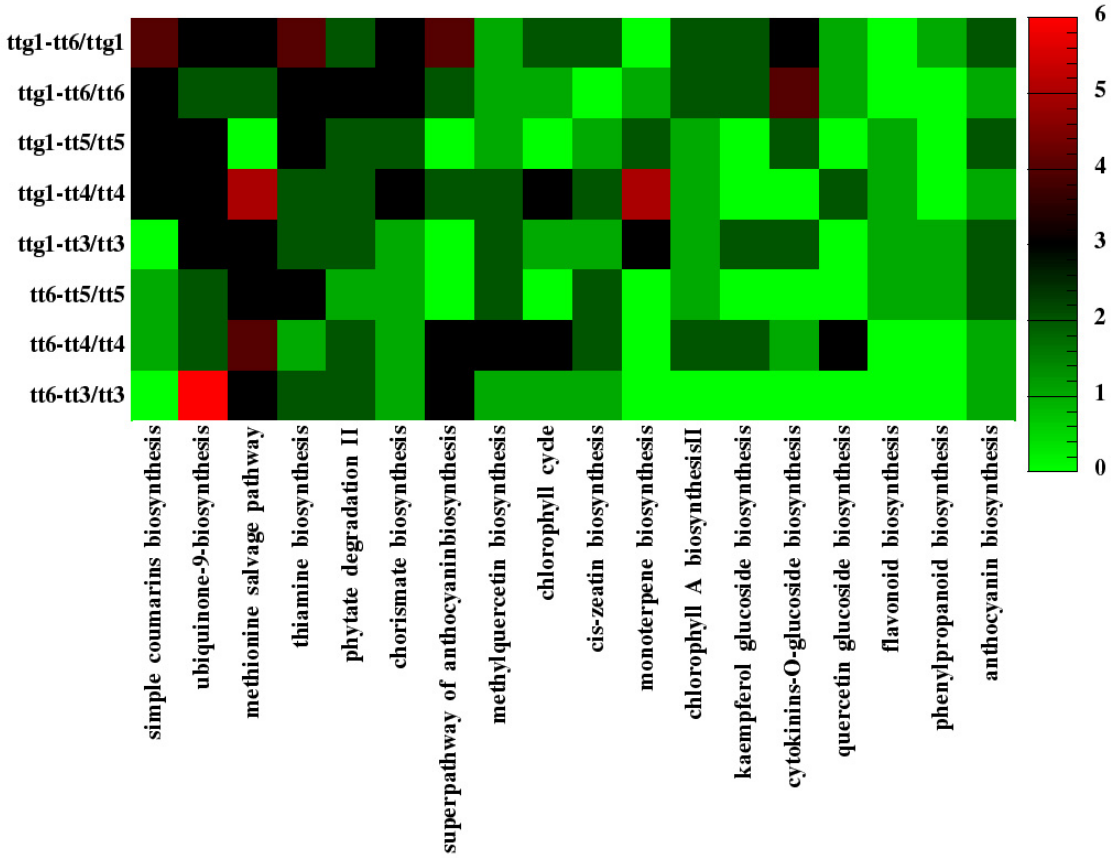


Fig 4.18 Heat map representing number of differential metabolites from a particular metabolic pathway. Differential metabolites for single and double mutants filtered by p value <0.05 and fold-change >±2. Differential metabolites from *ttg1* and *tt6* double mutants w.r.t corresponding single mutants *tt3*, *tt4*, *tt5* and *tt6* were identified by MS2 fragmentation and matching fragmentation pattern with mass-bank database. Identified metabolites were then mapped to metabolic pathways, and number of metabolites differentially expressed from a particular metabolic pathways were visualized here as heat-map.

4.2.3 TT8 loss-of-function mutant lines had lower levels of glycosylated metabolites

Metabolic profiling of *tt* mutant lines revealed a unique metabolic property of TT8 loss-of-function mutant lines. Total spectra of Wt and *tt8* mutant line were superimposed which showed Wt spectra consisting of metabolite ions of all mass range (50-1200), while *tt8* mutant lines showed most of the metabolite ions from lower mass range (50-400) (Fig 4.19). This was interesting since none of the *tt* mutant lines other than tt8 showed such specific difference with Wt. One possibility of this behavior could be lowered conjugation processes in *tt8* mutant lines. Different conjugation processes such as glycosylation, methylation, acylation and others results in synthesis of derived compounds from a small set of parent metabolites which are of higher masses and brings large metabolic diversity. To identify if lowered conjugation process was the main reason for low mass range metabolites in *tt8* mutant lines and if that's the case, then what types of conjugation processes were affected in *tt8* mutant lines, differential metabolites from *tt8* mutant line w.r.t Wt were identified. Differential metabolites were fragmented by MS2 and fragmentation patterns were matched with standards or mass-bank. Total of 159 metabolites were identified which then were mapped to their associated metabolic pathways.

Metabolic pathways such as phenylpropanoid metabolic network, purine biosynthesis, pyrimidine biosynthesis, carotenoid biosynthesis, steroid biosynthesis pathways and nucleotide sugar metabolism and brassinosteroid biosynthesis pathways were affected in tt8 mutant lines (Fig 4.21). Phenylpropanoid metabolic pathway was significantly affected in tt8 mutant line, showing highest number of differential metabolites among all metabolic pathways affected. Interestingly, metabolites from phenylpropanoid

network, such quercetin, kaempferol, cyanidin, pelargonidine that undergo several types of sugar conjugation processes were affected in tt8 mutant lines, showing a specific glycosylation pattern for these aglycones. Glycosylated form of quercetin and cyanidin such as cyanidin-3,5-diglucoside, cyanidin-3-O-beta-D-glucoside, quercetin 3,7-O-diglucoside, quercetin 3-O-rhamnoside-7-O-glucoside and quercetin-3-O-glucoside-7-O-rhamnoside were down-regulated in *tt8* mutant lines while that of kaempferol and pelargonidin such as kaempferol-7-O-glucoside, kaempferol-3-glucoside and pelargonidin-3,5-diglucoside with exception of kaempferol 3,7-O-diglucoside were up-regulated in tt8 mutant lines (Fig 4.20). Rhamnoside conjugated quercetin such as quercetin3,7-rhamnoside and quercetin 3-O-rhamnoside were up-regulated in *tt8* mutant lines while that with kaempferol such as kaempferol-3-rhamnoside, kaempferol-3-O-gentiobioside-7-O-rhamnoside and kaempferol-3-rhamnoside-7-rhamnoside were down-regulated in tt8 mutant lines (Table 4.1). Quercetin and kaempferol have been shown to compete for glycosylation process and our results showed that *TT8* loss-of-function results in down-regulation of several of their derivatives with showing a specific pattern for glucosides and rhamnosides derivatives. Levels of aglycones such as quercetin and cyanidine were up-regulated in *tt8* mutant line, consistent with earlier report which also showed increased quercetin levels in tt8 mutant line (Pelletier et al., 1999). Using HPLC based approach, Pelletier et al. 1999 suggested *tt8* mutant to be impaired in glycosylation of quercetin, which was consistent with our study, and using high mass accuracy mass spectrometry, we identified several metabolites other than quercetin that were affected in glycosylation. Glycosylation of several nucleotides were also affected with different glycosylated forms of UDP such as UDP-arabinose, UDP-galactose, UDP-xylose, UDP-glucuronate and UDP-apiose were down-regulated in tt8 mutant lines (Fig 4.20).

When all these observations are taken together, we can conclude that *tt8* mutant line was affected in glycosylation where conjugation of sugar molecules to several metabolites were affected. Specificity in types of glycosylation affected suggested TT8 playing an important role in regulating glycosylation process. We created inducible over-expression and silencing transgenic lines for TT8 to identify direct targets of TT8 and to understand the mechanism through which TT8 regulates glycosylation. Results related to these transgenic lines are discussed in Section 4.3.

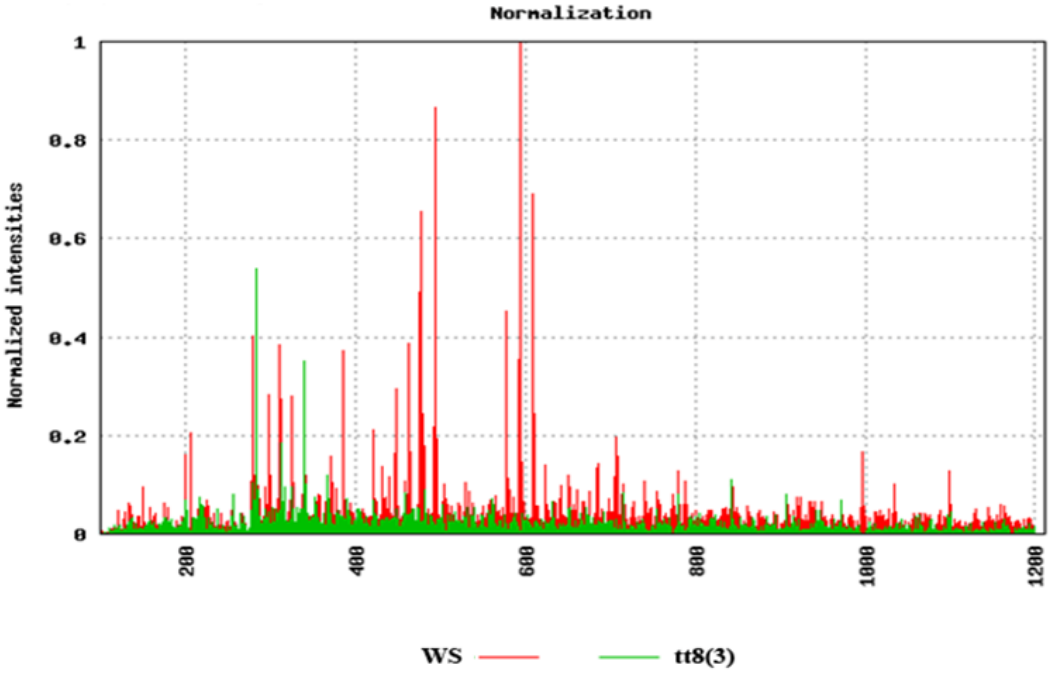


Fig 4.19 Metabolic profiles of WS (Wt background for *tt8(3)*) and *tt8(3)* mutant line over-lapped to highlight differences in the metabolites. Mass spectra of wild type showed all ranges of metabolite ions while tt8 mutant lines showed mainly lower mass range metabolites, a property specific to *tt8* mutant line. This suggested conjugation of metabolites affected in *TT8* loss-of-function transgenic line.

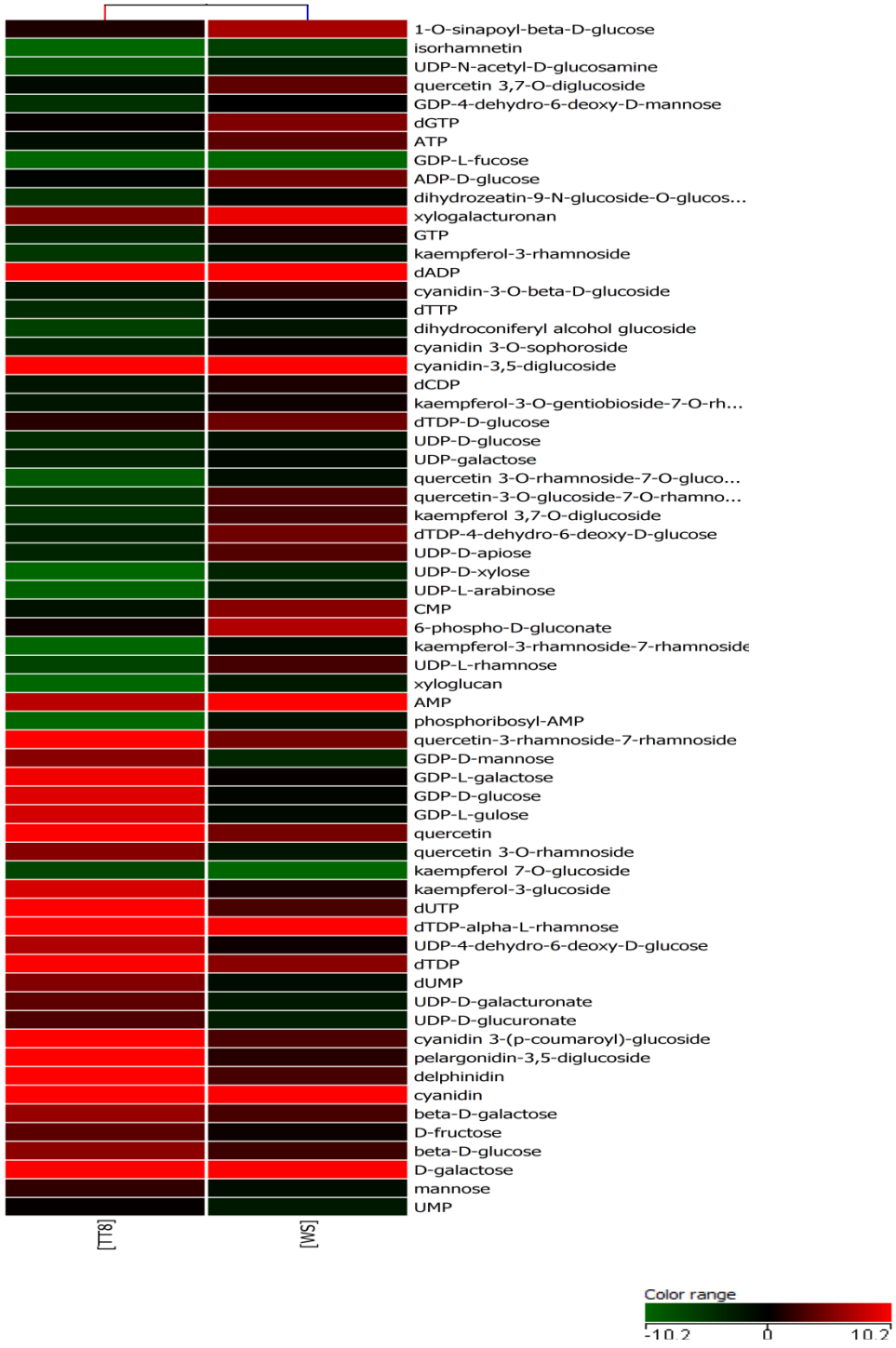


Fig 4.20 Heat map showing differential metabolites affected in glycosylation in *tt8* mutant line. Differential metabolites with p-value 0.05 or below and fold change ±2 w.r.t Ws were filtered. Only those metabolites that were associated with glycosylation have been shown in this heat map. All metabolites were confirmed by MS2 fragmentation and matching daughter ions with pure standards or mass-bank (http://www.massbank.jp/index.html?lang=en) database. Colour bar represents normalized intensity of metabolites. Complete list of differential metabolites is shown in Appendix III.

Table 4.1 Selected glycosylated metabolites differentially expressed in *tt8* mutant line.

Selected glycosylated metabolites in TT8 loss-of-function mutant lines w.r.t Wt	
Up-regulated	**Down-regulated**
kaempferol 7-O-glucoside	**kaempferol 3,7-O-diglucoside**
kaempferol-3-glucoside	**kaempferol-3-O-gentiobioside-7-O-rhamnoside**
kinetin-7-N-glucoside	**kaempferol-3-rhamnoside**
kinetin-9-N-glucoside	**kaempferol-3-rhamnoside-7-rhamnoside**
pelargonidin-3,5-diglucoside	**quercetin 3,7-O-diglucoside**
quercetin 3-O-rhamnoside	**quercetin 3-O-rhamnoside-7-O-glucoside**
quercetin-3-rhamnoside-7-rhamnoside	**quercetin-3-O-glucoside-7-O-rhamnoside**
	cyanidin-3,5-diglucoside
	cyanidin-3-O-beta-D-glucoside
	cyanidin 3-O-sophoroside

Glucoside derivative of kaempferol and pelargonidin were up-regulated while that of quercetin and cyanidin were down-regulated. On the other hand, rhamnoside derivative of kaempferol were down-regulated while that of quercetin were up-regulated. All metabolites shown here were confirmed by MS2 fragmentation and comparing with either pure standard or mass-bank database. All metabolites were with p-value 0.05 or below and fold change ±2 or above.\

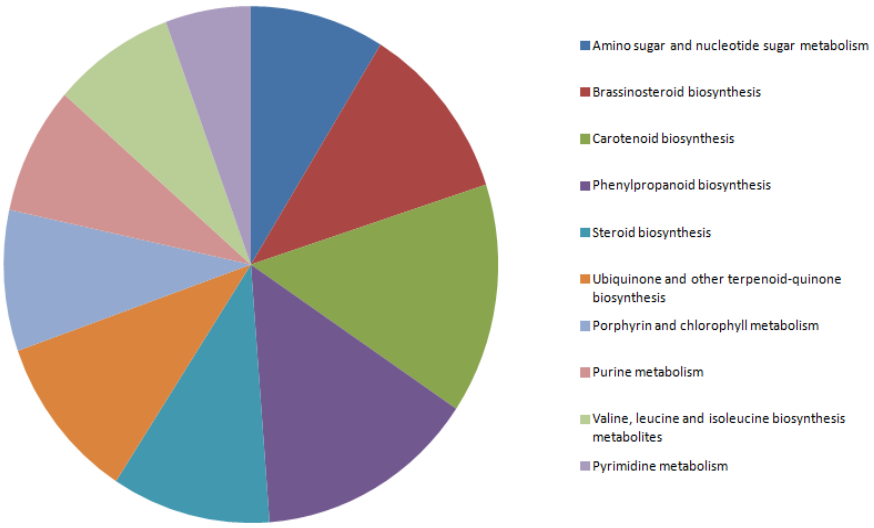


Fig 4.21 Metabolic pathways significantly affected in *TT8* loss-of-function mutant lines. Differential metabolites were identified and mapped on metabolic pathways. Phenylpropanoid metabolic pathways showed highest number of metabolites affected among other pathways.

4.2.4 Discussion

Previous study on characterization of QuoA transgenic lines showed several pathways to be altered by introduced perturbation at phenylpropanoid pathways (Reuben, 2008). These observed effects were at different branches of isoprenoid metabolic network, such as in carotenoid and brassinosteroid pathways. QuoA-mediated perturbation in transgenic lines at isoprenoid metabolic network was at both metabolites and gene expression levels. QuoA activity in transgenic lines resulting in isoprenoid metabolic network can be explained in two possible ways. First, there could be a regulatory system, connecting both phenylpropanoid and isoprenoid metabolic network. Second, QuoA may also act on metabolites of isoprenoid metabolic network resulting in its perturbation along with phenylpropanoid network. Second possibility is unlikely, as QuoA transgenic lines showed specific perturbational pattern, with increase in QuoA activity corresponding to increase in metabolites accumulation and phenylpropanoid pathways associated phenotypes of coloration and stiffness. This suggests QuoA specificity with its substrate quercetin.

To identify regulatory network that might link phenylpropanoid with isoprenoid metabolic network, we analyzed metabolic profile of ten mutant lines from lower phenylpropanoid metabolic network. Comparison of metabolic profile of *tt* mutant lines against *Q10* identified *tt6* and *ttg1* as two mutant lines, which showed similar crosstalk between phenylpropanoid and isoprenoid metabolic network (Fig 4.11). Perturbation in phenylpropanoid metabolic network due to loss of F3H and due to QuoA activity results in accumulation of same metabolite naringenin which is restricted to upper phenylpropanoid metabolic network in case of *tt6* while in *Q10,* it may be channelled to both upper and lower pathways. Thus, *tt6* mutant line showing

similar metabolic perturbational pattern corroborated with the results of immediate effects of QuoA-mediated perturbation in *Q10* transgenic lines. On the other hand, ttg1 showing similar metabolic perturbational behavior as that of *tt6* and *Q10* was a novel finding, since the point of perturbation was different than that of *tt6* and *Q10*.

Hierarchical clustering based on metabolic profiling of *tt* mutant lines showed formation of two groups, one of mutants of transcription factors and second of mutants of structural genes (Fig 4.10A). Only exception was *ttg1*, a transcriptional factor, which was grouped with structural genes and showed high similarity with tt6. Similar metabolic profiles of *ttg1*, *tt6* and *Q10* transgenic lines suggests disruption of similar regulatory mechanism that links phenylpropanoid and isoprenoid metabolic networks. RNA levels of *TTG1* and *F3H* were down-regulated in *tt6*, *ttg1* and *Q10* transgenic lines in 6-day-old seedlings (Fig 4.13).

To study if mutation of *ttg1* or metabolic perturbation due to *tt6* loss of function is the reason of the observed crosstalk, we performed metabolic profiling of double mutant lines for *ttg1*/*tt6* with *tt3, tt4* and *tt5*. Loss-of-function mutant lines *tt3, tt4* and *tt5* did not show significant perturbation in isoprenoid metabolic network. We intended to test if introducing mutation in *TTG1* or *F3H* in these mutant lines will result in perturbation in isoprenoid metabolic network. Metabolic profiling of 6-day-old seedlings of double mutant lines *ttg1*/*tt3, ttg1*/*tt4, ttg1*/*tt5, ttg1*/*tt6, tt6*/*tt3, tt6*/*tt4* and *tt6*/*tt5* were performed and compared with single mutants of *tt3, tt4* and *tt5* to differential metabolites due to introduced TTG1 or F3H mutation. Hierarchical clustering of metabolic profile of double mutants showed *ttg1* and all double mutants of *ttg1* with *tt3, tt4* and *tt6* were grouped together except *ttg1-tt5* (Fig 4.14). PCA plot

reflected the same trend (Fig 4.15), suggesting TTG1 as the dominating factor in deciding final metabolic state of transgenic lines. On the other hand, *tt6* mutant was grouped with *ttg1* mutant while its double mutants showed formation of two groups. TT4 and TT5 act up-stream while TT3 acts down-stream to F3H in phenylpropanoid pathways (Fig 4.1). Double mutants of *tt6* with *tt3*, *tt4* and *tt5* showed small number of differential metabolites, suggesting metabolic channelling to be less important for creating metabolic perturbation (Fig 4.16). While double mutation of *tt6* with *tt3*, *tt4* and *tt5* did not show much difference with single mutants, ttg1 double mutants showed significant metabolic changes compared to single mutant. This was reflected by the fact that *ttg1*/*tt3*, *ttg1*/*tt4*, *ttg1*/*tt6* and *ttg1* were grouped together and also showed significantly high number of differential metabolites. TTG1, being a transcription factor thus seems to shape the metabolic state of *tt3*, *tt4*, *tt5* and *tt6*. This was an interesting observation since while tt4 mutants are devoid of any flavonoids, *tt3* mutant blocks pathways which lead to synthesis of anthocyanins and proanthocyanins. Although, *tt3* and *tt4* mutant lines showed highly different metabolic profile with no significant effect on isoprenoid metabolic network, double mutant lines showed similar metabolic state. This suggests TTG1 to have more targets than what is known so far.

Differential metabolites from single and double mutants were mapped to metabolic pathways. Results showed *ttg1* and its double mutants to be perturbed in isoprenoid metabolic network with mevalonate pathway, methylerythritol phosphate pathway, PRPP biosynthesis, IAA biosynthesis, nonaprenyl diphosphate biosynthesis, geranyl diphosphate biosynthesis being significantly perturbed in *ttg1*, *ttg1*/*tt3*, *ttg1*/*tt4*, *ttg1*/*tt6* and *ttg1*/*tt5* (Fig 4.17 and 4.18). On the other hand, *tt6* and its double mutants

showed less effect on isoprenoid metabolic network. Effects on several other pathways such as anthocyanin biosynthesis, simple coumarins biosynthesis, quercetin and kaempferol glucoside biosynthesis were of different extent in all of its double mutants, thus not showing any specific pattern.

Till date, there have been no reports suggesting interaction between phenylpropanoid and isoprenoid metabolic network. Characterization of sixteen independent QuoA transgenic lines showed perturbation of isoprenoid metabolic network while perturbing phenylpropanoid metabolic network (Reuben, 2008). Our results showed similar cross-talk in TTG1 and F3H loss-of-function mutant lines, suggesting a common link between these three transgenic lines showing metabolic network cross-talk. Comparison of metabolic profiling of *QuoA* transgenic lines, *tt* mutant lines and double mutant lines thus conclusively showed *ttg1* as a possible link between phenylpropanoid and isoprenoid metabolic network.

TTG1 and TT8 have been shown to form ternary complex with TT2 to regulate expression levels of BAN and DFR (Baudry et al., 2004), two biosynthetic genes from phenylpropanoid metabolic network. Metabolic profiling of TT8 loss-of-function mutant line showed unique metabolic state feature compared to that of the other transparent testa mutant. The glycosylation process was impaired in TT8 loss-of-function lines with down-regulation of several glycosylated metabolites. This effect was specific for glycosylation of quercetin, cyaniding, pelarogonidin and kaempferol and exhibited a unique pattern. While quercetin and cyanidin glucosides were down-regulated in *tt8* mutant line, the kaempferol and pelarogonidin glucosides were up-regulated. On the other hand, the rhamnoside conjugates of quercetin and cyaniding

were up-regulated in *tt8* mutant line while down-regulated in kaempferol. Earlier, HPLC based profiling of *tt* mutants showed increased accumulation of quercetin in *tt8* mutant with speculation of impaired glycosylation process (Peltier et al., 1999). Previously, metabolic profiling of root exudates from *tt8* mutant line showed significantly down-regulation of glycosylated metabolites compared to Wt and other *tt* mutant lines (Narasimhan, 2003). Several sugar conjugated forms of UDP and GDP metabolites were affected in tt8 mutant line with glycosylated UDP were don-regulated.

Nucleotides serve as acceptor as well as donor of sugar molecules, thus facilitating process of glycosylation and its regulation (Butler, 2006; Vaistij et al., 2009). This, along with transferases, hydrolases and sugar transporters regulates process of glycosylation. In our study, several of the metabolites having affected glycosylation process were identified. Metabolic profiling data showed tt8 mutant to be defective in glycosylation. Since glycosylation involves a complex regulatory mechanism and participates in the toxification and detoxification of several secondary metabolites, understanding its process is of great importance. Therefore, regulation of glycosylation process was further studied using tt8 as a model system and is described in Section 4.3.

4.3 Identification of metabolic pathways and genes regulated by TTG1 and TT8

Our study on metabolic profiling of *tt* mutant lines showed *ttg1* and *tt6* mutants as biochemical phenocopy of *QuoA* transgenic lines, showing perturbation to isoprenoid metabolic network in response to perturbation in phenylpropanoid metabolic network. To test whether *ttg1* or *F3H* loss-of-function could be a reason for this perturbational behaviour, we tested double mutants of *TTG1* and *F3H* with BAN (*tt3*), CHS (*tt4*) and CHI (*tt5*). Metabolic profiling of single and double mutants of *ttg1* showed metabolic perturbation in isoprenoid metabolic network, which otherwise was unaffected in the single mutants *tt3*, *tt4* and *tt5* transgenic lines. These results suggested a possible role of TTG1 in regulation of members from isoprenoid metabolic network, thus explaining cross-talk between phenylpropanoid to isoprenoid metabolic.

To further confirm if TTG1 directly regulates isoprenoid pathways and to identify direct targets of TTG1, we created transgenic lines expressing 2x35S:TTG1:GR (called as TTG1:GR). Since TT8 has been shown to form a complex with TTG1 to regulate its target, we also created inducible over-expression lines of TT8 (called as TT8:GR) to identify those targets that are common to both TTG1 and TT8 (See experimental workflow Fig 4.22). Metabolic profiling and gene expression profiling of *ttg1*, *tt8* and *tt2* loss-of-function mutant lines were performed to identify genes differentially expressed in all three mutant lines. These were than compared with inducible over-expression transgenic lines to identify metabolic pathways and processes that were directly regulated by TTG1 or TTG1-TT8 complex. Results were validated using RT-PCR approach and direct targets were tested against in-vivo binding of TTG1 by ChIP experiment.

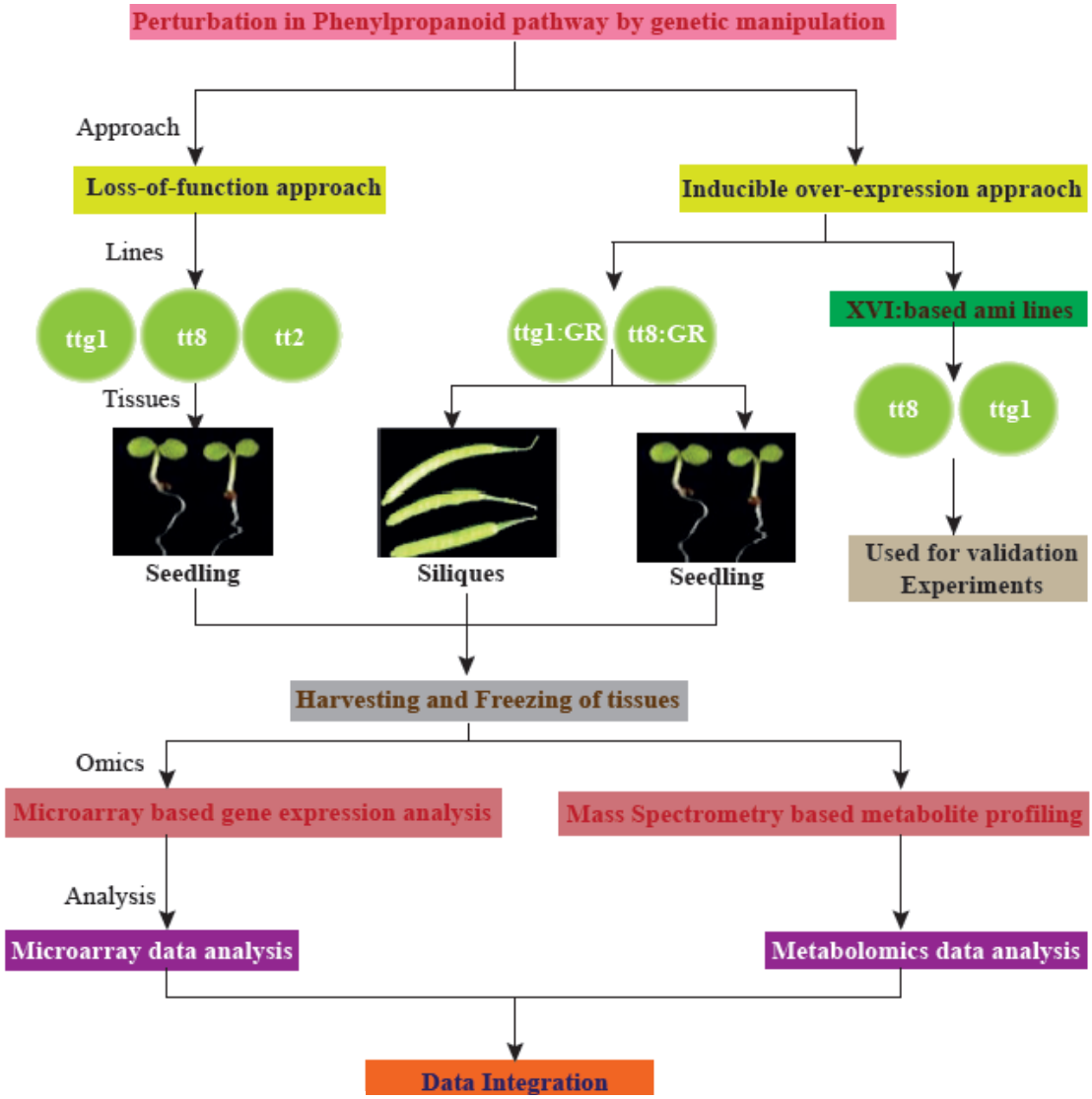


Fig 4.22 Experimental work flow for functional characterization of TTG1 and TT8 using inducible over-expression gain-of-function and loss-of-function mutant lines. Dex inducible TT8:GR and TTG1:GR were treated with dex or mock and samples were harvested to perform microarray and metabolic profiling based studies. List of genes showing reciprocal expression profile in loss-of-function and inducible transgenic lines were used to further analyze if they are direct targets of TTG1 and TT8. TTG1:GR and TT8:GR transgenic lines were then treated with mock or dex and/or chx to induce activity of TT8 and TTG1 in respective transgenic lines to identify direct targets of these transcriptional factors, which were then validated using ChIP assays.

4.3.1 TTG1 and TT8 fusion with GR is functional and inducible

Dex-inducible transgenic lines were created for inducible over-expression of TTG1 and TT8 by transforming Col-0 with pGreenII vector with insert 2x35S-TTG1:GR and 2x35S-TT8:GR. To test if fusion of TT8 and TTG1 with GR were functional, positively selected transgenic lines for the insert were tested for associated phenotypes for TT8 and TTG1. TTG1 has been associated with trichome initiation and regulation. It is also associated with proanthocyanin formation which gets accumulated in seeds and oxidation of which gives rise to brownish coloured tannin, giving characteristic brown coloration of seeds. *ttg1* loss-of-function transgenic lines shows phenotypes such as no trichome formation on leaf surface or stem base, yellow seeds due to absence of brown pigment in seed coat (testa) and no anthocyanins in leaves, stems and all tissues. TT8 loss-of-function results in yellow seed coloration due to absence of brown pigmentations, although it had no effect on trichome number or accumulation of anthocyanin.

Inducible over-expression transgenic lines were grown on MS agar medium for 5 days before transferring them to MS agar medium with dex or equivalent amount of its solvent and incubated for 24 hours before observing if increased activity of TTG1 or TT8 fussed with GR resulted in expected phenotype. Dex treatment of seedlings resulted in ectopic trichome formation on petiolate, while there was no trichome formation on seedlings treated with mock (Fig 4.23 and 4.24). Col-O also did not show any formation of trichome with or without dex treatment, suggesting trichome formation as a result of TTG1 or TT8 induced activity in respective transgenic lines. TTG1 has been described earlier as essential for trichome initiation, thus ectopic trichome formation was expected. However, TT8 has not been shown to be associated

with trichome formation. Presence of its homologs such as EGL3 and GL3 has been shown to be more important bHLH transcription factor than TT8 that regulates trichome formation (Zhang et al., 2003). Triple mutants of *egl3-gl1-tt8* showed *ttg1* phenotype of no trichome formation (Zhang et al., 2003; Zhao et al., 2007). Thus induced expression of TT8 leading to ectopic trichome formation at petiolates confirms that TT8 also participates in trichome formation. We then counted total number of trichome on first two true leaves of 6-day-old seedlings with or without DEX treatment. As expected, TTG1:GR and TT8:GR transgenic lines showed increased trichomes on adaxial side of leaves, while Col-0 and tt8 mutant did not showed any changes in trichome numbers. *ttg1* loss-of-function mutant did not show any trichome formation on the leaves (Fig 4.25). As expected, dex treatment to 6-day-old seedlings showed accumulation of anthocyanin in hypocotyl region, suggesting increased formation of anthocyanins in response to induced activity of TTG1 and TT8 (Fig 4.26).

Adult siliques of transgenic lines were then treated with dex or mock to observe any increase in seed coloration. As expected, both TTG1:GR and TT8:GR inducible lines showed increased brown pigmentation in seeds by dex treatment compared to mock or wild type seeds. There was no increase in seed coloration of Col-0 by dex or mock treatment, suggesting increased tannin formation in dex induced TTG1:GR or TT8:GR transgenic lines were due to TTG1 and TT8 activity (Fig 4.27 and 4.28). Dex treatment of 6-day-old inducible transgenic lines for TTG1 and TT8 also showed up-regulation of BAN and DFR, two known direct targets of TTG1 and TT8 (Fig 4.29). These results thus establish that TTG1 and TT8 fusion with GR was fully functional.

Fig 4.23 Ectopic trichome formation by dex treatment on Col-0 transformed with 2x35S:TTG1-GR or 2x35S:TT8-GR construct. (A, B) TT8:GR transgenic lines, (C, D) TTG1:GR transgenic lines, (E, F) Wt. All images were taken at 80x magnification under a dissection microscope with bright field illumination.

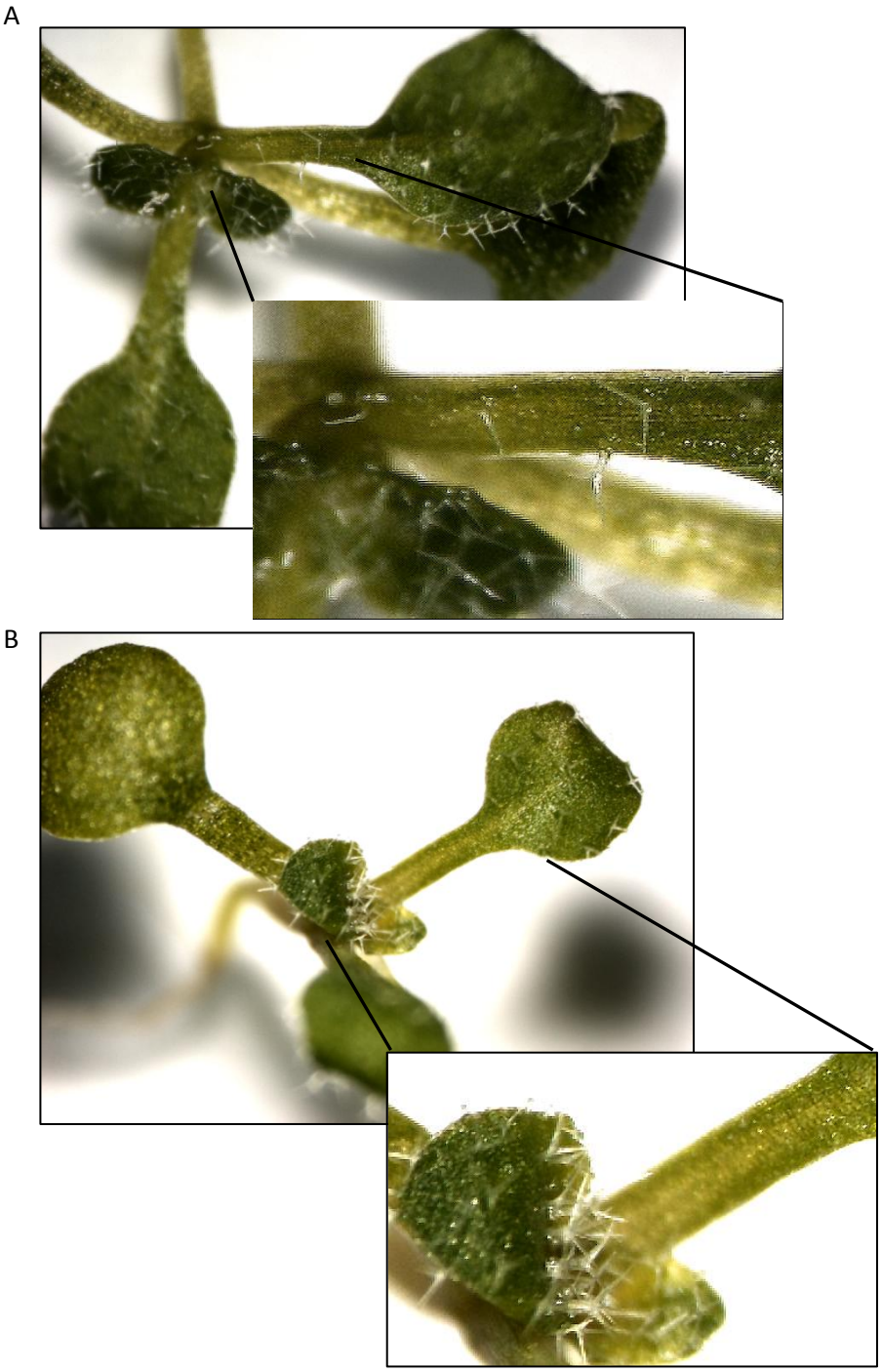


Fig 4.24 Formation of ectopic trichome in response to dex treatment in 2x35S:TTG1-GR or 2x35S:TT8-GR transgenic lines to the tissues which do not produce trichome. All images were taken at 80x magnification under a dissection microscope with bright field illumination.

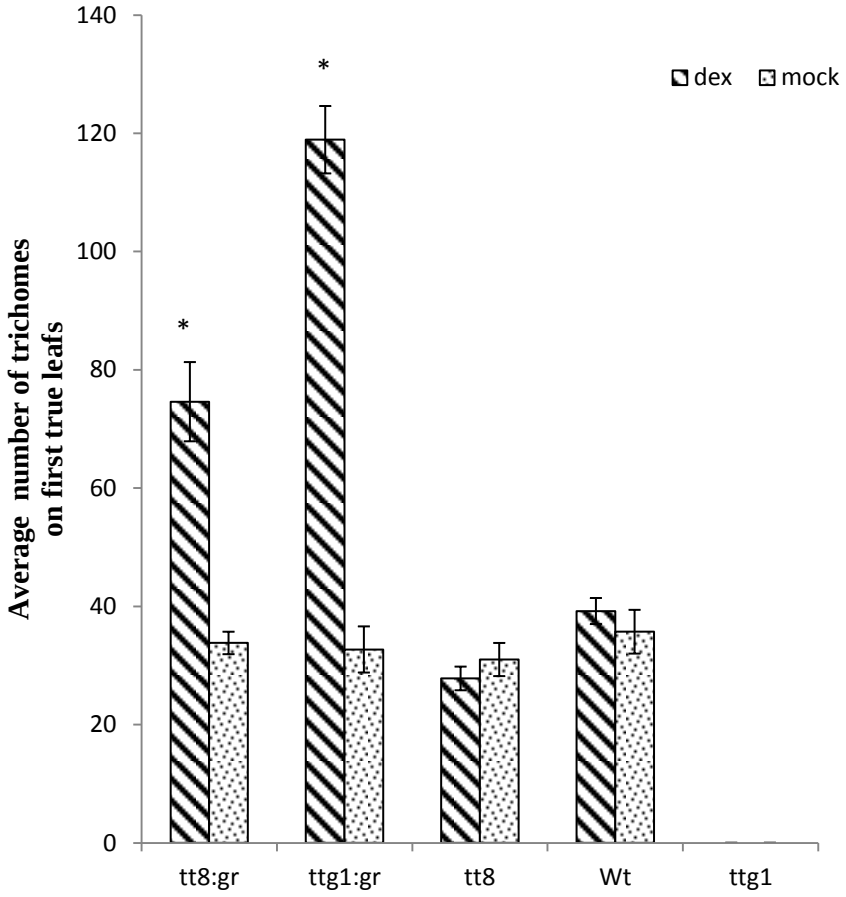


Fig 4.25 dex treatment of TT8:GR and TTG1:GR transgenic lines showed increase in total adaxial trichomes on first two true leaves. Total numbers of trichomes were counted on adaxial side of first two true leaves of 7 day old seedlings after treating them with dex or mock for 24 hours. Data are given as mean ±SD (n > 15). * $P < 0.01$ compared with the wild type.

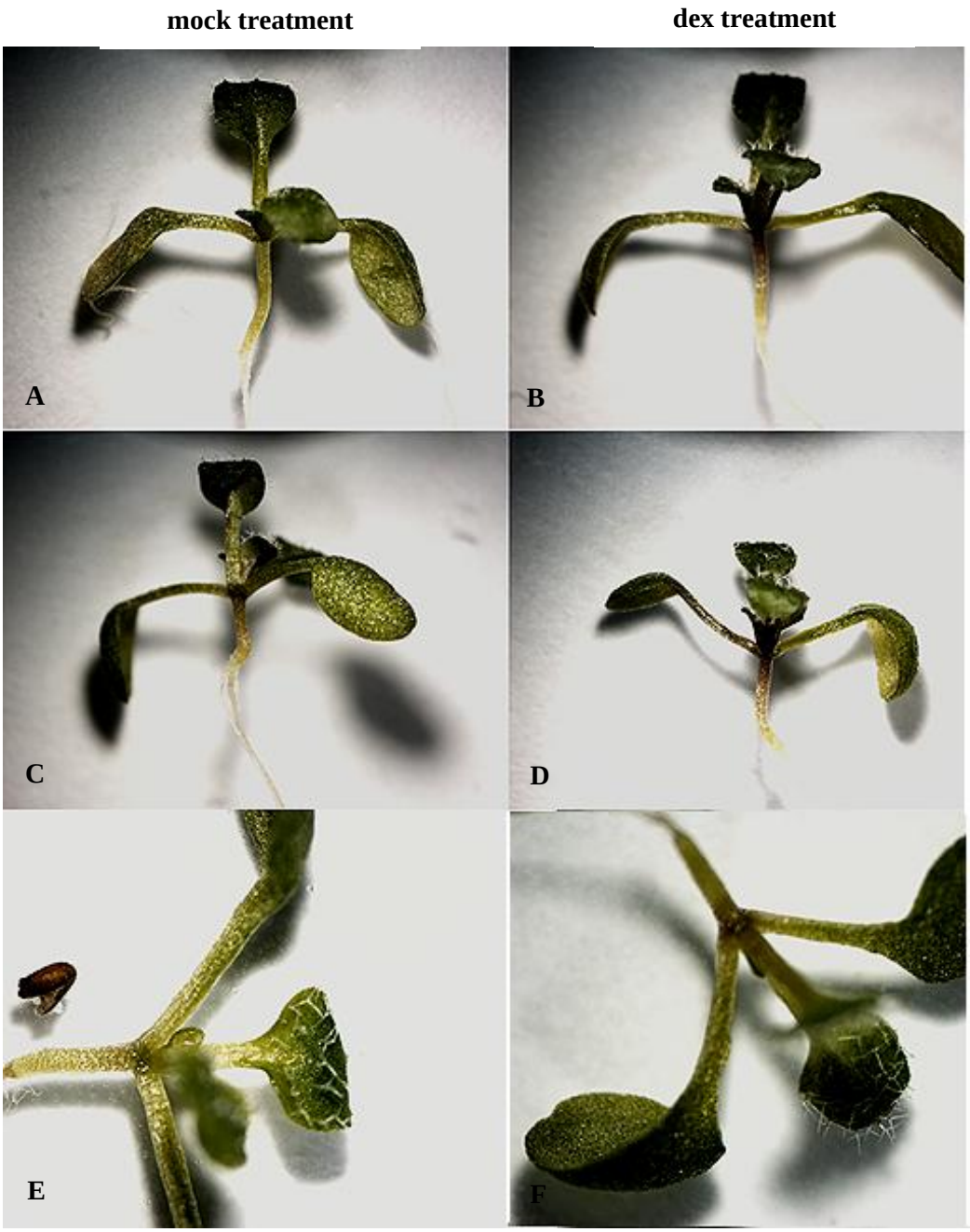


Fig 4.26 Anthocyanin accumulation in hypocotyl region of dex and mock treated transgenic lines. (A,B) TT8:GR transgenic lines, (C,D) TTG1:GR transgenic lines, (E,F) Wt. All images were taken at 80x magnification under a dissection microscope with bright field illumination.

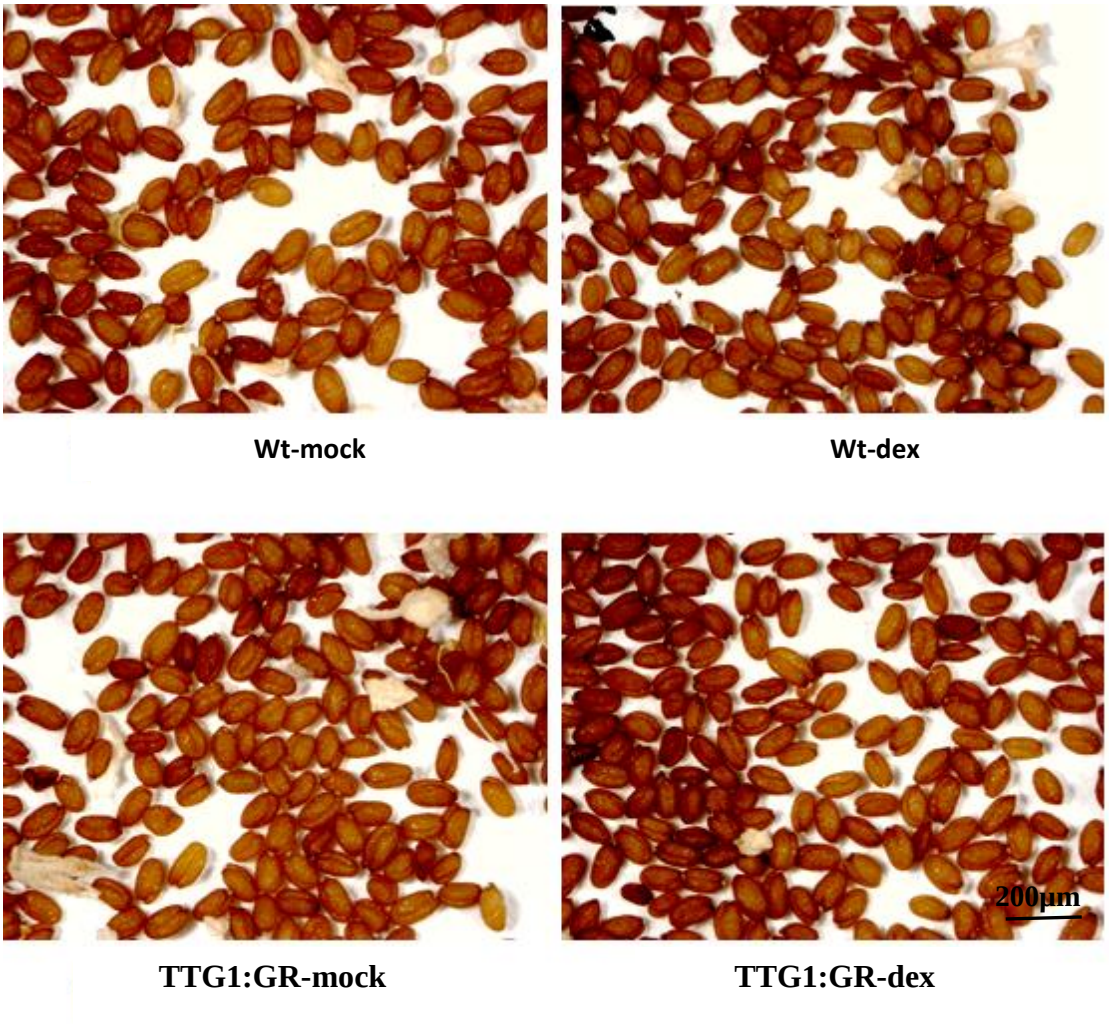


Fig 4.27 Increased seed coloration due to dex treatment to TTG1:GR inducible transgenic lines. Transgenic lines and Wild type arabidopsis lines were watered with 10µM final concentration of dex for 1 week before harvesting the seeds. Images were taken under a dissection microscope with bright field illumination. All images were taken at 80x magnification under a dissection microscope with bright field illumination.

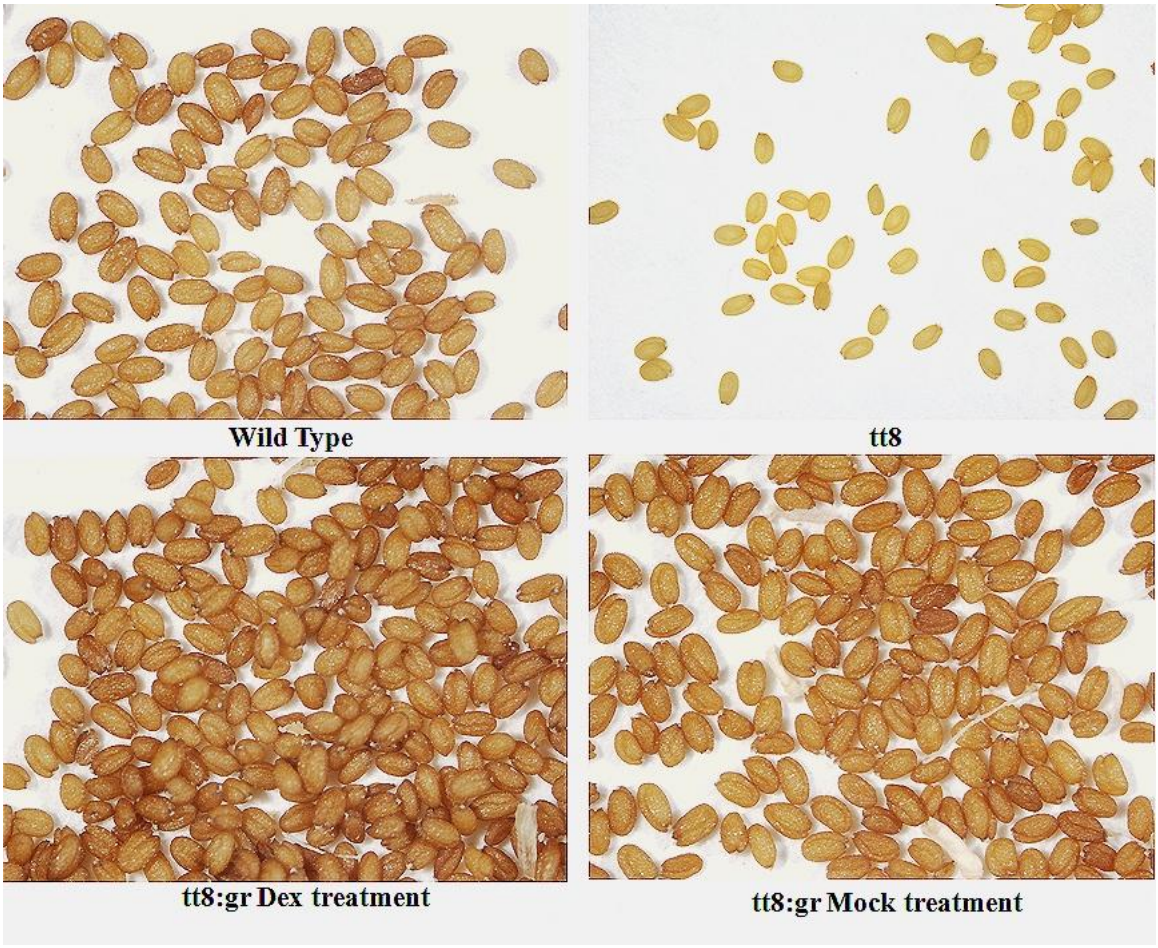


Fig 4.28 Increased seed coloration due to dex treatment to TT8:GR inducible transgenic lines. Transgenic lines were watered with 10µM final concentration of dex for 1 week before harvesting the seeds. Images were taken under a dissection microscope with bright field illumination. All images were taken at 80x magnification under a dissection microscope with bright field illumination.

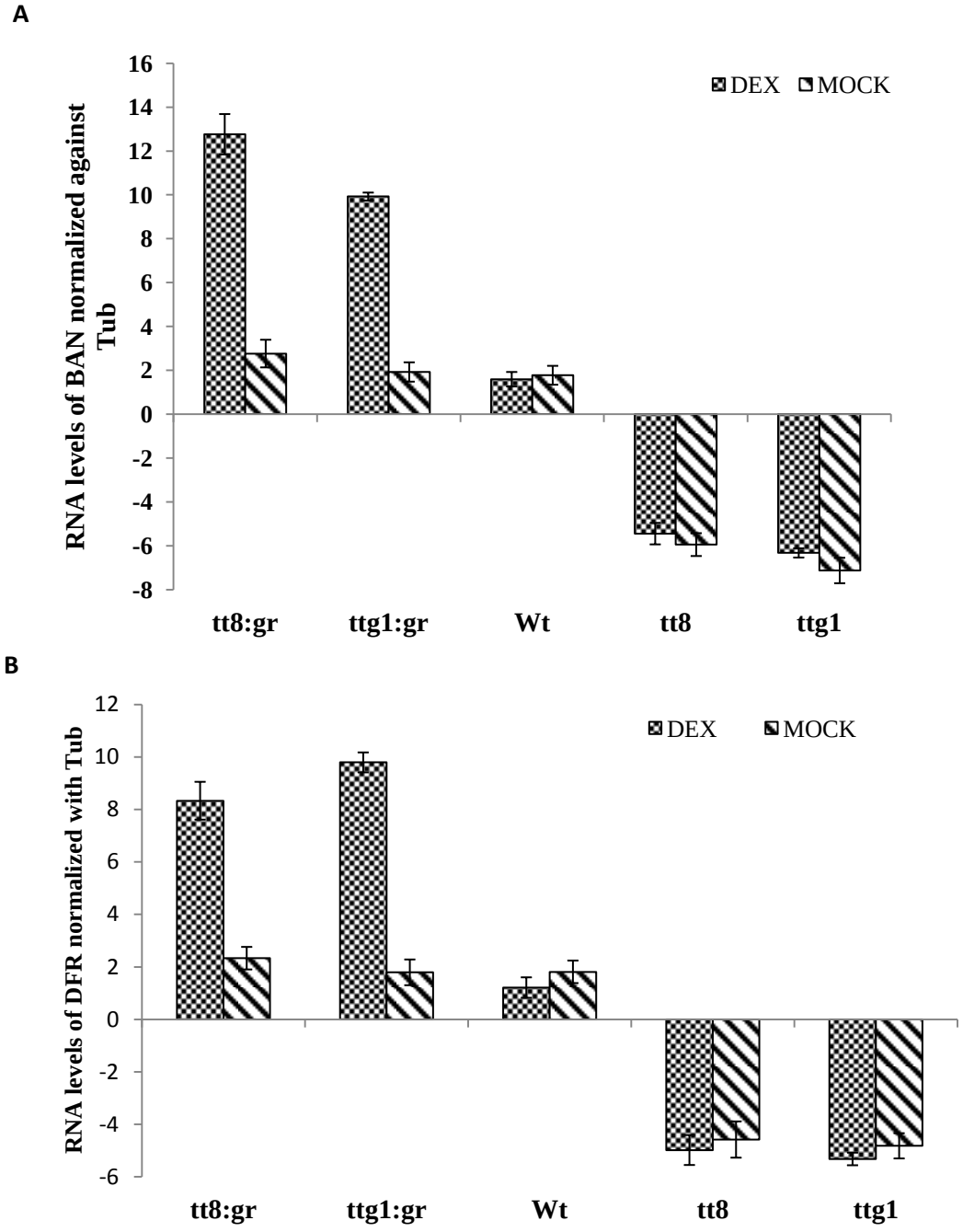


Fig 4.29 Quantitative Real time PCR showed increased RNA levels of BAN and DFR in dex-treated TTG1:GR and TT8:GR transgenic lines while were down-regulated in tt8 and ttg1 mutant lines. 6 day old seedlings were used for real-time PCR based quantification. BAN and DFR are two biosynthetic genes shown to be directly regulated by TT8 and TTG1 complex with TT2, thus this shows TTG1:GR and TT8:GR as functional transgenic lines. All fold changes values were of p-value 0.05 or bellow. Each transgenic line was used with 3 biological-replicates, with each replicate consisting with 45 seedlings. Dex treatment was performed and seedlings were harvested after 24 hours of treatment.

4.3.2 Non-targeted metabolic profiling of inducible over-expression transgenic lines showed up-regulation of metabolites from phenylpropanoid metabolic network

Metabolites extracted from 6-day-old seedlings of inducible over-expression transgenic lines after dex or mock treatment for 24 hours were subjected to untargeted metabolic profiling using UHPLC-TOF. Raw data were processed using Mass Hunter software from Agilent and ions were selected from raw data by Mass Extraction feature parameters set and described in Material and Methods Section (Chapter 3). Extracted metabolic data were subjected to alignment, mass correction, RT correction, noise removal, normalized, filtered based on abundance and scaled using Gene Spring 12.0 software from Agilent and were used for further multi-variant data analysis.

Metabolic profiles of transgenic lines with and without treatment were plotted using PCA which showed transgenic lines forming three groups (Fig 4.30). Mock treated inducible over-expression transgenic lines showed similar metabolic profile as that of Wt and were grouped together in PCA plot. Other two groups formed in PCA plot were loss-of function mutants and dex induced transgenic lines (Fig 4.30). PCA plot showed that metabolic profiling of transgenic lines was determined by loss or induced expression of *ttg1* or *tt8.*

Metabolites with p-value <0.05 and fold change $\geq \pm 2$ were filtered for further identification and mapping on to metabolic pathways. Differential metabolites were identified by matching fragmentation pattern of metabolite (MS2) with standards or MS2 data deposited in MassBank (http://www.massbank.jp/?lang=en). Since TTG1 and TT8 have been shown to physically interact to form a complex and co-regulate gene-expression of genes in phenylpropanoid metabolic network, we first compared metabolic profile of differential metabolites for TTG1:GR and TT8:GR transgenic

lines with and without dex treatment. Dex treatment resulted in 1234 differential metabolites in TTG1:GR transgenic lines and 763 differential metabolites in TT8:GR transgenic lines as compared to corresponding metabolic profile of mock treatment, with 533 metabolites common to both transgenic lines after dex treatment (Fig 4.31). Out of total 1464 differential metabolites together in TTG1:GR and TT8:GR, we were able to identify 368 metabolites based on MS2 fragmentation pattern and mapped to metabolic pathways, part of which is shown in Table 4.2 and complete list is in appendix.

Overall metabolic perturbation showed up-regulation of metabolites in response to dex-induced activity of TTG1 and TT8, with 167 metabolites up-regulated and 73 down-regulated in TTG1:GR and 185 metabolites up-regulated and 87 down-regulated in TT8:GR. Among identified metabolites, 194 metabolites showed similar trend of perturbation in TTG1:GR and TT8:GR, with 139 metabolites showed up-regulation and 55 metabolites were down-regulation in response to dex treatment. As expected, several metabolites from phenylpropanoid metabolic network, such as dihydrokaempferol, quercetin, 3,7-di-O-methylquercetin, cyanidin-3-O-beta-D-glucoside, pelargonidin-3-O-beta-D-glucoside, pelargonidin-3,5-diglucoside, caffeoylquinate, 4-coumarate were all up-regulated in TTG1:GR and TT8:GR transgenic lines by dex treatment (Table 4.2). Levels of metabolites were consistent with observed phenotype of anthocyanin accumulation in young seedlings of transgenic lines in response to dex treatment (Fig 4.26). Conjugation processes such as methylation, sulphates, glycosylation of metabolites from phenylpropanoid metabolic network were up-regulated in transgenic lines. Methylated and sulphate form of quercetin such 3,7,4'-tri-O-methylquercetin, 3,7-di-O-methylquercetin,

quercetin-3,7,3',4'-tetrasulphate, quercetin-3,3',7-trissulfate and quercetin-3,3'-bisulfate were all up-regulated in dex induced TTG1:GR and TT8:GR transgenic lines. Glycosylation of metabolites from phenylpropanoid metabolic network such as quercetin, cyanidin, sinapaldehyde, delphinin and pelargonidin were also up-regulated in dex induced TT8:GR and TTG1:GR. Interestingly, none of kaempferol conjugated forms were differentially expressed in either of TT8:GR and TTG1:GR transgenic lines. This is particularly important since quercetin and kaempferol competes for different conjugation processes and these results showed TT8 and TTG1 favouring conjugation of quercetin than of kaempferol. Glycosylation of metabolites other than those from phenylpropanoid metabolic network were also up-regulated particularly in TT8:GR lines. Many metabolic pathways other than phenylpropanoid metabolic network were also perturbed in response to dex treatment in TTG1:GR and TT8:GR. These pathways and perturbational pattern has been described in next section.

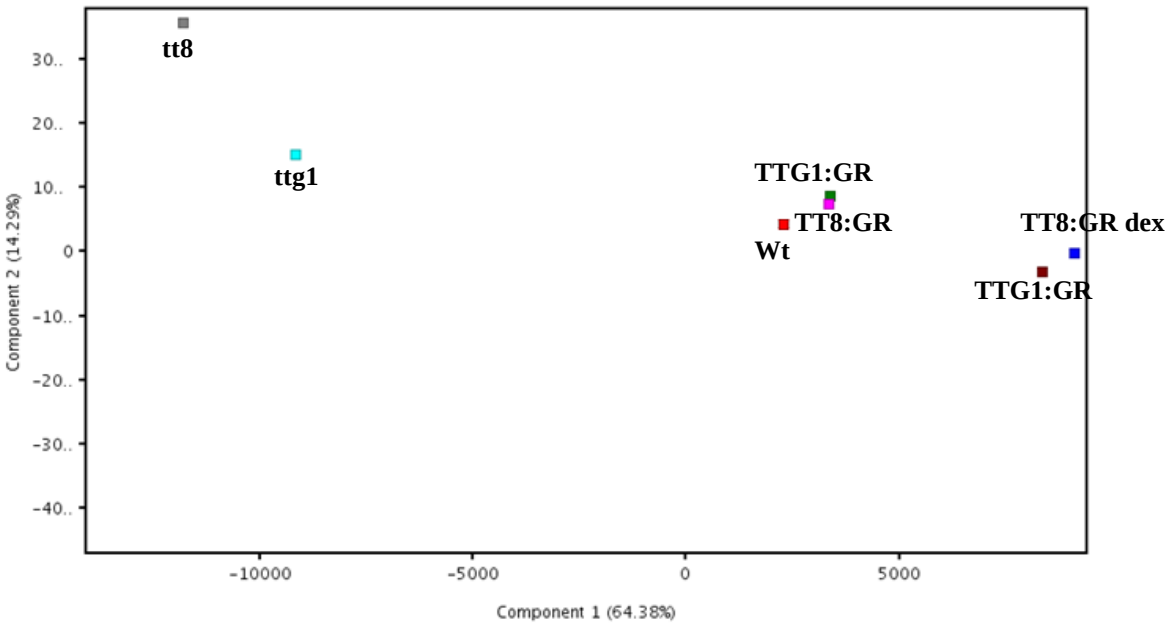


Fig 4.30 PCA plot on metabolic profile of TT8 and TTG1 loss-of-function mutant lines and inducible over-expression transgenic lines with dex or mock treatment. Metabolic profiling for each treatment was done for 4 biological replicates and each replicates were injected 3 times to record metabolic profile. PCA showed mock treated TTG1:GR, TT8:GR grouped together with Wt while mutants and dex-treated lines were grouped together.

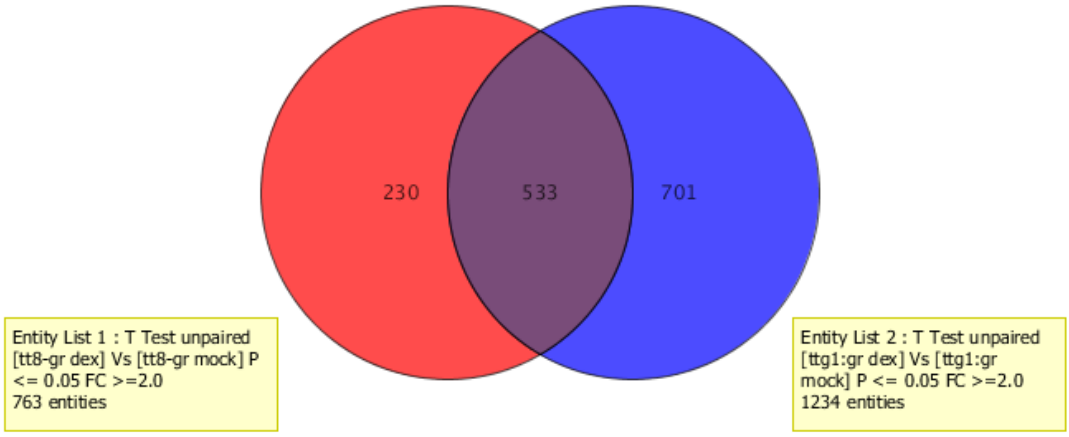


Fig 4.31 Venn Diagram showing distribution of differentially expressed metabolites in dex-treated TTG1:GR and TT8:GR transgenic lines.

Table 4.2 Differential metabolites from phenylpropanoid metabolic network in dex-treated TTG1:GR or TT8:GR w.r.t mock treated transgenic lines.

Metabolites	Dex-treated TTG1:GR	Dex-treated TT8:GR
Quercetin	7.59	6.34
Dihydrokamepferol	-3.05	6.13
Caffeoylquinate	3.58	3.06
4-coumarate	0.18	0.74
caffeate	10.74	12.36
cyanidin	5.31	3.44
3,7-di-O-methylquercetin	6.64	10.18
Cyanidin-3-O-beta-D-glucoside	2.19	1.79
cyanidin 3-(p-coumaroyl)-glucoside	2.8	9.7
Pelargonidin-3-O-beta-D-glucoside	0.62	5.5
Pelargonidin-3,5-diglucoside	-12.61	-1.92
3,7,4'-tri-O-methylquercetin	2.31	3.97
3,7-di-O-methylquercetin	4.6	4.2
Quercetin-3,7,3',4'-tetrasulphate	1.9	3.2
Quercetin-3,3',7-trissulfate	2.1	2.91
Quercetin-3,3'-bisulfate	3.12	2.09
4-coumaroylshikimate	-0.01	-1.9
Shikimate	3.22	2.1
Sinapoyl-(S)-malate	1.42	1.06
Sinapaldehyde glucoside	5.41	11.26
UDP-L-arabinose	4.2	4.83
UDP-D-apiose	3.71	4.09
UDP-4-dehydro-6-deoxy-D-glucose	5.33	3.95
GDP-L-gulose	-2.01	-3.66
UDP-galactose	2.1	3.8

Metabolites from phenylpropanoid metabolic network and its conjugation were up-regulated in both TTG1:GR and TT8:GR transgenic lines in response to dex treatment.

4.3.3 Metabolites from isoprenoid metabolic pathways and its branches are up-regulated in dex-treated TTG1:GR

Perturbation in metabolic pathways in response to induced activation of TTG1 and TT8 was not limited to phenylpropanoid metabolic network, but other pathways too showed up-regulated response. Metabolic pathways other than phenylpropanoid metabolic network, which showed up-regulation of 4 or more metabolites from a pathways in both TTG1 and TT8 induced activation included mevalonate biosynthesis pathways, methylerythritol phosphate pathway (non mevalonate isoprenoid biosynthesis), trans-zeatin biosynthesis, methylquercetin biosynthesis, de novo biosynthesis of purine and pyrimidine nucleotides, brassinosteroid biosynthesis II, gibberellin biosynthesis I (non C-3, non C-13 hydroxylation) and chlorophyll a biosynthesis II (Table 4.3). Except de novo biosynthesis of purine and pyrimidine, rest all of the up-regulated metabolic pathways forms different branches of isoprenoid biosynthesis pathways and are collectively called as isoprenoid metabolic network.

Isopentenyl diphosphate (IPP) and its isoform dimethylallyl-pyrophosphate (DMAPP) serves as building blocks for all isoprenoids and are acted upon by different enzymes leading to synthesis of cytokinins, sterols, polyprenoids, carotenoids and other isoprenoids. In plants, IPP and DMAPP are synthesized by two alternative pathways, namely mevalonate (MVP) that is localized in cytosol and methylerythritol phosphate pathway (MEP) which is localized in plastids. IPP and DMAPP were up-regulated by 13.31 and 3.52 fold change respectively in dex-treated TTG1:GR transgenic lines. The cytoplasmic MVP begins with condensation of 3 units of acetyl-CoA to 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) and continues by reduction to mevalonate (MVA), followed by two successive phosphorylation steps at C-5 of

mevalonate, and a decarboxylation/elimination step leading to IPP (Croteau et al., 2000). In mevalonate pathways, levels of acetyl-CoA and mevalonate-diphosphate were up-regulated by more than 2 fold in dex-treated TTG1:GR while HMG-CoA, mevalonate-phosphate and mevalonate showed moderate up-regulation, suggesting over-all up-regulation of mevalonate pathways in dex-treated TTG1:GR (Fig 4.32).

MEP pathway, also referred as non-mevalonate pathway is alternate pathway for IPP and DMAPP biosynthesis in plastids (chloroplast). The initial reaction of MEP pathway involves condensation of (hydroxyl ethyl) thiamine with the C-1 aldehyde group of D-glyceraldehyde-3-phosphate (GA-3P) resulting in production of 1-Deoxy-D-xylulose 5 phosphate (DXP). DXP then undergoes intramolecular rearrangement and reduction leading to formation of 2-C-methyl-D-erythritol 4-phosphate (MEP). MEP thus formed undergoes several rearrangement and phosphorylation and finally leads to production of a mixture of IPP and DMAPP. Out of nine metabolic intermediates of MEP pathway, six of them were up-regulated by more than 2 fold change in dex-treated ttg1: gr transgenic lines (Fig 4.33). Metabolites such as 2-phospho-4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol,4-(cytidine-5'diphospho)-2-C-methyl-D-erythritol,1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate, 1-deoxy-D-xylulose 5-phosphate, dimethylallyl-pyrophosphate, 2-C-methyl-D-erythritol-4-phosphate and CMP from MEP pathway were all up-regulated in both TTG1:GR and TT8:GR. Induced activity of TTG1 showed stronger effect as up-regulation of metabolites from isoprenoid metabolic network with higher fold change in all metabolites compared to TT8. Metabolites from isoprenoid metabolic network such as geranyl diphosphate, mevalonate-5-diphosphate and isopentenyl diphospahte which

are very important steps in isoprenoid metabolic network were up-regulated only in TTG1:GR and not in TT8:GR (Fig 4.34).

Synthesis of IPP and DMAPP can be conceptually viewed as the first module of isoprenoid metabolic network. The second order of complexity in isoprenoid synthesis requires condensation of IPP subunits (C5) to generate linear polymer of defined chain length (i.e. e.g. C10, C15, C12……. $C_{n)}$. These molecules serve as direct precursors for synthesis of different classes of isoprenoid compounds. For instance, DMAPP (C5) acts as precursor for synthesis of isoprene, cytokinins and other hemiterpenoids; geranyl diphosphate (GPP: C10) serves as precursor for monoterpenoids; geranyl geranyl diphosphate (GGPP: C20) is used as a precursor for synthesis of gibberellic acid and of side chains of chlorophyll, phylloquinones and tocopherols; squalene (C30) serves as precursor for sterols; brassinosteroids and other triterpenoids, phytoene (C40) is a precursor of abscisic acid (ABA), carotenoids and other apocarotenoids.

Up-regulation of IPP and DMAPP resulted in increased synthesis of several of these isoprenoids. Metabolites from brassinosteroids biosynthesis pathway, namely 6-Oxocampesterol, (22α) - hydroxycampesterol and (22α) -hydroxycampestan-3-one showed up-regulation in response to dex-treated ttg1: GR transgenic lines. Gibberellin biosynthesis pathway was also up-regulated with up-regulation of GA13, GA17, GA19, GA25 and GA36 by more than fold change of 2. Biosynthesis of essential isoprenoids synthesized from IPP and DMAPP such as cytokinins, IAA and abscisic acid (ABA) were also up-regulated in dex-treated TTG1:GR transgenic lines.

Intermediates of isoprenoid metabolic network serves as starting point for synthesis of several metabolic pathways such as brassinosteroids, carotenoids, cytokinins, dolichols, polyprenols, phytosteroids, sesquiterpenes, thiamine, isoprenes, ubiquinones, monoterpenes, chlorophylls, tocopherols, plastoquinones, phylloquinones, ABA and gibberellins biosynthetic pathways. Among up-regulated metabolic pathways, methylerythritol phosphate pathways, brassinosteroid biosynthesis, chlorophyll a biosynthesis, trans-zeatin biosynthesis, gibberellin biosynthesis and cytokinins biosynthesis forms part of isoprenoid metabolic network (Table 4.3). Metabolites such as cis- zeatin, trans- zeatin riboside diphosphate, trans-zeatin riboside triphosphate, trans- zeatin riboside, kinetin-7-N-glucoside and , kinetin-9-N-glucoside from cytokinins pathway and 9'- cis- neoxanthin, L-leucine and indole-3-glycerol-phosphate from IAA biosynthesis pathway were up-regulated in dex-treated TTG1:GR transgenic lines (Table 4.3). Abscisic acid biosynthesis was also up-regulated with the up-regulation in the level of intermediate metabolites such as trans-neoxanthin, 9'- cis-neoxanthin, biolaxanthin, phaseic acid and 8'-hydroxyabscisate.

Compared to dex-treated TTG1:GR, ttg1 mutant lines showed down-regulation of metabolites from isoprenoid metabolic network. Metabolites from both MVA and MEP pathways leading to synthesis of IPP and DMAPP were down-regulated in ttg1 mutant lines (Fig 4.35). Overall, metabolites from isoprenoids and its branch pathways showed were significantly down-regulated while dex-treated TTG1:GR transgenic lines showed up-regulated trend in these pathways. Metabolic profiling of dex-treated TTG1:GR transgenic lines showed up-regulation in biosynthesis of

several isoprenoids, thus, suggesting important role of TTG1 in regulation of isoprenoid metabolic network.

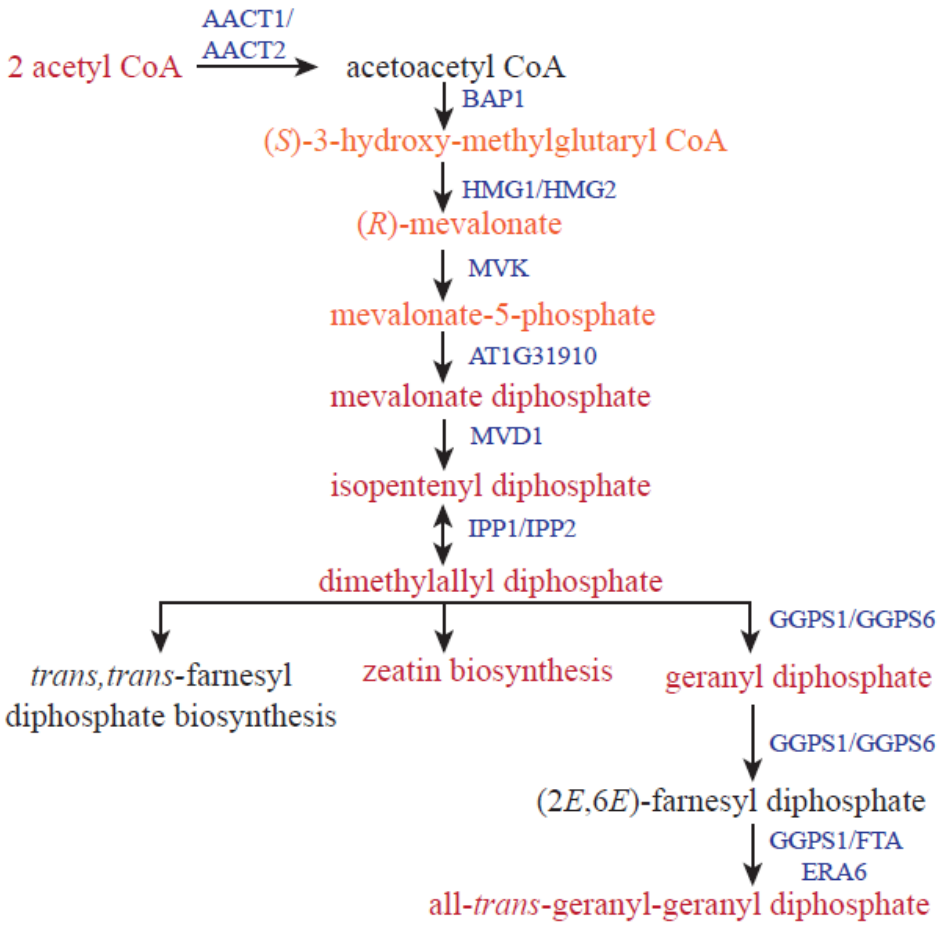


Fig 4.32 Altered metabolites from mevalonate biosynthesis pathway. Mevalonate (MVA) pathway is localized to the cytosol and is responsible for the synthesis of sterols, certain sesquiterpenes, and the side chain of ubiquinone. All coloured metabolite shown in this pathways were differentially expressed metabolites in dex-treated TTG1:GR w.r.t mock treatment with p-value 0.05 or bellow, red coloured metabolites showed up-regulation with 2 or above fold change while those coloured orange were above 1.5 but less than 2 fold change, black coloured metabolites showed no significant difference in dex and mock treated TTG1:GR transgenic lines. AACT1/AACT2- acetoacetyl- CoA thiolase; BAP1- 3-hydroxy-3-methylglutaryl coenzyme A synthase; HMG1/HMG2- 3-hydroxy-3-methylglutaryl coenzyme A reductase; MVK- mevalonate kinase; AT1G31910- phosphomevalonate kinase; MVD1- mevalonate diphosphate decarboxylase; IPP1/IPP2- isopentenyl diphosphate isomerase; GGPS1/GGPS6- geranylgeranyl pyrophosphate synthase. This pathways is adapted from aracyc (http://pmn.plantcyc.org/ARA/NEW-IMAGE?type=PATHWAY&object=PWY-922). Data for these metabolites are shown in Table 4.3.

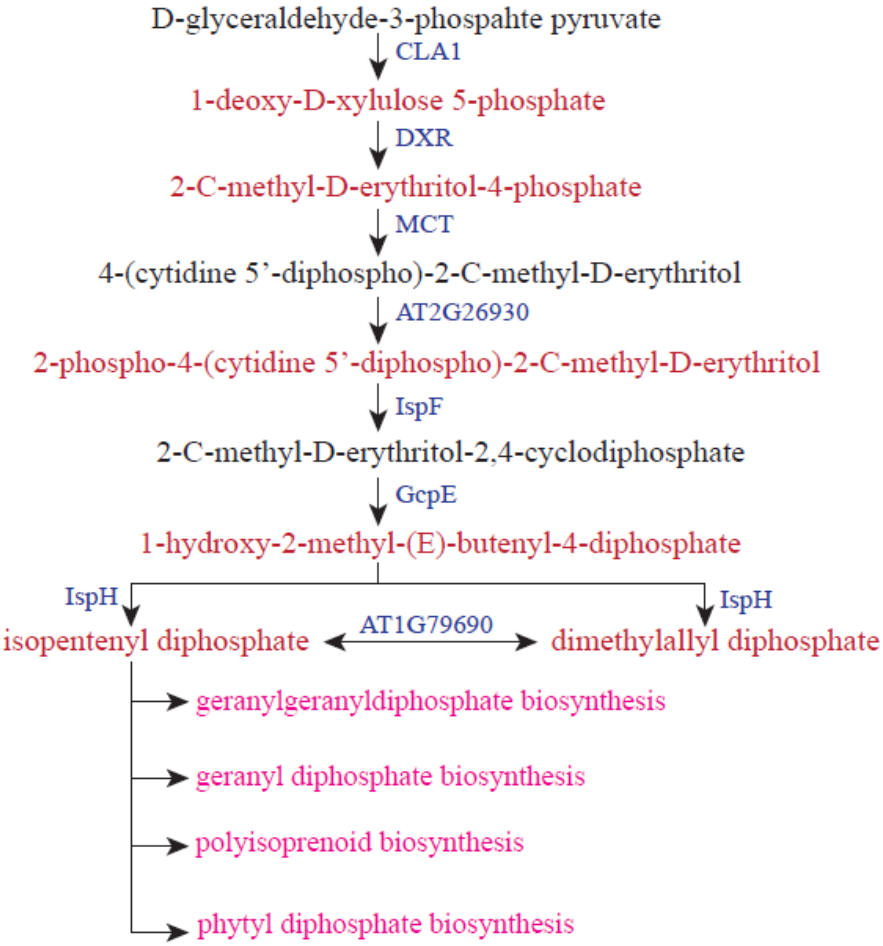


Fig 4.33 Altered metabolites from non-mevalonate biosynthesis pathway (methylerthritol biosynthesis pathways). Non-mevalonate biosynthesis pathway operates in plastids and is involved in providing the precursors for monoterpenes, certain sesquiterpenes, diterpenes, carotenoids, and the side chains of chlorophylls and plastoquinone. All coloured metabolite shown in this pathways were differentially expressed metabolites in dex-treated TTG1:GR w.r.t mock treatment with p-value≤0.05, red coloured metabolites showed up-regulation with 2 or above fold change, black coloured metabolites showed no significant difference in dex and mock treated TTG1:GR transgenic lines. CLA1- 1-deoxy-D-xylulose-5-phosphate synthase; DXR- 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT- 4-diphosphocytidyl-2C-methyl-D-erythritol synthase; AT2G26930- 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol kinase; IspF- 2C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; GcpE- 4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase; IspH- 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate reductase. This pathway map is adapted from aracyc(http://pmn.plantcyc.org/ARA/NEWIMAGE?type=PATHWAY&object=NONMEVIPP-PWY&detail-level=2&detail-level=1&detail-level=2). Data for these metabolites are shown in Table 4.3.

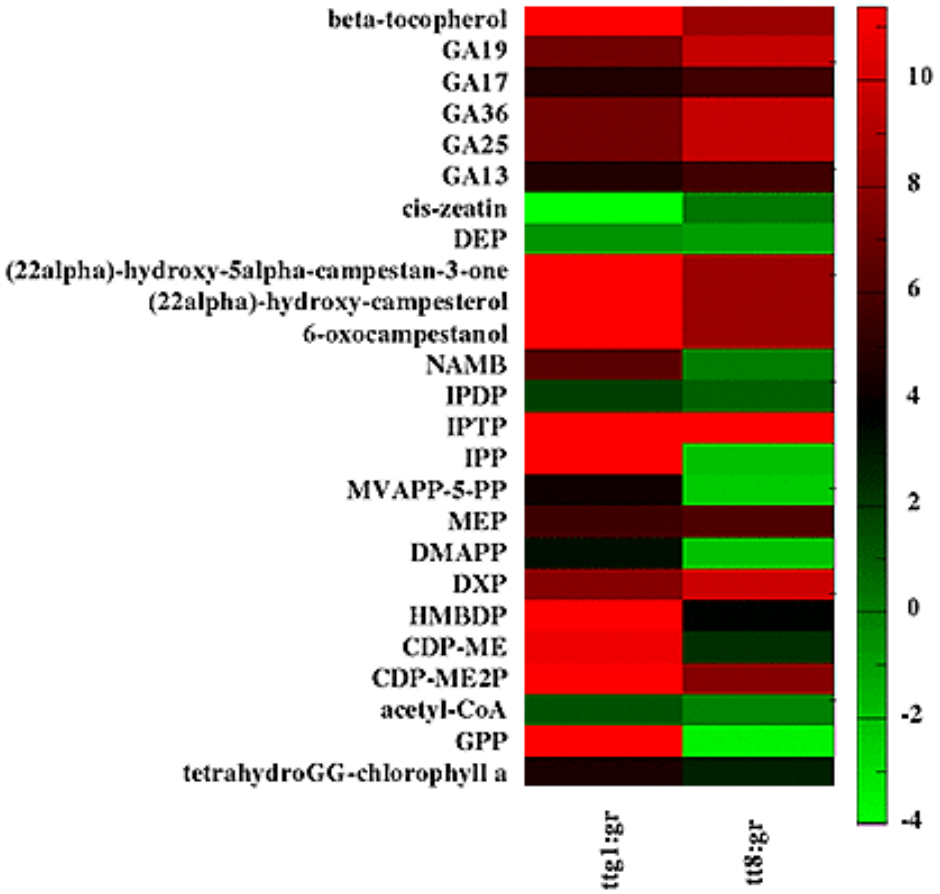


Fig 4.34 Differentially metabolites from Isoprenoid metabolic network (MVP and MEP) in dex-treated TTG1:GR and TT8:GR w.r.t their mock treated transgenic lines. Here colour code represents fold change of dex-treated TTG1:GR and TT8:GR with their mock treated transgenic lines. Metabolites with p-value≤0.05 and fold change±2 and mapped to isoprenoid metabolic network are presented here. GA- gibberellic acid; DEP- 1-deoxy-D-xylulose 5-phosphate; IPDP- isopentenyladenosine-5'-diphosphate; IPTP- isopentenyladenosine-5'-triphosphate; IPP- isopentenyl diphosphate; MVAPP-5PP- dimethylallyl diphosphate; MEP- 2-C-methyl-D-erythritol-4-phosphate; DMAPP- dimethylallyl diphosphate; DXP- 1-deoxy-D-xylulose 5-phosphate; HMBDP- (E)-4-hydroxy-3-methylbut 2-en-1-yl diphosphate; CDP-ME- 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol; GPP- geranyl diphosphate. Data for these metabolites are shown in Table 4.3.

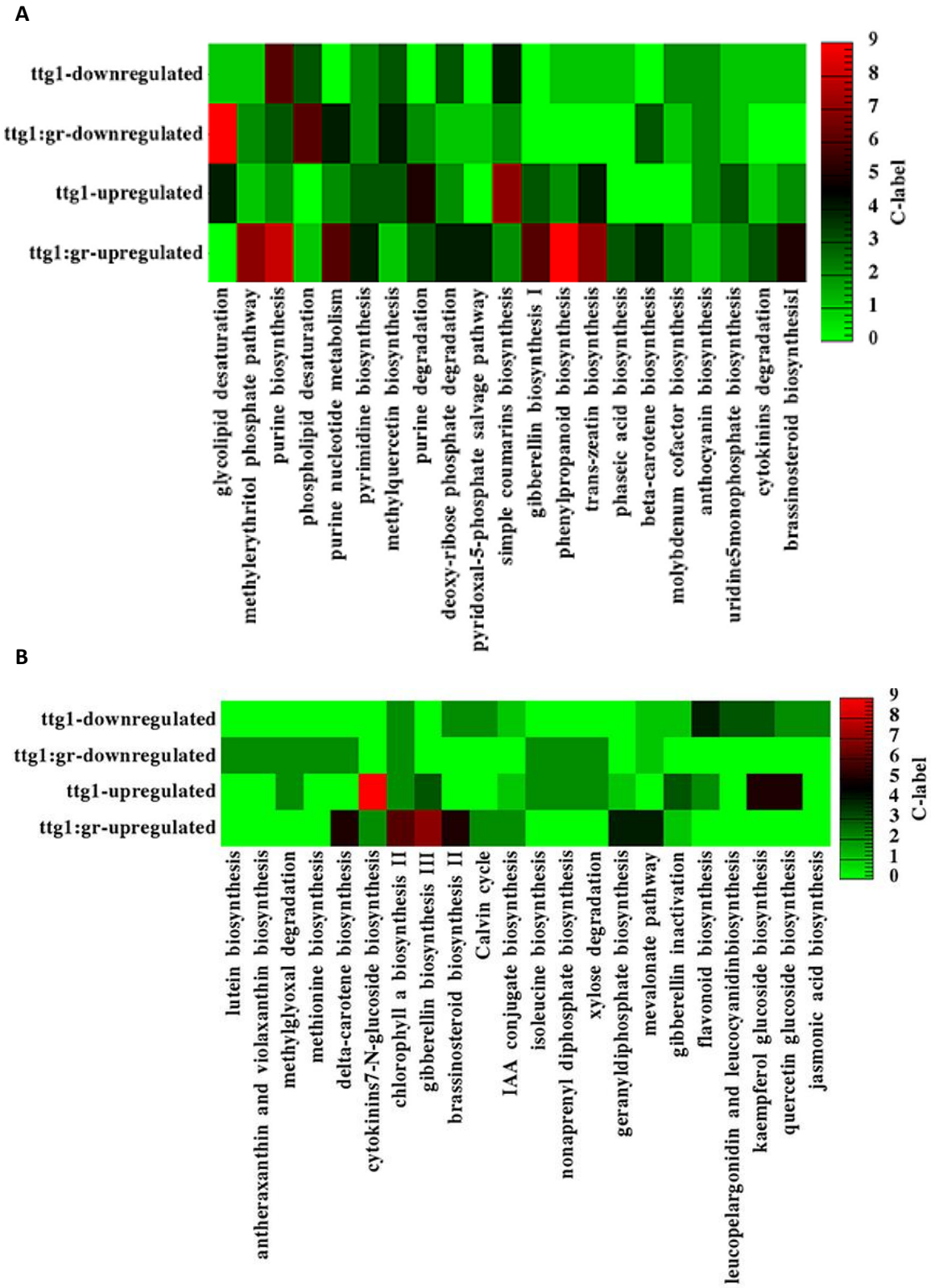


Fig 4.35 Heat map showing number of differential metabolites from metabolic pathways affected in TTG1 loss-of-function mutant line and dex inducible transgenic lines. Metabolites that are up-regulated and down-regulated in ttg1 mutant line and dex-treated TTG1:GR respectively, are here plotted as heat-map

Table 4.3 Differential metabolites from isoprenoid branch pathways in dex-treated TTG1:GR and TT8:GR transgenic lines w.r.t mock treatment.

Metabolite	Dex-treated TTG1:GR	Dex-treated TT8:GR	Pathway
Uridine	5.35	7.9	(deoxy)ribose phosphate degradation
Cytidine	8.09	8.17	(deoxy)ribose phosphate degradation
deoxyribose-5-phosphate	7.77	9.83	(deoxy)ribose phosphate degradation
deoxyribose-1-phosphate	7.77	9.83	(deoxy)ribose phosphate degradation
Thymine	16.94	16.94	(deoxy)ribose phosphate degradation
Thymidine	-0.09	-4.78	(deoxy)ribose phosphate degradation
trans-neoxanthin	0.06	1.75	abscisic acid biosynthesis
9'-cis-neoxanthin	0.06	1.75	abscisic acid biosynthesis
cis-zeatin riboside	13.54	0	cis-zeatin biosynthesis
trans-beta-D-Glucosyl-2-hydroxycinnamate	4.64	9.46	coumarin biosynthesis
UDP	0.38	0.7	coumarin biosynthesis
coumarinic acid-beta-D-glucoside	4.64	9.46	coumarin biosynthesis
kinetin-7-N-glucoside	8.39	1.46	cytokinins 7-N-glucoside biosynthesis
kinetin-9-N-glucoside	8.39	1.46	cytokinins 9-N-glucoside biosynthesis
trans-zeatin riboside	13.54	0	cytokinins degradation
4-carboxyaminoimidazole ribonucleotide	7.19	11.82	biosynthesis of purine nucleotides
5-aminoimidazole ribonucleotide	8.09	8.41	biosynthesis of purine nucleotides
5-phospho-beta-D-ribosyl-amine	8.91	3.44	biosynthesis of purine nucleotides
Caffeate	10.74	12.36	esculetin biosynthesis
cis-caffeate	10.74	12.36	esculetin biosynthesis
GDP-4-dehydro-6-deoxy-D-mannose	-2.28	-1.29	GDP-L-fucose biosynthesis I
GA8-catabolite	7.11	9.56	gibberellin inactivation
indole-3-glycerol-phosphate	-10.84	-3.64	IAA biosynthesis I
L-leucine	-8.83	-4.92	IAA biosynthesis II
indole-3-acetyl-beta-1-D-glucose	11.05	10.58	IAA conjugate biosynthesis
UTP	6.58	0	inositol oxidation pathway
3-methylcrotonyl-CoA	0.8	4.16	leucine degradation
UDP-N-acetyl-D-glucosamine	0.97	1.24	lipid-A-precursor biosynthesis
CMP	12.4	13.92	methylerythritol phosphate pathway
sinapaldehyde glucoside	5.41	11.26	monolignol glucosides biosynthesis
2,4-dihydroxycinnamate	10.74	12.36	simple coumarins biosynthesis
Shikimate	2.67	1.89	simple coumarins biosynthesis
sinapoyl-(S)-malate	1.42	1.06	sinapate ester biosynthesis
beta-tocopherol	13.77	8.33	vitamin E biosynthesis
6-oxocampestanol	6.91	3.11	brassinosteroid biosynthesis II
(22alpha)-hydroxy-campesterol	3.8	1.9	brassinosteroid biosynthesis II

(22alpha)-hydroxy-5alpha-campestan-3-one	**4.1**	**2.7**	**brassinosteroid biosynthesis II**
GA25	**2.33**	**3.77**	**gibberellin biosynthesis I**
GA13	**3.71**	**1.69**	**gibberellin biosynthesis I**
GA36	**3.23**	**-1.76**	**gibberellin biosynthesis II**
GA19	**4.12**	**-2.11**	**gibberellin biosynthesis III**
GA17	**2.19**	**1.98**	**gibberellin biosynthesis III**
GA8-catabolite	**1.66**	**2.33**	**gibberellin inactivation**
3-methyl-4-cis-hydroxy-2-butenal	**1.78**	**1.94**	**cytokinins degradation**
trans-lycopene	**2.37**	**2.34**	**delta-carotene biosynthesis**
isopentenyladenosine-5'-triphosphate	**3.11**	**1.8**	**trans-zeatin biosynthesis**
zeatin riboside triphosphate	**2.73**	**2.11**	**trans-zeatin biosynthesis**
zeatin riboside	**2.56**	**2.97**	**cis-zeatin biosynthesis**
zeatin riboside diphosphate	**5.63**	**2.54**	**trans-zeatin biosynthesis**
isopentenyladenosine-5'-diphosphate	**2.77**	**2.09**	**trans-zeatin biosynthesis**
GMP	**6.97**	**0.02**	**purine degradation**
dGDP	**-5.52**	**-1.7**	**purine nucleotide metabolism**
dGTP	**1.71**	**3.1**	**purine nucleotide metabolism**
dADP	**4.85**	**7.9**	**purine nucleotide metabolism**
ADP	**-5.52**	**-1.7**	**pyrimidine nucleotide metabolism**
dTTP	**-6.49**	**-7.77**	**biosynthesis of pyrimidine deoxyribonucleotides**
dTDP	**-6.4**	**-3.1**	**biosynthesis of pyrimidine deoxyribonucleotides**
dUTP	**-6.54**	**-0.38**	**biosynthesis of pyrimidine deoxyribonucleotides**
dUDP	**1.59**	**0.98**	**biosynthesis of pyrimidine deoxyribonucleotides**
GDP	**-6.85**	**1.59**	**dolichyl-diphosphooligosaccharide biosynthesis**
dTMP	**3.46**	**5.01**	**formylTHF biosynthesis**
dUMP	**-4.17**	**-2.44**	**formylTHF biosynthesis**
dTDP-alpha-L-rhamnose	**8.99**	**4.7**	**dTDP-L-rhamnose biosynthesis**
dTDP-4-dehydro-6-deoxy-D-glucose	**3.84**	**3.09**	**dTDP-L-rhamnose biosynthesis**

All metabolites showed here were filtered with p-value≤0.05 and fold change ±2. Complete list of differential metabolites has been shown in appendix.

4.3.4 Microarray analysis of loss-of-function mutant lines tt2, tt8 and ttg1, and dex-treated TTG1:GR and TT8:GR

Comparison of metabolic profiles of single transparent testa mutants and double mutant combinations of ttg1 and tt6 with tt3, tt4 and tt5 revealed TTG1 as playing an important role in isoprenoid metabolic network regulation as described in Section 4.2. Loss-of-function of TTG1 showed down-regulation of several metabolites from branch pathways of isoprenoid metabolic network. We next investigated whether the affect of TTG1 perturbation in isoprenoid pathway (1) might involve coordinated changes at RNA level and (2) might be direct responses of TTG1 perturbation. To establish pathways affected directly due to TTG1 activity, we created dex inducible TTG1:GR transgenic lines and analyzed their metabolic profiles. Non-targeted metabolic profiling of dex-treated TTG1:GR showed up-regulation of metabolites from isoprenoid metabolic network with several metabolites down-regulated in ttg1 mutant lines were up-regulated in response to dex treatment (Section 4.3.3).

In order to identify transcriptome changes directly in response to TTG1 and TT8 perturbation, we performed gene microarray analysis of RNA from 6-days-old seedlings with dex or mock treated TTG1:GR, TT8:GR and loss-of-function transgenic lines of TT2, TT8 and TTG1. Gene expression profiles of all transgenic lines were plotted on 2D-PCA which showed formation of two clusters (Fig 4.36). Dex inducible transgenic lines formed one cluster in which mock treated lines were grouped together while dex-treated transgenic lines were far apart. Second group was formed of wild type and tt2, tt8 and ttg1 mutant lines with tt2 being closer to wild type. Differentially expressed genes were filtered based on p-value less than 0.05 and

2 or above fold change. Hierarchical clustering of differentially expressed genes showed that profile of dex inducible transgenic lines treated with mock were similar to Wt, suggesting that inducible system behaves like wild type until it is treated with dex (Fig 4.37). Similar to PCA plot, loss-of-function transgenic lines tt8 and ttg1were clustered together showing similar expression profile, while dex-treated TTG1:GR and TT8:GR showed similar expression pattern.

We then compared gene expression of all loss-of-function transgenic lines with their Wt-Col line as control. Differential genes were filtered based on corrected p-value ≤0.05 and fold change ±2 or above. Among, mutant lines, tt8, ttg1 and tt2 showed 2480, 2453 and 849 differentially expressed genes respectively, with 302 genes common to all three (Fig 4.38). Differential genes common to all three mutant lines showed high similarity in expression pattern with 255 out of 302 differentially expressed genes similar trend. TTG1 and TT8 loss-of-function mutant lines showed high similarity with 743 differentially expressed genes out of 985 common genes showing similar expression pattern.

Consistent with earlier reports, BAN and DFR, two biosynthetic genes regulated by TTG1-TT8-TT2 complex were down-regulated by more than 2-fold in all three mutant lines (Baudry et al., 2004, 2006). Other genes such as *PCC1* (Pathogen and circadian controlled 1), *STP1* (Sugar transporter 1), *UGT85A1, RPS5* (Resistant to *P. Syringae*), *ATMYB16, GL2* (GLABRA 2), *UGT74D1, UGT73B5* and *ERD5* (Early responsive to dehydration 5) were significantly down-regulated in *tt2, tt8* and *ttg1* mutant lines. Several biosynthetic genes from phenylpropanoid metabolic network such as *TTG2* (Transparent testa glabra 2), *LDOX* (Leucocyanidin oxygenase), *TT7*

(Transparent testa 7), *TT5* (Transparent testa 5) and *TTG1* were significantly down-regulated in *ttg1, tt8* and *tt2* mutant lines.

TTG1 loss-of-function transgenic lines showed several genes from isoprenoid metabolic network such as geranyl diphosphate synthase (GPPS), p-hydroxyphenylpyruvate synthase (HPPS), farnesyl diphosphate synthase (FPPS2), gibberellin 20-oxidase, sterol C 22-desaturase (CYP710A1), sterol C5-desaturase (DWF7-2), isopentenyl transferase (IPT4), sesquiterpenoid/diterpenoid synthase TPS18 (TPS16), sesquiterpenoid synthase TPS27 (TPS03), dehydrodolichol diphosphate synthase (DDPPS), 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR1), phosphomevalonate kinase (PMK), dehydrodolichol diphosphate synthase (DDPPS) and farnesyl transferase/geranyl geranyl transferase I, alpha subunit (FTase/GGTase1A) significantly down-regulated (Fig 4.40). Biosynthetic genes from isoprenoid branch pathways such as cytokinins, brassinosteroids, carotenoids, GA biosynthesis and ABA biosynthesis were also down-regulated in ttg1 and tt8 mutant line. Table of entire list of differentially expressed genes and comparison for ttg1, tt8 and tt2 mutant lines have been shown in Appendix1. TT2 loss-of-function transgenic lines showed no differentially expressed gene from isoprenoid metabolic network while TT8 loss-of-function transgenic line showed few down-regulated genes. This was consistent with metabolic profiling of loss-of-function transgenic lines that showed lowered levels of several metabolic intermediates from isoprenoid metabolic network in TTG1 loss-of-function transgenic lines (Section 4.2 and 4.3.3). Major metabolic pathways other than phenylpropanoid and isoprenoid metabolic network that showed differential expressed genes includes sciadonic acid biosynthesis, pyrimidine ribonucleotides *de novo*biosynthesis, jasmonic acid biosynthesis,

glutathione-mediated detoxification II, coumarin biosynthesis (via 2-coumarate), quercetin glucoside biosynthesis and sucrose degradation III.

To identify direct target genes of TTG1 and TT8, we analyzed microarray-based gene expression profile of TTG1:GR and TT8:GR with dex and mock treatment. Our rationale was that differentially expressed genes down-regulated in loss-of-function transgenic lines and up-regulated in dex induced transgenic lines would be ideal candidates for being a direct target. As many as 1,253 and 1,283 differentially expressed genes were identified from dex-treated TT8:GR and TTG1:GR, respectively, with 137 genes common to both transgenic lines. Interestingly, only 4 out of 137 common differentially expressed genes differed in expression pattern, suggesting many common targets but not all among TTG1 and TT8. This was expected since both transcriptional factors have been shown to physically interact (baudry et al., 2006).

Gene ontology of differentially expressed genes of loss-of-function mutants and dex-treated TTG1:GR and TT8:GR transgenic lines were plotted for molecular functions and biological processes using http://www.arabidopsis.org/tools/bulk/go/index.jsp. Biological processes such as metabolic process, stress response, stress response to biotic and abiotic stress, developmental process, protein metabolism, cell organization and biogenesis, DNA-RNA metabolism were enriched in all transgenic lines. TTG1 loss-of-function and dex induced transgenic lines showed high enrichment of developmental processes compared to rest transgenic lines which is consistent with several reports suggesting TTG1 involved in various developmental processes in *Arabidopsis*. GO subcategories terms of differentially expressed genes for all transgenic lines are given in Appendix. For example, GO subcategories term for

oxidation-reduction reaction, secondary metabolic process and phenylpropanoid biosynthesis process was enriched in both dex-treated TT8:GR and TTG1:GR while it was comparatively less in other transgenic lines.

After GO enrichment analysis, we analyzed genes that showed reciprocal behavior, i.e down-regulated in mutant lines and up-regulated in dex-treated transgenic lines and vice-versa for TT8 and TTG1. For TTG1, there were 261 common differentially expressed genes between ttg1 mutant and dex-treated TTG1:GR, with 185 genes showing reciprocal behaviour. Among 185 common genes with reciprocal expression pattern, 113 were down-regulated in ttg1 mutant and were up-regulated in dex-treated TTG1:GR. Overall, dex-treated TTG1:GR transgenic lines showed increased expression of genes that were down-regulated in TTG1 loss-of-function. Among common genes showing reciprocal expression pattern in TTG1 loss-of-function and dex-induced transgenic lines, genes such as *GL2* (Glabra 2), *AT1G77210* (sugar transporter), *AT3G11930* (universal stress protein family protein), *ATPSK4* (Phytosulfokine 4 precursor), ERD5 (Early responsive to dehydration), *BGL2* (Beat-1,3-glucanase 2), *VSP1* (Vegetative storage protein 1), *PCC1* (Pathogen and circadian controlled 1), *PR1*, *PR5*, *DFR*, *BAN*, *TT7*, *TT10*, *CYP710A1* (Sterol C 22-desaturase), *UPPS* (Undecaprenyl diphosphate synthase), *At5g56300* (methyltransferases; predominant methylation of C19-GAs), *PPH* (Pheophytine pheophorbide hydrolase), *GPPS* (geranyl diphosphate synthase), *HPPD* (p-hydroxyphenylpyruvate dioxygenase), *TPS18* (Sesquiterpenoid/diterpenoid synthase TPS13), *HPPS* (p-hydroxyphenylpyruvate synthase) and *KAO* (ent-kaurenoic acid oxidase) were down-regulated in ttg1 and up-regulated in dex-treated TTG1:GR transgenic lines (Fig 4.40). As expected, BAN and DFR from phenylpropanoid metabolic network were up-regulated in dex-treated TTG1:GR and down-regulated in ttg1 mutant lines.

Interestingly, expression levels of TT7 and TT10 from phenylpropanoid metabolic network expression showed dependence upon TTG1 which has not been reported earlier. Genes related to pathogenesis and circadian controllers such as *PCC1, PR1* and *PR5* were significantly down-regulated in ttg1 mutant and were up-regulated in TTG1:GR lines. Several genes from isoprenoid metabolic network genes such as *GPPS, UPPS, CYP710A1, At5g56300, PPH, HPPD, HPPS* and *KAO* showed increased expression levels in dex-treated TTG1:GR while were down-regulated in ttg1 mutants. Thus genes from phenylpropanoid, isoprenoid metabolic network and from biotic-abiotic stress response showed reciprocal expression profile for ttg1 mutant and dex-treated TTG1:GR transgenic lines.

Apart from genes showing reciprocal behaviour in *ttg1* loss-of-function and dex-induced transgenic lines, there were several other genes that showed increased expression in dex-treated TTG1:GR while were unaffected or *slightly down-regulated (-1 to -1.4-fold change) in ttg1 mutant lines. Genes such as UGT76E12, HMG1, HMG2, SUS2, SUS3, SUS4, DMR6, LOX1, LOX5, AOC2, A1CSGL2, YAP169, AT1G22400, SAG13, AOXID, CYP71A13, XDH1, XGD1, IBR3, AAO1, ACS2, VSP2, DDTG1, DXR* and *DXPS3* among several others were up-regulated in dex-treated TTG1:GR while unaffected in TTG1 loss-of-function transgenic lines. Out of above mentioned genes, *HMG1, HMG2, DXR, DXPS3, YAP169, UGT76E12* and *IBR3* are from isoprenoid metabolic network while *LOX1, LOX5* and *AOC2* are associated with jasmonic acid biosynthesis. Overall, isoprenoid metabolic network, phenylpropanoid metabolic network, jasmonic acid biosynthesis, scaidonic acid biosynthesis and galactose degradation pathways were among major metabolic pathways that were up-regulated in TTG1:GR transgenic lines in response to dex treatment. Metabolic

profiling of dex induced TTG1:GR transgenic lines also showed increased accumulation of metabolites from phenylpropanoid and isoprenoid metabolic network which was down-regulated in TTG1 loss-of-function mutant lines. Thus, gene expression pattern of metabolic pathways was consistent with our metabolic profiling results.

Twenty seven genes were selected based on showing reciprocal expression pattern for both TT8 and TTG1 to validate micro-array expression profile (Table 4.4). These 27 genes were grouped under three major categories based on their function, namely genes related to metabolic pathway, genes related to biotic stress response and glucoside transporters, transferases and hydrolases. In TTG1:GR, all genes grouped under metabolic pathway genes were up-regulated in dex-treated TTG1:GR and down-regulated in TTG1 loss-of-function mutant lines while only BAN, DFR, BR6OX1, CBB2 and GGR showed such reciprocal behaviour in TT8 (Fig 4.41). Also, as expected the expression level of BAN and DFR were up-regulated in both dex-treated TTG1:GR and TT8:GR transgenic lines and down-regulated in their respective mutant lines. The expression pattern of only BR6OX1, CBB2 and GGR genes grouped under metabolic pathway genes were similar to that observed for BAN and DFR while others such as GL2, TT10, HMG2, ISPD and GPPS showed such reciprocal behaviour only for TTG1. These results suggests TTG1 to be a universal regulator possibly involved in expression regulation for several of these biosynthetic genes while TT8 serves as one of the interacting partners of TTG1 which together regulates expression of some of these genes. Again, all genes grouped under biotic stress resistance genes and glucoside transporters, transferases and hydrolases were up-regulated in dex-treated TTG1:GR transgenic lines and down-regulated in TTG1

loss-of-function mutant lines while in TT8, all genes except *PR1* grouped under biotic stress resistance genes and all glucoside transporters, transferases and hydrolases genes showed reciprocal behaviour. This suggests involvement of both TTG1 and TT8 in regulation of biotic stress response in *Arabidopsis.*

To test whether the genes showing reciprocal behaviour are direct targets of TTG1 and TT8, their expression levels were tested in chx (Cycloheximide), dex+chx, dex and mock treated 6-day-old seedlings of TTG1:GR and TT8:GR transgenic lines. The expression levels of those genes that would be up-regulated in both dex and dex+chx treatments would possibly be direct targets of TTG1 and TT8. In TTG1:GR, the expression levels of genes such as *CYP76C2, CBB2, PR1* and *SUS4* were up-regulated in dex-treated seedlings while down-regulated in dex+chx treated seedlings suggesting TTG1 to be indirectly related to these genes (Fig 4.42F). The expression levels of genes such as *GPPS, DOGT1, SUC7, NIT4, UGT75B1* and *CSLE1* in TTG1:GR were up-regulated by dex treatment while moderately up-regulated by dex+chx treatment suggesting TTG1 to be involved in regulating the expression level of these genes (Fig 4.42-D, E, F). Genes such as *GL2, TT10, BAN, DFR, HMG2, ISPD, GPPS, BR6OX1, GGR, PR5, LURP1, PCC1, PDR12, SAG13, WAK1, WAK3, SUS2* and *UGT85A1* showed significant up-regulated expression levels in both dex and dex+chx treated TTG1:GR seedlings suggesting TTG1 to be directly involved in regulation of expression levels of these genes (Fig 4.42).

In TT8:GR transgenic lines, genes such as *BAN, DFR, GPPS, BR6OX1, GGR, LURP1, DOGT1, PCC1, PDR12, SAG13, CY76C2, WAK1, WAK3, SUS2, SUS4, SUC7, UGT85A1* and *UGT75B1* showed significant up-regulation of expression levels for both dex and dex+chx treatment (Fig 4.42-A, B, C). Genes such as *BAN, DFR, GPPS, BR6OX1, GGR, WAK1, WAK3, PCC1, SUS2, SUC7, UGT85A1, LURP1*

and *SAG13* were significantly up-regulated for dex and dex+chx treatment in both TTG1:GR and TT8:GR transgenic lines suggesting TTG1 and TT8 transcriptional complex as possible regulator of these genes.

Our microarray results suggested TTG1 to be involved in regulation of several genes from isoprenoid and phenylpropanoid metabolic networks as well as genes related to biotic stress response. Quantitative real-time PCR showed genes such as *TT10, BAN, BR6OX1, CBB2, HMG2, GGR, ISPD, PCC1, GPPS* and *PDR12* among several others to be directly associated with TTG1 expression level and activity. To further test if TTG1 physically interacts with promoters of these biosynthetic genes, chromatin Immunoprecipitation assay (ChIP) was performed. Anti-GR antibodies were used to pull down by chromatin bound GR-tagged TTG1 complexes and using quantitative real-time PCR, enrichment of promoters in mock and dex-treated TTG1:GR transgenic lines were tested. Western blotting results for anti GR antibodies showed specific single band suggesting specificity of ChIP experiment (Fig 4.43A). Western blotting results showed increased levels of GR tagged ttg1 in the nucleus compared to mock treated transgenic lines while it was undetected in WT. This suggested that dex treatment was effective and transgenic lines were not leaky i.e., only dex-treated samples showed the protein in nucleus. ChIP samples were then tested for promoter enrichment analysis for 10 selected genes using qRT- PCR approach along with actin as internal control. Promoter fragment for each gene was selected from their respective 0 to -1000 bp regions and 5 sets of primers were designed spanning the selected promoter regions of the target genes. Chip-Quantitative qRT-PCR showed promoter enrichment for all 10 genes in dex-treated TTG1:GR transgenic lines while no effect was observed in mock treated TTG1:GR transgenic lines (Fig 4.43B). This confirmed physical interaction of TTG1 with promoters of these genes, and that TTG1

as direct regulator of these structural genes. As expected, for BAN, enriched promoter were in agreement of bio-informatics predicted regions as well as cis-region predicted by promoter fragmentation and GUS assay (Debeaujon et al., 2003). This result was consistent with our qRT- PCR data suggesting these genes to be directly associated with TTG1. Thus, combining all these results, we have identified new targets of TTG1 transcriptional factor. These target genes are related with phenylpropanoid, isoprenoid metabolic network as well as pathogen responsive genes, showing the diverse roles played by TTG1 in the regulation of several processes.

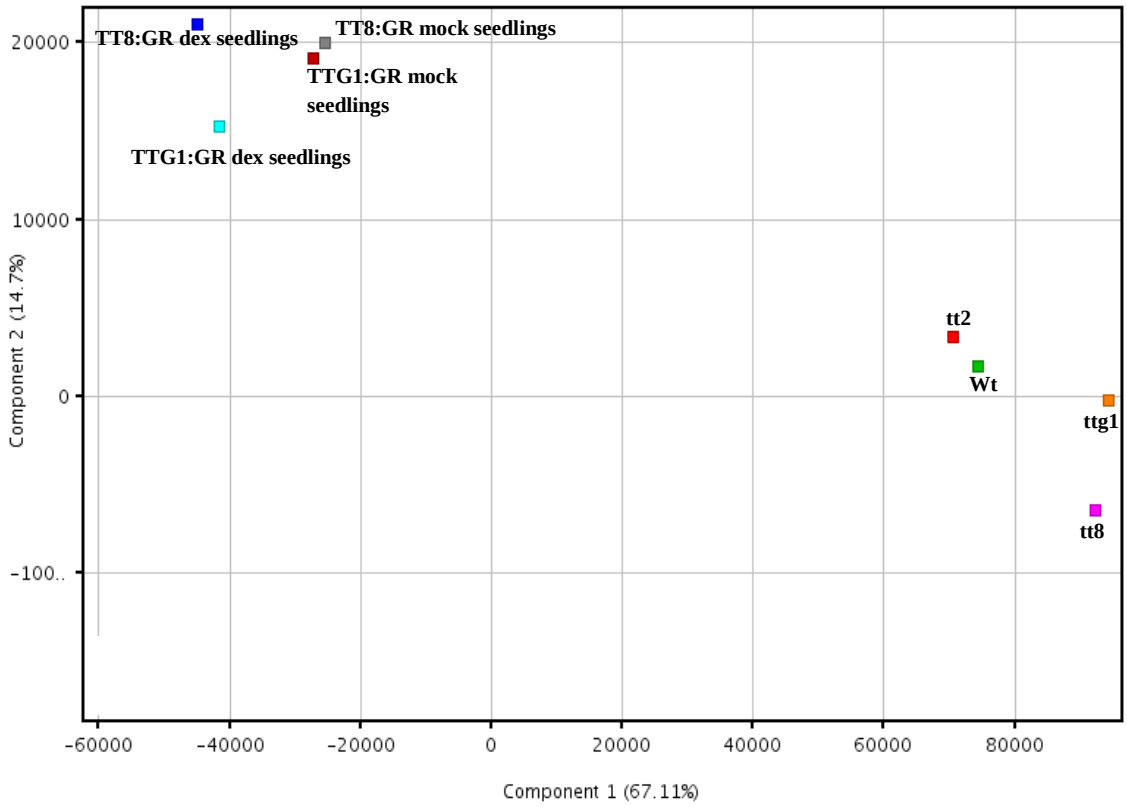


Fig 4.36 2D-PCA plot representing gene expression profile of loss of function and inducible over-expression lines with or without dex treatment. Gene expression profiling for each transgenic lines were performed for 3 biological replicates. PCA showed mock treated TTG1:GR, TT8:GR grouped together while mutants and dex-treated lines were grouped together.

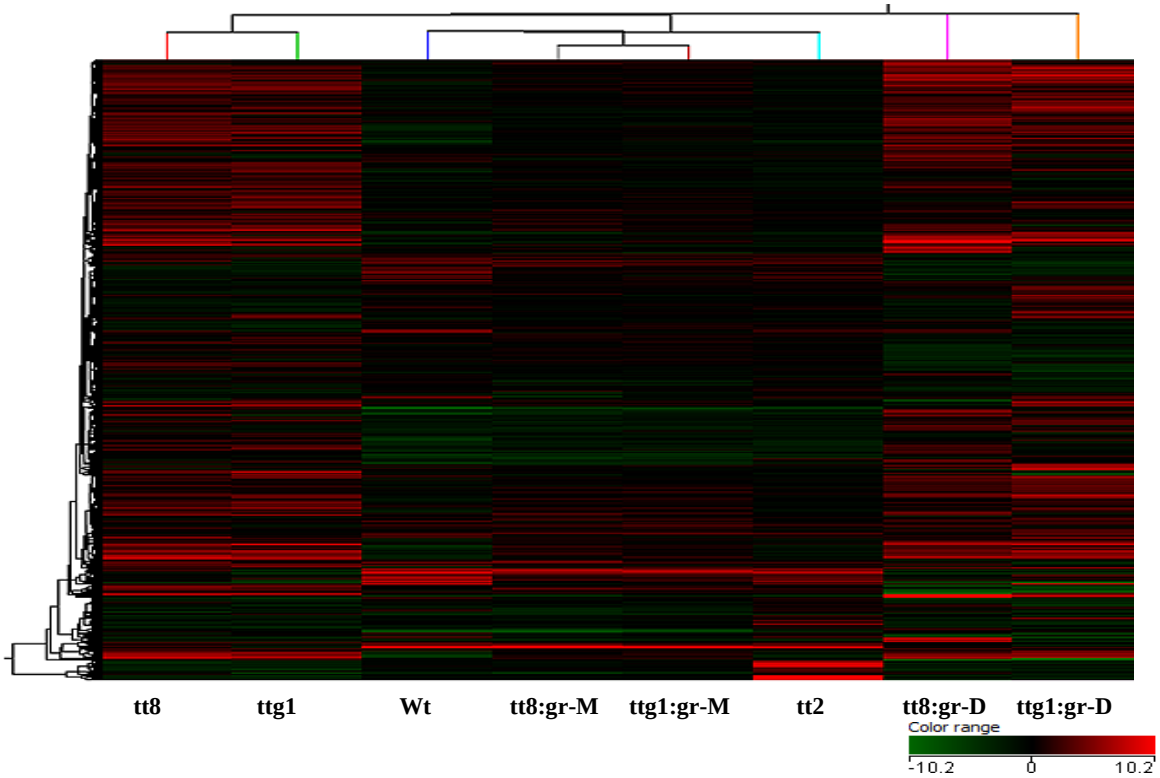


Fig 4.37 Hierarchical clustering of differentially expressed genes of loss of function and inducible over-expression lines with or without dex treatment. Colour code here represents normalized intensity values of each gene represented in the rows. Differential genes were filtered based on p-value≤0.05 and fold change ±2 in at least one of all transgenic lines. Expression values were averaged across biological replicates and each transgenic line had 3 biological replicates that were analyzed using microarray.

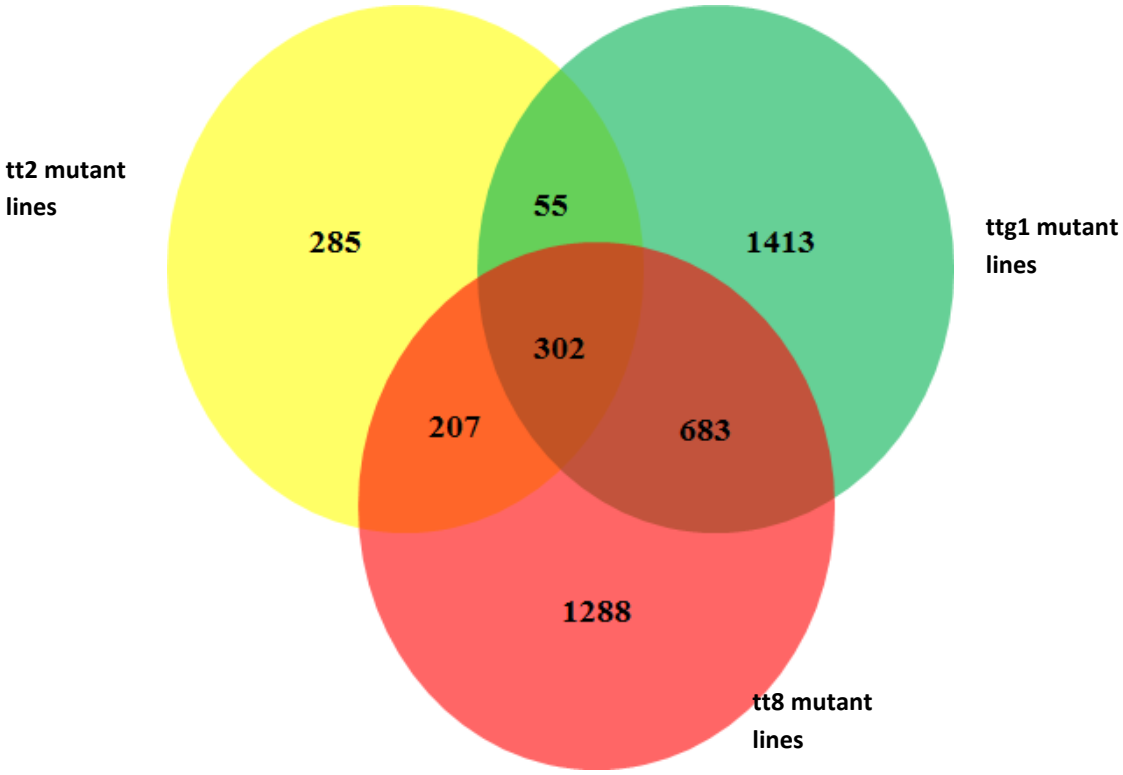


Fig 4.38 Differentially expressed genes between three loss-of-function transgenic lines. Differential genes were filtered based on p-value≤0.05 and fold change ±2 or above. Each transgenic line had 3 biological replicates used for microarray hybridization and samples for all loss-of-function lines were hybridized on single glass slide array.

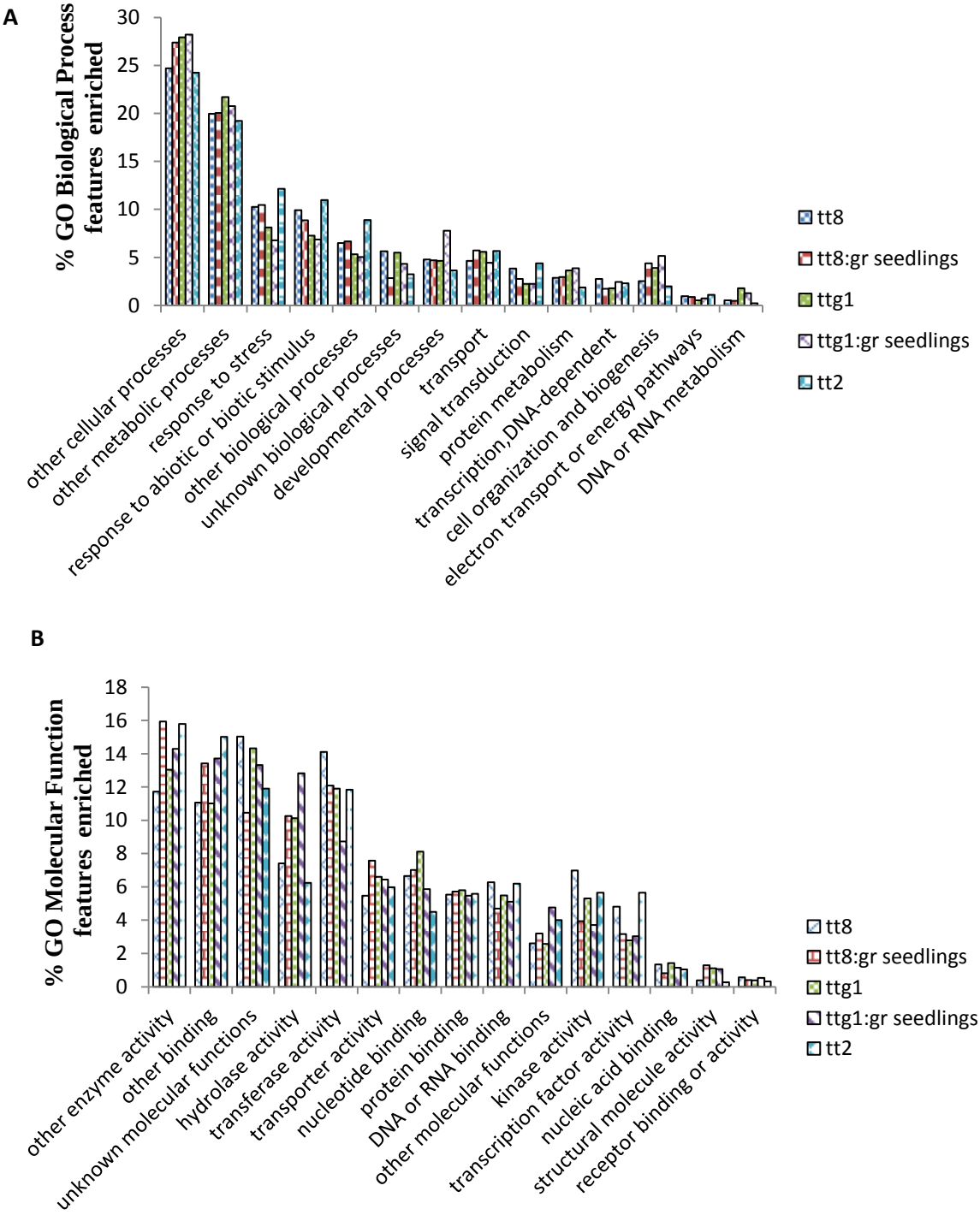
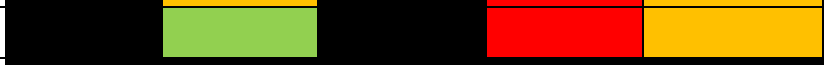
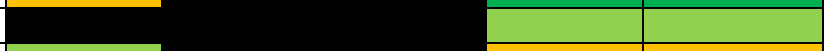
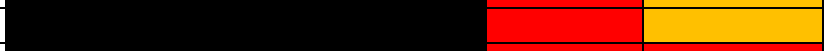
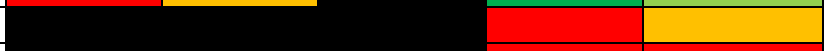
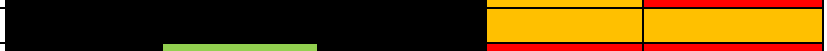
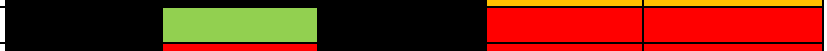
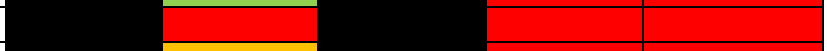
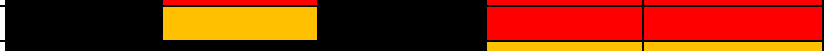
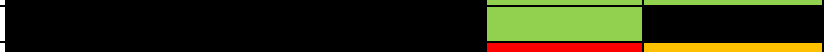


Fig 4.39 Broad GO categories of differentially expressed genes from mutants and inducible over-expression lines. Differentiallly expressed genes of transgenic lines were analyzed using online tool (http://www.arabidopsis.org/tools/bulk/go/index.jsp) and % value of each feature enriched was plotted to visualize biological and molecular function features enriched in different transgenic lines. Here, differential genes were filtered based on p-value≤0.05 and fold change ±2 w.r.t respective controls.

Symbol	tt2 mutant	tt8 mutant	ttg1 mutant	tt8 dex seedlings	ttg1 dex seedling
GGR					
GGPS6					
GPPS					
UGT73B2					
AT1G05680					
UGT73B4					
PKT3					
ATASE2					
UGT75B1					
SAG13					
ADH1					
PDC2					
AOX1D					
UGT72E2					
UGT71D1					
ASP4					
CBB3					
BR6OX1					
ROT3					
BAS1					
UGT73C5					
PAD3					
CYP71A12					
CYP71A13					
AT1G12640					
ATCSLE1					
ATCSLG2					
PCB2					
SQP2					
SQS1					
SQS2					
AT1G73600					
PYK10					
BGLU21					
BGLU10					
BGLU27					
DIN2					
NIT4					
UGT73C5					
CKX4					
UGT85A1					
AT2G36780					
UGT76E12					
CAT1					
ACS2					
CAC1					
GPP2					
DMR6					

ATSRG1	
SUS4	
SPS2	
ATGA2OX2	
AT1G14130	
PGM	
AT2G25450	
CYP79F2	
NSP1	
NSP5	
GDH2	
GAD4	
PPDK	
ATGPX6	
GST20	
GST6	
ATGSTU11	
ATGSTU4	
ADS1	
PFK3	
XDH1	
ATPMEPCRB	
AT5G49180	
ATPMEPCRD	
AAO1	
NIT2	
IAR3	
IBR3	
AT4G18440	
BCAT4	
ACAT2	
ACX1	
OPR2	
BEN1	
AT1G09500	
CYP89A5	
AT1G76470	
LUT5	
ISPD	
DXPS1	
ISPF	
GST8	
HMG1	
AT1G31910	
HMG2	
PAP3	
PSAD-2	
AT5G39050	
ELI3	
CAD1	
CAD5	
PSBQ-2	

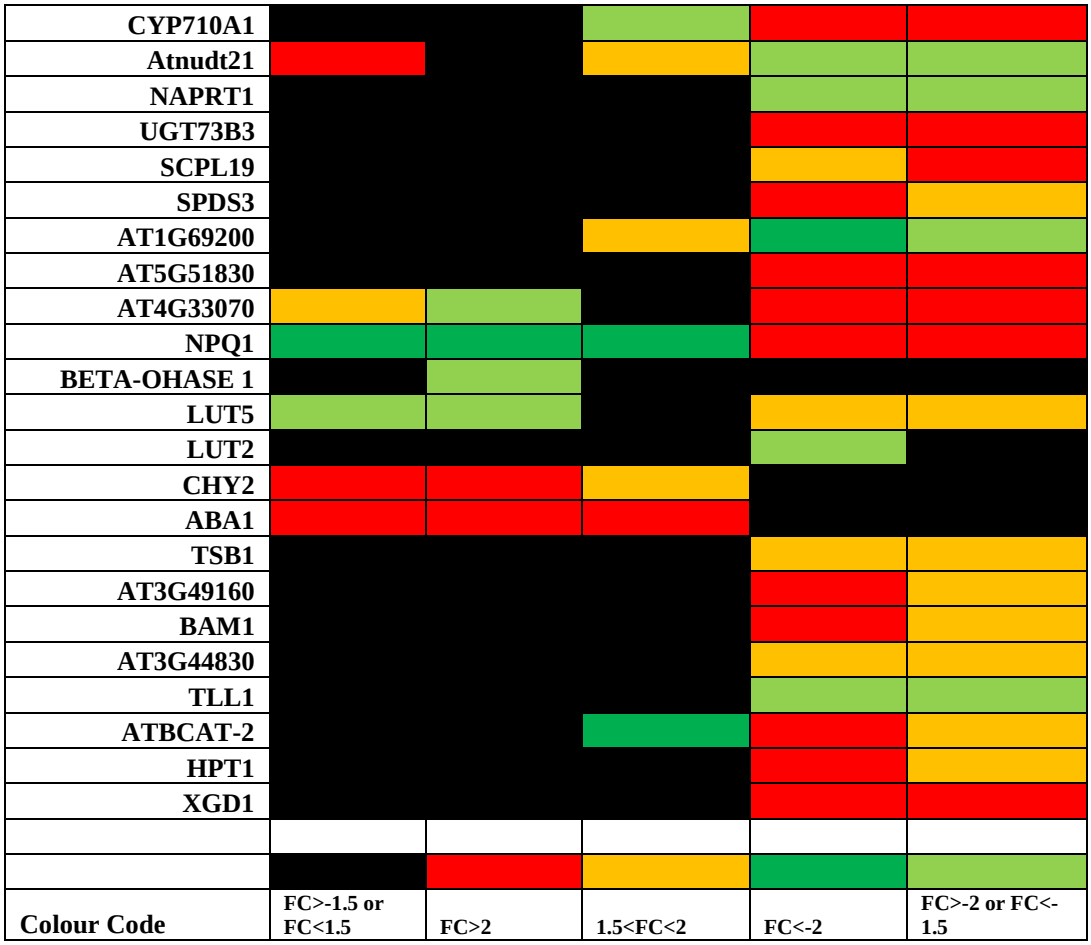


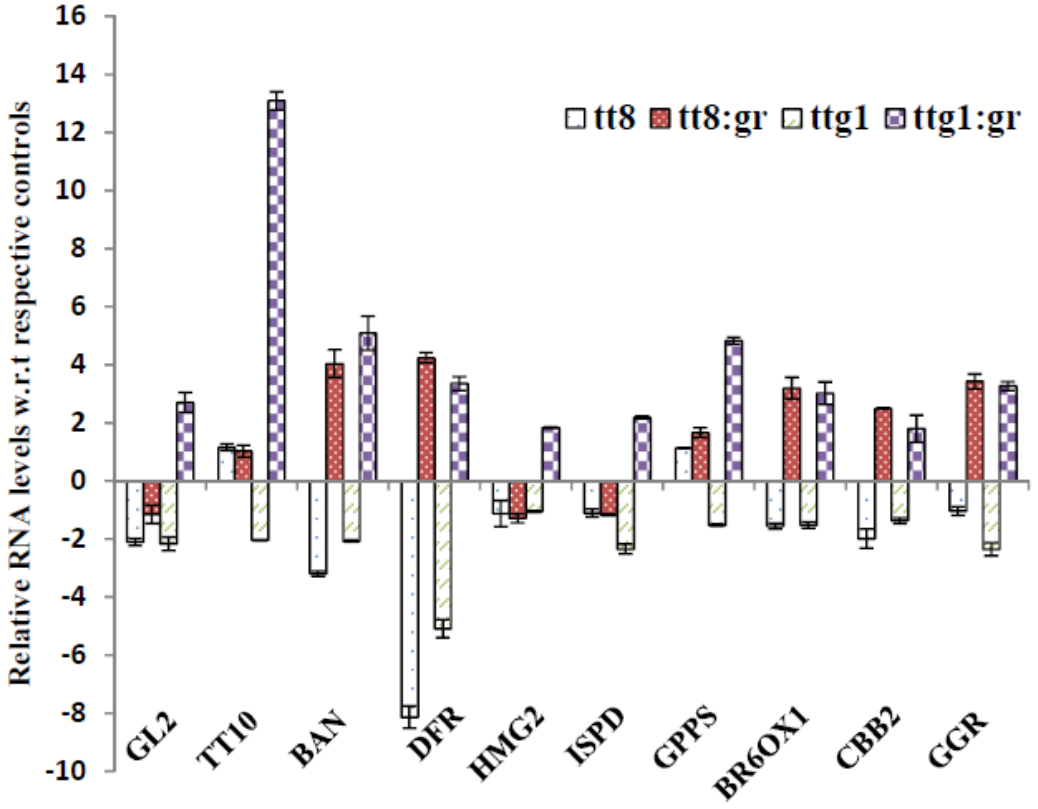
Fig 4.40 Heat map of differential gene list from metabolic pathways showing reciprocal expression pattern between loss-of-function and dex induced transgenic lines. FC stands for fold change. Genes shown here are associated with either metabolic pathways or other biological features. Genes that are associated with metabolic pathways have been shown in Appendix II.

Table 4.4 List of differentially expressed genes selected for quantitative real time PCR analysis.

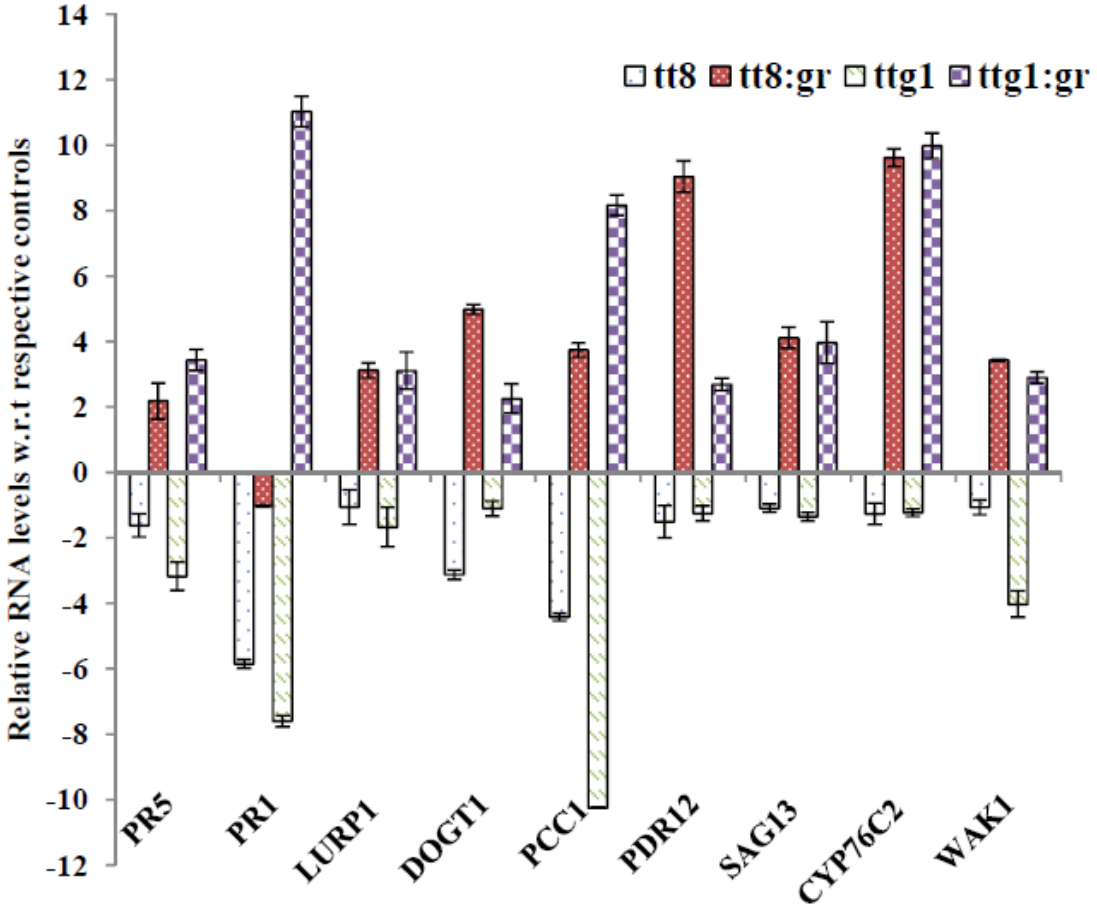
Fold change calculated w.r.t respective controls for transgenic lines							
symbol	tt2 mutant	tt8 mutant	ttg1 mutant	tt8 dex seedlings	ttg1 dex seedling	ttg1 dex siliq	tt8 dex siliq
GL2	1.37	-2.10	-2.17	-1.15	1.70	-1.16	-1.11
TT10	1.35	1.16	-1.04	1.02	44.10	1.08	1.20
SUS2	-1.04	-1.32	-1.06	1.08	5.10	-1.11	1.13
SUS4	1.09	-1.54	-1.40	2.25	1.55	1.01	-1.16
PR5	-1.90	-1.62	1.18	1.18	3.44	4.08	1.89
ATSUC7	-1.10	1.10	-1.37	3.18	2.33	2.57	2.61
PR1	-7.59	-5.85	-7.60	-1.03	11.03	9.59	6.66
LURP1	-2.93	-1.06	-1.67	3.12	3.11	6.12	2.46
DOGT1	-2.47	-3.13	-1.11	4.98	2.25	1.24	1.22
PCC1	-3.94	-44.42	-110.25	3.74	18.17	6.01	1.04
WAK1	-1.68	-1.07	-4.03	3.44	1.90	2.31	1.59
UGT85A1	-2.13	-4.78	-3.65	2.53	1.73	1.28	1.11
NIT4	-1.62	-1.73	-1.78	10.04	5.21	1.74	1.64
WAK3	-1.67	-1.42	-1.73	3.30	2.33	2.21	3.23
CSLE1	-2.01	-2.17	-1.67	2.00	1.63	1.62	-1.10
UGT75B1	-1.60	-2.66	-1.46	2.67	1.59	-1.02	1.20
SAG13	-1.11	-1.10	-1.35	22.11	3.96	2.62	2.05
PDR12	-1.35	-1.51	-1.25	9.04	2.69	1.09	1.26
CYP76C2	-1.38	-1.27	-1.23	9.62	9.98	1.35	1.62
GPPS	1.04	1.13	1.53	1.66	1.14	-1.09	1.20
BIN2	-1.09	-1.74	-1.59	2.03	3.04	-1.05	2.08
BR6OX1	-1.58	-1.55	-1.52	3.19	3.02	1.12	1.90
CBB2	-1.10	-1.99	-1.36	2.50	1.80	1.01	1.08
GGR	-3.01	-1.03	-2.35	3.43	2.27	1.73	1.02
BAN	-2.14	-3.19	-2.07	4.04	5.09	1.86	1.55
DFR	-2.02	-8.14	-5.09	4.24	3.36	1.69	1.99
myb96	1.94	5.66	1.11	-2.20	-1.32	-1.38	-1.16

Selected genes showed reciprocal expression pattern, i.e. down-regulated in tt8, tt2 and ttg1 mutant lines while up-regulated in dex-treated TT8:GR and TTG1:GR transgenic lines in two tissues, 6 day old seedlings and siliqs. Complex list of genes differentially expressed among transgenic lines is shown in Appendix II. Differential genes were filtered based on p-value≤0.05 w.r.t respective controls.

A


Relative RNA levels w.r.t respective controls
16
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
tt8
tt8:gr
ttg1
ttg1:gr
GL2
TT10
BAN
DFR
HMG2
ISPD
GPPS
BR6OX1
CBB2
GGR

B


Relative RNA levels w.r.t respective controls
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
tt8
tt8:gr
ttg1
ttg1:gr
PR5
PR1
LURP1
DOGT1
PCC1
PDR12
SAG13
CYP76C2
WAK1

C

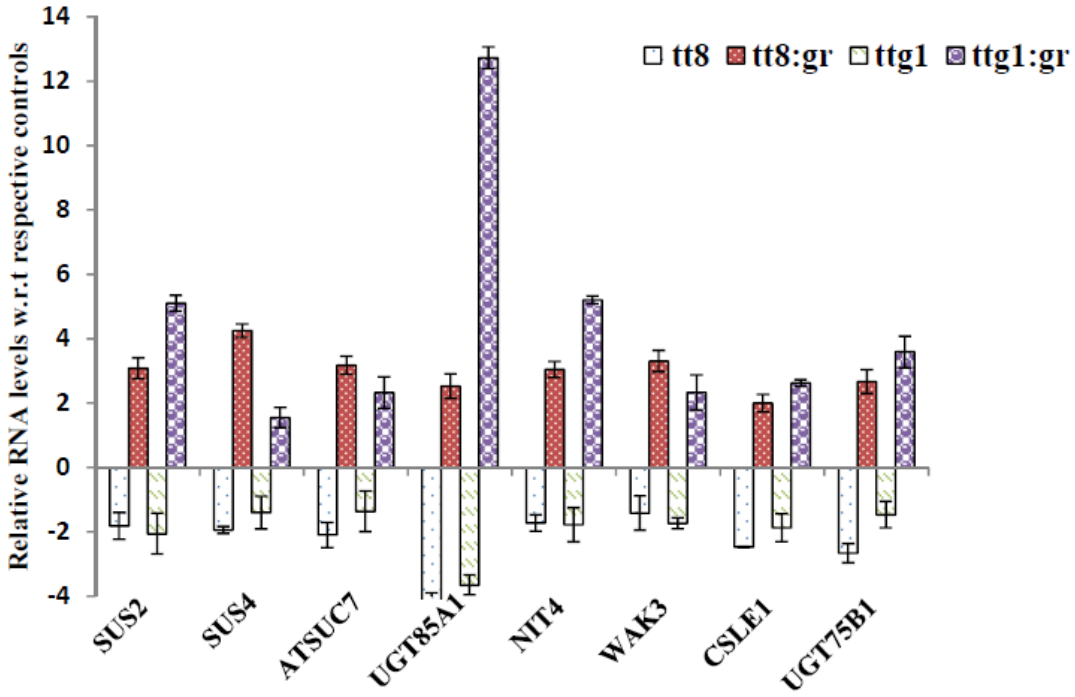
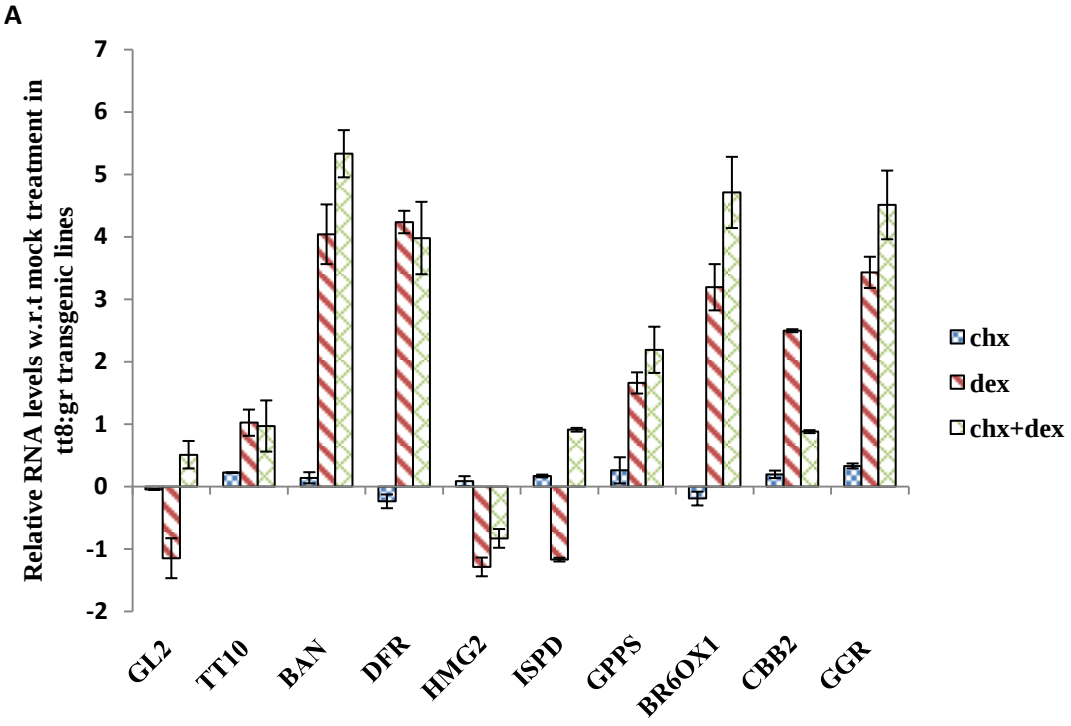
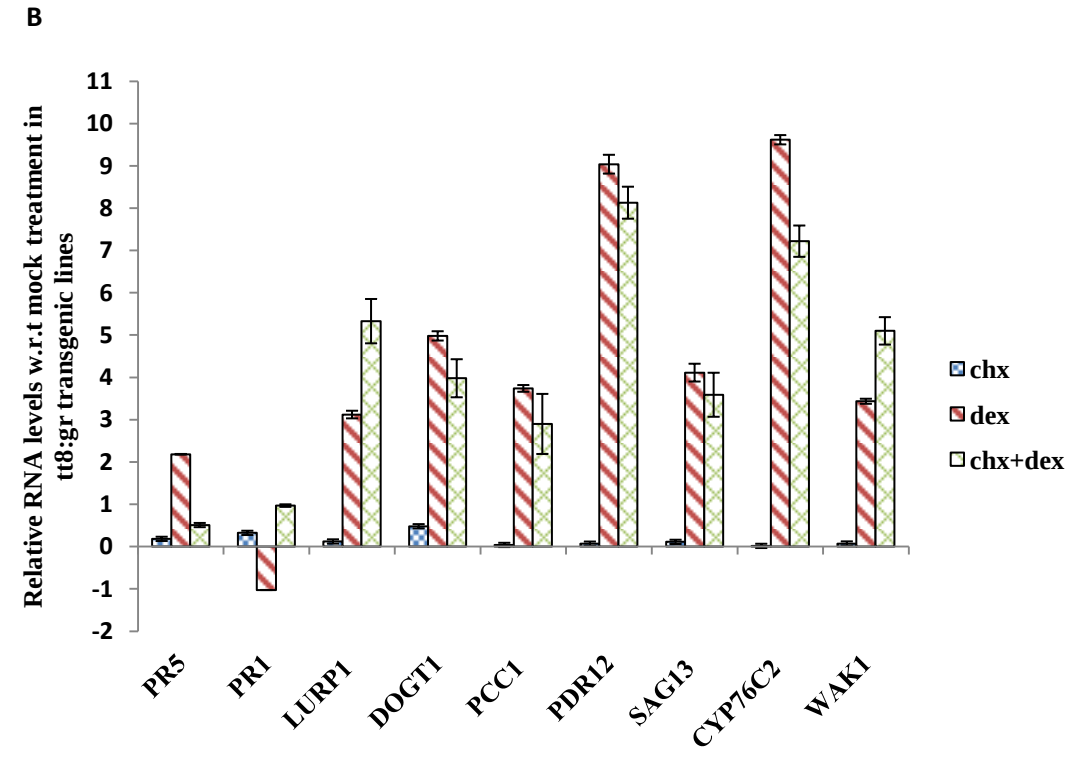
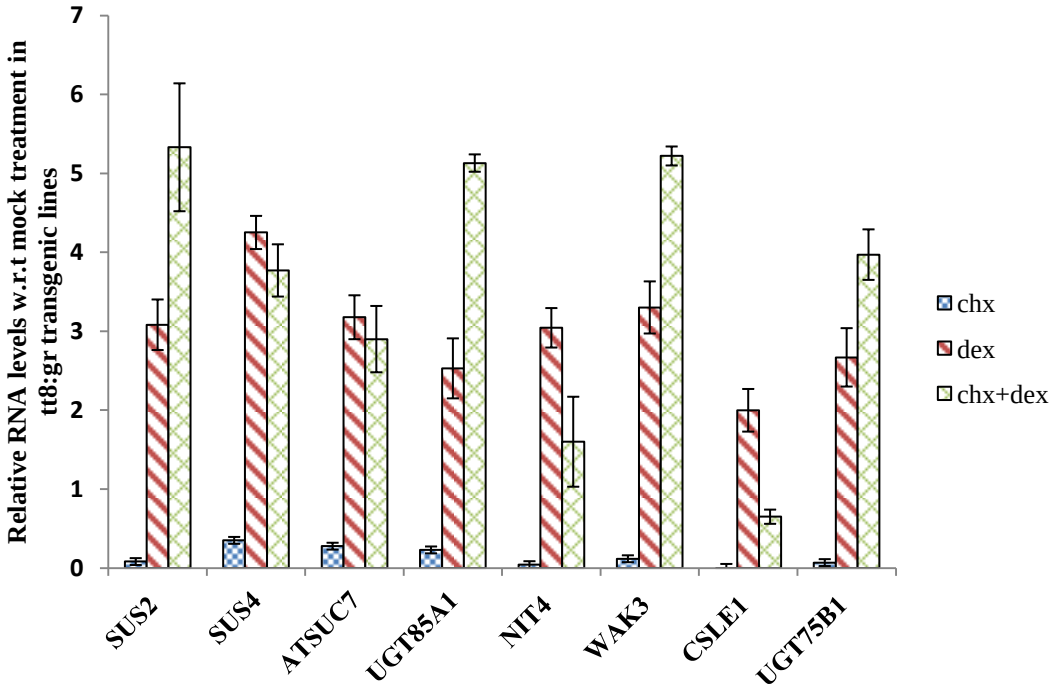


Fig 4.41 Quantitative real time PCR for 27 selected genes to confirm reciprocal expression pattern in tt8, ttg1 mutant lines and dex-treated TTG1:GR and TT8:GR transgenic lines w.r.t respective controls. A, represents genes from phenylpropanoid and isoprenoid metabolic network; B, represents genes related to stress response and C, represents genes from other metabolic pathways such as transferases or sugar transporters. These genes were tested for RNA levels in tt8 and ttg1 mutant lines and dex-treated TTG1:GR and TT8:GR w.r.t their controls. C_t values were normalized against Tubulin and fold change was calculated using $\Delta\Delta C_t$ method. Data are mean ± SD obtained from biological triplicates.

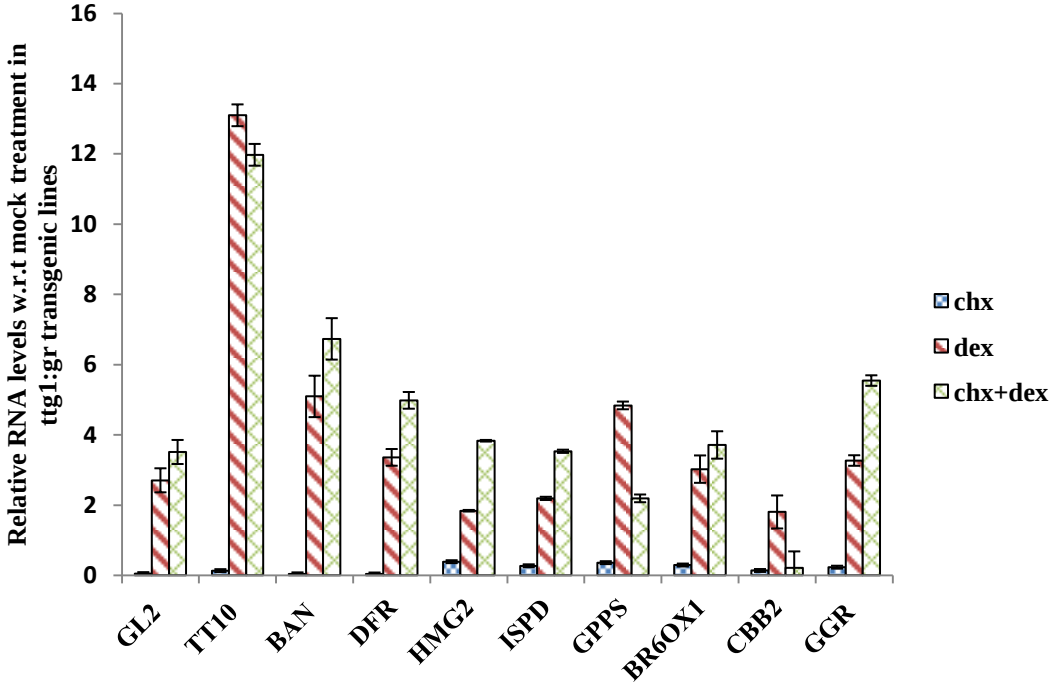

A
Relative RNA levels w.r.t mock treatment in tt8:gr transgenic lines
7
6
5
4
3
2
1
0
-1
-2
GL2
TT10
BAN
DFR
HMG2
ISPD
GPPS
BR6OX1
CBB2
GGR
chx
dex
chx+dex

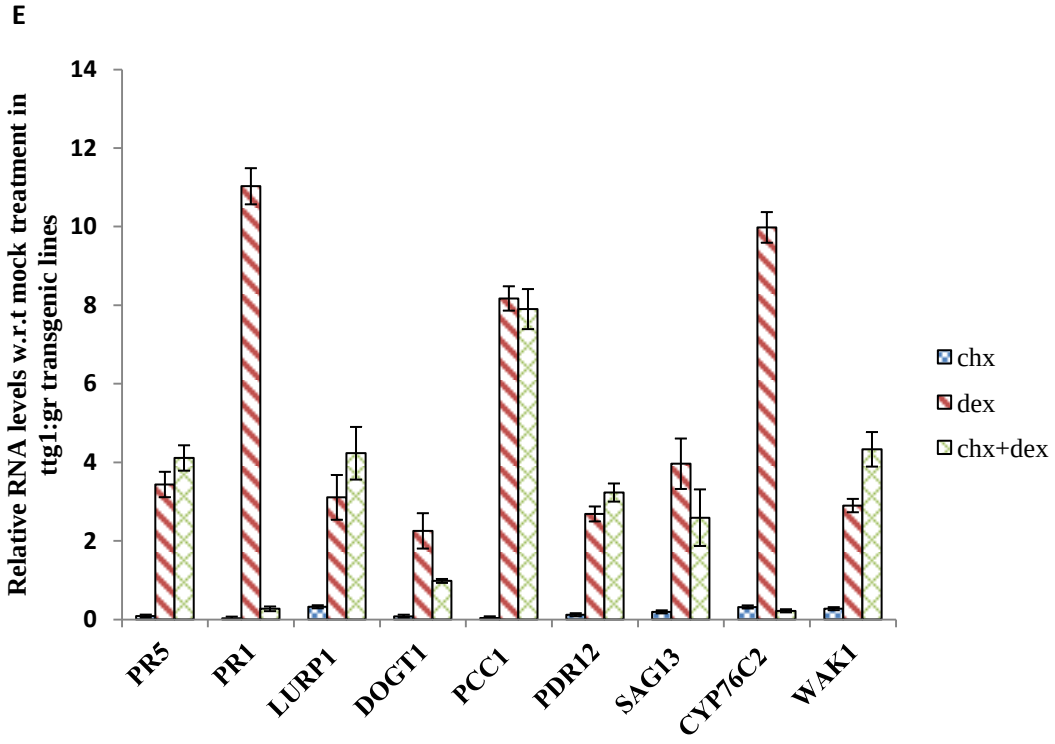

B
Relative RNA levels w.r.t mock treatment in tt8:gr transgenic lines
11
10
9
8
7
6
5
4
3
2
1
0
-1
-2
PR5
PR1
LURP1
DOGT1
PCC1
PDR12
SAG13
CYP76C2
WAK1
chx
dex
chx+dex

C


Relative RNA levels w.r.t mock treatment in tt8:gr transgenic lines
0
1
2
3
4
5
6
7
SUS2
SUS4
ATSUC7
UGT85A1
NIT4
WAK3
CSLE1
UGT75B1
chx
dex
chx+dex

D


Relative RNA levels w.r.t mock treatment in ttg1:gr transgenic lines
0
2
4
6
8
10
12
14
16
GL2
TT10
BAN
DFR
HMG2
ISPD
GPPS
BR6OX1
CBB2
GGR
chx
dex
chx+dex


E
Relative RNA levels w.r.t mock treatment in ttg1:gr transgenic lines
14
12
10
8
6
4
2
0
PR5
PR1
LURP1
DOGT1
PCC1
PDR12
SAG13
CYP76C2
WAK1
chx
dex
chx+dex

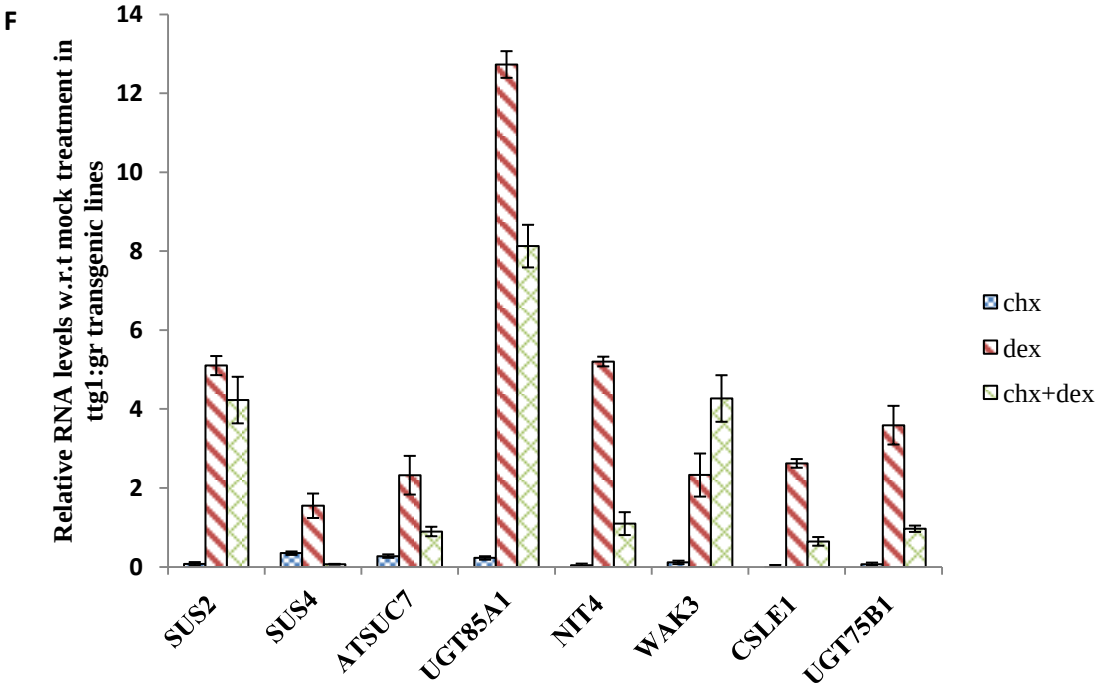
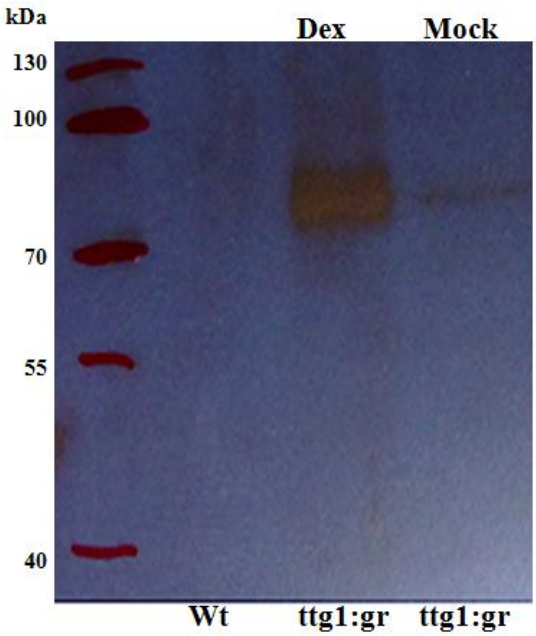
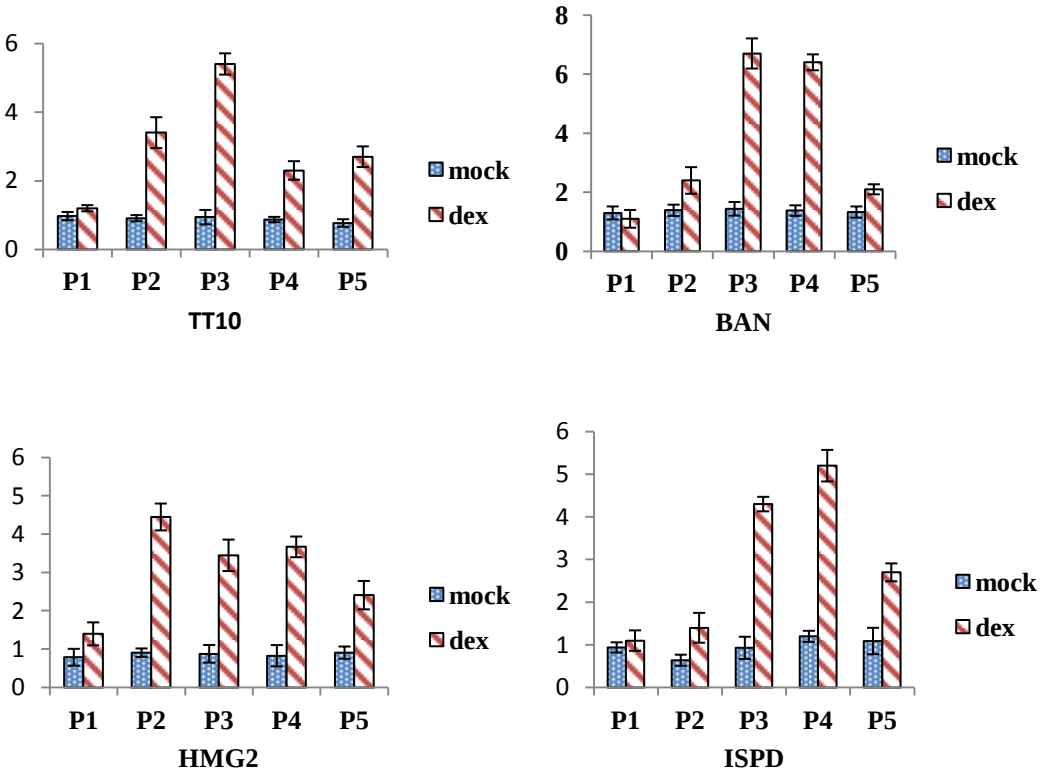


Fig 4.42 Quantitative real time PCR for 27 selected genes tested in 6 day old seedlings of TTG1:GR and TT8:GR transgenic lines with dex, dex+chx, chx and mock treatment to identify direct targets of TTG1 and TT8. A,D, represents genes from phenylpropanoid and isoprenoid metabolic network; B,E, represents genes related to stress response; C,F, represents genes from other metabolic pathways such as transferases or sugar transporters. A, B and C represents RNA levels of selected genes in TT8:GR transgenic lines tested for dex, dex+chx, chx and mock treatment. D, E and F represents RNA levels of selected genes in TTG1:GR transgenic lines tested for dex, dex+chx, chx and mock treatment. These genes were tested for RNA levels in mock treated TT8:GR and TTG1:GR transgenic lines. C_t values were normalized against Tubulin and fold change was calculated using $\Delta\Delta C_t$ method. Data are mean ± SD obtained from biological triplicates.

A


kDa
Dex
Mock
130
100
70
55
40
Wt
ttg1:gr
ttg1:gr

B


TT10
BAN
HMG2
ISPD
P1
P2
P3
P4
P5
mock
dex

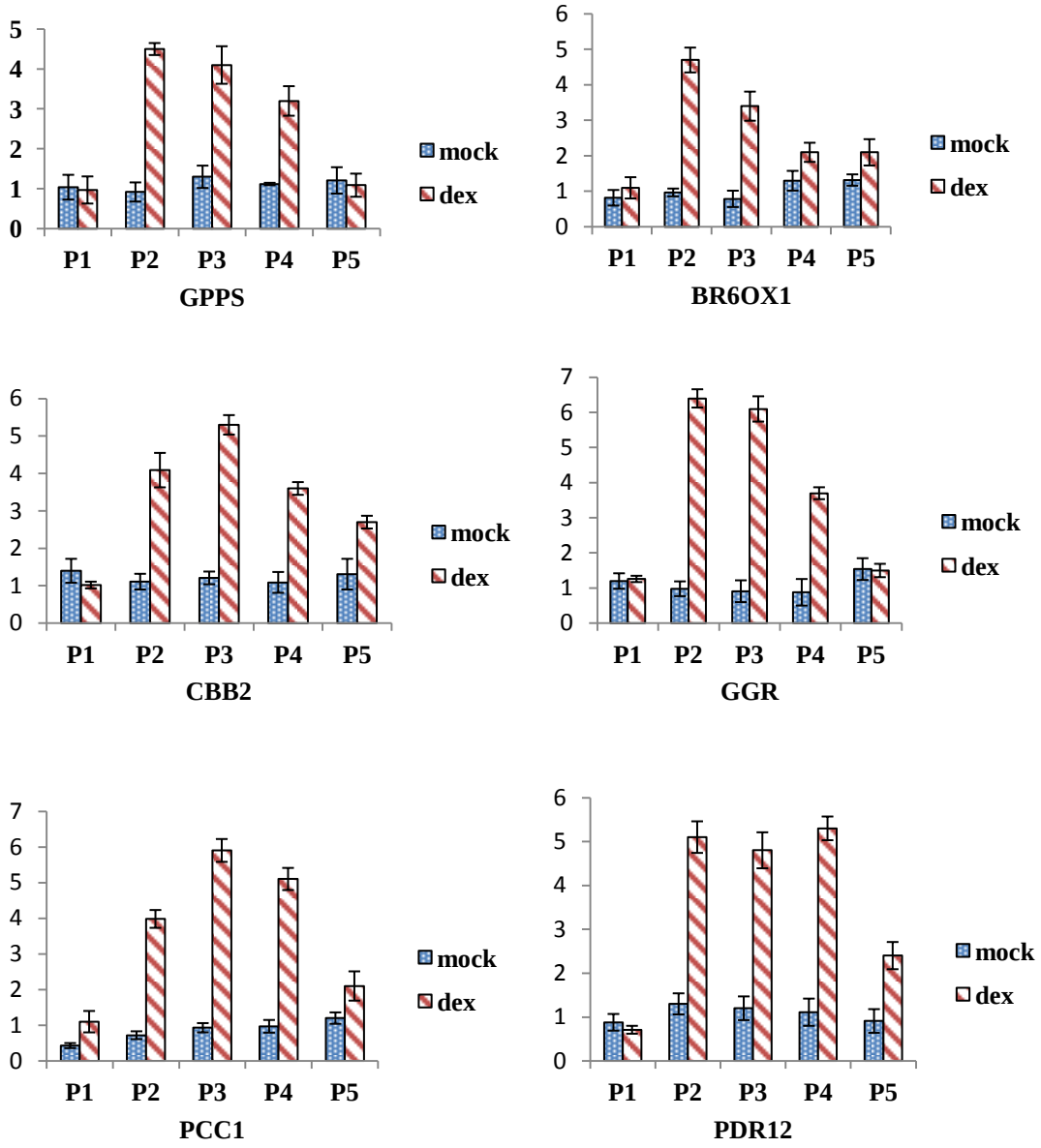


Fig 4.43 Protein gel blot analysis and ChIP-quantitative real time PCR assay for 10 selected genes from isoprenoid, phenylpropanoid metabolic pathways and stress response process. A, Western blotting showed specificity of used antibodies and that only dex-treated transgenic lines showed tagged protein in the nucleus. B, ChIP data showed enrichment of promoters of 10 selected genes showing them to be direct targets of TTG1. 5 set of primers were designed to cover -1 to -1000bps (P1-P5, where P1 is near coding region) of promoter region of all selected genes and were used to perform quantitative real time PCR using ChIP samples as template. Y-axis represents normalized C_t value of promoter region enriched in ChIP samples while X-axis represents promoter's fragments enriched by 5 primer set designed for each gene promoter (P1-P5). C_t values of Actin were used to normalize data. Data are mean ± SD obtained from biological triplicates.

4.3.5 Discussion

Transparent testa glabra1 (*TTG1*) encodes a protein containing WD40 repeats (Walker et al., 1999), a domain to promote interaction with other proteins and acts as a transcriptional factor, TTG1 regulates several developmental and biochemical pathways in *Arabidopsis*. This regulator does not bind directly to chromatin but helps in stabilizing complex that binds to chromatin to regulate gene expression. TTG1 has been shown to be expressed in all developmental stages in *Arabidopsis*, corroborating its role in more than one biological process. It has been shown to participate in regulation of processes such as trichome (formation of hairs on leaves, stems, and roots) and the production of seed mucilage and anthocyanin pigments (Walker et al., 1999; Payne et al., 2000; Gonzalez et al., 2008; Zhao et al., 2008). Several studies have shown TTG1 along with GL1, EGL1, GL3, CPC and TRY physically interact and form a complex to initiate and regulate trichome formation (Zhang et al., 2003; Cheng, 2009). TTG1 is also involved in the regulation of flavonoid production. In *ttg1* mutants, anthocyanins are absent (Shirley et al., 1995) and so are the dense brown tannin produced by the inner layer of the integument of the seed coat (Debeaujon et al., 2000). As a consequence, the testa of seeds appears transparent in ttg1 mutants. The TTG1 protein forms a complex with bHLH and MYB proteins encoded by TT8 (Nesi et al., 2000) and TT2 (Nesi et al., 2001) to regulate expression of the genes encoding enzymes BAN and DFR.

TTG1 is distantly related to the β-subunits of heterotrimeric G proteins, which suggests role for *TTG1* in signal transduction to multiple downstream transcription factors (Walker et al., 1999). Indeed, metabolic profiling results for TTG1 loss-of-function mutants and its double mutants with tt3, tt4, tt5 and tt6 mutant lines showed

TTG1 loss as an important factor resulting in isoprenoid metabolic network perturbation along with phenylpropanoid metabolic network, suggesting TTG1 to have more targets than the two that have been described till date. Similarly, metabolic profiling for tt8 mutant line showed defected glycosylation, a biochemical phenotype which was specific to loss-of function of TT8, suggesting more targets of TT8. In order to identify direct targets of TTG1 and TT8, dex-inducible TTG1 and TT8 over-expression transgenic lines and estrogen inducible TTG1 and TT8 silencing transgenic lines were generated using amiRNA approach and systems biology approach were used to characterize the responses. Due to time constraint, we were not able to use inducible transgenic lines to test reciprocal expression pattern and at metabolic profiling, as these transgenic lines are at T2 generation. These lines will be used in future to further validate the results.

Induced activity of TTG1 in transgenic lines showed ectopic trichome formation on petioles and significant increase in number of adaxial trichomes as compared to that of mock treatment, which was expected since TTG1 initiates trichome development. Interestingly, induced activity of TT8 in transgenic lines also showed ectopic trichome formation with increased trichome number, although the effect was less pronounced than from TTG1 induction (Fig 4.23-25). Previous studies on trichome formation have reported formation of a ternary complex between myb-bHLH-ttg1transcriptional factors. GL2 as myb factor and GL3, EGL3 as bHLH factor were shown as members of ternary complex with TTG1 (Qi et al., 2011). TT8, a bHLH transcriptional factor has been shown to work as EGL3 and GL3, although effect of TT8 on trichome formation has not been described yet (Zhang et a., 2003; Maes et al., 2008). Previous studies reported that the triple mutant of tt8/gl3/egl3 has a similar

trichome-less phenotype as that of *ttg1* loss-of-function mutant line, while *tt8* was shown to have minimal or no effect on trichome number as compared to that of gl3 and egl3 mutant lines, which suggested TT8 to be less important for trichome formation (Zhang et al., 2003). Our results showed that induced activation of TT8 alone was sufficient for increased trichome formation, although the effect was less pronounced than that observed by TTG1 activation. This result further validates that GL3, EGL3 and TT8 have functional redundancies because increased activity of TT8 is sufficient to increase trichome formation. It would be interesting to test if such activation for GL3 or EGL3 also results in ectopic trichome formation and if its effect is stronger than that of TT8.

Differential metabolites from induced verses non-induced levels of TTG1 and TT8 in 6-day-old seedlings of transgenic lines showed high similarity between the two lines. Of the 368 metabolites mapped to metabolic pathways, 194 metabolites showed similar perturbational trends in both sets of induced transgenic lines with 139 metabolites up-regulated and 55 being down-regulated. As TTG1 and TT8 share common targets for regulation in a single complex, such high similarity of common metabolites was indeed expected. Several metabolites from phenylpropanoid metabolic network were up-regulated by induction of the two regulators. Metabolites such as dihydrokaempferol, quercetin, glucoside derivatives of cyanidin, pelarogonidin, quercetin and metabolites intermediates from lignin branch of phenylpropanoid metabolic network were up-regulated in these lines. Hence known pathways of phenylpropanoid metabolic network were perturbed, thus validating the experimental set-up.

Interestingly, none of the glucoside or any other sugar conjugated form of kaempferol were differentially expressed in both TTG1 and TT8 induced lines. Metabolic profiling of tt8 and ttg1 mutant lines showed several sugar conjugated form of kaempferol affected specifically in tt8 mutant line. Sugar conjugates of quercetin, cyanidin and pelarogonidin showed increased level in both TTG1 and TT8 induced lines as compared to that of mock treated transgenic lines, which were down-regulated in tt8 mutant line. These observations thus suggest role of TT8 in regulation of glycosylation process involving quercetin, cyanidin and pelarogonidin as substrate, while having minimal or no effect on kaempferol. Methylated and sulphated forms of quercetin were up-regulated by TTG1 and TT8 up-regulation suggesting members of these processes as potential targets regulated by the two transcriptional factors. Quercetin and kaempferol compete for conjugation with sugar molecules (Burbulis et al., 1999), which appears to be favoured to quercetin in TT8-induced transgenic lines. Glycosylation of metabolites were also affected by TTG1 induction, but only few metabolites were up-regulated compared to those upon TT8 induction. Since glycosylation occurs through activity of UGTs, which shows specificity both the aglycone substrate to be glycosylated as well as type of sugar molecule conjugated, one of the possible explanation for this specificity could be TT8 acting as regulator of glucoside transferases, which act specifically on quercetin aglycone. Reverse genetics using the advantage of a huge amount of microarray data publically available through sources such as Genevestigator and TAIR could be used to identify if any quercetin-specific glucoside transferases have correlation of expression with TT8. We are intending to perform ChIP-seq for samples from TT8-induced transgenic lines and compare with non-induced forms to further identify carbohydrate active genes that might be involved in glycosylation process.

Metabolic perturbation in TTG1 and TT8 induced lines were not limited to phenylpropanoid metabolic network, but also included novel targets of isoprenoid and its branch pathways. Both MVP and MEP pathways that synthesize, precursors of isoprenoids (IPP and DMAPP) were up-regulated in TTG1 induced lines while down-regulated in ttg1 mutant line. Branch pathways of isoprenoid metabolic network such as trans-zeatin biosynthesis, and biosynthetic pathways of brassinosteroids , carotenoids , abscisic acid, tocopherols , cytokinins , gibberellin (non C-3, non C-13 hydroxylation), and chlorophyll a along with other metabolic pathways were up-regulated upon induction of TTG1 while down-regulated in ttg1 mutant line (Fig 4.35). Isoprenoids are linked very tightly and are formed by several condensation reactions of a single precursor (Vranova´ et al., 2012). Further studies on MVP and MEP pathways have shown to have a crosstalk between these two cytosolic and plastid localized pathways (Laule et al., 2003). Interestingly, some of these pathways occur in different compartments. MVP pathways, being cytosolic, could be explained by possibilities that expressions of key structural genes from MVP pathways are regulated by TTG1. MEP pathways being located in plastids and up-regulated by TTG1 could be the result of crosstalk with MVP as described by Laule et al. (2003). Another possibility could be TTG1 directly or indirectly regulating genes from plastids. It is not clear if TTG1 may enter plastids to regulate their structural genes, or it can influence gene expression in plastids by any indirect mechanism. Experiments can be designed to track intracellular localization of TTG1 to answer parts of this question. Detailed knowledge of chloroplast gene expression and the nucleus-encoded proteins that influence it are prerequisites for understanding nuclear-organellar cross talk and chloroplast evolution, and will aid in optimizing transgene expression in the

plastid compartment. TT8 induction also showed increased accumulation of metabolites from isoprenoid metabolic network, but its effect was not as pronounced as that from TTG1 induction (Fig 4.34). For instance, IPP and DMAPP, precursor of isoprenoid metabolic network were down-regulated upon TT8 induction, while metabolites from mevalonate and methylerythritol phosphate pathway showed increased accumulation. Further, three out of nine metabolites from methylerythritol phosphate pathway were up-regulated compared to six metabolites being up-regulated in dex-treated TTG1:GR transgenic lines. These results shows TTG1 induced transgenic lines to have higher levels of isoprenoid compared to that of TT8 induced transgenic lines.

Metabolic profiling of gain- and loss-of-function transgenic lines of TTG1 and TT8 revealed several important metabolic pathways being affected by these transcriptional factors. TTG1 and TT8 being the two transcriptional factors, regulates metabolic pathways by exerting a control over expression levels of biosynthetic genes associated with these metabolic pathways. To identify target genes of TT8 and TTG1, microarray-based gene expression analysis of dex and mock treated TTG1:GR, TT8:GR and loss-of-function TT2, TT8 and TTG1 mutant lines were performed. Differential genes for tt8, ttg1 and tt2 showed high similarity with similar expression pattern for 273 genes out of 332 common genes. Since complex formed by TT2, TT8 and TTG1 have been shown to regulate expression of BAN and DFR (Baudry et al., 2004, 2006), some level of similarity of expression pattern was expected. The *tt2* mutant line showed fewer differentially expressed genes as compared to *tt8* and *ttg1* mutant lines. This could be possibly due to TT2 being involved in regulation of few processes and its specific interaction with TT8 and TTG1. On the other hand, TT8 and

TTG1 has been shown to play a major role in regulation of several developmental and metabolic pathways, thus, their loss-of-function mutant lines showing higher number of differentially expressed genes was expected. Gene expression profile of loss-of-function of TTG1, TT8 and their induced forms in transgenic lines revealed similar expression pattern of several genes. GO analysis of loss-of-function of TT2 and TTG1 and induced forms of TTG1 and TT8 in transgenic lines showed enrichment of several features in these transgenic lines (Fig 4.39). TTG1 loss-of-function mutant line and dex-treated TTG1:GR showed enrichment of several developmental processes. This observation was in accordance with earlier studies, which reported TTG1 to be involved in several developmental processes such as trichome development, root hair development and elongation, seed coat formation and several others. The tt8 mutant line and its induced over-expression lines showed enrichment of hydrolases and transferases process, some of which have been reported previously (Zhang et al., 2003; Zhang et al., 2007; Qi et al., 2011). Interestingly, TT8 and TTG1 loss-of-function and their induced activity in transgenic lines also showed enrichment of biotic and abiotic stress response processes. GO enrichment analysis thus implicated role of TT8 and TTG1 to be involved in metabolic processes and extending to other processes related to biotic and abiotic responses. Isoprenoid and phenylpropanoid metabolic networks contribute to those metabolites, which play a major role in conferring biotic and abiotic stress responses. For instance, flavonoids and camalexins from phenylpropanoid metabolic network participate in UV damage and pathogenic infection responses in plants. Thus, enrichment of biotic and abiotic stress response in dex-treated TTG1:GR and TT8:GR indirectly relates TTG1 and TT8 to stress responsive metabolic pathways including phenylpropanoid and isoprenoid metabolic network.

We then focused on those genes that showed reciprocal expression patterns between loss-of-function and dex induced transgenic lines, and gave more weightage for genes showing similar expression pattern for TT8 and TTG1 loss or gain-of-function transgenic lines. Comparison of gene expression profiles of tt8, ttg1, dex-treated TTG1:GR and TT8:GR helped in identifying 185 genes showing reciprocal behavior pattern along with several other genes that were up-regulated in dex-treated TTG1:GR and TT8:GR transgenic lines but were not differentially expressed in ttg1 and tt8 mutant lines. Metabolic pathwayassociated genes such as *PBS3, TT10, GL2, BAN, DFR, CBB2, BIN2, HMG2, BR6OX1, GGR, GPPS, ISPD, UGT85A1* and *UGT75B1* showed reciprocal expression pattern. BAN and DFR showing reciprocal expression patterns was expected since these genes are reported to be direct targets of TT8 and TTG1. Genes such as *CBB2, BIN2, BR6OX1, HMG2, GGR, GGPS* and *ISPD* associated with isoprenoid metabolic network were up-regulated in dex-treated TTG1:GR transgenic lines while were down-regulated in ttg1 mutant line. Apart from metabolic genes, other genes related to biotic and abiotic stress responses such as *PR1, PR5, PCC1, PBS3, WAK1* and *PDR12* also showed reciprocal behavior for TTG1. Genes such as *SUC7, SUS2, SUS4, CSLE1* serve as sugar transporter and transferases and were seen to be up-regulated in TTG1 or TT8 induced transgenic lines and down-regulated in ttg1 and tt8 mutant lines. Functions of all these genes have been described previously using mutant associated phenotypes (Sauer et al., 2004), but their regulation is still not well understood.

From the list of genes showing reciprocal behavior, twenty seven genes were selected and further tested for gene expression in lines carrying loss- or induction of TTG1 and TT8. Quantitative real-time PCR data for these genes showed similar expression

pattern as that was observed by microarray gene expression. Cycloheximide (chx) treatment combined with TTG1 and TT8 induction helped us to identify genes affected at transcriptional level. Genes such as *GL2, TT10, BAN, DFR, HMG2, ISPD, GPPS, BR6OX1, GG2, PR5, LURP1, PCC1, PDR12, SAG13, WAK1, UGT85A1* and *SUS2* are likely to be direct targets of TTG1 (Fig 4.40). For TT8, genes such as *SUS4, SUC7, UGT75B1* and *CSLE1* were observed as direct targets, while these genes were not detected in TTG1 set of transgenic lines suggesting them to be indirect targets of TTG1. These observations revealed that apart from *BAN, DFR, UGT85A1, PCC1, PDR12,* and *WAK1*, which seemed to be direct targets of both TTG1 and TT8, there were several other genes that were direct target for one but not for the other. Although TTG1 has been shown to interact with other transcriptional factors such as GL3 and EGL3, TT8 has only been shown to interact with TTG1 and TT2 to regulate the expression of target genes. Our data suggests possibility of additional partners of TT8 other than TTG1 for regulation of sugar transporters and transferases, which are direct targets of TT8. Hence, their possible regulation by TTG1 and TT8 opens new research avenues to pursue.

From the list of genes shown to be the direct targets of TTG1, ten genes were selected to perform the ChIP experiment in order to confirm whether TTG1 directly binds to promoter region of these selected genes. TTG1 has been reported to bind to the promoter of BAN; hence BAN was used as a control. ChIP-RT-PCR results showed the enrichment of promoters of all ten genes, namely *TT10, HMG2, GGR, BAN, PCC1, BR6OX1, GPPS, ISPD, PDR12* and *CBB2* (Fig 4.41). Among these, *GGR, HMG2, GPPS, ISPD* and *BR6OX1*, are all associated with MVP pathway from isoprenoid metabolic network, while *BAN* and *TT10* are from phenylpropanoid

metabolic network. We intentionally did not select genes specific to MEP pathways, since only DXR from MEP pathways showed reciprocal expression patterns in TTG1 and TT8, and we do not have any experimental proof if TTG1 could be a direct target for genes from MEP pathways. Thus, using systems biology approach, this study has identified nine new targets of TTG1. Our study confirmed the involvement of TTG1, having four WD40 repeat domains, in regulation of several processes. ChIP-Seq of ChIP samples of TTG1:GR would play an important role in identification of different motifs and target genes to further improve our knowledge of TTG1-mediated regulation of different processes.

These results taken together confirmed TTG1 as a common regulator for genes from both isoprenoid and phenylpropanoid metabolic networks, and explain the crosstalk observed between these two secondary metabolism pathways (Fig 4.42). ChIP assay also confirmed PCC1 and PDR12, of biotic and abiotic stress response pathways to be regulated by TTG1. This is an interesting finding since there is no systems level study that reports direct regulation of abiotic and biotic responsive genes by a known regulator of secondary metabolism pathway (such as TTG1 in this case). Apart from these two, there were other genes such as *PR1, PR5, WAK1, WAK3* and *SAG13*, which also showed reciprocal gene expression behavior, and could be direct targets of TTG1. This study has thus identified a common regulatory mechanism for stress-response pathways and distant secondary metabolic networks. Therefore, a longstanding question of how secondary metabolism network is triggered during stress response seems to have been partially answered. Further studies involving this mechanism will be useful in better understanding of how plants coordinate their large-scale responses to various biotic and abiotic stresses.

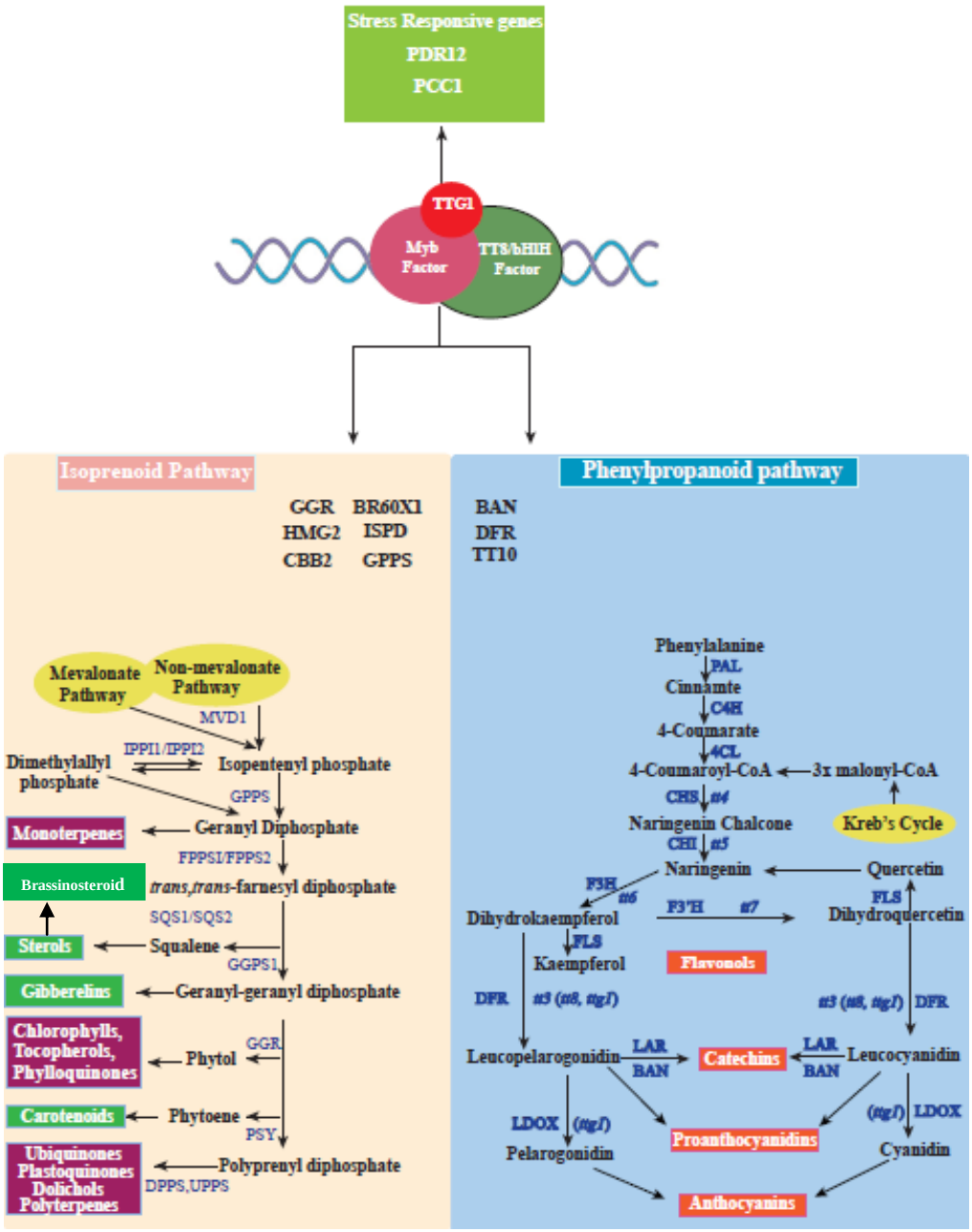


Fig 4.44 Cross-talk between isoprenoid and phenylpropanoid metabolic network mediated by TTG1 through transcriptionally regulating biosynthetic genes from these pathways. Genes shown connected to TTG1 are those which were confirmed through ChIP as directly associated with TTG1. Enzymes full name described here has been mentioned in Fig 4.1, 4.32 and 4.33.

CHAPTER 5

CONCLUSIONS, INFERENCES AND FUTURE DIRECTIONS

5.1 Conclusions and inferences

In this chapter the main conclusions derived from this study are presented, as well as the possible future research which should contribute to elucidation of some of the issues which were not covered here.

Major conclusions from this research are summarized as follows.

- Metabolic profiling of *QuoA* transgenic lines and comparative analysis with *tt6* mutant lines showed that lignification of transgenic lines resulted at the expense of phenolamide metabolic pathway. Further study on association of phenolamide and lignin pathway can give us an important tool to manipulate lignin metabolic pathway.
- Gene-metabolite association in phenylpropanoid metabolic network in QuoA transgenic lines showed metabolic network to be tightly regulated at the branch points. All five branch points of phenylpropanoid metabolic network showed increased levels of metabolites and down-regulation of associated genes, suggesting possibility of feed-back regulation at all the branch points.
- Metabolic profiling of dex-treated TTG1:GR showed increased levels of metabolite intermediates from phenylpropanoid and isoprenoid metabolic network, thus showing for the first time direct association of two secondary metabolic pathways via a common transcriptional regulator.

- Metabolic profiling of *tt8* mutant line showed specificity to glycosylation of quercetin, kaempferol, cyanidin and pelarogonidin with quercetin and cyanidin showing reciprocal behavior as compared to that of kaempferol and pelarogonidin. On the other hand, dex-treated TT8:GR showed increase in the glycosylation of several metabolites such as quercetin, cyanidin, pelarogonidin, zeatin, cytokinins and others, although no effect was observed for kaempferol and its glycosylated metabolites.
- Several genes from different metabolic pathways such as *TT10*, *GL2*, *BAN*, *DFR* from phenylpropanoid pathway and *GGR*, *HMG2*, *ISPD*, *GPPS*, *CBB2*, *BIN2*, *BR6OX1* from isoprenoid metabolic network were up-regulated in dex treated TTG1:GR and TT8:GR transgenic lines while down-regulated in *ttg1* and *tt8* mutant lines. Dex-treated TT8:GR transgenic lines also showed up-regulation of several transferases while hydrolases were down-regulated.
- Several genes related to pathogen resistance and response process such as *PCC1*, *PR1*, *PR5*, *WAK3*, *PAD3* were up-regulated in dex treated TTG1:GR and TT8:GR transgenic lines while down-regulated in their mutant lines.
- ChIP analysis for TTG1:GR transgenic lines using antiGR antibody and quantitative RT-PCR based approach showed secondary metabolic genes and stress responsive genes as direct targets of TTG1. This is the first report showing common regulatory mechanism for stress-response pathways and distant secondary metabolic networks. Therefore, a longstanding question of how secondary network triggered during stress response seems to have been partially answered.

5.2 Future directions

Some specific lines of future investigations are listed below

- *QuoA* activity resulted simultaneous accumulation of lignin and anthocyanin. It would be interesting to test this system in model crop species, such as Rice to study if perturbation resulting from *QuoA* leads to similar phenotype. Increased anthocyanin and lignification are very useful traits and hence will improve rice variety.
- The mechanisms to regulate isoprenoid pathways described in this study can be used to develop synthetic compounds in yeast or other heterologous hosts.
- The stress responsive of transgenic lines with altered expression of *TTG1* or *TT8* can be studied. Outcomes of these investigation will have develop improve stress responsiveness of plants.
- ChIP-Seq could be performed for TTG1 and TT8 to identify the type of domains and target genes that are being regulated by these transcriptional factors. Identifying direct targets of these two regulators will help in modifying or enhancing stress responsiveness in plant.

Taken together, these future directions will help not only in improving our understanding of relationship between stress response and secondary metabolism in plants, but it will also be useful in improving plant performance and in biotech application in synthetic biology.

REFERENCES

Abian J. The coupling of gas and liquid chromatography with mass spectrometry. Journal of mass spectrometry 1999, 34:157-168.

Allen E, Moing A, Ebbels TM, Maucourt M, Tomos AD, Rolin D, Hooks MA. Correlation network analysis reveals a sequential reorganization of metabolic and transcriptional states during germination and gene-metabolite relationships in developing seedlings of *Arabidopsis*. BMC Systems Biology 2010, 4:62.

Allen J, Bisbee PA, Darnell RL, Kuang A, Levine LH, Musgrave ME, van Loon JJWA. Gravity control of growth in *Brassica rapa* and *Arabidopsis thaliana* (Brassicaceae): Consequences for secondary metabolism. American Journal of Botany 2009, 96: 652–660.

Almaas E, Oltvai ZN, Barabasi A-L. The Activity reaction core and plasticity of metabolic networks. PLoS Computational Biology 2005, 1:e68.

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, et al. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 2003, 301:653-657.

Andersen SU , Cvitanich C , Hougaard BK , Roussis A , Gronlund M , Jensen DB , Frokjaer LA , Jensen EO. The glucocorticoid-inducible GVG system causes severe growth defects in both root and shoot of the model legume *Lotus japonicus*. Molecular Plant-Microbe Interactions 2003, 16:1069-1076.

Anthony JR, Anthony LC, F Nowroozi, Kwon G, Newman JD, Keasling JD. Optimization of the mevalonate-based isoprenoid bio-synthetic pathway in *Escherichia coli* for production of the anti-malarial drug precursor amorpha-4,11-diene. Metabolic Engineering 2009, 11:13–19.

Arsenault PR, Vail D, Wobbe KK, Erickson K, Weathers PJ. Reproductive development modulates gene expression and metabolite levels with possible feedback inhibition of artemisinin in *Artemisia annua*. Plant Physiology 2010, 154:958-968.

Baillie GS, Scott JD, Houslay MD. Compartmentalisation of phosphodiesterases and protein kinase A: opposites attract. FEBS Letters 2005, 579:3264–3270.

Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nature Reviews Genetics 2004, 5:101-13.

Bazurto JV, Downs DM. Plasticity in the purine–thiamine metabolic network of *Salmonella*. Genetics 2011, 187:623-631.

Bedair M, Sumner LW. Current and emerging mass-spectrometry technologies for metabolomics. Trends in Analytical Chemistry 2008, 27:238–250.

Begley P, Francis-McIntyre S, Dunn WB, Broadhurst DI, Halsall A, Tseng A, Knowles J, Consortium H, Goodacre R and Kell DB. Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale nontargeted metabolomic studies of human serum. Analytical Chemistry 2009, 81:7038–7046.

Bernards MA, Lewis NG. The macromolecular aromatic domain in suberized tissue: a changing paradigm. Phytochemistry 1998, 47:915–933.

Bernards MA, Lopez ML, Zajicek J, Lewis NG. Hydroxycinnamic acid-derived polymers constitute the polyaromatic domain of suberin. The Journal of Biological Chemistry 1995, 270:7382–7386.

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M. Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. The Plant Cell 2007, 19:148-162.

Bleeker PM, Diergaarde PJ, Ament K, Guerra J, Weidner M, Schütz S, de Both MTJ, Haring MA, Schuurink RC. The Role of Specific Tomato Volatiles in Tomato-Whitefly Interaction. Plant Physiology 2009, 151:923-935.

Bloemberg GV, O'Toole GA, Lugtenberg BJ, Kolter R. Green fluorescent protein as a marker for pseudomonas spp. Applied Environmental Microbiology 1997, 63: 4543-4551.

Blount JW, Korth KL, Masoud SA, Rasmussen S, Lamb C, Dixon RA. Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for

a feedback loop at the entry point into the phenylpropanoid pathway. Plant Physiology 2000, 122:107-116.

Bollina V, Kumaraswamy GK, Kushalappa AC, Choo TM, Dion Y, Rioux S, Faubert D, Hamzehzarghani H. Mass spectrometry based metabolomics application to identify quantitative resistance related metabolites in barley against Fusarium head blight. Molecular Plant Pathology 2010, 11:769–782.

Bordbar A, Jamshidi N, Palsson BO. iAB-RBC-283:A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states. BMC Systems Biology 2011, 5:110.

Bordel S, Agren R, Nielsen J. Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. plos computational biology 2010, 6:e1000859.

Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C. Activation tagging identifies a conserved myb regulator of phenylpropanoid biosynthesis. The Plant Cell 2000, 12:2383-2394.

Borghi L. Inducible gene expression systems for plants. Methods in Molecular Biology 2010, 655:65-75.

Broeckling CD, Huhmann DV, Farag MA, Smith JT, May GD, Mendes P, Dixon RA, Sumner LW. Metabolic profiling of medicago truncatula cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. Journal of Experimental Botany 2005, 56:323-336.

Brown PD, Tokuhisa JG, Reichelt M, Gershenzon J. Variation of glucosinolate accumulation among different organs and developmental stages of *Arabidopsis thaliana*. Phytochemistry 2003, 62:471-481.

Burbulis IE, Shirley BW. Interactions among enzymes of the *Arabidopsis* flavonoid biosynthetic pathway. Proceedings of the National Academy of Sciences of the United States of America 1999, 96:12929-12934.

Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A,

Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. Nucleic Acids Research 2012, 40 (Database issue), D742-D753.

Chemler J, Koffas M. Metabolic engineering for plant natural product biosynthesis in microbes. Current Opinion in Biotechnology 2008:597-605.

Clough SJ, Bent AF. Floral dip: A simplified method for agrobacterium-mediated transformation of *Arabidopsis thaliana*. The Plant Journal 1998, 16:735-743.

Coleman JOD, Randall R, Blake-Kalff MMA. Detoxification of xenobiotics in plant cells by glutathione conjugation and vacuolar compartmentalization: a fluorescent assay using monochlorobimane. Plant Cell and Environment 1997, 20:449–460.

Conrado RJ, Varner JD, DeLisa MP. Engineering the spatial organization of metabolic enzymes: Mimicking nature's synergy. Current Opinion in Biotechnology 2008, 19:492-499.

Cooke TJ, Poli DB, Sztein AE, Cohen JD. Evolutionary patterns in auxin action. Plant Molecular Biology 2002, 49:319-338.

Dalmolin RJ, Castro MA, Rybarczyk Filho JL, Souza LH, de Almeida RM, Moreira JC. Evolutionary plasticity determination by orthologous groups distribution. Biology Direct 2011, 6:22

Datta K, Baisakh N, Ganguly M, Krishnan S, Yamaguchi Shinozaki K, Datta SK. Overexpression of *Arabidopsis* and rice stress genes' inducible transcription factor confers drought and salinity tolerance to rice. Plant Biotechnolgy Journal 2012, 10:579-586.

D'Auria JC, Gershenzon J. The secondary metabolism of *Arabidopsis* thaliana: growing like a weed. Current Opinion in Plant Biology 2005, 8:308–316.

Desbrosses GG, Kopka J, Udvardi MK. *Lotus japonicus* Metabolic Profiling. Development of Gas Chromatography-Mass Spectrometry Resources for the Study of Plant-Microbe Interactions. Plant Physiology 2005, 137:1302-1318.

Dietmair S, Hodson MP, Quek L-E, Timmins NE, Gray P, Nielsen LK. A multi-omics analysis of recombinant protein production in Hek293 Cells. PLoS ONE 2012 7:e43394.

Dixon DP, Skipsey M, Edwards R. Roles for glutathione transferases in plant secondary metabolism. Phytochemistry 2010, 71:338-350.

Drubin DA, Way JC, Silver PA. Designing biological systems. Genes and Development 2007, 21:242–254.

Dunn WB, Bailey NJ, Johnson HE. Measuring the metabolome: current analytical technologies. Analyst 2005, 130:606–625.

Espinoza C, Degenkolbe T, Caldana C, Zuther E, Leisse A, Willmitzer L, Hincha DK, Hannah MA. Interaction with diurnal and circadian regulation results in dynamic metabolic and transcriptional changes during cold acclimation in *arabidopsis*. PLoS ONE 2010, 5:e14101.

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie AR, Galili G. *Arabidopsis* seed development and germination is associated with temporally distinct metabolic switches. Plant Physiology 2006, 142:839-854.

Faust K, Dupont P, Callut J, van Helden J. Pathway discovery in metabolic networks by subgraph extraction. Bioinformatics 2010, 26:1211-1218.

Feist AM, Palsson BO. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. Nature Biotechnology 2008, 26:659-667.

Fiehn O, Kloska S, Altmann T. Integrated studies on plant biology using multiparallel techniques. Current Opinion in Biotechnology 2001, 12:82–86.

Fluck RA, Leber PA, Lieser JD,Szczerbicki SK, Varnes JG, Vitale MA, Wolfe EE. Choline conjugates of auxins. I. Direct evidence for the hydrolysis of choline-auxin conjugates by pea cholinesterase. Plant Physiology and Biochemistry 2000, 38:301-308.

Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C. Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of *Arabidopsis*. The Plant Journal 2002, 30:47-59.

Gat-Viks I, Tanay A, Shamir R. Modeling and analysis of heterogeneous regulation in biological networks. Journal of Computational Biology 2004, 11:1034-1049.

Geng X, Mackey D. Dose-response to and systemic movement of dexamethasone in the GVG-inducible transgene system in *Arabidopsis*. Methods in Molecular Biology 2011, 712:59-68.

Giakountis A, Coupland G. Phloem transport of flowering signals. Current Opinion in Plant Biology 2008, 11:687-694.

Glinski M, Weckwerth W. The role of mass spectrometry in plant systems biology. Mass Spectrometry Reviews 2006, 25:173-214.

Gomes LC, Manuel S. 13C Metabolic Flux Analysis: From the Principle to Recent Applications. Current Bioinformatics 2012, 7:77-86.

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. Regulation of the anthocyanin biosynthetic pathway by the ttg1/bhlh/myb transcriptional complex in *Arabidopsis* seedlings. The Plant Journal 2008 53:814-827.

Goodacre R, Broadhurst D, Smilde AK, Kristal BS, Baker JD, Beger R, Bessant C, Connor S, Capuani G, Craig A, Ebbels T, Kell DB, Manetti C, Newton J, Paternostro G, Somorjai R, Sjöström M, Trygg J, Wulfert F. Proposed minimum reporting standards for data analysis in metabolomics. Metabolomics 2007, 3:231–241.

Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB. Metabolomics by numbers: acquiring and understanding global metabolite data. Trends in Biotechnology 2004, 22:245-252.

Goodman CD, Casati P, Walbot V. A Multidrug Resistance–Associated Protein Involved in Anthocyanin Transport in Zea mays. The Plant Cell 2004, 16:1812-1826.

Graf A, Schlereth A, Stitt M, Smith AM. Circadian control of carbohydrate availability for growth in *Arabidopsis* plants at night. Proceedings of the National Academy of Sciences of the United States of America 2010, 107:9458-9463.

Graf A, Smith AM. Starch and the clock: the dark side of plant productivity. Trends in Plant Science 2011, 16:169-175.

Gutteridge A, Kanehisa M, Goto S. Regulation of metabolic networks by small molecule metabolites. BMC Bioinformatics 2007, 8:88.

Hall R, Beale M, Fiehn O, Hardy N, Sumner L, Bino R. Plant metabolomics: the missing link in functional genomics strategies. The Plant Cell 2002, 14:1437–1440.

Hanada K, Sawada Y, Kuromori T, Klausnitzer R, Saito K, Toyoda T, Shinozaki K, Li W-H, Hirai MY. Functional compensation of primary and secondary metabolites by duplicate genes in *Arabidopsis thaliana*. Molecular Biology and Evolution 2011, 28:377–82.

Harrigan GG, LaPlante RH, Cosma GN, Cockerell G, Goodacre R, Maddox JF, Luyendyk JP, Ganey PE, Roth RA. Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity. Toxicology Letters 2004, 146:197-205.

Hart Y, Mayo AE, Milo R, Alon U. Robust Control of PEP Formation Rate in the Carbon Fixation Pathway of C4 Plants by a Bi-functional Enzyme. BMC Systems Biology 2011, 5:171.

Hartmann T. From waste products to ecochemicals: Fifty years research of plant secondary metabolism. Phytochemistry 2007, 68:2831-2846.

Haynes KA, Silver PA. Eukaryotic systems broaden the scope of synthetic biology. The Journal of Cell Biology 2009, 187:589-596.

Heim KE, Tagliaferro AR, Bobilya DJ. Flavonoid antioxidants: chemistry, metabolism and structure-activity relationships. The Journal of Nutritional Biochemistry 2002, 13:572-584.

Hennegan KP, Danna KJ. Pbin20: An improved binary vector for agrobacterium-mediated transformation. . Plant Molecular Biology Reporter 1998, 16:129-131.

Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nature Biotechnology 2010, 28:977–982.

Herrmann KM. The Shikimate pathway: Early steps in the biosynthesis of aromatic compounds. The Plant Cell 1995, 7:907-919.

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, Goda H, Nishizawa OI, Shibata D, Saito K. Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proceedings of the National Academy of Sciences of the United States of America 2007, 104:6478-6483.

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.* Proceedings of the National Academy of Sciences of the United States of America 2004, 101: 10205-10210.

Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M. Silencing of hydroxycinnamoyl-coenzymeA shikimate/quinate hydroxycinnamoyl-transferase affects phenylpropanoid biosynthesis. The Plant Cell 2004, 16:1446–1465.

Hong-Wu M, An-Ping Z. The connectivity structure, giant strong component and centrality of metabolic networks. Bioinformatics 2003, 19:1423-1430.

Horai H, Arita M, Kanaya S, Nihei Y, Ikeda T, Suwa K, Ojima Y, Tanaka K, Tanaka S, Aoshima K, et al (2010) MassBank: a public repository for sharing mass spectral data for life sciences. Journal of Mass Spectrometry 2010, 45: 703–714

Howell KA, Narsai R, Carroll A, Ivanova A, Lohse M, Usadel B, Millar AH, Whelan J. Mapping metabolic and transcript temporal switches during germination in rice highlights specific transcription factors and the role of RNA instability in the germination process. Plant Physiology 2009, 149:961-980.

Ito H, Grau WM. A gain-of-function mutation in the *arabidopsis* pleiotropic drug resistance transporter pdr9 confers resistance to auxinic herbicides. Plant Physiology 2006, 142:63-74.

Jakubowska A, Kowalczyk S. A specific enzyme hydrolyzing 6-O(4-O)-indole-3-ylacetyl-beta-D-glucose in immature kernels of Zea mays. Plant Physiology 2005, 162:207-213.

Jennifer LR, Palsson BO. Thirteen years of building constraint-based in silico models of *Escherichia coli*. Journal of Bacteriology 2003, 185:2692-2699.

Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C. Transcriptional repression by atmyb4 controls production of uv-protecting sunscreens in *Arabidopsis*. EMBO Journal 2000, 19:6150-6161.

Kaddurah-Daouk R, Kristal BS, Weinshilboum RM. Metabolomics: A global biochemical approach to drug response and disease. Annual review of Pharmacology and Toxicology 2008, 48:653-683.

Kim TY, Sohn SB, Kim HU, Lee SY. Strategies for systems-level metabolic engineering. Biotechnology Journal 2008, 3:612–623.

Klein C, Marino A, Sagot M-F, Milreu PV, Brilli M. Structural and dynamical analysis of biological networks. Briefings in Functional Genomics 2012, 11:420-433.

Kooke R, Keurentjes JJB. Multi-dimensional regulation of metabolic networks shaping plant development and performance. Journal of Experimental Botany 2011, 63:3353-3365.

Kreimer A, Borenstein E, Gophna U, Ruppin E. The evolution of modularity in bacterial metabolic networks. Proceedings of the National Academy of Sciences of the United States of America 2008, 105:6976-6981.

Laule O, Furholz A, Chang H-S, Zhu T, Wang X, Heifetz PB, Gruissem W, Lange BM. Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 2003, 100:6866-6871.

Lewis NG, Davin LB, Sarkanen S. The nature and function of lignins. In: Barton, Sir D.H.R., Nakanishi, K., Meth-Cohn, O. (Eds.), Comprehensive Natural Products Chemistry1999, 3:617–745.

Lewis NG, Davin LB. Evolution of lignan and neolignan biochemical pathways. In: Nes, W.D. (Ed.), Isopentenoids and Other Natural Products: Evolution and Function, . ACS Symposium Series1994, 562:202–246.

Lewis NG, Davin LB. Lignans: biosynthesis and function. In: Barton, Sir D.H.R., Nakanishi, K., Meth-Cohn, O. (Eds.), Comprehensive Natural Products Chemistry 1999, 1:639–712.

Lewis NG, Yamamoto E. Lignin: occurrence, biogenesis and biodegradation. Annual Review of Plant Physiology and Plant Molecular Biology 1990, 41:455–496.

Li, X., Weng, J.K., and Chapple C. Improvement of biomass through lignin modification. The Plant Journal 2008, 54: 569–581.

Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, Liou YC, Yu H. Direct interaction of *AGL24* and *SOC1* integrates flowering signals in *Arabidopsis*. Development 2008, 135:1481–1491.

Ludwig-Müller J. Auxin conjugates: their role for plant development and in the evolution of land plants. Journal of Experimental Botany 2011, 62:1757-1773.

Macel M, Van Dam NM, Keurentjes JJ. Metabolomics: the chemistry between ecology and genetics. Molecular Ecology Resources 2010, 10:583-93.

Masclauxa F, Charpenteaua M, Takahashib T, Pont-Lezicaa R, Galaud J-P. Gene silencing using a heat-inducible RNAi system in *Arabidopsis*. Biochemical and Biophysical Research Communications 2004, 321:364-369.

Mayer F, Takeoka GR, Buttery RG, Whitehand LC, Naim M, Rabinowitch HD. Studies on the aroma of five fresh tomato cultivars and the precursors of cis- and trans-4,5-epoxy-(E)-2-decenals and methional. Journal of Agricultural and Food Chemistry 2008, 56:3749-3757.

Mayer MJ, Narbad A, Parr AJ, Parker ML, Walton NJ, Mellon FA, Michael AJ. Rerouting the plant phenylpropanoid pathway by expression of a novel bacterial enoyl-coa hydratase/lyase enzyme function. The Plant Cell 2001, 13:1669-1682.

Mehrtens F, Kranz H, Bednarek P, Weisshaar B. The *Arabidopsis* transcription factor myb12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. Plant Physiology 2005, 138:1083-1096.

Messerli G, Nia VP, Trevisan M, Kolbe A, Schauer N, Geigenberger P, Chen J, Davison AC, Fernie AR, Zeeman Sc. Rapid classification of phenotypic mutants of *arabidopsis* via metabolite fingerprinting. Plant Physiology 2007, 143:1484-1492.

Meyerson M, Gabriel S, Getz G. Advances in understanding cancer genomes through second-generation sequencing. Nature Reviews Genetics 2010, 11:685–696.

Mintz-Orona S, Meira S, Malitskya S, Ruppin E, Aharoni A, Shlomi T. Reconstruction of *Arabidopsis* metabolic network models accounting for subcellular compartmentalization and tissue-specificity. Proceedings of the National Academy of Sciences of the United States of America 2011, 109:339-344.

Mithani A, Preston GM, Hein J. Rahnuma: hypergraph-based tool for metabolic pathway prediction and network comparison. Bioinformatics 2009, 25:1831-1832.

Morohashi K, Zhao M, Yang M, Read B, Lloyd A, Lamb R, Grotewold E. Participation of the *Arabidopsis* bHLH factor GL3 in trichome initiation regulatory events. Plant Physiology 2007, 145:736–746.

Morris WL, Ducreux LJM, Hedden P, Millam S, Taylor MA. Overexpression of a bacterial 1-deoxy-D-xylulose 5-phosphate synthase gene in potato tubers perturbs the isoprenoid metabolic network: implications for the control of the tuber life cycle. Journal of Experimental Botany 2006, 57:3007-3018.

Moxley JF, Jewettb MC, Antoniewicza MR, Villas-Boas SG,1,3, Alper H, Wheeler RT, Tong L, Hinnebusch AG, Ideker T, Nielsen J, Stephanopoulos G. Linking high-resolution metabolic flux phenotypes and transcriptional regulation in yeast modulated by the global regulator Gcn4p. Proceedings of the National Academy of Sciences of the United States of America 2009, 106:6477-6482.

Mueller LA, Goodman CD, Silady RA, Walbot V. AN9, a Petunia glutathione s-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. Plant Physiology 2000, 123:1561-1570.

Mur LAJ, Kenton P, Atzorn R, Miersch O, Wasternack C. The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. Plant Physiology 2006, 140:249–262.

Narasimhan K, Basheer C, Bajic VB, Swarup S. Enhancement of plant-microbe interactions using a rhizosphere metabolomics-driven approach and its application in the removal of polychlorinated biphenyls. Plant Physiology 2005 132:146-153.

Natalia Dudareva, Eran Pichersky, and Jonathan Gershenzon. Biochemistry of plant volatiles. Plant Physiology 2004, 135 :1893–1902

Nevoigt E. Progress in metabolic engineering of *Saccharomyces cerevisiae*. Microbiology and Molecular Biology Review 2008, 72:379–412.

Newman DJ, Cragg GM. Natural products as sources of new drugs over the last 25 years. Journal of Natural Products 2007, 70:461-477.

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. Analytical Chemistry. 2006,78:3289–3295.

Nordström A, Want E, Northen T, Lehtiö J, Siuzdak G. Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics. Analytical Chemistry 2008, 80:421–429.

Oberhardt MA, Palsson BO, Papin JA. Applications of genome-scale metabolic reconstructions. Molecular Systems Biology 2009, 5:320.

Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. Trends in Biotechnology 1998, 16:373–378.

Ossowski S, Schneeberger K, Clark RM, Lanz C, Warthmann N, Weigel D. Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Research 2008, 18:2024 - 2033.

Ossowski S, Schwab R, Weigel D. Gene silencing in plants using artificial microRNAs and other small RNAs. The Plant Journal 2008, 53:674-690.

Ouwerkerk PBF, Kam RJ, Hoge HC, Meijer AH. Glucocolticoid-inducible gene expression in rice. Planta 2001, 213:370–378.

Palsson BO. Systems biology: properties of reconstructed networks. Cambridge University Press 2006.

Papdi C, Abraham E, Joseph MP, Popescu C, Koncz C,Szabados L. Functional identification of *Arabidopsis* stress regulatory genes using the controlled cdna overexpression system. Plant Physiology 2008, 147:528-542.

Parter M, Kashtan N, Alon U. Environmental variability and modularity of bacterial metabolic networks. BMC Evolutionary Biology 2007, 7:169.

Patil KR, Rocha I, Förster J, Nielsen J. Evolutionary programming as a platform for *in silico* metabolic engineering. BMC Bioinformatics 2005, 6:308.

Peer WA, Brown DE, Tague BW, Muday GK, Taiz L, Murphy AS. Flavonoid accumulation patterns of transparent testa mutants of *Arabidopsis*. Plant Physiology 2001, 126:536-548.

Peregrín-Alvarez JM, Xiong X, Su C, Parkinson J. The modular organization of protein interactions in *Escherichia coli*. PLoS Computational Biology 2009, 5:e1000523.

Pfeiffer T, Soyer OS, Bonhoeffer S. The evolution of connectivity in metabolic networks. PLOS Biology 2005, 3:e228.

Pillai BVS, Swarup S. Elucidation of the flavonoid catabolism pathway in *Pseudomonas putida* pml2 by comparative metabolic profiling. Applied and Environmental Microbiology 2002, 68:143-151.

Quinones MP, Kaddurah-Daouk R. Metabolomics tools for identifying biomarkers for neuropsychiatric diseases Neurobiology of Disease 2009, 35:165–176.

Ramos I, Vivas EI, Downs DM. Mutations in the tryptophan operon allow PurF-independent thiamine synthesis by altering flux in vivo. Journal of Bacteriology 2008, 190 815–822.

Rasmussen S, Dixon RA. Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. The Plant Cell 1999, 11:1537-1552.

Reichard P. Ribonucleotide reductases: the evolution of allosteric regulation. Archives of Biochemistry and Biophysics 2002, 397:149–155.

Riedelsheimer C, Lisec J,Eysenberg AC, Sulpice R, Flis A, Grieder C, Altmann T, Stitt M, Willmitzer L, Melchinger AE. Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. Proceedings of the National Academy of Sciences of the United States of America 2012, 109:8872-8877.

Ro DK, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 2006.440:940–943.

Rodrigues JFM, Wagner A. Evolutionary plasticity and innovations in complex metabolic reaction networks. PLoS Computational Biology 2009, 5:e1000613.

Roslan HA, Salter MG, Wood CD, White MRH, Croft KP, Robson F, Coupland G, Doonan J, Laufs P, Tomsett AB, Caddick MX. Characterization of the ethanol-inducible alc gene-expression system in *Arabidopsis thaliana*. The Plant Journal 2001, 28:225–235.

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

Sakakibara H. Cytokinins: activity, biosynthesis, and translocation. Annual Review of Plant Biology 2006, 57:431–439.

Saslowsky D and Winkel-Shirley B. Localization of favonoid enzymes in *Arabidopsis* roots .The Plant Journal 2001, 27:37-48.

Schenk PM, Kazan K, Wilson I, Anderson JP, Richmond T, Somerville SC, Manners JM. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. Proceedings of the National Academy of Sciences of the United States of America 2000 97:11655–11660.

Schilling CH, Covert MW, Famili I, Church GM, Edwards JS, Palsson BO. Genome-scale metabolic model of *Helicobacter pylori* 26695. Journal of Bacteriology 2002, 184:4582-4593.

Simão RCG, Susin MF, Martinez CEA, Gomes SL. Cells lacking ClpB display a prolonged shutoff phase of the heat shock response in Caulobacter crescentus. Molecular Microbiology 2005, 57:592–603.

Smith AM, Stitt M. Coordination of carbon supply and plant growth. Plant, Cell and Environment 2007, 30:1126-1149.

Soyer OS, Pfeiffer T. Evolution under fluctuating environments explains observed robustness in metabolic networks. PLoS Computational Biology 2010, 6:e1000907.

Stark J, Callard R, Hubank M. From the top down: towards a predictive biology of signalling networks. Trends in Biotechnology 2003, 21:290–293.

Staswick PE, Serban B, Rowe M, Tiryaki I, Maldonado MT, Maldonado MC, Suza W. Characterization of an *Arabidopsis* enzyme family that conjugates amino acids to indole-3-acetic acid. The Plant Cell 2005, 17:616-627.

Stitt M, Sulpice R, Keurentjes J. Metabolic networks: how to identify key components in the regulation of metabolism and growth. Plant Physiology 2010, 152:428-444.

Sweetlove LJ, Fell D, Fernie AR. Getting to grips with the plant metabolic network. Biochemical Journal 2008, 409:27–41.

Sztein AE, Cohen JD, Garcia de la Fuente I, Cooke TJ. Auxin metabolism in mosses and liverworts. American Journal of Botany 1999, 86:1544-1555.

Takahashi, Y., and U. Tokumoto. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. Journal of Biological Chemistry 2002, 277 28380–28383.

Tanaka Y, Tsuda S, Kusumi T. Metabolic engineering to modify flower color. Plant Cell Physiology 1998, 39:1119-1126.

Travis S. Walker, Harsh Pal Bais, Erich Grotewold, and Jorge M. Vivanco. Root exudation and rhizosphere biology. Plant Physiology 2003, 132:44–51

Tweeddale H, Notley-McRobb L, Ferenci T. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool ('metabolome') analysis. Journal of Bacteriology 1998, 180:5109–5116.

Tzin V, Malitsky S, Aharoni A, Galili G. Expression of a bacterial bi-functional chorismate mutase/prephenate dehydratase modulates primary and secondary metabolism associated with aromatic amino acids in *Arabidopsis*. The Plant Journal 2009, 60:156-167.

Van Wees SCM, De Swart EAM, Van Pelt JA, Van Loon LC, Pieterse CMJ. Enhancement of induced disease resistance by simultaneous activation of salicylate- and jasmonate-dependent defense pathways in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 2000, 97: 8711–8716.

Venkatesh KV, Bhartiya S, Ruhela A. Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in Escherichia coli. FEBS Letters 2004, 563:234–240.

Vieira G, Sabarly V, Bourguignon P-Y, Durot M, Le Fèvre F, Mornico D, Vallenet D, Bouvet O, Denamur E, Schachter V, Médigue C. Core and panmetabolism in *Escherichia coli.* Journal of Bacteriology 2011, 193:1461-1472.

Vogt T. Phenylpropanoid Biosynthesis. Molecular Plant 2010, 3:2-20.

Vranova E, Coman D, Gruissem W. Structure and dynamics of the isoprenoid pathway network. Molecular Plant 2012, 5:318–333.

Wagner A, Fell DA. The small world inside large metabolic networks. Proceedings of the Royal Society of London Series B 2001, 268:1803-1810.

Wang J, Zhang Y, Marian C, Ressom HW. Identification of aberrant pathways and network activities from high-throughput data. Briefings in Bioinformatics 2012, 13:406-419.

Wang W, Vinocur B, Altman A. Plant responses to drought, salinity and extreme temperatures: Towards genetic engineering for stress tolerance. Planta 2003, 218:1-14

Wang YC, Zhang CH, Deng NY, Wang Y. Kernel-based data fusion improves the drug-protein interaction prediction. Computational Biology and Chemistry 2011, 35:353-362.

Weckwerth W, Loureiro ME, Wenzel K, Fiehn O. Differential metabolic networks unravel the effects of silent plant phenotypes. Proceedings of the National Academy of Sciences of the United States of America 2004, 101:7809–14

Werner E, Croixmarie V, Umbdenstock T, Ezan E, Chaminade P, Tabet JC, Junot C. Mass spectrometry-based metabolomics: accelerating the characterization of discriminating signals by combining statistical correlations and ultrahigh resolution. Analytical Chemistry 2008, 80:4918–4932.

Wiermann R, Gubatz S. Pollen wall and sporopollenin. International Review of Cytology 1992, 140:35–72.

Wille A, Zimmermann P, Vranova E, Fürholz A, Laule O, Bleuler S, Hennig L, Prelic A, von Rohr P, Thiele L, Zitzler E, Gruissem W, Bühlmann P. Sparse graphical Gaussian modeling of the isoprenoid gene network in *Arabidopsis thaliana*. Genome Biology 2004, 5:R92.

Williams TCR, Miguet L, Masakapalli SK, Kruger NJ, Sweetlove LJ, Ratcliffe RG. Metabolic network fluxes in heterotrophic *Arabidopsis* cells: stability of the flux distribution under different oxygenation conditions. Plant Physiology 2008, 148:704–718.

Winkel BS. 2004. Metabolic channeling in plants. Annual Review of Plant Biology 2004, 55:85-107.

Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiology 2001, 126:485–493.

Winkler WC. Metabolic monitoring by bacterial mRNAs. Archives of Microbiology 2005, 183:151–159.

Xu H, Zhang J, Zeng J, Jiang L, Liu E, Peng C, He Z, Peng X. Inducible antisense suppression of glycolate oxidase reveals its strong regulation over photosynthesis in rice. Journal of Experimental Botany 2009, 60:1799–1809.

Xu Li, Nicholas D. Bonawitz, Jing-Ke Weng and Clint Chapple. The Growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. The Plant Cell 2010, 22:1620–1632.

Yeang C-H, Vingron M. A joint model of regulatory and metabolic networks. BMC Bioinformatics 2006, 7:332.

Yin R, Messner B, Faus-Kessler T, Hoffmann T, Schwab W, Hajirezaei MR, von Saint Paul V, Heller W, Schäffner AR. Feedback inhibition of the general phenylpropanoid and flavonol biosynthetic pathways upon a compromised flavonol-3-O-glycosylation. Journal of Experimental Botany 2012, 63:2465 - 2478.

Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K. Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis*. The Plant Cell 2009, 20:2160–2176.

Yonekura-Sakakibara K, Tohge T, Niida R, Saito K. Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in *Arabidopsis* by transcriptome coexpression analysis and reverse genetics. The Journal of Biological chemistry 2007, 282:14932-14941.

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. Development 2003, 130:4859–4869.

Zhang W, Li F, Nie L. Integrating multiple ‘omics’ analysis for microbial biology: application and methodologies. Microbiology 2010, 156:287-301.

Zhang ZD, Paccanaro A, Fu Y, Weissman S, Weng Z, Chang J, Snyder M, Gerstein MB. Statistical analysis of the genomic distribution and correlation of regulatory elements in the encode regions. Genome Research 2007, 17:787-797.

Zhao J, Ding G-H, Tao L, Yu H, Yu Z-H, Luo J-H, Cao Z-W , Li Y-X. Modular co-evolution of metabolic networks. BMC Bioinformatics 2007, 8:311.

Zhao J, Yu H, Luo JH, Cao ZW, Li YX. Hierarchical modularity of nested bow-ties in metabolic networks. BMC Bioinformatics 2006, 7:386.

Zhong R, Ripperger A, Ye ZH. Ectopic deposition of lignin in the pith of stems of two *Arabidopsis* mutants. Plant Physiology 2000, 123:59-69.

APPENDIX I

GO subcategories enriched in differentially expressed gene list from mutants and inducible over-expression lines

GO term	tt8 mutant_%	ttg1 mutant_%	tt2 mutant_%	tt8:gr seedlings_%	ttg1:gr seedlings _%
oxidation reduction	5.548098	5.398111	5.409091	9.05982906	9.333333
secondary metabolic process	2.46085	2.519118	3.977273	4.273504274	3.111111
response to water	2.639821	1.079622	2.159091	1.709401709	2.888889
response to water deprivation	2.595078	1.079622	2.159091	1.709401709	2.444444
phenylpropanoid biosynthetic process	0.805369	0.899685	0	1.452991453	1.333333
response to organic substance	7.785235	0	9.886364	9.05982906	7.777778
response to endogenous stimulus	6.353468	0	7.5	7.692307692	7.333333
response to hormone stimulus	6.085011	0	6.931818	7.094017094	6.444444
response to oxidative stress	1.744966	1.979307	3.863636	5.299145299	0
response to temperature stimulus	2.818792	2.384166	3.75	4.957264957	0
response to radiation	3.176734	2.834008	5.113636	4.444444444	0
response to light stimulus	3.176734	2.699055	5.113636	4.358974359	0
intracellular signaling cascade	4.250559	0	0	4.273504274	0
response to heat	0.760626	1.124606	1.477273	3.931623932	0
response to osmotic stress	3.131991	2.024291	3.181818	2.991452991	0
cellular response to hormone stimulus	2.550336	0	0	2.991452991	0
hormone-mediated signaling	2.550336	0	0	2.991452991	0
response to salt stress	2.639821	0	3.068182	2.820512821	0
response to bacterium	1.47651	0	2.386364	2.735042735	0
cellular amino acid derivative biosynthetic process	1.073826	1.304543	1.477273	2.051282051	0
response to ethylene stimulus	1.923937	0	2.5	1.709401709	0
defense response to bacterium	1.252796	0	1.818182	1.709401709	0
response to carbohydrate stimulus	1.655481	0	3.181818	1.623931624	0
response to high light intensity	0.357942	0	0.795455	1.623931624	0
response to red or far red light	1.655481	1.169591	2.840909	1.452991453	0
toxin catabolic process	0.357942	0.494827	0.681818	1.367521368	0
toxin metabolic process	0.357942	0.494827	0.681818	1.367521368	0
cellular polysaccharide metabolic process	0.939597	1.124606	1.590909	1.282051282	0
response to chitin	1.252796	0	2.386364	1.196581197	0
response to salicylic acid stimulus	1.073826	0.854701	1.363636	1.111111111	0
response to gibberellin stimulus	1.252796	0	1.25	1.025641026	0
cellular glucan metabolic process	0.760626	0.989654	1.590909	1.025641026	0
response to red light	0.805369	0.62978	1.25	0.598290598	0
glycosinolate metabolic process	0.357942	0.359874	0	0.512820513	0

S-glycoside metabolic process	0.357942	0.359874	0	0.512820513	0
glucosinolate metabolic process	0.357942	0.359874	0	0.512820513	0
lignin biosynthetic process	0.357942	0	0	0.512820513	0
heat acclimation	0.223714	0	0	0.512820513	0
response to reactive oxygen species	0	1.124606	2.045455	3.846153846	1.555556
response to hydrogen peroxide	0	0.989654	1.931818	3.760683761	1.333333
oxygen and reactive oxygen species metabolic process	0	0.674764	1.136364	2.307692308	1.777778
cellular response to reactive oxygen species	0	0.674764	1.136364	2.136752137	1.555556
cellular response to oxidative stress	0	0.674764	1.136364	2.136752137	1.555556
cellular response to hydrogen peroxide	0	0.584795	1.136364	2.136752137	1.333333
hydrogen peroxide catabolic process	0	0.584795	1.136364	2.136752137	1.333333
hydrogen peroxide metabolic process	0	0.584795	1.136364	2.136752137	1.333333
lipid localization	0	0.989654	1.704545	1.794871795	5.333333
lipid transport	0	0.989654	1.590909	1.794871795	4
phenylpropanoid metabolic process	0	1.034638	0	1.709401709	1.555556
response to cold	2.192394	1.214575	2.272727	0	2
aging	0	0.494827	0	0	1.111111
systemic acquired resistance	0	0.31489	0	0	0.888889
cell wall organization	0	0	0	2.905982906	3.333333
external encapsulating structure organization	0	0	0	2.905982906	3.333333
plant-type cell wall organization	0	0	0	2.222222222	1.111111
cell wall modification	0	0	0	2.051282051	2
response to oxygen levels	0	0	0	0.341880342	0.666667
response to hypoxia	0	0	0	0.341880342	0.666667
response to abiotic stimulus	8.456376	6.252812	10.79545	10.76923077	7.111111

APPENDIX II

Heat map representing differentially expressed genes in transgenic lines related to metabolic pathways.

symbol	tt2 mutant	tt8 mutant	ttg1 mutant	tt8 dex seedlings	ttg1 dex seedling	Pathways
GGR						farnesyl diphosphate biosynthesis
GGPS6						farnesyl diphosphate biosynthesis
GPPS						farnesyl diphosphate biosynthesis
UGT73B2						2,4,6-trinitrotoluene degradation
AT1G05680						2,4,6-trinitrotoluene degradation
UGT73B4						2,4,6-trinitrotoluene degradation
PKT3						4-hydroxybenzoate biosynthesis V
ATASE2						ribonucleotide biosynthesis I
UGT75B1						abscisic acid glucose ester biosynthesis
SAG13						acetaldehyde biosynthesis I
ADH1						acetaldehyde biosynthesis I
PDC2						acetaldehyde biosynthesis I
AOX1D						aerobic respiration
UGT72E2						anthocyanin biosynthesis
UGT71D1						anthocyanin biosynthesis
ASP4						aspartate biosynthesis
CBB3						brassinosteroid biosynthesis I
BR6OX1						brassinosteroid biosynthesis I
ROT3						brassinosteroid biosynthesis I
BAS1						brassinosteroids inactivation
UGT73C5						brassinosteroids inactivation
PAD3						camalexin biosynthesis
CYP71A12						camalexin biosynthesis
CYP71A13						camalexin biosynthesis
AT1G12640						CDP-diacylglycerol biosynthesis I
ATCSLE1						cellulose biosynthesis
ATCSLG2						cellulose biosynthesis
PCB2						chlorophyllide a biosynthesis I
SQP2						cholesterol biosynthesis I
SQS1						cholesterol biosynthesis I
SQS2						cholesterol biosynthesis I
AT1G73600						choline biosynthesis I
PYK10						coumarin biosynthesis (via 2-coumarate)
BGLU21						coumarin biosynthesis (via 2-coumarate)
BGLU10						coumarin biosynthesis (via 2-coumarate)
BGLU27						coumarin biosynthesis (via 2-coumarate)
DIN2						coumarin biosynthesis (via 2-coumarate)
NIT4						cyanide detoxification II

UGT73C5						cytokinins glucoside biosynthesis
CKX4						cytokinins degradation
UGT85A1						cytokinins glucoside biosynthesis
AT2G36780						cytokinins glucoside biosynthesis
UGT76E12						cytokinins glucoside biosynthesis
CAT1						ethanol degradation IV
ACS2						ethylene biosynthesis I
CAC1						fatty acid biosynthesis initiation I
GPP2						flavin biosynthesis I
DMR6						flavonoid biosynthesis
ATSRG1						flavonoid biosynthesis
SUS4						galactose degradation III
SPS2						geranylgeranyldiphosphate biosynthesis
ATGA2OX2						gibberellin inactivation I
AT1G14130						gibberellin inactivation I
PGM						gluconeogenesis I
AT2G25450						glucosinolate biosynthesis
CYP79F2						glucosinolate biosynthesis
NSP1						glucosinolate breakdown
NSP5						glucosinolate breakdown
GDH2						glutamate degradation I
GAD4						glutamate degradation IV
PPDK						glutamine biosynthesis III
ATGPX6						glutathione redox reactions I
GST20						glutathione-mediated detoxification II
GST6						glutathione-mediated detoxification II
ATGSTU11						glutathione-mediated detoxification II
ATGSTU4						glutathione-mediated detoxification II
ADS1						glycolipid desaturation
PFK3						glycolysis I
XDH1						guanosine nucleotides degradation I
ATPMEPCRB						homogalacturonan degradation
AT5G49180						homogalacturonan degradation
ATPMEPCRD						homogalacturonan degradation
AAO1						IAA biosynthesis I
NIT2						IAA biosynthesis I
IAR3						IAA biosynthesis II
IBR3						IAA biosynthesis VII
AT4G18440						inosine-5'-phosphate biosynthesis II
BCAT4						isoleucine biosynthesis I
ACAT2						isoleucine degradation I
ACX1						jasmonic acid biosynthesis
OPR2						jasmonic acid biosynthesis
BEN1						leucodelphinidin biosynthesis
AT1G09500						leucodelphinidin biosynthesis

CYP89A5						leucodelphinidin biosynthesis
AT1G76470						leucodelphinidin biosynthesis
LUT5						lutein biosynthesis
ISPD						methylerythritol phosphate pathway
DXPS1						methylerythritol phosphate pathway
ISPF, MECPS						methylerythritol phosphate pathway
GST8						methylglyoxal degradation I
HMG1						mevalonate pathway I
AT1G31910						mevalonate pathway I
HMG2						mevalonate pathway I
PAP3						NAD/NADH
PSAD-2						oxygenic photosynthesis
AT5G39050						phenolic malonylglucosides biosynthesis
ELI3						phenylpropanoid biosynthesis
CAD1						phenylpropanoid biosynthesis
CAD5						phenylpropanoid biosynthesis
PSBQ-2						photosynthesis light reactions
CYP710A1						plant sterol biosynthesis
atnudt21						pyridine nucleotide cycling (plants)
NAPRT1						pyridine nucleotide cycling (plants)
UGT73B3						quercetin glucoside biosynthesis (Arabidopsis)
SCPL19						sinapate ester biosynthesis
SPDS3						spermidine biosynthesis I
AT1G69200						sucrose degradation III
AT5G51830						sucrose degradation III
AT4G33070						sucrose degradation to ethanol and lactate (anaerobic)
NPQ1						superpathway of carotenoid biosynthesis
BETA-OHASE 1						superpathway of carotenoid biosynthesis
LUT5						superpathway of carotenoid biosynthesis
LUT2						superpathway of carotenoid biosynthesis
CHY2						superpathway of carotenoid biosynthesis
ABA1						superpathway of carotenoid biosynthesis
TSB1						phenylalanine, tyrosine and tryptophan biosynthesis
AT3G49160						starch degradation to pyruvate
BAM1						sucrose and starch metabolism II
AT3G44830						triacylglycerol biosynthesis
TLL1						triacylglycerol degradation
ATBCAT-2						valine biosynthesis
HPT1						vitamin E biosynthesis
XGD1						xylogalacturonan biosynthesis
	-1.5<FC<1.5	**FC>2**	**1.5<FC<2**	**FC<-2**	**-2<FC<-1.5**	

APPENDIX III

List of differential metabolites in tt8 mutant line w.r.t Wild type. A part of this list has been described in figure 4.20. These metabolites were identified and mapped to metabolic pathways using Aracyc databases

Mass	fold	Common_Name	Pathway
290.163	-19.73	D-sedoheptulose-7-phosphate	(deoxy)ribose phosphate degradation
227.219	6.44	deoxycytidine	(deoxy)ribose phosphate degradation
214.111	-3.56	deoxyribose-1-phosphate	(deoxy)ribose phosphate degradation
214.111	-3.56	deoxyribose-5-phosphate	(deoxy)ribose phosphate degradation
953.701	-16.00	1,4-dihydroxy-2-naphthoyl-CoA	1,4-dihydroxy-2-naphthoate biosynthesis II (plants)
971.716	-3.14	4-(2'-carboxyphenyl)-4-oxobutyryl-CoA	1,4-dihydroxy-2-naphthoate biosynthesis II (plants)
744.418	-74.12	NADP+	2,3-cis-flavanols biosynthesis
301.232	32.51	delphinidin	2,3-cis-flavanols biosynthesis
348.2	4.07	2-O-(6-phospho-alpha-mannosyl)-D-glycerate	2-O-alpha-mannosyl-D-glycerate degradation
103.121	-2.68	4-aminobutyrate	4-aminobutyrate degradation I
600.88	-16.00	9'-cis-neoxanthin	abscisic acid biosynthesis
384.558	16.00	C25-allenic-apo-aldehyde	abscisic acid biosynthesis
426.463	-4.84	abscisic acid glucose ester	abscisic acid glucose ester biosynthesis
666.583	-2.72	stachyose	ajugose biosynthesis (galactinol-dependent)
180.157	2.56	myo-inositol	ajugose biosynthesis (galactinol-dependent)
104.111	-2.85	3-cyanopyridine	aldoxime degradation
122.126	-2.19	E-pyridine-3-aldoxime	aldoxime degradation
122.126	-2.19	nicotinamide	aldoxime degradation
134.091	2.22	(S)-ureidoglycolate	allantoin degradation to ureidoglycolate (ammonia producing)
867.608	-2.90	succinyl-CoA	alpha-ketoglutarate dehydrogenase complex
174.11	9.21	L-dehydro-ascorbate	antheraxanthin and violaxanthin biosynthesis
287.248	2.73	cyanidin	anthocyanin biosynthesis (pelargonidin 3-O-glucoside, cyanidin 3-O-glucoside)
507.183	-4.13	ATP	arginine biosynthesis II (acetyl cycle)
113.116	-3.70	pyrroline 5-carboxylate	arginine degradation II
60.055	-2.04	urea	arginine degradation II
605.345	16.00	GDP-L-gulose	ascorbate biosynthesis II (L-gulose pathway)
180.157	2.56	L-gulose	ascorbate biosynthesis II (L-gulose pathway)
605.345	16.00	GDP-L-galactose	ascorbate biosynthesis III (VTC2 cycle)

605.345	16.00	GDP-D-glucose	ascorbate biosynthesis IV (extended VTC2 cycle)
90.079	-2.13	3-hydroxy-propionate	beta-alanine biosynthesis II
132.119	2.09	3-ureidopropionate	beta-alanine biosynthesis III
418.702	11.21	(6alpha)-hydroxycampestanol	brassinosteroid biosynthesis I
450.701	-8.94	6-deoxocastasterone	brassinosteroid biosynthesis I
416.686	-8.56	6-oxocampestanol	brassinosteroid biosynthesis I
432.685	-5.13	3-dehydro-6-deoxoteasterone	brassinosteroid biosynthesis I
432.685	-5.13	cathasterone	brassinosteroid biosynthesis I
400.687	3.29	(5alpha)-campestan-3-one	brassinosteroid biosynthesis I
434.701	-2.97	6-deoxoteasterone	brassinosteroid biosynthesis I
434.701	-2.97	6-deoxotyphasterol	brassinosteroid biosynthesis I
448.685	-2.63	teasterone	brassinosteroid biosynthesis I
448.685	-2.63	typhasterol	brassinosteroid biosynthesis I
466.7	2.04	6alpha-hydroxy-castasterone	brassinosteroid biosynthesis I
418.702	11.21	6-deoxocathasterone	brassinosteroid biosynthesis II
416.686	-8.56	(22alpha)-hydroxy-5alpha-campestan-3-one	brassinosteroid biosynthesis II
416.686	-8.56	(22alpha)-hydroxy-campesterol	brassinosteroid biosynthesis II
414.67	4.67	(22alpha)-hydroxy-campest-4-en-3-one	brassinosteroid biosynthesis II
400.687	3.29	campest-4-en-3beta-ol	brassinosteroid biosynthesis II
180.157	2.56	2-carboxy-arabinitol	CA1P biosynthesis
180.157	2.56	hamamelose	CA1P biosynthesis
340.117	-2.26	hamamelose-21,5-bisphosphate	CA1P biosynthesis
186.058	20.53	3-phosphoglycerate	Calvin cycle
370.143	-2.11	D-sedoheptulose-1,7-bisphosphate	Calvin cycle
893.503	3.52	chlorophyll a	chlorophyll a biosynthesis I
891.487	-5.40	tetrahydrogeranylgeranyl-chlorophyll a	chlorophyll a biosynthesis II
887.455	2.91	geranylgeranyl-chlorophyll a	chlorophyll a biosynthesis II
630.982	-2.50	7-hydroxy-chlorophyllide a	chlorophyll cycle
598.983	-3.19	Mg-protoporphyrin monomethyl ester	chlorophyllide a biosynthesis
396.655	-11.07	4alpha-methyl-5alpha-cholesta-8,24-dien-3-one	cholesterol biosynthesis
428.654	-6.39	4alpha-carboxy-5alpha-cholesta-8,24-dien-3beta-ol	cholesterol biosynthesis
428.697	-6.39	4alpha-hydroxymethyl-4beta-methyl-5alpha-cholesta-8,24-dien-3beta-ol	cholesterol biosynthesis
410.682	-6.33	4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-beta-ol	cholesterol biosynthesis

414.67	4.67	4alpha-hydroxymethyl-5alpha-cholesta-8,24-dien-3beta-ol	cholesterol biosynthesis
382.628	2.51	5alpha-cholesta-8,24-dien-3-one	cholesterol biosynthesis
61.083	-2.41	ethanolamine	choline biosynthesis II
75.11	-2.07	N-monomethylethanolamine	choline biosynthesis II
254.133	56.98	shikimate-3-phosphate	chorismate biosynthesis
172.137	-5.33	3-dehydro-shikimate	chorismate biosynthesis
190.152	-2.56	3-dehydroquinate	chorismate biosynthesis
174.153	-2.25	shikimate	chorismate biosynthesis
324.18	2.18	5-enolpyruvyl-shikimate-3-phosphate	chorismate biosynthesis
431.341	-5.22	cis-zeatin riboside monophosphate	cis-zeatin biosynthesis
785.556	17.20	FAD	citrulline biosynthesis
141.02	-3.34	carbamoyl-phosphate	citrulline degradation
318.174	-8.15	3-deoxy-D-manno-octulosonate 8-P	CMP-KDO biosynthesis II (from D-ribulose 5-phosphate)
687.554	-23.51	dephospho-CoA	coenzyme A biosynthesis
60.032	-2.04	carbamate	cyanate degradation
147.13	-4.75	O-acetyl-L-serine	cysteine biosynthesis
377.356	16.00	kinetin-7-N-glucoside	cytokinins 7-N-glucoside biosynthesis
543.53	16.00	trans-zeatin-O-glucoside-7-N-glucoside	cytokinins 7-N-glucoside biosynthesis
377.356	16.00	kinetin-9-N-glucoside	cytokinins 9-N-glucoside biosynthesis
335.362	-2.34	isopentenyl adenosine	cytokinins degradation
545.545	-3.85	dihydrozeatin-9-N-glucoside-O-glucoside	cytokinins-O-glucoside biosynthesis
314.188	6.37	5'-phosphoribosyl-N-formylglycineamide	de novo biosynthesis of purine nucleotides
338.213	-3.31	AICAR	de novo biosynthesis of purine nucleotides
339.198	3.27	4-carboxyaminoimidazole ribonucleotide	de novo biosynthesis of purine nucleotides
463.297	2.06	adenylo-succinate	de novo biosynthesis of purine nucleotides
468.144	8.58	dUTP	de novo biosynthesis of pyrimidine deoxyribonucleotides
308.184	6.70	dUMP	de novo biosynthesis of pyrimidine deoxyribonucleotides
402.191	6.53	dTDP	de novo biosynthesis of pyrimidine deoxyribonucleotides
482.171	-3.29	dTTP	de novo biosynthesis of pyrimidine deoxyribonucleotides
387.179	-2.56	dCDP	de novo biosynthesis of pyrimidine deoxyribonucleotides
390.071	3.36	5-phosphoribosyl 1-pyrophosphate	de novo biosynthesis of uridine-5'-monophosphate
368.193	2.05	orotidine-5'-phosphate	de novo biosynthesis of uridine-5'-

			monophosphate
90.079	-2.13	D-lactate	D-lactate + UQ or MQ = pyruvate + UQH2 or MQH2
607.358	-4.80	UDP-N-acetyl-D-glucosamine	dolichyl-diphosphooligosaccharide biosynthesis
546.317	-9.30	dTDP-4-dehydro-6-deoxy-D-glucose	dTDP-L-rhamnose biosynthesis
548.333	6.57	dTDP-alpha-L-rhamnose	dTDP-L-rhamnose biosynthesis
564.333	-2.26	dTDP-D-glucose	dTDP-L-rhamnose biosynthesis
450.448	-10.44	ent-copalyl diphosphate	ent-kaurene biosynthesis
258.121	-3.67	2-keto-3-deoxy-6-phospho-gluconate	Entner-Doudoroff pathway I
410.725	-6.33	squalene	epoxysqualene biosynthesis
586.684	-5.29	presqualene diphosphate	epoxysqualene biosynthesis
256.257	2.81	isoliquiritigenin	flavonoid biosynthesis
316.267	-4.74	isorhamnetin	flavonol biosynthesis
335.331	-2.34	S-formylglutathione	formaldehyde oxidation (glutathione-dependent)
459.461	-5.57	5-methyl-THF	formylTHF biosynthesis
180.16	2.56	caffeate	free phenylpropanoid acid biosynthesis
180.157	2.56	D-galactose	galactose degradation II
566.305	-2.06	UDP-galactose	galactose degradation II
129.115	2.33	5-oxoproline	gamma-glutamyl cycle
523.183	-3.79	GTP	GDP-D-mannose biosynthesis
589.345	-4.24	GDP-L-fucose	GDP-L-fucose biosynthesis I (from GDP-D-mannose)
587.33	-4.04	GDP-4-dehydro-6-deoxy-D-mannose	GDP-L-fucose biosynthesis I (from GDP-D-mannose)
346.422	-6.07	GA24	gibberellin biosynthesis I (non C-3, non C-13 hydroxylation)
348.438	-3.79	GA14	gibberellin biosynthesis II (early C-3 hydroxylation)
348.438	-3.79	GA15	gibberellin biosynthesis II (early C-3 hydroxylation)
118.089	-3.03	succinate	gibberellin biosynthesis II (early C-3 hydroxylation)
348.395	-3.79	GA1	gibberellin biosynthesis III (early C-13 hydroxylation)
346.379	-5.10	GA29-catabolite	gibberellin inactivation
346.379	-5.10	GA34-catabolite	gibberellin inactivation
348.438	-3.79	GA110	gibberellin inactivation
348.395	-3.79	GA29	gibberellin inactivation
348.395	-3.79	GA34	gibberellin inactivation
348.438	-3.79	GA53	gibberellin inactivation
168.043	-3.27	phosphoenolpyruvate	gluconeogenesis
170.058	-2.53	D-glyceraldehyde-3-phosphate	gluconeogenesis
104.105	-2.85	4-hydroxybutyrate	glutamate degradation II

186.058	20.53	2-phosphoglycerate	glycolysis I (plant cytosol)
340.117	-2.26	fructose-1,6-bisphosphate	glycolysis I (plant cytosol)
192.125	6.80	citrate	glyoxylate cycle
562.667	-3.90	protoporphyrin IX	heme biosynthesis from uroporphyrinogen
580.289	5.84	UDP-D-galacturonate	homogalacturonan biosynthesis
546.392	-9.30	1,4-alpha-D-galacturonosyl	homogalacturonan degradation
973.732	-10.98	sinapoyl-CoA	hydroxycinnamic acid tyramine amides biosynthesis
343.379	-5.50	sinapoyltyramine	hydroxycinnamic acid tyramine amides biosynthesis
172.074	-5.33	sn-glycerol-3-phosphate	IAA biosynthesis I
156.187	-4.31	indole-3-acetonitrile	IAA biosynthesis I
203.197	2.64	indole-3-pyruvate	IAA biosynthesis I
175.187	2.00	indole-3-acetate	IAA biosynthesis II
966.743	-7.09	3-keto-indole-3-butyryl-CoA	IAA biosynthesis III
274.121	7.73	alpha-D-glucuronate 1-phosphate	inositol oxidation pathway
130.143	2.25	2-keto-3-methyl-valerate	isoleucine biosynthesis
865.635	-19.85	2-methylacetoacetyl-CoA	isoleucine degradation
849.636	-10.29	tiglyl-CoA	isoleucine degradation
787.572	-6.48	FADH2	isoleucine degradation
867.651	-2.90	2-methyl-3-hydroxybutyryl-CoA	isoleucine degradation
294.433	-4.44	3-oxo-2-(cis-2'-pentenyl)-cyclopentane-1-octanoate	jasmonic acid biosynthesis
578.526	-16.00	kaempferol-3-rhamnoside-7-rhamnoside	kaempferol glucoside biosynthesis (Arabidopsis)
448.382	9.14	kaempferol 7-O-glucoside	kaempferol glucoside biosynthesis (Arabidopsis)
448.382	9.14	kaempferol-3-glucoside	kaempferol glucoside biosynthesis (Arabidopsis)
610.524	-8.95	kaempferol 3,7-O-diglucoside	kaempferol glucoside biosynthesis (Arabidopsis)
432.383	-3.13	kaempferol-3-rhamnoside	kaempferol glucoside biosynthesis (Arabidopsis)
756.667	-2.20	kaempferol-3-O-gentiobioside-7-O-rhamnoside	kaempferol glucoside biosynthesis (Arabidopsis)
314.211	6.37	geranyl-PP	linalool biosynthesis
216.127	53.00	2-C-methyl-D-erythritol-4-phosphate	methylerythritol phosphate pathway
521.311	10.41	4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol	methylerythritol phosphate pathway
246.093	-6.17	isopentenyl diphosphate	methylerythritol phosphate pathway
214.111	-3.56	1-deoxy-D-xylulose 5-phosphate	methylerythritol phosphate pathway
346.293	-5.10	3,7-di-methylquercetagetin	methylquercetin biosynthesis
316.267	-4.74	3-O-methylquercetin	methylquercetin biosynthesis

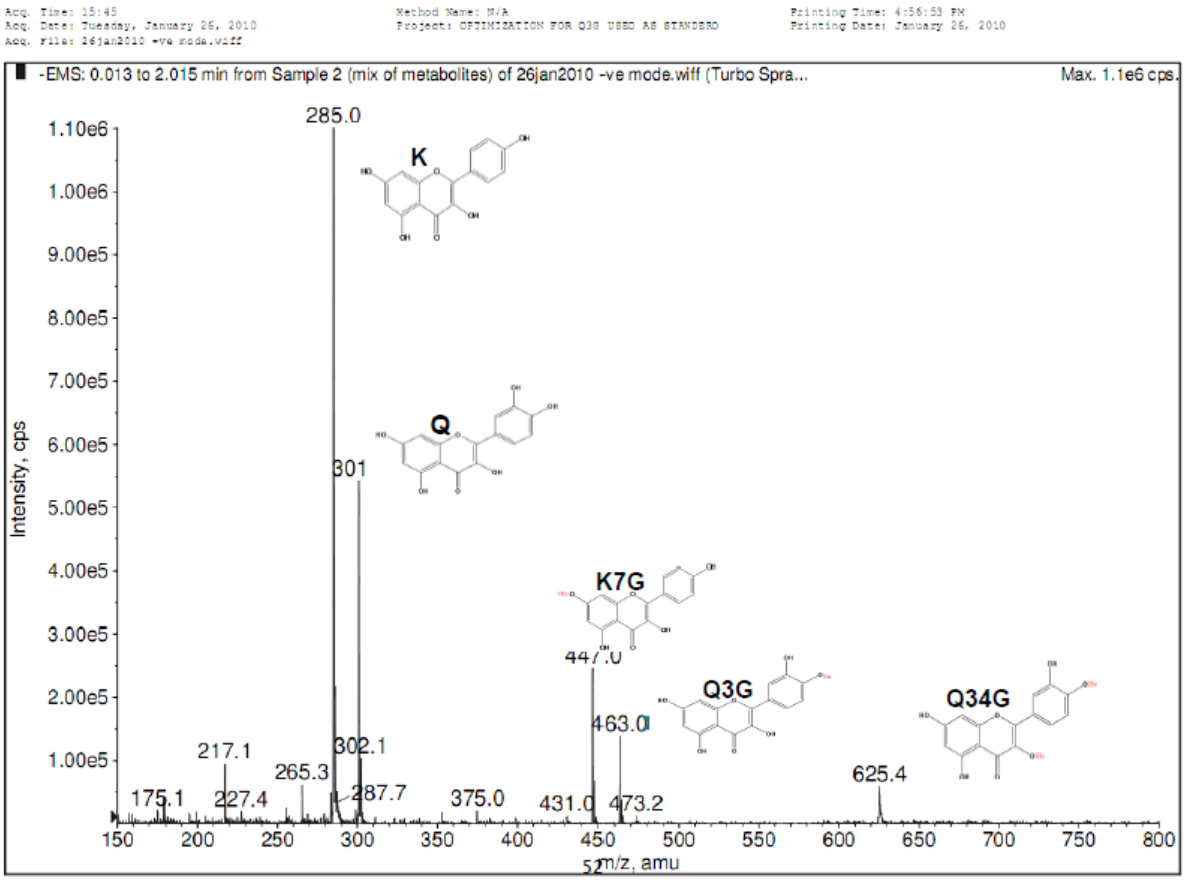
344.32	-3.49	3,7,4'-tri-O-methylquercetin	methylquercetin biosynthesis
308.118	6.70	mevalonate-5-PP	mevalonate pathway
586.684	-5.29	all-trans-hexaprenyl diphosphate	nonaprenyl diphosphate biosynthesis
518.566	-2.58	geranylfarnesyl diphosphate	nonaprenyl diphosphate biosynthesis
200.085	-2.69	D-erythrose-4-phosphate	non-oxidative branch of the pentose phosphate pathway
189.168	-9.50	N-acetyl-L-glutamate	ornithine biosynthesis
173.168	-2.55	N-acetyl-L-glutamate 5-semialdehyde	ornithine biosynthesis
276.136	-7.01	6-phospho-D-gluconate	oxidative branch of the pentose phosphate pathway
258.121	-3.67	D-glucono-delta-lactone-6-phosphate	oxidative branch of the pentose phosphate pathway
595.533	35.02	pelargonidin-3,5-diglucoside	pelargonidin conjugates biosynthesis
280.32	-5.63	8'-hydroxyabscisate	phaseic acid biosynthesis
280.32	-5.63	phaseic acid	phaseic acid biosynthesis
264.321	3.68	(+)-abscisate	phaseic acid biosynthesis
208.213	-11.37	sinapaldehyde	phenylpropanoid biosynthesis
192.168	6.80	quinate	phenylpropanoid biosynthesis
460.273	-4.36	CDP-N-methylethanolamine	phosphatidylcholine biosynthesis III
155.09	-2.65	N-methylethanolamine phosphate	phosphatidylcholine biosynthesis III
446.247	-4.96	CDP-ethanolamine	phosphatidylethanolamine biosynthesis II
141.063	-3.34	phosphoryl-ethanolamine	phosphatidylethanolamine biosynthesis II
323.199	-8.23	CMP	phosphatidylglycerol biosynthesis I (plastid)
184.152	10.44	phosphoryl-choline	phospholipases
104.062	-2.85	hydroxypyruvate	photorespiration
436.676	-17.92	demethylphylloquinone	phylloquinone biosynthesis
450.703	-8.94	phylloquinone	phylloquinone biosynthesis
336.215	4.12	nicotinic acid mononucleotide	pyridine nucleotide cycling (plants)
256.235	2.81	nicotinate riboside	pyridine nucleotide cycling (plants)
123.111	-2.10	nicotinate	pyridine nucleotide cycling (plants)
249.16	6.13	pyridoxine-5'-phosphate	pyridoxal 5'-phosphate salvage pathway
248.175	-2.37	pyridoxamine 5'-phosphate	pyridoxal 5'-phosphate salvage pathway
112.088	-2.35	uracil	pyrimidine salvage pathway
90.079	-2.13	lactate	pyruvate fermentation to lactate
550.306	-16.00	UDP-L-rhamnose	quercetin glucoside biosynthesis
302.24	10.25	a quercetin	quercetin glucoside biosynthesis
448.382	9.14	quercetin 3-O-rhamnoside	quercetin glucoside biosynthesis
610.524	-8.95	quercetin 3-O-rhamnoside-7-O-glucoside	quercetin glucoside biosynthesis
610.524	-8.95	quercetin-3-O-glucoside-7-O-rhamnoside	quercetin glucoside biosynthesis
596.541	16.00	quercetin-3-rhamnoside-7-rhamnoside	quercetin glucoside biosynthesis
626.524	-4.08	quercetin 3,7-O-diglucoside	quercetin glucoside biosynthesis

354.213	-39.07	5-amino-6-(5'-phosphoribosylamino)uracil	riboflavin and FMN and FAD biosynthesis
184.085	10.44	3,4-dihydroxy-2-butanone-4-P	riboflavin and FMN and FAD biosynthesis
276.249	-9.84	5-amino-6-ribitylamino-2,4(1H,3H)-pyrimidinedione	riboflavin and FMN and FAD biosynthesis
310.091	-2.12	D-ribulose-1,5-bisphosphate	Rubisco shunt
915.695	4.14	hydroxycinnamoyl-CoA	salicylic acid biosynthesis
704.672	2.99	tetrahydropteroyltri-L-glutamate	SAM cycle
943.705	-17.95	feruloyl-CoA	scopoletin biosynthesis
180.16	2.56	2,4-dihydroxycinnamate	simple coumarins biosynthesis
336.298	2.45	caffeoylshikimate	simple coumarins biosynthesis
386.355	-5.10	1-O-sinapoyl-beta-D-glucose	sinapate ester biosynthesis
396.655	-11.07	5-dehydro episterol	sterol biosynthesis
410.682	-6.33	(4alpha)-methyl-(5alpha)-ergosta-8,14,24(28)-trien-3beta-ol	sterol biosynthesis
410.682	-6.33	5-dehydroavenasterol	sterol biosynthesis
414.713	-4.66	sitosterol	sterol biosynthesis
400.687	3.29	24-epi-campesterol	sterol biosynthesis
400.687	3.29	campesterol	sterol biosynthesis
165.191	-6.92	L-phenylalanine	suberin biosynthesis
566.305	-2.06	UDP-D-glucose	sucrose biosynthesis
170.058	-2.53	dihydroxyacetone phosphate	sucrose degradation to ethanol and lactate (anaerobic)
595.535	35.02	cyanidin 3-(p-coumaroyl)-glucoside	superpathway of anthocyanin biosynthesis (from cyanidin and cyanidin 3-O-glucoside)
449.39	-3.27	cyanidin-3-O-beta-D-glucoside	superpathway of anthocyanin biosynthesis (from cyanidin and cyanidin 3-O-glucoside)
611.532	-2.74	cyanidin 3-O-sophoroside	superpathway of anthocyanin biosynthesis (from cyanidin and cyanidin 3-O-glucoside)
611.532	-2.74	cyanidin-3,5-diglucoside	superpathway of anthocyanin biosynthesis (from cyanidin and cyanidin 3-O-glucoside)
838.868	3.16	heme o	superpathway of proto- and siroheme biosynthesis
616.498	-2.55	protoheme IX	superpathway of proto- and siroheme biosynthesis
536.279	-7.86	UDP-D-xylose	UDP-D-xylose biosynthesis
180.157	2.56	D-fructose	UDP-glucose biosynthesis (from sucrose)
536.279	-7.86	UDP-L-arabinose	UDP-L-arabinose biosynthesis I (from UDP-xylose)
548.29	6.57	UDP-4-dehydro-6-deoxy-D-glucose	UDP-L-rhamnose biosynthesis
301.189	32.51	N-acetyl-D-glucosamine-6-phosphate	UDP-N-acetyl-D-glucosamine biosynthesis
301.189	32.51	N-acetyl-glucosamine-1-phosphate	UDP-N-acetyl-D-glucosamine biosynthesis

348.208	4.07	inosine-5'-phosphate	urate biosynthesis
152.112	-2.71	xanthine	urate biosynthesis
183.103	-45.71	5-hydroxyisourate	urate degradation to allantoin
201.118	-3.10	5-hydroxy-2-oxo-4-ureido-2,5-dihydro-1H imidazole-5-carboxylate	urate degradation to allantoin
167.104	-2.72	urate	urate degradation to allantoin
416.686	-8.56	2,3-dimethyl-6-phytyl-1,4-benzoquinone	vitamin E biosynthesis
416.686	-8.56	beta-tocopherol	vitamin E biosynthesis
416.686	-8.56	gamma-tocopherol	vitamin E biosynthesis
767.534	2.49	coenzyme A	volatile benzenoid ester biosynthesis
580.289	5.84	UDP-D-glucuronate	xylan biosynthesis

APPENDIX IV

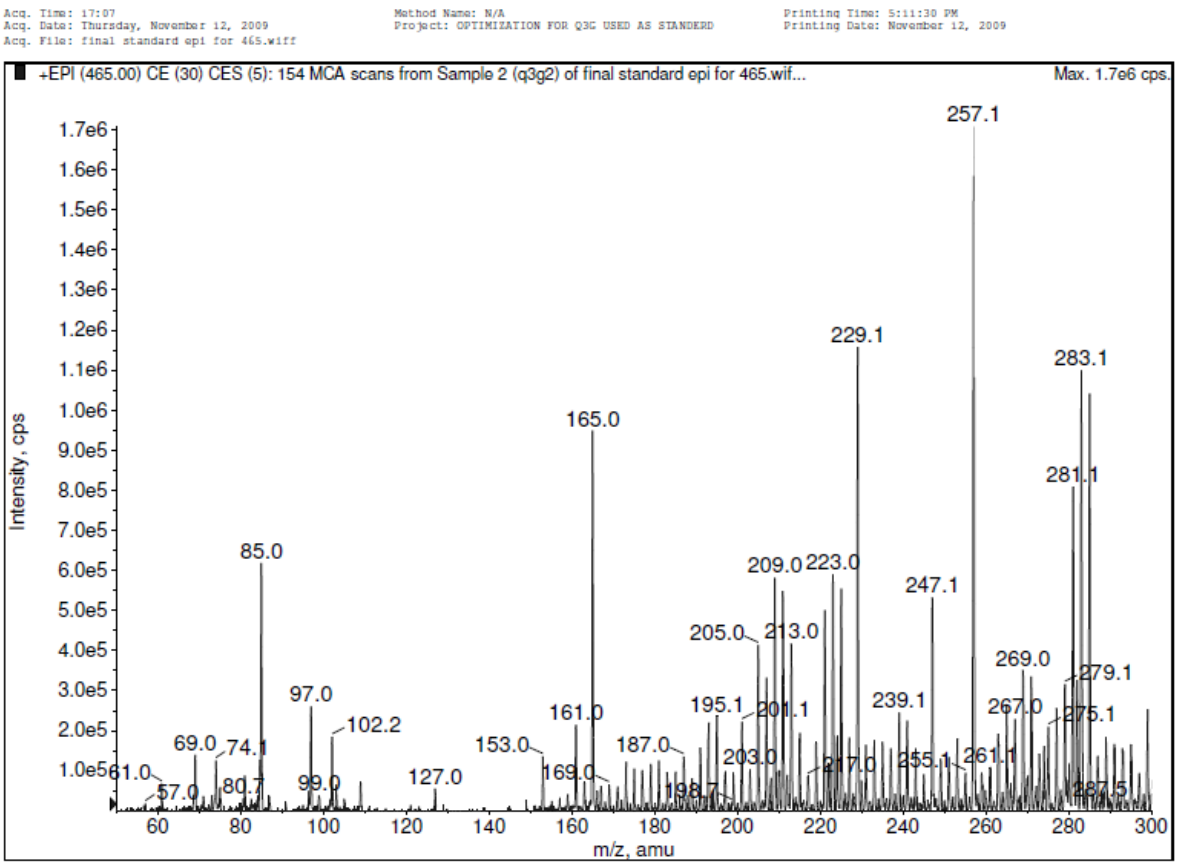
MS and MS2 profiling of 5 standards used as glycosylated derivatives of quercetin and kaempferol



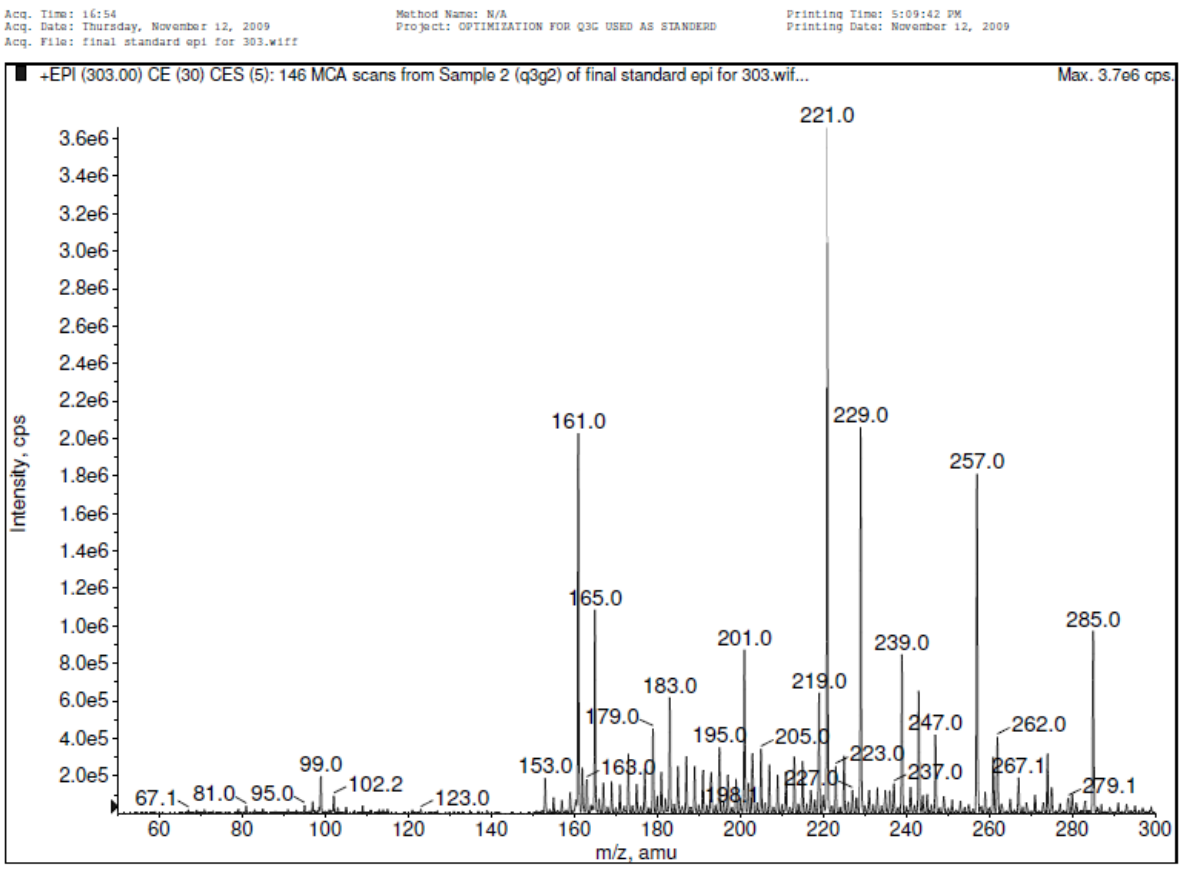
Enhanced mass scan of mix of 5 metabolites, 2 of them are aglycone while rest 3 are glucoside derivative

Q-quercetin Q3G- quercetin-3 glucoside Q34G- quercetin-3,4 glucosie

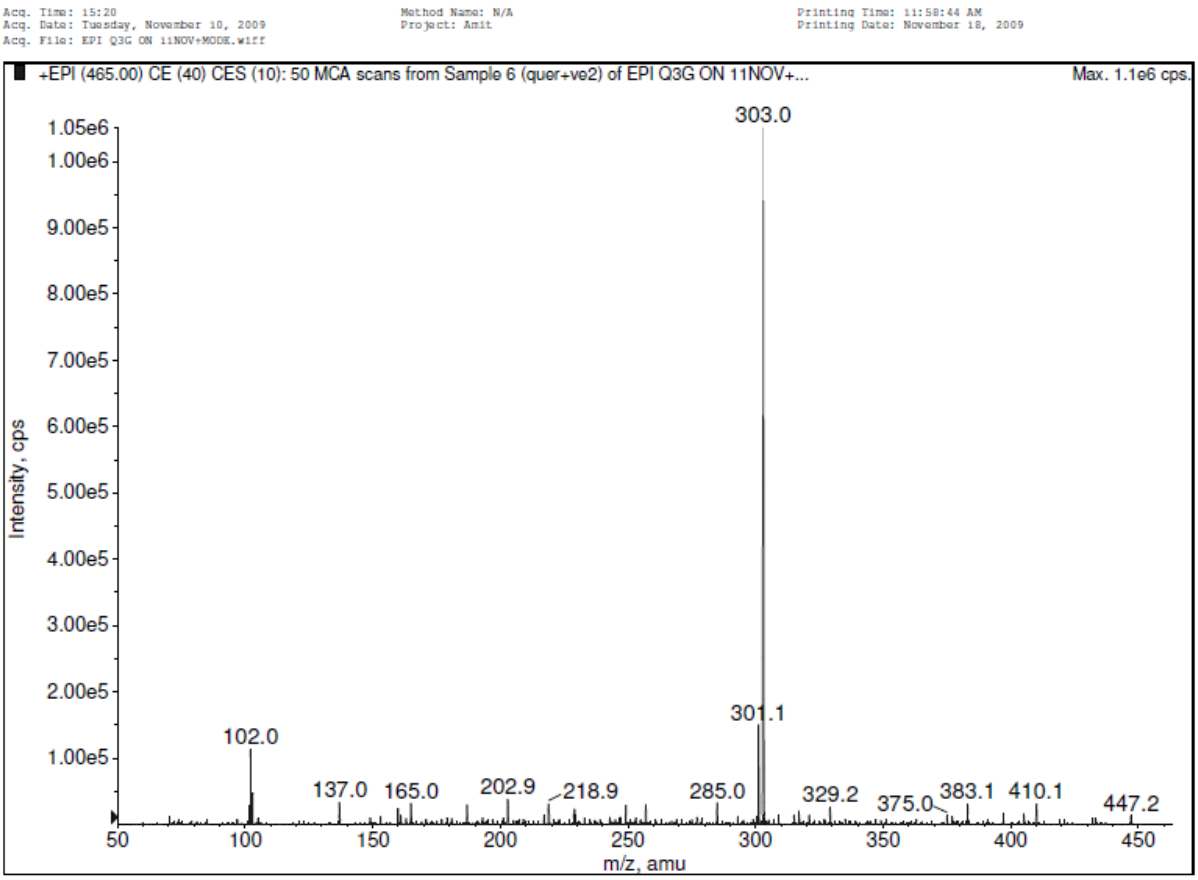
K- kaempferol K7G- kaempferol-7 glucoside



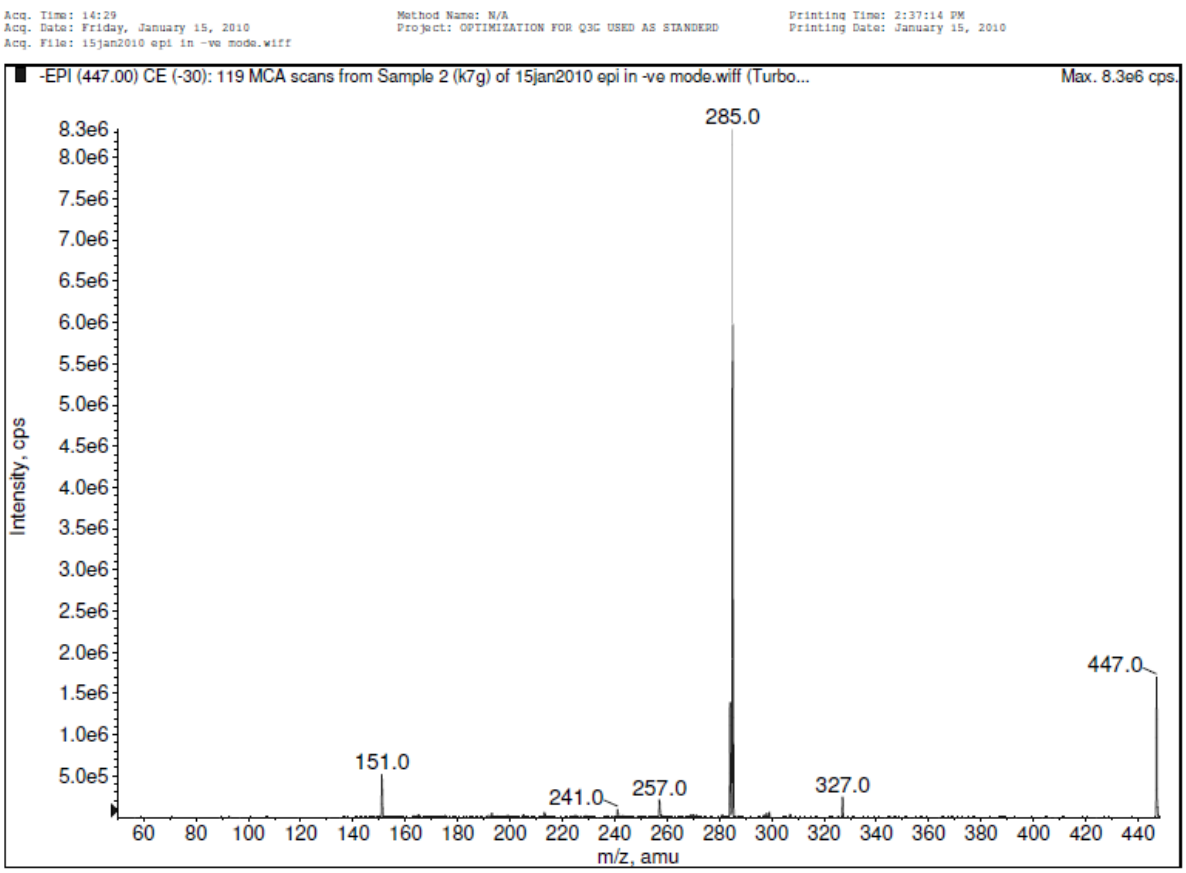
Product Ion scan for Q3G at collision energy 30ev.



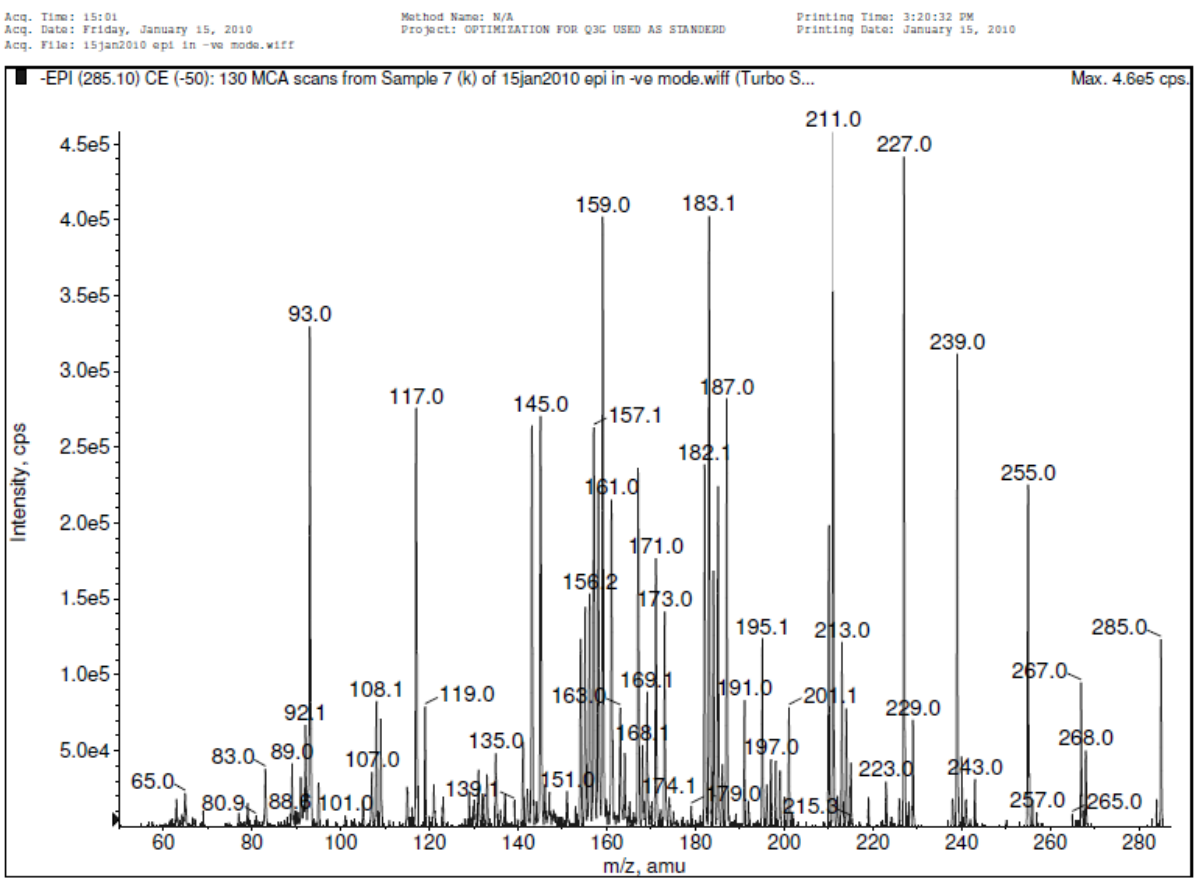
Product Ion scan for Q at collision energy 30ev.



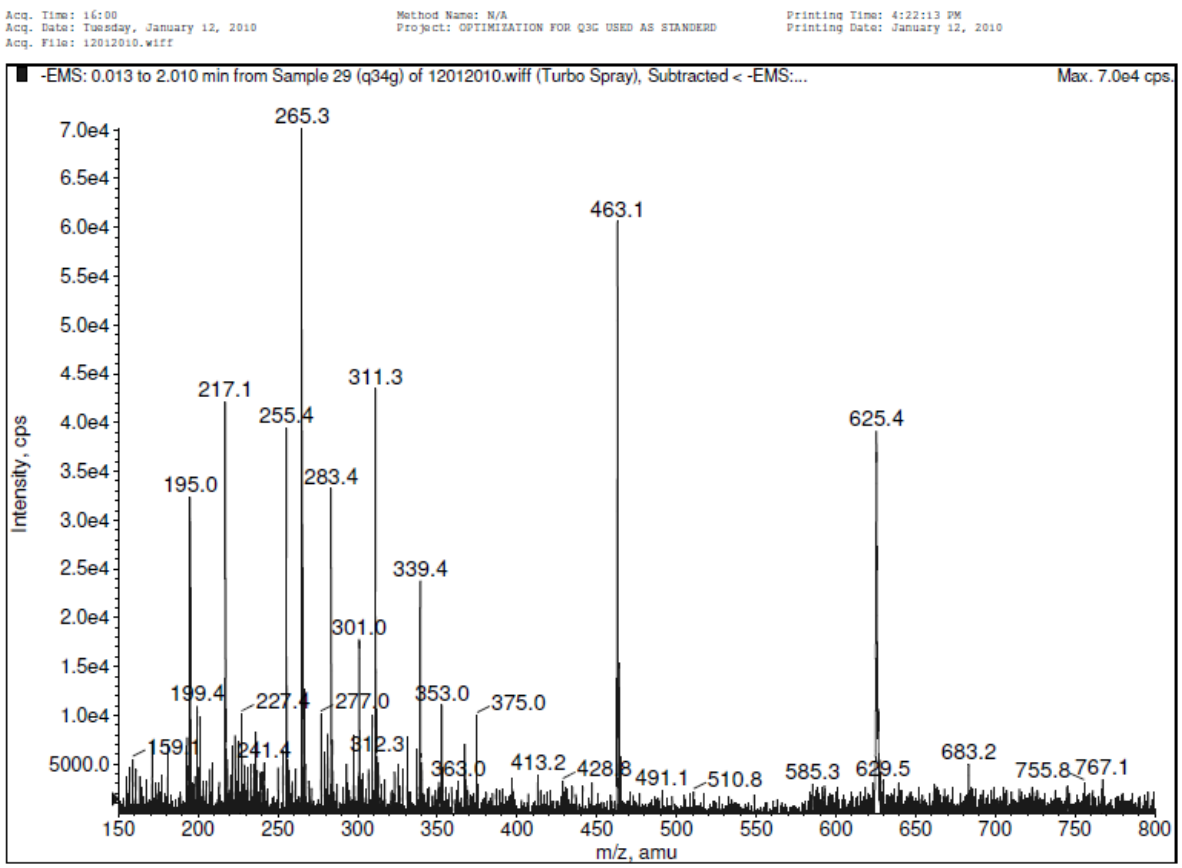
Product Ion scan for Q3G at collision energy 30ev. This shows release of Q, and shows that it is derivative of quercetin



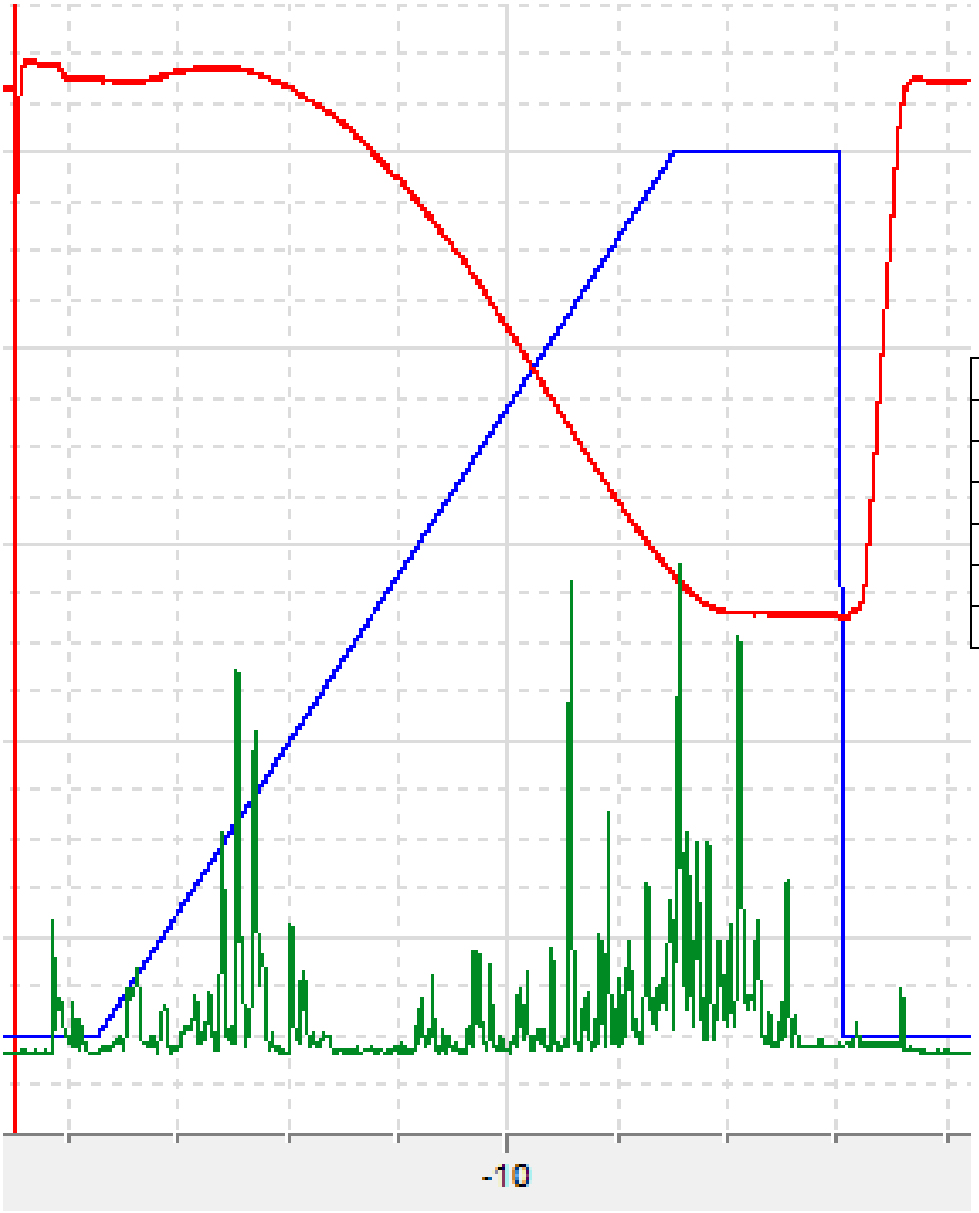
Product Ion scan for k7G at collision energy 30ev. 285 mass represents Kaempferol.



Product Ion scan for k at collision energy 30ev.

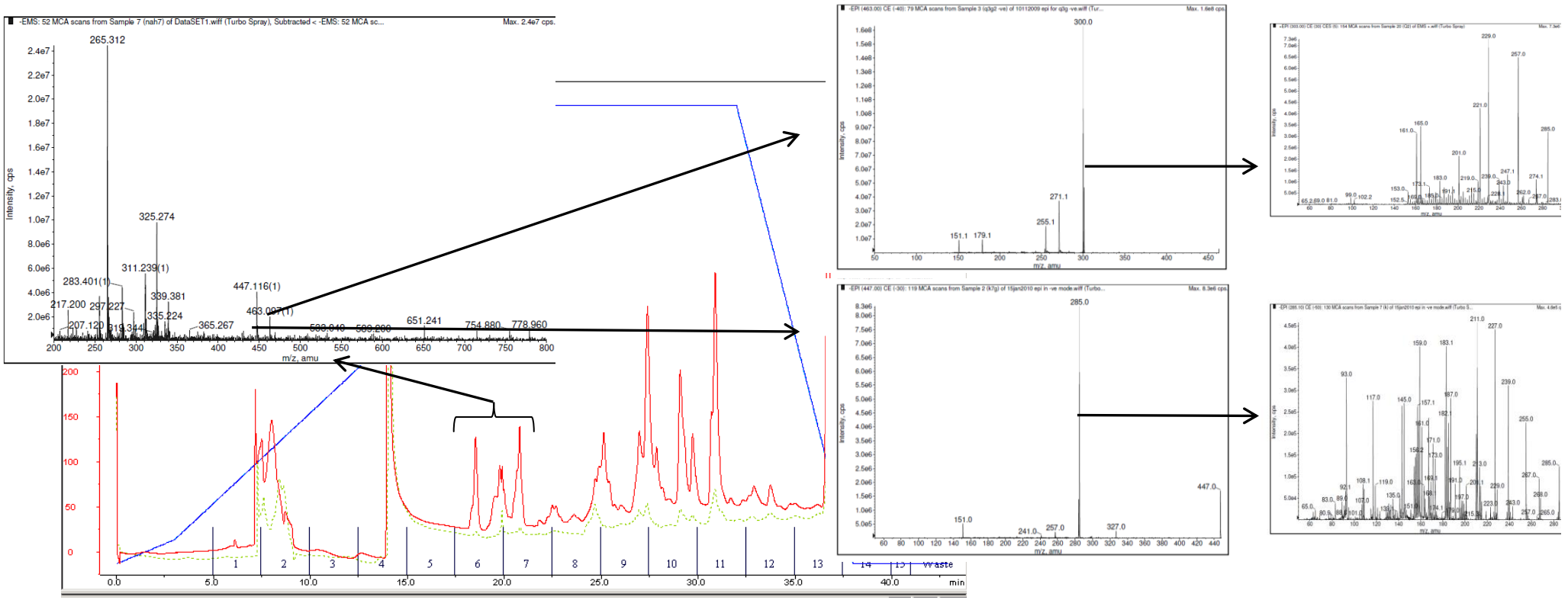


Product Ion scan for Q34G at collision energy 30ev.



time	%B
0	10
1.5	10
12	100
15	100
15.1	10
18	10

HPLC method established for LC-MS runs for all the samples described in this thesis. Here green graph represents total ion chromatogram, blue represents gradient for organic solvent (%B refers to acetonitrile) and red refers to pressure recorded for the run.



HPLC profile of Wt (red) superimposed upon that of tt8 mutant line (Green), showing several peaks absent in tt8 mutant lines. Selecting fraction 6 and 7th of HPLC run and MS scan showed several glycosylated metabolites absent in tt8 mutant lines. Here in this example, mass ion 447.11 and 463.09 were fragmented and showed derivative of kaempferol and quercetin which was then confirmed by MS2 of 285 (kaempferol) and 303(quercetin) which confirmed these metabolites and their derivatives.