Promotor: Professor Alain Goossens

Identification of Potential Regulators of Jasmonate-Modulated Secondary Metabolism in *Medicago truncatula* Azra Gholami

Identification of Potential Regulators of Jasmonate-Modulated Secondary Metabolism in *Medicago truncatula*

Azra Gholami

June 2013

Ghent University - Faculty of Sciences
Department of Plant Biotechnology and Bioinformatics
VIB - Department of Plant Systems Biology

Identification of Potential Regulators of Jasmonate-Modulated Secondary Metabolism in *Medicago truncatula*

Azra Gholami

Thesis submitted in partial fulfillment of the requirements
for the degree of Doctor (PhD) in Sciences: Biotechnology

Academic year: 2012-2013
Promotor: Prof. Alain Goossens

This work was conducted in the VIB Department of Plant Systems Biology, Ghent University.

Board of Examiners

Prof. Wout Boerjan (Chair)

VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

Prof. Alain Goossens (Promotor)
VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

Prof. Sofie Goormachtig *
VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

Prof. Bartel Vanholme*
VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

Prof. Jan Van Bocxlaer*
Department of Bio-analysis, Faculty of Pharmaceutical Sciences, Ghent University

Prof. Rosa Maria Cusido*

Departament de Productes Naturals, Biologia Vegetal i Edafologia, University of Barcelona

Prof. Tom Beeckman
VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

Dr. Jacob Pollier

VIB Department of Plant Systems Biology, Department of Plant Biotechnology and Bioinformatics, Faculty of Sciences, Ghent University

***** Members of the Reading Committee

Table of contents

Commonly used abbreviations and gene names

40S	40S ribosomal protein S8
4CL	4-coumarate: CoA ligase
AAC	Alcohol acyl-transferase
ABA	Abscisic acid
ABC	ATP binding cassette
ACC	Acetyl-CoA carboxylase
ACN	Acetonitrile
ADS	Amorpha-4,11-diene synthase
ADH	Alcohol dehydrogenase
AFLP	Amplified fragment length polymorphism
AM	Arbuscular mycorrhization
ANR	Anthocyanidin reductase
ANS	Anthocyanidin synthase
AOC	Allene oxide cyclase
AOS	Allene oxide synthase
AP2/ERF	APETALA2/ETHYLENE Response Factor
AVIs	Anthocyanic vacuolar inclusions
BAN	BANYULS gene
BAS	β-Amyrin synthase
bHLH	basic helix–loop–helix
BLAST	Basic local alignment search tool
bp	Base pair
C4H	Cinnamate 4-hydroxylase
CAD1	(1)- δ-cadinene synthase
CAF1	Ccr4-associated factor1
CAS	Cycloartenol synthase
Ccr4-Not complex	Carbon catabolite repressor protein 4 - Negative on TATA complex
CHI	Chalcone isomerase
CHR	Chalcone reductase
CHS	Chalcone synthase

CID	Collision-induced dissociation
CMC	Critical micelle concentration
COI1	Coronatine insensitive 1
Cou	Coumaric acid
CPC	CAPRICE
CYP450	Cytochrome P450
DAT	Acetyl-coA:4-*O*-deacetylvindoline 4-*O*-acetyltransferase
DFR	Dihydroflavonol-4-reductase
DMAPP	Dimethylallyl diphosphate
DMID	Dihydroxy-4'-methoxy- isoflavanol dehydratase
Dof	DNA-binding with one finger
EAR	ERF-associated amphiphilic repression
EEP	Endonuclease–exonuclease–phophatase
EGL3	Enhancer of GL3
ELF1α	Elongation factor 1α
ERF	Ethylene response factor
ER	Endoplasmic reticulum
ESI-IT MS	Electrospray ionization ion-trap mass spectrometry
EST	Expressed sequence tag
ε_{max}	Maximum molar extinction coefficients
F3H	Flavanone 3-hydroxylase
F3′5′H	Flavonoid 3′,5′-hydroxylase
Fer	Ferrulic acid
FL- ORF	Full length- open reading frame
FLS	Flavonol synthase
FNS	Flavone synthase
FPKM	Fragments per kilobase of transcript per million fragments mapped
FPS	Farnesyl pyrophosphate synthase
FT-ICR MS	Fourier transform ion cyclotron resonance mass spectrometry
G10H	Geraniol 10-hydroxylase
GC- MS	Gas chromatography- mass spectrometry
GL3	Glabra3

Glc	Glucose
GlcA	Glucuronic acid
GlcAPyr	Glucuronopyranosyl
GLS	Glucosinolates
GLVs	Green leaf volatiles
GST	Glutathione *S*-transferase
HD- ZIP	Homeodomain-leucine zipper
HMGR	3-hydroxy-3-methylglutaryl-CoA reductase
HDAC	Histone deacetylases
HPL	Hydroperoxide lyase
HPLC	High-performance liquid chromatography
HPOT	Hydroperoxy linolenic acid
hpRNAi	Hairpin RNA-mediated Interference
I2'H	Isoflavone 2'-hydroxylase
I3'H	Isoflavone 3'-hydroxylase
IFR	Isoflavone reductase
IFS	Isoflavone synthase
IPP	Isopentenyl pyrophosphate
JA	Jasmonate
JA–Ile	JA–isoleucine
JAM1	JA factor-stimulating MAPKK 1
JAZ	Jasmonate ZIM-domain
JMT	Jasmonate *O*-methyltransferase
LAP1	Legume anthocyanin production 1
LAR	Leucoanthocyanidin reductase
LIP	Lipase
LiF	Laser-induced fluorescence
LOX	Lipoxygenase
m/z	Mass-to-charge ratio
Mal	Malonic acid
MAPKK	Mitogen-activated protein kinase kinase
MaT	Malonyltransferases

MATE	Vacuole-localized multidrug and toxic extrusion
MeJA	Methyl jasmonate
MEKK1	MAPK/ERK kinase kinase 1
MEP pathway	2-C-methyl-D-erythritol 4-phosphate pathway
MAPK	Mitogen-activated protein kinase
MtPAR	*M. truncatula* proanthocyanidin regulator
MVA	Mevalonic acid pathway
NFs	Nod factors
Nod gene	Nodulin gene
NINJA	Novel Interactor of JAZ
NMR	Nuclear magnetic resonance
NPAAS	Nonprotein amino acids
ORCA	Octadecanoid-derivative responsive catharanthus AP2-domain
OSC	Oxidosqualene cyclase
PAL	Phenylalanine ammonia-lyase
PA	Proanthocyanidin
PAP1	Production of anthocyanin pigment 1
PCA	Principal component analysis
PMT	Putrescine *N*-methyltransferase
PR1a	Pathogenesis-related protein 1a
PRX	Peroxidase
PSK-α	Phytosulfokine-α
PVCs	Prevacuolar compartments
PVs	Plant volatiles
QRT-PCR	Quantitative Real-time Polymerase Chain Reaction
QToF	Quadrupole time-of-flight
SCF^{COI1}	Skp–Cullin–F-box-type E3 ubiquitin ligase complex
SD-U	Synthetic defined medium lacking URA
SGD	Strictosidine *β*-d-glucosidase
SQE	Squalene epoxidase
SEM	Standard error of the mean
SQS	Squalene synthase

SNAREs	Soluble N-ethylmaleimide-sensitive factor Attachment protein Receptors
TAP	Tandem affinity purification
TCP	Teosinte Branched/Cycloidea/Pcf
TDC	Tryptophan decarboxylase
TFs	Transcription factors
TGN	Trans-Golgi network
TIA	Terpenoid indole alkaloid
TIC	Total ion current
TPI	Trypsin protease inhibitor
TPL	Topless
Trp	Tryptophan
TT8	Transparent testa 8
TTG1	Transparent Testa Glabra 1
UGT	UDP-dependent glycosyltransferase
UPLC	Ultra performance liquid chromatography
UTR	Untranslated region
VR	Vestitone reductase
YE	Yeast elicitor

Abstract

Accumulation of secondary metabolites often occurs in plants imposed to biotic and abiotic stress signals. Perception of the stress signals triggers a signal transduction network that leads to the activation or *de novo* biosynthesis of transcription factors. Then transcription factors regulate the expression of the genes encoding enzymes that catalyze the biosynthesis of target secondary metabolites. Here, we focused on the identification of the regulators of secondary metabolite biosynthesis in the model legume *Medicago truncatula*. Considering the conserved role of JAs in the induction of secondary metabolite biosynthesis in plants, we exploited a gene list generated previously through a genome-wide transcript profiling of MeJA-elicited cell cultures of *M. truncatula*. In this PhD thesis, several of the potential regulators identified in the previous study were selected for further characterization. Through a reverse genetics approach two regulatory proteins were found with a putative role in secondary metabolite production in *M. truncatula* hairy roots. Overexpression of *Mt148* encoding a CCR4-associate factor1 (Caf1) protein led to the downregulation of some secondary metabolism genes in *M. truncatula* hairy roots. Overexpression of *Mt061* codifying for an R2R3-type MYB family transcription factor, demonstrated that it is a regulator of green leaf volatiles (GLV) biosynthesis. This MYB-transcription factor can also modulate defense response processes in *M.truncatula*.

Furthermore, we studied the putative role of small signaling peptides in the signaling transduction pathways towards the regulation secondary metabolism in *M. truncatula*. We found that Taximin, a small signaling peptide previously identified in *Taxus baccata*, as well as the *M. truncatula* taximin homologs can affect secondary metabolite biosynthesis in *M. truncatula*.

1

Transcriptional machineries in jasmonate-elicited plant secondary metabolism

Nathan De Geyter, Azra Gholami*, Sofie Goormachtig and Alain Goossens*

Published in Trends in Plant Science (IF: 11.047)
De Geyter, N.*, Gholami, A.*, Goormachtig, S., and Goossens, A. *(2012). Transcriptional machineries in jasmonate-elicited plant secondary metabolism. Trends in Plant Science* ***17,*** *349-359.*
** These authors contributed equally to this work. (Parts 5, 6 and 7 were written by the candidate).*

Abstract

Jasmonates (JAs) act as conserved elicitors of plant secondary metabolism. JA perception triggers an extensive transcriptional reprogramming leading to the concerted activation of entire metabolic pathways. This observation triggered numerous quests for 'master' regulators capable of enhancing the production of specific sets of valuable plant metabolites. Many transcription factors (TFs), often JA-activated themselves, with a role in the JA-modulated regulation of metabolism were discovered. At the same time, it became clear that metabolic reprogramming is subjected to complex control mechanisms integrated in robust cellular networks. Here, we will discuss the current knowledge on the effect of JA-modulated TFs in the elicitation of secondary metabolism in the model plant Arabidopsis and a range of medicinal plant species with structurally divergent secondary metabolites. We draw parallels with the regulation of secondary metabolism in fungi and consider the remaining challenges to map and exploit the transcriptional machineries that drive JA-mediated elicitation of plant secondary metabolism.

1 JAs: ubiquitous and conserved elicitors of plant secondary metabolism

JAs are oxylipin-derived phytohormones that regulate a wide variety of physiological plant processes ranging from growth and development to reproduction and defence. Originally, JAs were labelled as secondary metabolites present in the scent of jasmine flowers (*Jasminum* spp.). Now, it has become clear that they themselves act as elicitors of the production of secondary metabolites across the plant kingdom, from angiosperms to gymnosperms (Zhao et al., 2005; Wasternack, 2007; Browse, 2009; Pauwels et al., 2009). This suggests that the signalling machinery underlying JA-mediated secondary metabolite elicitation may be conserved and that it was installed early in the higher plant lineage, which seems to be supported by the existence of a conserved module for JA perception and subsequent 'primary' signal transduction (Browse, 2009; Chini et al., 2009a; Memelink, 2009; Pauwels and Goossens, 2011). Nevertheless, there is a pronounced degree of species specificity with respect to the metabolic pathways that are elicited by JAs. This is reflected by the specific JA-mediated compendium of bioactive metabolites of a wide structural variety and different biochemical origin that can be found in each plant species. Broadly, three major classes of plant secondary metabolites can be defined: the terpenoids, alkaloids and phenylpropanoids; however, more exist. JAs can induce the synthesis of molecules in all these classes (Zhao et al., 2005; Pauwels et al., 2009). In addition, JAs can modulate particular primary metabolic pathways to supply connected secondary metabolite pathways with the necessary substrates (Pauwels et al., 2009;

Spitzer-Rimon et al., 2010). Hence, downstream of a conserved elicitation mechanism, species-specific secondary metabolic pathways have evolved under JA control.

Several genome-wide transcript profiling studies have demonstrated that JA treatment triggers an extensive transcriptional reprogramming of metabolism. The expression of genes encoding enzymes involved in one particular secondary metabolic pathway often displayed a marked concerted upregulation after JA elicitation, leading to the recognition of so-called 'transcriptional regulons' (Pauwels et al., 2009). This prompted many researchers worldwide to launch gene discovery projects to identify so-called 'master switches' of plant secondary metabolism: that is, proteins capable of activating expression of all or most of the genes encoding the enzymes involved in one particular metabolic pathway. This knowledge could ultimately be converted into powerful generic tools for plant metabolic engineering programmes. The identification, at the end of the past millennium, of TFs such as OCTADECANOID-DERIVATIVE RESPONSIVE CATHARANTHUS AP2-DOMAIN 2 and 3 (ORCA2 and ORCA3), driving terpenoid indole alkaloid (TIA) synthesis in Madagascar periwinkle (*Catharanthus roseus*) (reviewed in (Memelink et al., 2001)), or PRODUCTION OF ANTHOCYANIN PIGMENT 1 (PAP1) and C1/R, steering anthocyanin biosynthesis in *Arabidopsis* (*Arabidopsis thaliana*) and maize (*Zea mays*), respectively (reviewed in (Dubos et al., 2010; Petroni and Tonelli, 2011)), raised expectations tremendously. This review will list the endeavours of the quests conducted since then in a range of different plant species (Table 1) and will discuss the remaining challenges to map and exploit the transcriptional machineries that drive JA-mediated elicitation of secondary metabolism. We focus on the effect of JA-modulated TFs and their role in JA-mediated secondary metabolism in the model plant *Arabidopsis* and a range of medicinal plant species with structurally divergent metabolites. Finally, we also speculate on possible analogy between the elicitation of secondary metabolism in plants and fungi.

Table 1. Overview of TFs recruited by JA signalling to steer secondary metabolite biosynthesis

TF name	Accession Number	Plant species	Secondary metabolite	Reference
AP2/ERF				
ORCA2	AJ238740	*Catharanthus roseus*	Terpenoid indole alkaloids	(Menke et al., 1999)
ORCA3	EU072424	*C. roseus*	Terpenoid indole alkaloids	(van der Fits and Memelink, 2000)
ERF189	-	*Nicotiana tabacum*	Nicotine	(Shoji et al., 2010)
ERF221/ORC1	CQ808982	*N. tabacum*	Nicotine	(De Sutter et al., 2005; Shoji et al., 2010)
ERF1	JN162091	*Artemisia annua*	Artemisinin	(Yu et al., 2012)
ERF2	JN162092	*A. annua*	Artemisinin	(Osbourn, 2010)
bHLH				
MYC2/At1g32640	NM_102998	*Arabidopsis thaliana*	Indole glucosinolates and anthocyanins	(Dombrecht et al., 2007)
MYC2	AF283507	*C. roseus*	Terpenoid indole alkaloids	(Zhang et al., 2011)
MYC2a	HM466974	*N. tabacum*	Nicotine	(Zhang et al., 2012)
MYC2b	HM466975	*N. tabacum*	Nicotine	(Zhang et al., 2012)
NbbHLH1	GQ859152	*Nicotiana benthamiana*	Nicotine	(Todd et al., 2010)
NbbHLH2	GQ859153	*N. benthamiana*	Nicotine	(Todd et al., 2010)
GL3/At5g41315	NM_148067	*A. thaliana*	Anthocyanins	(Qi et al., 2011)
EGL3/At1g63650	NM_105042	*A. thaliana*	Anthocyanins	(Qi et al., 2011)
TT8/At4g09820	NM_117050	*A. thaliana*	Anthocyanins	(Qi et al., 2011)
R2R3-MYB				
PAP1/At1g56650	NM_104541	*A. thaliana*	Anthocyanins	(Qi et al., 2011)
MYB14	DQ399056	*Pinus taeda*	Flavonoids and Isoprenoids	(Hirai et al., 2007)
MYB29/At5g07690	NM_120851	*A. thaliana*	Aliphatic glucosinolates	(Bedon et al., 2010)
MYBJS1	AB236951	*N. tabacum*	Phenylpropanoids	(Gális et al., 2006)
MYB8	GU451752	*N. attenuata*	Phenylpropanoids	(Onkokesung et al., 2012)
WRKY				
WRKY1	AY507929	*Gossypium arboretum*	Gossypol	(Xu et al., 2004)
WRKY1	FJ390842	*A. annua*	Artemisinin	(Ma et al., 2009)
WRKY1	HQ646368	*C. roseus*	Terpenoid indole alkaloids	(Suttipanta et al., 2011)
WRKY3	AY456271	*N. attenuata*	Volatile terpenes	(Skibbe et al., 2008)
WRKY6	AY456272	*N. attenuata*	Volatile terpenes	(Skibbe et al., 2008)
WRKY33/At2g38470	NM_129404	*A. thaliana*	Camalexin	(Mao et al., 2011)
NAC				
ANAC042/ At2g43000	NM_129861	*A. thaliana*	Camalexin	(Saga et al., 2012)
DOF				
OBP2/Dof1.1/At1g07640	NM_001035911	*A. thaliana*	Indole glucosinolates	(Skirycz et al., 2006)
DOF4;2/At4g21030	NM_118221	*A. thaliana*	Flavonoids	(Skirycz et al., 2007)

HD-ZIP				
HAHB4	AF339748	*H. annuus*	Green leaf volatiles	(Manavella et al., 2008)
TFIIIA zinc finger				
ZCT1	AJ632082	*C. roseus*	Terpenoid indole alkaloids	(Pauw et al., 2004)
ZCT2	AJ632083	*C. roseus*	Terpenoid indole alkaloids	(Pauw et al., 2004)
ZCT3	AJ632084	*C. roseus*	Terpenoid indole alkaloids	(Pauw et al., 2004)

2 Oxylipins: ancient signals of distress?

JAs are oxylipins that originate from free radical oxidation of lipids (Wasternack, 2007). Oxylipins have been suggested to be ancient signals of tissue damage that activate general stress response pathways (Mueller, 2004). This is indeed the case in plants, where JAs elicit production of secondary metabolites that serve a vital role in the interaction between a plant and its surrounding environment and are important players in the constitutive and/or inducible plant defences against a wide variety of attackers. Oxylipins have also been discovered in prokaryotes, mosses, fungi, algae, and invertebrate and vertebrate animals (Andreou et al., 2009). Eicosanoids for instance, which the human prostaglandins also belong to, are biologically important oxylipins that function as signalling molecules in eukaryotic microbes, and invertebrate and vertebrate animals. In insects, eicosanoids mediate cellular immunity to microbial and metazoan challenge, and act in the response to infection (Stanley, 2005). Mammalian eicosanoids have crucial functions in the inflammatory process, both in allergic responses and in general reactions to infection (Das, 2011). In the model fungus *Aspergillus nidulans*, the 'psi factor', which represents a mixture of different oxylipins, is reported to regulate fungal reproduction and secondary metabolite synthesis. Fungi produce a wide variety of secondary metabolites that can either be beneficial (e.g. antibiotics) or detrimental (e.g. mycotoxins) to humans (Yu and Keller, 2005). *A. nidulans* mutants with defects in the psi factor biosynthesis genes (the *ppo* genes) have been shown to manifest a decreased expression of the mycotoxin biosynthesis genes and are unable to produce mycotoxins. This indicates that, as in plants, fungal oxylipins are capable of regulating secondary metabolite production at the transcriptional level. The *ppo* genes are widespread in saprophytic and pathogenic Asco- and Basidiomycetes (Tsitsigiannis and Keller, 2007), underscoring the evolutionary importance of oxylipins in this process.

3 Core JA module?

Despite the conserved importance of oxylipins in the elicitation of secondary metabolism, to our knowledge, no molecular overlaps have yet been encountered in

the plant and fungal oxylipin signalling machineries. Essential in the 'core JA signalling module' in plants (Figure 1) is the F-box protein CORONATINE INSENSITIVE 1 (COI1), which is part of a Skp/Cullin/F-box-type E3 ubiquitin ligase complex (SCF^{COI1}), to which it provides substrate specificity. The targets of the SCF^{COI1} complex are the JA ZIM domain (JAZ) family of repressor proteins. JAZ and COI1 proteins directly interact in the presence of the bioactive JA-isoleucine (JA-Ile) conjugate to form a co-receptor complex. Although ubiquitination of the JAZ proteins by SCF^{COI1} remains to be proven, this interaction ultimately triggers the degradation of the JAZ proteins by the 26S proteasome (Browse, 2009; Chini et al., 2009a; Memelink, 2009; Pauwels et al., 2009). The JAZ proteins contain a highly conserved TIFY motif within the ZIM domain (Vanholme et al., 2007) that mediates homo- and heterodimeric interactions between different JAZ proteins (Chini et al., 2009b; Chung and Howe, 2009). The ZIM domain also functions to recruit transcriptional co-repressors, such as TOPLESS (TPL), through the Novel Interactor of JAZ (NINJA) protein (Pauwels et al., 2010). The JAZ proteins are further characterized by a conserved C-terminal Jas domain, which is required for the interaction with both COI1 and a broad array of TFs (reviewed in (Pauwels and Goossens, 2011)). JA-triggered JAZ degradation releases these TFs, which each modulate expression of specific sets of JA-responsive genes and thereby the production of specific sets of secondary metabolites (Figure 1).

In *Arabidopsis*, the basic helix–loop–helix (bHLH) factor MYC2 is the best known target of the JAZ proteins (Figure 1 and Figure 2a). MYC2 has been shown to be both directly and indirectly involved in regulating secondary metabolite induction. It positively regulates TFs and biosynthetic enzymes of flavonoid biosynthesis but negatively controls tryptophan (Trp)-derived indole glucosinolate synthesis (Dombrecht et al., 2007). The *C. roseus* MYC2 homologue regulates the expression of the *ORCA* TFs by direct binding to the 'on/off switch' in the promoter of the *ORCA3* gene, and thereby controlling expression of several TIA biosynthesis genes (Zhang et al., 2011) (see below, Figure 2b). In common tobacco (*Nicotiana tabacum*), MYC2a/b proteins upregulate the *ORCA*-related *NIC2* locus APETALA2/ETHYLENE Response Factor (AP2/ERF) TFs that regulate nicotine biosynthesis (see below, Figure 2c). In parallel, they also directly bind the target promoters of several nicotine biosynthesis genes (Shoji and Hashimoto, 2011; Zhang et al., 2012). Accordingly, co-expression with the MYC2 TFs stimulated the functionality of at least one of these *NIC2* locus AP2/ERFs, whereas co-expression with the JAZ proteins reduced it (De Boer et al., 2011). In the related species *N.*

benthamiana, the MYC2-homologues NbbHLH1 and NbbHLH2 also function as positive regulators in the JA-mediated activation of nicotine biosynthesis (Todd et al., 2010).

JAZ proteins also directly interact with other TFs with a well-established role in the synthesis of secondary metabolites, such as the bHLH TFs GLABRA3 (GL3), ENHANCER OF GL3 (EGL3) and TRANSPARENT TESTA8 (TT8), and the R2R3-MYB TF PAP1 [28], which together compose transcriptional activator complexes that control anthocyanin biosynthesis and are conserved in the plant kingdom (reviewed in (Dubos et al., 2010; Petroni and Tonelli, 2011) (Figure 1 and Figure 2a). Analogous to the MYC-type bHLHs, the JAZ proteins repress the activity of these TFs and JAs elevate this repression in a *COI1*-dependent manner (Qi et al., 2011).

These examples indicate that the highly conserved COI1-JAZ co-receptor complex is central in the JA-mediated metabolic reprogramming in a variety of plant species. As well as direct JAZ interactors, many more TFs with a proven role in JA-mediated elicitation of a specific metabolic pathway exist (Table 1). Insights into the molecular mechanisms that govern the link between the conserved module and the plethora of (species-specific) regulators are increasing, but the full picture on how the central module exerts control over evolutionary distant metabolic pathways, leading to natural products of a wide structural variety, is still lacking.

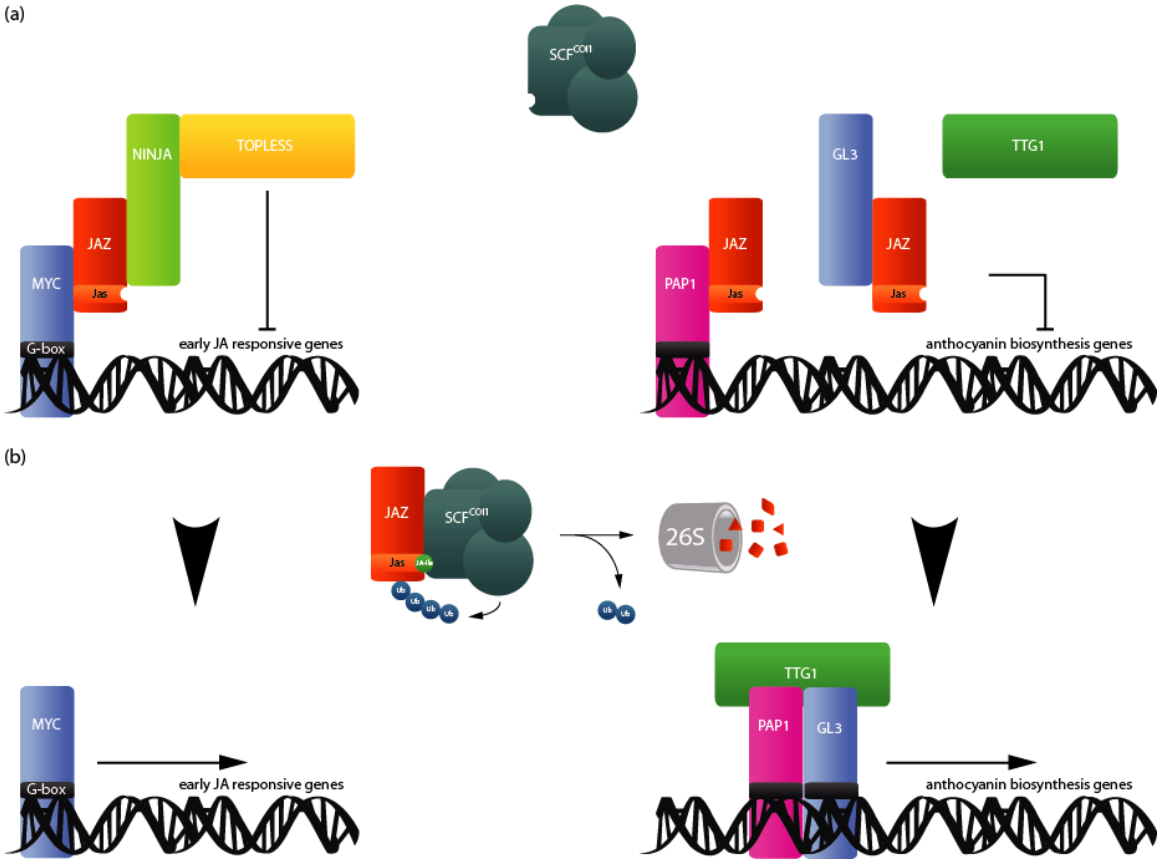


Figure 1. The jasmonate (JA) perception and early signalling modules that elicit secondary metabolism in plants. **(a)** In the absence of JA-isoleucine (JA-Ile), JA ZIM domain (JAZ) proteins interact with co-repressor complexes containing Novel Interactor of JAZ (NINJA) and/or TOPLESS. Through binding with transcription factors (TFs), such as MYC2, GLABRA3 (GL3) and PRODUCTION OF ANTHOCYANIN PIGMENT 1 (PAP1), the JAZ proteins block TF activity and repress JA-responsive gene expression and anthocyanin biosynthesis. **(b)** Upon stress or developmental cues (i.e. when JA-Ile levels rise), the bioactive hormone binds the CORONATINE INSENSITIVE 1 (COI1) receptor in the Skp/Cullin/F-box-type E3 ubiquitin ligase (SCF^{COI1}) complex, thereby recruiting the JAZ proteins, targeting them for their degradation by the 26S proteasome, and ultimately leading to the release of the TFs that can modulate the expression of JA-responsive and anthocyanin synthesis genes. Abbreviation: TTG1, TRANSPARENT TESTA GLABRA 1.

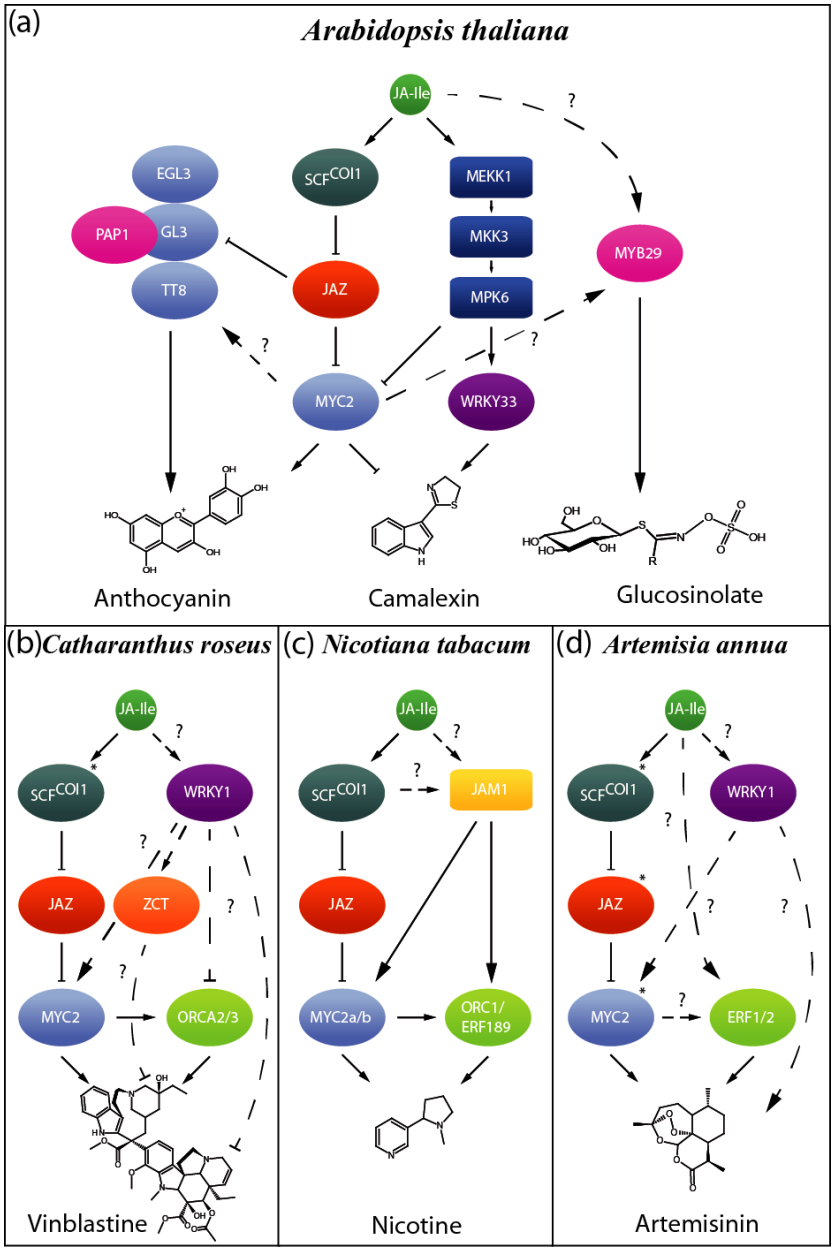


Figure 2. Transcriptional networks that regulate secondary metabolism in model and medicinal plants. Jasmonate (JA)-modulated regulation of: **(a)** anthocyanin, camalexin and glucosinolate synthesis in *Arabidopsis thaliana*; **(b)** terpenoid indole alkaloid synthesis in *Catharanthus roseus*; **(c)** nicotine synthesis in *Nicotiana tabacum*; and **(d)** artemisinin synthesis in *Artemisia annua.* Solid and broken lines indicate proven and hypothetical (yet to be experimentally established) links, respectively, that can be either direct or indirect. Arrows indicate positive interactions; T-bars indicate negative interactions. Asterisks indicate that the identity (i.e. sequence) of the COI1, JAZ or MYC2 proteins from a given species has not yet been determined. Abbreviations: EGL3, ENHANCER OF GL3; JAM1, JA factor stimulating MAPKK 1; MEKK1, MAPK/ERK KINASE KINASE 1; ORCA2/3, OCTADECANOID-DERIVATIVE RESPONSIVE CATHARANTHUS AP2-DOMAIN 2 and 3; TT8, TRANSPARENT TESTA 8.

4 Importance of MYB factors in secondary metabolite biosynthesis

MYB TFs are characterized by one or more copies of a highly conserved MYB DNA-binding domain that consists of imperfect sequence repeats of about 52 amino acids

(aa) (Feller et al., 2011). The largest subfamily, the R2R3-MYBs, has members involved in the regulation of diverse metabolic pathways, including many phenolics, such as anthocyanins, proanthocyanidins, flavonols, lignins and volatile benzenoids, in a wide range of different plant species (Dubos et al., 2010; Spitzer-Rimon et al., 2010; Feller et al., 2011; Petroni and Tonelli, 2011). Some of these R2R3-MYBs are JA-responsive, such as PAP1, which regulates the expression of anthocyanin biosynthesis genes, and thereby induces anthocyanin accumulation in *Arabidopsis* (Borevitz et al., 2000; Shan et al., 2009) (Figure 2a). It has been well established that R2R3 MYB proteins interact and exert a combinatorial regulation with bHLH TFs to activate phenolic biosynthesis. Within these protein complexes, the R2R3-MYBs confer the specificity for the downstream effects and these interactions seem to be conserved across the plant kingdom (Feller et al., 2011; Petroni and Tonelli, 2011). For instance, in *Arabidopsis*, the bHLH TFs TT8, GL3 and EGL3 can all interact with PAP1. JAs can affect the abundance and activity of these bHLH and MYB proteins, both at the transcriptional and post-translational level, through induced expression of the corresponding TF genes and interaction with the JAZ proteins, respectively (Maes et al., 2008; Qi et al., 2011) (Figure 1 and Figure 2a). The latter regulatory aspect raises another intriguing question. Earlier to the finding that JA treatment can induce the activity of the 'PAP1 complex' through the depletion of the inhibitory JAZ proteins (Qi et al., 2011), it was demonstrated that also the R1-MYB protein MYBL2 is a strong negative regulator of anthocyanin biosynthesis (Dubos et al., 2008; Matsui et al., 2008). It remains to be determined whether *MYBL2* expression is also regulated by JAs or whether this repressor protein is redundant to the JAZ, and acts concomitantly (e.g. in the same organs) and/or additively with the JAZ proteins to block PAP1 activity. A similar reflection can be made on the role of the R1 MYB proteins TRIPTYCHON (TRY), CAPRICE (CPC), ENHANCER OF TRY AND CPC1 (ETC1), and ETC2 that negatively regulate trichome formation, another JA-modulated process that is dependent on the TT8, GL3 and EGL3 bHLH TFs, and the expression of which has already been demonstrated to be influenced by JA, either in a positive (CPC) or negative manner (TRY, ETC1, and ETC2) (Maes et al., 2008).

In *Arabidopsis*, other R2R3 MYB TFs involved in phenolic synthesis have been described, such as MYB11, MYB12 and MYB111 for flavonols or MYB123 for proanthocyanidin biosynthesis (Baudry et al., 2004; Stracke et al., 2007; Dubos et al., 2010). However, so far, a possible link with the JA response awaits further characterization. By contrast, in tobacco, the JA-inducible R2R3 MYB, MYBJS1, was shown to induce phenylpropanoid biosynthetic genes and the accumulation of

phenylpropanoid-polyamine conjugates during stress (Gális et al., 2006). A related MYB, MYB8, was recently found to control inducible phenolamide levels in *Nicotiana attenuata* (Onkokesung et al., 2012).

Besides phenolic compounds, MYB TFs control the biosynthesis of another important class of secondary metabolites in *Arabidopsis*, the glucosinolates (GLS). MYB34, MYB51 and MYB122 control the indole GLS pathway and MYB28, MYB29 and MYB76 regulate the aliphatic GLS pathway (Gigolashvili et al., 2007b; Gigolashvili et al., 2007a; Hirai et al., 2007; Gigolashvili et al., 2008; Dubos et al., 2010). At least one of them, MYB29, has been shown to play an accessory role in the JA-mediated elicitation of aliphatic GLS synthesis (Hirai et al., 2007) (Figure 2a). Finally, a JA-responsive R2R3 MYB TF from the loblolly pine (*Pinus taeda*), MYB14, was identified as a putative regulator of a broad defence response implicating flavonoids and isoprenoids. Overexpression of pine *MYB14* in transgenic white spruce (*Picea glauca*) has been shown to impact the terpenoid-, flavonoid-, and JA-related transcriptome and to stimulate terpene and anthocyanin accumulation (Bedon et al., 2010).

5 Importance of AP2/ERFs in JA-modulated alkaloid biosynthesis

The AP2/ERF TFs are characterized by their AP2/ERF DNA-binding domain and several members of the ERF-subfamily have proven roles in JA-responsive gene expression (Memelink, 2009). The previously mentioned Madagascar periwinkle JA-responsive ORCA TFs, in particular ORCA2 and ORCA3, are the most renown members of the subfamily and interact with the promoters of their target genes (e.g. strictosidine synthase, *Str*) via sequence-specific binding to the GCC-element, a hallmark of the ERF-subfamily (Memelink et al., 2001; Van Der Fits and Memelink, 2001; Memelink, 2009; De Boer et al., 2011) (Figure 2b). As such, ORCA3 controls expression of multiple genes encoding enzymes involved in all branches of the TIA pathway, including the primary plastidial isopentenyl pyrophosphate pathway and the periwinkle-specific secondary TIA pathways. However, because not all TIA pathway genes are under ORCA3 control, *ORCA3* overexpression was not sufficient to elicit TIA synthesis in transgenic *C. roseus* cells (van der Fits and Memelink, 2000). It remains to be determined whether the target genes of ORCA2 might be different from those of ORCA3 or whether they might act redundantly (Memelink et al., 2001).

Transcript profiling of tobacco mutants with deficient nicotine biosynthesis (*nic* mutants) revealed at least seven ERF genes involved in nicotine biosynthesis, including *ERF189* and *ORC1/ERF221*, which cluster together in the *NIC2* locus and

are upregulated by JA elicitation (Figure 2c). These TFs specifically activate all known structural genes in the nicotine pathway and at least one TF, ERF189, can recognize a GCC-box element in the promoter of the gene encoding the enzyme that catalyses the first committed step in nicotine biosynthesis, putrescine *N*-methyltransferase (PMT). Correspondingly, *ERF189* overexpression increases nicotine biosynthesis gene expression in transgenic hairy roots (Shoji et al., 2010) (Figure 2c). ORC1 is a close homolog of *C. roseus* ORCA3 and its overexpression stimulated alkaloid biosynthesis in stably transformed tobacco plants and tree tobacco (*Nicotiana glauca*) root cultures (De Boer et al., 2011). Furthermore, the activity of both ORC1 and the MYC-type bHLH proteins can be post-translationally upregulated by a JA-modulated phosphorylation cascade, in which a specific mitogen-activated protein kinase kinase, JA factor stimulating MAPKK 1 (JAM1), is active (De Boer et al., 2011) (Figure 2c).

It is likely that more AP2/ERF TFs involved in JA-elicited secondary metabolite production will be revealed. Recently, two JA-responsive ERF TFs were isolated from *Artemisia annua*, ERF1 and ERF2, which can bind and transactivate the promoters of the genes encoding amorpha-4,11-diene synthase (ADS) and a cytochrome P450 monooxygenase (CYP71AV1), which are both involved in the biosynthesis of the antimalarial sesquiterpene lactone artemisinin (Figure 2d). Transgenic *A. annua* plants overexpressing either TF showed elevated expression of both synthesis genes and an increased accumulation of artemisinin and precursors thereof (Yu et al., 2012). The finding that phylogenetically distant plant species have all recruited closely related TF genes to control expression of JA-inducible enzymes catalysing their respective specific metabolic pathways, further corroborates the hypothesis that the JA elicitor signalling machinery seems to be conserved and installed early in the higher plant lineage and evolved to control evolutionary distinct secondary metabolic pathways. This hypothesis awaits further confirmation while other (medicinal) plant species are being investigated.

6 Other JA-responsive TFs that activate secondary metabolite biosynthesis

Transcriptional regulators of two other TF families have been reported to be involved in transcriptional reprogramming of secondary metabolite pathways in a JA-inducible manner, but their exact position in the JA signalling cascades and/or their interaction with the JA core module remains unclear.

The WRKY TFs, characterized by their highly conserved 60-aa long WRKY domain, composed of the conserved heptapeptide sequence WRKYGQK, have been implicated in a diverse range of stress tolerance and development programmes (Rushton et al., 2010; Agarwal et al., 2011). Several WRKY TFs may regulate secondary metabolism biosynthesis in response to JA elicitation as suggested by their (fast) upregulation by JA treatment. Examples include: (i) cotton (*Gossypium arboretum*) WRKY1, which can transactivate the promoter of the (1)-δ-cadinene synthase (*CAD1*) gene, and so might participate in the regulation of sesquiterpene phytoalexin biosynthesis (Xu et al., 2004); (ii) *A. annua* WRKY1, which can transactivate the promoter of the *ADS* gene, and so might participate in the regulation of artemisinin synthesis (Ma et al., 2009) (Figure 2d); and (iii) Madagascar periwinkle WRKY1, which may participate in the regulation of TIA biosynthesis through an as yet undefined manner (Suttipanta et al., 2011) (Figure 2b). In all cases, the WRKY TFs were shown to bind the W-box in the promoters of the respective biosynthetic genes. Furthermore, JA-responsive WRKY factors may also regulate accumulation of lignin or other phenolics in rice (*Oryza sativa*), *Medicago truncatula* and tobacco (Wang et al., 2007; Naoumkina et al., 2008).

Some WRKY genes may enable cross-talk between JA and other hormone or stress response signalling pathways, and thereby provide additional ways to modulate secondary metabolite synthesis. WRKY33, a pathogen-inducible TF, functions downstream of two pathogen-responsive mitogen-activated protein kinases, MPK3 and MPK6, to change the expression of camalexin biosynthetic genes, and to drive production of camalexin, the major phytoalexin in *Arabidopsis*. The MPK3/MPK6 phosphorylation cascade regulates both *WRKY33* expression and WRKY33 activity (Mao et al., 2011). MPK6 is a downstream target of a MAPK kinase, MKK3, and the MKK3/MPK6 cascade can be activated in response to JA and regulate, among others, *MYC2* expression (Takahashi et al., 2007) (Figure 2a). Silencing of two insect-responsive (but not JA-responsive) WRKY genes from the native tobacco *Nicotiana attenuata*, *WRKY3* and *WRKY6*, makes plants highly vulnerable to herbivores by impairing JA accumulation and synthesis of sesquiterpene volatiles such as *cis*-α-bergamotene (Skibbe et al., 2008).

Recently, induction of camalexin synthesis was also demonstrated to be positively controlled by ANAC042, a member of the NAM, ATAF1/2 and CUC2 (NAC) TFs. ANAC042-mediated control of camalexin synthesis likely occurs via transcriptional regulation of the genes encoding the cytochrome P450 proteins CYP71A12, CYP71A13, and CYP71B15/PAD3 in an unknown signalling pathway distinct from

the WRKY33 pathway. *ANAC042* expression can be downregulated by JA treatment, enabling JA-modulated control of camalexin synthesis (Saga et al., 2012).

To date, two members of the plant-specific DNA-binding with one finger (Dof) family of TFs have been characterized as JA-modulated regulators of secondary metabolism. The Dof TFs carry a highly conserved DNA-binding domain that is thought to include a single C_2–C_2 zinc finger (Yanagisawa, 2002, 2004). The JA-inducible OBP2/DOF1.1 plays a positive role in mediating indole glucosinolate biosynthesis in *Arabidopsis* (Skirycz et al., 2006), whereas the JA-repressed DOF4;2 influences *Arabidopsis* phenylpropanoid metabolism in an environmental and tissue-specific manner (Skirycz et al., 2007).

7 Positive and negative feedback loops boost the system but not at all cost

A tight and coordinated control of hormone biosynthesis and signalling is required to fine-tune the broad effects that hormones have. In the case of JAs, and more specifically for JA-mediated induction of secondary metabolism, a strong and rapid induction is vital, particularly during defence responses. The JA signal needs to persist or even to intensify as long as the plant is under attack. As a consequence, the plants have evolved a positive feedback system with loops at various control points, termed the autoregulatory JA loop (Wasternack, 2007). First, a control point is situated at the genes encoding JA biosynthesis enzymes that are all JA-inducible and controlled by JA-responsive TFs, such as MYC2 and ORA47, allowing bioactive JA synthesis to be boosted (Wasternack, 2007; Pauwels et al., 2008; Pauwels et al., 2009). Second, many genes encoding TFs involved in the primary JA response, such as MYC2, are themselves rapidly induced by the same signal and can modulate their own expression (Dombrecht et al., 2007; Pauwels et al., 2009). Third, other sets of regulators, such as the TEOSINTE BRANCHED/CYCLOIDEA/PCF (TCP) TFs, and the auxin response factors ARF6 and ARF8, can positively influence JA biosynthesis in a developmental context (Nagpal et al., 2005; Schommer et al., 2008). The machineries involved in the autoregulatory JA loop have mainly been studied in *Arabidopsis*, but the phenomenon is conserved among plants (Wasternack, 2007; Pauwels et al., 2009). Correspondingly, a member of the sunflower (*Helianthus annuus*) homeodomain-leucine zipper (HD-ZIP) subfamily, HAHB4, was identified as a positive regulator of the synthesis of JAs and green leaf volatiles (GLVs), in the defence responses against (a)biotic stresses. HAHB4 upregulates the transcript levels

of several genes involved in JA and GLV biosynthesis and *HAHB4* expression itself is stimulated by JAs (Manavella et al., 2008).

In addition to positive loops, negative feedback is also needed to shut down plant defence and stress responses when they are not needed, to avoid energy being wasted. Multiple regulatory mechanisms have been developed to keep such energy-consuming responses silent during normal conditions. As a result, negative regulators are key components in the control of stress-related gene expression. ERF-associated amphiphilic repression (EAR) domain-containing proteins have been identified as characteristic elements of transcriptional repression of gene expression in plants (Kazan, 2006). Some EAR proteins have been shown to be active in metabolic pathways. In *C. roseus*, three members of the EAR-domain containing TF IIIA-type zinc finger protein family, ZCT1, ZCT2 and ZCT3, were reported to bind the promoters of the *STR* and *Tryptophan decarboxylase* (*TDC*) genes and thereby repress their expression, at least in transient promoter activity assays (Pauw et al., 2004) (Figure 2b). Acting early and conserved in the JA signalling pathway, the NINJA protein, as well as some of the JAZ proteins themselves (e.g. JAZ5 and JAZ8), contain an EAR motif, which interacts with TPL to repress the activity of JAZ-bound TFs (Kagale et al., 2010; Pauwels et al., 2010) (Arabidopsis Interactome Mapping Consortium, 2011). Upon JA perception, expression of JAZ, NINJA and ZCT genes is increased (notably, in the case of the JAZ genes by MYC2 itself) to produce negative feedback loops, which guarantee the instalment of balanced defence responses (Pauw et al., 2004; Chini et al., 2007; Pauwels and Goossens, 2008; Pauwels et al., 2009; Kagale et al., 2010; Pauwels et al., 2010; Figueroa and Browse, 2012).

8 Concluding remarks and future prospects

JAs have an evolutionarily conserved role in the reprogramming of plant secondary metabolism in response to various environmental or developmental stimuli. An important aspect herein is the concerted transcriptional activation of the genes encoding the enzymes that catalyse the secondary metabolic reactions. Often, JAs simultaneously induce all known biosynthetic genes from a particular pathway, as illustrated by the TIA, nicotine and artemisinin pathways in *C. roseus*, *N. tabacum* and *A. annua*, respectively (van der Fits and Memelink, 2000, 2001; Maes et al., 2011). The discovery of arrays of TFs that are activated early during JA elicitation triggered the hope of finding master regulators that could be used to boost the production of specific sets of valuable natural products. Although overexpression of several discussed TFs could stimulate synthesis of some secondary metabolites, no

master switches have been found that can mimic the full JA spectrum, neither quantitatively nor qualitatively, or replace JAs in plant engineering programmes.

The concerted transcriptional activation of whole pathways by JAs is not necessarily evoked by the action of a single TF. On the contrary, a combinatorial role for several TFs in the regulation of different enzymes or suites of enzymes seems currently more plausible to account for the control of biosynthetic pathways by JAs. These TFs can either belong to the same or different families and/or be subjected to different environmental or developmental cues. A classic example is the finding that ORCA3 controls the expression of the genes encoding the TDC, STR, CytP450 reductase (CPR) and desacetoxyvindoline 4-hydroxylase (D4H) enzymes but not those corresponding to the geraniol 10-hydroxylase (G10H), strictosidine β-d-glucosidase (SGD) and acetyl-CoA:4-*O*-deacetylvindoline 4-*O*-acetyltransferase (DAT) enzymes involved in the *C. roseus* TIA pathway (van der Fits and Memelink, 2000; Memelink et al., 2001). TF(s) that regulate expression of the latter three genes still await discovery. Combinatorial action of TFs has already been demonstrated in the JA-mediated elicitation of tobacco nicotine biosynthesis, which involves and requires the concerted action of AP2/ERF and bHLH factors (De Boer et al., 2011), and likely additional, but yet unknown, TFs that determine the cell and organ specificity of the nicotine synthesis pathway.

Indeed, biosynthesis of secondary metabolites is spatially often very strictly regulated. For instance, nicotine and artemisinin biosynthesis occur exclusively in specific cell layers of tobacco roots and *A. annua* trichomes, respectively, even after JA elicitation. Usually, this correlates well with the expression patterns of the genes encoding the enzymes that catalyse for instance committed steps in the pathway, such as of PMT in nicotine and ADS in artemisinin synthesis, respectively (Shoji et al., 2000; Kim et al., 2008). However, little is known on the role of the JA-modulated TFs in the determination of this tissue or organ specificity. On the contrary, none of the TFs known to be involved in the regulation of nicotine and artemisinin biosynthesis (Table 1) are expressed exclusively in the roots or trichomes, respectively, and many of them seem to be ubiquitously expressed throughout the plants. Similarly, JA-mediated JAZ degradation and subsequent induction of TF expression occurs ubiquitously in the plant, suggesting that other, perhaps non JA-modulated TFs or other regulatory mechanisms are also at play.

It is now clear that the regulation of plant secondary metabolism constitutes more than just an on/off-switch but rather that it is subjected to complex control mechanisms integrated in robust cellular networks. This then leads on to the question of how future

gene discovery programmes should be designed to enable an increase in the understanding of plant metabolism and the generation of new generic tools for plant metabolic engineering. The booming of functional genomics technologies that increase the resolution and coverage of genome, transcriptome, proteome, interactome, as well as metabolome analysis offers unprecedented ways of listing all the possible players involved in the regulation of plant metabolism. Ever more important will be the design of original screens that not only reveal the identity of as yet unknown regulators, but also detect as yet unknown regulatory mechanisms. A nice example of the latter was the recent finding that JAZ proteins not only directly interact with the MYC-type TFs, but also with the related bHLH proteins TT8, GL3 and EGL3, and the R2R3 MYB protein PAP1 (Qi et al., 2011). As such, JAs exert post-translational control over anthocyanin accumulation, as JA-induced degradation of JAZ proteins abolishes the interactions between the JAZ and the bHLH and MYB factors, which in turn releases the latter to activate the downstream signal cascades that trigger anthocyanin accumulation. Past and ongoing screens suggest that JAZ can interact with a broad array of TFs that each control specific downstream processes (reviewed in (Pauwels and Goossens, 2011)).

Screens for TFs that transactivate particular biosynthesis genes or for proteins that interact with, and thereby modulate the activity of, known activator TFs, are likely to keep delivering new (transcriptional) regulators of plant secondary metabolism. As a result, a more in-depth view on the JA signalling cascade might be obtained and factors might be found that specifically control one (or more) secondary metabolic pathway(s), and do not affect plant viability and growth when overexpressed in transgenic plants or plant cultures, an undesired but frequently occurring 'side-effect' for many of the known JA-responsive TFs. However, inspiration for future screens should also be sought beyond these boundaries, in the biological context of JA-triggered secondary metabolite synthesis, for instance. Plants produce species-specific bioactive or protective compounds to face particular acute biotic or abiotic stresses or to ensure an appropriate fitness-cost response when they experience prolonged stress periods. These responses not only demand the action of JAs, but also that of many other hormones that crosstalk with the JAs, such as abscisic acid (ABA), ethylene or salicylic acid. Hence, screens for hormonal crosstalk points might reveal new checkpoints that control the reprogramming of plant metabolism, both quantitatively and/or qualitatively, as illustrated by a recent study on the interaction between JA and ABA (Lackman et al., 2011).

Similarly, we can also learn from evolution and signalling in other, non-plant organisms. As mentioned above, oxylipins also regulate fungal secondary metabolism (Yu and Keller, 2005; Tsitsigiannis and Keller, 2007) but the exact mechanisms by which this occurs remain to be determined. Hallmarks of fungal secondary metabolism are the gene clusters that contain cluster-specific TFs functioning to co-activate the biosynthetic genes present in their respective cluster (Palmer and Keller, 2010). These TFs can be (de)activated in response to a variety of environmental stimuli such as light, pH, nutrients and temperature via signal transduction cascades that involve other TFs with a broader action range. However, an important aspect in the control of fungal secondary metabolism is dependent on the 'locality', that is, on the chromosomal location of the gene cluster because of the action of a conserved global regulator complex, the velvet complex, which is involved in chromatin remodelling at the cluster loci (Palmer and Keller, 2010) (Figure 3a). Interestingly, examples of secondary metabolic gene clusters have now also been discovered in plants (Osbourn, 2010), and cell type-specific chromatin decondensation has been observed for the avenacin gene cluster in oat (*Avena strigosa*), which has provided new insights into the regulation of secondary metabolism in plants (Wegel et al., 2009). Similar chromatin effects may represent the *modus operandi* of the JAZ proteins, which (in)directly recruit co-repressors such as the TPL proteins, which have been linked with histone deacetylases and demethylases (Long et al., 2006; Macrae and Long, 2011) (Figure 3b). Whether the order of events is "chromatin remodelling allowing TF activation or TF binding allowing chromatin remodelling" (Palmer and Keller, 2010), and whether this link between chromatin remodelling and TF activity is a conserved mechanism in oxylipin signalling to control secondary metabolism in eukaryotes, are questions that still need further investigation.

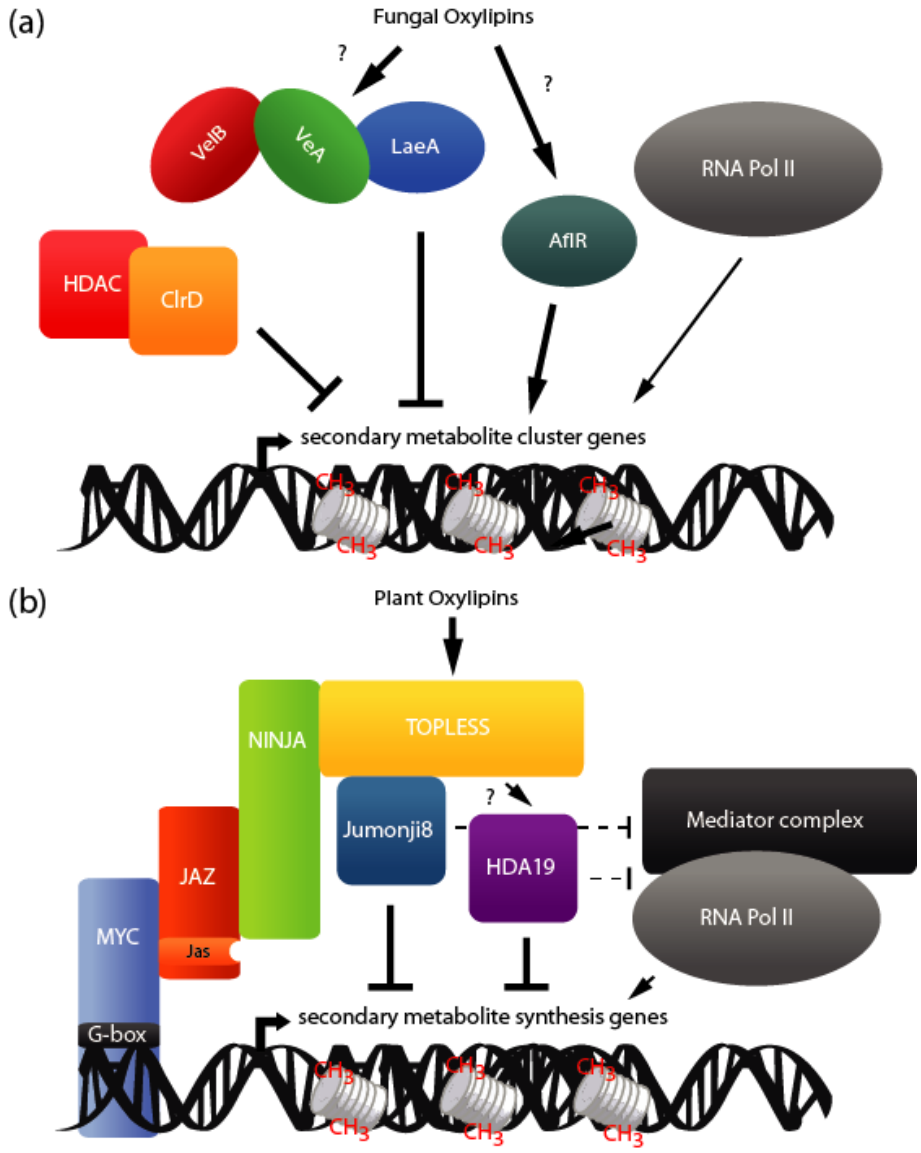


Figure 3. Oxylipin-responsive transcriptional networks that modulate secondary metabolism in plants and fungi. (a) An integrated model for oxylipin- and chromatin-mediated control of secondary metabolite gene clusters in fungi. Environmental stimuli cause production of fungal oxylipins (e.g. the psi factor) that trigger signal transduction cascades towards the production of secondary metabolites. These signals involve the velvet complex (containing the LaeA, VeA and VelB proteins), which activates gene transcription of secondary metabolite cluster genes (including the cluster-specific transcription factor AflR) by mediating chromatin remodelling and facilitating RNA polymerase II (Pol II) action. Conversely, the activity of histone deacetylases (HDAC) and H3K9 methyltransferases (ClrD) is associated with silencing of the gene cluster (Palmer and Keller, 2010). (b) A proposed mechanistic model for jasmonate (JA) ZIM domain (JAZ)-mediated transcriptional repression of the JA response in plants. The JAZ–NINJA-TPL complex represses the transcriptional activity of transcription factors, such as MYC2 (Pauwels et al., 2010). It is postulated that co-repressors such as TPL restrain gene expression by recruiting histone deacetylases (e.g. HDA19) and demethylases (e.g. Jumonji8) to remodel chromatin into a silent state, and/or inhibiting the activity of the RNA Pol II (Long et al., 2006; Macrae and Long, 2011), but the importance of the latter proteins for the regulation of plant secondary metabolism remains to be determined. It has also been postulated that JAZ proteins might directly interact with HDA proteins (Zhu et al., 2011) or use different repression mechanisms for different TF sets (Pauwels and Goossens, 2011), theories that require further investigation in the frame of the regulation of plant secondary metabolism.

Acknowledgements

We thank Annick Bleys and Jacob Pollier for help in preparing the manuscript. Nathan De Geyter and Azra Gholami are indebted to the Agency for Innovation by Science and Technology and the Iranian Ministry for Health and Medical Education, respectively, for predoctoral fellowships. This work was supported by the European FP7 project SmartCell (222716).

References

Agarwal, P., Reddy, M.P., and Chikara, J. (2011). WRKY: its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. Molecular Biology Reports **38,** 3883-3896.

Andreou, A., Brodhun, F., and Feussner, I. (2009). Biosynthesis of oxylipins in non-mammals. Progress in Lipid Research **48,** 148-170.

Arabidopsis Interactome Mapping Consortium. (2011). Evidence for Network Evolution in an Arabidopsis Interactome Map. Science **333,** 601-607.

Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. The Plant Journal **39,** 366-380.

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S.D., Schmidt, A., Gershenzon, J., Grima-Pettenati, J., Séguin, A., and MacKay, J. (2010). Subgroup 4 R2R3-MYBs in conifer trees: gene family expansion and contribution to the isoprenoid- and flavonoid-oriented responses. Journal of Experimental Botany **61,** 3847-3864.

Borevitz, J.O., Xia, Y., Blount, J., Dixon, R.A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. Plant Cell **12,** 2383-2393.

Browse, J. (2009). Jasmonate passes muster: a receptor and targets for the defense hormone. Annual Review of Plant Biology **60,** 183-205.

Chini, A., Boter, M., and Solano, R. (2009a). Plant oxylipins: COI1/JAZs/MYC2 as the core jasmonic acid-signalling module. FEBS Journal **276,** 4682-4692.

Chini, A., Fonseca, S., Chico, J.M., Fernández-Calvo, P., and Solano, R. (2009b). The ZIM domain mediates homo- and heteromeric interactions between Arabidopsis JAZ proteins. The Plant Journal **59,** 77-87.

Chini, A., Fonseca, S., Fern, G., ndez, Adie, B., Chico, J.M., Lorenzo, O., Garc, G., a, C., I, L., pez, V., Lozano, F.M., Ponce, M.R., Micol, J.L., and Solano, R. (2007). The JAZ family of repressors is the missing link in jasmonate signalling. Nature **448,** 666-671.

Chung, H.S., and Howe, G.A. (2009). A critical role for the TIFY motif in repression of jasmonate signaling by a stabilized splice variant of the JASMONATE ZIM-Domain protein JAZ10 in Arabidopsis. Plant Cell **21,** 131-145.

Das, U. (2011). Essential fatty acids and their metabolites as modulators of stem cell biology with reference to inflammation, cancer, and metastasis. Cancer Metastasis Review **30,** 311-324.

De Boer, K., Tilleman, S., Pauwels, L., Vanden Bossche, R., De Sutter, V., Vanderhaeghen, R., Hilson, P., Hamill, J.D., and Goossens, A. (2011). APETALA2/ETHYLENE RESPONSE FACTOR and basic helix–loop–helix tobacco transcription factors cooperatively mediate jasmonate-elicited nicotine biosynthesis. The Plant Journal **66,** 1053-1065.

De Sutter, V., Vanderhaeghen, R., Tilleman, S., Lammertyn, F., Vanhoutte, I., Karimi, M., Inzé, D., Goossens, A., and Hilson, P. (2005). Exploration of jasmonate signalling via automated and standardized transient expression assays in tobacco cells. The Plant Journal **44,** 1065-1076.

Dombrecht, B., Xue, G.P., Sprague, S.J., Kirkegaard, J.A., Ross, J.J., Reid, J.B., Fitt, G.P., Sewelam, N., Schenk, P.M., Manners, J.M., and Kazan, K. (2007). MYC2 differentially modulates diverse jasmonate-dependent functions in Arabidopsis. Plant Cell **19,** 2225-2245.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. Trends In Plant Science **15,** 573-581.

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., Routaboul, J.-M., Alboresi, A., Weisshaar, B., and Lepiniec, L. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. The Plant Journal **55,** 940-953.

Feller, A., Machemer, K., Braun, E.L., and Grotewold, E. (2011). Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. The Plant Journal **66,** 94-116.

Figueroa, P., and Browse, J. (2012). The Arabidopsis JAZ2 promoter contains a G-Box and thymidine-rich module that are necessary and sufficient for jasmonate-dependent activation by MYC transcription factors and repression by JAZ proteins. Plant and Cell Physiology **53,** 330-343.

Gális, I., Šimek, P., Narisawa, T., Sasaki, M., Horiguchi, T., Fukuda, H., and Matsuoka, K. (2006). A novel R2R3 MYB transcription factor NtMYBJS1 is a methyl jasmonate-dependent regulator of phenylpropanoid-conjugate biosynthesis in tobacco. The Plant Journal **46,** 573-592.

Gigolashvili, T., Yatusevich, R., Berger, B., Müller, C., and Flügge, U.-I. (2007a). The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. The Plant Journal **51,** 247-261.

Gigolashvili, T., Engqvist, M., Yatusevich, R., Müller, C., and Flügge, U.I. (2008). HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana*. New Phytologist **177,** 627-642.

Gigolashvili, T., Berger, B., Mock, H.-P., Müller, C., Weisshaar, B., and Flügge, U.-I. (2007b). The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. The Plant Journal **50,** 886-901.

Hirai, M.Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., Araki, R., Sakurai, N., Suzuki, H., Aoki, K., Goda, H., Nishizawa, O.I., Shibata, D., and Saito, K. (2007). Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. Proceeding National Academy of Sciences of the United State of America **104,** 6478-6483.

Kagale, S., Links, M.G., and Rozwadowski, K. (2010). Genome-wide analysis of ethylene-responsive element binding factor-associated amphiphilic repression motif-containing transcriptional regulators in Arabidopsis. Plant Physiology **152,** 1109-1134.

Kazan, K. (2006). Negative regulation of defence and stress genes by EAR-motif-containing repressors. Trends in Plant Science **11,** 109-112.

Kim, S.-H., Chang, Y.-J., and Kim, S.-U. (2008). Tissue specificity and developmental pattern of Amorpha-4,11-diene synthase (ADS) proved by ads promoter-driven GUS expression in the heterologous plant, *Arabidopsis thaliana*. Planta Medica **74,** 188-193.

Lackman, P., González-Guzmán, M., Tilleman, S., Carqueijeiro, I., Pérez, A.C., Moses, T., Seo, M., Kanno, Y., Häkkinen, S.T., Van Montagu, M.C.E., Thevelein, J.M.,

Maaheimo, H., Oksman-Caldentey, K.-M., Rodriguez, P.L., Rischer, H., and Goossens, A. (2011). Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in Arabidopsis and tobacco. Proceedings of the National Academy of Sciences **108,** 5891-5896.

Long, J., Ohno, C., Smith, Z., and Meyerowitz, E. (2006). TOPLESS regulates apical embryonic fate in Arabidopsis. Science (New York) **312,** 1520-1523.

Ma, D., Pu, G., Lei, C., Ma, L., Wang, H., Guo, Y., Chen, J., Du, Z., Wang, H., Li, G., Ye, H., and Liu, B. (2009). Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the Amorpha-4,11-diene synthase gene, a key gene of artemisinin biosynthesis. Plant and Cell Physiology **50,** 2146-2161.

Macrae, R., and Long, J. (2011). Interactions between TOPLESS and histone-modifying enzymes. In 22nd International Conference on Arabidopsis Research, T.N.A.A.S.C. ed), ed (The North American Arabidopsis Steering Committee), pp. 483.

Maes, L., Inzé, D., and Goossens, A. (2008). Functional specialization of the TRANSPARENT TESTA GLABRA1 network allows differential hormonal control of laminal and marginal trichome initiation in Arabidopsis rosette leaves. Plant Physiology **148,** 1453-1464.

Maes, L., Van Nieuwerburgh, F.C.W., Zhang, Y., Reed, D.W., Pollier, J., Vande Casteele, S.R.F., Inzé, D., Covello, P.S., Deforce, D.L.D., and Goossens, A. (2011). Dissection of the phytohormonal regulation of trichome formation and biosynthesis of the antimalarial compound artemisinin in *Artemisia annua* plants. New Phytologist **189,** 176-189.

Manavella, P.A., Dezar, C.A., Bonaventure, G., Baldwin, I.T., and Chan, R.L. (2008). HAHB4, a sunflower HD-Zip protein, integrates signals from the jasmonic acid and ethylene pathways during wounding and biotic stress responses. The Plant Journal **56,** 376-388.

Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., and Zhang, S. (2011). Phosphorylation of a WRKY Transcription Factor by Two Pathogen-Responsive MAPKs Drives Phytoalexin Biosynthesis in Arabidopsis. Plant Cell **23,** 1639-1653.

Matsui, K., Umemura, Y., and Ohme-Takagi, M. (2008). AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. The Plant Journal **55,** 954-967.

Memelink, J. (2009). Regulation of gene expression by jasmonate hormones. Phytochemistry **70,** 1560-1570.

Memelink, J., Verpoorte, R., and Kijne, J.W. (2001). ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism. Trends In Plant Science **6,** 212-219.

Menke, F.L., Champion, A., Kijne, J.W., and Memelink, J. (1999). A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. The EMBO Journal **18,** 4455-4463.

Mueller, M.J. (2004). Archetype signals in plants: the phytoprostanes. Current Opinion in Plant Biology **7,** 441-448.

Nagpal, P., Ellis, C.M., Weber, H., Ploense, S.E., Barkawi, L.S., Guilfoyle, T.J., Hagen, G., Alonso, J.M., Cohen, J.D., Farmer, E.E., Ecker, J.R., and Reed, J.W. (2005). Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. Development **132,** 4107-4118.

Naoumkina, M., He, X., and Dixon, R. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. BMC Plant Biology **8,** 1-14.

Onkokesung, N., Gaquerel, E., Kotkar, H., Kaur, H., Baldwin, I.T., and Galis, I. (2012). MYB8 controls inducible phenolamide levels by activating three novel hydroxycinnamoyl-

coenzyme a:polyamine transferases in *Nicotiana attenuata*. Plant Physiology **158,** 389-407.

Osbourn, A. (2010). Gene clusters for secondary metabolic pathways: an emerging theme in plant biology. Plant Physiology **154,** 531-535.

Palmer, J.M., and Keller, N.P. (2010). Secondary metabolism in fungi: does chromosomal location matter? Current Opinion in Microbiology **13,** 431-436.

Pauw, B., Hilliou, F.A.O., Martin, V.S., Chatel, G., de Wolf, C.J.F., Champion, A., Pré, M., van Duijn, B., Kijne, J.W., van der Fits, L., and Memelink, J. (2004). Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. Journal of Biological Chemistry **279,** 52940-52948.

Pauwels, L., and Goossens, A. (2008). Fine-tuning of early events in the jasmonate response. Plant Signaling & Behavior **3,** 846-847.

Pauwels, L., and Goossens, A. (2011). The JAZ proteins: a crucial interface in the jasmonate signaling cascade. Plant Cell **23,** 3089-3100.

Pauwels, L., Inzé, D., and Goossens, A. (2009). Jasmonate-inducible gene: what does it mean? Trends in Plant Science **14,** 87-91.

Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., Inzé, D., and Goossens, A. (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured Arabidopsis cells. Proceedings of the National Academy of Sciences **105,** 1380-1385.

Pauwels, L., Barbero, G.F., Geerinck, J., Tilleman, S., Grunewald, W., Perez, A.C., Chico, J.M., Bossche, R.V., Sewell, J., Gil, E., Garcia-Casado, G., Witters, E., Inze, D., Long, J.A., De Jaeger, G., Solano, R., and Goossens, A. (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature **464,** 788-791.

Petroni, K., and Tonelli, C. (2011). Recent advances on the regulation of anthocyanin synthesis in reproductive organs. Plant Science **181,** 219-229.

Qi, T., Song, S., Ren, Q., Wu, D., Huang, H., Chen, Y., Fan, M., Peng, W., Ren, C., and Xie, D. (2011). The Jasmonate-ZIM-Domain proteins interact with the WD-Repeat/bHLH/MYB complexes to regulate jasmonate-mediated anthocyanin accumulation and trichome initiation in *Arabidopsis thaliana*. Plant Cell **23,** 1795-1814.

Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J. (2010). WRKY transcription factors. Trends in Plant Science **15,** 247-258.

Saga, H., Ogawa, T., Kai, K., Suzuki, H., Ogata, Y., Sakurai, N., Shibata, D., and Ohta, D. (2012). Identification and characterization of ANAC042, a transcription factor family gene involved in the regulation of camalexin biosynthesis in Arabidopsis. Molecular Plant-Microbe Interactions **25,** 684-696.

Schommer, C., Palatnik, J.F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E.E., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 Targets. PLOS Biology **6,** 1991-2001.

Shan, X., Zhang, Y., Peng, W., Wang, Z., and Xie, D. (2009). Molecular mechanism for jasmonate-induction of anthocyanin accumulation in Arabidopsis. Journal of Experimental Botany **60,** 3849-3860.

Shoji, T., and Hashimoto, T. (2011). Tobacco MYC2 regulates jasmonate-inducible nicotine biosynthesis genes directly and by way of the NIC2-locus ERF genes. Plant and Cell Physiology **52,** 1117-1130.

Shoji, T., Yamada, Y., and Hashimoto, T. (2000). Jasmonate induction of putrescine N-methyltransferase genes in the root of *Nicotiana sylvestris*. Plant and Cell Physiology **41,** 831-839.

Shoji, T., Kajikawa, M., and Hashimoto, T. (2010). clustered transcription factor genes regulate nicotine biosynthesis in tobacco. Plant Cell **22,** 3390-3409.

Skibbe, M., Qu, N., Galis, I., and Baldwin, I.T. (2008). Induced plant defenses in the natural environment: *Nicotiana attenuata* WRKY3 and WRKY6 coordinate responses to herbivory. Plant Cell **20,** 1984-2000.

Skirycz, A., Jozefczuk, S., Stobiecki, M., Muth, D., Zanor, M.I., Witt, I., and Mueller-Roeber, B. (2007). Transcription factor AtDOF4;2 affects phenylpropanoid metabolism in *Arabidopsis thaliana*. New Phytologist **175,** 425-438.

Skirycz, A., Reichelt, M., Burow, M., Birkemeyer, C., Rolcik, J., Kopka, J., Zanor, M.I., Gershenzon, J., Strnad, M., Szopa, J., Mueller-Roeber, B., and Witt, I. (2006). DOF transcription factor AtDof1.1 (OBP2) is part of a regulatory network controlling glucosinolate biosynthesis in Arabidopsis. The Plant Journal **47,** 10-24.

Spitzer-Rimon, B., Marhevka, E., Barkai, O., Marton, I., Edelbaum, O., Masci, T., Prathapani, N.K., Shklarman, E., Ovadis, M., and Vainstein, A. (2010). EOBII, a gene encoding a flower-specific regulator of phenylpropanoid volatiles' biosynthesis in petunia. Plant Cell **22,** 1961-1976.

Stanley, D. (2005). Prostaglandins and other eicosanoids in insects: biological significance. Annual Review of Entomology **51,** 25-44.

Stracke, R., Ishihara, H., Huep, G., Barsch, A., Mehrtens, F., Niehaus, K., and Weisshaar, B. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. The Plant Journal **50,** 660-677.

Suttipanta, N., Pattanaik, S., Kulshrestha, M., Patra, B., Singh, S.K., and Yuan, L. (2011). The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. Plant Physiology **157,** 2081-2093.

Takahashi, F., Yoshida, R., Ichimura, K., Mizoguchi, T., Seo, S., Yonezawa, M., Maruyama, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2007). The mitogen-activated protein kinase cascade MKK3–MPK6 is an important part of the jasmonate signal transduction pathway in Arabidopsis. Plant Cell **19,** 805-818.

Todd, A.T., Liu, E., Polvi, S.L., Pammett, R.T., and Page, J.E. (2010). A functional genomics screen identifies diverse transcription factors that regulate alkaloid biosynthesis in *Nicotiana benthamiana*. The Plant Journal **62,** 589-600.

Tsitsigiannis, D.I., and Keller, N.P. (2007). Oxylipins as developmental and host fungal communication signals. Trends in microbiology **15,** 109-118.

van der Fits, L., and Memelink, J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science **289,** 295-297.

Van Der Fits, L., and Memelink, J. (2001). The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. The Plant Journal **25,** 43-53.

Vanholme, B., Grunewald, W., Bateman, A., Kohchi, T., and Gheysen, G. (2007). The TIFY family previously known as ZIM. Trends in Plant Science **12,** 239-244.

Wang, H., Hao, J., Chen, X., Hao, Z., Wang, X., Lou, Y., Peng, Y., and Guo, Z. (2007). Overexpression of rice WRKY89 enhances ultraviolet B tolerance and disease resistance in rice plants. Plant Molecular Biology **65,** 799-815.

Wasternack, C. (2007). Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. Annals of Botany **100,** 681-697.

Wegel, E., Koumproglou, R., Shaw, P., and Osbourn, A. (2009). Cell type–specific chromatin decondensation of a metabolic gene cluster in oats. Plant Cell **21,** 3926-3936.

Xu, Y.-H., Wang, J.-W., Wang, S., Wang, J.-Y., and Chen, X.-Y. (2004). Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-δ-Cadinene synthase-a. Plant Physiology **135,** 507-515.

Yanagisawa, S. (2002). The Dof family of plant transcription factors. Trends in Plant Science **7,** 555-560.

Yanagisawa, S. (2004). Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. Plant and Cell Physiology **45,** 386-391.

Yu, J.-H., and Keller, N. (2005). Regulation of secondary metabolism in filamentous fungi. Annual Review of Phytopathology **43,** 437-458.

Yu, Z.-X., Li, J.-X., Yang, C.-Q., Hu, W.-L., Wang, L.-J., and Chen, X.-Y. (2012). The Jasmonate-responsive AP2/ERF transcription factors AaERF1 and AaERF2 positively regulate artemisinin biosynthesis in *Artemisia annua* L. Molecular Plant **5,** 353-365.

Zhang, H.-B., Bokowiec, M.T., Rushton, P.J., Han, S.-C., and Timko, M.P. (2012). Tobacco transcription factors NtMYC2a and NtMYC2b form nuclear complexes with the NtJAZ1 repressor and regulate multiple jasmonate-inducible steps in Nicotine biosynthesis. Molecular Plant **5,** 73-84.

Zhang, H., Hedhili, S., Montiel, G., Zhang, Y., Chatel, G., Pré, M., Gantet, P., and Memelink, J. (2011). The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. The Plant Journal **67,** 61-71.

Zhao, J., Davis, L.C., and Verpoorte, R. (2005). Elicitor signal transduction leading to production of plant secondary metabolites. Biotechnology Advances **23,** 283-333.

Zhu, Z., An, F., Feng, Y., Li, P., Xue, L., A, M., Jiang, Z., Kim, J.-M., To, T.K., Li, W., Zhang, X., Yu, Q., Dong, Z., Chen, W.-Q., Seki, M., Zhou, J.-M., and Guo, H. (2011). Derepression of ethylene-stabilized transcription factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in Arabidopsis. Proceedings of the National Academy of Sciences **108,** 12539-12544.

2

Natural product biosynthesis in *Medicago* species

Azra Gholami, Nathan De Geyter, Jacob Pollier, Alain Goossens

Publication status: Manuscript in preparation
Target journal (Natural Product Reports, IF: 9.790)
[Parts 1, 2.1-2.5 and 3.3, 4 and all figures were prepared by the candidate].

Abstract

The genus Medicago, a member of the legume (Fabaceae) family, comprises 87 species flowering plants, including forage crop alfalfa (Medicago sativa) and the model legume M. truncatula. As legume, Medicago species can establish symbiotic interaction with nitrogen- fixing bacteria in root nodules. To facilitate this interaction and to deter pathogen and herbivore, Medicago species accumulate a variety of bioactive natural products. For human, these compounds possess promising pharmaceutical activities. Here we provide a detailed of flavonoids and triterpene saponin, as the two most important classes of natural products produced by Medicago species. For each class we review the accumulating compounds, their biosynthesis, regulation and the advances in their genetic engineering and metabolic profiling. Furthermore, the biological role of these small molecules in terms of plant fitness, symbiosis and defense and ultimately their impact on human health are discussed.

1 Introduction

Legumes constitute a highly diverse plant family that encompasses economically important crops and provides about one- third of humankind's protein intake, fodder and forage for livestock, and raw materials for industries. They are unique among cultivated plants because of their ability to establish symbiotic interactions with the nitrogen-fixing bacteria, which leads to the formation of root nodules. In the nodules the bacteria differentiate into bacteroids and catalyze a symbiotic nitrogen fixation process referred to the reduction of atmospheric di-nitrogen into biologically useful ammonia (Jones et al., 2007). In addition, legumes are rich sources of health promoting phytochemicals such as isoflavones and triterpene saponins (Dixon and Sumner, 2003; Zhu et al., 2005).

The genus *Medicago* comprises 87 species of flowering plants belong to the legume family (Fabaceae) and includes the widely cultivated crop species alfalfa (*Medicago sativa*) and the model legume barrel medic (*M. truncatula*) (Steele et al., 2010; Sanders et al., 2011; Small, 2011). Due to the presence of agriculturally and economically important species, such as alfalfa, the *Medicago* genus has been the subject of numerous studies (Small, 2011). Alfalfa is the fourth largest crop economically (after corn, soybean, and wheat respectively) in North America and the temperate world's most forage crop (Small, 2011). It is a valuble food for livestock and alfalfa sprouts are a very popular human food. There has been a great interest recently to use it as a potential source of protein for human food and protein-based pharmaceuticals. In addition, alfalfa is likely the world's most environmentally friendly crop and has great potential to reduce negative ecological

aspects of agriculture. Beside alfalfa, other *Medicago spp*. are being used as medicine, human food, green manure, sources of industrial enzymes in biotechnology, model genomic species, and model systems for the study of nitrogen fixation (Young and Udvardi, 2009; Small, 2011).

Alfalfa is an obligate outcrossing and tetraploid species which make genetic and genomic studies difficult. Therefore, *Medicago truncatula* was developed as a model legume (Choi et al., 2004; Zhou et al., 2011). *M. truncatula* is an annual, diploid and autogamous legume with a moderate genome size (500-550 Mbp) and it is a close relative of alfalfa. These properties collectively establish *M. truncatula* as a model plant for legume genetics and genomics studies. Furthermore, the availability of a wide range of genomic (Cannon et al., 2006; Young et al., 2011) and genetic resources (Tadege et al., 2009) for this species makes it an invaluable model for studying secondary metabolism of legumes at the molecular genetic level (Zhou et al., 2011).

Improved functional genomic technologies such as transcriptomics, proteomics and metabolomics provide opportunities for an in-depth understanding of secondary metabolism in plants. In the model legume *M. truncatula* an integrated functional genomics approach were used to study natural product biosynthesis (Vorwerk et al., 2004; Achnine et al., 2005; Broeckling et al., 2005; Suzuki et al., 2005; Naoumkina et al., 2008). Such an integrated functional genomics approaches led to the reconstruction of the metabolic map of the pathways of the first legume pathway database for *M. truncatula* "MedicCyc" (http://www.noble.org/MedicCyc/). MedicCyc was generated based on new genomic sequence data and more than 225000 *M. truncatula* expressed sequence tags (ESTs) (Urbanczyk-Wochniak and Sumner, 2007). It currently represents over 400 pathways including the involved genes, enzymes and the resultant metabolites.

Among the secondary metabolites produced by *Medicago spp*., triterpene saponins and phenylpropanoid-derived isoflavonoids are of high interest and have been explored more thoroughly than the others (Dixon and Sumner, 2003). Lignins are the other phenylpropanoid- derived compounds that are found in all higher plants, and an important factor affecting cell wall digestibility in forage legumes including alfalfa (Dixon and Sumner, 2003).

In addition, the occurrence of alkaloids like stachydrine and trigonelline (Phillips et al., 1995a), cyanogenic glycosides (cyanogens) and nonprotein amino acids (NPAAS) such as canavanine (Wink et al., 2010) have been also reported in *Medicago* species.

This review provides a comprehensive overview of flavonoids and triterpene saponin, as the two most important classes of natural products produced by the genus *Medicago* and their impact on plant survival and human health. Here, for each class, the biosynthesis, their regulation and the most recent advances in the genetic engineering are discussed. Furthermore, we also describe their effects on plant structure, growth and development, their role in plant defense and ultimately their impact on human health.

2 Flavonoid

Flavonoids comprise one of the largest groups of secondary metabolites found ubiquitously in plants. Their wide occurrence, complex diversity and various functions have made them attractive for chemical, genetic and biological studies. They play various important biochemical and physiological roles by affecting several developmental processes. Flavonoids are responsible for much of the red, blue, and purple pigmentation found in plants and flowers and different classes of these compounds and their conjugates are involved in the interactions of plant with the environment, both in biotic and abiotic stress conditions (Dixon and Paiva, 1995; Shirley, 1996). In addition, the evidences of their beneficial effects on human health are growing. Flavonoids are characterized by polyphenolic structure, i.e. the basic C6-C3-C6 structural skeleton consisting of a double ring attached by a single bond to a third one. C ring is the heterocyclic benzopyran ring fused to the aromatic A ring, and the B is attached to the phenyl ring (Fig. 1). Based on different alterations and further modifications of this common backbone, flavonoids have been classified into different subclasses.

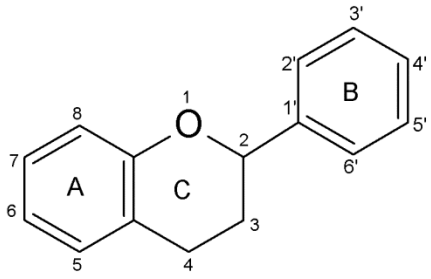


Figure 1. Flavonoid general structure.

2.1 Flavonoids in *Medicago* species

M. sativa and other *Medicago* species accumulate a variety of flavonoids. The main subclasses of flavonoids in *Medicago* comprise of chalcones, flavones, flavanones, isoflavones, flavonols, pterocarpan, aurones, anthocyanidin glycosides (anthocyanin) and

proanthocyanidins (PAs)(flavan-3-ol polymers) (Fig. 2). Table 1 presents the list of the flavonoid compounds have been reported in *Medicago* species, so far.

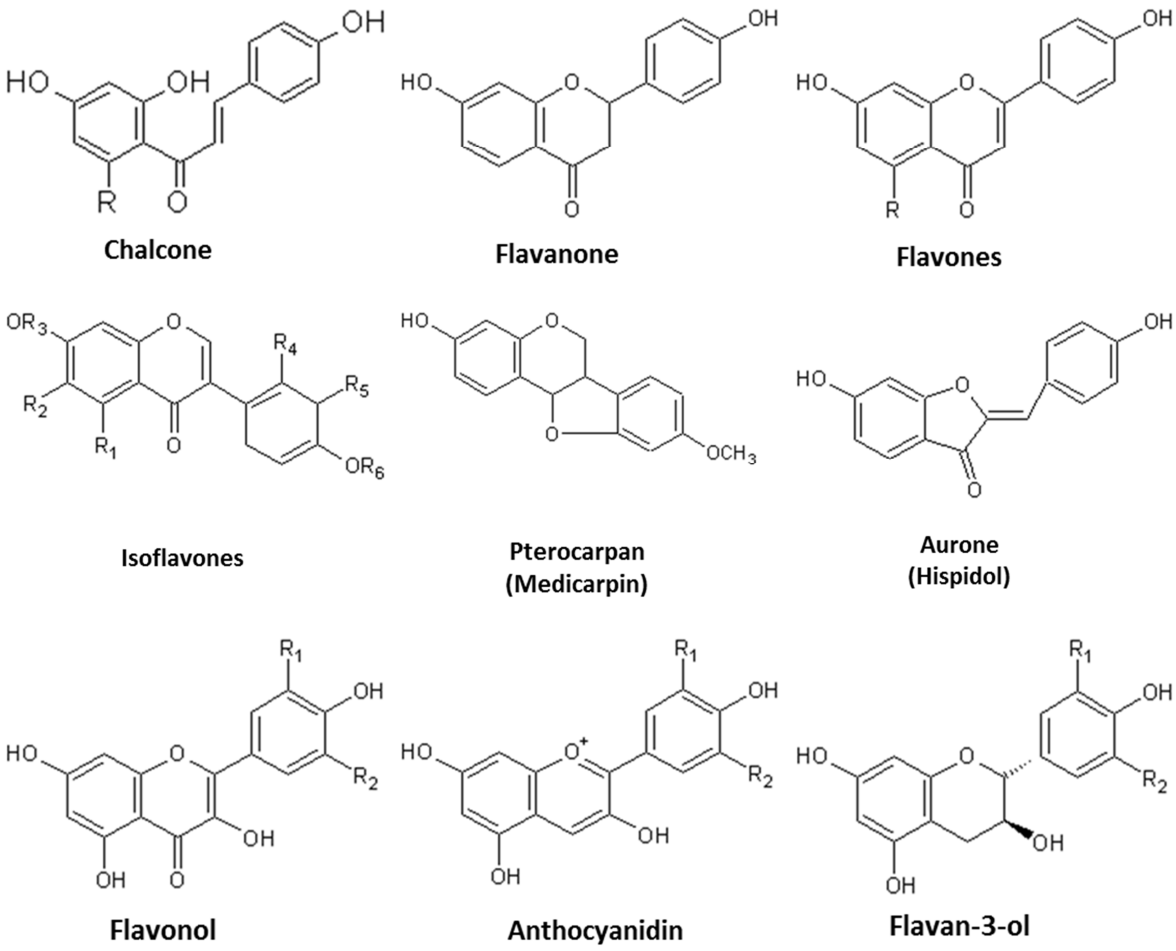


Figure 2 The main subclasses of flavonoids found in *Medicago spp.*

Table 1 Flavonoids and their glycoconjugates identified in different species of *Medicago* genus. Abbreviations of sugars and acyl groups: GlcA-glucuronic acid, Glc-glucose, Mal-malonic acid, Fer-ferrulic acid, Cou-Coumaric acid, GlcAPyr: glucuronopyranosyl

No.	Compound	Source	Tissue	References
Chalcones				
1	Isoliquiritigenin	*M. truncatula*	Cell cultures	(Farag et al., 2007)
2	Isoliquiritigenin	*M. sativa*	Leave	(Deavours and Dixon, 2005)
			Sprout	(Hong et al., 2011)
Flavanones				
3	Naringenin	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszkow et al., 2011)
			Cell cultures	(Frag et al., 2007
4	Liquiritigenin	*M. truncatula*	Hairy roots , suspension root cell cultures	Staszków, et al. 2011)
			Cell cultures	(Frag et al., 2007
		M. sativa	Leave	Deavours et al., 2005)
			Sprout	(Hong,2011)
5	*p*-Hydroxybenzaldehyde Liquiritigenin Glc	*M. truncatula*	Cell cultures	(Farag et al., 2008)
Flavones				
6	3', 4',7-Tihydroxyflavone7-Glc	*M. arabica*	Whole plant	(Saleh et al., 1982)
7	3', 4',7-Tihydroxyflavone7-GlcA	*M. arabica*	Whole plant	(Saleh, 1982)
8	3', 4',7-Trihydroxyflavone	*M. arabica*	Whole plant	(Saleh, 1982)
9	3',5-Dimethoxyluteolin Glc Mal	*M. truncatula*		(Farag et al., 2008)
10	4',7-Dihydroxyflavone	*M. truncatula*		(Farag et al., 2008)
		M. arabica	Whole plant	(Saleh, 1982)
		M. sativa	Whole plant	(Saleh, 1982)
		M. polymorpha	Whole plant	(Saleh, 1982)
		M. radiata	Whole plant	(Saleh, 1982)
		M. sativa	Leave	(Deavours et al., 2005)
		M. truncatula	Cell cultures	(Frag et al., 2007)
11	4',7-Dihydroxyflavone Glc	*M. truncatula*		(Farag et al., 2008)
12	4',7-Dihydroxyflavone7- GlcA	*M. arabica*	Whole plant	(Saleh, 1982)
		M. polymorpha	Whole plant	(Saleh, 1982)
		M. truncatula	Roots	(Staszków, et al. 2011)
13	4',7-Dihydroxyflavone7-GlcGlc	*M. arabica*	Whole plant	(Saleh, 1982)
		M. sativa	Whole plant	(Saleh, 1982)
		M. polymorpha	Whole plant	(Saleh, 1982)
14	5, 3'-Dimethoxyluteolin *β -D*-Glc	*M. truncatula*	Cell cultures	(Frag et al., 2007)
15	5, 3'-Dimethoxyluteolin *β -D*-Glc-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
16	5,3'-Dimethoxyluteolin	*M. truncatula*	Cell cultures	(Farag et al., 2007)
17	6,8-Dihydroxyflavone-7-*O*-*β*-*D*-GlcA	*M. sativa*	Plant extract	(Liang et al., 2011a)
18	6-methoxy-8-hydroxy-flavone-7-*O*-*β*-*D*-GlcA	*M. sativa*	Plant extract	(Liang, 2011)
19	Apigenin	*M. sativa*	Leave	(Deavours et al., 2005)
		M. truncatula	Root, seed	(Pang et al., 2009)
			Leave	(Marczak et al., 2010)
	Apigenin 7-*O*-Glc	*M. truncatula*	flower	(Pang, 2009)
20	Apigenin 4'-*O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001a; Golawska et al., 2010),

				(Golawska, et al. 2010)
21	Apigenin 4'-*O*-[2'-*O*-Fer-*β*-*D*-GlcAPyr(1→2)]-*O*-*β*-*D*- GlcAPyr]	*M. sativa*	Aerial parts	(Golawska, et al. 2010)
22	Apigenin 4'-O-[2'-*O*-*E*-Fer- *O*- *β* -GlcAPyr (1→2)- *O*- *β* -GlcAPyr]	*M. sativa* var. Artal	Plant extract	(Stochmal et al., 2001b)
23	Apigenin 7- GlcA	*M. arabica*	Whole plant	(Saleh, 1982)
24	Apigenin 7- *O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
25	Apigenin 7- *O*- *β* -GlcAPyr -4'-O-[2'-*O*-*E*-Fer- *O*-*β* -GlcAPyr (1→2)- *O*- *β* -GlcAPyr]	*M. sativa* var. Artal	Plant extract	(Stochmal et al. c, 2001)
26	Apigenin 7-*O*-[*β* -*D*-GlcAPyr(1→2)-*O*- *β* -*D*-GlcAPyr]-4'-*O*- *β* -*D*- GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
27	Apigenin 7-*O*-[2'-*O*-Cou-GlcAPyr-(1→2)- *O*-GlcAPyr]	*M. truncatula*	Leave	(Jasinski et al., 2009), Marczak, 2010)
28	Apigenin 7-*O*-[2'-*O*-Fer-GlcAPyr-(1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
29	Apigenin 7-*O*-[2'-*O*-sinapoyl-GlcAPyr-(1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Marczak, 2010)
30	Apigenin 7-*O*-[2-*O*-Fer- *β* -*D*-GlcAPyr(1→2)-*O*- *β* -*D*-GlcAPyr]-4'-*O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001, Golawska, et al. 2010)
31	Apigenin 7-*O*-[2'-O-sinapoyl-*β*-*D*-GlcAPyr-(1→2)-*O*-*β*-*D*-GlcAPyr],	*M. truncatula*	Aerial parts	(Kowalska et al., 2007)
32	Apigenin 7-*O*-[*β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
33	Apigenin 7-*O*-{2'-*O*-Fer-[GlcAPyr-(1→3)]-*O*-GlcAPyr-(1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Marczak, 2010)
34	Apigenin 7-O-{2-*O*-*E*-Fer-[*O*- *β* -GlcAPyr (1→3)]- *O*- *β* -GlcAPyr (1→2)- *O*- *β* -GlcAPyr }Madmememsattti	*M. ivanivivaiva* var. Artal	Plant extract	(Stochmal et al. 2001)
35	Apigenin 7-*O*-{2-*O*-Fer-[*β* -*D*-GlcAPyr (1→3)]-*O*- *β*-*D*-GlcAPyr (1→2)-*O*-*β* -*D*-GlcAPyr}	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
		M. truncatula	Aerial parts	(Kowalska, 2007)
36	Apigenin 7-*O*-{2-*O*-*p*-Cou-[*β* - GlcAPyr (1→3)]-*O*- *β*-*D*-GlcAPyr (1→2)-*O*-*β* -*D*-GlcAPyr}	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
37	Apigenin 7-*O*-GlcAPyr-(1→2)-O- GlcAPyr	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
38	Apigenin 7-*O*-GlcAPyr-(1→2)-*O*-GlcAPyr-(1→2)-*O*-Glc	*M. truncatula*	Leave	(Marczak, 2010)
39	Apigenin 7-*O*-GlcAPyr-(1→3)-*O*-GlcAPyr-(1→2)-*O*- GlcAPyr	*M. truncatula*	Leave	(Marczak, 2010)
40	Apigenin 7-*O*-*β*-*D*-GlcAPyr	*M. sativa*	Aerial parts	(Golawska, et al. 2010)
41	Apigenin 7-*O*-*β*-*D*-GlcAPyr-(1→3)-*O*-*β*-*D*-GlcAPyr-(1→2)-*O*-*β*-*D*-GlcAPyr,	*M. truncatula*	Aerial parts	(Kowalska, 2007)
42	Apigenin 7-*O*-*β*-*D*-GlcAPyr-4'-*O*-[2'-*O*-Fer-*O*-*β*-*D*-GlcAPyr(1→2)-*O*-*β*-*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Golawska, et al. 2010)
43	Apigenin 7-*O*-*β*-*D*-GlcAPyr-4'-*O*-[2'-*O*-p-Cou-*O*-*β*-*D*-GlcAPyr(1→2)-*O*-*β*-*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al. 2001, Golawska, et al. 2010)
44	Chrysoeriol	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
45	Chrysoeriol 7- GlcA	*M. arabica*	Whole plant	(Saleh, 1982)
46	chrysoeriol 7- GlcA	*M. polymorpha*	Whole plant	(Saleh, 1982)
		M. radiata	Whole plant	(Saleh, 1982)
47	Chrysoeriol GlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
48	Chrysoeriol 7- GlcAGlcA	*M. arabica*	Whole plant	(Saleh, 1982)
		M. sativa	Whole plant	(Saleh, 1982)
		M. polymorpha	Whole plant	(Saleh, 1982)
49	Chrysoeriol 7- GlcAGlcAGlcA	*M. sativa*	Whole plant	(Saleh, 1982)

50	Chrysoeriol 7- *O*- *β* -*D*-GlcAPyr -4'-*O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
51	Chrysoeriol 7-*O*- GlcAPyr	*M. truncatula*	Leave	(Marczak, 2010)
52	Chrysoeriol 7-*O*-[2'-*O*-Fer *O*- *β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
53	Chrysoeriol 7-*O*-[2'-*O*-Cou-GlcAPyr- (1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
54	Chrysoeriol 7-*O*-[2'-*O*-Fer-GlcAPyr-(1→2)- *O*-GlcAPyr]	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
55	Chrysoeriol 7-*O*-{2'-*O*-Fer-[*O*- *β* -*D*-GlcAPyr (1→3)]- *O*- *β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr}	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
56	• Chrysoeriol 7-*O*-{2'-*O*-*p*-Cou-[*β*-*D*-GlcAPyr-(1→3)]-*O*-*β*-*D*-GlcAPyr(1→2)-*O*-*β*-*D*-GlcAPyr}	*M. truncatula*	Aerial parts	(Kowalska, 2007)
57	Chrysoeriol 7-*O*-GlcAPyr-(1→2)-*O*- GlcAPyr	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
58	Chrysoeriol 7-*O*-GlcAPyr-(1→2)-*O*- GlcAPyr	*M. truncatula*	Leave	(Marczak, 2010)
59	Chrysoeriol 7-*O*-*β*-*D*-GlcAPyr-(1→2)-*O*-*β*-*D*-GlcAPyr	*M. truncatula*	Aerial parts	(Kowalska, 2007)
60	Iirosolidone or dimethoxyluteolin	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
61	Luteolin	*M. sativa* *M. falcate* *M. polymorpha* *M. orbicularis*	seeds	(Phillips et al., 1995b; Prati et al., 2007)
62	Luteolin 7- GlcA	*M. arabica* *M. sativa* *M. polymorpha* *M. radiata*	Whole plant Whole plant Whole plant Whole plant	(Saleh, 1982) (Saleh, 1982) (Saleh, 1982) (Saleh, 1982)
63	Luteolin 7- *O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
64	Luteolin 7-*O*- GlcAPyr	*M. truncatula*	Leave	(Marczak, 2010)
65	Luteolin 7-*O*-[2-*O*-Fer- *β* -*D*-GlcAPyr(1→2)-*O*- *β* -*D*-GlcAPyr]-4'-*O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
66	Luteolin 7-*O*-GlcAPyr	*M. truncatula*	Leave	(Jasiński, et al., 2009)
67	Luteolin Glc	*M. truncatula*	Roots, hairy roots Flower	(Staszków, et al. 2011) (Pang, 2009)
68	Luteolin GlcA	*M. truncatula*	Roots, hairy roots	(Staszków, et al. 2011)
69	Luteolin GlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
70	Luteolin-7-*O*-Glc	*M. sativa* *M. falcate* *M. polymorpha* *M. orbicularis*	seeds	(Phillips et al, 1995)
71	Mal 3',5-Dimethoxyluteolin Glc	*M. truncatula*	Cell cultures	(Farag et al., 2008)
72	Rutin	*M. truncatula*	Aerial parts	(Kowalska, 2007)
73	Tricetin 7- *O*- *β* -*D*-GlcAPyr -3'-*O*-methyl	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
74	Tricin	*M. truncatula*	Cell cultures	(Frag et al., 2007, 2008)
75	Tricin 7- GlcA	*M. sativa*	Whole plant	(Saleh, 1982)
76	Tricin 7- GlcAGlcA	*M. sativa*	Whole plant	(Saleh, 1982)
77	Tricin 7- GlcAGlcAGlcA	*M. sativa*	Whole plant	(Saleh, 1982)
78	Tricin 7- *O*- *β* -*D*-GlcAPyr	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
79	Tricin 7-*O*-[2'-*O*- Fer-*β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)

80	Tricin 7-*O*-[2'-*O*-*p*-Cou- *β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
81	Tricin 7-*O*-[2'-*O*-sinapoyl-*β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
82	Tricin 7-*O*-[2'-*O*-Cou-GlcAPyr-(1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
83	Tricin 7-*O*-[2'-*O*-Fer-GlcAPyr-(1→2)-*O*-GlcAPyr]	*M. truncatula*	Leave	(Marczak, 2010)
84	Tricin 7-*O*-[2'-*O*-Fer-*β*-*D*-GlcAPyr-(1→2)-*O*-*β*-*D*-glucopyranoside]	*M. truncatula*	Aerial parts	(Kowalska, 2007)
85	Tricin 7-*O*-[*β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr]	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
86	Tricin 7-*O*-{2'-*O*-Fer-[*β* -*D*-GlcAPyr (1→3)]- *O*- *β* -*D*-GlcAPyr (1→2)- *O*- *β* -*D*-GlcAPyr }	*M. sativa*	Aerial parts	(Stochmal et al., 2001)
87	Tricin 7-*O*-{2'-*O*-*p*-Cou-[*β*-*D*-GlcAPyr-(1→3)]-*O*-*β*-*D*-GlcAPyr(1→2)-*O*-*β*-*D*-GlcAPyr	*M. truncatula*	Aerial parts	(Kowalska, 2007
88	Tricin 7-*O*-GlcAPyr-(1→2)-*O*- GlcAPyr	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
89	Tricin 7-*O*-*β*-*D*-GlcAPyr-4'-*O*- GlcAPyr	*M. truncatula*	Aerial parts	(Kowalska, 2007)
	Hyperoside	*M. sativa*	Seed	(Prati, 2007)
Isoflavones				
90	2'-Hydroxyformononetin	*M. truncatula*	Cell cultures	(Farag et al., 2007)
91	2'-Hydroxyformononetin MalGlc	*M. truncatula*	Roots, hairy roots	(Staszków, et al. 2011)
			Cell cultures	(Frag et al., 2007, Farag et al., 2008)
92	2'-Hydroxyformononetinb-*D*-Glc	*M. truncatula*	Cell cultures	(Frag et al., 2007, Farag et al., 2008)
93	7,4'-dihydroxyflavoneononin	*M. sativa*	Roots	(Coronado et al., 1995)
94	Afromosin,	*M. sativa*	Leave	(Deavours et al., 2005)
95	Afrormosin-7-*O*-Glc	*M. sativa*	Cell cultures	(Kessmann et al., 1990)
96	Afrormosin-7-*O*-Glc-6"-*O*-mal	*M. sativa*	Cell cultures	(Kessmann et al., 1990)
		M. truncatula	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
			Cell cultures	(Farag et al., 2007, Farag et al., 2008)
97	Afrormosin 7-*O*-*β*-*D*-Glc	*M. truncatula*	Cell cultures	(Frag et al., 2007, Farag et al., 2008)
98	Afrormosin 7-*O*-*β*-*D*-Glc-6"-*O*-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
99	Afrormosin 7-*O*-*β*-*D*-Glc-Mal (isomer)	*M. truncatula*	Cell cultures	(Frag et al., 2007; Farag et al., 2008,
100	Afrormosin Glc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	Staszków, et al. 2011)
101	Afrormosin MalGlc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
102	Alfalone	*M. truncatula*	Cell cultures	(Farag et al., 2008)
103	Biochanin A	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
104		*M. truncatula*	Cell cultures	(Farag et al., 2007)
		M. sativa	Leave	(Deavours et al., 2005)

105	Biochanin A 7- *O*- *β* -GlcAPyr	*M. littoralis*	Aerial parts	(Alessandra et al., 2010)
106	Biochanin A 7-*O*-*β*-*D*-Glc	*M. truncatula*	Cell cultures Root	(Frag et al., 2007) (Pang 2009)
107	Biochanin A 7-*O*-*β*-*D*-Glc-6"-*O*-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
108	Biochanin A MalGlc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
109	Biochanin A MalGlcGlc	*M. truncatula*	Hairy roots	(Staszków, et al. 2011)
110	Biochanin A *β*-*D*- Glc Glc	*M. truncatula*	Cell cultures	(Frag et al., 2007, 2008)
111	Biochanin A *β*-*D*- Glc Glc –Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007, 2008)
112	Daidzein	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
113		*M. sativa*	Leave	(Deavours et al., 2005)
114		*M. truncatula*	Cell cultures	(Farag et al., 2007)
115		*M. arabica*	Plant extract	(Saleh, 1982)
116	Daidzein 7-*O*-*β*-D-Glc (Daidzin)	*M. truncatula*	Cell cultures	(Frag et al., 2007)
117	Daidzein 7-*O*-*β*-*D*-Glc-6"-*O*-Mal (Daidzin Mal)	*M. truncatula*	Cell cultures	(Frag et al., 2007)
118	Daidzein CouGlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
119	Daidzein FerGlcAGlcA	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
120	Daidzein GlcA	*M. truncatula*	Suspension root cell cultures	(Staszków, et al. 2011)
121	Daidzein GlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
122	Daidzein MalGlc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
123	Formononetin	*M. truncatula*	Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
			Cell cultures	(Farag et al., 2007)
			Root	(Aloui, 2012)
		M. sativa	Leave	(Deavours et al., 2005)
124	Formononetin 7-*O*- *β*-(6'-*O*-MalGlc) (Malononin)	*M. truncatula*	Root	(Aloui et al., 2012)
125	Formononetin 7-*O*-Glc	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010, Pang 2009)
			Root	(Pang 2009)
126	Formononetin 7-*O*-Glc Malated	*M. truncatula*	Leave	(Jasiński, et al., 2009, Marczak, 2010)
127	Formononetin- 7-*O*-*P*- *β*-*D*-glycoside (Ononin)	*M. sativa*	Roots	(Coronado et al., 1995)
		M. truncatula	Cell cultures	(Farag et al., 2007)
			Roots, hairy roots	(Staszków, et al. 2011)
128	Formononetin 7-*O*-*β*-*D*-Glc-6"-*O*-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
129	Formononetin MalGlc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
130	Formononetin-7-*O*-*P*-*β*-*D*-Glc-6"-Mal methyl ester	*M. sativa*	Roots	(Coronado et al., 1995)

131	Genisitin	*M. truncatula*	Cell cultures	(Farag et al., 2007, Farag et al., 2008)
			Hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
132	Genistein,	*M. sativa*	Leave	(Deavours et al., 2005)
133	Genistein 7- *O*- *β* -GlcAPyr	*M. littoralis*	Aerial parts	(Alessandra et al., 2010)
134	Genistein 7-*O*-*β*-*D*- Glc-6”-*O*-Mal	*M. truncatula*	Cell cultures	(Farag et al., 2007)
135	Genistein 7-*O*-*β*-*D*-Glc (genistin)	*M. truncatula*	Cell cultures	(Farag et al., 2007)
136	Genistein CouGlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
137	Genistein FerGlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
138	Genistein Glc Mal	*M. truncatula*	Cell cultures	(Farag et al., 2008)
139	Genistein Glc Mal (isomer)	*M. truncatula*	Cell cultures	(Farag et al., 2008)
			Hairy roots	(Staszków, et al. 2011)
140	Genistein GlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
			Flower	(Pang et al., 2009)
141	Genistein GlcAGlcA	*M. truncatula*	Roots	(Staszków, et al. 2011)
142	Genistein *β*-*D*- Glc Glc	*M. truncatula*	Cell cultures	(Farag et al., 2007)
143	Genistein *β*-*D*- Glc Glc Mal	*M. truncatula*	Cell cultures	(Farag et al., 2007)
144	Irilone	*M. truncatula*	Cell cultures	(Farag et al., 2007)
145	Irilone 4”-*O*-*β*-*D*-Glc-6”-*O*-Mal	*M. truncatula*	Cell cultures	(Farag et al., 2007)
146	Irisolidone	*M. truncatula*	Cell cultures	(Farag et al., 2007)
147	Irisolidone 7-*O*-*β*-*D*-Glc	*M. truncatula*	Cell cultures	(Farag et al., 2007)
148	Irisolidone 7-*O*-*β*-*D*-Glc-6”-O-Mal	*M. truncatula*	Cell cultures	(Farag et al., 2007)
149	Irisolidone MalGlc	*M. truncatula*	Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
150	Pratensein	*M. sativa*	Seed	(Prati, 2007)
151	Isoflav-3-ene-Glc Mal	*M. truncatula*		(Farag et al., 2008)
Coumestan				
152	Coumestrol	*M. truncatula*	Root	(Aloui, 2012)
153	Coumestrol	*M. sativa*	Sprout	(Hong et al., 2011)
			Cotyledon	(O'Neill, 1996)
Flavonol				
154	Kaempferol	*M. truncatula*	Flower	(Pang et al., 2009)
155	Kaempferol 3- Glc	*M. polymorpha*	Plant extract	(Saleh, 1982)
156	Kaempferol 3,7- Glc Glc	*M. radiata*	Plant extract	(Saleh, 1982)
157	Kaempferol 3-O-rutinoside	*M. truncatula*	Seed	(Pang et al., 2009)
158	laricitrin 3,5'- *O*- *β* -*D*-GlcAPyr	*M. littoralis*	Aerial parts	(Alessandra et al., 2010)
159	Laricitrin 3,5‘-GlcGlc	*M. truncatula*	Aerial parts	(Kowalska, 2007)
160	Laricitrin 3,7,5‘-GlcGlcGlc	*M. truncatula*	Aerial parts	(Kowalska, 2007)
161	Laricitrin 3-*O*-g GlcAPyr--5′-*O*- GlcAPyr-l-7-*O*-Glc	*M. truncatula*	Leave	(Marczak, 2010)
162	Myricetin	*M. sativa*	leave	(LeRoy et al., 2002)
163	Quercetin	*M. sativa*	Seed	(Prati, 2007)
Pterocarpans				
164	Medicarpin	*M. truncatula*	Root	(Aloui, 2012)

			Cell cultures	(Frag et al., 2007)
			Suspension root cell cultures	(Staszków, et al. 2011)
		M. sativa	Cotyledone	(O'Neill, 1996)
165	Medicarpin-3-*O*-Glc-6"-*O*-mal	*M. sativa*	Cell cultures	(Kessmann et al., 1990)
166	Medicarpin 3-*O*-*β*-*D*-Glc	*M. truncatula*	Cell cultures	(Frag et al., 2007)
167	Medicarpin 3-*O*-*β*-*D*-Glc-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
168			Root	(Aloui, 2012)
			Roots, hairy roots , suspension root cell cultures	(Staszków, et al. 2011)
Isoflavans				
169	Sativan	*M. sativa*	Leaves	(INGHAM and MILLAR, 1973)
			Cotyledon	(O'Neill, 1996)
170	Vestitol	*M. sativa*	Cotyledon	(O'Neill, 1996)
171	Vestitol *β*-*D*-Glc-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
Aurone				
172	Hispidol 4'-*O*-*β*-*D*- Glc,	*M. truncatula*	Cell cultures	(Frag et al., 2007)
173	Hispidol 4'-*O*-*β*-*D*- Glc-Mal	*M. truncatula*	Cell cultures	(Frag et al., 2007)
174	Hispidol	*M. truncatula*	Cell cultures	(Farag et al., 2007)
Anthocyanidin (glycosides)				
175	Cyanidin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
176	Cyanidin 3-*O*-glucoside	*M. truncatula*	Flower, seed	(Pang, et al., 2009)
177	Delphinidin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
178	Pelargonidin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
179	Pelargonidin-3-*O*-glucoside	*M. truncatula*	Root	(Pang, et al., 2009)
Flavan-3-ol				
180	Epicatechin	*M. truncatula*	Seed coat, flower	(Pang, et al., 2007, 2009)
181	Epicatechin 3'-*O*-glucoside	*M. truncatula*	Flower, seed	(Pang, et al., 2007)
182	Catechin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
183	Gallocatechin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
184	Epigallocatechin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
185	Afzelechin	*M. truncatula*	Seed coat	(Pang, et al., 2007)
186	Epiafzelechin	*M. truncatula*	Seed coat	(Pang, et al., 2007)

2.2 Flavonoid biosynthesis in *M. truncatula*

The flavonoid biosynthesis pathway and the structural genes involved in the pathway are well characterized in plants. Here we first describe the early steps of flavonoid biosynthesis in legumes up to the biosynthesis of 5-hydroxy and 5-deoxy-flavonoids, the immediate precursors of isoflavonoids, flavones and dihydroflavonol. Then the flavonoid

pathway leading to the production of isoflavonoids, flavonols, flavones, aurones, anthocyanins and PAs have been described in separate sections with a short description of each class.

2.2.1 Early steps of the flavonoid biosynthetic pathway

The biosynthetic pathway leading to the production of flavonoids derived from phenylpropanoid pathway. It begins with the formation of 4-coumaroyl-CoA through three enzymatic conversions catalyzed by sequential activities of phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate: CoA ligase (4CL) on the one hand, and cytosolic formation of malonyl-CoA from acetyl-CoA by acetyl-CoA carboxylase (ACC), on the other hand. Then, one 4-coumaroyl-CoA and three molecules of malonyl-CoA serve as substrates for chalcone synthase (CHS) to run a series of sequential decarboxylation and condensation reactions leading to the formation of a polyketide intermediate. This intermediate undergoes cyclization and aromatization reactions to form the A-ring and the resultant chalcone (naringenin chalcone). In *Medicago spp*., CHS is encoded by 8-12 genes (Junghans et al., 1993; McKhann and Hirsch, 1994). Chalcone reductase (CHR) is found only in leguminous plants. It removes the hydroxyl group of the second malonyl-CoA during chalcone biosynthesis, which together with the CHS lead to the biosynthesis of 5-deoxyflavonoids (Ballance and Dixon, 1995).

Then, two flavanones, naringenin and liquiritigenin which are precursors of 5-hydroxy- and 5-deoxy-flavonoids, respectively, are produced via enzymatic activity of chalcone isomerase (CHI). Two classes of CHIs have been described based on substrate specificity. Type I CHIs are found in both legumes and nonlegumes, and use naringenin chalcone to convert it into naringenin. Type II CHIs are legume-specific enzyme which can isomerize both naringenin chalcone and isoliquiritigenin to naringenin and liquiritigenin, respectively (Fig. 3) (Shimada et al., 2003; Ralston et al., 2005). Both CHIs type I and II involved in flavonoid biosynthetic pathway have been identified in *Medicago* (Davies and Schwinn, 2005).

The resulting flavonoids, naringenin and liquiritigenin, then serve as the substrates of isoflavone synthase (IFS), flavone synthase (FNS) and flavanone 3-hydroxylase (F3H) to produce isoflavonoids, flavones and dihydroflavonol, respectively.

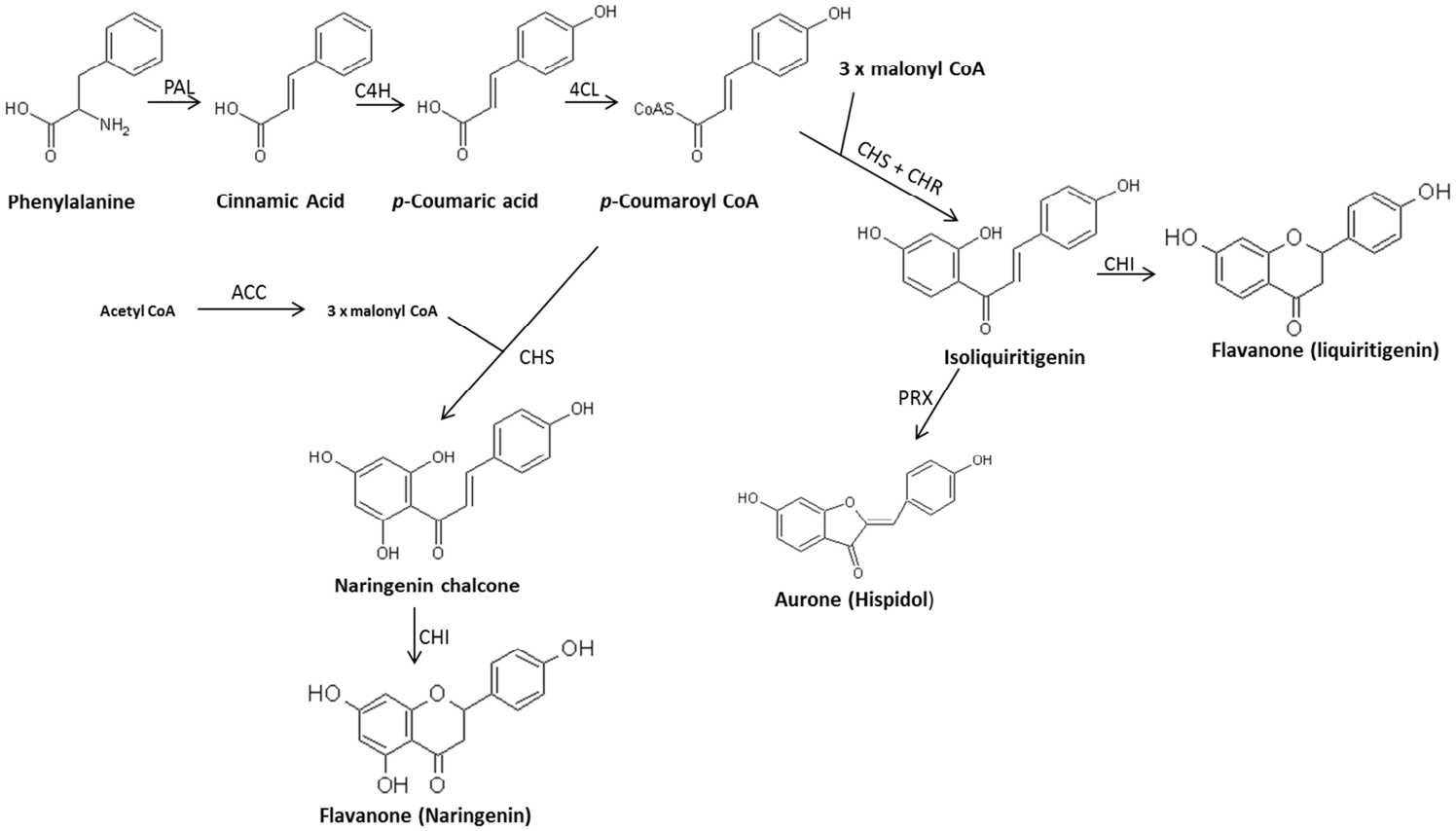


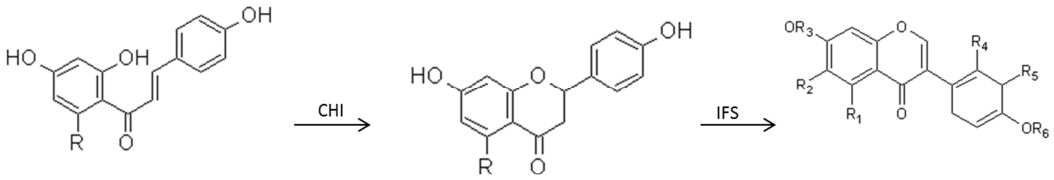
Figure 3. Early steps of the flavonoid biosynthetic pathway in *M. truncatula*. CHS and CHR co-act together to produce 5-deoxyflavonoids. PAL; phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate: CoA ligase; ACC, acetyl-CoA carboxylase; CHS, chalcone synthase; CHI, chalcone isomerase; CHR, chalcone reductase, PRX: peroxidase.

2.2.2 Isoflavonoid biosynthesis

In isoflavonoids the B-ring is attached to the C-ring via the C-3 rather than the C-2 position**.** The first committed step of the isoflavonoid biosynthesis begins with the 2-hydroxylation and aryl migration, catalyzed by a cytochrome P450 (CYP450) enzyme, 2-hydroxyisoflavanone synthase (2HIS, also known as isoflavone synthase, IFS). Subsequently, dehydration of the 2-hydroxyisoflavanone intermediates, 2,5,7,4'-tetrahydroxyisoflavanone and 2,7,4'-trihydroxyisoflavanone, by 2-hydroxyisoflavanone dehydratase (2HID) forms the isoflavones, genistein or daidzein, respectively. The 4'-*O*-methylation of genistein and daidzein catalyzed by hydroxyisoflavanone-4'-*O*-methyltransferase (HI4'OMT) leads to the formation of biochanin A and formononetin, respectively, which are the most abundant isoflavonoid aglycones in *M. truncatula* roots (Zhang et al., 2009b) (Fig. 4).

Flavonoids are often accumulated as malonylated or acetylated glucoconjugates (Table1). Malonylated and acetylated isoflavonoids are considered as 'storage forms' that accumulate in the vacuoles, serving as a pool of biosynthetic precursors or inactive forms of phytoalexins (Dixon, 1999; Naoumkina et al., 2007). Several glycosyltransferases (UGTs) involved in isoflavonoid glycosylation have been characterized in *Medicago sp*. Of more than 100 UGTs characterized in *M. truncatula*, UGT71G1, UGT85H2, and UGT78G1 have been identified with potential activity on isoflavonoid aglycones (Achnine et al., 2005; He et al., 2006; Li et al., 2007; Modolo et al., 2009). UGT78G1 has a broad activity on formononetin, kaempferol and the anthocyanidins pelargonidin and cyanidin (Modolo et al., 2009). UGT71G1 was shown to be capable of glycosylation of flavonols, quercetin and isoflavonoid, genistein and saponin (Achnine et al., 2005; Shao et al., 2005). UGT85H2 is a multifunctional flavonoid glycosyltransferase with the ability to glycosylate several flavonoid-related compounds, such as isoflavones (biochanin A), flavonols (kaempferol), and chalcones (isoliquiritigenin) (Li et al., 2007).

The malonyl residues substituted on the sugar moiety, protect isoflavonoids from enzymatic degradation of the glucoconjugates, change their lipophilicity and act as a molecular tag promoting efficient vacuolar uptake of the conjugates (Markham et al., 2000). Three characterized malonyltransferase from *M. truncatula* named as MtMaT1, 2 and 3, have shown that are capable to catalyze the malonylation of a range of isoflavone 7-*O*-glucosides *in vitro.* Of which MtMaT1 and/or 2 function were suggested as malonyl CoA:isoflavone 7-*O*-malonyltransferases *in vivo* (Yu et al., 2008).



Chalcone	R
Naringenin chalcone	OH
Isoliquiritigenin	H

Flavanone	R
Naringenin	OH
liquiritigenin	H

Isoflavone	R1	R2	R3	R4	R5	R6
Biochanin A	OH	H	H	H	H	CH3
Genistein	OH	H	H	H	H	H
Irilone	OH	-O-CH2-		H	H	H
Irisolidone	OH	OCH3	H	H	H	CH3
Afrormosin	H	OCH3	H	H	H	CH3
Daidzein	H	H	H	H	H	H
Formononetin	H	H	H	H	H	CH3
2'-Hydroxy Formononetin	H	H	H	OH	H	CH3
Alfalone	OH	H	CH3	H	H	CH3
Glycitein	H	OCH3	H	H	H	H

Figure 4. Flavanone and isoflavonoid biosynthesis in *Medicago spp*. CHI, chalcone isomerase; IFS, isoflavone synthase.

2.2.3 Pterocarpan biosynthesis

The isoflavones undergo a series of reactions leading to the production of pterocarpans such as medicarpin. The first step in the biosynthesis of pterocarpans from isoflavones is catalyzed by isoflavone 2'-hydroxylase (I2'H) and isoflavone 3'-hydroxylase (I3'H) that catalyze hydroxylation of C2' and C3' of B-ring, respectively. *M. truncatula* I2'H (CYP81E7) catalyzes the hydroxylation of formononetin at the C2' position which provides the hydroxyl-group that is later used for the formation of the ether linkage present in the medicarpin structure (Liu et al., 2003). *M. truncatula* I3'H was shown to be able to catalyze the hydroxylation of both biochanin A and formononetin. However, pratensein, a naturally occurring product of the 3'-hydroxylation of biochanin A in chickpea (*Cicer arietinum*) has not been reported in *M. truncatula* (Liu et al., 2003). In the next step, the resulting 2'-hydroxy-formononetin is reduced to the corresponding 2'-hydroxylated isoflavanone vestitone by the NADPH- dependent isoflavone reductase (IFR). The 2-step conversion of vestitone into medicarpin is catalyzed by vestitone reductase that reduces vestitone to 7, 2'-dihydroxy-4'-methoxy- isoflavanol (DMI), which is subsequently dehydrated by DMI dehydratase there by forming the ether linkage of the pterocarpan skeleton (Guo et al., 1994; Shao et al., 2007) (Fig. 5).

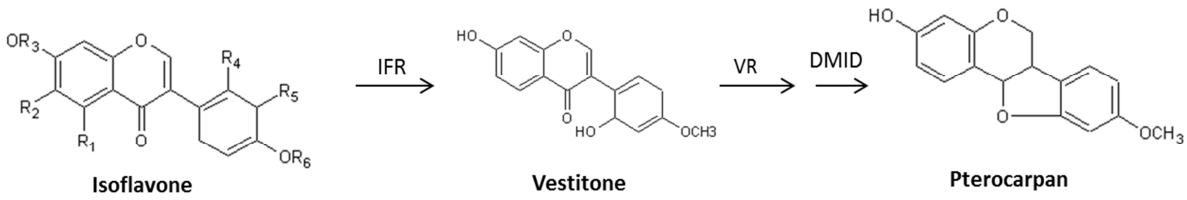


Figure 5. Pterocarpan biosynthesis. IFR, isoflavone reductase; VR, vestitone reductase; DMID, dihydroxy-4'-methoxy- isoflavanol dehydratase.

2.2.4 Flavonols biosynthesis

Flavonols are the most widespread of the flavonoids in higher plants. Flavonol synthase (FLS) catalyzes the key step in the biosynthesis of flavonols. This enzyme converts the dihydroflavonols, dihydrokaempferol, dihydroquercetin, and dihydromyricetin to the corresponding flavonols, kaempferol, quercetin, and myricetin, respectively (Fig. 6). FLS has been characterized in several plant species including *A. thaliana* (Pelletier et al., 1997; Wisman et al., 1998), *Citrus unshiu* Marc. (Lukacin et al., 2003), parsley (Martens et al., 2003), soybean (Takahashi et al., 2007) and strawberry (Almeida et al., 2007). However, the FLS encoding gene in *M. truncatula* has not been characterized yet.

The 3'-*O*-methylation of myricetin catalyzed by myricetin-*O*-methyltransferase leads to the formation of laricitrin (Fig.6).

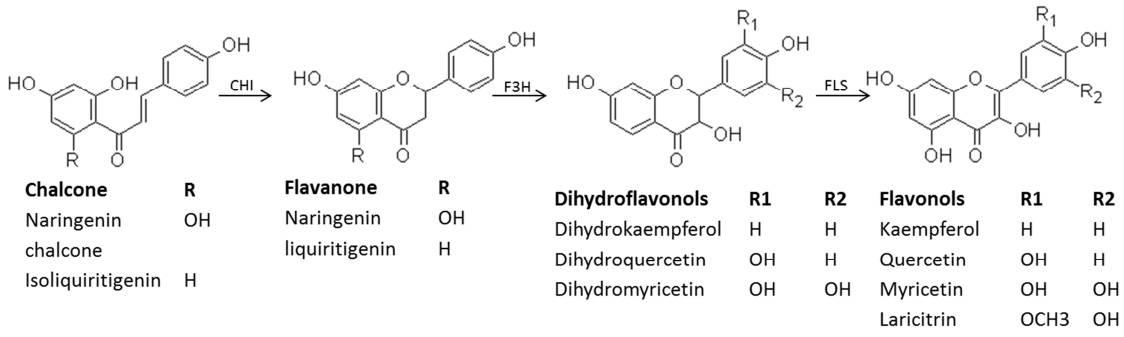


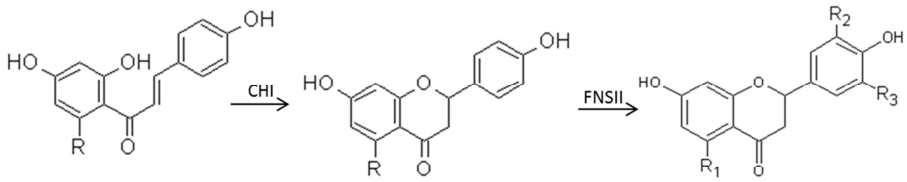
Figure 6. Flavonols biosynthesis. CHI: chalcone isomerase; F3H: Flavanone 3-hydroxylase; FLS, flavonol synthase.

2.2.5 Flavones biosynthesis

The biosynthesis of flavones from flavanone occurs via introducing a double bond between C2 and C3 that is catalyzed by the flavone synthase (FNS) enzyme (Fig. 7). Two different FNS (FNSI and FNSII) have been characterized. FNSI, a soluble 2-oxoglutarate-dependent and Fe^{2+}-dependent dioxygenase is mainly characterized in the members of Apiaceae family as well as monocotyledonous plants (Martens et al., 2001; Kim et al., 2008). Whereas FNSII, a NADPH-dependent CYP450 monooxygenase widespread

among higher plants (Martens and Mithofer, 2005). All known FNS II proteins belong to the plant CYP450 subfamily CYP93B and two different mechanisms for their catalytic activity have been reported. FNS II from the legume licorice converts flavanone substrates to flavones via a 2-hydroxyflavanone intermediate, whereas FNS II from the nonlegume *Gerbera hybrida*, converted flavanones to flavones directly (Humphreys and Chapple, 2002; Choi et al., 2004).

Two *FNS II* genes, *MtFNSII-1* (*CYP93B10*) and *MtFNSII-2* (*CYP93B11*) have been characterized in *M. truncatula.* Both MtFNSII-1 and MtFNSII-2 convert flavanones to 2-hydroxyflavanones instead of flavones and have a distinct tissue-specific expression pattern. *MtFNSII-1* is mostly expressed in roots and seeds where the major accumulated flavones are 7,4'-dihydroxyflavone and apigenin/luteolin, respectively, while *MtFNSII-2* is highly expressed in flowers and siliques (Zhang et al., 2007).



Chalcone	R
Naringenin chalcone	OH
Isoliquiritigenin	H

Flavanone	R
Naringenin	OH
liquiritigenin	H

Flavones	R1	R2	R3
Apigenin	OH	H	H
7,4'-dihydroxyflavone	H	H	H
Tricin	OH	OCH3	OCH3
Luteolin	OH	OH	H
Chrysoeriol	OH	OCH3	OH

Figure 7. Flavone biosynthesis in *M. truncatula.* CHI: chalcone isomerase; FNS II, flavone synthase II.

2.2.6 Aurones biosynthesis

Structurally, aurones are the isomers of flavones and widely distributed in fruits and flowers. They are responsible for yellow pigmentation in their sites of accumulation. Accumulation of the aurone hispidol and its glycoside derivative, hispidol-4'-*O*-glucoside was reported in yeast elicitor (YE)-induced cell cultures of *M. truncatula* (Fig. 3). Three peroxidases, MtPRX1, MtPRX2, and MtPRX3, which were induced parallel to accumulation of hispidol, were suggested to be involved in hispidol biosynthesis. MtPRX1 and MtPRX2 were shown to have aurone synthase activity *in vitro* (Farag et al., 2009)

2.2.7 Anthocyanin biosynthesis

Anthocyanins are the largest and most important group of water-soluble plant pigments in nature. The presence of a positive charge in the anthocyanin structure at acidic pH, called flavylium cation (2-phenylbenzopyrylium), makes anthocyanins different from other subgroups of flavonoids with the same C6–C3–C6 skeleton (Fig.8). Anthocyanidins are the aglycone form and are the chromophore responsible for the color variation. The number and position of hydroxy and methoxy groups on the anthocyanidin skeleton; the identity, number, and positions at which sugars are attached; and the extent of sugar acylation and the identity of the acylating agent are all responsible for the variation in anthocyanins (Prior and Wu, 2006). Color differences between anthocyanins are largely determined by the substitution pattern of the B-ring of the anthocyanidin, the pattern of glucosylation, the degree and nature of esterification of the sugars with aliphatic or aromatic acids as well as by the pH, temperature, type of solvent and the presence of co-pigments (Mazza, 2007).

Genes controlling the B-ring hydroxylation in early steps of flavonoid biosynthesis pathway of naringine, such as those encoding flavonoid 3′-hydroxylase (F3′H) and flavonoid 3′,5′-hydroxylase (F3′5′H) are key determinants for the structural fate of the resulting anthocyanins as well as 2,3-*cis*-flavan- 3-ol (Grotewold, 2006).

The first dedicated step towards biosynthesis of anthocyanins, and PAs, is controlled by dihydroflavonol-4-reductase (DFR) that uses the same substrates as flavonol synthase. Dihydrokaempferol, dihydroquercetin, and dihydromyricetin which differ only in the hydroxylation pattern of their B-ring, are common substrates of DFR and are converted into leucopelargonidin, leucocyanidin and leucodelphinidin, respectively (Xie et al., 2004a).

In *M. truncatula*, two DFRs, MtDFR1 and MtDFR2, have been characterized which contain some differences in their amino acid sequences and their enzymatic properties like distinct substrate preferences. Overexpression of *MtDFR1* induced changes in flower anthocyanin profiles in tobacco, while *MtDFR2* overexpression didn't induce such changes (Xie et al., 2004a). Based on the expression pattern of the two characterized MtDFRs, it has been proposed that both are involved in anthocyanin biosynthesis. However, MtDFR1 has a more dramatic role in anthocyanin (cyanidin glucoside) biosynthesis in the leave (Xie et al., 2004a).

The next step in anthocyanin biosynthesis is steered by anthocyanidin synthase (ANS) that catalyzes the conversion of leucoanthocyanidins (leucocyanidin, leucopelargonidin,

leucodelphinidin) to the corresponding anthocyanidins (cyanidin, pelargonidin, and delphinidin, respectively) (Fig. 8) (Nakajima et al., 2001). *In vitro* assays characterized MtANS as a bifunctional enzyme involved in conversion of leucocyanidin to cyanidin in the anthocyanidin/PA biosynthesis, and dihydroquercetin to the flavonol quercetin in flavonol biosynthetic pathway (Pang et al., 2007). *MtANS* is mainly expressed in the seed coat of *Medicago*, but it is also expressed in other tissues to play its role in anthocyanin/PA biosynthesis. Also, down-regulation of *MtANS* resulted in reduced levels of anthocyanins in *Medicago* leaves and of soluble and insoluble PAs in seeds (Pang et al., 2007).

Anthocyanidin glycosylation enhances their water-solubility and is critical for the transport and sequestration of these compounds in the vacuole. Anthocyanidin *O*-glycosylation is catalyzed by family I glycosyltransferases (UGTs), that comprise a superfamily in the plant kingdom (Yonekura-Sakakibara and Hanada, 2011; Yonekura-Sakakibara et al., 2012) and recognize the hydroxyl groups of a wide variety of flavonoid aglycones including anthocyanidins.

As mentioned before, UGT78G1 has been shown to have a role in the conversion of anthocyanidins (cyanidin and pelargonidin) to their corresponding 3-*O*-glucosides. Although it had been shown that isoflavones were the preferred *in vitro* substrates, yielding the corresponding 7-*O*-glucosides (Modolo et al., 2007). Increased accumulation of anthocyanin in transgenic alfalfa overexpressing *UGT78G1* as well as by retrotransposon insertion lines of *M. truncatula* signify its role in the glycosylation of anthocyanidins pigment. In addition, strong up-regulation of *UGT78G1* by Legume Anthocyanin Production 1 (LAP1), a transcription factors (TF) involved in anthocyanin biosynthesis but not isoflavone biosynthesis (Peel et al., 2009) indicated its role on glycosylation of anthocyanidin. Moreover, *UGT78G1* is expressed in anthocyanin accumulating sites (i.e. flowers, leaves, and buds) rather than in roots, which is the major site of accumulation of isoflavone in *Medicago* (Modolo et al., 2007).

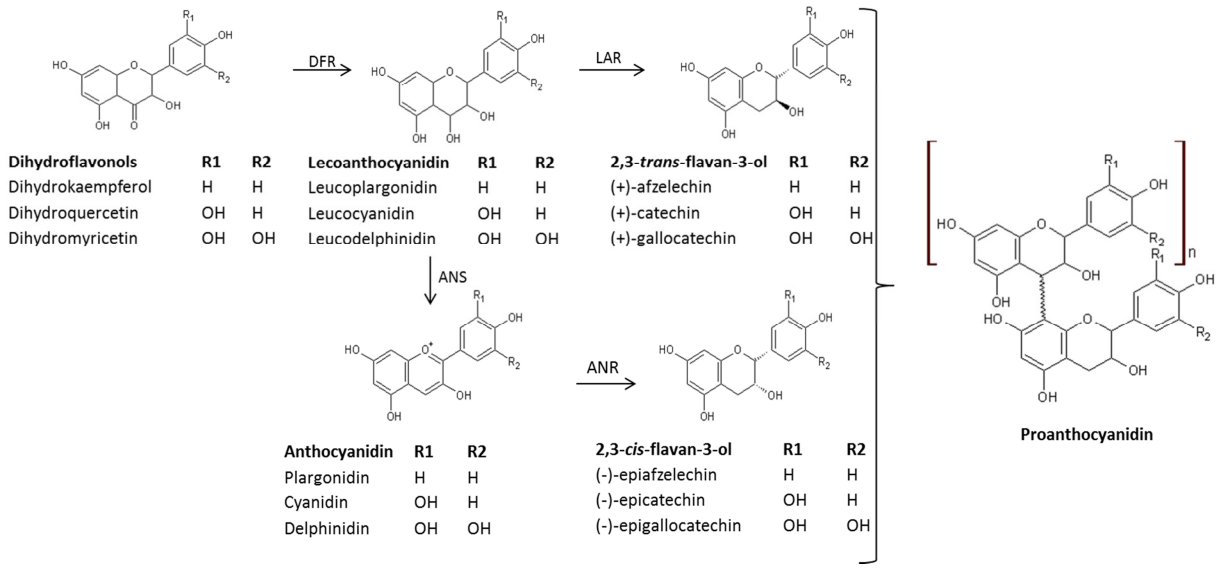


Figure 8. Anthocyanin and PA biosynthesis pathway in *M. truncatula.* DFR, dihydroflavonol-4-reductase; ANS, anthocyanidin synthase; LAR, leucoanthocyanidin reductase; ANR, anthocyanidin reductase.

2.2.8 PAs biosynthesis

PAs comprise mixtures of oligomers or polymers of flavan-3-ol units. Similar to all flavonoids, flavan-3-ols possess the typical C6-C3-C6 flavonoid skeletons (Fig. 8). Here also the hydroxylation pattern of the aromatic rings destines the structure of the resulting PAs. Hydroxylation on the A-ring determines stereochemistry at C-4. Stereochemistry of the building blocks have great importance in the biosynthesis of PA, since all chiral intermediates of PA biosynthesis pathway up to (+)-catechin possess the 2,3- *trans* stereochemistry. On the other hands, (-)-epicathecin with 2, 3-*cis* stereochemistry is arisen from their achiral anthocyanidin precursors providing another starter unit of PA (McKhann and Hirsch, 1994; Dixon et al., 2005).

Structural diversity of PAs depends on the stereochemistry and hydroxylation pattern of the flavan-3-ol units, the position and stereochemistry of the interflavanyl linkage between the monomeric units, the extent of polymer, and the type of modifications like various methyl, acyl or glycosyl substituents of the monomeric units (McKhann and Hirsch, 1994; Dixon et al., 2005). According to the location and stereochemistry of the linkage between the monomer units of flavan-3-ol and the extent of their polymerization PAs fall into three groups of A, B and C (McKhann and Hirsch, 1994; Dixon, 2005).

PAs share the same upstream biosynthetic pathway as anthocyanins and flavonols. In *M. truncatula* the structural genes encoding ANS, leucoanthocyanidin reductase (LAR), anthocyanidin reductase (ANR), and epicatechin 3'-*O*-glucosyltransferase (UGT72L1) all

involved in PAs biosynthesis have been characterized (Xie et al., 2003; Pang et al., 2007; Pang et al., 2008).

LAR is a member of the plant reductase-epimerase-dehydrogenase (RED) supergene family and closely related to IFR (Tanner et al., 2003). It uses the same substrates as ANS, flavan-3,4- diols (leucocyanidin, leucopelargonidin, leucodelphinidin) to convert them to their corresponding PA "starter units" 2,3-*trans*-flavan-3-ols (catechin, afzelechin and gallocatechin, respectively) in a NADPH-dependent manner. This first committed step in PA biosynthesis diverges from the pathway common with anthocyanins (Xie et al., 2003). In *M. truncatula*, the *MtLAR* gene is expressed in flowers, pods, and seed coats (Pang et al., 2007).

ANR catalyzes the conversion of the anthocyanidin resulting from ANS activity, cyanidin, pelargonidin, delphinidin, into 2,3-*cis*-flavan-3-ol, (-)-epicatechin (-)-epiafzelechin and (-)-epigallocatechin, respectively. ANR enzyme is encoded by the *BANYULS (BAN)* gene, identified through characterization of the banyuls locus in *Arabidopsis* (Devic et al., 1999). Later, the *BAN* gene and the ANR enzyme were described in detail in *M. truncatula* and *A. thaliana* (Xie et al., 2003; Xie et al., 2004b). In *M. truncatula* a single ANR enzyme encoded by a single *MtBAN* gene is responsible for the formation of epicatechin, epigallocatechin or epiafzelechin (Xie et al., 2003).

Similar to LAR, ANR also belongs to the isoflavone reductase-like (IFR-like) group of the plant RED super family. Although, the only difference between the two PA starter units, 2,3-*cis*-2R,3R-(-)- epicatechin and 2,3-*trans*-2R,3S-(+)-catechin, is the *cis*- or *trans*- stereochemical configuration, they are synthesized from two distinct biosynthetic pathways (Fig. 8) (Dixon et al., 2005).

Since PAs are typically sequestered in vacuoles in malonylated and glycosylated form, glycosylation has been proposed to be critical for the transport and storage of these compounds at their final destinations in the vacuole. UGT72L1, is a *Medicago* glycosyltransferase responsible for glycosilation of the PA precursor (-)-epicatechin leading to epicatechin 3′-*O*-glucoside. It was shown that expression pattern of *UGT72L1* in developing seeds is correlated with the presence of epicatechin glucoside and accumulation of PAs (Pang et al., 2008).

Three malonyltransferases, MaT4, MaT5 and MaT6 that are localized into the cytoplasm, endoplasmic reticulum (ER), and cytoplasm and ER respectively, are likely important for anthocyanin/flavonoid malonylation and their transport facilitation.

2.3 Regulation of the flavonoids biosynthesis

(Iso)flavonoids are known to play a significant role in responses against several environmental factors, including both biotic (such as pathogen attack and wounding) and abiotic stresses (such as UV-light, salt stress and nutrient deficiencies) (Dixon and Paiva, 1995; Shirley, 1996; Deavours and Dixon, 2005; Young and Udvardi, 2009; Zhou et al., 2011). In *Medicago* species, such as *M. sativa* and *M. truncatula*, isoflavonoid phytoalexins and related compounds accumulate in response to yeast or fungal elicitors (Shimada et al., 2003; Broeckling et al., 2005; Suzuki et al., 2005; Naoumkina et al., 2007; Farag et al., 2008; Naoumkina et al., 2008). Fungal pathogen infections are capable of massively changing the expression of phenylpropanoid biosynthesis genes. The fungus *Phymatotrichopsis omnivore*, responsible of destructive root rot disease in many dicot species including alfalfa, can induce phenylpropanoid biosynthesis genes in alfalfa during infection (Marek et al., 2009). In *M. truncatula*, increased levels of medicarpin precursors, formononetin 7-*O*-glucoside and malonylated formononetin 7-*O*-glucoside, were observed when infected with the fungal pathogen *Phoma medicaginis* (Jasinski et al., 2009). Accordingly, establishment of mycorrhizal fungus *Glomus versiforme* symbiosis in *M. truncatula* and *M. sativa* caused a transient increase in medicarpin levels (Harrison and Dixon, 1994).

Several studies have shown that in *M. truncatula* cell cultures the biosynthesis of the genes involved in the early steps of phenylpropanoid/isoflavonoid biosynthesis are induced by YE elicitation, while they are not induced by methyl jasmonate (MeJA) elicitation. However, the downstream pathway genes specific for medicarpin formation were induced by both elicitors (Naoumkina et al., 2007).

YE mimics pathogen attack mechanisms in the plant cell, while MeJA induces wound signaling cascades. This is consistent with the results obtained from phytoalexin accumulation during fungal infection and likely reflects different transcriptional regulatory systems orchestrating metabolic fluxes through specific secondary metabolites (Broeckling et al., 2005; Suzuki et al., 2005; Naoumkina et al., 2007; Farag et al., 2008; Naoumkina et al., 2008). Two different strategies have been proposed for the induction of medicarpin in response to pathogen (YE) and wound signals (MeJA) in *M. truncatula*. In non-stress condition, glycosylated and malonylated formononetin are sequestrated in the vacuoles. When MeJA or wound stresses are applied, formononetin glucoside are converted to free formononetin isoflavonoids and transferred to the cytosol. Concomitantly, the downstream enzymes are induced by MeJA and formononetin is

converted to medicarpin. Elicitation with YE on the other hand leads to the elevation of medicarpin levels via *de novo* biosynthesis (Naoumkina et al., 2007)

In *M. truncatula* four YE-induced WRKY genes (*W100577*, *W100630*, *W108715*, and *W109669*) have been characterized that are correlated with the genes involved in the central phenylpropanoid pathway and in the downstream steps of medicarpin biosynthesis. Ectopic expression of these WRKY TFs in tobacco (*Nicotiana tabacum*) also led to higher levels of flavonoids and other phenolic compounds (Naoumkina et al., 2008).

Most studies on the regulation of the flavonoid biosynthesis in various plant species have been conducted on anthocyanin and PA biosynthesis transcriptional regulation. This transcriptional regulation is accomplished via ternary complexes composed of the combinatorial interaction of three families of TFs R2R3-MYB domain, basic helix–loop–helix (bHLH) domain and conserved WD40 repeat (MBW) (Hichri et al., 2011).

Up to now, the only characterized regulator orchestrating the biosynthesis of anthocyanins in *M. truncatula* is *LAP1* gene (Peel et al., 2009). Overexpression of *LAP1* in *M. truncatula* and its closely related species including alfalfa and white clover (*Trifolium repens*) led to the accumulation of anthocyanin pigments. Constitutive expression of *LAP1* resulted in the up-regulation of a large number of genes associated with anthocyanin biosynthesis including the glucosyltransferase *UGT78G1* (Peel et al., 2009).

Two major components of the regulatory complex related to PA biosynthesis have been found in *M. truncatula,* so far: *M. truncatula* WD40-1 (MtWD40-1), a single-copy WDR TF (Pang et al, 2009) and *M. truncatula* proanthocyanidin regulator (MtPAR), an R2R3–MYB-type TF (Verdier et al., 2012).

MtWD40-1, is a single WD40-repeat TF orthologous to *Arabidopsis* TTG1 and has been implicated as a positive regulator of PA biosynthesis in *M. truncatula* seeds (Pang et al., 2008).

MtPAR plays a positive regulatory role on the genes involved in PA pathway (Verdier et al., 2012). The gene encodes a R2R3 MYB TF and is only expressed in the seed coat, where PAs are accumulated. In addition, expression of *MtPAR* in *M. sativa* was shown to cause the accumulation of PAs in shoots. Gene expression analyses of *Medicago par* mutants suggested that MtPAR acts upstream of MtWD40-1 and regulates its expression probably via direct interaction with WD40-1 (Verdier et al., 2012).

2.4 Flavonoid localization

Vacuole sequestration of flavonoid seems to be necessary for the majority of flavonoids such as isoflavonoids, anthocyanins and PAs to store them under certain physiological conditions. Once synthesized, anthocyanin pigments accumulate inside the vacuole where the vacuolar acidic environment and co-pigments determine anthocyanin-mediated floral pigment (Verweij et al., 2008). It is also believed that the PA subunits are polymerized and subsequently converted to brown oxidation products in the vacuoles. Isoflavonoids such as formononetin were also found to be sequestrated in the vacuoles in glycosylated and malonylated forms and they converted to free formononetin in response to certain stress and then move to the cytosol for conversion to medicarpin (Naoumkina et al., 2007).

Metabolons are multi-enzymes complex consist of specific interaction between several soluble enzymes involved in a secondary metabolite biosynthetic pathway. The enzymes centered around a membrane either by membrane-bound structural proteins that serve as nucleation sites or by membrane-anchored proteins such as CYP450s that directly catalyze one or more of the sequential channeled reactions performed by the metabolons. In flavonoid biosynthetic pathway the all CYP450 enzyme including C4H, F3'H, F3',5'H, IFS and 2HID are membrane-spanning proteins and thought to provide nucleation sites or platforms for self-assembling of soluble subunits on ER (Jorgensen et al., 2005). The ER multi-enzyme complex can facilitate the transport of flavonoid products through membrane trafficking, however, the general mechanisms of transport are still poorly understood (Zhao and Dixon, 2010).

The vesicle trafficking-mediated transport and membrane transporter-mediated transport have been generally proposed as two major mechanisms for flavonoid transportation (Grotewold and Davies, 2008; Zhao and Dixon, 2010).

Anthocyanins are accumulated first in vesicle-like structures, named anthocyanoplasts or in the anthocyanic vacuolar inclusions (AVIs) which are membrane-less proteinaceous matrices (Markham et al., 2000; Conn et al., 2003). Then, anthocyanoplasts and AVIs are covered by prevacuolar compartments (PVCs), an endocytic multi-vesicle compartment involved in ER-Golgi-vacuole vesicle trafficking and import into the central vacuole. PVCs contain Soluble N-ethylmaleimide-sensitive factor Attachment protein Receptors (SNAREs), which play an essential role in vesicle-mediated transport events (Hong, 2005). Specific interaction between v-SNAREs on the surface of transport vesicles (PVCs) and t-SNAREs on target compartment (Hong, 2005) governs general fusion

processes of vesicle-mediated transport (Zhao and Dixon, 2010). Besides, a direct Trans-Golgi network (TGN)- independent trafficking pathway from the ER to PVC exists in trafficking that enables the transport of anthocyanins through interaction of protein storage vacuoles (PSVs) and the anthocyanin-containing vesicle-like structures (Zhao and Dixon, 2010).

Two major mechanisms have been suggested for the membrane transporter-mediated transport of anthocyanins: primary transport mediated by multidrug resistance-associated protein (MRP)-type ABC [(the ATP binding cassette (ABC)] transporters (Goodman et al., 2004; Verrier et al., 2008) and proton gradient -dependent secondary transport that is mainly driven by V-ATPase and vacuolar H^+-pyrophosphatase (Zhao and Dixon, 2010). The vacuole-localized multidrug and toxic extrusion (MATE) transporters have been identified as the secondary transporter involved in flavonoid/H^+ exchange (Yazaki, 2005). Generally, the MATE superfamily is composed of 9 to 12 membrane- spanning domains (Schwacke et al., 2003) that use electrochemical gradients of protons or sodium ion gradients across membranes as a force to drive waste or toxic compounds out of the cytoplasm. In plants MATE transporters are responsible for the detoxification of xenobiotics and the transportation of a wide range of metabolites, including cations, organic acids and secondary metabolites.

In *M. truncatula* two MATE- type transporters known as MATE1 and MATE2 had been identified as H^+-gradient-dependent transporters of PAs and anthocyanin/ flavonol, respectively (Yu et al., 2008; Zhao et al., 2011). MATE1 was shown to transport epicatechin 3'-*O*-glucoside and its expression is confined to the seed coat of *M. truncatula. MATE1* is an ortholog of *Arabidopsis TT12* and could complement the seed PA deficiency phenotype of the *tt12* mutation in *Arabidopsis* (Yu et al., 2008). MATE2 has 12 putative transmembrane domains in the same pattern as other known MATE transporters. *MATE2* is expressed primarily in leaves and flowers and is involved in vacuolar sequestration of anthocyanins and other flavonoids in flowers and leaves. It has higher transport capacity for anthocyanins than for other flavonoid glycosides and in spite of its high similarity to MATE1, the PA transporter, it cannot effectively transport PAs precursors (Zhao et al., 2011).

The anthocyanins mostly accumulate as acylated forms in vacuole (Zhang et al., 2006) and it was reported that MATE2 specially transports glycosylated and malonyl-glycosylated flavonoid compounds that are ubiquitously found in *M. truncatula* and generated by the malonyltransferases (Zhao et al., 2011). It seems that malonylation

increases both the MATE2 affinity and transport efficiency of glycosylated flavonoid compounds (Zhao et al., 2011). Co-regulation of *MATE2* with *UGT78G1* and *MaT4* by *LAP1* in *M. truncatula* (Modolo et al., 2007; Peel et al., 2009) can demonstrate the role of MATE2 in transport of malonylated anthocyanin to vacuoles (Zhao et al., 2011).

In both vesicle trafficking and transporter-mediated systems, a glutathione *S*-transferase (GST) is involved in flavonoid transport. Although the full mechanism of action of GSTs in anthocyanin transport has not been shown yet, it has been suggested that it can bind to PAs, anthocyanins or flavonols to form a GST-anthocyanin or GST-flavonol complex protecting them from oxidation and/or guiding them to the central vacuole (Zhao and Dixon, 2010; Gomez et al., 2011).

2.5 Engineering of flavonoids biosynthesis in *Medicago*

Due to various functions in plants, animals and in human health, modifying the plant's ability, both quantitatively and qualitatively, to biosynthesize these bioactive natural compounds via metabolic engineering approaches is very important. In addition the putative role of isoflavonoids as pharmaceutical and health promoting compounds, has led to a great deal of interest to introduce them in non-legume plants such as *Arabidopsis*, tobacco, corn and tomato as well as an improvement of the nutritional value of soybean, a main source of isoflavonoids as dietary supplement, by modifying the total level of isoflavonoids in seeds. In addition, the critical role of isoflavonoids as signaling compound in legume-rhizobia interactions rises attention to introduce them in other economical important crops such as rice to develop nitrogen fixing root nodules (Ladha and Reddy, 2003; Sreevidya et al., 2006). Above all, flavonoids act as phytoalexins involved in plant defense and boosting the production of phytoalexins in legumes can increases their resistance to pathogens and ultimately lead to low-pesticide/fungicide farming (Deavours and Dixon, 2005).

Basic strategies that can be applied for genetic modulation of flavonoid pathways in plants are: (I) increasing endogenous flavonoid levels using up-regulation of structural or regulatory genes; (II) overexpression of heterologous structural and regulatory genes that are not present in the gene pool of the target plant (to open the pathway to new metabolites); and (III) blocking of specific steps in the flavonoid biosynthetic pathway by RNA interference strategies to switch the metabolic flux towards desired end products (Martens et al., 2003; Dixon, 2005).

Influence of the overexpression of several of the maize flavonoid regulatory genes, MYC/bHLH (*B-Peru*, *Sn*, *Lc*) and MYB (*C1*), on accumulation of anthocyanins in heterologous plants were investigated (de Majnik et al., 2000; Bovy et al., 2002; Robbins et al., 2003). However, the expression of *B-Peru* gene (a bHLH) and the MYB type *C1* (*colourless1*) failed to induce anthocyanins in alfalfa and the transgenic alfalfa populations expressing the *Lc* (*leaf color*) gene showed a slight induction of anthocyanin production when the plants were exposed to the abiotic stress (Ray et al., 2003).

The Arabidopsis Production of Anthocyanin Pigment 1 (PAP1) TF is a general regulator of the anthocyanin biosynthetic pathway, and its ectopic over-expression activates the anthocyanin biosynthetic pathway in *Arabidopsis* and *Nicotiana tabacum* (Borevitz et al., 2000; Tohge et al., 2005). However, PAP1 was not able to induce anthocyanin biosynthesis in *M. truncatula* or in *M. sativa*. Unlike PAP1, the *Medicago* LAP1 induced anthocyanin production in transgenic *M. sativa*, *M. truncatula*, or *Trifolium repens* when ectopically expressed. Hence, LAP1 could be used as a general regulator for anthocyanin engineering in legumes (Peel et al., 2009).

MYB TFs play a central role in the PA biosynthetic pathway. Ectopic expression of *AtTT2* (encoding an R2R3 MYB TF involved in the regulation of PA biosynthesis) in *M. trunculata* hairy roots caused a massive accumulation of PAs as a result of the activation of the genes involved in the biosynthesis, transport and oligomerization of PAs (Pang et al., 2008). Recently it was shown that ectopic expression of *MtPAR* in *M. truncatula* hairy roots increased the expression of the genes involved in PA/anthocyanin biosynthesis, such as *CHS*, *F3H*, *ANS*, and *ANR*, which consequently led to a massive accumulation of PAs (Verdier et al., 2012). Overexpression of *MtPAR* in alfalfa also resulted in the accumulation of PAs in the shoots (Verdier et al., 2012). Some efforts to engineer PAs were done using more than one TF or a combination of TFs and biosynthetic genes to increase PA yields in transgenic plants. Co-expression of *PAP1* and *MtBAN* in tobacco resulted in a significant decrease in the anthocyanin levels, and higher content of the PA precursors (Xie et al., 2006).

The attempts toward induction of isoflavonoid production are mostly focused on the overexpression of the gene encoding the key enzyme of the isoflavonoid biosynthesis, IFS, in transgenic plants. The expression of *M. truncatula IFS1* in *M. sativa* resulted in the accumulation of genistein glycosides in the transgenic plants (Deavours and Dixon, 2005; Shih et al., 2006; Sreevidya et al., 2006; Dhaubhadel, 2010).

2.6 Flavonoid biological function in *Medicago*

Flavonoids are very widespread in nature and are one of the best characterized and largest groups of natural plant products. Since they were already present in the most early plant lineages and evolved together with the migration of plants out of the sea to conquer the land, they have adopted many different crucial functions in plants. This holds also true for *Medicago* species, of which the most important functions of flavonoids in plants are highlighted below.

2.6.1 Role in plant

2.6.1.1 Nutrient acquisition

Flavonoids can affect the nutrient acquisition of a plant through soil chemical changes, either via direct or indirect contributions to the availability of nitrogen (N), phosphorous (P) and iron (Fe) (Cesco et al., 2010) and *vice versa*. N, P and Fe supply in the soil each affect flavonoid biosynthesis (Hassan and Mathesius, 2012; Zamboni et al., 2012). The indirect involvement of flavonoids in enhancing the uptake of two of the most important macronutrients, N and P, by stimulating plant-microbe symbiosis will be discussed in another section (see 2.6.1.4. Plant-microbe interactions). Here, we will only discuss the direct biological function of these phenolic compounds to nutrient acquisition. Flavonoids are likely exuded from roots through an active mechanism, often in response to elicitors, but can also be released passively from decomposing root cap and border cells (Hassan and Mathesius, 2012). The active process may involve plasma membrane-localized ATP-binding cassette transporters (Dixon and Pasinetti, 2010). Since flavonoids can act as metal chelators, they have a putative function for chemical mobilization of scarcely soluble soil-P forms. Possible mechanisms include (I) exchanging chelation at the cationic P-binding sites, (II) occupation of P-binding sites and (III) iron-phosphate splitting by Fe reduction due to reducing properties of flavonoids (Cesco et al., 2010). For example, an isoflavonoid identified in root exudates of *M. sativa* was able to dissolve ferric phosphate, thus making both phosphate and iron available to the plant (Masaoka et al., 1993). In all higher plants, except Gramineae, phenolic compounds are reported to be the main components of root exudates in response to Fe-deficiency (Cesco et al., 2010). Flavonoids, including the isoflavone genistein and the flavonols quercetin and kaempferol, can alter iron availability by reducing Fe(III) to Fe(II) and by chelating iron otherwise unavailable in iron oxides and/or poorly soluble iron minerals (Cesco et al., 2010). Even though the reduction of copper (Cu) and its complexation by flavonoids has

clearly been demonstrated and chemically characterized, the contribution of this phenomenon to Cu availability and plant uptake has not yet been evaluated (Cesco et al., 2010). Although, the role of flavonoids in plant-soil interactions and nutrient acquisition has already been widely established since a long time, a lot of open questions and points to be addressed still remain. For example, detailed information on the dynamic composition, the concentrations, the microbial modifications and persistence of flavonoid profiles released in the rhizosphere, is often lacking or contradicting (Cesco et al., 2012).

2.6.1.2 Plant pigmentation

Flavonoids and especially their colored class, anthocyanins, contribute to a lot of biological functions within plants, of which the most obvious one is in reproduction as attractants for pollinators via flower pigmentation and for the dispersal of seeds via brightly colored fruits. On the contrary, pollination actually serves as a rapid trigger for flavonoid biosynthesis (Dong et al., 1998). Structural studies on pelargonidin, cyanidin and delphinidin, the major anthocyanidins in *M. truncatula*, revealed that an increase in hydroxyl groups confers a color shift towards the longer wavelength end of the spectrum i.e. more blue: pelargonidin tends to yield orange to intense red, cyanidin yields red to magenta and delphinidin yields blue to violet (Tanaka et al., 2010). Other modifications giving rise to their structural diversity also influence their color. Methylation of the 3' or 5'-hydroxyl group or glycosylation of anthocyanins causes a shift towards a slightly redder color, while aromatic acylation does the opposite and increases their stability (Tanaka et al., 2008). Even though anthocyanins are the most substantial ones, other flavonoid classes also influence flower color. Aurones and chalcones provide yellow color (absorption of blue wavelengths) and some flavonols, flavones, flavan-3-ols and isoflavones are pale yellow (Di Meo et al., 2012). However, most of the latter are colorless to the human eye, but can be detected by bees and other insects, which see much farther in the ultraviolet (UV) range than humans. As such they are often responsible for the formation of symmetrical patterns of stripes, spots or circles called nectar guides (Sasaki and Takahashi, 2002). In addition, stacking of anthocyanins with co-pigments such as flavones and flavonols results in a bathochromic shift towards a more intense and blue color. Only recently, using quantum chemical calculations to provide a detailed molecular orbital picture researchers have tried to obtain a complete spectroscopic understanding of anthocyanin - flavonol copigmentation (Tanaka et al., 2010; Di Meo et al., 2012). The production of many colors is dependent on the ability of flavonoids to

form complexes with metal ions, e.g. Fe (III) or Al (III) (Tanaka et al., 2010). pH is another important factor influencing anthocyanin color. Acidic vacuolar pH gives a redder color, while a neutral vacuolar pH tends to yield a bluer one. In addition, anthocyanidins are unstable at neutral pH and must be stabilized by glycosylation and acylation, both processes influencing their color (Tanaka et al., 2010).

2.6.1.3 Photoprotective function

2.6.1.3.1 Photoprotective screen versus anti-oxidant function

Most flavonoids and more specifically several flavonols play an important role in protecting the plant against oxidative damage associated with exposure to highly energetic, short wavelength solar light. Even under normal conditions the photosynthetic electron transport system in plants is a major source of active oxygen species. In theory, flavonols could accomplish this role via their two important characteristics, i.e. ultraviolet (UV) screen and antioxidant properties. It was long believed that this photoprotection was primarily due to the UV screening properties of certain flavonoids and their predominant localization in epidermal cell layers (Treutter, 2006). In addition, enhanced accumulation of many flavonoid classes, e.g. chalcones, flavonols and isoflavones, can be observed after UV treatment. Over-expression of IFS from *M. truncatula* in *M. sativa* resulted in an additional accumulation of the isoflavones formononetin and daidzein after UV-B treatment (Deavours and Dixon, 2005). However, because of some recently obtained insights this view is slowly changing. The spectrum of UV radiation reaching the Earth's atmosphere can be divided according to the different wavelengths into low energy UV-A (320-400 nm), higher energy UV-B (280-320 nm) and high energy UV-C (254-280 nm). Because UV-C light cannot penetrate the Earth's protective ozone layer, UV-B light poses the highest threat and causes the most damage to plant tissues and cellular processes. UV-B-responsive flavonols, like quercetin, display the greatest antioxidant potential, but not the greatest UV-B attenuating capacity (Pollastri and Tattini, 2011). UV-B responsive flavonols generally have the maximum molar extinction coefficients (ε_{max}) in the wave bands higher than 335 nm, as compared to e.g. hydroxycinnamic acid derivatives with ε_{max} in the 280-320 nm wave bands. Therefore, they are not as efficient as other flavonoids to absorb wavelengths in the 280-320 nm spectral region and hence do not equip the leaf with the most effective shield against UV-B irradiance (Agati and Tattini, 2010). The ability of flavonoids to act as antioxidants depends on their molecular structure and the position of hydroxyl groups. Structural studies determined the main

features required for efficient radical scavenging: (I) an *ortho*-dihydroxy (catechol) structure in the 3',4'-position of the B-ring for electron delocalization, (II) an unsaturated 2,3 double-bond in conjugation with a 4-*oxo* function in the C-ring that provides electron delocalization from the B-ring for flavonoids lacking a catechol group, and (III) hydroxyl groups at positions 3 and 5 to provide hydrogen bonding to the oxo group (Cesco et al., 2010; Bodewes et al., 2011; Prochazkova et al., 2011). For example, quercetin, having the greatest antioxidant potential, is a flavonol that fulfils all the above mentioned criteria. Additionally, an OH-group in the 3-position on the C-ring of the flavonoid skeleton is a key structural feature responsible for the ability of flavonols to chelate transition metal ions, therefore inhibiting the generation of free radicals and reducing reactive oxygen species (ROS) once formed (Pollastri and Tattini, 2011). Glycosylation of this OH-group has strongly suppressive effects on the antioxidant activity (Prochazkova et al., 2011). In addition, these flavonols also accumulate in the mesophyll, not only in the epidermal cell layers as was originally reported to be the only site of flavonoid accumulation. Moreover, this accumulation could be seen in leaves exposed to full sunlight, in the presence or absence of UV irradiance (Agati and Tattini, 2010). This finding leads to the interesting hypothesis that excess light-induced oxidative damage may regulate the biosynthesis of flavonoids, irrespective of the proportion of solar wavelengths reaching and penetrating the leave. Nevertheless, the widely accepted antioxidant function of flavonoids in plants is still a matter of great debate. In established *in vitro* antioxidant tests the antioxidant capacities of flavonoids are several-fold higher than those of ascorbate (vitamin C) or α-tocopherol (vitamin E), two well known *in planta* antioxidants (Hernandez et al., 2009). However, an *in vivo* antioxidant can only be truly functional if the end-products of the oxidation process are harmless for the plant cell. Since many different flavonoid oxidative intermediates have been identified, it is difficult to determine the end-products of a specific flavonoid (Hernandez et al., 2009). Furthermore, many oxidation processes generate flavonoid radicals, which are highly reactive and subject to further oxidation, yielding, among others products, the more stable flavonoid quinones. These are still reactive but can be stabilized by conjugation with nucleophiles, such as glutathione, cysteine or nucleic acids, thus being responsible for their prooxidant activities (Prochazkova et al., 2011). It should be taken into account that once flavonoids are stored in the vacuole, they are separated from the main sources of ROS, decreasing their functional relevance as *in planta* antioxidants (Hernandez et al., 2009). In conclusion, experimental evidences suggest that the general antioxidant role of flavonoids in plants is

actually only limited to a few individual flavonoids under specific experimental and developmental conditions and that several flavonoids even might have prooxidant activities and cause DNA damage, which puts a toll on their other beneficial functions (Hernandez et al., 2009; Prochazkova et al., 2011).

2.6.1.3.2 Energy escape-valve

Aside from the already mentioned roles of flavonoids as UV screens and scavengers of ROS in protecting the plant cells against oxidative damage associated with excess light, they might also have a third role as energy escape valves. Most abiotic stresses, which also induce flavonoid biosynthesis, create an imbalance between the amount of energy that is received and the photosynthetic capacity to process it (Hernandez and Van Breusegem, 2010). This excess light can produce ROS and by-products that can potentially cause photo-oxidative damage and inhibit photosynthesis (Li et al., 2009). It was already proposed that phenylpropanoid metabolism might provide an additional layer of photoprotection due to its ability to consume photochemical reducing power and act as an alternative carbon sink under excess light conditions (Grace and Logan, 2000). More specifically, flavonoids might act as energy escape outlets by dissipating such excess energy. The biosynthesis of flavonoids serves as a more efficient excess energy combustion system compared to the synthesis of other phenylpropanoids, because it consumes more energy (as reduction equivalents and ATP) and photoassimilates without sequestering any N or P. According to the carbon-nutrient balance hypothesis photosynthate is used primarily for the production of carbon-based metabolites such as phenolics when the carbon-to-nutrient ratio is high (Grace and Logan, 2000). As such, in contrast to other molecules with high carbon-to-nutrient ratios such as starch or sucrose, flavonoids have additional anti-stress activities and thus serve multifunctional roles in general and especially excess light-induced stress responses (Hernandez and Van Breusegem, 2010).

2.6.1.4 Plant-microbe interactions

Flavonoids have an important role as signalling molecules in beneficial plant-microbe interactions, of which the two most significant ones are the process of nodulation and the process of arbuscular mycorrhization (AM). Nodule formation is a complex process resulted from symbiotic relation between plant and nitrogen fixing rhizobia, in which the bacteria fix atmospheric nitrogen for the plant in exchange for organic sugars. Symbiosis mainly involves two main groups of molecules: the plant produces Nodulin (*nod)* gene-

inducing flavonoids and in response the bacteria secrete mitogenic lipo-chito-oligosaccharide nod factors (NFs) (Cooper, 2007; Jones et al., 2007). These NFs trigger early plant responses involved in root infection and nodule formation (Cesco et al., 2010). Even though around ten thousand of flavonoids have been identified so far, only about 30 *nod* gene-inducing flavonoids have been isolated from nine legume genera and one of the first discovered to possess *nod* gene inducing activities was luteolin, isolated from *M. sativa* (Peters et al., 1986; Cooper, 2007; Cesco et al., 2010). Flavones were shown to be the most potent inducers of *nod* genes in *Sinorhizobium meliloti,* which colonizes *M. truncatula* (Wasson et al., 2006). In addition, mixtures can be more effective than single compounds and some molecules act as inducers for certain rhizobia and as anti-inducers for others (Cooper, 2007; Hassan and Mathesius, 2012). Aside from their *nod* gene-inducing activity most of these flavonoids also act as chemo-attractants, thereby concentrating compatible rhizobia at the root surface (Cooper, 2007). In a study with transgenic *M. truncatula* roots with different flavonoid profiles flavonoid-deficient roots exhibited a near complete loss of nodulation, whereas flavone-depleted roots had reduced nodulation and isoflavone-deficient roots nodulated normally (Zhang et al., 2009a). It was previously shown that some isoflavonoids from *M. sativa*, such as medicarpin and coumestrol, even repress NF production (Zuanazzi et al., 1998). Whereas, formononetin and its glycoside ononin isolated from *M. sativa* were able to counteract the autoregulation of nodulation (Catford et al., 2006). The difference in nodulation between flavonoid- and flavone-depleted roots could be attributed to the role of flavonols, in particular kaempferol (Zhang et al., 2009a). Nonetheless, with the exception of certain photosynthetic *Bradyrhizobia,* rhizobial symbiosis is strictly dependent on the production of bacterial NFs that is induced by flavonoids (Giraud et al., 2007). Flavonoids are detected by rhizobia through an assortment of NodD proteins, which belong to the LysR family of transcriptional regulators. The flavonoids interact with the constitutively expressed *nodD* gene products to form a protein-phenolic (NodD-flavonoid) complex, which binds to highly conserved DNA motifs, called *nod* boxes, located in the promoters of inducible *nod* operons of *nod* genes. Secretion of appropriate flavonoids by the plant and the ability of the rhizobia to perceive and transduce this signal to NF biosynthesis are the earliest steps that determine host-specificity (Subramanian et al., 2007; Hassan and Mathesius, 2012). Flavonoids have also been shown to regulate a number of other rhizobial genes important for nodulation, such as those for exopolysaccharide production and type III secretion systems (Hassan and Mathesius, 2012). Aside from this essential

role, in some legumes that form indeterminate nodules, such as *M. truncatula,* flavonoids cause an inhibition in auxin transport causing local auxin accumulation, which precedes the initiation of cell division and is thus needed for correct nodule primordium development. Moreover, these plant molecules act as endogenous Nod signal inducers inside the plant roots, which is believed to be responsible for an additional level of host specificity. Experiments in *M. truncatula* indicated that both flavones and flavonols play distinct roles during nodulation, the former as internal inducers of rhizobial *nod* genes and the latter as auxin transport regulators (Subramanian et al., 2007; Zhang et al., 2009a).

The other important form of plant-microbe symbiosis is the process of arbuscular mycorrhization (AM), occurred between almost all land plants and fungi of the *phylum Glomeromycota*. The fungal partner extends the underground root system of the host, thereby greatly enhancing the uptake of water and nutrients, such as phosphate and nitrogen. In return, the plant invests up to 20 % of its fixed carbon into the fungus (Parniske, 2008). Some flavonoids compounds exuded from host plant stimulate processes promoting AM, e.g. spore germination, hyphal growth, hyphal branching in the soil, root colonization and infection (Hassan and Mathesius, 2012). In studies with *M. sativa,* roots began to accumulate flavonoids prior to colonization by *Glomus intraradices,* indicating elicitation by an AM fungi-derived signal. Moreover, in contrast to non-colonized roots, flavonoid profiles changed over the time-course of colonization in *M. sativa* and *M. truncatula* colonized with *Glomus mosseae* and *Glomus versiforme*, respectively (Shaw et al., 2006). However, contradicting effects have been reported based on the type of flavonoid molecule, the nutrient status, the developmental stage, the degree of colonization and the target organism under study (Cesco et al., 2012). Hyperoside obtained from *M. sativa* seeds, could stimulate spore germination in *Glomus etunicatum* and *Glomus macrocarpum,* while other flavones were only active in the *Glomus etunicatum*, and formononetin inhibited spore germination in both (Tsai and Phillips, 1991). Ononin can counteract autoregulation of AM in *M. sativa*, while formononetin is not able to suppress autoregulation (Catford et al., 2006). Interestingly, in contrast to nodulation some isoflavonoids have also been shown to promote AM. For example, coumestrol has been identified as an active stimulator of hyphal growth in *M. truncatula* and a mutant hyperaccumulating the compound was found to be hyperinfected by its mycorrhizal symbiont (Morandi et al., 2009). Additionally, there are some evidences that flavonoids have been implicated in the same effects on ectomycorrhizal fungi (Cesco et al., 2012).

However, some contradicting reports about the flavonoid-mediated regulation of beneficial plant-microbe interactions indicate the process appears to be highly complex and requires more detailed investigations.

2.6.1.5 Defense

Not all plant microbe-interactions are beneficial and flavonoids are also known to play an important role in the defense against plant pathogens. In general, a distinction has been made between compounds that are produced *de novo* upon pathogen challenge, the phytoalexins, and those that are already preformed and stored as a baseline defense, the phytoanticipins. Flavonoids are used by the plant as both and protect against a wide range of pathogens, ranging from bacteria and fungi to insects and nematodes (Treutter, 2006). Simple isoflavone compounds, such as daidzein, glycitein and formononetin glycosides behave as phytoanticipins, because they accumulate constitutively and the corresponding aglycones restrict the growth of microbial pathogens (Shaw et al., 2006). Global gene expression profiling has shown an elevation in flavonoid biosynthesis when *M. truncatula* plants were challenged by *Phymatotrichopsis omnivore* (Uppalapati et al., 2009). Isoflavonoids represent a major class of phytoalexins in legume plant species. Furthermore pterocarpans, such as medicarpin and maackiain, have antimicrobial properties and are produced either constitutively or after induction by pathogens or endogenous elicitors (Shaw et al., 2006). For example, medicarpin from *M. sativa* protects the plant from the pathogenic fungus *Rhizoctonia solani* (Naoumkina et al., 2010a). DNA microarray analysis revealed that (iso)flavonoid pathway in *M. truncatula* cell cultures can be induced by MeJA and YE, mimicking wounding and pathogen attack, respectively (Naoumkina et al., 2007). Moreover, inoculation of transgenic alfalfa overexpressing isoflavone *O*-methyltransferase with *Phoma medicaginis* resulted in a more rapid and increased production of medicarpin, which decreased disease symptoms (He and Dixon, 2000). Over-expression IFS from *M. truncatula* in alfalfa confirmed these results, with an additional accumulation of the isoflavones formononetin and daidzein after *P. medicaginis* infection (Deavours and Dixon, 2005). In addition, alfalfa seedlings challenged with the fungal pathogen *Colletotrichum trifolii* exhibited a defense response that was accompanied with increased expression of flavonoid biosynthesis genes and accumulation of medicarpin and sativan (Saunders and O'Neill, 2004). The same was observed from microarray analysis when *M. truncatula* infected with *Erysiphe pisi,* the causative agent of powdery mildew. Seven of the eleven enzymes required for medicarpin

biosynthesis were strongly upregulated and associated with the hypersensitive response (HR) a defense response that leads to localized cell death after certain avirulence gene products from the pathogen are recognized by plants carrying the corresponding resistance (*R*) genes (Foster-Hartnett et al., 2007). On top, the flavanone naringenin was shown to interfere with the quorum sensing– controlled production of virulence factors from *Pseudomonas aeruginosa* PAO1 (Vandeputte et al., 2011). Flavonols, such as quercetin, also have strong antimicrobial properties (Naoumkina et al., 2010a). In addition, flavonoids are known to be highly effective against some insects and nematodes. Several insects are sensitive to flavonoids and are deterred in feeding tests, because they can behave as antifeedants, digestibility reducers and toxins (Treutter, 2006). For instance, the phytoalexin medicarpin inhibited the motility of *Pratylenchus penetrans* in alfalfa (Baldridge et al., 1998) and accumulation of isoflavonoids in response to infection with the stem nematode *Ditylenchus dipsaci* correlated with resistance and was even induced systemically (Edwards et al., 1995).

2.6.1.6 Allelopathy

Aside from their well-known role in plant defense and in interacting with other organisms flavonoids also play an important role as allelochemicals in plant-plant interactions (Cesco et al., 2012). Only a few detailed studies have been performed on the direct inhibitory effects of flavonoid exudates on the growth and development of neighboring plants in leguminous species (Bido et al., 2010). However, because of the limited number of studies their direct role or mode of action is still not unequivocally known. Allelopathic interactions are not always caused by direct toxicity of the allelochemicals themselves, but can also be induced by biotic or abiotic structural modifications caused by flavonoids in the rhizosphere, because they induce redox reactions in soils and selectively influence the growth of soil microorganisms. This in turn influences the hormonal balance, enzymatic activity, availability of phytonutrients and competition between neighboring plants. As a result of this dynamic and ever-changing interaction, the structure, the chemistry and the microbial composition of the soil are altered significantly by the release of these plant phenolics (Bhattacharya et al., 2010; Cesco et al., 2012). For example, flavonoids can represent an important carbon source to those microorganisms in possession of appropriate enzymes to degrade them. Shaw *et al.* (Shaw et al., 2006) summarized some of these studies that have quantified and characterized aerobic flavonoid biodegradation for a number of bacterial species, both beneficial as

pathogenic including *Rhizobia, Pseudomonas, Agrobacterium,* etc. Biodegradation of flavonoids is one mechanism by which 'non-target' bacteria may cope with the toxic concentrations of flavonoids; however, many others have evolved an inducible resistance mechanism. For instance, *Agrobacterium tumefaciens* also possess an isoflavonoid-inducible isoflavonoid efflux pump which contributes significantly to its colonization of alfalfa roots (Palumbo et al., 1998). Since both strategies give an ecological advantage in microbial rhizosphere competiveness, the exudation of flavonoids in the rhizosphere puts a selective pressure on soil microbiota and thus contributes in shaping the rhizosphere ecosystem. Another well-known phenomenon illustrating this is called replant disease, also known as autotoxicity or soil sickness. It is an example of intraspecific allelopathy, whereby the plant produces allelochemicals when it starts decaying in the soil and these allelochemicals are detrimental to the establishment of new seedlings of the same species. The observed damping-off of *M. sativa* seedlings by the fungal pathogens *Pythium spp*. and *Rhizoctonia solani* has been attributed to autotoxicity of undecomposed *M. sativa* plant residues (Bonanomi et al., 2011).

Although not that many studies have been actually performed on allelopathy in *Medicago spp.,* all above given examples will likely also hold true for the *Medicago* family of leguminous plants.

2.6.1.7 Developmental regulators

As it turns out, antioxidant flavonoids, aside from their important role in ROS homeostasis, are also key developmental regulators. They are the most effective inhibitors of basipetal auxin transport (Agati and Tattini, 2010), as already mentioned in a previous section to also be important for correct nodule development (Zhang et al., 2009a). RNAi mediated silencing of flavonoid biosynthesis in *M. truncatula* hairy roots led to increased auxin transport, indicating that flavonoids also act as auxin transport inhibitors in this species (Wasson et al., 2006; Subramanian et al., 2007). However, although they modulate auxin transport and interact with auxin transporters, they are not specific regulators. Otherwise they would not have so many potential modes of action and targets in interfering with auxin transport, such as auxin transporters, kinases and trafficking machinery (Peer et al., 2011). For example, it has been shown that flavonols are able to bind to and thereby inhibit the activity of ABC subfamily B (ABCB) transporter proteins, which are necessary for directional auxin transport. However, it has also been shown that they disrupt the proper folding of ABCBs and inhibit ATPase activity via interference

with phosphorylation or allosteric binding (Peer et al., 2011). In particular, ABCBs are involved in naphthalene-1-acetic acid (NAA) transport and it has recently been shown that flavonols affect the export of NAA, but not of the native auxin, indole-3-acetic acid (IAA) (Kuhn et al., 2011). Additionally, flavonoids are also versatile modulators of the distribution of another class of auxin efflux transporters, i.e. the auxin efflux facilitating PIN proteins (Kuhn et al., 2011; Peer et al., 2011). On the other hand, high sunlight induces the synthesis of both auxin and quercetin derivatives and increases the activity of phenol-oxidizing peroxidases (Pollastri and Tattini, 2011). Quercetin displays a great capacity for fine regulating auxin gradients as well as local auxin concentrations, which represent the actual determinants for different morphological responses (Pollastri and Tattini, 2011). As a result, some have suggested that the high-light induced biosynthesis of antioxidant flavonoids may have a role in regulating whole-plant and individual organ architecture (Agati and Tattini, 2010). This makes flavonols good candidates to affect the 'flight response' of sessile plants, i.e. the stress-induced redistribution of growth (Pollastri and Tattini, 2011).

2.6.2 Flavonoid pharmacological properties

Besides the important biological functions of flavonoids in plants, they are significant components of the human diet and exert wide range of pharmacological properties (Cazarolli et al., 2008). This is already been known for a long time, since in many cultures plant-derived infusions, poultices, balms and spices containing flavonoids as active constituents have been used in traditional medicine for centuries. They use conventionally for prevention and treatment of various infectious and toxin-mediated diseases, such as sores, infected wounds, acne, respiratory infections, gastrointestinal disease and urinary tract infections (Cushnie and Lamb, 2011). Since flavonoids are key nutraceuticals of the human diet and also since they have been implicated in the prevention of a wide range of physiological disorders and diseases, numerous studies are conducting on flavonoids and their impact on human and animal health. Flavonoids have been reported as anti-inflammatory, antibacterial, antiviral, antiallergic (Cook and Samman, 1996; Cushnie and Lamb, 2005), cytotoxic antitumor, anti-hepatotoxic and anti-ulcer agents. Flavonoids have shown that are effective in neurodegenerative diseases, exert vasodilator action and it seems that they provide protection against cardiovascular mortality (Chebil et al., 2006; Chang et al., 2010; Tsuchiya, 2010). They are known to inhibit lipid-peroxidation, platelet aggregation, capillary permeability and fragility, cyclo-

oxygenase and enzyme activities. In addition, flavonoids are potent antioxidants and have free radical scavenging abilities (Cook and Samman, 1996; Middleton et al., 2000; Chebil et al., 2006). They are also reported to inhibit enzymes such as aldose reductase and xanthine oxidase (Cazarolli et al., 2008; Rashidi and Nazemiyeh, 2010).

This potential resides in a number of biological properties, including their antioxidant abilities, their interactions with intracellular signaling pathways, regulation of cell survival/apoptotic genes and mitochondrial function (Pan et al., 2010; Spencer et al., 2012).

Among *Medicago* species, *M. sativa* is important as forage and also as an industrial source of leaf protein concentrate used in animal diets. In addition, alfalfa sprouts are often consumed as vegetable salad. Therefore, many studies on the biological activity of the *Medicago* natural products, especially on their effect on animals, have been carried out on the metabolites obtained from this species. The extracts from alfalfa sprouts, leaves, and roots have been reported to reduce cholesterol levels in animal and human studies (Lee et al., 2005; Tava and Avato, 2006). In addition, alfalfa sprouts or leaves has been traditionally used for treating of arthritis, kidney problems, and boils (reviewed in (Hong et al., 2009a). The ethyl acetate extract of alfalfa sprouts have been shown to ameliorates the autoimmune-prone disease of lupus in mice, probably by attenuating cytokine and inflammatory responses and also can alleviate acute inflammatory hazards (Hong et al., 2009a; Hong et al., 2009b). Although many flavonoid compounds with putative pharmacological properties reported from other plant species, such as quercetin and luteolin etc., are found in *Medicago* species, to the best of our knowledge there are limited reports about pharmacological properties of flavonoids from *Medicago* species. Bickoff et al. (Bickoff et al., 1964) showed that tricin obtained from *M. sativa* caused smooth muscle relaxation on intestinal strips of guinea pig. However, it showed slight activity as an antioxidant. The most well-known pharmacological property of flavonoids derived from *Medicago* species is related to their isoflavonoid profile that exert phytoesterogenic activities, as described below.

2.6.2.1 Phytoestrogenic

Isoflavonoids are a unique subgroup of flavonoids found predominantly in the species of the Leguminosae family. They possess a chemical structure that is similar to the hormone estrogen (Fig. 9), and are able to interact with estrogen receptor. Thereby isoflavonoids are often referred to as phytoestrogens (McCue and Shetty, 2004). Soybean (*Glycine max*)

is the legume plant species containing isoflavonoids via which humans have been traditionally exposed to the most. Therefore, numerous epidemiological and experimental studies are available that imply a preventive effect of soy and its associated isoflavones against chronic diseases such as osteoporosis, menopausal disorders and breast- and prostate cancer (Franke et al., 2009). Diet supplementation with soybean phytoestrogens has been reported to reduce hot flashes and other symptoms of post-menopausal women (McCue and Shetty, 2004). The major isoflavones found in soybean are daidzein and genistein that are also produced by *M. truncatula* and *M. sativa* (Urbanczyk-Wochniak and Sumner, 2007; Cazarolli et al., 2008; Bora and Sharma, 2011). The major phytoestrogenic isoflavone from alfalfa sprouts, another common food for humans, is coumestrol. However, in a recent study researchers identified other phytoestrogenic flavonoids in alfalfa, such as (iso)liquiritigenin that was able to transactivate the estrogen receptors α and β (Hong et al., 2011). Since estrogen normally affects calcium metabolism positively, common disorders associated with menopause are bone thinning, bristle-bone disease and osteoporosis. These can be counteracted by hormone (estrogen) replacement therapy, but unfortunately this also increases the risk for estrogen-linked cancers. It has been reported that genistein and daidzein are selective estrogen receptor modulators that can act against bone loss without the negative side-effects associated with estrogen treatment (McCue and Shetty, 2004).

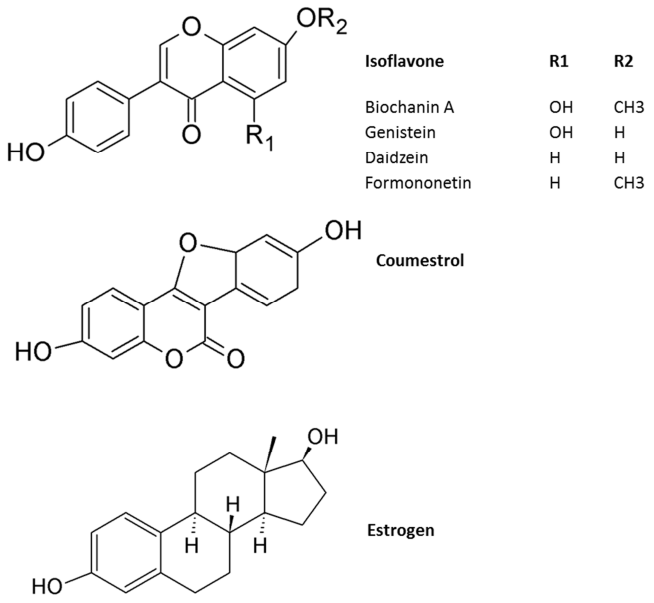


Figure 9. Chemical structures of the most common phytoestrogen found in *Medicago spp*. compared with estrogen found in animal.

2.7 Detection methods

Many different techniques can be used to identify and/or quantify flavonoids and a current overview of the commonly used and new emerging separation and identification techniqueshas recently been reviewed (Valls et al., 2009; Qiao et al., 2011). These techniques range from simple methods to the use of sophisticated instrumentation such as mass spectrometry (MS), nuclear magnetic resonance (NMR) spectroscopy and laser-induced fluorescence (LiF) detection, each having unique advantages and disadvantages (Lei et al., 2011). For example, NMR is considered as the Holy Grail in structural elucidation and is highly selective and non-destructive, but suffers from a relatively lower sensitivity. In contrast, LiF is one of the most sensitive techniques, but cannot be used for structural identification because it lacks chemical selectivity. MS offers a good combination of both sensitivity and selectivity, making it one of the prime methods of choice (Lei et al., 2011). Coupling chromatography to MS offers an extra level of information to flavonoid elucidation and is ideal for the analysis of complex plant tissue samples. Advances in sample extraction, chromatographic separation, detection and structural analysis of flavonoids have dramatically influenced the evolution of flavonoid discovery and improved flavonoid research in general (Qiao et al., 2011). High-performance liquid chromatography (HPLC) coupled to UV detection (HPLC-UV) and to MS (HPLC-UV-MS) systems are well recognized methods to profile flavonoid conjugates in plant tissue extracts. As such, various classes of flavones, flavonol and isoflavone glycoconjugates and aglycones have been identified in *Medicago spp*. However, these methods are insufficient to distinguish isomeric and isobaric compounds, which require tandem mass spectrometry with collision induced dissociation (CID MS/MS) for functional characterization. In one study, they used an integrated approach utilizing HPLC-UV coupled to electrospray ionization (ESI)-MS and gas chromatography (GC)-MS to elucidate the flavonoid profiles from *M. truncatula* root and cell cultures. A quadrupole time-of-flight (QToF) MS device was used for all structural identifications, but where the stereochemistry of sugar conjugates was uncertain, they used enzymatic hydrolysis followed by GC-MS to assign a correct sugar stereochemical configuration (Farag et al., 2007). The same lab also used reverse phase high-performance liquid chromatography coupled to UV photodiode array detection and electrospray ionization ion-trap mass spectrometry (HPLC-PDA-ESI-ITMS) to analyze the intra- and extracellular phenylpropanoid and isoflavonoid metabolome of *M. truncatula* cell cultures in response to YE or MeJA (Farag et al., 2008). More recently, this technique was

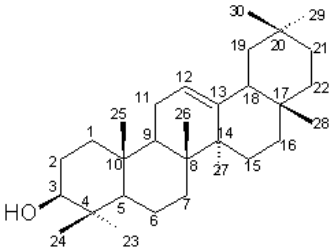
validated in the same *M. truncatula* plant system to identify the biosynthetic mechanism for the aurone hispidol (Farag et al., 2009). In a study to discover the changes of flavonoid accumulation in *M. truncatula* leaves infected with *P. medicaginis,* in which they also used a HPLC-UV-MS system (Jasinski et al., 2009). In a comparable study, both a low resolution ion trap (IT) and a high resolution tandem QToF LC-UV-MS system was used to evaluate the fragmentation pathways of *M. truncatula* flavonoids after CID experiments. They concluded that although some decent fragmentations of minor compounds could be obtained using the low resolution IT, it does not always allow proper identification of all molecules in the sample. For that reason and because of the presence of equal nominal masses of some different substituents, a high resolution instrument could only be used to correctly analyze these derivatives (Marczak et al., 2010). In addition, they used the same high resolution methodology to compare the flavonoid profiles from *M. truncatula* seedling roots, hairy roots and suspension root cell cultures (Staszkow et al., 2011). In another study, they metabolically profiled mycorrhizal roots of *M. truncatula,* in which the polar metabolites were analyzed by GC-TOF-MS, the nonpolar by LC-MS and the cell-wall bound components by reverse phase HPLC coupled to a photodiode array detector (DAD). As mass analyzer they used the very sensitive fourier transformation ion cyclotron resonance (FT-ICR) MS device (Schliemann et al., 2008). In a study a new stop-and-go two-dimensional chromatography was used for the preparative separation of flavonoids from *M. sativa*, combining counter-current chromatography and liquid chromatography (2D CCC x LC). Moreover, they identified two new flavonoids coupling this technique to ESI-MS, ESI-TOF-MS and 1D and 2D NMR (Liang et al., 2011b).

3 Saponins

Saponins are a structurally diverse class of amphipathic glycosides with a lipophilic steroid, steroidal alkaloid, or triterpenoid aglycone backbone or sapogenin that is covalently linked to one (monodesmosidic) or more (di- or tridesmosidic) hydrophilic sugar chains via a glycosidic bond. The name saponin is derived from the Latin word for soap, *sapo*, and points to an important physicochemical property of the compounds, their ability to form a colloidal solution in water that forms stable foam when shaken. This makes them useful as emulsifiers and foaming agents in the food and beverage industries. Saponins also possess various pharmacological properties, making them commonly used in phytotherapy and cosmetics (Osbourn et al., 2011; Pollier et al., 2011a).

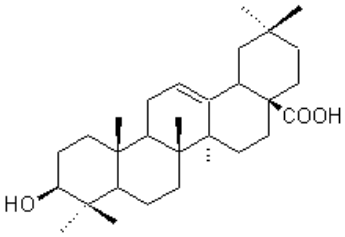
The saponins that are present in the various *Medicago* species all have pentacyclic oleanane-type sapogenins that are characterized by the presence of a C-12–C-13 double bond, and an oxygen atom at the 3β-position (Fig.10A) (Tava et al., 2011). The sapogenins are derived from β-amyrin through various oxidative modifications, and based on the oxidation pattern; two distinct types of aglycones can be distinguished in *Medicago*. A first class consists of sapogenol aglycones that are oxidized at the C-28 position, which is frequently accompanied with oxidation at the C-23 position. Aglycones of this class in *Medicago* are oleanolic acid, hederagenin, bayogenin, medicagenic acid, and zanhic acid (Fig. 10B). A second class of sapogenins, called soyasapogenols, possess a hydroxy group at the C-24 position, which excludes oxidation at the C-28 position (Tava et al., 2011). In *Medicago*, sapogenins of this class include the soyasapogenols A, B and E (Fig. 10C). The type of aglycone is correlated with the biological properties of the saponins. For instance, the haemolytic activity of the saponins in *M. truncatula* is due to the presence of sapogenins that are oxidized at the C-28 position while the non-haemolytic soyasapogenols are hydroxylated at their C-24 position (Carelli et al., 2011).

A

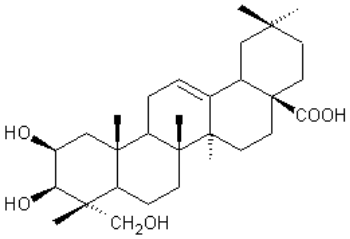


β-amyrin

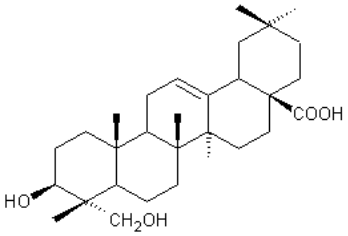
B



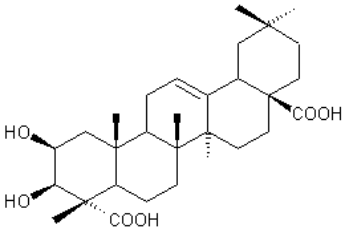
Oleanolic acid



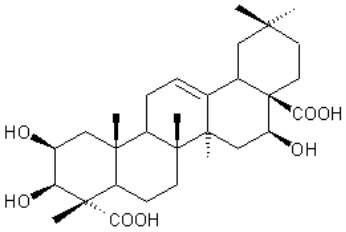
Bayogenin



Hederagenin

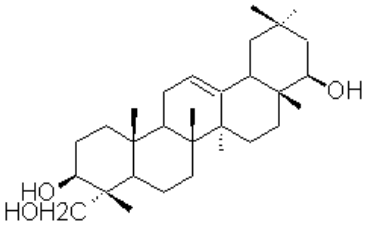


Medicagenic acid

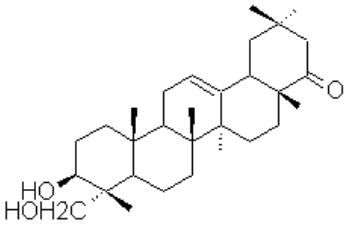


Zanhic acid

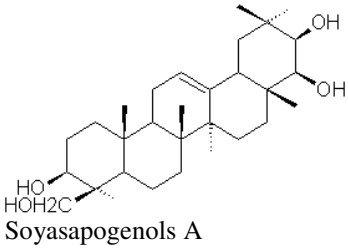
C



Soyasapogenols B



Soyasapogenols E



Soyasapogenols A

Figure 10. The triterpene aglycones saponins found in *Medicago* species; A) β-amyrin: basic structure of triterpene saponin found in *Medicago sp.* B) sapogenols C) soyasapogenols.

3.1 Saponin biosynthesis in *M. truncatula*

As secondary metabolites, the triterpene saponins present in *Medicago* share a common biogenic origin with the primary sterol metabolism. 2,3-oxidosqualene, the last common precursor molecule between the sterol metabolism and the secondary triterpenoid metabolism, is synthesized in the cytosol from isopentenyl pyrophosphate (IPP) derived

from the mevalonate (MVA) pathway (Chappell, 2002; Xu et al., 2004; Phillips et al., 2006). In *M. truncatula*, five isoforms of the key MVA pathway enzyme 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR) have been characterized (Kevei et al., 2007), for the remaining steps of the MVA pathway, candidate genes were identified *via* a functional genomics approach (Naoumkina et al., 2010b).

The end-product of the mevalonate pathway, IPP, is isomerised by IPP isomerase (IPPI) to yield the allylic isomer dimethylallyl pyrophosphate (DMAPP) (Croteau et al., 2000). Subsequently, the prenyl transferase farnesyl pyrophosphate synthase (FPS) catalyzes the sequential condensation reactions of DMAPP with two units of IPP to form farnesyl pyrophosphate (FPP) (McGarvey and Croteau, 1995). Neither IPPI nor FPS have been functionally characterized in *M. truncatula*, but candidate genes can be found in the genome (Naoumkina et al., 2010b). In the next step of the biosynthetic pathway leading to 2,3-oxidosqualene, two molecules of FPP are coupled head-to-head to form squalene. This reaction is catalyzed by squalene synthase (SQS), of which the single copy that can be found in the *M. truncatula* genome has been characterized (Suzuki et al., 2002). Finally, squalene is oxidized by squalene epoxidase (SQE) to 2,3-oxidosqualene. In *M. truncatula*, at least three *SQE* genes are present (Naoumkina et al., 2010b), 2 of which have been characterized (Suzuki et al., 2002).

The cyclization of 2,3-oxidosqualene forms the branch-point between the primary sterol and the secondary triterpene saponin metabolism. Cycloartenol (Fig. 11), the tetracyclic plant sterol precursor, is synthesized through cyclization of 2,3-oxidosqualene by cycloartenol synthase (CAS) (Corey et al., 1993), the ancestral enzyme of all plant oxidosqualene cyclases (OSCs) involved in the secondary metabolism (Phillips et al., 2006). For the biosynthesis of triterpene saponins in *Medicago*, 2,3-oxidosqualene is cyclized to the pentacyclic oleanane-type triterpene backbone β-amyrin by the OSC β-amyrin synthase (BAS) (Fig. 11) (Suzuki et al., 2002; Iturbe-Ormaetxe et al., 2003). Next to the characterized *M. truncatula BAS* gene, several other *OSCs* can be identified in the *M. truncatula* genome. Among these are genes that cluster with OSCs that catalyze the formation of other types of triterpene backbones, such as lupeol, of which derivatives have not been detected yet in *M. truncatula* (Naoumkina et al., 2010b).

After the synthesis of β-amyrin, the competitive action of two enzymes causes another branching of the saponin biosynthetic pathway. The CYP450 enzyme CYP716A12 catalyzes the carboxylation of β-amyrin at the C-28 position, which seems to exclude hydroxylation of β-amyrin at the C-24 position catalyzed by the CYP450 enzyme

CYP93E2, or *vice versa* (Fig. 11) (Carelli et al., 2011; Fukushima et al., 2011; Tava et al., 2011). Carboxylation of β-amyrin at the C-28 position leads to oleanolic acid, which is further modified by CYP72A61v2 to yield haemolytic saponins. In another branch CYP72A68v2 catalyzes the C-22 hydroxylation of 24-hydroxy- β-amyrin leading to the production of the non-haemolytic soyasapogenol glycosides (Tava et al., 2011; Fukushima et al., 2013).

Next to oxidative modifications at the C-24 and C-28 positions, several other positions of the β-amyrin backbone are oxidatively modified in *Medicago*. The specific enzymes catalyzing these reactions have not been reported yet, but most likely they are also CYP450s.

After oxidation of the β-amyrin backbone, the resulting sapogenins are glycosylated at different positions. In *Medicago*, glycosylation occurs mainly at the C-3 hydroxy and the C-28 carboxy groups of the aglycone. Rarely, additional glycosylation is observed at the C-23 position (Tava and Avato, 2006). In *M. truncatula*, several glucosyltransferases that catalyze the transfer of glucosyl residues have been identified. In a first study, UGT71G1 and UGT73K1 were found to transfer glucosyl residues to different sapogenins *in vitro*, however, the specific position to which the glucosyl residues were transferred to were not determined (Achnine et al., 2005). Also, UGT71G1 recognizes isoflavones and the flavonol quercetin as substrates, and glucosylates these compounds with higher efficiencies than triterpenes (Achnine et al., 2005), suggesting UGT71G1 might not correspond to a saponin-specific biosynthetic enzyme. Later, another glucosyltransferase, UGT73F3, was found to catalyze the glucosylation of sapogenins at the C-28 carboxy group *in vitro*, an effect confirmed *in vivo* by genetic loss-of-function studies (Naoumkina et al., 2010b). Next to glucose, several other sugars, including glucuronic acid, galactose, rhamnose, arabinose and xylose are found to be part of the sugar chains of *Medicago* saponins (Tava and Avato, 2006). To date, however, no enzymes catalyzing the transfer of these sugars residues to the aglycones have been discovered in *Medicago*.

Other types of modifications occurring on the *Medicago* saponins include binding of malonyl and methyl functionalities on the sugar residues, and in *M. sativa*, the occurrence of soyasaponin VI, containing a C-22 maltol functionality has been reported (Massiot et al., 1992; Tava and Avato, 2006). No specific enzymes for these modifications have been reported yet.

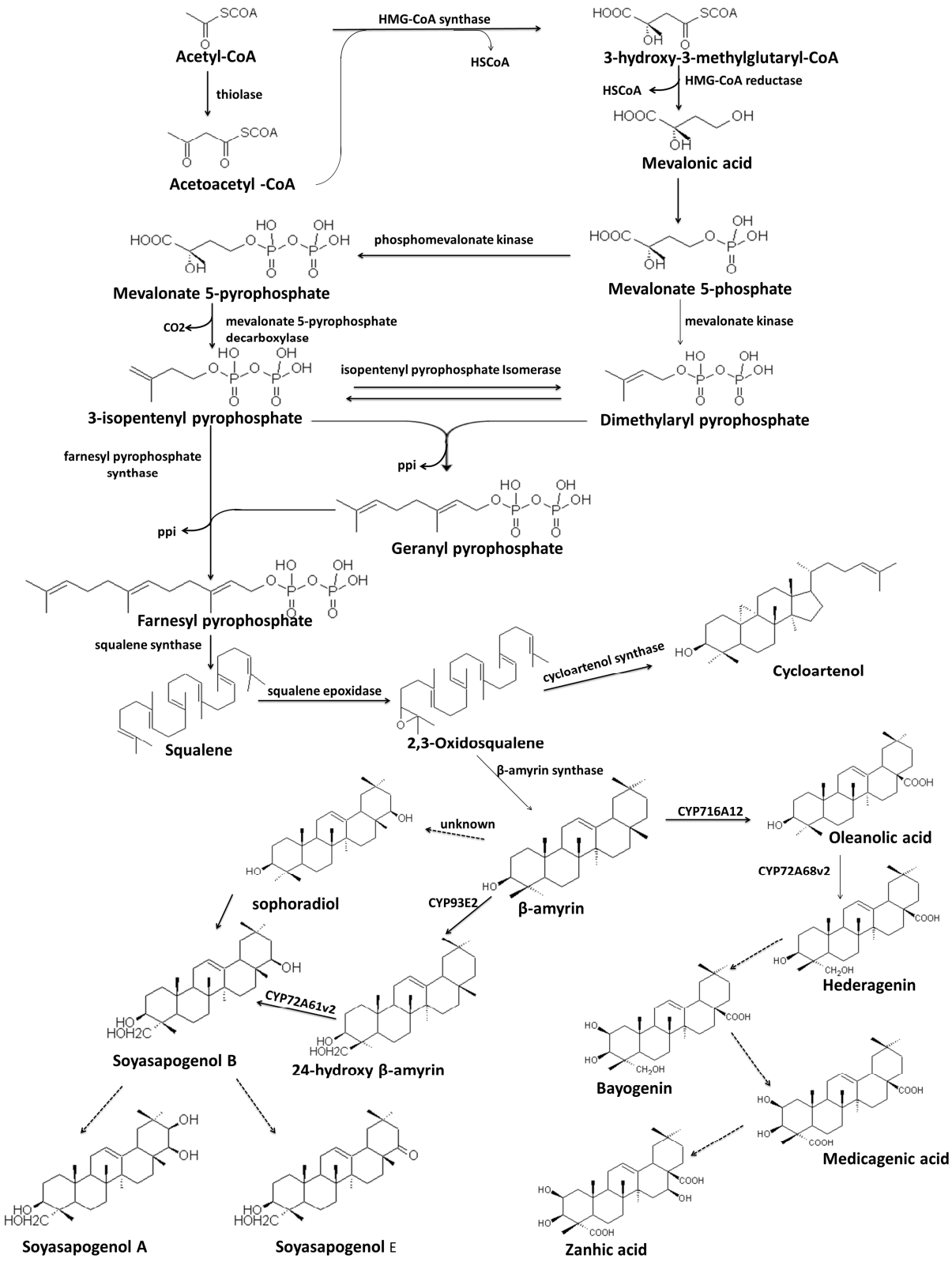


Figure 11. Proposed sapogenin biosynthetic pathway in *M. truncatula*.

3.2 Determination of *Medicago* saponins

The saponins present in *Medicago* occur in complex mixtures of structurally related high-molecular weight compounds, and individual saponins are often present in low concentrations *in planta* (Huhman and Sumner, 2002; Oleszek and Bialy, 2006; Pollier et al., 2011b). This significantly complicates the detection and quantification of saponins, and hinders purification of individual compounds, which is necessary for structure elucidation and biological activity testing. In addition, saponins lack chromophores that allow for detection in UV, limiting the choice of detection methods that can be employed for analytical purposes (Oleszek, 2002; Oleszek and Bialy, 2006).

The screening of crude extracts using hyphenated techniques such as LC-MS provides information on the composition of the saponin mixture present in the investigated plant material. In a pioneering study, a large number of saponins in root extracts of *M. sativa* and *M. truncatula* were separated with HPLC, and 15 and 27 saponins of *M. sativa* and *M. truncatula*, respectively, were tentatively identified based on fragmentation data under negative ionization (Huhman and Sumner, 2002). More recently, by coupling HPLC to the highly accurate Fourier transform ion cyclotron resonance mass spectrometry; reliable prediction of the molecular formula of the detected saponins was possible. This has led, in combination with the fragmentation spectra under negative ionization, to the tentative identification of 79 saponins in *M. truncatula* hairy roots (Pollier et al., 2011b), and further underscores the complexity of the saponin mixture present in *Medicago*.

This qualitative profiling of the saponin mixture, however, does not provide conclusive evidence for the absolute chemical structures of the saponins present in *Medicago*, which can only be obtained by subjecting purified compounds to a combination of analytical methods, including MS and NMR (Tava and Avato, 2006). For instance, the purification of saponins from the aerial parts of *M. truncatula* has led to the chemical characterization of 15 individual compounds (Kapusta et al., 2005b). A detailed overview of the different saponin structures reported in various *Medicago* species can be found elsewhere (Tava and Avato, 2006). So far, however, the purification of compounds has only led to the identification of the dominant compounds in the saponin mixture; the purification of low abundant compounds remains a challenge.

The purified and characterized compounds can be used as standards to generate response curves that can be used for the absolute quantification of saponins with HPLC/MS. In the first study in *Medicago*, differential accumulation of saponins in the various organs of *M. truncatula* was reported, with higher levels of medicagenic acid conjugates present in

leaves and seeds, and higher levels of soyasapogenol conjugates present in the roots. This suggests a tissue-specific biosynthesis and biological role for the saponins *in planta* (Huhman et al., 2005). In a similar approach, saponins in the aerial parts of three different *M. truncatula* cultivars were quantified, showing very similar saponin mixtures in the three cultivars, with medicagenic acid, zanhic acid and soyasapogenol glycosides being the dominant compounds. The observed total saponin concentration in *M. truncatula* is very similar to the observed total saponin concentration in *M. sativa*, however, *M. truncatula* has a higher concentration of zanhic acid glycosides, and a lower concentration of soyasapogenol glycosides as compared to *M. sativa* (Kapusta et al., 2005a). The high levels of zanhic acid glycosides (> 40% of the total saponin content) observed in the latter study are in disagreement with the first study, which reports zanhic acid glycosides to be only 0.6% of the total saponin content. This difference is attributed to the lack of appropriate standards in the first study, in which no zanhic acid glycosides were available as standards for absolute quantification (Huhman et al., 2005; Kapusta et al., 2005a).

Next to HPLC/MS, GC/MS is used for the quantification of *Medicago* saponins. However, being high-molecular weight compounds, saponins are not volatile, and hence, the GC/MS analysis is performed on hydrolyzed saponins, and thus only provides information on the present aglycones. Furthermore, hydrolysis is often not complete or can lead to artefacts which influence the final result (Tava et al., 1993; Oleszek, 2002; Carelli et al., 2011).

3.3 Biological roles of *Medicago* saponins

Saponins are widely distributed through plant kingdom and display several biological properties. Studies on the structure-activity relationships of saponins show that nature of the aglycone, the functional groups on the aglycone backbone and identity and number of the sugars reflect on their physicochemical and biological properties (Liu and Henkel, 2002; Güçlü-Üstündağ and Mazza, 2007). Due to their amphipathic nature, saponins are surface active compounds possessing detergent, wetting, emulsifying, and foaming properties (Güçlü-Üstündağ and Mazza, 2007). In aqueous solutions saponins form micelles above a critical concentration called critical micelle concentration (cmc). Therefore, they can be used in cosmetic, pharmaceutical or food formulations to increase solubility, bioavailability, bioactivity and extraction yields (Güçlü-Üstündağ and Mazza, 2007). Various pharmacological and biological effects reported for saponins have been

attributed to their ability to interact with sterols, particularly those that are related to membrane properties like haemolytic and hypocholesterolemic activities (Kortner et al., 2012). Haemolytic saponins are capable to rupture erythrocytes, causing an increase in membrane permeability and a loss of hemoglobin; consequently saponins can also be toxic to monogastric animals, act as anti-palatability factors, or negatively impact forage digestibility in ruminants (Sen et al., 1998; Tava and Avato, 2006).

3.3.1 Saponins roles in plants

Saponins are likely involved in plant defense mechanisms against potential plant pathogens and can be classified in a large group of protective molecules, namely phytoprotectants. They can be either produced upon a pathogen attack or stress as phytoalexins (Augustin et al., 2011) or produced during normal growth and development as phytoanticipins (Papadopoulou et al., 1999; Lambert et al., 2011). They are known to possess a wide range of biological activities in plants including antimicrobial, antiviral, molluscicidal, nematocidal, insecticidal, antiparasitic and allelopathic activities (Sparg et al., 2004; Tava and Avato, 2006; Augustin et al., 2011; D'Addabbo et al., 2011). Antimicrobial, nematocidal, insecticidal and allelopathic activities were described for saponins obtained from different species of *Medicago*.

3.3.1.1 Antimicrobial activity

Several bioassay reports have indicated the role of saponins isolated from *Medicago* species in plant defense mechanisms against phytopathogenic fungi and bacteria (Saniewska et al., 2006; Tava and Avato, 2006). For instance, antifungal efficacy of saponins from *Medicago* on the model fungus *Trichoderma viride* revealed that growth of *T. viride* was inhibited by *Medicago* saponins (Tava and Avato, 2006). Accordingly, incorporation of alfalfa plant material in soil inoculated with *Phytophthora capsici* reduced *Phytophthora* blight of capsicum in pots and field trials (Demirci and Dolar, 2006). Therefore, the antifungal activity of alfalfa saponins can cause to reduce the presence of phytopathogenic fungi in the amended soil. It seems that medicagenic acid is the major aglycone contributing to the antifungal activity (Gestetner et al., 1971; Levy et al., 1989).

3.3.1.2 Nematicidal activity

The study of nematicidal activity of different concentrations of saponins from *M. arborea*, *M. arabica* and *M. sativa* against the plant-parasitic nematode *Xiphinema index*

has showed that all possess nematicidal activity (Argentieri et al., 2008; D'Addabbo et al., 2009; D'Addabbo et al., 2011). The nematicidal activity of saponins allows using *Medicago* species dry material as soil amendments to control phytoparasitic nematods. For instance, applying dry material from top and root of *M. sativa* and/or *M. arborea* as soil amendments was shown to suppress root and soil population density of the root- knot nematode *Meloidogyne incognita* and the cyst nematode *Globodera rostochiensis* in potting mixes (D'Addabbo et al., 2009).

The exact mechanism of nematotoxic activity of saponin is not yet fully understood. However, based on biological effects of saponins, it is ascribed to their specific interaction with cell membranes which lead to changes in cell permeability (Tava and Avato, 2006). Furthermore, it has been demonstrated that saponins are able to interact with proteins (Potter et al., 1993). Nematode possesses a protective cuticle which is critical structure for its viability. This structure is primarily composed of collagen proteins assembled into higher order complexes (Page and Winter, 2003). Therefore, it is supposed that saponins may interact with cuticle collagen proteins which results in the observed nematotoxic effects.

3.3.1.3 Insecticidal activities

Saponins are also involved in plant protection against insect attack and their levels increase in leaves of damaged plants. Saponins from *M. sativa*, *M. arabica*, *M. hybrida* and *M. murex* showed insecticidal properties against several classes of insects and pests (reviewed in (Tava and Avato, 2006). Three alfalfa saponins (zanhic acid tridesmoside, 3-GlcA,28-AraRhaXyl medicagenic acid glycoside, and 3-GlcA,28-AraRha medicagenic acid glycoside) were found to inhibit *A. pisum* feeding (Goławska, 2007). Similarly, 3-GlcA-28-AraRhaxyl-medicagenate isolated from the seed flour of *M. truncatula* seeds was reported to be highly toxic for the rice weevil *Sitophilus oryzae* (Da Silva et al., 2012). It is speculated that saponins causes higher mortality levels, lower food intake, weight reduction, disturbances in development and decreased reproduction in pest insects. The mechanism underlying their insecticidal activity is, however, still largely unknown, but it is proposed that saponins act via multiple mechanisms. The main hypotheses are that saponins could either make the food less attractive to eat (repellent/deterrent activity) or they cause digestive problems by slowing down the passage of food through the insect gut, reducing food digestibility by inhibiting the secretion of digestive enzymes (proteases) or by formation of hardly digestible saponin- protein complexes (Potter et al.,

1993). Collectively, these can lead to an obstruction of alimentary canal that would limit or inhibit food uptake. Another possible mechanism can be due to the interaction of saponin with sterol which blocked sterol uptake. Insects are not able to synthesize sterol (Bellés et al., 2005). But, they take steroids from foods and used them for the synthesis of steroids like cholesterol and the insect molting hormone 20-hydroxyecdysone (20E) (Bellés et al., 2005). It is suggested that saponins can block sterol uptake by formation of insoluble complexes with sterols and make all cholesterol in the food unusable for the insects. Moreover, when larvae feed on a saponin-rich food, saponins may form a complex with cholesterol in their body, and thus disturbs the biosynthesis of ecdysteroids. Finally, the insecticidal effect of saponins can be due to their membrane-permeabilising ability. They can interact with membrane which permeabilize the small intestine mucosal cells of insects, leading to a marked reduction in their ability to transport nutrients (Francis et al., 2002).

3.3.2 Pharmacological activity

Beside their role in plant defense, the valuable pharmacological properties of saponin draw many attentions. Saponins exert a wide range of pharmacological activities including antiinflammatory, vasoprotective, hypocholesterolemic, immunomodulatory, hypoglycaemic, antifungal, antiparasitic and many others (Sparg et al., 2004; Sahu et al., 2008). Recent studies have reported anti-cholesterolemic, antimicrobial, antifungal, anti-cancer and cytotoxic properties of saponins in *Medicago spp*. (Houghton et al., 2006; Tava and Avato, 2006; Tava and Pecetti, 2012). Some of the most important are highlighted below.

3.3.2.1 Antimicrobial activity

Antimicrobial activity of saponins from *M. sativa*, *M. arabica* and *M. arborea* was explored against the most medically important yeasts (*Candida albicans, C. tropicalis, Saccharomyces cerevisiae, Cryptococcus laurentii and Blastomyces capitatus*) and Gram-positive and Gram-negative bacteria (Avato et al., 2006). The results showed that *S. cerevisiae* was the most susceptible and showed high growth inhibition when treated with the sapogenin mixtures from the different species of *Medicago* (Avato et al., 2006). Saponins from *Medicago* species did not show activity (MICs > 500 µg/mL) against Gram-negative bacteria, while they have a broader spectrum of efficacy against Gram-positive bacteria. The largest inhibitory effect observed for the sapogenin mixtures obtained from the acid hydrolysis of *M. arabica* (tops and roots) saponins. The observed

antimicrobial properties of *M. sativa* and *M. arborea* were related to the content of medicagenic acid, while hederagenin seems to contribute to the bioactivity of *M. arabica* total sapogenins (Avato et al., 2006). Moreover, antimicrobial activity of saponins obtained from roots and aerial parts of four *Medicago* species, *Medicago sativa*, *M. murex*, *M. arabica* and *M. hybrida* have been shown to have inhibitory effects on the growth the dermatophytic fungi such as *Microsporum gypseum*, *Trichophyton interdigitale* and *T. tonsurans (Houghton et al., 2006).*

3.3.2.2 Cytotoxic activity

Triterpene saponins are also known for their cytotoxic activities. For instance, saponins from *M. sativa* leaves caused to growth inhibition *in vitro* against human leukemic cell line K562 (Tava and Odoardi, 1996). Several compositions of sapogenin moieties isolated from 12 annual *Medicag* species were shown to have cytotoxic activity against the brine shrimp *Artemia salina* (Tava and Pecetti, 2012). The cytotoxic activity of the species was ascribed to different content of mono- and bidesmoside triterpene saponins, as monodesmoside saponins are more biologically active than the bidesmoside (Tava and Pecetti, 2012). Moreover, hederagenin and bayogenin glycosides were identified as the main saponin compound in the species with the highest cytotoxic activity such as *M. arabica* and *M. rigidula* with the LD50 value of 10.1 and 4.6 µg/ml, respectively. Whereas, the species with low cytotoxic activity (with the LD50 value comprised between 114.5 and 181.3 µg/ml) including, *M. aculeate*, *M. littoralis* and *M. doliata* accumulate mainly medicagenic acid, zanhic acid and soyasapogenols B (Tava and Pecetti, 2012).

3.3.2.3 Hypocholesterolemic activity

Triterpene saponins have a specific affinity for cholesterol (Bangham and Horne, 1962), they can form a complex with intestinal bile salts and cholesterol (Messina, 1999), thus decreasing intestinal cholesterol absorption. Moreover, membranolytic action of saponins can increase intestinal cell turnover rate and loss of cell membrane cholesterol from shed cells (Kortner et al., 2012). In *Medicago* species, hypocholestrolemic activity of saponins has been reported from alfalfa saponins, in which rats fed with alfalfa saponins at levels of 1% in the diet for up to 26 weeks had a significant decrease in the levels of their serum cholesterol and triglycerides (reviewed in (Tava and Avato, 2006). Similarly, in the non-human primate, *Macaca fascicularis*, *M. sativa* saponins derived from stem and leaves have been shown to decrease plasma cholesterol concentrations without changing the

level of HDL –cholesterol and, therefore, reduce the total cholesterol/HDL- cholesterol ratio. Saponins also decreased intestinal absorption of cholesterol, increased fecal excretion of endogenous and exogenous neutral steroids and bile acids, and decreased the percentage distribution of fecal deoxycholic and lithocholic acids (reviewed in (Tava and Avato, 2006)). Hypocholesterolemic activity of alfalfa root saponins has also been reported in monkeys receiving a high-cholesterol diet (Reviewed in (Bora and Sharma, 2011)). Furthermore, *in vitro* studies indicated that saponins from *M. sativa* roots and aerial parts are capable to increase lipolytic activity of pancreatic enzymes while they do not influence on the proteolytic and amylolytic activities (reviewed in (Tava and Avato, 2006)).

4 Concluding remarks

The presence of a wide range of natural products in the genus indicates that *Medicago* species have potential to be used as a platform for the development of novel bioactive compounds. Exploration of the elements involved in several aspects of secondary metabolite production in the plant provides the basis for developing such a platform. The current information available on the biology, genomics, biochemistry and metabolic engineering of *Medicago* species shows that considerable progress has been made in the past decade, although a lot of questions remain still unsolved.

Nowadays, with the aid of functional and structural genomics, in association with metabolomics, exploring of plant natural product biosynthesis pathway is more straightforward. In addition, advances in genomics area have a huge influence on exploring the transcriptional regulators involved in regulation of natural product biosynthetic enzymes. Despite some reports of identification of regulatory factors controlling metabolite biosynthesis genes in *Medicago* species, regulatory mechanism for the production of these natural compounds are still largely unknown. Identification of the mechanisms controlling the production of bioactive natural product in the genus can secure a stable, reliable and large scale production of these compounds whether it is through manipulating the wildtype species or serving the biosynthesis machinery for production in heterologous expression systems.

Furthermore, taking into account the beneficial effects of secondary metabolites from *Medicago* species in agriculture industry and their potential roles in animal and human health, *Medicago* species seems to hold great potential for in depth investigation for

various biological activities. Therefore, it is necessary to exploit its maximum potential in the field of medicinal and pharmaceutical sciences for novel and fruitful application.

References

Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., and Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant Journal **41,** 875-887.

Agati, G., and Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection. New Phytologist **186,** 786-793.

Alessandra, B., Ciccarelli, D., Garbari, F., and Pistelli, L. (2010). Flavonoids isolated from *Medicago littoralis* Rhode (Fabaceae): their ecological and chemosystematic significance. Caryologia **63,** 106-114.

Almeida, J.R.M., D'Amico, E., Preuss, A., Carbone, F., de Vos, C.H.R., Deiml, B., Mourgues, F., Perrotta, G., Fischer, T.C., Bovy, A.G., Martens, S., and Rosati, C. (2007). Characterization of major enzymes and genes involved in flavonoid and proanthocyanidin biosynthesis during fruit development in strawberry (*Fragaria x ananassa*). Archives of Biochemistry and Biophysics **465,** 61-71.

Aloui, A., Dumas-Gaudot, E., Daher, Z., van Tuinen, D., Aschi-Smit, S., and Morandi, D. (2012). Influence of arbuscular mycorrhizal colonisation on cadmium induced *Medicago truncatula* root isoflavonoid accumulation. Plant Physiology and Biochemistry **60,** 233-239.

Argentieri, M., D'Addabbo, T., Tava, A., Agostinelli, A., Jurzysta, M., and Avato, P. (2008). Evaluation of nematicidal properties of saponins from *Medicago spp.* European Journal of Plant Pathology **120,** 189-197.

Augustin, J.M., Kuzina, V., Andersen, S.B., and Bak, S. (2011). Molecular activities, biosynthesis and evolution of triterpenoid saponins. Phytochemistry **72,** 435-457.

Avato, P., Bucci, R., Tava, A., Vitali, C., Rosato, A., Bialy, Z., and Jurzysta, M. (2006). Antimicrobial activity of saponins from *Medicago sp.*: structure-activity relationship. Phytotherapy Research **20,** 454-457.

Baldridge, G.D., O'Neill, N.R., and Samac, D.A. (1998). Alfalfa (*Medicago sativa* L.) resistance to the root-lesion nematode, Pratylenchus penetrans: defense-response gene mRNA and isoflavonoid phytoalexin levels in roots. Plant Mol Biol **38,** 999-1010.

Ballance, G.M., and Dixon, R.A. (1995). MEDICAGO-SATIVA CDNAS ENCODING CHALCONE REDUCTASE. Plant Physiology **107,** 1027-1028.

Bangham, A.D., and Horne, R.W. (1962). Action of saponin on biological cell membranes. Nature **196,** 952-953.

Bellés, X., Martín, D., and Piulachs, M.-D. (2005). The mevalonate pathway and the synthesis of juvenile hormone in insects. Annual Review of Entomology **50,** 181-199.

Bhattacharya, A., Sood, P., and Citovsky, V. (2010). The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. Molecular Plant Pathology **11,** 705-719.

Bickoff, E.M., Livingston, A.L., and Booth, A.N. (1964). Tricin from alfalfa—isolation and physiological activity. Journal of Pharmaceutical Sciences **53,** 1411-1412.

Bido, G.D., Ferrarese, M.D.L., Marchiosi, R., and Ferrarese, O. (2010). Naringenin inhibits the growth and stimulates the lignification of soybean root. Brazilian Archives of Biology and Technology **53,** 533-542.

Bodewes, T.C.F., Luttikhold, J., van Stijn, M.F.M., Visser, M., van Norren, K., Vermeulen, M.A.R., and van Leeuwen, P.A.M. (2011). Antioxidative properties of flavonoids. Current Organic Chemistry **15,** 2616-2626.

Bonanomi, G., Antignani, V., Barile, E., Lanzotti, V., and Scala, F. (2011). Decomposition of *Medicago sativa* residues affects phytotoxicity, fungal growth and soil-borne pathogen diseases. Journal of Plant Pathology **93,** 57-69.

Bora, K.S., and Sharma, A. (2011). Phytochemical and pharmacological potential of *Medicago sativa*: A review. Pharmaceutical Biology **49,** 211-220.

Borevitz, J.O., Xia, Y.J., Blount, J., Dixon, R.A., and Lamb, C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. The Plant Cell **12,** 2383-2393.

Bovy, A., de Vos, R., Kemper, M., Schijlen, E., Pertejo, M.A., Muir, S., Collins, G., Robinson, S., Verhoeyen, M., Hughes, S., Santos-Buelga, C., and van Tunen, A. (2002). High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes *LC* and *C1*. Plant Cell **14,** 2509-2526.

Broeckling, C.D., Huhman, D.V., Farag, M.A., Smith, J.T., May, G.D., Mendes, P., Dixon, R.A., and Sumner, L.W. (2005). Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. Journal of Experimental Botany **56,** 323-336.

Cannon, S.B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., Wang, X.H., Mudge, J., Vasdewani, J., Scheix, T., Spannagl, M., Monaghan, E., Nicholson, C., Humphray, S.J., Schoof, H., Mayer, K.F.X., Rogers, J., Quetier, F., Oldroyd, G.E., Debelle, F., Cook, D.R., Retzel, E.F., Roe, B.A., Town, C.D., Tabata, S., Van de Peer, Y., and Young, N.D. (2006). Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. Proceedings of the National Academy of Sciences of the United States of America **103,** 14959-14964.

Carelli, M., Biazzi, E., Panara, F., Tava, A., Scaramelli, L., Porceddu, A., Graham, N., Odoardi, M., Piano, E., Arcioni, S., May, S., Scotti, C., and Calderini, O. (2011). *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins. Plant Cell **23,** 3070-3081.

Catford, J.G., Staehelin, C., Larose, G., Piche, Y., and Vierheilig, H. (2006). Systemically suppressed isoflavonoids and their stimulating effects on nodulation and mycorrhization in alfalfa split-root systems. Plant and Soil **285,** 257-266.

Cazarolli, L.H., Zanatta, L., Alberton, E.H., Figueiredo, M., Folador, P., Damazio, R.G., Pizzolatti, M.G., and Silva, F. (2008). Flavonoids: prospective drug candidates. Mini-Reviews in Medicinal Chemistry **8,** 1429-1440.

Cesco, S., Neumann, G., Tomasi, N., Pinton, R., and Weisskopf, L. (2010). Release of plant-borne flavonoids into the rhizosphere and their role in plant nutrition. Plant and Soil **329,** 1-25.

Cesco, S., Mimmo, T., Tonon, G., Tomasi, N., Pinton, R., Terzano, R., Neumann, G., Weisskopf, L., Renella, G., Landi, L., and Nannipieri, P. (2012). Plant-borne flavonoids released into the rhizosphere: impact on soil bio-activities related to plant nutrition. A review. Biology and Fertility of Soils **48,** 123-149.

Chang, C.-L., Wang, G.-J., Zhang, L.-J., Tsai, W.-J., Chen, R.-Y., Wu, Y.-C., and Kuo, Y.-H. (2010). Cardiovascular protective flavonolignans and flavonoids from *Calamus quiquesetinervius*. Phytochemistry **71,** 271-279.

Chappell, J. (2002). The genetics and molecular genetics of terpene and sterol origami. Curr. Opin. Plant Biol. **5,** 151-157.

Chebil, L., Humeau, C., Falcimaigne, A., Engasser, J.-M., and Ghoul, M. (2006). Enzymatic acylation of flavonoids. Process Biochemistry **41,** 2237-2251.

Choi, H.-K., Mun, J.-H., Kim, D.-J., Zhu, H., Baek, J.-M., Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G.B., Young, N.D., and Cook, D.R. (2004). Estimating genome conservation

between crop and model legume species. Proceedings of the National Academy of Sciences of the United States of America **101,** 15289-15294.

Conn, S., Zhang, W., and Franco, C. (2003). Anthocyanic vacuolar inclusions (AVIs) selectively bind acylated anthocyanins in *Vitis vinifera* L. (grapevine) suspension culture. Biotechnology Letters **25,** 835-839.

Cook, N.C., and Samman, S. (1996). Flavonoids—Chemistry, metabolism, cardioprotective effects, and dietary sources. The Journal of Nutritional Biochemistry **7,** 66-76.

Cooper, J.E. (2007). Early interactions between legumes and rhizobia: disclosing complexity in a molecular dialogue. Journal of Applied Microbiology **103,** 1355-1365.

Corey, E.J., Matsuda, S.P.T., and Bartel, B. (1993). Isolation of an *Arabidopsis thaliana* gene encoding cycloartenol synthase by functional expression in a yeast mutant lacking lanosterol synthase by the use of a chromatographic screen. Proc. Natl. Acad. Sci. U.S.A. **90,** 11628-11632.

Coronado, C., Zuanazzi, J.A.S., Sallaud, C., Quirion, J.C., Esnault, R., Husson, H.P., Kondorosi, A., and Ratet, P. (1995). ALFALFA ROOT FLAVONOID PRODUCTION IS NITROGEN REGULATED. Plant Physiology **108,** 533-542.

Croteau, R., Kutchan, T.M., and Lewis, N.G. (2000). Natural Products (Secondary Metabolites). In Biochemistry & Molecular Biology of Plants, B. Buchanan, W. Gruissem, and R. Jones, eds (American Society of Plant Biologists), pp. 1250-1318.

Cushnie, T.P.T., and Lamb, A.J. (2005). Antimicrobial activity of flavonoids. International Journal of Antimicrobial Agents **26,** 343-356.

Cushnie, T.P.T., and Lamb, A.J. (2011). Recent advances in understanding the antibacterial properties of flavonoids. International Journal of Antimicrobial Agents **38,** 99-107.

D'Addabbo, T., Carbonara, T., Leonetti, P., Radicci, V., Tava, A., and Avato, P. (2011). Control of plant parasitic nematodes with active saponins and biomass from *Medicago sativa*. Phytochemistry Reviews **10,** 503-519.

D'Addabbo, T., Avato, P., and Tava, A. (2009). Nematicidal potential of materials from Medicago spp. Eur J Plant Pathol **125,** 39-49.

Da Silva, P., Eyraud, V., Carre-Pierrat, M., Sivignon, C., Rahioui, I., Royer, C., and Gressent, F. (2012). High toxicity and specificity of the saponin 3-GlcA-28-AraRhaxyl-medicagenate, from *Medicago truncatula* seeds, for *Sitophilus oryzae*. BMC Chemical Biology **12,** 1-9.

Davies, K.M., and Schwinn, K.E. (2005). Molecular biology and biotechnology of flavonoid biosynthesis. In Chemistry, Biochemistry and Applications, M. Andersen and K.R. Markham, eds (Taylor & Francis), pp. 143–218.

de Majnik, J., Weinman, J.J., Djordjevic, M.A., Rolfe, B.G., Tanner, G.J., Joseph, R.G., and Larkin, P.J. (2000). Anthocyanin regulatory gene expression in transgenic white clover can result in an altered pattern of pigmentation. Functional Plant Biology **27,** 659-667.

Deavours, B.E., and Dixon, R.A. (2005). Metabolic engineering of isoflavonoid biosynthesis in Alfalfa. Plant Physiology **138,** 2245-2259.

Demirci, F., and Dolar, F.S. (2006). Effects of some plant materials on *Phytophthora* blight (*Phytophthora capsici* Leon.) of pepper. Turk. J. Agric. For. **30,** 247-252.

Devic, M., Guilleminot, J., Debeaujon, I., Bechtold, N., Bensaude, E., Koornneef, M., Pelletier, G., and Delseny, M. (1999). The *BANYULS* gene encodes a DFR-like protein and is a marker of early seed coat development. Plant Journal **19,** 387-398.

Dhaubhadel, S. (2010). Regulation of isoflavonoid biosynthesis in soybean seeds. In Soybean - Biochemistry, Chemistry and Physiology, N. Tzi-Bun, ed (Croatia: InTech), pp. 243-258.

Di Meo, F., Garcia, J.C.S., Dangles, O., and Trouillas, P. (2012). Highlights on anthocyanin pigmentation and copigmentation: a matter of flavonoid pi-stacking complexation to be described by DFT-D. Journal of Chemical Theory and Computation **8,** 2034-2043.

Dixon, R.A. (1999). Isoflavonoids: biochemistry, molecular biology, and biological functions. In Comprehensive natural products chemistry U. Sankawa, ed (Amsterdam: Elsevier), pp. 773-823.

Dixon, R.A. (2005). Engineering of plant natural product pathways. Current Opinion in Plant Biology **8,** 329-336.

Dixon, R.A., and Paiva, N.L. (1995). Stress-induced phenylpropanoid metabolism. Plant Cell **7,** 1085-1097.

Dixon, R.A., and Sumner, L.W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. Plant Physiology **131,** 878-885.

Dixon, R.A., and Pasinetti, G.M. (2010). Flavonoids and isoflavonoids: from plant biology to agriculture and neuroscience. Plant Physiology **154,** 453-457.

Dixon, R.A., Xie, D.Y., and Sharma, S.B. (2005). Proanthocyanidins - a final frontier in flavonoid research? New Phytologist **165,** 9-28.

Dong, Y.H., Kvarnheden, A., Yao, J.L., Sutherland, P.W., Atkinson, R.G., Morris, B.A., and Gardner, R.C. (1998). Identification of pollination-induced genes from the ovary of apple (*Malus domestica*). Sexual Plant Reproduction **11,** 277-283.

Edwards, R., Mizen, T., and Cook, R. (1995). Isoflavonoid conjugate accumulation in the roots of lucerne (*Medicago sativa*) seedlings following infection by the stem nematode (*Ditylenchus dipsaci*). Nematologica **41,** 51-66.

Farag, M.A., Huhman, D.V., Lei, Z.T., and Sumner, L.W. (2007). Metabolic profiling and systematic identification of flavonoids and isoflavonoids in roots and cell suspension cultures of *Medicago truncatula* using HPLC-UV-ESI-MS and GC-MS. Phytochemistry **68,** 342-354.

Farag, M.A., Huhman, D.V., Dixon, R.A., and Sumner, L.W. (2008). Metabolomics reveals novel pathways and differential mechanistic and elicitor-specific responses in phenylpropanoid and isoflavonoid biosynthesis in *Medicago truncatula* cell cultures. Plant Physiology **146,** 387-402.

Farag, M.A., Deavours, B.E., de Fatima, A., Naoumkina, M., Dixon, R.A., and Sumner, L.W. (2009). Integrated metabolite and transcript profiling identify a biosynthetic mechanism for hispidol in *Medicago truncatula* cell cultures. Plant Physiology **151,** 1096-1113.

Foster-Hartnett, D., Danesh, D., Penuela, S., Sharopova, N., Endre, G., Vandenbosch, K.A., Young, N.D., and Samac, D.A. (2007). Molecular and cytological responses of *Medicago truncatula* to *Erysiphe pisi*. Molecular Plant Pathology **8,** 307-319.

Francis, G., Kerem, Z., Makkar, H.P.S., and Becker, K. (2002). The biological action of saponins in animal systems: a review. British Journal of Nutrition **88,** 587-605.

Franke, A.A., Halm, B.M., Kakazu, K., Li, X.N., and Custer, L.J. (2009). Phytoestrogenic isoflavonoids in epidemiologic and clinical research. Drug Testing and Analysis **1,** 14-21.

Fukushima, E.O., Seki, H., Sawai, S., Suzuki, M., Ohyama, K., Saito, K., and Muranaka, T. (2013). Combinatorial biosynthesis of legume natural and rare triterpenoids in engineered yeast. Plant Cell Physiology **54,** 740-749.

Fukushima, E.O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., Saito, K., and Muranaka, T. (2011). CYP716A subfamily members are multifunctional oxidases in triterpenoid biosynthesis. Plant Cell Physiol. **52,** 2050-2061.

Gestetner, B., Assa, Y., Henis, Y., Birk, Y., and Bondi, A. (1971). Lucerne saponins. IV.- Relationships between their chemical constitution, and heamolytic and antifungal activities. J. Sci. Food Agr. **22,** 168-172.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W.L., Chang, W.S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol,

D., Stacey, G., Emerich, D., Vermeglio, A., Medigue, C., and Sadowsky, M. (2007). Legumes symbioses: Absence of Nod genes in photosynthetic bradyrhizobia. Science **316,** 1307-1312.

Golawska, S., Lukasik, I., Golawski, A., Kapusta, I., and Janda, B. (2010). Alfalfa (*Medicago sativa* L.) apigenin glycosides and their effect on the pea aphid (*Acyrthosiphon pisum*). Polish Journal of Environmental Studies **19,** 913-919.

Goławska, S. (2007). Deterrence and toxicity of plant saponins for the pea aphid acyrthosiphon *Pisum Harris*. Journal of Chemical Ecology **33,** 1598-1606.

Gomez, C., Conejero, G., Torregrosa, L., Cheynier, V., Terrier, N., and Ageorges, A. (2011). *In vivo* grapevine anthocyanin transport involves vesicle-mediated trafficking and the contribution of anthoMATE transporters and GST. Plant Journal **67,** 960-970.

Goodman, C.D., Casati, P., and Walbot, V. (2004). A multidrug resistance-associated protein involved in anthocyanin transport in *Zea mays*. Plant Cell **16,** 1812-1826.

Grace, S.C., and Logan, B.A. (2000). Energy dissipation and radical scavenging by the plant phenylpropanoid pathway. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences **355,** 1499-1510.

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. In Annual Review of Plant Biology, pp. 761-780.

Grotewold, E., and Davies, K. (2008). Trafficking and sequestration of anthocyanins. Natural Product Communications **3,** 1251-1258.

Güçlü-Üstündağ, Ö., and Mazza, G. (2007). Saponins: properties, applications and processing. Critical Reviews in Food Science and Nutrition **47,** 231-258.

Guo, L., Dixon, R.A., and Paiva, N.L. (1994). Conversion of vestitone to medicarpin in alfalfa (*Medicago sativa* L.) is catalyzed by two independent enzymes. Identification, purification, and characterization of vestitone reductase and 7,2'-dihydroxy-4'-methoxyisoflavanol dehydratase. Journal of Biological Chemistry **269,** 22372-22378.

Harrison, M.J., and Dixon, R.A. (1994). Spatial patterns of expression of flavonoid/isoflavonoid pathway genes during interactions between roots of Medicago-truncatula and the mycorrhizal fungus glomus versiforme. Plant Journal **6,** 9-20.

Hassan, S., and Mathesius, U. (2012). The role of flavonoids in root-rhizosphere signalling: opportunities and challenges for improving plant-microbe interactions. Journal of Experimental Botany **63,** 3429-3444.

He, X.Z., and Dixon, R.A. (2000). Genetic manipulation of isoflavone 7-*O*-methyltransferase enhances biosynthesis of 4 '-*O*-methylated isoflavonoid phytoalexins and disease resistance in alfalfa. Plant Cell **12,** 1689-1702.

He, X.Z., Wang, X.Q., and Dixon, R.A. (2006). Mutational analysis of the *Medicago* glycosyltransferase *UGT71G1* reveals residues that control regioselectivity for (Iso) flavonoid glycosylation. Journal of Biological Chemistry **281,** 34441-34447.

Hernandez, I., and Van Breusegem, F. (2010). Opinion on the possible role of flavonoids as energy escape valves: Novel tools for nature's Swiss army knife? Plant Science **179,** 297-301.

Hernandez, I., Alegre, L., Van Breusegem, F., and Munne-Bosch, S. (2009). How relevant are flavonoids as antioxidants in plants? Trends in Plant Science **14,** 125-132.

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. Journal of Experimental Botany **62,** 2465-2483.

Hong, W.J. (2005). SNARES and traffic. Biochimica Et Biophysica Acta-Molecular Cell Research **1744,** 120-144.

Hong, Y.H., Chao, W.W., Chen, M.L., and Lin, B.F. (2009a). Ethyl acetate extracts of alfalfa (*Medicago sativa* L.) sprouts inhibit lipopolysaccharide-induced inflammation *in vitro* and *in vivo*. Journal of Biomedical Science **16,** 64-76.

Hong, Y.H., Huang, C.J., Wang, S.C., and Lin, B.F. (2009b). The ethyl acetate extract of alfalfa sprout ameliorates disease severity of autoimmune-prone MRL-lpr/lpr mice. Lupus **18,** 206-215.

Hong, Y.H., Wang, S.C., Hsu, C., Lin, B.F., Kuo, Y.H., and Huang, C.J. (2011). Phytoestrogenic compounds in alfalfa sprout (*Medicago sativa*) beyond coumestrol. Journal of Agricultural and Food Chemistry **59,** 131-137.

Houghton, P., Patel, N., Jurzysta, M., Biely, Z., and Cheung, C. (2006). Antidermatophyte activity of medicago extracts and contained saponins and their structure-activity relationships. Phytotherapy Research **20,** 1061-1066.

Huhman, D.V., and Sumner, L.W. (2002). Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. Phytochemistry **59,** 347-360.

Huhman, D.V., Berhow, M.A., and Sumner, L.W. (2005). Quantification of saponins in aerial and subterranean tissues of Medicago truncatula. J. Agric. Food Chem. **53,** 1914-1920.

Humphreys, J.M., and Chapple, C. (2002). Rewriting the lignin roadmap. Current Opinion in Plant Biology **5,** 224-229.

INGHAM, J.L., and MILLAR, R.L. (1973). Sativin: an Induced Isoflavan from the Leaves of Medicago sativa L. Nature **242,** 125 - 126.

Iturbe-Ormaetxe, I., Haralampidis, K., Papadopoulou, K., and Osbourn, A.E. (2003). Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus*. Plant Mol. Biol. **51,** 731-743.

Jasinski, M., Kachlicki, P., Rodziewicz, P., Figlerowicz, M., and Stobiecki, M. (2009). Changes in the profile of flavonoid accumulation in *Medicago truncatula* leaves during infection with fungal pathogen *Phoma medicaginis*. Plant Physiology and Biochemistry **47,** 847-853.

Jones, K.M., Kobayashi, H., Davies, B.W., Taga, M.E., and Walker, G.C. (2007). How rhizobial symbionts invade plants: the Sinorhizobium-Medicago model. Nature Reviews Microbiology **5,** 619-633.

Jorgensen, K., Rasmussen, A.V., Morant, M., Nielsen, A.H., Bjarnholt, N., Zagrobelny, M., Bak, S., and Moller, B.L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Current Opinion in Plant Biology **8,** 280-291.

Junghans, H., Dalkin, K., and Dixon, R.A. (1993). Stress responses in alfalfa (*Medicago sativa* L) .15. characterization and expression patterns of members of a subset of the chalcone synthase multigene family. Plant Mol Biol **22,** 239-253.

Kapusta, I., Janda, B., Stochmal, A., and Oleszek, W. (2005a). Determination of saponins in aerial parts of barrel medic (Medicago truncatula) by liquid chromatography-electrospray ionization/mass spectrometry. J. Agric. Food Chem. **53,** 7654-7660.

Kapusta, I., Stochmal, A., Perrone, A., Piacente, S., Pizza, C., and Oleszek, W. (2005b). Triterpene saponins from barrel medic (Medicago truncatula) aerial parts. J. Agric. Food Chem. **53,** 2164-2170.

Kessmann, H., Edwards, R., Geno, P.W., and Dixon, R.A. (1990). Stress responses in alfalfa (*Medicago sativa* L.) Plant Physiol. **94,** 227-232.

Kevei, Z., Lougnon, G., Mergaert, P., Horváth, G.V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G.B., Kondorosi, A., Endre, G., Kondorosi, E., and Ané, J.M. (2007). 3-hydroxy-3-methylglutaryl coenzyme a reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. Plant Cell **19,** 3974-3989.

Kim, J.H., Cheon, Y.M., Kim, R.G., and Ahn, J.H. (2008). Analysis of flavonoids and characterization of the OsFNS gene involved in flavone biosynthesis in rice. Journal of Plant Biology **51,** 97-101.

Kortner, T.M., Skugor, S., Penn, M.H., Mydland, L.T., Djordjevic, B., Hillestad, M., Krasnov, A., and Krogdahl, A. (2012). Dietary soyasaponin supplementation to pea protein

concentrate reveals nutrigenomic interactions underlying enteropathy in Atlantic salmon (*Salmo salar*). Bmc Veterinary Research **8,** 101-118.

Kowalska, I., Stochmal, A., Kapusta, I., Janda, B., Pizza, C., Piacente, S., and Oleszek, W. (2007). Flavonoids from Barrel Medic (*Medicago truncatula*) Aerial Parts. Journal of Agricultural and Food Chemistry **55,** 2645-2652.

Kuhn, B.M., Geisler, M., Bigler, L., and Ringli, C. (2011). Flavonols accumulate asymmetrically and affect auxin transport in Arabidopsis. Plant Physiology **156,** 585-595.

Ladha, J.K., and Reddy, P.M. (2003). Nitrogen fixation in rice systems: state of knowledge and future prospects. Plant and Soil **252,** 151-167.

Lambert, E., Faizal, A., and Geelen, D. (2011). Modulation of Triterpene Saponin Production: In Vitro Cultures, Elicitation, and Metabolic Engineering. Appl Biochem Biotechnol **164,** 220-237.

Lee, S.O., Simons, A.L., Murphy, P.A., and Hendrich, S. (2005). Soyasaponins lowered plasma cholesterol and increased fecal bile acids in female golden Syrian hamsters. Experimental Biology and Medicine **230,** 472-478.

Lei, Z.T., Huhman, D.V., and Sumner, L.W. (2011). Mass spectrometry strategies in metabolomics. Journal of Biological Chemistry **286,** 25435-25442.

LeRoy, A., Potter, E., Woo, H.H., Heber, D., and Hirsch, A.M. (2002). Characterization and identification of alfalfa and red clover dietary supplements using a PCR-based method. Journal of Agricultural and Food Chemistry **50,** 5063-5069.

Levy, M., Zehavi, U., Naim, M., and Polacheck, I. (1989). Isolation, structure determination, synthesis, and antifungal activity of a new native alfalfa-root saponin. Carbohydr. Res. **193,** 115-123.

Li, L.N., Modolo, L.V., Escamilia-Trevino, L.L., Achnine, L., Dixon, R.A., and Wang, X.Q. (2007). Crystal structure of *Medicago truncatula* UGT85H2 - Insights into the structural basis of a multifunctional (Iso)flavonoid glycosyltransferase. Journal of Molecular Biology **370,** 951-963.

Li, Z.R., Wakao, S., Fischer, B.B., and Niyogi, K.K. (2009). Sensing and responding to excess light. Annual Review of Plant Biology **60,** 239-260.

Liang, J., Yang, Z., Cao, X., Wu, B., and Wu, S. (2011a). Preparative isolation of novel antioxidant flavonoids of alfalfa by stop-and-go counter-current chromatography and following on-line liquid chromatography desalination. Journal of Chromatography A **1218,** 6191-6199.

Liang, J.L., Yang, Z., Cao, X.J., Wu, B., and Wu, S.H. (2011b). Preparative isolation of novel antioxidant flavonoids of alfalfa by stop-and-go counter-current chromatography and following on-line liquid chromatography desalination. Journal of Chromatography A **1218,** 6191-6199.

Liu, C.-J., Huhman, D., Sumner, L.W., and Dixon, R.A. (2003). Regiospecific hydroxylation of isoflavones by cytochrome P450 81E enzymes from *Medicago truncatula*. The Plant Journal **36,** 471-484.

Liu, J.K., and Henkel, T. (2002). Traditional Chinese medicine (TCM): Are polyphenols and saponins the key ingredients triggering biological activities? Current Medicinal Chemistry **9,** 1483-1485.

Lukacin, R., Wellmann, F., Britsch, L., Martens, S., and Matern, U. (2003). Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. Phytochemistry **62,** 287-292.

Marczak, Ł., Stobiecki, M., Jasiński, M., Oleszek, W., and Kachlicki, P. (2010). Fragmentation pathways of acylated flavonoid diglucuronides from leaves of *Medicago truncatula*. Phytochemical Analysis **21,** 224-233.

Marek, S.M., Hansen, K., Romanish, M., and Thorn, R.G. (2009). Molecular systematics of the cotton root rot pathogen, *Phymatotrichopsis omnivora*. Persoonia - Molecular Phylogeny and Evolution of Fungi **22,** 63-74.

Markham, K.R., Gould, K.S., Winefield, C.S., Mitchell, K.A., Bloor, S.J., and Boase, M.R. (2000). Anthocyanic vacuolar inclusions - their nature and significance in flower colouration. Phytochemistry **55,** 327-336.

Martens, S., and Mithofer, A. (2005). Flavones and flavone synthases. Phytochemistry **66,** 2399-2407.

Martens, S., Forkmann, G., Matern, U., and Lukacin, R. (2001). Cloning of parsley flavone synthase I. Phytochemistry **58,** 43-46.

Martens, S., Knott, J., Seitz, C.A., Janvari, L., Yu, S.N., and Forkmann, G. (2003). Impact of biochemical pre-studies on specific metabolic engineering strategies of flavonoid biosynthesis in plant tissues. Biochemical Engineering Journal **14,** 227-235.

Masaoka, Y., Kojima, M., Sugihara, S., Yoshihara, T., Koshino, M., and Ichihara, A. (1993). Dissolution of ferric phosphate by alfalfa (*Medicago sativa* L) root exudates. Plant and Soil **155,** 75-78.

Massiot, G., Lavaud, C., Benkhaled, M., and Le Men-Olivier, L. (1992). Soyasaponin VI, a new maltol conjugate from alfalfa and soybean. J. Nat. Prod. **55,** 1339-1342.

Mazza, G.J. (2007). Anthocyanins and heart health. Annali dell'Istituto Superiore di Sanita **43.,** 369-374.

McCue, P., and Shetty, K. (2004). Health benefits of soy isoflavonoids and strategies for enhancement: A review. Critical Reviews in Food Science and Nutrition **44,** 361-367.

McGarvey, D.J., and Croteau, R. (1995). Terpenoid metabolism. Plant Cell **7,** 1015-1026.

McKhann, H.I., and Hirsch, A.M. (1994). ISOLATION OF CHALCONE SYNTHASE AND CHALCONE ISOMERASE CDNAS FROM ALFALFA (MEDICAGO-SATIVA L) - HIGHEST TRANSCRIPT LEVELS OCCUR IN YOUNG ROOTS AND ROOT-TIPS. Plant Mol Biol **24,** 767-777.

Messina, M.J. (1999). Legumes and soybeans: overview of their nutritional profiles and health effects. The American Journal of Clinical Nutrition **70,** 439s-450s.

Middleton, E., Kandaswami, C., and Theoharides, T.C. (2000). The effects of plant flavonoids on mammalian cells:implications for inflammation, heart disease, and cancer. Pharmacological Reviews **52,** 673-751.

Modolo, L.V., Blount, J.W., Achnine, L., Naoumkina, M.A., Wang, X.Q., and Dixon, R.A. (2007). A functional genomics approach to (iso)flavonoid glycosylation in the model legume *Medicago truncatula*. Plant Mol Biol **64,** 499-518.

Modolo, L.V., Li, L.N., Pan, H.Y., Blount, J.W., Dixon, R.A., and Wang, X.Q. (2009). Crystal structures of glycosyltransferase *UGT78G1* reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. Journal of Molecular Biology **392,** 1292-1302.

Morandi, D., le Signor, C., Gianinazzi-Pearson, V., and Duc, G. (2009). A *Medicago truncatula* mutant hyper-responsive to mycorrhiza and defective for nodulation. Mycorrhiza **19,** 435-441.

Nakajima, N., Ishihara, K., and Hamada, H. (2001). Functional glucosylation of kojic acid and daidzein,with the Eucalyptus membrane-associated UDP-glucosyltransferase reaction system. Journal of Bioscience and Bioengineering **92,** 469-471.

Naoumkina, M., Farag, M.A., Sumner, L.W., Tang, Y.H., Liu, C.J., and Dixon, R.A. (2007). Different mechanisms for phytoalexin induction by pathogen and wound signals in *Medicago truncatula*. Proceedings of the National Academy of Sciences of the United States of America **104,** 17909-17915.

Naoumkina, M.A., He, X.Z., and Dixon, R.A. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. BMC Plant Biol **8,** 132-146.

Naoumkina, M.A., Zhao, Q.A., Gallego-Giraldo, L., Dai, X.B., Zhao, P.X., and Dixon, R.A. (2010a). Genome-wide analysis of phenylpropanoid defence pathways. Molecular Plant Pathology **11,** 829-846.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L.W., and Dixon, R.A. (2010b). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell **22,** 850-866.

O'Neill, N.R. (1996). Defense expression in protected tissues of *Medicago sativa* is enhanced during compatible interactions with *Colletotrichum trifolii*. Phytopathology **86,** 1045-1050.

Oleszek, W. (2002). Chromatographic determination of plant saponins. J. Chromatogr. A. **967,** 147-162.

Oleszek, W., and Bialy, Z. (2006). Chromatographic determination of plant saponins - an update (2002-2005). J. Chromatogr. A. **1112,** 78-91.

Osbourn, A., Goss, R.J.M., and Field, R.A. (2011). The saponins: polar isoprenoids with important and diverse biological activities. Nat. Prod. Rep. **28,** 1261-1268.

Page, A.P., and Winter, A.D. (2003). Enzymes involved in the biogenesis of the nematode cuticle. Advances in Parasitology **53,** 85-148.

Palumbo, J.D., Kado, C.I., and Phillips, D.A. (1998). An isoflavonoid-inducible efflux pump in *Agrobacterium tumefaciens* is involved in competitive colonization of roots. Journal of Bacteriology **180,** 3107-3113.

Pan, M.H., Lai, C.S., and Ho, C.T. (2010). Anti-inflammatory activity of natural dietary flavonoids. Food & Function **1,** 15-31.

Pang, Y., Peel, G.J., Wright, E., Wang, Z., and Dixon, R.A. (2007). Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. Plant Physiology **145,** 601-615.

Pang, Y., Wenger, J.P., Saathoff, K., Peel, G.J., Wen, J., Huhman, D., Allen, S.N., Tang, Y., Cheng, X., and et al. (2009). A WD40 repeat protein from *Medicago truncatula* is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development. Plant Physiology **151,** 1114-1129.

Pang, Y.Z., Peel, G.J., Sharma, S.B., Tang, Y.H., and Dixon, R.A. (2008). A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proceedings of the National Academy of Sciences of the United States of America **105,** 14210-14215.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., and Osbourn, A.E. (1999). Compromised disease resistance in saponin-deficient plants. Proceedings of the National Academy of Sciences **96,** 12923-12928.

Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nature Reviews Microbiology **6,** 763-775.

Peel, G.J., Pang, Y., Modolo, L.V., and Dixon, R.A. (2009). The LAP1 MYB transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in *Medicago*. The Plant Journal **59,** 136-149.

Peer, W.A., Blakeslee, J.J., Yang, H.B., and Murphy, A.S. (2011). Seven things we think we know about auxin transport. Molecular Plant **4,** 487-504.

Pelletier, M.K., Murrell, J.R., and Shirley, B.W. (1997). Characterization of flavonol synthase and leucoanthocyanidin dioxygenase genes in Arabidopsis - Further evidence for differential regulation of ''early'' and ''late'' genes. Plant Physiology **113,** 1437-1445.

Peters, N.K., Frost, J.W., and Long, S.R. (1986). A plant flavone, luteolin, induces expression of rhizobium-meliloti nodulation genes. Science **233,** 977-980.

Phillips, D.A., Wery, J., Joseph, C.M., Jones, A.D., and Teuber, L.R. (1995a). RELEASE OF FLAVONOIDS AND BETAINES FROM SEEDS OF 7 MEDICAGO SPECIES. Crop Science **35,** 805-808.

Phillips, D.A., Wery, J., Joseph, C.M., Jones, A.D., and Teuber, L.R. (1995b). Release of flavonoids and betaines from seeds of seven *Medicago* species. Crop Sci. **35,** 805-808.

Phillips, D.R., Rasbery, J.M., Bartel, B., and Matsuda, S.P.T. (2006). Biosynthetic diversity in plant triterpene cyclization. Curr. Opin. Plant Biol. **9,** 305-314.

Pollastri, S., and Tattini, M. (2011). Flavonols: old compounds for old roles. Annals of Botany **108,** 1225-1233.

Pollier, J., Moses, T., and Goossens, A. (2011a). Combinatorial biosynthesis in plants: a (p)review on its potential and future exploitation. Nat. Prod. Rep. **28,** 1897-1916.

Pollier, J., Morreel, K., Geelen, D., and Goossens, A. (2011b). Metabolite profiling of triterpene saponins in *Medicago truncatula* hairy roots by liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry. J. Nat. Prod. **74,** 1462-1476.

Potter, S.M., Jimenezflores, R., Pollack, J.A., Lone, T.A., and Berberjimenez, M.D. (1993). Protein saponin interaction and its influence on blood-lipids. Journal of Agricultural and Food Chemistry **41,** 1287-1291.

Prati, S., Baravelli, V., Fabbri, D., Schwarzinger, C., Brandolini, V., Maietti, A., Tedeschi, P., Benvenuti, S., Macchia, M., Marotti, I., Bonetti, A., Catizone, P., and Dinelli, G. (2007). Composition and content of seed flavonoids in forage and grain legume crops. Journal of Separation Science **30,** 491-501.

Prior, R.L., and Wu, X. (2006). Anthocyanins: Structural characteristics that result in unique metabolic patterns and biological activities*. Free Radical Research **40,** 1014-1028.

Prochazkova, D., Bousova, I., and Wilhelmova, N. (2011). Antioxidant and prooxidant properties of flavonoids. Fitoterapia **82,** 513-523.

Qiao, X., Yang, W.Z., Guo, D.A., and Ye, M. (2011). Extraction, separation, detection, and structural analysis of flavonoids. Current Organic Chemistry **15,** 2541-2566.

Ralston, L., Subramanian, S., Matsuno, M., and Yu, O. (2005). Partial reconstruction of flavonoid and isoflavonoid biosynthesis in yeast using soybean type I and type II chalcone isomerases. Plant Physiology **137,** 1375-1388.

Rashidi, M.R., and Nazemiyeh, H. (2010). Inhibitory effects of flavonoids on molybdenum hydroxylases activity. Expert Opinion on Drug Metabolism & Toxicology **6,** 133-152.

Ray, H., Yu, M., Auser, P., Blahut-Beatty, L., McKersie, B., Bowley, S., Westcott, N., Coulman, B., Lloyd, A., and Gruber, M.Y. (2003). Expression of anthocyanins and proanthocyanidins after transformation of alfalfa with maize *Lc*. Plant Physiology **132,** 1448-1463.

Robbins, M.P., Paolocci, F., Hughes, J.W., Turchetti, V., Allison, G., Arcioni, S., Morris, P., and Damiani, F. (2003). *Sn*, a maize bHLH gene, modulates anthocyanin and condensed tannin pathways in *Lotus corniculatus*. Journal of Experimental Botany **54,** 239-248.

Sahu, N.P., Banerjee, S., Mondal, N.B., and Mandal, D. (2008). Steroidal Saponins. In Fortschritte der Chemie organischer Naturstoffe / Progress in the Chemistry of Organic Natural Products (Springer Vienna), pp. 45-141.

Saleh, N.A.M., Boulos, L., El-Negoumy, S.I., and Abdalla, M.F. (1982). A comparative study of the flavonoids of *Medicago radiata* with other *Medicago* and related *Trigonella* species. Biochemical Systematics and Ecology **10,** 33-36.

Sanders, I., Sukharnikov, L., Najar, F., and Roe, B. (2011). Medicago. In Wild crop relatives: Genomic and breeding resources, C. Kole, ed (Springer Berlin Heidelberg), pp. 207-222.

Saniewska, A., Jarecka, A.Y., Zbigniew, B.I.A., and Jurzysta, M. (2006). Antifungal activity of saponins originated from *Medicago hybrida* against some ornamental plant pathogens. ACTA AGROBOTANICA **59,** 51-58.

Sasaki, K., and Takahashi, T. (2002). A flavonoid from *Brassica rapa* flower as the UV-absorbing nectar guide. Phytochemistry **61,** 339-343.

Saunders, J.A., and O'Neill, N.R. (2004). The characterization of defense responses to fungal infection in alfalfa. Biocontrol **49,** 715-728.

Schliemann, W., Ammer, C., and Strack, D. (2008). Metabolite profiling of mycorrhizal roots of *Medicago truncatula*. Phytochemistry **69,** 112-146.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flugge, U.I., and Kunze, R. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. Plant Physiology **131,** 16-26.

Sen, S., Makkar, H.P.S., and Becker, K. (1998). Alfalfa saponins and their implication in animal nutrition. Journal of Agricultural and Food Chemistry **46,** 131-140.

Shao, H., Dixon, R.A., and Wang, X. (2007). Crystal structure of vestitone reductase from alfalfa (*Medicago sativa* L.). Journal of Molecular Biology **369,** 265-276.

Shao, H., He, X.Z., Achnine, L., Blount, J.W., Dixon, R.A., and Wang, X.Q. (2005). Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from *Medicago truncatula*. Plant Cell **17,** 3141-3154.

Shaw, L.J., Morris, P., and Hooker, J.E. (2006). Perception and modification of plant flavonoid signals by rhizosphere microorganisms. Environmental Microbiology **8,** 1867-1880.

Shih, C.H., Chu, I.K., Yip, W.K., and Lo, C. (2006). Differential expression of two flavonoid 3 '-hydroxylase cDNAs involved in biosynthesis of anthocyanin pigments and 3-deoxyanthocyanidin phytoalexins in sorghum. Plant and Cell Physiology **47,** 1412-1419.

Shimada, N., Aoki, T., Sato, S., Nakamura, Y., Tabata, S., and Ayabe, S.-i. (2003). A Cluster of Genes Encodes the Two Types of Chalcone Isomerase Involved in the Biosynthesis of General Flavonoids and Legume-Specific 5-Deoxy(iso)flavonoids in Lotus japonicus. Plant Physiology **131,** 941-951.

Shirley, B.W. (1996). Flavonoid biosynthesis: 'new' functions for an 'old' pathway. Trends in Plant Science **1,** 377-382.

Small, E. (2011). Alfalfa and relatives: Evolution and classification of *Medicago*. (Ottawa, Ontario, Canada: NRC Research Press).

Sparg, S.G., Light, M.E., and van Staden, J. (2004). Biological activities and distribution of plant saponins. Journal of Ethnopharmacology **94,** 219-243.

Spencer, J.P.E., Vafeiadou, K., Williams, R.J., and Vauzour, D. (2012). Neuroinflammation: Modulation by flavonoids and mechanisms of action. Molecular Aspects of Medicine **33,** 83-97.

Sreevidya, V.S., Rao, C.S., Sullia, S.B., Ladha, J.K., and Reddy, P.M. (2006). Metabolic engineering of rice with soybean isoflavone synthase for promoting nodulation gene expression in rhizobia. Journal of Experimental Botany **57,** 1957-1969.

Staszkow, A., Swarcewicz, B., Banasiak, J., Muth, D., Jasinski, M., and Stobiecki, M. (2011). LC/MS profiling of flavonoid glycoconjugates isolated from hairy roots, suspension root cell cultures and seedling roots of *Medicago truncatula*. Metabolomics **7,** 604-613.

Steele, K.P., Ickert-Bond, S.M., Zarre, S., and Wojciechowski, M.F. (2010). Phylogeny and character evolution in *Medicago* (Leguminosae): Evidence from analyses of plastid trnK/matK and nuclear GA3ox1 sequences. American Journal of Botany **97,** 1142-1155.

Stochmal, A., Piacente, S., Pizza, C., De Riccardis, F., Leitz, R., and Oleszek, W. (2001a). Alfalfa (Medicago sativa L.) Flavonoids. 1. Apigenin and Luteolin Glycosides from Aerial Parts. Journal of Agricultural and Food Chemistry **49,** 753-758.

Stochmal, A., Simonet, A.M., Macias, F.A., Oliveira, M.A., Abreu, J.M., Nash, R., and Oleszek, W. (2001b). Acylated apigenin glycosides from alfalfa (Medicago sativa L.) var. Artal. Phytochemistry **57,** 1223-1226.

Subramanian, S., Stacey, G., and Yu, O. (2007). Distinct, crucial roles of flavonoids during legume nodulation. Trends in Plant Science **12,** 282-285.

Suzuki, H., Achnine, L., Xu, R., Matsuda, S.P.T., and Dixon, R.A. (2002). A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. Plant J. **32,** 1033-1048.

Suzuki, H., Reddy, M., Naoumkina, M., Aziz, N., May, G., Huhman, D., Sumner, L., Blount, J., Mendes, P., and Dixon, R. (2005). Methyl jasmonate and yeast elicitor induce

differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. Planta **220,** 696-707.

Tadege, M., Wang, T.L., Wen, J.Q., Ratet, P., and Mysore, K.S. (2009). Mutagenesis and Beyond! Tools for Understanding Legume Biology. Plant Physiology **151,** 978-984.

Takahashi, R., Githiri, S.M., Hatayama, K., Dubouzet, E.G., Shimada, N., Aoki, T., Ayabe, S., Iwashina, T., Toda, K., and Matsumura, H. (2007). A single-base deletion in soybean flavonol synthase gene is associated with magenta flower color. Plant Mol Biol **63,** 125-135.

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. Plant Journal **54,** 733-749.

Tanaka, Y., Brugliera, F., Kalc, G., Senior, M., Dyson, B., Nakamura, N., Katsumoto, Y., and Chandler, S. (2010). Flower color modification by engineering of the flavonoid biosynthetic pathway: Practical perspectives. Bioscience Biotechnology and Biochemistry **74,** 1760-1769.

Tanner, G.J., Francki, K.T., Abrahams, S., Watson, J.M., Larkin, P.J., and Ashton, A.R. (2003). Proanthocyanidin biosynthesis in plants - Purification of legume leucoanthocyanidin reductase and molecular cloning of its cDNA. Journal of Biological Chemistry **278,** 31647-31656.

Tava, A., and Odoardi, M. (1996). Saponins from *Medicago Spp*.: chemical characterization and biological activity against insects. Advances in Experimental Medicine and Biology **405,** 97-109.

Tava, A., and Avato, P. (2006). Chemical and biological activity of triterpene saponins from *Medicago* species. Natural Product Communications **1,** 1159-1180.

Tava, A., and Pecetti, L. (2012). Chemical Investigation of Saponins from Twelve Annual Medicago Species and their Bioassay with the Brine Shrimp Artemia salina. Natural Product Communications **7,** 837-840.

Tava, A., Scotti, C., and Avato, P. (2011). Biosynthesis of saponins in the genus *Medicago*. Phytochem. Rev. **10,** 459-469.

Tava, A., Oleszek, W., Jurzysta, M., Berardo, N., and Odoardi, M. (1993). Alfalfa saponins and sapogenins: Isolation and quantification in two different cultivars. Phytochem. Anal. **4,** 269-274.

Tohge, T., Nishiyama, Y., Hirai, M.Y., Yano, M., Nakajima, J., Awazuhara, M., Inoue, E., Takahashi, H., Goodenowe, D.B., Kitayama, M., Noji, M., Yamazaki, M., and Saito, K. (2005). Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. Plant Journal **42,** 218-235.

Treutter, D. (2006). Significance of flavonoids in plant resistance: a review. Environmental Chemistry Letters **4,** 147-157.

Tsai, S.M., and Phillips, D.A. (1991). Flavonoids released naturally from alfalfa promote development of symbiotic Glomus spores *in vitro*. Applied and Environmental Microbiology **57,** 1485-1488.

Tsuchiya, H. (2010). Structure-dependent membrane interaction of flavonoids associated with their bioactivity. Food Chemistry **120,** 1089-1096.

Uppalapati, S.R., Marek, S.M., Lee, H.K., Nakashima, J., Tang, Y., Sledge, M.K., Dixon, R.A., and Mysore, K.S. (2009). Global gene expression profiling during *Medicago truncatula*-phymatotrichopsis omnivora interaction reveals a role for jasmonic acid, ethylene, and the flavonoid pathway in disease development. Molecular Plant-Microbe Interactions **22,** 7-17.

Urbanczyk-Wochniak, E., and Sumner, L.W. (2007). MedicCyc: a biochemical pathway database for *Medicago truncatula*. Bioinformatics **23,** 1418-1423.

Valls, J., Millan, S., Marti, M.P., Borras, E., and Arola, L. (2009). Advanced separation methods of food anthocyanins, isoflavones and flavanols. Journal of Chromatography A **1216,** 7143-7172.

Vandeputte, O.M., Kiendrebeogo, M., Rasamiravaka, T., Stevigny, C., Duez, P., Rajaonson, S., Diallo, B., Mol, A., Baucher, M., and El Jaziri, M. (2011). The flavanone naringenin reduces the production of quorum sensing-controlled virulence factors in *Pseudomonas aeruginosa* PAO1. Microbiology-Sgm **157,** 2120-2132.

Verdier, J., Zhao, J., Torres-Jerez, I., Ge, S.J., Liu, C.G., He, X.Z., Mysore, K.S., Dixon, R.A., and Udvardi, M.K. (2012). MtPAR MYB transcription factor acts as an on switch for proanthocyanidin biosynthesis in *Medicago truncatula*. Proceedings of the National Academy of Sciences of the United States of America **109,** 1766-1771.

Verrier, P.J., Bird, D., Buria, B., Dassa, E., Forestier, C., Geisler, M., Klein, M., Kolukisaoglu, U., Lee, Y., Martinoia, E., Murphy, A., Rea, P.A., Samuels, L., Schulz, B., Spalding, E.P., Yazaki, K., and Theodoulou, F.L. (2008). Plant ABC proteins - a unified nomenclature and updated inventory. Trends in Plant Science **13,** 151-159.

Verweij, W., Spelt, C., Di Sansebastiano, G.P., Vermeer, J., Reale, L., Ferranti, F., Koes, R., and Quattrocchio, F. (2008). An H^+ P-ATPase on the tonoplast determines vacuolar pH and flower colour. Nature Cell Biology **10,** 1456-1462.

Vorwerk, S., Somerville, S., and Somerville, C. (2004). The role of plant cell wall polysaccharide composition in disease resistance. Trends in Plant Science **9,** 203-209.

Wasson, A.P., Pellerone, F.I., and Mathesius, U. (2006). Silencing the flavonoid pathway in *Medicago truncatula* inhibits root nodule formation and prevents auxin transport regulation by *Rhizobia*. Plant Cell **18,** 1617-1629.

Wink, M., Botschen, F., Gosmann, C., Schäfer, H., and Waterman, P.G. (2010). Chemotaxonomy seen from a phylogenetic perspective and evolution of secondary metabolism. In Annual Plant Reviews Volume 40: Biochemistry of Plant Secondary Metabolism (Wiley-Blackwell), pp. 364-433.

Wisman, E., Hartmann, U., Sagasser, M., Baumann, E., Palme, K., Hahlbrock, K., Saedler, H., and Weisshaar, B. (1998). Knock-out mutants from an En-1 mutagenized *Arabidopsis thaliana* population generate phenylpropanoid biosynthesis phenotypes. Proceedings of the National Academy of Sciences of the United States of America **95,** 12432-12437.

Xie, D.-Y., Jackson, L.A., Cooper, J.D., Ferreira, D., and Paiva, N.L. (2004a). Molecular and Biochemical Analysis of Two cDNA Clones Encoding Dihydroflavonol-4-Reductase from Medicago truncatula. Plant Physiology **134,** 979-994.

Xie, D.-Y., Sharma, S.B., Wright, E., Wang, Z.-Y., and Dixon, R.A. (2006). Metabolic engineering of proanthocyanidins through co-expression of anthocyanidin reductase and the PAP1 MYB transcription factor. The Plant Journal **45,** 895-907.

Xie, D.Y., Sharma, S.B., and Dixon, R.A. (2004b). Anthocyanidin reductases from *Medicago truncatula* and *Arabidopsis thaliana*. Archives of Biochemistry and Biophysics **422,** 91-102.

Xie, D.Y., Sharma, S.B., Paiva, N.L., Ferreira, D., and Dixon, R.A. (2003). Role of anthocyanidin reductase, encoded by *BANYULS* in plant flavonoid biosynthesis. Science **299,** 396-399.

Xu, R., Fazio, G.C., and Matsuda, S.P.T. (2004). On the origins of triterpenoid skeletal diversity. Phytochemistry **65,** 261-291.

Yazaki, K. (2005). Transporters of secondary metabolites. Current Opinion in Plant Biology **8,** 301-307.

Yonekura-Sakakibara, K., and Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. The Plant Journal **66,** 182-193.

Yonekura-Sakakibara, K., Fukushima, A., Nakabayashi, R., Hanada, K., Matsuda, F., Sugawara, S., Inoue, E., Kuromori, T., Ito, T., Shinozaki, K., Wangwattana, B., Yamazaki, M., and Saito, K. (2012). Two glycosyltransferases involved in anthocyanin modification

delineated by transcriptome independent component analysis in *Arabidopsis thaliana*. The Plant Journal **69,** 154-167.

Young, N.D., and Udvardi, M. (2009). Translating *Medicago truncatula* genomics to crop legumes. Current Opinion in Plant Biology **12,** 193-201.

Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., and al., e. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature **480,** 520-524.

Yu, X.-H., Chen, M.-H., and Liu, C.-J. (2008). Nucleocytoplasmic-localized acyltransferases catalyze the malonylation of 7-*O*-glycosidic (iso)flavones in *Medicago truncatula*. The Plant Journal **55,** 382-396.

Zamboni, A., Zanin, L., Tomasi, N., Pezzotti, M., Pinton, R., Varanini, Z., and Cesco, S. (2012). Genome-wide microarray analysis of tomato roots showed defined responses to iron deficiency. Bmc Genomics **13,** 101-115.

Zhang, H.B., Wang, L., Deroles, S., Bennett, R., and Davies, K. (2006). New insight into the structures and formation of anthocyanic vacuolar inclusions in flower petals. BMC Plant Biol **6,** 29-43.

Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009a). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. Plant Journal **57,** 171-183.

Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009b). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. The Plant Journal **57,** 171-183.

Zhang, J.A., Subramanian, S., Zhang, Y.S., and Yu, O. (2007). Flavone synthases from *Medicago truncatula* are flavanone-2-hydroxylases and are important for nodulation. Plant Physiology **144,** 741-751.

Zhao, J., and Dixon, R.A. (2010). The 'ins' and 'outs' of flavonoid transport. Trends in Plant Science **15,** 72-80.

Zhao, J., Huhman, D., Shadle, G., He, X.Z., Sumner, L.W., Tang, Y.H., and Dixon, R.A. (2011). MATE2 mediates vacuolar sequestration of flavonoid glycosides and glycoside malonates in *Medicago truncatula*. Plant Cell **23,** 1536-1555.

Zhou, C., Han, L., Pislariu, C., Nakashima, J., Fu, C., Jiang, Q., Quan, L., Blancaflor, E.B., Tang, Y., Bouton, J.H., Udvardi, M., Xia, G., and Wang, Z.-Y. (2011). From model to crop: Functional analysis of a STAY-GREEN gene in the model legume *Medicago truncatula* and effective use of the gene for alfalfa improvement. Plant Physiology **157,** 1483-1496.

Zhu, H., Choi, H.-K., Cook, D.R., and Shoemaker, R.C. (2005). Bridging Model and Crop Legumes through Comparative Genomics. Plant Physiology **137,** 1189-1196.

Zuanazzi, J.A.S., Clergeot, P.H., Quirion, J.C., Husson, H.P., Kondorosi, A., and Ratet, P. (1998). Production of *Sinorhizobium meliloti* nod gene activator and repressor flavonoids from *Medicago sativa* roots. Molecular Plant-Microbe Interactions **11,** 784-794.

3

Scope of Research

Plants produce a wide variety of secondary metabolites with highly complex structures. Secondary metabolites fulfill a major role in plant defense against pathogens and herbivores and many of them are used as food, pharmaceutical compounds, dye, flavor, fragrances and so forth. Therefore, their biosynthetic pathways have been extensively studied and, accordingly, genetic engineering strategies have been developed to improve their production. Production of secondary metabolites is regulated by different families of regulatory proteins, some of which are elicited by extra- and intracellular signaling molecules. Overexpression of regulatory genes that coordinately control several genes of a biosynthetic pathway may offer a potentially powerful means to increase the production of secondary metabolites. Nowadays, the availability of several genetic and molecular approaches provides an opportunity to reveal fascinating insights into the complex regulatory cascades that govern secondary metabolite production.

The overall aim of this study is to better understand the molecular events underlying secondary metabolite production and to identify and characterize potential regulators involved in secondary metabolite biosynthesis in the barrel medic (*Medicago truncatula*). *M. truncatula* is a model system for legume functional genomics (Choi et al., 2004; Zhou et al., 2011). Legumes such as *M. truncatula* produces an array of bioactive natural products and the most important of those are the triterpene saponins and isoflavonoids (Dixon and Sumner, 2003). *M. truncatula* has a diverse profile of triterpene saponins (Huhman and Sumner, 2002), with the potential to use as starting materials for industrial and pharmacological applications. Although, due to the diverse biological activities and beneficial properties, the number of studies on the biosynthesis of saponins has been increasing, our molecular knowledge about their biosynthetic pathway is still linear and insufficient, and virtually inexistent when concerning its regulation. Instead, the enzymes involved in biosynthesis of the isoflavonoids, the second important group of legume natural products, are well characterized (Dixon and Paiva, 1995; Dixon, 1999). Nonetheless, the molecular events underlying the regulatory network(s) steering isoflavonoid biosynthesis pathway remain mostly unknown as well.

M. truncatula cell suspension cultures exhibit different transcriptional and metabolic responses to elicitors of plant secondary metabolism, such as yeast elicitor (YE) and

methyl jasmonate (MeJA) (Suzuki et al., 2005; Naoumkina et al., 2008). Targeted metabolic profiling showed that MeJA elicits the production of triterpene saponins in cell *M. truncatula* cultures (Achnine et al., 2005; Suzuki et al., 2005; Naoumkina et al., 2008), whereas the accumulation of isoflavonoids compounds is mostly induced by YE elicitation (Naoumkina et al., 2008). It has been proposed that these differences are orchestrated by rapid induction of different sets/combinations of transcription factors activated by YE and MeJA in *M. truncatula* cell cultures (Suzuki et al., 2005; Naoumkina et al., 2008).

To date, only a few regulators of secondary metabolism from legumes have been identified. In YE-elicited cell cultures of *M. truncatula* four WRKY genes (*W100577*, *W100630*, *W 108715*, and *W109669*) have been characterized with a regulatory roles in the biosynthesis of phenolics (Naoumkina et al., 2008). However, to the best of our knowledge, no MeJA-modulated transcriptional regulator of triterpene saponin biosynthesis in *M. truncatula* has been identified so far.

Functional genomics approaches are powerful tools to facilitate the understanding of secondary metabolism in plants. For example, cDNA-AFLP-based transcript profiling in combination with targeted metabolite analysis has been applied for the discovery of genes involved in secondary metabolism in tobacco (*Nicotiana tabacum*) cells (Goossens et al., 2003). In the PhD study of Jacob Pollier, which founded the basis for this study, a genome wide transcript profiling cDNA-AFLP analysis on MeJA elicited cell *M. truncatula* cultures showed that several regulatory proteins with a potential function in the jasmonate (JA) signaling machinery such as a MYC-like transcription factor, JAZ repressor proteins, RING E3 ligases, F box proteins as well as several other tags corresponding to genes potentially involved in regulation of gene expression are upregulated in response to MeJA, prior or concomitant with genes encoding the known triterpene saponin synthesis enzymes.

In this thesis, we extend this study by a functional screening and characterization of several of these MeJA-modulated genes encoding proteins with potential regulatory functions in secondary metabolism.

As MeJA mediated induction of the biosynthesis of several classes of secondary metabolites is conserved over multiple plant species (De Geyter et al., 2012), the

related regulators may have a conserved role across the plant kingdom as well. Therefore, identification of the regulatory elements controlling the production of bioactive natural products from a model legume such as *M. truncatula* may provide novel tools to facilitate specific and efficient engineering of the production of valuable plant secondary metabolites in general.

References

Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., and Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant Journal **41,** 875-887.

Choi, H.K., Mun, J.H., Kim, D.J., Zhu, H.Y., Baek, J.M., Mudge, J., Roe, B., Ellis, N., Doyle, J., Kiss, G.B., Young, N.D., and Cook, D.R. (2004). Estimating genome conservation between crop and model legume species. Proceedings of the National Academy of Sciences of the United States of America **101,** 15289-15294.

Dixon, R.A. (1999). Plant natural products: the molecular genetic basis of biosynthetic diversity. Current opinion in biotechnology **10,** 192-197.

Dixon, R.A., and Paiva, N.L. (1995). Stress-induced phenylpropanoid metabolism. Plant Cell **7,** 1085-1097.

Dixon, R.A., and Sumner, L.W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. Plant Physiology **131,** 878-885.

Goossens, A., Häkkinen, S.T., Laakso, I., Seppänen-Laakso, T., Biondi, S., De Sutter, V., Lammertyn, F., Nuutila, A.M., Söderlund, H., Zabeau, M., Inzé, D., and Oksman-Caldentey, K.-M. (2003). A functional genomics approach toward the understanding of secondary metabolism in plant cells. Proceedings of the National Academy of Sciences **100,** 8595-8600.

Huhman, D.V., and Sumner, L.W. (2002). Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. Phytochemistry **59,** 347-360.

Naoumkina, M.A., He, X.Z., and Dixon, R.A. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. BMC Plant Biology **8,** 132-146.

Suzuki, H., Reddy, M., Naoumkina, M., Aziz, N., May, G., Huhman, D., Sumner, L., Blount, J., Mendes, P., and Dixon, R. (2005). Methyl jasmonate and yeast elicitor induce differential transcriptional and metabolic re-programming in cell suspension cultures of the model legume *Medicago truncatula*. Planta **220,** 696-707.

Zhou, C., Han, L., Pislariu, C., Nakashima, J., Fu, C., Jiang, Q., Quan, L., Blancaflor, E.B., Tang, Y., Bouton, J.H., Udvardi, M., Xia, G., and Wang, Z.-Y. (2011). From model to crop: Functional analysis of a STAY-GREEN gene in the model legume *Medicago truncatula* and effective use of the gene for alfalfa improvement. Plant Physiology **157,** 1483-1496.

4

Ccr4-associated factor1 (Caf1) is a component of a large regulatory complex potentially involved in *Medicago truncatula* secondary metabolism

Azra Gholami, Jacob Pollier, Tessa Moses, Alain Goossens

Author contribution:

Molecular cloning, generation of M. truncatula hairy roots, qRT-PCR transcript profiling, MeJA elicitation, data interpretation.

Abstract

Plants are excellent manufacturers of secondary metabolites that have served as sources of pharmaceuticals and industrial materials. An in-depth understanding of the metabolite biosynthetic pathways is necessary for the improvement of the yield of secondary metabolites. Plant metabolites are produced through complex processes often including multiple enzymatic steps, branched pathways and regulation by a number of functionally redundant transcription factors (TFs). TFs are well-known as master regulators of various plant functions and many of them are characterized as the key regulators of metabolic pathways. Here, we have studied a number of TFs and other potential regulatory proteins that may act as jasmonate-modulated master switches towards the biosynthesis of secondary metabolites. To identify putative regulators of secondary metabolism in M. truncatula a reverse-genetics screen was performed, starting from a genome-wide cDNA-AFLP transcript profiling analysis of MeJA elicited cell cultures of M. truncatula. In this study, we focused on the potential role(s) of the Mt148 gene by assessing the effect of gain-of- function on secondary metabolite production in M. truncatula hairy roots. Mt148 encodes a CCR4-associate factor1 (Caf1) protein presumably involved in mRNA metabolism and post- transcriptional gene regulation. We found that overexpression of Mt148 caused downregulation of some secondary metabolism genes in M. truncatula hairy roots by a yet unknown molecular mechanism.

1 Introduction

Plants are excellent manufacturers of secondary metabolites that have served as sources of pharmaceuticals, food additives, flavors, and other industrial materials. Furthermore, secondary metabolism plays an important role in the survival of the plant in its environment and can be a defense response against biotic and abiotic stresses. Through stress response processes, a large number of metabolic genes undergo transcriptional regulation associated with dramatic changes in their expression level. Transcriptional control requires specific signals to be transduced to the cell nucleus where the specific sets of the target genes are regulated. The transcription regulation network is a subset of a larger genetic regulatory network, which in addition to transcriptional regulation includes translational regulation, RNA editing, and so forth.

Messenger RNA (mRNA) degradation is also an essential process that allows rapid changes in gene expression profile of a cell, especially in response to stress signals. In both mammalian and yeast cells, the initial and rate-limiting step of the mRNA

turnover is degradation of the poly (A) tail at the 3′ end of the mRNA (deadenylation) by a variety of deadenylases. The deadenylated mRNA can trigger the removal of the cap at the 5′ end of the mRNA by the decapping enzyme, Dcp1 and Dcp2, and subsequently, the Xrn1 exoribonuclease hydrolyses the RNA body from its 5′ end. Alternatively, deadenylated mRNAs can be degraded in a 3′ to 5′ direction by the cytoplasmic exosome complex (Chiba and Green, 2009)

"The carbon catabolite repressor protein 4 (Ccr4) - Negative on TATA (Not) complex" is a global regulatory complex conserved from yeast to human (Collart and Panasenko, 2012; Miller and Reese, 2012). In *Saccharomyces cerevisiae*, the Ccr4-Not complex comprises a core complex of 9 subunits: Ccr4, Ccr4 associated factor1 (Caf1), Not1 to Not5, Caf130 and Caf40. Not1 is the scaffold of the complex. Ccr4 and Caf1 bind to the N-terminal region of Not1, and the association of Ccr4 is critically dependent upon Caf1. Not2 binds to the most C-terminal portion of Not1. Not4 and Not5 bind to the same region of Not1, just N-terminally to Not2. Not3 is also thought to bind to the C-terminal portion of Not1 (Collart, 2003) (Fig. 1). The complex affects gene expression at two levels, in the nucleus and the cytoplasm. In the nucleus, it regulates the basal transcription machinery, nuclear receptor-mediated transcription and histone modifications via transcription initiation and ubiquitination activities provided by the C-terminal elements of the complex (Fig. 1). In the cytoplasm, the complex is entangled with mRNA turnover through its two associated deadenylases, Ccr4 and Caf1. As such, the Ccr4–Not complex is involved in regulation of several aspects of mRNA metabolism including initiation, elongation, deadenylation and subsequent degradation.

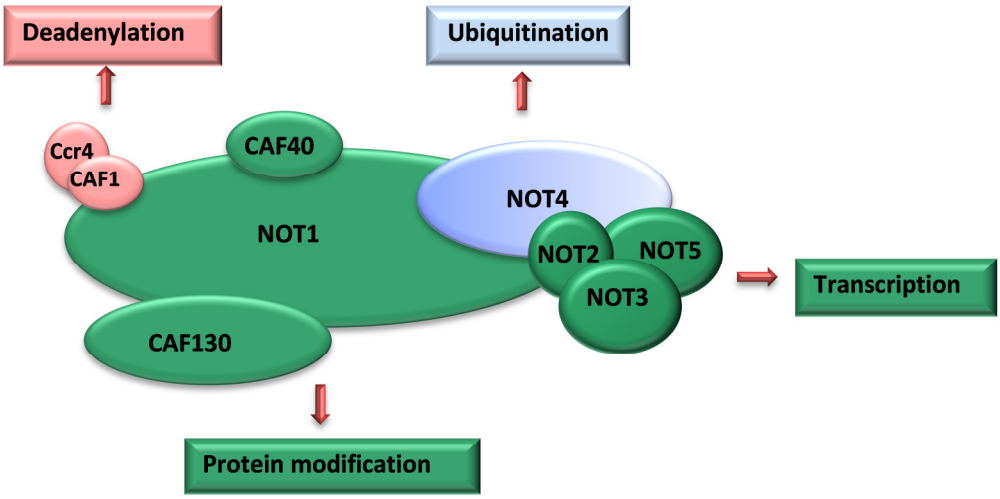


Figure 1. Scheme of the Ccr4–Not complex. Different subunits of the Ccr4–Not complex are grouped in color codes according to their function: subunits of the complex responsible for deadenylation activity (Ccr4 and Caf1) are indicated in pink and that for ubiquitination activity (Not4) in blue. The other subunits of the complex are indicated in green.

Caf1 and Ccr4 are directly involved in mRNA deadenylation. Ccr4 contains an exonuclease domain, belongs to the endonuclease–exonuclease–phophatase (EEP) superfamily and has a 3′ exoribonuclease activity with a preference for poly (A) substrates (Chen et al., 2002). Caf1 proteins belong to the DEDDh subgroup of the DEDD family of nucleases (Zuo and Deutscher, 2001). Caf1 links Ccr4 to the core of the Ccr4–Not complex (Tucker et al., 2002). It is also associated with Dhh1 which is a putative RNA helicase and a component of the mRNA decapping complex (Collart and Panasenko, 2012). Although the yeast Caf1 shows deadenylase activity *in vitro*, its deadenylase activity *in vivo* remains obscure so far.

Whereas Ccr4-Caf1 has been extensively investigated in non-plant species (including yeast and human) (Tucker et al., 2001; Chen et al., 2002; Temme et al., 2004; Bianchin et al., 2005; Morris et al., 2005; Yamashita et al., 2005), only a limited number of studies have been conducted on the Ccr4-Caf1 complex in plants. In *Arabidopsis*, two hormone and stress inducible *CAFs*, *AtCAF1a* and *AtCAF1b*, exhibited deadenylase activity *in vitro* and could partially complement the growth defect of the yeast *caf1* deletion strain (Liang et al., 2009). These observations as well as their roles in pathogen resistance and stress-tolerance responses suggested a role as regulators of stress-responsive genes (Liang et al., 2009). Furthermore, overexpression of *CaCAF1*, a *CAF1* homolog from pepper (*Capsicum annuum*) led to growth enhancement and resistance to the oomycete pathogen, *Phytophthora infestans*, in tomato. Conversely, silencing of the *CaCAF1* gene resulted in a

significant growth retardation and plants were more susceptible to pathogens (Sarowar et al., 2007).

In this study, from a genome-wide study of candidate regulatory genes that are inducible by MeJA prior to, or simultaneous with secondary metabolite biosynthesis genes in MeJA- elicited *M. truncatula* cell cultures, a set of 11 MeJA- induced genes encoding for potential regulatory factors were chosen for functional characterization. Among these MeJA-induced tags, we identified *Mt148* gene encoding a putative Caf1 protein. Overexpression of *Mt148* in transgenic *M. truncatula* hairy roots led to transcriptional repression of some secondary metabolite biosynthesis genes. Further experimental analyses were undertaken but we could not clarify the molecular mechanism behind this effect.

2 Results

2.1 Selection of genes potentially involved in transcriptional regulation of *M. truncatula* secondary metabolism

It has been previously shown that exposure of *M. truncatula* cell cultures to MeJA led to the increased accumulation of various triterpene saponin compounds orchestrated via a complex signaling network in which regulatory elements are essential components (Broeckling et al., 2005; Naoumkina et al., 2010). Previous to this study, a cDNA-AFLP transcript profiling was performed on MeJA-elicited cell cultures of *M. truncatula* (Pollier, 2011). Here, a set of 11 genes encoding regulatory proteins and proteins of unknown function potentially involved in the regulation of secondary metabolite biosynthesis in *M. truncatula* were selected based on their induction in response to MeJA prior to or simultaneous with the secondary metabolite biosynthesis genes (Fig. 2). An overview of the selected genes is given in Table 1.

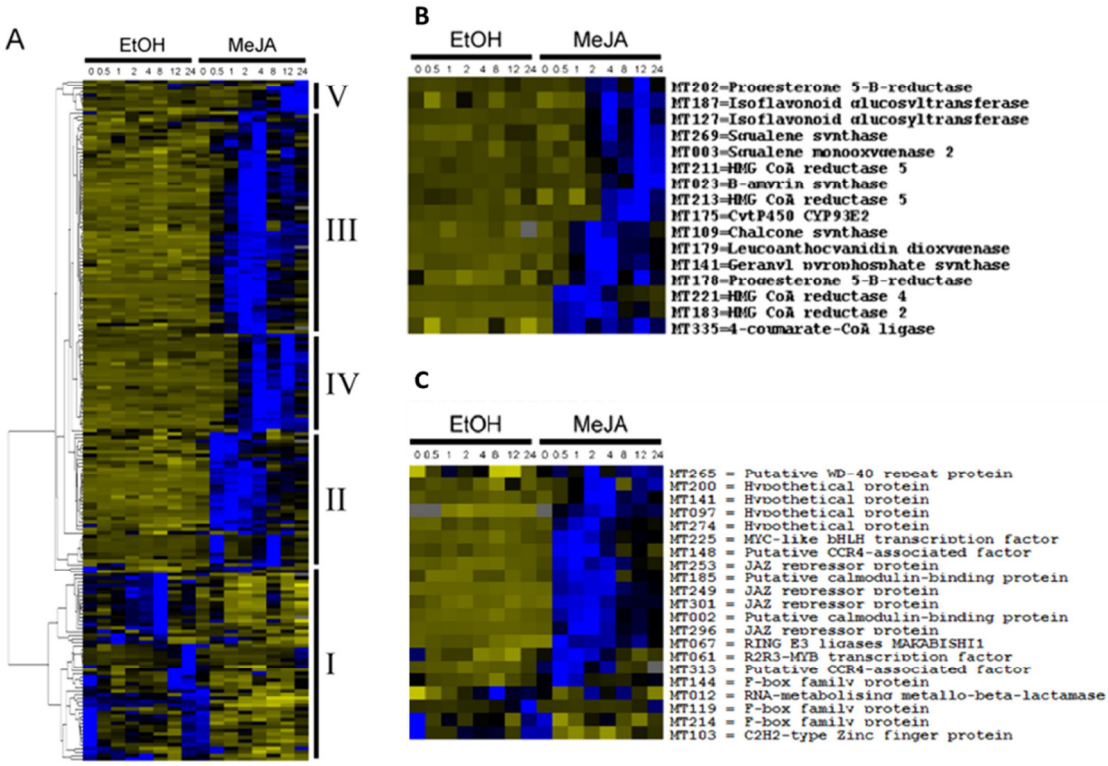


Figure 2. Transcriptome of MeJA-elicited *M. truncatula* cells. A. General view on the average linkage hierarchical clustering of *M. truncatula* gene tags. B. Tags corresponding to genes reported to be involved in secondary metabolite biosynthesis, or with high sequence similarity to such genes. C. Tags corresponding to genes encoding putative regulatory proteins with a potential role in regulation of secondary metabolism or JA signaling cascade. Treatments and time points (h) are indicated at the top. Blue and yellow boxes reflect transcriptional activation and repression by methyl jasmonate (MeJA) relative to the average expression level, respectively. Gray boxes correspond to missing time points. The figure is adapted from Pollier (2011).

Table 1. List of *M. truncatula* genes selected through cDNA-AFLP transcript profiling of MeJA-elicited cell cultures and their responses to MeJA elicitation checked by qPCR analysis. HR: hairy root cultures; CS: cell suspension cultures, MI: MeJA- inducible; MR: MeJA repressed; NE: No effect of MeJA; NT: Not tested.

cDNA-AFLP tag	(putative) annotation	Response to MeJA elicitation	
		HR	CS
Mt002	Putative calmodulin-binding protein	MI	NT
Mt012	RNA-metabolising metallo-beta-lactamase	MR	NE
Mt061	MYB transcription factor	NE	MI
Mt097	Hypothetical protein	MI	MI
Mt141	Hypothetical protein	NE	NE
Mt148	Putative Ccr4-associated factor	MI	NT
Mt185	Putative calmodulin-binding protein	MI	NT
Mt200	Hypothetical protein	NE	MI
Mt265	Putative WD-40 repeat protein	NE	NE
Mt274	Hypothetical protein	MI	NT
Mt313	Putative Ccr4-associated factor	MI	NT

2.2 Expression levels of regulatory genes were increased in MeJA-elicited *M. truncatula* hairy roots prior/ concomitant to the onset of secondary metabolite

To confirm the MeJA responsiveness of the selected regulatory genes, a MeJA elicitation experiment with an extensive time course was set up on hairy root cultures. A transcript profiling was then performed on the elicited cultures by means of qRT-PCR analysis. Following MeJA treatment of root, 6 genes out of the 11 selected genes showed rapid induction following MeJA elicitation (Fig. 3A; the black graphs), while only one (*Mt012*, encoding RNA-metabolising metallo-beta-lactamase) was repressed (Fig. 3A, the white graph). Next, the steady-state levels of the genes that did not alter in MeJA-elicited hairy roots (Fig. 3A, the gray graphs) were checked in MeJA-elicited cell cultures. The results revealed that following MeJA elicitation, the steady state levels of *Mt061* and *Mt200* were rapidly increased in the elicited cell suspension cultures (Fig. 3B). Two genes (*Mt141* and *Mt265* encoding a hypothetical protein and putative WD-40 repeat protein, respectively) did not respond to MeJA elicitation neither in hairy roots nor in suspension cell cultures of *M. truncatula* (Fig. 3A, B). A summary is given in Table 1.

Based on this analysis, from these MeJA-inducible putative regulatory genes, the genes encoding a MYB (*Mt061*) transcription factor, a putative calmodulin-binding (*Mt185*) protein, a putative Ccr4-associated factor (*Mt148*) and three hypothetical proteins (*Mt097*, *Mt200* and *Mt274*) were selected for further analysis. All six selected genes had maximal transcriptional upregulation within 30 min to 2 h after elicitation, which was prior or simultaneous to the onset of the genes involved in triterpene saponin biosynthesis pathway (Pollier, 2011) (Fig. 3A & B). For all selected genes, a full-matching expressed sequence tag (EST) was encountered in the DFCI *Medicago* Gene Index (http://compbio.dfci.harvard.edu/cgibin/tgi/gimain.pl?gudb=medicago).

A

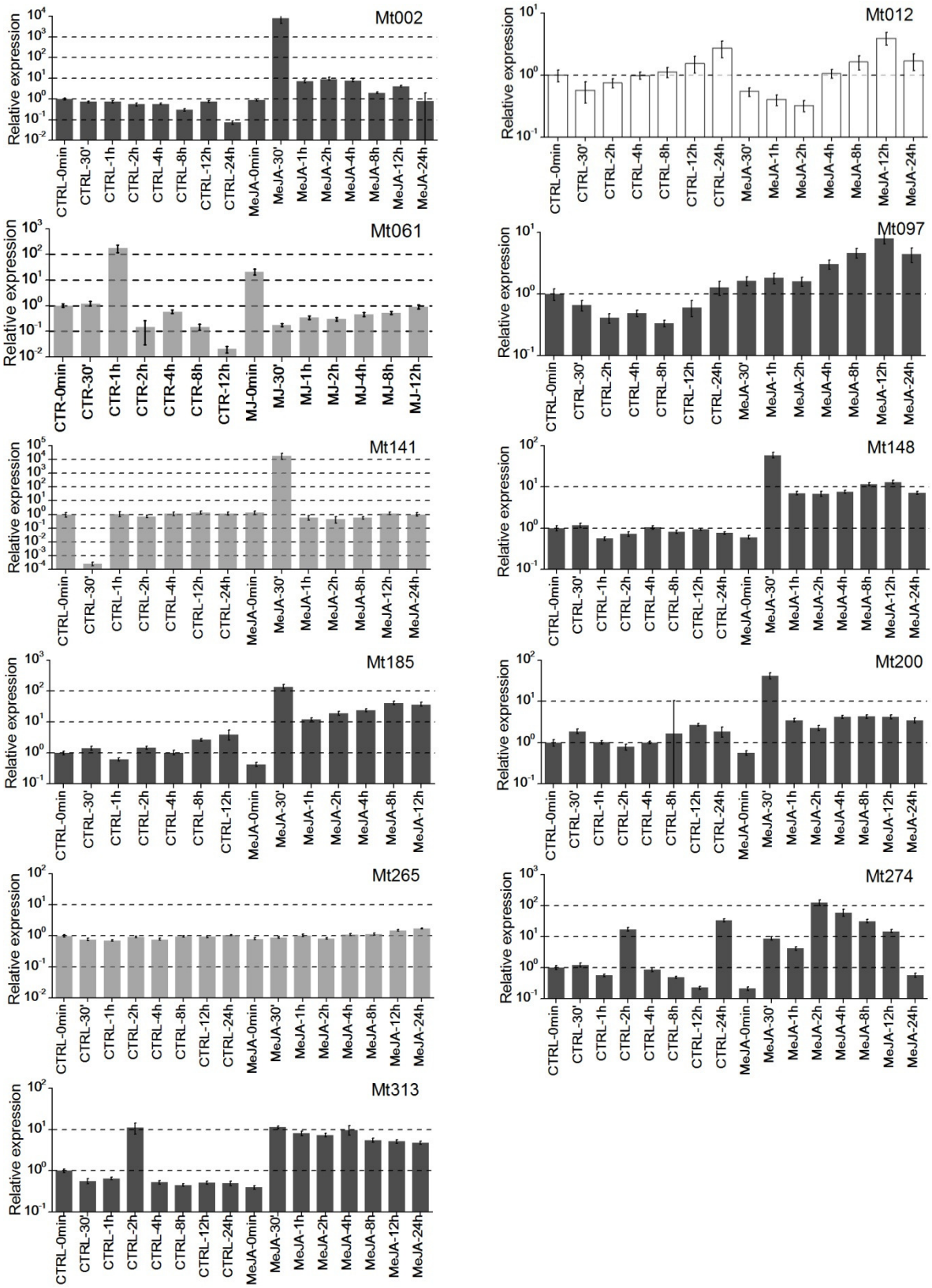


Figure 3A. Transcript analysis of *M. truncatula* putative regulatory genes in MeJA-elicited hairy root cultures. Treatments and time-points are indicated at the bottom of the graphs. Black: MeJA-induced, white: MeJA-repressed; gray: no effect of MeJA. Y-axis represents relative expression ratio as compared to the control (CTRL) treatment at time 0. Error bars represent the standard error of the mean (SEM) of three technical repeats.

B

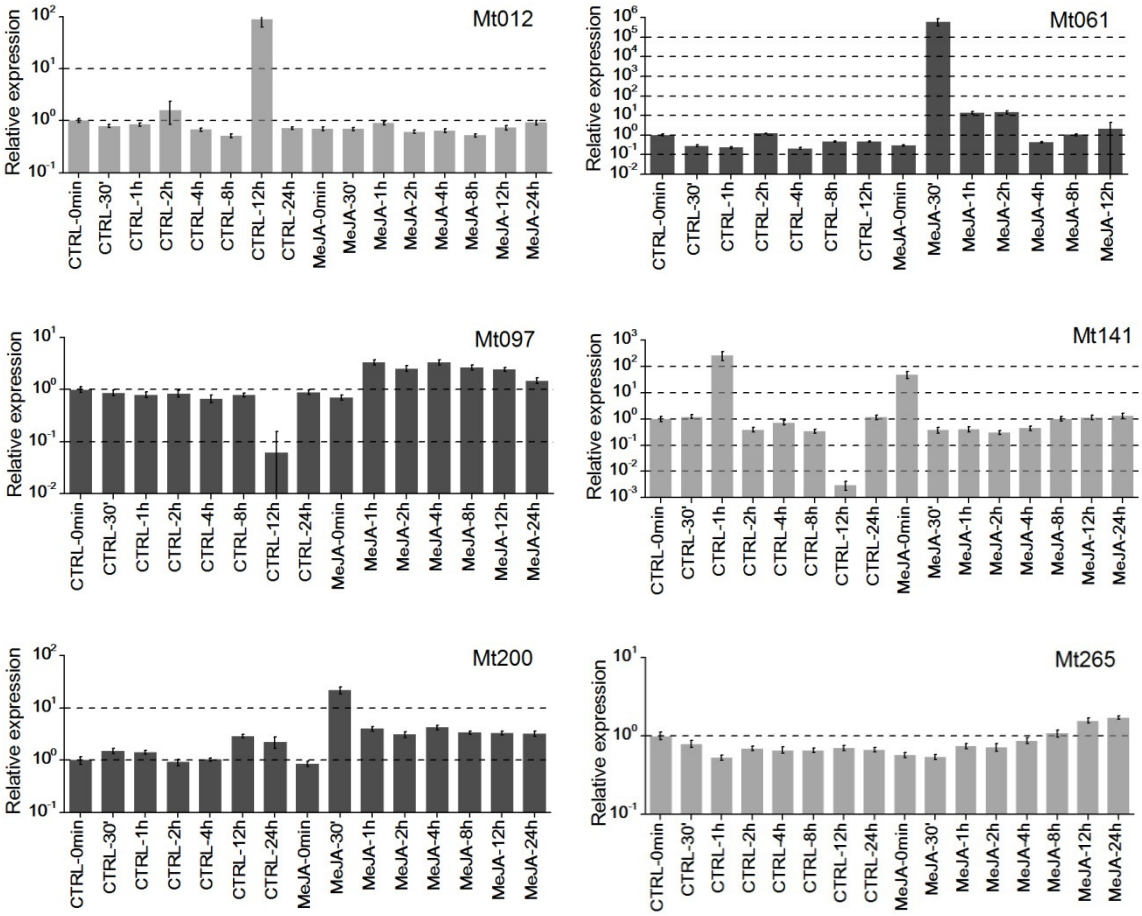


Figure 3B. Transcript analysis of *M. truncatula* putative regulatory genes in MeJA- elicited cell suspension cultures. Treatments and time-points are indicated at the bottom of the graphs. Black: MeJA-induced, gray: no effect of MeJA. Y-axis represents relative expression ratio as compared to the control (CTRL) treatment at time 0. Error bars represent SEM of three technical repeats.

2.3 Overexpression of the *Mt148* gene resulted in a repression of the secondary metabolite biosynthetic genes

To functionally characterize the candidate regulatory genes, a reverse genetics screen was launched, in which gain- and loss-of-function of the candidate genes was pursued by constitutive CaMV 35S-mediated overexpression of the full-length open reading frame (FL- ORF) or of a hairpin RNA-mediated interference (hpRNAi) construct, respectively, in *M. truncatula* hairy roots. Overexpression and knock-down of the transgene in hairy roots were confirmed via qRT-PCR analyses. Three control lines and three independent transgenic lines for each construct were selected (Fig. 4). Since most selected genes showed low expression levels in *M. truncatula* hairy roots, the RNAi lines were generated only for *Mt148* and *Mt185*. However, the *Mt148* knock-

down lines were not stable and in subsequent subcultures they showed the same expression levels as the control lines. Results for all 6 genes, except *Mt061* are shown in this chapter. The characterization of *Mt061* is discussed in chapter 5.

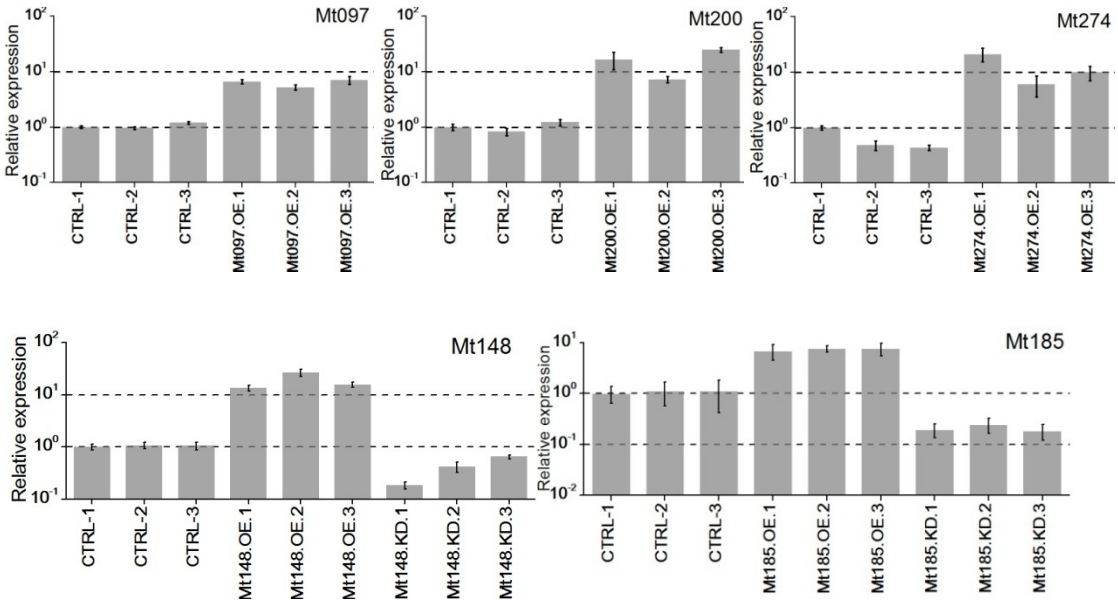


Figure 4. qRT-PCR analysis of the transgenic hairy roots harboring overexpression or hpRNAi constructs. Numbers in the Y-axis represent relative expression ratio as compared to the control line 1. Error bars represent SEM of three technical repeats.

Subsequently, to assess the potential role of the candidate genes in the regulation of secondary metabolite biosynthesis, the steady state levels of the genes encoding the enzymes catalyzing saponin and (iso)flavonoid biosynthesis, the two major secondary metabolite groups in legumes (Dixon and Sumner, 2003), were measured in transgenic hairy roots (Fig. 5). For none of the transgenic lines, except the Mt148^{OE} lines, any remarkable effect was observed. These lines were therefore not considered for further characterization.

Transcript profiling of the Mt148^{OE} hairy roots revealed that overexpression of the *Mt148* gene led to a slight transcriptional downregulation of several saponin and isoflavonoid biosynthesis genes including those encoding phenylalanine ammonia-lyase (PAL) and isoflavone synthase (IFS), catalysts in the isoflavonoid pathway, and beta-amyrin synthase (BAS), squalene synthase (SQS), squalene epoxide (SQE2), and UGT73K1 involved in triterpene saponin biosynthesis (Fig. 5).

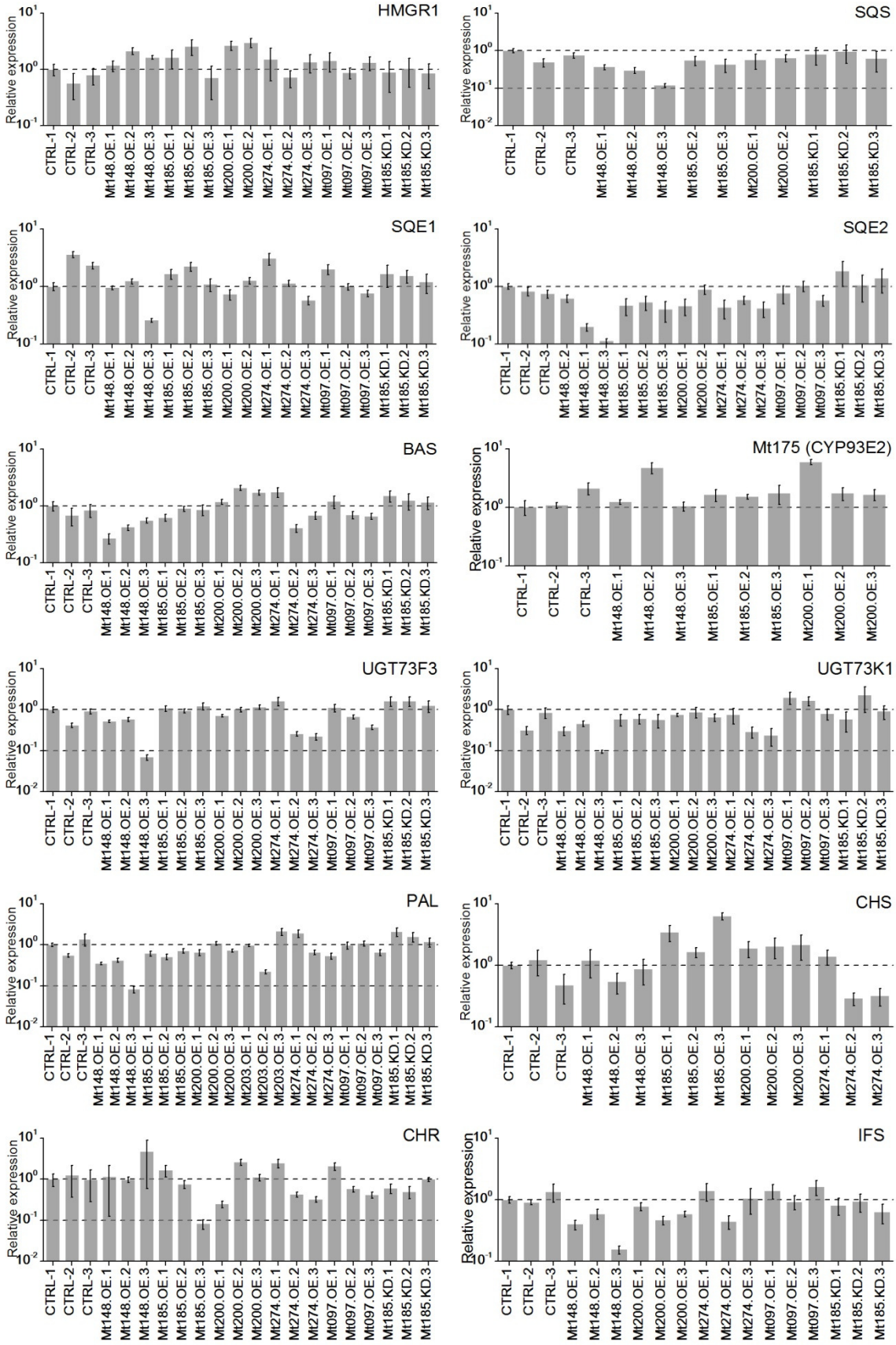


Figure 5. Transcript profiling of transgenic *M. truncatula* hairy roots. CTRL, control lines; OE, overexpression lines, KD, knockdown lines. Numbers in the Y-axis represent relative expression ratio as compared to the control line 1. Data and error bars represent means ±SEM of three technical repeats.

2.4 *Mt148* encodes a Ccr4 associated factor1 (CAF1) proteins

Mt148 (MTR_4g006800) is annotated as the *Ccr4 associated factor1 (CAF1)* gene. It encodes a putative Caf1 protein with 277 amino acids and a calculated mass of 31.653 kDa and a predicted pI of 4.77. BLAST searches revealed that there are seven additional putative CAF1 sequences have been identified in the *Medicago* genome, so far (Fig. 6). Phylogenetic analysis of Mt148 with *M. truncatula*, *Arabidopsis* and *Saccharomyces cerevisiae* CAF1 (Pop2p) homologs shows that the closest homologs of Mt148 are MTR_3g106180, At2G32070 and At1G80780 with 82%, 80% and 80% amino acid similarities, respectively (Fig. 6).

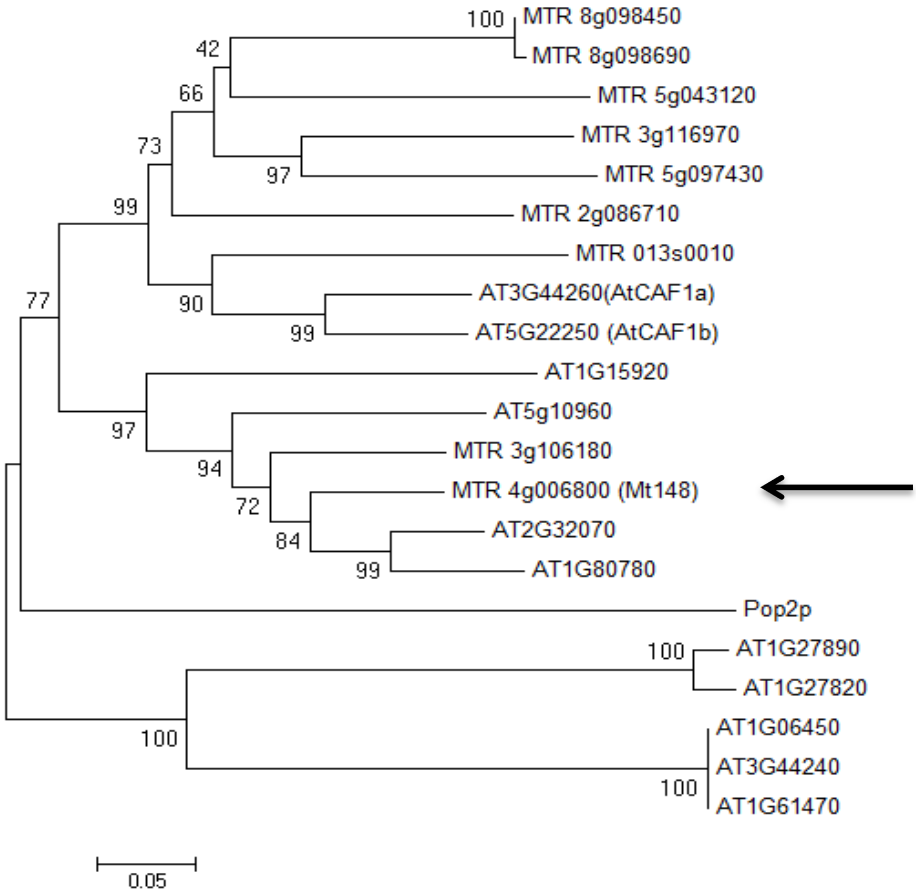


Figure 6. Phylogenetic analysis of *Mt148* and related *Arabidopsis*, *Saccharomyces cerevisiae* (Pop2p) and *M. truncatula* homologs. Neighbor-joining tree showing the evolutionary relationships between *Mt148* and its closest homologs. The percentage of replicate trees that clustered together in the bootstrap test is shown next to the branches. Gene names or protein accession numbers are given.

The results obtained from transcriptional analysis pointed towards an overall repression of metabolic biosynthesis in the Mt148^{OE} lines. Therefore, we investigated the role of the *Mt148* in the regulation of secondary metabolite biosynthesis more thoroughly. SQS and SQE are involved in biosynthesis of oxidosqualene as a

precursor common to the biosynthesis of both steroids and triterpenoids (Abe et al., 1993). Since both *SQS* and *SQE2* transcript levels are decreased in Mt148^{OE} lines, we also measured the transcript levels of the putative *M. truncatula* sterol biosynthesis genes that are homologs of the genes encoding sterol biosynthesis enzymes in *Arabidopsis*, including cycloartenol synthase (CAS), sterol methyltransferase1 (C24MT), sterol methyltransferase2 (CVP1), obtusifoliol 14α-demethylase (CYP51G1), sterol C- 14 reductase (FACKEL), sterol C-8,7 isomerase (HYDRA1), and CYP710A15 (Carland et al., 2002; Schrick et al., 2004; Morikawa et al., 2006). Among these, only CVP1 and CYP710A15 were slightly decreased (Fig. 7), suggesting *Mt148* might have a broader effect than on secondary metabolism only.

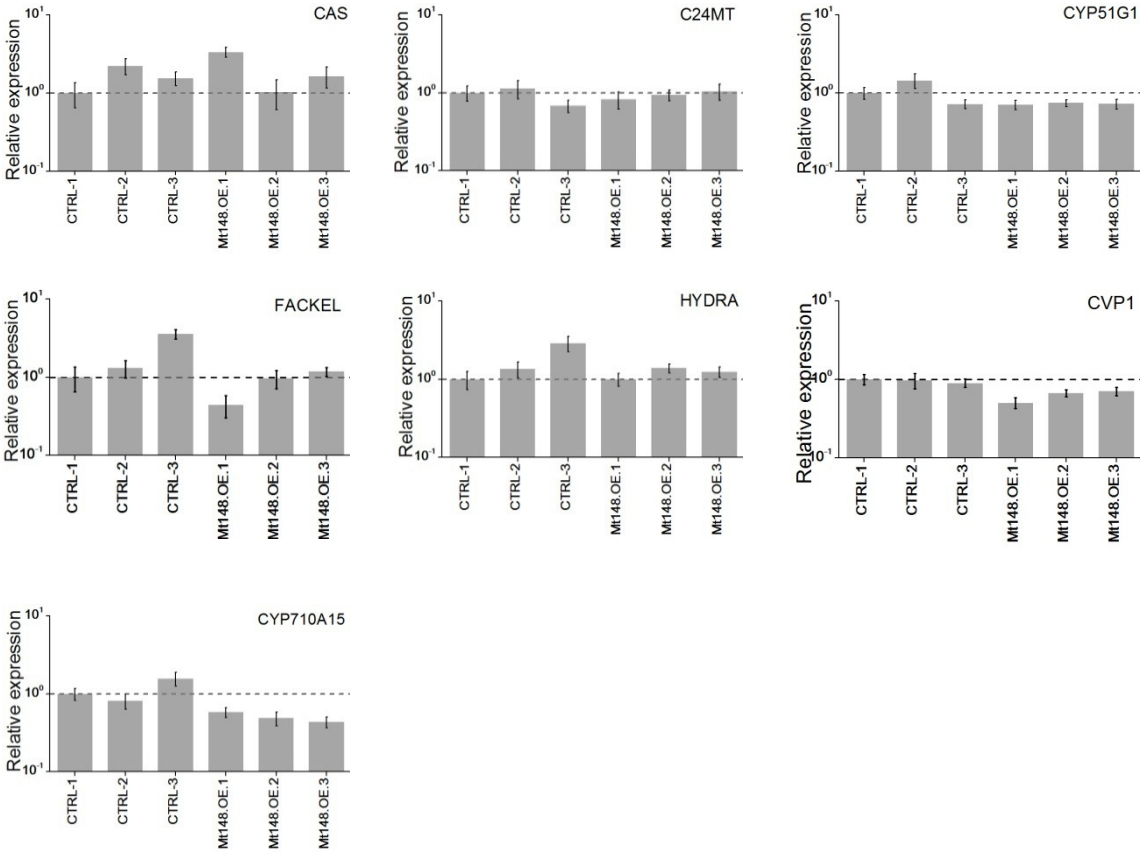


Figure 7. qRT-PCR analysis of sterol biosynthetic genes in Mt148^{OE} hairy roots. CTRL, control lines; Mt148OE, *Mt148* overexpression lines. Numbers in the Y-axis represent relative expression ratio as compared to the control line1. Data and error bars represent means ±SEM of three technical repeats.

2.5 *Mt148* is not involved in RNA degradation

The Caf1 protein, in association with the Ccr4 protein is presumably involved in cytoplasmic mRNA deadenylation, which subsequently leads to mRNA degradation. Considering the downregulation of several primary and secondary metabolite biosynthesis genes in Mt148^{OE} hairy roots (Fig. 5 and 7), we explored whether Mt148

is involved in mRNA degradation of the corresponding transcripts. Therefore, we performed an RNA degradation assay in which changes in the transcription levels of *BAS* and *Mt148* transcripts were measured over a time course following application of plant RNA synthesis inhibitors. RNA degradation rate can depend on the inhibitor of RNA synthesis used (Holtorf et al., 1999). Therefore, two separate experiments were launched using two different RNA inhibitors, cordycepin and actinomycin D (Figs. 8 and 9).Cordycepin, also known as 3′-deoxyadenosine, is a known polyadenylation inhibitor whereas actinomycin D is an inhibitor of DNA-directed RNA synthesis. To confirm that transcription was efficiently inhibited, we tested the ability of cordycepin and actinomycin D to prevent induction of transcription upon addition of MeJA, known to result in an increase in transcriptional level of *BAS* and *Mt148*. Samples were harvested in four time points. The rate of possible RNA degradation was scored via qRT-PCR. The results showed that pretreatment of hairy root cultures with cordycepin prevented rapid induction of *BAS* and *Mt148* observed after addition of MeJA in both $MT148^{OE}$ and control lines indicating efficient inhibition of transcription by cordycepin (Fig. 8A and B). MeJA caused a rapid induction in transcript levels of the *BAS* gene in both control and $Mt148^{OE}$ lines followed by a rapid decline in the *BAS* transcript levels in $Mt148^{OE}$ lines whereas in control lines its transcript levels were steadily increased, However, when the experiment was executed with actinomycin D the trend did not repeat (Fig. 9A) indicating that despite some variation no marked changes has been observed in transcript level of the genes in $Mt148^{OE}$ lines as compared to control lines.

Effective transcription inhibition by actinomycin D was also confirmed by the prevented induction in transcriptional level of *Mt148* in control lines observed after addition of MeJA (Fig. 9B). However, the same result was not observed on *BAS* transcription level that might indicate that actinomycin D is not effective on the *BAS* gene (Fig. 9A). Consistence with the experiment with cordycepin, no consistent and remarkable differences were observed in transcription levels of the gene in $Mt148^{OE}$ and control lines (Fig. 9) . Therefore, according to these results it seems unlikely that Mt148 affects degradation of e.g. *BAS* transcripts.

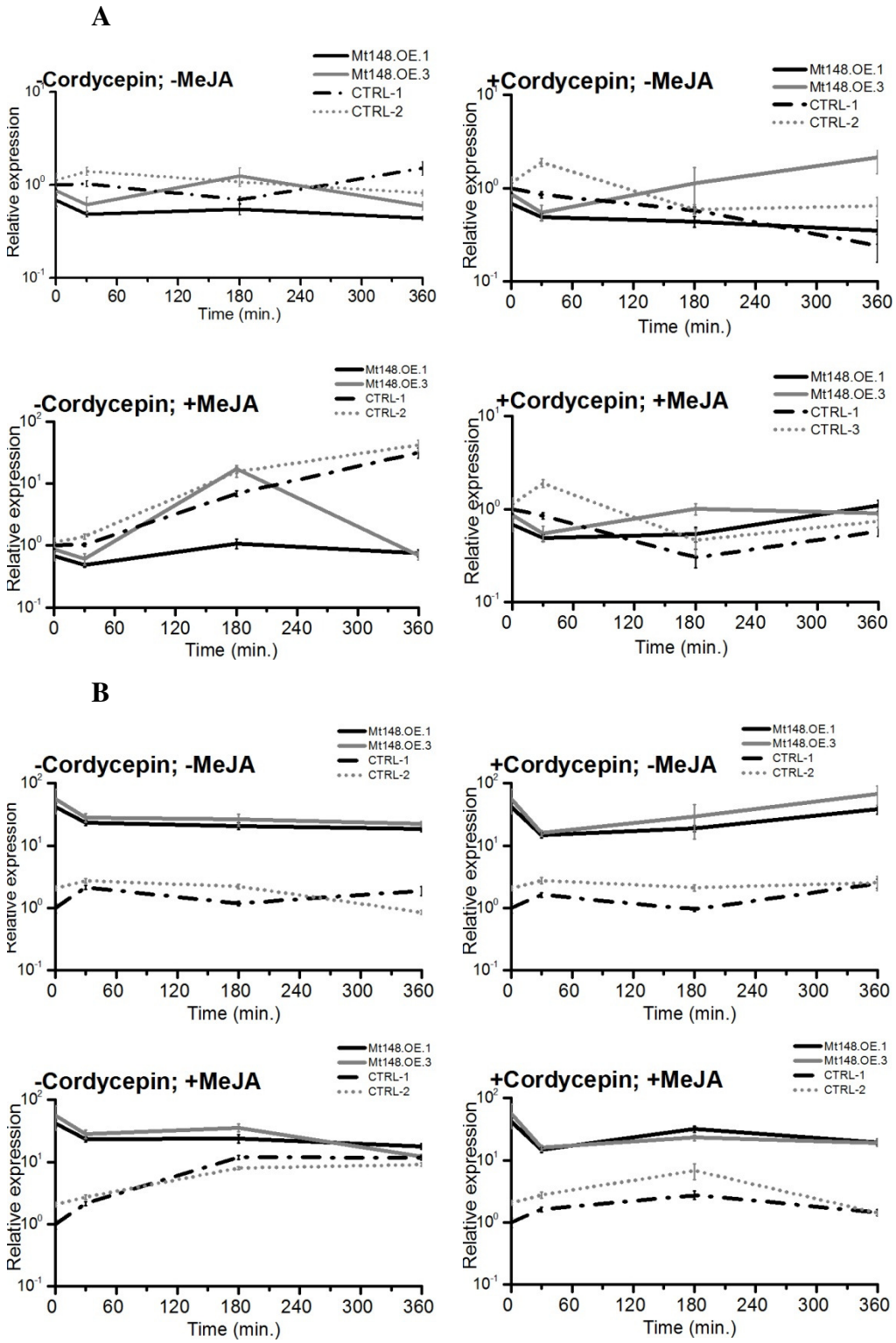


Figure 8. qRT-PCR analysis of (A) *BAS* and (B) *Mt148* transcripts in *M. truncatula* Mt148^{OE} and control (CTRL) hairy roots treated with cordycepin and/or MeJA. Expression level at different time points is represented as a ratio relative to the CTRL-1 at time 0, which is set to 1. Error bars represent SEM of three technical repeats.

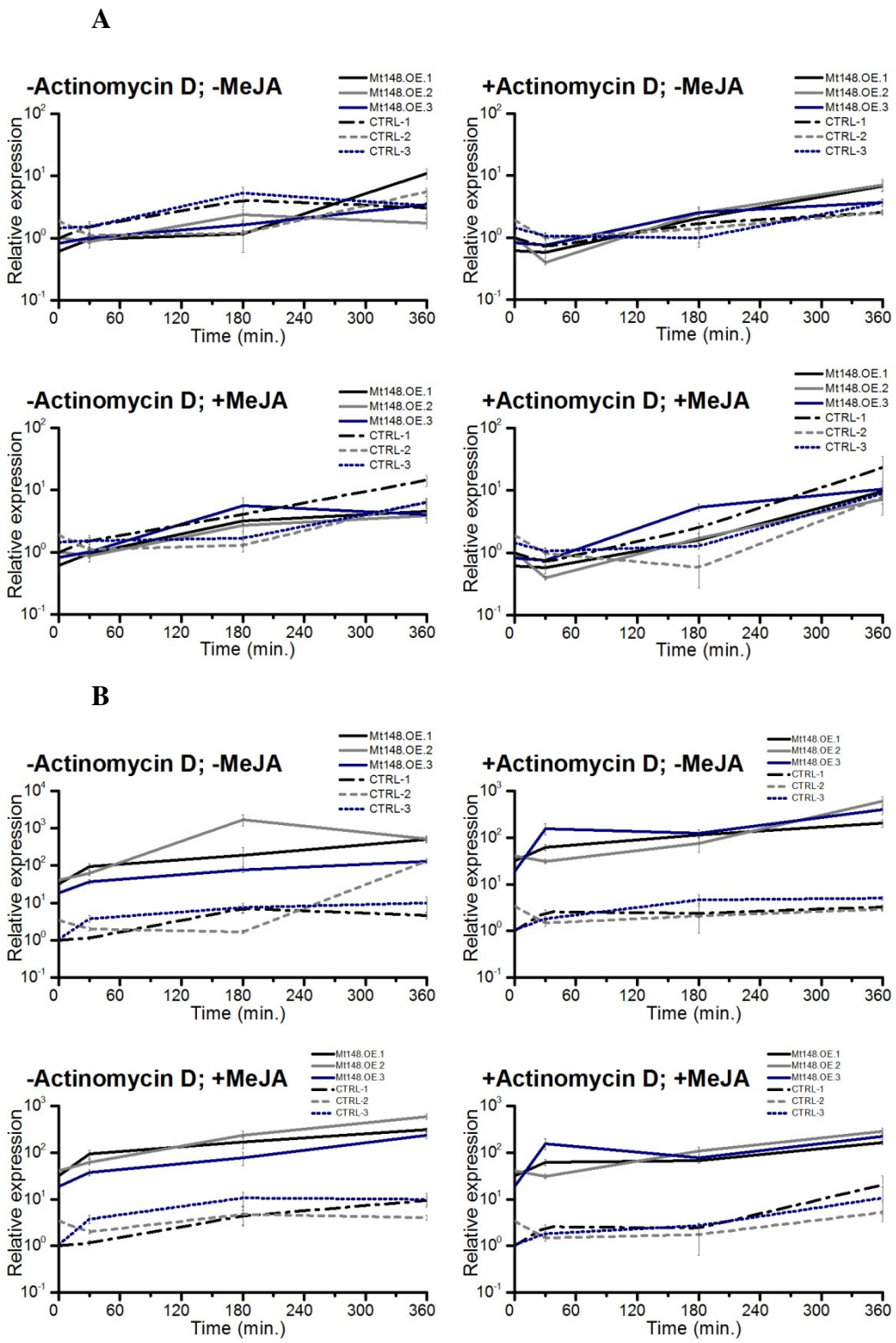


Figure 9. qRT-PCR analysis of (a) *BAS* and (b) *Mt148* transcripts in *M. truncatula* Mt148^{OE} and control (CTRL) hairy roots treated with actinomycin D and/or MeJA. Expression level at different time points is represented as a ratio relative to the CTRL-1 at time 0, which is set to 1. Error bars represent SEM of three technical repeats.

2.6 *Mt148* gene induces growth defect in yeast

Yeast *caf1* mutants are sensitive to high temperature and caffeine (Liang et al., 2009). To examine the functional properties of *Mt148*, the full length gene was first heterologously expressed in the yeast *caf1* mutant strain KY803-c1 and its wild type background KY803. The complementation was scored by determining the ability of the transformed strain to grow on caffeine containing medium. The results showed that the pAG426GPD-*Mt148* construct was unable to restore the growth defect of the *caf1* mutant. Surprisingly, it even caused a growth defect in the KY803 wild type strain (Fig. 10A). KY803 is a strain with *GCN4* deficiency and *GCN4* is a general transcriptional activator of amino acid biosynthetic genes (Wolfner et al., 1975). To exclude that this would interfere with *Mt148* activity, we re-launched the experiment with the BY4741 strain and its respective Δ*caf1* mutant that possess functional *GCN4*. Identical results were obtained with these strains as well (Fig. 10B). Hence, the results indicate that *Mt148* is not a functional homologue of yeast *CAF1*. However, it is not clear why *Mt148* interferes with yeast growth. Perhaps it might affect functioning of other elements of the endogenous yeast CCR4-NOT complex.

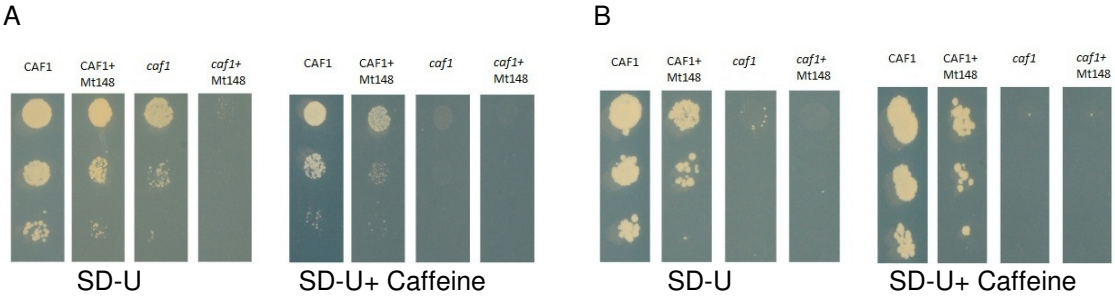


Figure 10. Effects of heterologous expression of the *M. truncatula Mt148* gene on the growth of, (A) the yeast *caf1* mutant, KY803-c1, and its respective wild type strain, KY803; (B) Y07123 *caf1* mutant and its corresponding wild type, BY4741. All strains harbored pAG426GPD;*Mt148* or empty pAG426GPD vector (as control) and were grown on SD-U plates with and without 5 mM caffeine at 28 °C. In each panel, from the top to bottom, 10-fold fewer cells were plated in each row. CAF1, KY803 or BY4741*; caf1*:*caf1* mutant*,* KY803-c1or Y07123.

2.7 *Mt148* overexpression does not influence metabolite accumulation

Considering the co-regulation of the *Mt148* and saponin biosynthesis genes in MeJA-elicited cell cultures of *M. truncatula* (Pollier, 2011), metabolite profiling was performed by liquid chromatography electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (LC-ESI-FT- ICRMS) to verify the effect of *Mt148* overexpression on the metabolite profile of transgenic hairy roots (Fig. 11A). Comparative analysis of the extracts of Mt148^{OE} and control hairy roots yielded a total of 5,506 *m/z* peaks, containing 65 peaks which were significantly changed ($p \leq$

10E-05). However, subsequent Principal Component Analysis (PCA) revealed that the variation was mostly due to experimental variation rather than differences between the two lines. Hence, *Mt148* overexpression did not seem to affect the metabolism of *M. truncatula* hairy roots (Fig. 11A).

We also profiled the metabolic response of the Mt148^{OE} lines to MeJA elicitation. Comparative data analysis of the LC-ESI-FT- ICRMS spectra from MeJA-elicited roots with that of non-elicited roots revealed a total of 10,617 *m/z* peaks, of which 2,129 m/z peaks discrete between elicited and non-elicited lines were observed, and 2,305 peaks were significantly changed ($p \leq$ 10E-05) (Fig. 11B). However, statistical analysis of data did not reveal marked general differences in MeJA response in terms of metabolite accumulation between Mt148^{OE} and control hairy root lines (Fig. 11B). Exogenous MeJA induced higher accumulation of known triterpene saponins, including soyasaponin I and Rha-Gal-GlcA-soyasapogenol E (2.39 fold, *p* value 4.4E-14 and 3.08 fold, *p* value 1.55E-12 higher induction, respectively) in both MT148^{OE} and control lines (Fig. 11C). Collectively, these results showed that the MeJA responses leading to the high accumulation of triterpene saponins in hairy roots are not affected by *Mt148* overexpression.

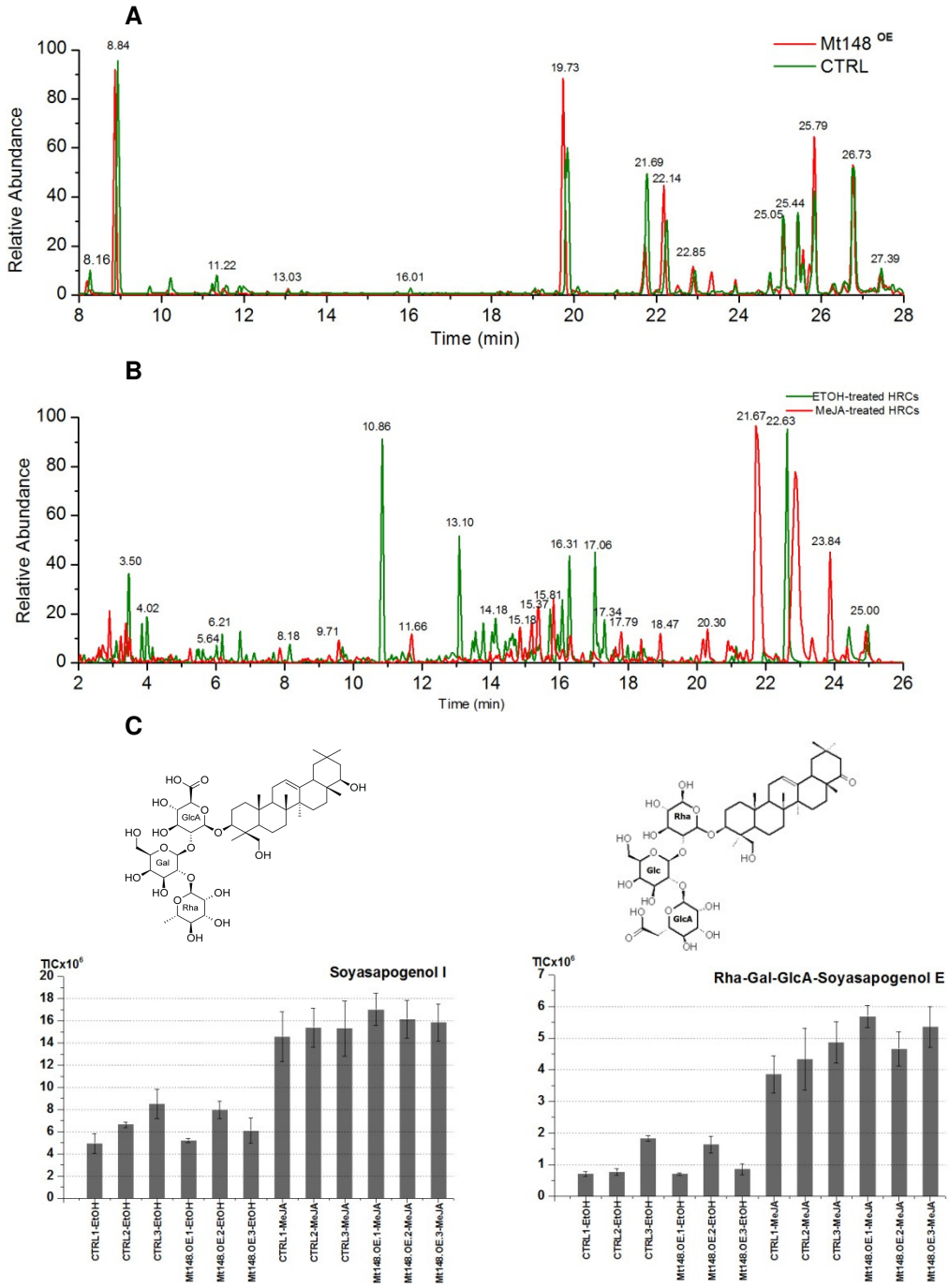


Figure 11. Saponin profiling of Mt148^{OE} and control hairy roots. (A) Comparison of the chromatogram of a control line (green) and a Mt148^{OE} line (red). No significant differences have been observed. (B) Detail of the full MS scan of a MeJA-treated line (red) and a mock (ETOH)-treated hairy root (green). The peak at t_R 21.67 and 23.84 represent Soyasaponin I and Rha-Gal-GlcA-Soyasapogenol E, respectively. The number indicated on each peak shows the retention time of the corresponding compound (C) Average Total Ion Current (TIC) of the main masses corresponding to Soyasaponin I (left) and Rha-Gal-GlcA-Soyasapogenol E (right). Error bars indicate SEM of five technical repeats.

3 Discussion

Biosynthesis of plant secondary metabolites is regulated through highly complex regulatory networks, mediating either an increase or a decrease in the accumulation of

secondary metabolites. Therefore, manipulation of the components of these regulatory networks can provide a great opportunity to get insights into the signal transduction toward secondary metabolite production in plants. From a previous study (Pollier, 2011), we identified several regulatory genes potentially involved in the regulation of secondary metabolite biosynthesis in the model legume *M. truncatula*. The function of the selected genes was assessed by a gain- and loss-of function approach in transgenic hairy root cultures. Phenotypic characterization of transgenic hairy roots pinpointed two potential regulatory genes involved in the regulation of secondary metabolism, *Mt148* (this chapter) and *Mt061* (see chapter 5).

Mt148 encodes for a Ccr4-associated factor1 (Caf1) protein that is a component of a large regulatory complex involved in several aspects of gene regulation. Diverse studies have addressed a role of the Ccr4–Not complex subunits in the regulation of gene expression in *S. cerevisiae* and other eukaryotes. Many cellular functions have been assigned to the Ccr4–Not complex including regulation of transcription initiation, mRNA degradation, protein ubiquitination and RNA Polymerase II-dependent transcript elongation (Tucker et al., 2001; Denis and Chen, 2003; Panasenko et al., 2006; Kruk et al., 2011; Collart and Panasenko, 2012). Ccr4 and Caf1 have been well recognized as regulatory elements of this complex and are involved in post-transcriptional mRNA deadenylation. In this study, transcript profiling of *Mt148* overexpressing hairy roots showed that the overexpression of *Mt148* resulted in the transcriptional repression of some genes involved in primary and secondary metabolite biosynthesis pathways (Figs. 5 and 7). Based on these results we speculated that *Mt148* may be involved in the transcriptional repression of plant secondary metabolite biosynthesis genes via an mRNA deadenylation process. However, after applying RNA synthesis inhibitors, cordycepin and actinomycin D, no significant differences were observed in the stability of *BAS* and *Mt148* transcripts between Mt148^{OE} and control lines (Figs. 8 and 9). Hence, the results obtained from the RNA degradation assay did not support this hypothesis. Although previous studies have shown that both Ccr4 and Caf1 show deadenylase activity *in vitro,* the deadenylase activity of Caf1 *in vivo* is still matter of debate (Tucker et al., 2001; Tucker et al., 2002). The advantage of having two deadenylases in the Ccr4-Not complex is not clear, and it seems that there are species specific differences in the relative contribution of these two subunits to the deadenylation process. In yeast, it is suggested that Caf1 is probably involved in other processes including maintaining the

integrity of the complex or recruiting Ccr4-Not to mRNAs by binding sequence specific RNA binding proteins rather than to have deadenylase activity (Collart, 2003; Goldstrohm et al., 2006; Goldstrohm et al., 2007). Here also, it seems unlikely that the Caf1 protein encoded by *Mt148* possess deadenylase activity.

Unlike yeast that contains only one Ccr4-Caf1, the *CAF1* family spreads out widely in angiosperm genomes. For example, there are four CAF1s in grape (*Vitis vinifera*), twelve in sorghum (*Sorghum bicolor*), sixteen in rice (*Oryza sativa*), nineteen in poplar (*Populus trichocarpa*), eleven in Arabidopsis and eight in *M. truncatula*. Among the eleven closely related sequences characterized in *Arabidopsis* (Walley et al., 2010), only two hormone and stress inducible *CAF1s*, i.e. *AtCAF1a* and *AtCAF1b* could partially complement the growth defect of the yeast *caf1* deletion strain (Liang et al., 2009). Therefore, the diversity in the *Ccr4* and *Caf1* functions can rise from the presence of different versions of *Ccr4* and *Caf1* in higher eukaryotes that perform specialized functions. This is further supported by the results obtained from the yeast complementation assay, which revealed that *Mt148* could not complement the phenotypic defect of the yeast *caf1* mutant. On the contrary, it induced a growth defect in intact (wild type) yeast under normal growth conditions (Fig. 10), that indicates that *Mt148* encodes a Caf1 protein which functionally differs from yeast Caf1.

Finally, metabolite profiling of Mt148^{OE} hairy roots did not reveal any differences in accumulation of metabolites between Mt148^{OE} lines and control lines (Fig. 11), neither in MeJA-elicited nor in non-elicited hairy roots. Altogether the results from this study do not support a clear role for *Mt148* in the regulation of secondary metabolism in *M. truncatula* hairy roots yet.

4 Material and methods

4.1 Generation of DNA constructs

To obtain FL-ORFs for cloning of the selected *M. truncatula* genes, the cDNA-AFLP tag sequences were used for BLASTn searches against the *Medicago truncatula* Gene Index database (http://compbio.dfci.harvard.edu/tgi/). The FL-ORF consensus sequences were PCR-amplified and by Gateway™ recombination cloned into the entry vector pDONR221. To obtain entry clones with stop codons, Gateway primers were designed according to Underwood (Underwood et al., 2006). All entry constructs were sequence-verified. For the overexpression experiments, Gateway

recombination was carried out with the pK7WG2D binary vector (Karimi et al., 2002), and the resulting clones transformed into the *Agrobacterium rhizogenes* strain LBA 9402/12 for generation of hairy roots. For hpRNAi, the cDNA-AFLP fragments were PCR-amplified and by Gateway™ recombination cloned into the binary vector pK7GWIWG2D(II) (Karimi et al., 2002). The resulting expression clones were transformed into *A. rhizogenes*. For the yeast complementation assays, the pAG426GPD vector (Alberti et al., 2007) was used as the destination vector for the *Mt148* gene.

4.2 Generation and cultivation of *Agrobacterium rhizogenes*-mediated transgenic hairy roots

Protocol for *A. rhizogenes*-mediated transformation of *M. truncatula* (ecotype Jemalong J5) hairy roots was adapted from Boisson-Dernier *et al.* (Boisson-Dernier et al., 2001), with modifications. After 5 minutes treatment with H_2SO_4, seeds were sterilized with 12% sodium hypochlorite for 2 minutes and rinsed in sterile water. Subsequently, seeds were treated with 1 µM 6- benzylaminopurine for 3 h and then allowed to germinate on wet, sterile Whatman paper at room temperature in the dark. After 2 days of germination, the seedlings were wounded by cutting approximately 2 mm from the root tip of the radicle. The wounded seedlings were infected with *A. rhizogenes* harboring binary vectors grown on solid YEB medium. Incubated seedling were placed on slanted agar plates containing Murashige and Skoog (MS) medium (pH 5.8) supplemented with vitamins (Duchefa). The plates were sealed with micropore tape, placed vertically, and cocultivation was allowed for 10 days under a 16 h/8 h light/dark regime at 22°C. After 10 days, the plants were transferred to new plates, containing 100 mg/L cefotaxime to prevent *Agrobacterium* growth and incubated under identical conditions. Ten days later, plants were screened for transgenic hairy roots characterizing by GFP fluorescent. Transgenic hairy roots were excised from the plants and transferred to liquid MS medium (pH 5.8) supplemented with vitamins, 1% sucrose (w/v), and 100 mg/L cefotaxime. The hairy roots were grown for 7 days in the dark at 24°C and shaking at 300 rpm. Subsequently, the roots were transferred to Petri dishes containing solid MS medium (pH 5.8) supplemented with 1% sucrose (w/v) and 100 mg/L cefotaxime, grown in the dark at 24°C, and subcultured every 3 weeks. After 3 weeks, young hairy root tissue was subcultured to solid medium without antibiotics. For maintenance, the newly grown *Agrobacterium*-

free hairy root cultures were subcultured every 3 weeks onto fresh plates. Hairy roots for metabolic profiling were grown for 21 days in liquid MS medium (pH 5.8) supplemented with vitamins and 1% (w/v) sucrose.

4.3 Transcript profiling

For quantitative Real Time PCR (qRT-PCR), total RNA were isolated with the RNeasy mini kit (Qiagen) according to manufacturer instruction. Quality control and quantification were performed with a Nanodrop spectrometer (Isogen, Hackensack, NJ). cDNA were prepared using SuperScript™ II Reverse Transcriptase (Invitrogen). Primers were designed with Beacon Designer version 4.0 (Premier Biosoft International, Palo Alto, CA, USA). QRT-PCR was quantified on a Lightcycler 480 (Roche) and SYBR Green was used for detection. All reactions were done in a 5 µl volume in triplicate in a 384-multiwell plates allow determination of mean of standard error of Ct. Data were normalized against *40S ribosomal protein S8* (*40S*) (TC160725, MGI) and *translation elongation factor 1α* (*ELF1α*) (TC148782, MGI). Reactions were done in triplicate and qBase was used for the relative quantification with multiple reference genes (Hellemans et al., 2007).

4.4 Metabolite extractions

For metabolite profiling, five biological repeats of three independent transgenic lines per transgene construct of *M. truncatula* hairy roots were harvested and washed with purified water under vacuum filtration. The roots were immediately frozen and ground in liquid nitrogen. One milliliter MeOH, was added to 400 mg of the ground material and incubated at room temperature for 10 min. Samples were then centrifuged for 10 min at 20800xg. Under vacuum, 500 µL of the supernatant was evaporated to dryness. The residue was dissolved in 600 µL of H_2O/cyclohexane (2:1, v/v), centrifuged (10 min at 20,800xg), and 200 µL of the aqueous phase was taken for analysis.

4.5 LC ESI FT-ICR MS

For reversed-phase LC, an Acquity UPLC BEH C18 column (150 x 2.1 mm, 1.7 µm; Waters, Milford, MA) was serially coupled to an Acquity UPLC BEH C18 column (100 x 2.1 mm, 1.7 µm) and mounted on an ultra-high-performance LC system consisting of a Accela pump (Thermo Electron Corporation, Waltham, MA, USA) and Accela autosampler (Thermo Electron Corporation). The Accela LC system was

hyphenated to a LTQ FT Ultra (Thermo Electron Corporation) via an electrospray ionization source. The following gradient was run using water:MeCN (99:1, v:v) acidified with 0.1% (v:v) HOAc (solvent A) and MeCN:water (99:1, v:v) acidified with 0.1% (v:v) HOAc (solvent B): time 0 min, 5% B; 30 min, 55% B; 35 min, 100% B. The loop size, flow, and column temperature were 25 µL, 300 µL/min and 80°C, respectively. Full loop injection was applied. Negative ionization was obtained with the following parameter values: capillary temperature 150°C, sheath gas 25 (arbitrary units), aux. gas 3 (arbitrary units), and spray voltage 4.5 kV. Full FT-MS spectra between m/z 120– 1400 were recorded at a resolution of 100,000. For identification, full MS spectra were interchanged with a dependent MS^2 scan event in which the most abundant ion in the previous full MS scan was fragmented, and two dependent MS^3 scan events in which the two most abundant daughter ions were fragmented. The collision energy was set at 35%. Elucidation of the MS^n spectra was according to Pollier et al (Pollier et al., 2011) for the saponins and Morreel et al. (Morreel et al., 2006) for the flavonoids. The resulting chromatograms were integrated and aligned with the XCMS package (Smith et al., 2006) in R version 2.6.1. with the following parameter values: xcmsSet(fwhm=8, max=300, snthresh=5, mzdiff=0.5), group (bw=8, max=300), rector (method=loess, family=symmetric). A second grouping was done with the same parameter values. Due to in- source fragmentation, multiple m/z peaks for each compound were often observed. The number of compounds was estimated with "peak groups" consisting of m/z peaks with the same retention time (window, x s) that were correlated (Pearson; threshold, 0.85) across all control samples: reproduced from (Pollier, 2011).

4.6 Phylogenetic analysis

Protein sequences were aligned with ClustalW and the phylogenetic tree was produced using MEGA 5.0.1 software (Tamura et al., 2011), by the Neighbor-Joining method, and bootstrapping was done with 10,000 replicates. The evolutionary distances were computed with the Poisson correction method, and all positions containing gaps and missing data were eliminated from the data set (complete deletion option).

4.7 Measurement of RNA Degradation

The RNA degradation assay was performed with three biological repeats of two and three independent transgenic hairy roots in cordycepin and actinomycin D inhibition experiments, respectively. Mt148^{OE} and control hairy roots were cultured in MS liquid medium and 21 days after subculturing into fresh medium, cordycepin (Sigma) (150 µg/mL, dissolved in DMSO) or actinomycin D (Sigma) (100 µg/mL, dissolved in DMSO) or an equivalent amount of DMSO as control was added to the media. After thirty minutes, MeJA (50 µM, Dissolved in DMSO) or an equivalent amount of DMSO was added. Samples were taken 0, 0.5, 1, 3, and 6 h after addition of cordycepin or actinomycin D and analyzed by qRT-PCR to check RNA degradation rates. Experiments with actinomycin D were conducted in a darkroom.

4.8 Yeast complementation

Yeast strains KY803 (*MATa leu2-PET56 trp1-Δ1 ura3-52 gal2 gcn4-Δ1),* KY803-c1 *(MATa leu2-PET56 trp1-Δ1 ura3-52 gal2 gcn4-Δ1 caf1::LEU2*), BY4741 (*BY4741; Mat a; his3D1; leu2D0; met15D0; ura3D0*) and Y07123 (*BY4741; Mat a; his3D1; leu2D0; met15D0; ura3D0; YNR052c::kanMX4*) were transformed with pAG426GPD-*Mt148*. Precultures were grown for 18-20h in 30°C at 250 rpm in 5 ml synthetic defined (SD) medium lacking *URA* (SD-U) or YPD medium. Then the cultures were grown on SD-U plates supplemented with or without 5 mM caffeine at 28°C.

Acknowledgments

Azra Gholami and Tessa Moses are indebted to the Iranian Ministry for Health and Medical Education and the VIB International PhD Program for predoctoral fellowships respectively.

References

Abe, I., Rohmer, M., and Prestwich, G.D. (1993). Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. Chemical Reviews **93,** 2189-2206.

Alberti, S., Gitler, A.D., and Lindquist, S. (2007). A suite of Gateway$^{(R)}$ cloning vectors for high-throughput genetic analysis in *Saccharomyces cerevisiae*. Yeast **24,** 913-919.

Bianchin, C., Mauxion, F., Sentis, S., Seraphin, B., and Corbo, L. (2005). Conservation of the deadenylase activity of proteins of the Caf1 family in human. RNA-a Publication of the RNA Society **11,** 487-494.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C., and Barker, D.G. (2001). *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the

study of nitrogen-fixing and endomycorrhizal symbiotic associations. Molecular Plant-Microbe Interactions **14,** 695-700.

Broeckling, C.D., Huhman, D.V., Farag, M.A., Smith, J.T., May, G.D., Mendes, P., Dixon, R.A., and Sumner, L.W. (2005). Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. Journal of Experimental Botany **56,** 323-336.

Carland, F.M., Fujioka, S., Takatsuto, S., Yoshida, S., and Nelson, T. (2002). The identification of CVP1 reveals a role for sterols in vascular patterning. Plant Cell **14,** 2045-2058.

Chen, J.J., Chiang, Y.C., and Denis, C.L. (2002). CCR4, a 3 '-5 ' poly(A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. EMBO Journal **21,** 1414-1426.

Chiba, Y., and Green, P. (2009). mRNA degradation machinery in plants. Journal of Plant Biology **52,** 114-124.

Collart, M.A. (2003). Global control of gene expression in yeast by the Ccr4-Not complex. In Gene, pp. 1-16.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4-Not complex. Gene **492,** 42-53.

Denis, C.L., and Chen, J.J. (2003). The CCR4-NOT complex plays diverse roles in mRNA metabolism. Progress in Nucleic Acid Research and Molecular Biology **73,** 221-250.

Dixon, R.A., and Sumner, L.W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. Plant Physiology **131,** 878-885.

Goldstrohm, A.C., Hook, B.A., Seay, D.J., and Wickens, M. (2006). PUF proteins bind Pop2p to regulate messenger RNAs. Nature Structural & Molecular Biology **13,** 533-539.

Goldstrohm, A.C., Seay, D.J., Hook, B.A., and Wickens, M. (2007). PUF protein-mediated deadenylation is catalyzed by Ccr4p. Journal of Biological Chemistry **282,** 109-114.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biology **8,** 19-33.

Holtorf, H., Schob, H., Kunz, C., Waldvogel, R., and Meins, F. (1999). Stochastic and nonstochastic post-transcriptional silencing of chitinase and beta-1,3-glucanase genes involves increased RNA turnover - Possible role for ribosome-independent RNA degradation. Plant Cell **11,** 471-483.

Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY$^{((TM))}$ vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science **7,** 193-195.

Kruk, J.A., Dutta, A., Fu, J.H., Gilmour, D.S., and Reese, J.C. (2011). The multifunctional Ccr4-Not complex directly promotes transcription elongation. Genes & Development **25,** 581-593.

Liang, W.X., Li, C.B., Liu, F., Jiang, H.L., Li, S.Y., Sun, J.Q., Wu, X.Y., and Li, C.Y. (2009). The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. Cell Research **19,** 307-316.

Miller, J.E., and Reese, J.C. (2012). Ccr4-Not complex: the control freak of eukaryotic cells. Critical Reviews in Biochemistry and Molecular Biology **47,** 315-333.

Morikawa, T., Mizutani, M., Aoki, N., Watanabe, B., Saga, H., Saito, S., Oikawa, A., Suzuki, H., Sakurai, N., Shibata, D., Wadano, A., Sakata, K., and Ohta, D. (2006). Cytochrome P450CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. Plant Cell **18,** 1008-1022.

Morreel, K., Goeminne, G., Storme, V., Sterck, L., Ralph, J., Coppieters, W., Breyne, P., Steenackers, M., Georges, M., Messens, E., and Boerjan, W. (2006). Genetical metabolomics of flavonoid biosynthesis in Populus: a case study. Plant Journal **47,** 224-237.

Morris, J.Z., Hong, A., Lilly, M.A., and Lehmann, R. (2005). twin, a CCR4 homolog, regulates cyclin poly(A) tail length to permit *Drosophila* oogenesis. Development **132,** 1165-1174.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y.H., Sumner, L.W., and Dixon, R.A. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell **22,** 850-866.

Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M.A. (2006). The yeast Ccr4-not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. Journal of Biological Chemistry **281,** 31389-31398.

Pollier, J. (2011). Combinatorial Biochemistry of Triterpene Saponins in Plants (Faculty of Sciences VIB Department of Plant Systems Biology Ghent University).

Pollier, J., Morreel, K., Geelen, D., and Goossens, A. (2011). Metabolite profiling of triterpene saponins in *Medicago truncatula* hairy roots by liquid chromatography fourier transform ion cyclotron resonance mass spectrometry. Journal of Natural Products **74,** 1462-1476.

Sarowar, S., Oh, H.W., Cho, H.S., Baek, K.H., Seong, E.S., Joung, Y.H., Choi, G.J., Lee, S., and Choi, D. (2007). *Capsicum annuum* CCR4-associated factor CaCAF1 is necessary for plant development and defence response. Plant Journal **51,** 792-802.

Schrick, K., Fujioka, S., Takatsuto, S., Stierhof, Y.D., Stransky, H., Yoshida, S., and Jurgens, G. (2004). A link between sterol biosynthesis, the cell wall, and cellulose in Arabidopsis. Plant Journal **38,** 227-243.

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

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution **10,** 2731–2739.

Temme, C., Zaessinger, S., Meyer, S., Simonelig, M., and Wahle, E. (2004). A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in *Drosophila*. EMBO Journal **23,** 2862-2871.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*. EMBO Journal **21,** 1427-1436.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J.J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. Cell **104,** 377-386.

Underwood, B.A., Vanderhaeghen, R., Whitford, R., Town, C.D., and Hilson, P. (2006). Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations. Plant Biotechnology Journal **4,** 317-324.

Walley, J.W., Kelley, D.R., Nestorova, G., Hirschberg, D.L., and Dehesh, K. (2010). Arabidopsis deadenylases AtCAF1a and AtCAF1b play overlapping and distinct roles in mediating environmental stress responses. Plant Physiology **152,** 866-875.

Wolfner, M., Yep, D., Messenguy, F., and Fink, G.R. (1975). Integration of amino acid biosynthesis into the cell cycle of *Saccharomyces cerevisiae*. Journal of Molecular Biology **96,** 273-290.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W.M., Zhong, Z.P., Chen, C.Y.A., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nature Structural & Molecular Biology **12,** 1054-1063.

Zuo, Y., and Deutscher, M.P. (2001). Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Research **29,** 1017-1026.

5

The MYB transcription factor Mt061 modulates defense response mechanisms in *Medicago truncatula*

Azra Gholami, Nathan De Geyter, Jacob Pollier, Tessa Moses, Alain Goossens

Author contribution:

Molecular cloning, generation of M. truncatula hairy roots, qRT-PCR transcript profiling, MeJA elicitation, data interpretation.

Abstract:

Plants produce green leaf volatiles (GLVs), consisting of six-carbon (C6) aldehydes, alcohols and their esters, as a defense response to insect or pathogen damage. GLV formation is thought to be regulated at the step of lipid hydrolysis, which provides free fatty acids to the pathway. Although the pathways for GLV biosynthesis have been characterized, how their production and emission are regulated is mostly unknown. By genome-wide cDNA-AFLP transcript profiling of MeJA elicited cell cultures, we identified Mt061, which encodes an R2R3-type MYB family transcription factor, as a candidate regulator of GLV biosynthesis in Medicago truncatula. In contrast to control lines, Mt061 overexpressing hairy roots produce and emit GLV compounds. Enhanced transcript levels of several defense genes such as trypsin protease inhibitor (TPI) and phenolic biosynthesis genes were detected in Mt061-overexpressing hairy roots. In addition, higher levels of phytoalexin compounds were detected in transgenic lines as compared to control lines, suggesting that Mt061 coordinates the regulation of direct and indirect defense response processes.

1 Introduction

Plants emit a wide array of volatile organic compounds such as jasmonates (JAs), green leaf volatiles (GLVs) and isoprenoids in response to various biotic and abiotic stresses (Dudareva et al., 2006). The volatiles released from leaves upon herbivore attack help plants to hinder herbivores or attract their predators (Arimura et al., 2011). Moreover, volatiles act as signal molecules in plants to induce genes associated with the defense response (Bate and Rothstein, 1998; Kishimoto et al., 2005).

Plant volatiles (PVs) are low-molecular-weight compounds (below 300 Da) that mostly emanate from three major biosynthetic pathways: the terpenoid pathway, the benzenoid pathway, and fatty acid metabolism (Dudareva et al., 2006). However, small quantities of amino acid- derived volatiles can also be present in the blends emitted from flowers and fruits. Various forms of enzymatic modifications such as hydroxylation, acetylation, and methylation increase the volatility of these compounds and cause a remarkable diversity of the released volatiles (Dudareva et al., 2006).

The largest class of PVs is composed of terpenoids derived from the universal C5 precursor isopentenyl diphosphate (IPP) and its allylic isomer, dimethylallyl diphosphate (DMAPP). Terpenoids are synthesized via two independent pathways in separate compartments of the plant cell. In the cytosol, IPP biosynthesis initiates with the condensation of three molecules of acetyl-CoA through the cytoplasmic

mevalonic acid (MVA) pathway (Newman and Chappell, 1999). In plastids, IPP is generated from pyruvate and glyceraldehyde 3- phosphate via the 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway (Rodriguez-Concepcion and Boronat, 2002).

Phenylpropanoids and benzenoids are derived from *L*-phenylalanine and are the second largest class of PVs. The first step of benzenoid biosynthesis is catalyzed by *L*-phenylalanine ammonia-lyase that converts *L*-phenylalanine into *trans*-cinnamic acid. The next steps are shared with the lignin and lignan biosynthetic pathway up to the formation of phenylpropenol (monolignol) which results in the biosynthesis of an array of hydroxycinnamic acids, aldehydes and alcohols derived from *trans*-cinnamic acid through a series of hydroxylation and methylation reactions (Verdonk et al., 2003; Schuurink et al., 2006). Shortening of the three-carbon side chain of hydroxycinnamates to one carbon also leads to aromatic building blocks such as benzoic acid and benzaldehyde (Boatright et al., 2004). In addition, phenylalanine is reduced at the C9 position to phenylacetaldehyde and phenylethylalcohol that are also benzenoid volatiles (Boatright et al., 2004).

The third group of PVs consists of fatty acid- derived volatiles such as *trans*-2-hexenal, *cis*-3-hexenol and methyl jasmonate (MeJA) that are commonly named the oxylipins. Biosynthesis of oxylipins is triggered by cell membrane disruption that leads to the release of the C18 unsaturated fatty acids α-linolenic acid and linoleic acid from the chloroplast membrane by the enzymatic activity of lipases. Then, subsequent actions of the enzymes lipoxygenase (LOX), hydroperoxide lyase (HPL) and/or allene oxide synthase (AOS) convert linoleic acid or linolenic acid into short chain volatiles with aldehyde and ketone moieties, that often serve as precursors of oxylipins (Dudareva et al., 2006)

LOXs are divided into two major classes: 9-LOXs and 13-LOXs, introducing an oxygen group at the C-9 or C-13 position to produce 9-hydroperoxy linolenic acid (9HPOT) and 13-hydroperoxy linolenic acid (13HPOT), respectively (Liavonchanka and Feussner, 2006). The resulting HPOT derivatives can be further metabolized by either AOS or HPL, which represent two branches in the lipoxygenase pathway. AOS converts 13HPOT to an unstable allene oxide intermediate, 12,13-epoxy octadecatrienoic acid. Then, a cascade of subsequent enzymatic reactions leads to the formation of jasmonic acid (JA) which in turn can be converted into the volatile ester, MeJA, by jasmonate *O*-methyltransferase (JMT) (Seo et al., 2001; Song et al., 2005). In another branch, HPL catalyzes the first committed step toward GLV biosynthesis.

This includes the oxidative cleavage of 9HPOT and 13HPOT by 9HPL and 13HPL, respectively which results in the formation of a short chain C6-compound ((*Z*)-3-hexenal) and a C12- compound (12-oxo-(*Z*)-9-dodecenoic acid), commonly referred to as GLVs. Moreover, 9HPOT can be cleaved by the 9/13HPL enzyme that can act on both 13- and 9-hydroperoxides with almost the same efficiency to form C9-aldehydes and C9-oxo acids (Matsui, 2006). (*Z*)-3-hexenal is the first C6-GLV compound formed after tissue damage (Matsui et al., 2000a) and is then converted to other GLVs like (*E*)-2-hexenal (leaf aldehyde), (*Z*)-3-hexenol (leaf alcohol) and (*Z*)-3-hexenyl acetate (leaf ester) (Shiojiri et al., 2006). Figure 1 shows an overview of the GLV and MeJA biosynthetic pathways.

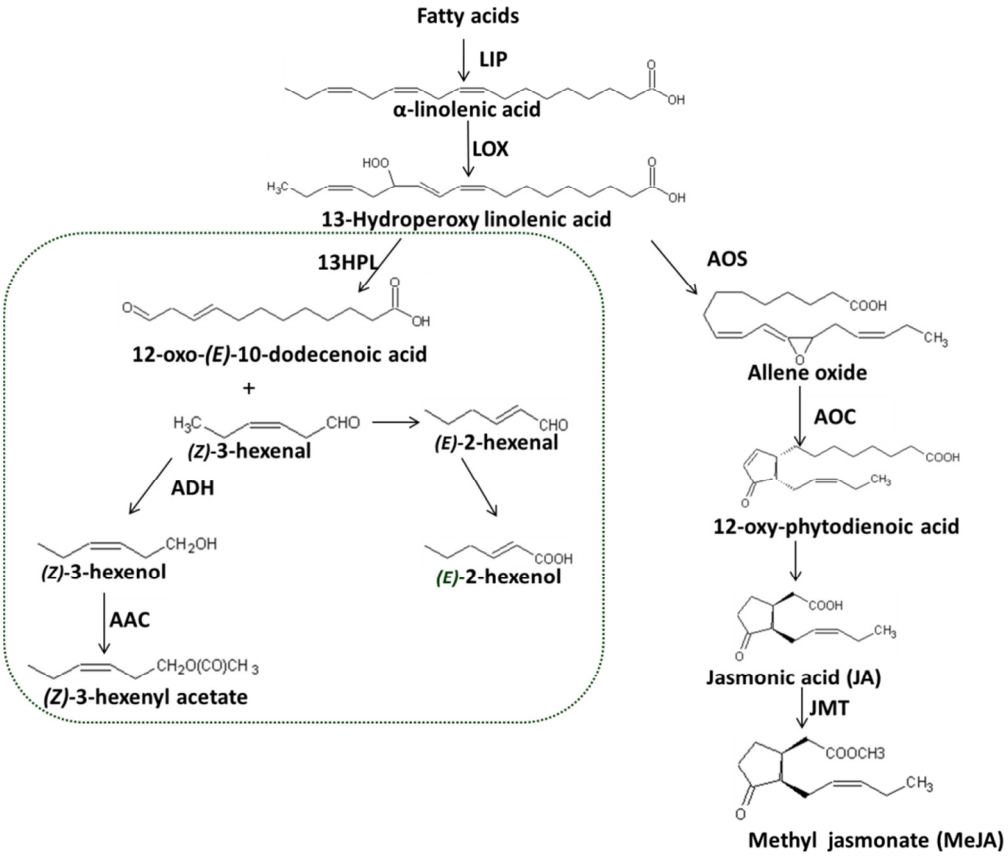


Figure 1. The biosynthetic pathways of GLVs and MeJA from C18 fatty acids compounds. LIP, lipase; LOX, lipoxygenase; 13-HPL,13-hydroperoxide lyase; ADH, alcohol dehydrogenase; AAC, alcohol acyl-transferase; AOS, allene oxide synthase; AOC, allene oxide cyclase; JMT, jasmonate O-methyltransferase.

Despite significant progress in the characterization of the genes and enzymes involved in the biosynthesis of plant volatiles and the approaches for manipulating the volatile spectrum in plants (Xie et al., 2004; Shao et al., 2007; Arimura et al., 2011), very little is known about the regulation of the plant volatiles at the molecular level. HAHB4, a

member of the sunflower (*Helianthus annuus*) homeodomain-leucine zipper (HD-ZIP) subfamily, was identified as a positive regulator of the synthesis of JAs and GLVs. HAHB4 upregulates the transcript levels of several genes involved in JA and GLV biosynthesis and *HAHB4* expression itself is stimulated by JAs (Manavella et al., 2008). Recently, a JAZ protein, named NaJAZh has been characterized in native tobacco (*Nicotiana attenuata*) and shown to be involved in the regulation of GLV biosynthesis (Oh et al., 2012).

In this study, we show that overexpression of the *Mt061* gene, encoding a MYB transcription factor (TF), in *M. truncatula* hairy roots results in the production and emission of GLVs. Furthermore, the accumulation of higher levels of phytoalexin compounds as well as upregulation of several phenolic biosynthesis and defense genes in *Mt061*-overexpressing hairy roots suggests that *Mt061* participates in the regulation of defense response processes in *M. truncatula*.

2 Results

2.1 Identification of *Mt061* as a MeJA-responsive regulatory gene

Based on the data obtained from the genome-wide cDNA-AFLP analysis of MeJA-elicited *M. truncatula* cells (Pollier, 2011), some genes with a putative regulatory role in secondary metabolite biosynthesis genes were selected for further analysis. Then, using cDNA obtained from the MeJA-elicited *M. truncatula* cells and hairy root cultures (see chapter 4), the transcript levels of the candidate genes were compared with those of control lines. The genes with elevated transcript levels between 30 min to 2 h after elicitation, which is prior or concomitant to the onset of the secondary metabolite biosynthesis genes, were selected (see chapter 4). Based on this rationale, *Mt061* was also selected (see chapter 4). *Mt061* encodes a protein with two SANT/MYB DNA-binding domains indicative of MYB-like TFs and has high similarity to proteins of the R2R3-type MYB family (Stracke et al., 2001). The transcript levels of *Mt061* were shown to increase between 30 min to 2 h after elicitation in MeJA-elicited cell cultures (Fig. 2A), while they did not respond to MeJA elicitation in hairy roots (Fig 2B). Subsequently, a reverse genetics approach was applied to interrogate the function of this potentially regulatory gene. Gain- of-function was pursued via expression of the *Mt061* full-length open reading frame (FL-ORF) under a constitutive CaMV35S promoter in *M. truncatula* hairy roots.

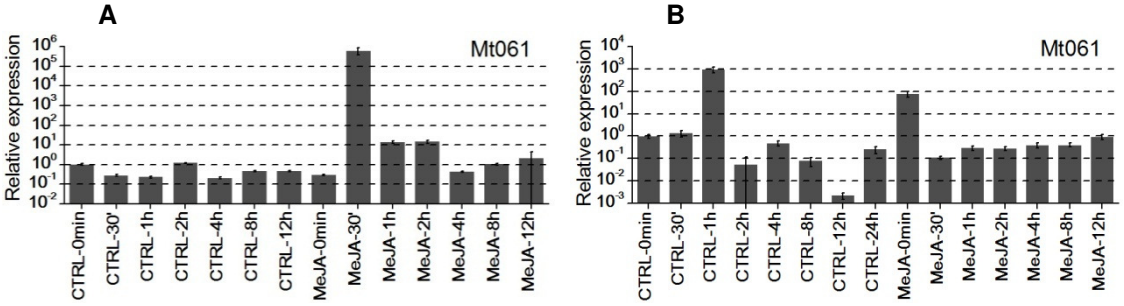


Figure 2. Expression analyses of *Mt061* in MeJA-treated *M. truncatula* (A) cell cultures and (B) hairy roots, using qRT-PCR. Treatments and time-points are indicated at the bottom of the graphs. Y-axis represent relative expression ratio as compared to the mock-treated line at time 0. Error bars represent SEM of three technical repeats.

When transgenic hairy roots were established and grown on the plates, the *Mt061*-overexpressing hairy roots (Mt061^{OE}) produced and emitted a penetrating scent of plant volatiles, perceivable by the human nose. Therefore we decided to explore the role(s) of this putative regulator in plant defense response and secondary metabolism more thoroughly.

2.2 Overexpression of *Mt061* in *M. truncatula* hairy roots is associated with GLV emission

Mt061^{OE} lines were used for volatile analysis by GC-MS. The GC chromatogram of Mt061^{OE} showed the presence of a new peak at 27.480 min (Fig. 3A) with electron ionization (EI) pattern corresponding to hexanal (Fig. 3C). This peak was absent from the control lines (Fig. 3B).

Hexanal and other six-carbon (C6-) volatiles, including the aldehydes *trans*- 2-hexenal and *cis*-3-hexenal, as well as their corresponding alcohols or esters, are well-known GLV compounds that are released as the earliest volatiles from damaged or wounded plant tissue. Therefore, we decided to explore GLV metabolism in Mt061^{OE} lines more thoroughly.

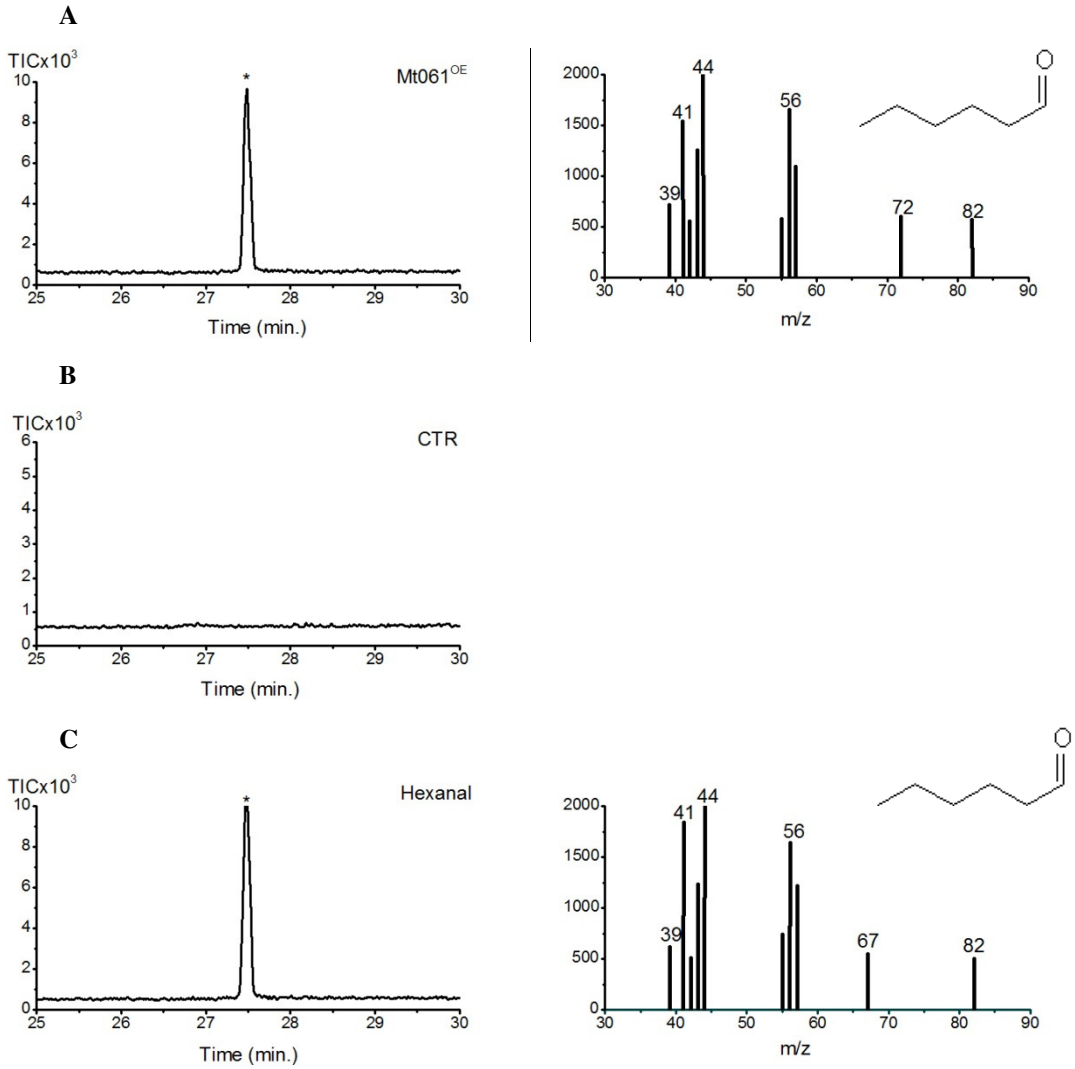


Figure 3. GC chromatograms corresponding to, (A) volatile extraction from the Mt061^{OE} line, (B) volatile extraction from control lines, (C) hexanal standard. Chromatograms are showing hexanal at 27.48 min. The X- and Y-axis of the GC chromatograms correspond to time in minutes and total ion current, respectively. Right panels show the mass spectra extracted from the indicated (*) peaks. The X- and Y-axis of the mass spectra correspond to mass-to-charge ratio (m/z) values and the signal intensity (abundance) for each of the detected fragments, respectively.

2.3 *Mt061* is coregulated with some genes of the lipoxygenase pathway

To verify the effect of *Mt061* overexpression on expression of GLV biosynthesis genes, we performed transcript profiling on Mt061^{OE} hairy roots with qRT-PCR. First, using data obtained from genome-wide cDNA-AFLP transcript profiling on MeJA-elicited cell cultures (Pollier, 2011), some major genes involved in biosynthesis of oxylipins which were shown to be coregulated with *Mt061* gene were chosen, including *Lipase* (*LIP; Mt201*: TC177401), two *LOXs* (*LOX1*; *Mt079*: TC181880 and *LOX2*; TC178353), *13HPL* (*HPL1*; TC173537), *9/13HPL* (*HPL2*; *Mt173*: TC173832) and *AOS* (TC174824) (Fig. 4).

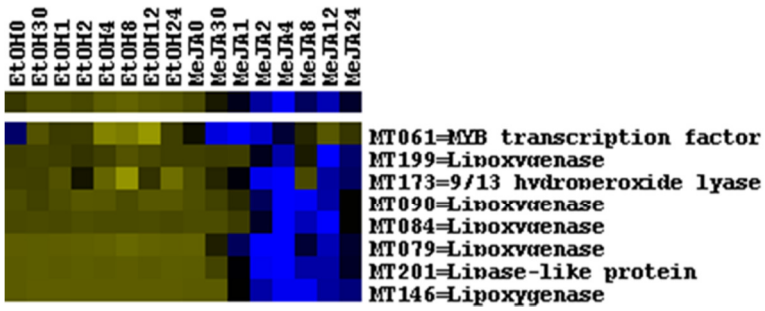


Figure 4. Average linkage hierarchical clustering of *M. truncatula* gene tags corresponding to the genes involved in oxylipins biosynthesis. Tags are identified by cDNA-AFLP analysis on MeJA-elicited *M. truncatula* cells (Pollier, 2011). Treatments and time points (h) are indicated at the top. Blue and yellow boxes reflect transcriptional activation and repression by methyl jasmonate (MeJA) relative to the average expression level, respectively.

This analysis revealed that *LIP*, and *LOX2* were upregulated in the Mt061^{OE} lines (Fig. 5), whereas the transcript levels of *HPLs* and *AOS* were not affected in Mt061^{OE} lines (Fig. 5).

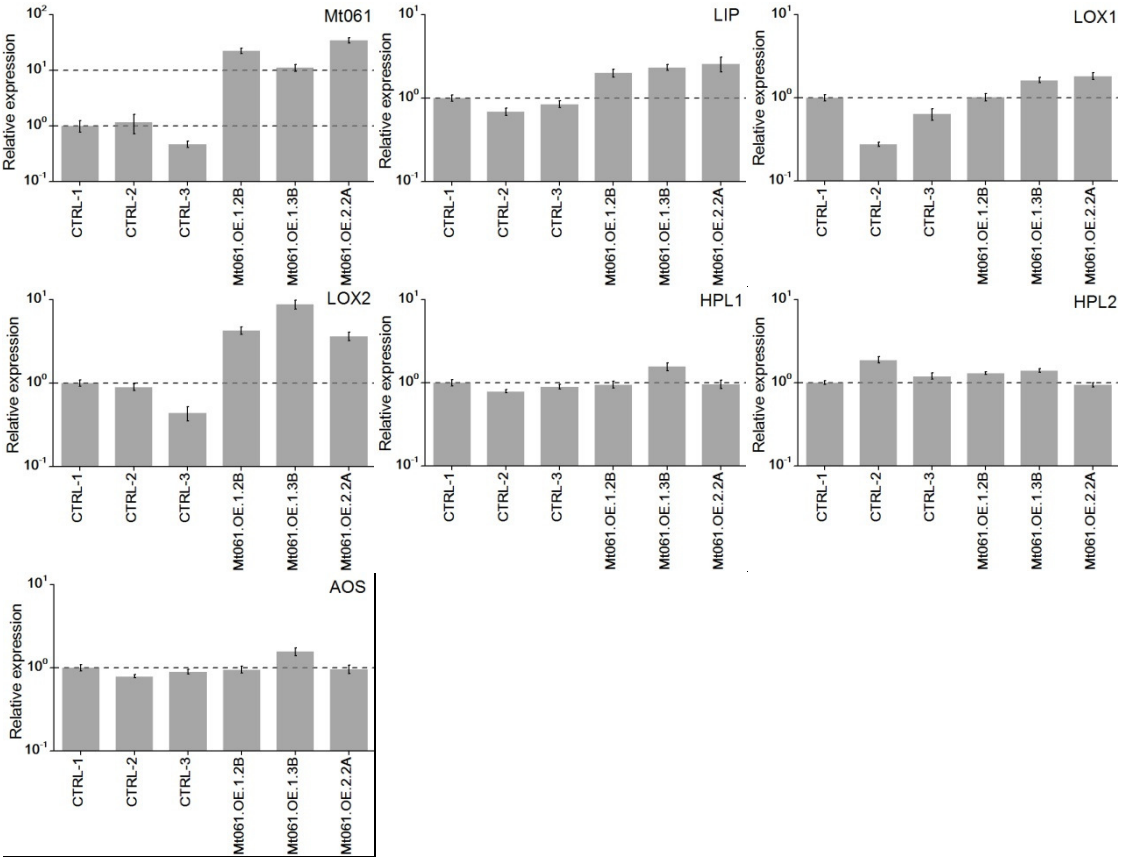


Figure 5. Expression analysis of GLV biosynthesis genes in Mt061^{OE} lines. CTRL, control lines; OE, Mt061^{OE} lines. *LIP*, *lipase*; *LOX*, *lipoxygenase*; *HPL*, *hydroperoxide lyase*; *AOS, allene oxide synthase*. Y-axis corresponds to relative expression ratio as compared to the control line 2. Error bars represent SEM of three technical repeats.

2.4 *Mt061* is not significantly induced under stress conditions in *M. truncatula* hairy roots

To further characterize the *Mt061* gene, we wanted to assess whether it might be associated with biotic and abiotic stress responses. Therefore, the quantitative changes in transcription of *Mt061* and GLV biosynthesis genes (*LIP*, *LOX2*, *HPL1* and *HPL2*) were assessed in *M. truncatula* hairy roots treated with NaCl (200mM) (Fig. 6) and MeJA (100 µM) (Fig. 7) over a 48 h time- course. This analysis showed that although some variations between mock treated and NaCl treated samples, neither Mt061 nor GLV biosynthesis genes were significantly induced by NaCl treatment. The only significant change was observed in transcription level of *HPL2* which was significantly upregulated ($p \leq 0.05$) after 24 hours (Fig. 6).

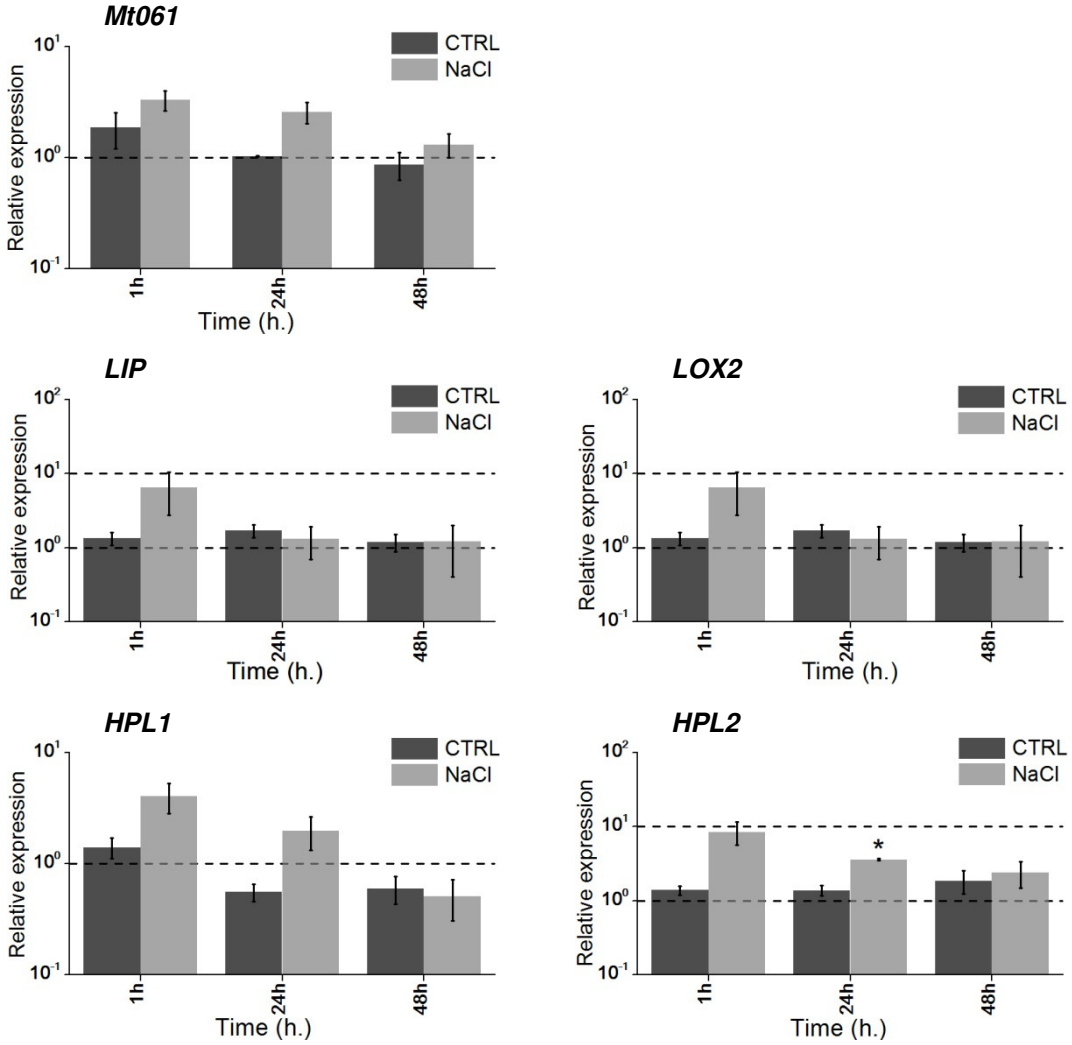


Figure 6. Expression patterns of *Mt061* and GLV biosynthesis genes *in M. truncatula* hairy roots treated with NaCl. *LIP, lipase; LOX2, lipoxygenase 2; HPL1* and *2, hydroperoxide lyase1 and 2*. Expression level at different time points is represented as a ratio relative to the mock-treated line 1 (CTRL-1) at time 1h, which is set to 1. Data and error bars represent Mean± SEM of three biological repeats. Asterisk indicates a significant difference on paired samples t-test with $p \leq 0.05$.

In contrast, MeJA elicitation resulted in the significant changes in the expression levels of *LIP*, *LOX* and *HPL2* genes with the *p value* of ≤0.05; ≤0.001 and ≤0.01 respectively. Whereas *HPL1* was not affected and *Mt061* was not induced with MeJA either, which is consistent with the results obtained in the earlier experiment (Fig. 2B).

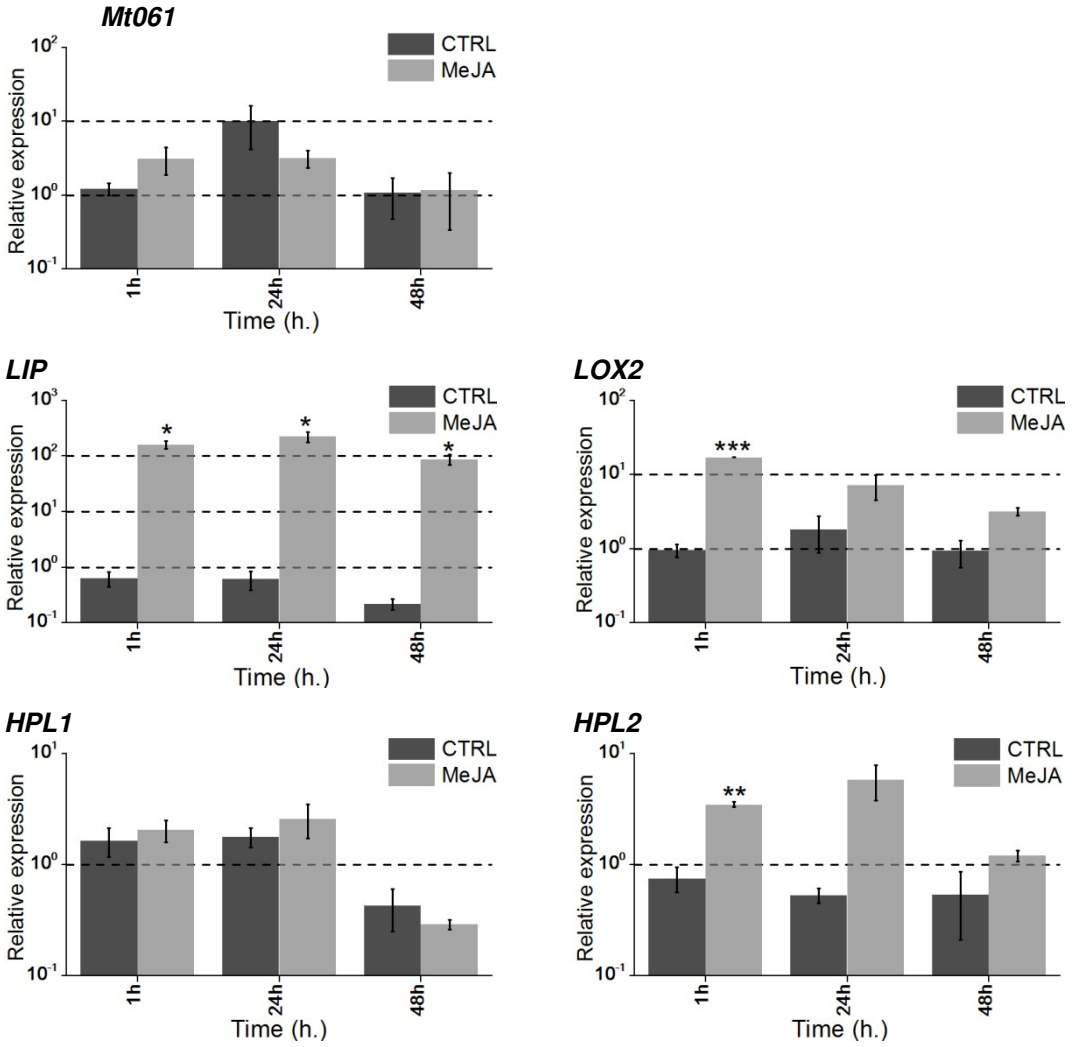


Figure 7. Expression patterns of *Mt061* and GLV biosynthesis genes in the presence of MeJA in *M. truncatula* hairy roots. *LIP*, *lipase*; *LOX2*, *lipoxygenase 2*; *HPL1* and *2*, *hydroperoxide lyase1* and *2*. Expression level at different time points is represented as a ratio relative to the mock-treated line 1 (CTRL-1) at time 1h, which is set to 1. Data and error bars represent Mean± SEM of three biological repeats. Asterisks indicate significant differences on paired samples t-test: * p≤0.05; ** p≤0.01; ***p≤0.001.

2.5 Isoflavonoid and pterocarpan accumulation are modulated by *Mt061* overexpression

Since *Mt061* and secondary metabolite biosynthesis genes were co-regulated in MeJA elicited cell cultures (Pollier, 2011), we explored the effect of *Mt061* overexpression on secondary metabolite profiles of *M. truncatula* hairy roots. Five biological repeats of three independent Mt061^{OE} lines were profiled with LC- ESI-FT-ICRMS and

compared with that of control lines. Comparative analyses of the root extracts of Mt061^{OE} and control lines yielded a total of 1045 *m/z* peaks corresponding to 90 differentially present peaks. Among the differentially present peaks, the peaks corresponding to some major known isoflavonoid compounds including malonyl-Hex-afrormosin, malonyl-Hex-formononetin and malonyl-Hex-biochanin A, as well as pterocarpan medicarpin 3-*O*-glucoside-6'-malonate were detected in higher abundance in Mt061^{OE} lines as compared to control lines (Fig. 8). Therefore, the results show that *Mt061* modulates isoflavonoid and pterocarpan accumulation in transgenic roots.

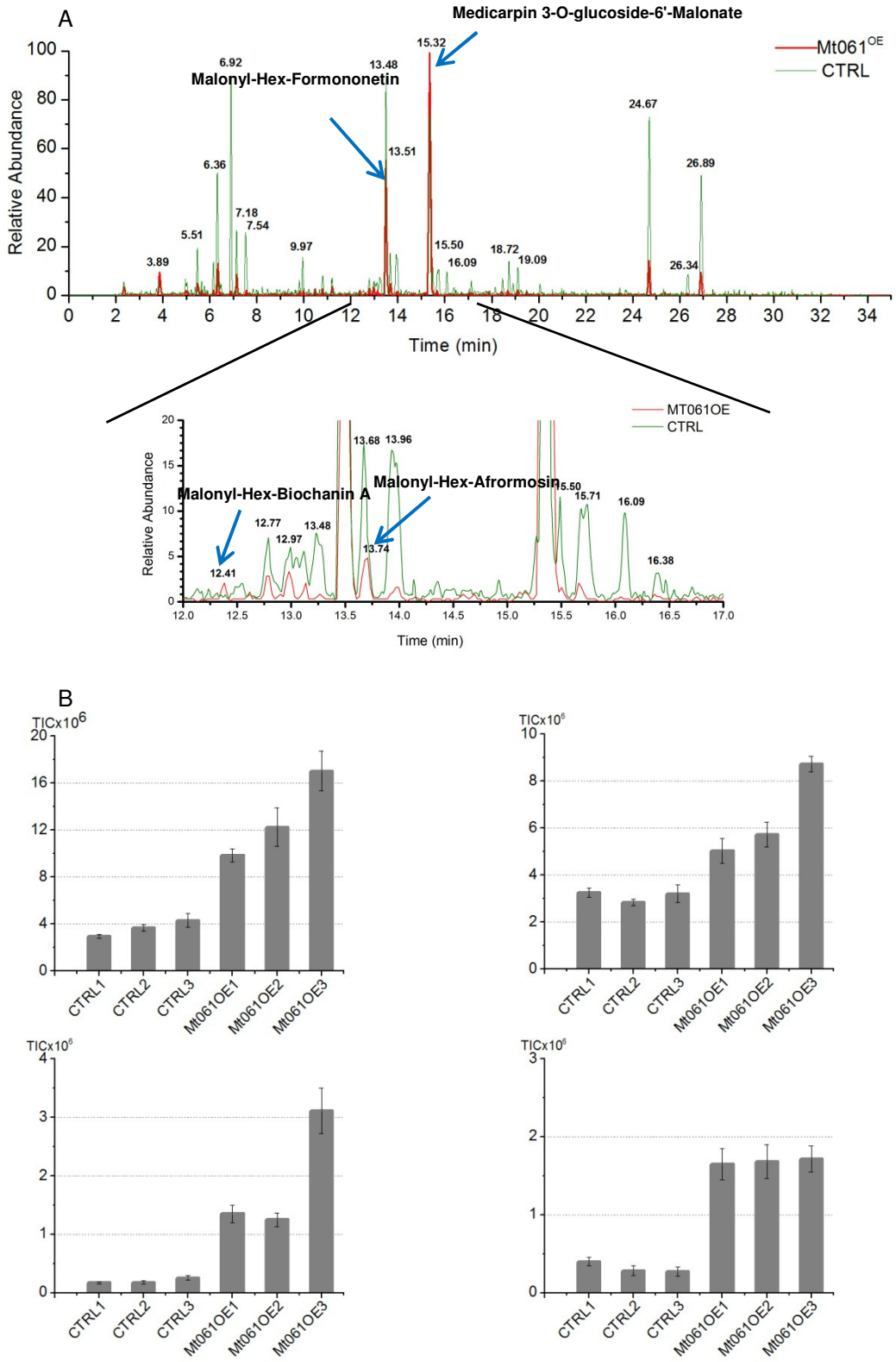


Figure 8. Metabolite profiling of Mt061^{OE} hairy roots. (A) Detail of the full MS scan of a control line (green) and a Mt061^{OE} line (red). The number indicated on each peak shows the retention time of the corresponding compound (B) Average Total Ion Current (TIC) of the main masses corresponding to medicarpin 3-*O*-glucoside-6'-Malonate(top, left), Malonyl-Hex-afrormosin (below, left), Malonyl-Hex-formononetin (top, right) and Malonyl-Hex-biochanin A (below, right). Error bars represent SEM of the five technical repeats.

2.6 RNA-Seq analysis

To gain a genome-wide view into the regulatory effect of *Mt061* we compared global gene expression between Mt061^{OE} and control lines using RNA sequencing transcript profiling (RNA-Seq analysis).

The transcript abundances were measured in fragments per kilobase of exon per million fragments mapped (FPKM) values of the different genes that are present on the genome. Gene annotations were retrieved using NCBI.

From the expressed genes in the list, we selected the genes with ≥5 fold induction (96 genes) and ≥ 10 fold reduction (43 genes), respectively (Table 1&2). In Table 1 and 2 we present the genes with increased and decreased transcript levels in Mt061^{OE} compared to control lines, respectively. The genes with enhanced transcript levels were grouped based on the putative function of the annotated genes. Since the genes with decreased transcript levels showed little functional annotations, we sorted them based on their fold of induction (Table 2). Thirty seven genes out of 96 genes with the highest fold of induction are related to direct or indirect defense responses, that can be categorized into three major groups: genes involved in volatile biosynthesis, entangled with phenolic biosynthesis and genes involved in direct defense response (Table 1).

1. Several genes potentially linked with GLV biosynthesis were upregulated in Mt061^{OE} lines. These genes can be further divided into two subgroups. Subgroup "A" consists of the genes that are directly involved in oxylipin production i.e. biosynthetic genes such as several *LOXs*, *JMT*, and those that are involved in fatty acid metabolism, including *long chain fatty acid-CoA ligase* and *fatty acid hydroxylase WAX2*. Subgroup "B" is composed of those that are involved in plastid formation, where the (precursors of) volatiles are produced. This subgroup includes the genes with putative roles in plastid formation like a GATA TF (Chiang et al., 2012), RNA polymerase sigma factor rpoD (Troxler et al., 1994; Kanamaru and Tanaka, 2004), a pentatricopeptide repeat-containing protein (Kotera et al., 2005; Schmitz-Linneweber et al., 2006), and a Dof zinc finger protein. A thylakoid lumenal protein which is involved in chloroplast thylakoid formation, where 13HPL is specifically localized (Feussner and Wasternack, 2002), is also upregulated in Mt061^{OE} lines. In addition, some plastid-specific primary metabolism

including *cysteine desulfurase* and two *pyruvate decarboxylase* genes are upregulated in Mt061^{OE} lines (Table 1).

2 Phenolics biosynthesis genes: This group is composed of monolignol biosynthesis genes such as *4-coumarate-CoA ligase* and *caffeic acid O-methyltransferase*, and genes involved in the flavonoid biosynthesis pathway including two genes encoding dihydroflavonol 4-reductase (Table 1).

3 Genes involved in defense responses, including several *trypsin protease inhibitor*, *cysteine protease inhibitor*, *pathogenesis-related protein 1a (PR1a)*, *callose synthase* and *cysteine-rich receptor-like protein kinase*. Some other resistance genes such as NBS resistance protein genes and metal tolerance protein genes were also over-expressed in Mt061^{OE} lines (Table 1).

Several genes with unknown function were up and down-regulated (25 and 26 genes, respectively). Transcript levels of some key enzymes involved in primary metabolism were also changed (Table 1 and 2).

Table 1 Genes significantly induced in the Mt061^{OE} hairy roots sorted according to their putative annotation (n=3).

Gene_id	Annotation	FPKM- normalized expression counts		
		control mean±SD	Mt061^{OE} mean±SD	Fold of induction
GLV biosynthesis genes				
Subgroup A				
MTR_3g079450	Lipoxygenase	0.88±0.2	13.06±4.50	14.9
MTR_4g018820	Jasmonate *O*-methyltransferase	31.3±14.5	363.9±184.6	11.6
MTR_7g109830	Long-chain-fatty-acid-CoA Ligase	1.09±0.48	9.15±2.321	8.41
MTR_7g090100	Fatty acid hydroxylase WAX2	0.66±0.43	4.35±1.404	6.6
MTR_8g083170	Long-chain-fatty-acid CoA ligase	0.53±0.26	3.18±0.69	5.95
MTR_8g018650	Lipoxygenase	22.2±18.5	105.1±35.83	4.74
Subgroup B				
MTR_5g012110	Thylakoid lumenal protein	0.11±0.17	1.77±0.36	15.8
MTR_4g072990	Pentatricopeptide repeat-containing protein	3±0.83	45.38±12.64	15.1
MTR_5g041400	Dof zinc finger protein	0.52±0.3	4.98±2.586	9.65
MTR_3g090900	RNA polymerase sigma factor rpoD	0.7±0.3	6.38±2.135	9.08
MTR_7g112330	GATA TF	4.72±1.58	40.1±12.07	8.5
MTR_7g102450	Cytokinin-O-glucosyltransferase	4.74±1.07	35.8±10.9	7.54
MTR_1g014320	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	58.2±20.5	373.4±98.48	6.41
MTR_5g046610	Pyruvate decarboxylase	31.2±6.36	196.7±39	6.3
MTR_8g093560	Cysteine desulfurase	40.3±21.7	249.5±63.5	6.19
MTR_5g046620	Pyruvate decarboxylase	4.38±3.18	23.34±4.95	5.33
Phenolic biosynthesis genes				
MTR_7g074870	Dihydroflavonol 4-reductase	0.04±0.06	1.24±0.76	32.6
MTR_5g027480	4-coumarate-CoA ligase	0.15±0.23	4.50±2.36	29.8

MTR_7g012070	Caffeate O-methyltransferase-1	1.15±0.79	13.3±4.54	11.5
MTR_7g074730	Dihydroflavonol-4-reductase	0.29±0.3	1.95±1.24	6.76
MTR_1g050220	Caffeic acid O-methyltransferase	0.51±0.19	2.60±0.38	5.1
Other genes involved in defense				
MTR_3g014870	Alpha-amylase/subtilisin inhibitor	0	1.01±0.77	-
MTR_3g027200	Cc-nbs resistance protein	0.05±0.07	1.21±0.50	26.3
MTR_6g078140	Trypsin proteinase inhibitor	3.65±1.72	43.0±18.4	11.8
MTR_6g078120	Trypsin proteinase inhibitor	3.13±1.84	35.4±16.4	11.3
MTR_5g012600	Pectinesterase	2.34±0.86	19.6±5.10	8.39
MTR_6g059650	Trypsin proteinase inhibitor	52.2±17.1	394.8±40.7	7.57
MTR_7g077180	Trypsin proteinase inhibitor	22.3±4.1	160.1±57.7	7.18
MTR_1g021610	Cysteine-rich receptor-like protein kinase	0.25±0.2	1.46±0.33	5.73
MTR_4g100990	Aspartic proteinase nepenthesin-2	4.59±2.36	26.23±8.17	5.72
MTR_2g012370	Pathogenesis-related protein 1a	1.64±0.56	9.34±2.28	5.69
MTR_1g101710	Callose synthase	0.22±0.18	1.26±0.29	5.63
MTR_3g084780	Cysteine proteinase inhibitor	12.8±5.43	66.59±37.8	5.19
MTR_5g084080	Nodule-specific glycine-rich protein	0.21±0.32	7.02± 2.77	33.7
Signal transduction genes				
MTR_5g075160	Epidermal patterning factor-like protein	0	3.47±1.07	-
MTR_5g034580	Rho GTPase	0.05±0.08	2.89±1.65	56.3
MTR_2g015920	Cysteine proteinase	0.07±0.11	0.74±0.5	10.9
MTR_4g006220	UBX domain-containing protein	0.22±0.35	10.11±3.88	45.1
MTR_5g092750	RING finger family protein	0.31±0.48	3.02±1.78	9.7
MTR_4g037130	Zinc finger SWIM domain-containing protein	0.75±0.62	5.01±1.23	6.71
MTR_7g117350	RING-H2 finger protein ATL1E	0.63± 0.19	3.49± 1.39	5.58
Gene expression and DNA repair				
MTR_5g036330	Structural maintenance of chromosomes protein	0.63±0.17	3.49±1.24	16.4
MTR_7g114050	Histone H2A	0.18±0.1	1.43±0.75	7.75

MTR_5g011780	CPD photolyase	0.28±0.22	2.14±0.32	5.33
Transcription factor				
MTR_3g031070	GL1-like protein	0.14±0.22	1.31±0.60	9.22
MTR_4g101970	B3 domain-containing protein	1.77±1.1	10.9±1.64	6.15
MTR_5g018260	PIF-like protein	0.25±0.2	1.48±0.56	5.99
MTR_7g075870	MADS-box TF	2.4±0.7	12.58±6.9	5.25
Auxin responsive protein				
MTR_1g063950	Auxin-induced protein 6B	1.54±1.23	13.51±4.93	8.78
MTR_5g021820	Auxin-induced protein 6B	5.06±2.7	29.19±10.18	5.76
MTR_3g084230	Auxin-induced protein-like protein	0.91±1.41	4.84±1.63	5.32
Other metabolism genes				
MTR_4g021210	Ribulose bisphosphate carboxylase/oxygenase activase	1.12±0.81	20.68±5.19	27.7
MTR_5g007460	Cytochrome P450	0.24±0.1	2.47±1.64	10.4
MTR_8g104080	Cytochrome P450	6.95±3.28	63.3±22.57	9.1
MTR_8g104100	Cytochrome P450	7.27±2.99	62.6±19.35	8.62
MTR_5g076200	2-succinylbenzoate-CoA ligase	0.19±0.29	1.50±0.122	8
MTR_7g044920	Albumin	0.24±0.37	8.22±1.32	7.2
MTR_5g012390	L-lactate dehydrogenase	3.07±1.98	20.5±3.15	6.71
MTR_5g027530	Phosphoribulokinase	1.33±0.7	8.22±2.13	6.17
MTR_2g019630	Phospho-N-acetylmuramoyl-pentapeptide-transferase	9.53±1.97	56.11±15.04	5.89
MTR_1g043220	Inorganic phosphate transporter 1-1	0.53±0.75	3.15± 2.64	5.93
Genes with no annotation, no known function				
MTR_4g133580	Hypothetical protein	0	8.00±3.56	-
MTR_8g070620	Hypothetical protein	0	15.33±2.12	-
MTR_4g121590	Hypothetical protein	0.06±0.09	1.39±1.11	23.3
MTR_2g103460	Hypothetical protein (Glyoxalase/Bleomycin	0.89±1.38	19.28±7.99	21.7

	resistance protein/Dioxygenase)			
MTR_1g095730	Hypothetical protein	3.96±2.97	48.01±28.01	12.1
MTR_1g019890	Hypothetical protein	0.22±0.3	2.21±1.48	10.1
MTR_4g114080	Hypothetical protein (Glyoxalase/Bleomycin resistance protein/Dioxygenase)	1.82±0.63	14.14±4.67	7.78
MTR_4g064360	Hypothetical protein	5.49±6.27	42.18±18.79	7.69
MTR_4g092780	Hypothetical protein (Galactose mutarotase)	1.32±1.05	9.79±0.68	7.42
MTR_5g089580	Hypothetical protein	1.62±0.65	11.58±1.62	7.15
MTR_7g108130	Hypothetical protein	3.17±0.26	22.06±4.35	6.97
MTR_7g082750	Hypothetical protein	0.6±0.9	4.13±0.67	6.92
MTR_3g091930	Hypothetical protein	0.93±0.73	6.35±1.12	6.82
MTR_5g058800	Hypothetical protein	0.58±0.39	3.86±2.73	6.63
MTR_2g100600	Hypothetical protein	1.64±0.39	10.41±8.29	6.35
MTR_4g083140	Hypothetical protein	6.69±5.21	38.62±12.86	5.77
MTR_5g094990	Hypothetical protein	0.18±0.14	1.04±0.83	5.76
MTR_8g021170	Hypothetical protein	0.48±0.74	2.68±0.58	5.65
MTR_5g006200	Hypothetical protein	2.81±4.3	15.19±9.49	5.41
MTR_8g061120	Hypothetical protein	1.67±1.16	8.88±4.65	5.33
MTR_8g005880	Hypothetical protein	10.9±3.55	57.78±28.28	5.28
MTR_2g063870	Hypothetical protein	10.9±3.55	57.78±28.28	5.28
MTR_2g029150	Hypothetical protein	0.39±0.4	2.04±1.001	5.17
MTR_8g106680	Hypothetical protein (NT-C2)	1.43±0.58	7.38±0.9	5.16
MTR_8g042200	Hypothetical protein	1.32±2.04	6.67±2.76	5.06
Transport and membrane trafficking				
MTR_4g006650	Aquaporin	1.87±0.07	30.4±11.7	16.2
MTR_4g133040	Importin alpha-1b subunit	0.11±0.09	1.25±0.51	10.9
MTR_8g009560	PRA1 family protein B1	0.54±0.38	5.02±2.026	9.23
MTR_3g111020	Multidrug and toxin extrusion protein	0.17±0.27	1.43±0.775	8.34
MTR_5g006070	Hexose transporter	108±33.7	563.2±53.4	5.19
MTR_5g071860	Potassium transporter	2.03±0.53	11.06±1.54	5.44

Table 2 Genes significantly repressed in the Mt061^{OE} hairy roots sorted according to fold of reduction (n=3).

		FPKM- normalized expression counts		
Gene_id	**Annotation**	**Control mean±SD**	**Mt061^{OE} mean±SD**	**Fold of induction**
MTR_2g045100	Germin-like protein	3.57±3.73	0	0
MTR_1g061980	Hypothetical protein	2.15±1.12	0	0
MTR_1g100870	Hypothetical protein	7.43±6.33	0	0
MTR_2g019360	Hypothetical protein	4.21±0.79	0	0
MTR_3g051620	Hypothetical protein	74.11±55.3	0	0
MTR_3g116570	Hypothetical protein	67.89±23.5	0	0
MTR_8g021280	Hypothetical protein	110.21±44.1	0	0
MTR_8g041260	Hypothetical protein (transposase)	11.86±9.18	0.03±0.028	0.002
MTR_2g008290	GRF zinc finger containing protein	16.62±11.9	0.31±0.47	0.018
MTR_1g045790	Hypothetical protein	11.56±4.31	0.30±0.24	0.026
MTR_4g010310	Hypothetical protein	39.07±23.52	1.44±1.58	0.037
MTR_1g016370	PIF-like protein	7.29±4.29	0.29±0.34	0.04
MTR_6g079020	Disease resistance-like protein	14.81±6.44	0.62±0.14	0.042
MTR_6g079000	Disease resistance-like protein	11.69±4.55	0.60±0.18	0.051
MTR_5g072920	Cytochrome P450 71B37	11.50±1.23	0.64±0.50	0.056
MTR_1g043330	Hypothetical protein	4.75±3.49	0.27±0.42	0.057
MTR_7g086390	Hypothetical protein	6.37±1.54	0.38±0.59	0.060
MTR_4g064640	Hypothetical protein	8.24±4.61	0.50±0.78	0.061
MTR_3g025470	Hypothetical protein	8.08±3.18	0.49±0.76	0.061
MTR_3g055630	Hypothetical protein	18.21±2.6	1.26±1.23	0.069
MTR_4g014860	NBS-containing resistance-like protein	17.98±2.9	1.26±1.23	0.069
MTR_4g058550	Hypothetical protein	17.08±3.9	1.24±1.21	0.073
MTR_8g036660	Aluminum activated citrate transporter	51.29±15.3	3.76±1.17	0.073
MTR_5g045710	Alcohol dehydrogenase class-3	9.35±4.19	0.69±0.42	0.074
MTR_5g074580	O-methyltransferase	4.66±1.70	0.35±0.32	0.075
MTR_6g059830	Hypothetical protein	4.47±1.04	0.33±0.41	0.075
MTR_3g032630	Hypothetical protein	1002.91±588.8	78.10±21.9	0.078
MTR_8g039620	Hypothetical protein (ankyrin repeat)	37.21±7.45	2.94±2.81	0.079

MTR_6g078990	Hypothetical protein (Cdt1_m & Cdt1_c)	9.07±4.51	0.74±0.64	0.081
MTR_6g044390	Aquaporin MIP family	8.19±3.40	0.67±0.54	0.082
MTR_7g006150	Hypothetical protein	2.72±1.58	0.23±0.19	0.084
MTR_1g006990	Subtilisin-like serine protease	73.34±17.8	6.32±7.95	0.086
MTR_4g090180	Hypothetical protein	36.31±26.2	3.14±1.014	0.086
MTR_7g086400	Hypothetical protein	18.51±6.56	1.61±2.49	0.087
MTR_5g073960	Leucine zipper protein	5.17±2.247	0.45±0.26	0.088
MTR_6g088140	Lanosterol synthase	12.38±6.52	1.10±0.15	0.089
MTR_2g033580	Desiccation-related protein PCC13-62	5.08±2.54	0.46±0.14	0.091
MTR_3g047450	Hypothetical protein	151.51±89.05	13.84±6.67	0.091
MTR_6g031240	GDSL esterase/lipase	2.55±1.28	0.24±0.19	0.092
MTR_3g025640	Hypothetical protein	6.46±0.27	0.62±0.56	0.096
MTR_6g012160	Quinone oxidoreductase-like protein	2.29±1.15	0.22±0.21	0.097
MTR_7g104460	Hypothetical protein (phosphatidylethanolamine-Binding Protein)	29.68±15.65	2.93±0.91	0.098
MTR_4g113590	Hypothetical protein	12.11±4.87	1.20±1.86	0.099

3 Discussion

3.1 *Mt061* stimulates GLV biosynthesis in *M. truncatula* hairy roots

GLVs including hexanol, *cis*-3-hexenol and their corresponding aldehydes are a valuable flavor class and widely used in food industry. Large scale production of these compounds is still a big challenge for producers and therefore, applying biotechnological approaches to improve the biosynthesis of these compounds could be beneficial.

In a healthy, undamaged plant small quantities of GLVs are emitted, while, large quantities can be released by plants immediately after wounding or herbivore attack (Turlings et al., 1995; D'Auria et al., 2007). Signaling pathways involved in plant defense responses against herbivores are complex and specific. It is known that certain TFs orchestrate anti-herbivore defense responses, but only a few have thus far been characterized. At present, we have very limited knowledge of the molecular factors regulating the expression of GLV biosynthesis genes (Manavella et al., 2008; Oh et al., 2012).

In this study, a cDNA-AFLP-based transcriptome analysis of a time series covering the early response of *M. truncatula* cells to MeJA elicitation led to the discovery of a MYB TF, *Mt061,* the overexpression of which in *M. truncatula* hairy roots resulted in the production and emission of GLV compounds such as hexanal (Fig. 3). Therefore, we considered it as a possible regulator of the genes encoding the enzymes catalyzing GLV biosynthesis. However, the results showed that only few GLV biosynthesis genes, one lipase and one LOX (Fig. 4) and only the jasmonate *O*-methyltransferase from the JA synthesis pathway were upregulated in Mt061^{OE} hairy roots (Table 1) which does not permit us to detail the direct relationship between *Mt061*overexpression and GLV biosynthesis gene regulation. For instance expression of the two *HPLs* (*HPL1* and *HPL2*) that have been reported from *M. truncatula* so far and corresponding to GLV-specific branches of the oxylipins pathway was not affected.

3.2 *Mt061* induces genes involved in direct defense responses and in phenolic biosynthesis pathway

The RNA-Seq data analysis also revealed that several phenolic biosynthesis genes were upregulated in Mt061^{OE} lines (Table 1). Moreover, we found accumulation of higher levels of phytoalexins and pterocarpan in Mt061^{OE} lines as compared to control

lines (Fig. 7). There are several reports from induction of phenolic biosynthesis genes and the accumulation of phytoalexin compounds in plants after applying GLV compounds such as *trans*-2-hexenal (Bate and Rothstein, 1998; Arimura et al., 2001; Kishimoto et al., 2005). Recently, it was reported that GLV compounds can prime JA responses in plants leading to the induction of higher anthocyanin accumulation in treated plants (Hirao et al., 2012). Therefore, it is plausible that phenolic gene induction and phytoalexin accumulation are consequences of GLV production in the transgenic hairy roots. Conversely however, previous studies have shown that when *M. truncatula* cell cultures were elicited with YE, flavonoid biosynthesis genes were mostly upregulated, while with MeJA elicitation the genes involved in biosynthesis of triterpene saponins were induced (Broeckling et al., 2005; Naoumkina et al., 2008). *In silico* expression analysis with the MtGEA tool showed that *Mt061* is also upregulated with YE elicitation. These results might support direct role of *Mt061* in flavonoid biosynthesis, which needs further investigation.

Furthermore, we observed that several genes involved in defense responses such as four *TPI, cysteine protease inhibitor* and *PR1a* were induced in $Mt061^{OE}$ lines (Table 1). In plants, protease inhibitors are produced as a defense mechanism in response to herbivore attack, to inhibit proteolysis and negatively affect the digestibility of ingested plant material in insect guts (Zavala et al., 2004; Hartl et al., 2010). Analogous results were reported for *HAHB4*, a member of the sunflower (*Helianthus annuus*) HD- ZIP subfamily that was previously characterized as a positive regulator of the synthesis of JAs and GLVs. *HAHB4* upregulates the transcript levels of several genes involved in JA and GLV biosynthesis as well as *TPI* and pathogen- related genes (Manavella et al., 2008). Collectively, these data suggest that apart from regulating the volatile emissions that are known to function as an indirect defense against herbivores in plants, *Mt061* might also regulate the accumulation of direct defense compounds such as TPIs and phytoalexins. Whether the induction of phytoalexin biosynthesis is the consequences of GLV production or, conversely, directly mediated by *Mt061*, as a multifunctional TFs involved in cross- talks between several biochemical pathways, remains to be determined.

3.3 *Mt061* affects genes involved in plastid formation and/or functions

Formation of plant volatiles can follow similar developmental patterns as other plant organs. For instance, volatiles are increasingly produced during the early stages of

organ development when leaves are young and not fully expanded, fruits are not yet mature, or when flowers are ready for pollination. Subsequently, their production can remain constant or decrease over the organs' life span (Bouwmeester et al., 1998; Dudareva and Pichersky, 2000; Gershenzon et al., 2000). Plastids are the organelles in which the precursors of GLVs are biosynthesized before being transferred to the cytosol for further modifications (Pichersky et al., 2006). Accordingly, it was shown that 9/13-HPL (HPL2 in this study), an isoform which is specifically involved in GLV biosynthesis, is localized in chloroplast thylakoids and in the outer membrane of the chloroplast envelope (Feussner and Wasternack, 2002). In this study we showed that several genes involved in plastid development including the thylakoid lumenal protein which is involved in chloroplast thylakoid formation were upregulated in Mt061^{OE} lines (Table 1). This suggests that *Mt061* might be involved in the development of plastids in *M. truncatula*.

Overall this indicates that *Mt061,* a MYB TF, might govern a suite of defense responses in *M. truncatula* leading to the production of defense-related metabolites, GLVs and phytoalexins. Nonetheless, much work is still required to fully understand the complex regulatory network by and within which this TF operates.

4 Materials and methods

4.1 Generation of DNA constructs

To obtain the FL-ORF construct specify to *Mt061*, the cDNA-AFLP tag sequence was used for BLASTn searches against the *Medicago truncatula* Gene Index database (http://compbio.dfci.harvard.edu/tgi/). PCR fragments corresponding to the FL-ORF of Mt061 were amplified and cloned into entry vector pDONR221 by Gateway™ recombination. To obtain an entry clones with stop codons, Gateway primers were designed according to Underwood (Underwood et al., 2006). The entry constructs were sequence-verified. For the overexpression experiment, Gateway recombination was carried out with the pK7WG2D binary vector (Karimi et al., 2002). The resulting clones were transformed into the *Agrobacterium rhizogenes* strain LBA 9402/12 for generation of hairy roots.

4.2 Generation and cultivation of transgenic *M. truncatula* hairy roots

Protocol was adapted from Boisson-Dernier et al (Boisson-Dernier et al., 2001) with some modifications as described in chapter 4 page 130-1.

4.3 *M. truncatula* hairy root culture maintenance and elicitation

For elicitation, the hairy roots were cultured in MS liquid medium and three weeks after subculturing into fresh medium, hairy roots were treated with 100 µM MeJA (dissolved in EtOH) or an equivalent amount of EtOH as a control. Samples were harvested and frozen in liquid nitrogen. For transcript profiling, samples were taken 0, 0.5, 1, 2, 4, 8, 12, and 24 h after elicitor or mock treatments, and 48 h after treatments for metabolite profiling. For NaCl treatment, hairy roots were subcultured in MS liquid medium and three weeks later, transferred to fresh medium containing 200mM NaCl (dissolved in water) or to medium without NaCl as a control.

4.4 Transcript profiling

For quantitative Real Time PCR (qRT-PCR), RNA from were isolated with the RNeasy mini kit (Qiagen) according to manufacturer instruction. Quality control and quantification were performed with a Nanodrop spectrometer (Isogen, Hackensack, NJ). cDNA were prepared using SuperscriptTM II Reverse Transcriptase (Invitrogen). Primers were designed with Beacon Designer version 4.0 (Premier Biosoft International, Palo Alto, CA, USA). qRT-PCR was carried out as described in chapter 4 page 131.

For the statistical analysis of the data “paired samples t-test” was performed and the statistical analyses were performed using SPSS v.21.0 (IBM Corp., Armonk, NY, USA).

4.5 Metabolite extractions

For metabolite profiling by LC ESI-FT-MS, *M. truncatula* hairy roots were harvested in five biological repeats of three independent transgenic lines per transgene construct and rinsed with purified H_2O under vacuum filtration. Metabolites were extracted as described in chapter 4 page 131.

For GC-MS analysis, hairy roots grown on the plates were transferred to the GC vials and crushed using sterile tweezers. The vials were immediately capped with air tight caps. Sample vials were heated at 50°C for 10min. 200µl of He (solvent phase) was injected into the vial and allowed to equilibrate for 30s. 200µl of the gas mixture (head-space) was subjected to GC-MS.

4.6 GC-MS analysis

For characterization of volatile compounds produced and emitted by $Mt061^{OE}$, the volatile blends from $Mt061^{OE}$ lines were subjected to GC-MS (GC model 6890, MS model 5973, Agilent). A 200 µl of gas mixture (head-space) was injected (split mode) to CP-Select 624 (ChromPac, Chromp99927) 6% Cyanopropylphenyl column and operated at a constant helium flow of 0.8 ml/min. The injector temperature was set to 45°C and the oven temperature was held at 40°C for 1 min post injection, ramped to 180°C at 2.5°C/min, ramped to 200°C at 15°C/min, held at 200°C for 1 min, and finally cooled down to 40°C at 50°C/min at the end of the run. For each sample the GC total run time was 80 min. The MS transfer line was set to 220°C, the MS ion source to 230°C and the quadruple to 180°C, throughout. For identification of metabolites a full mass spectra was generated by scanning the m/z range of 35-350 with a solvent delay of 2 min.

4.7 LC ESI FT-ICR MS

The LC ESI FT-ICR MS analysis was performed as described in chapter 4, pages 132.

4.8 RNA-Seq

Total RNA from three independent $Mt061^{OE}$ and control lines was extracted with the RNeasy mini kit (Qiagen) and cDNA prepared with SuperScript™ II Reverse Transcriptase (Invitrogen). RNA samples were sent to "Genomics core" (KUL, Leuven, Belgium) for mRNA purification, cDNA library construction and Illumina HiSeq 2000-based RNA sequencing with Solexa technology.

Read mapping on the *Medicago* genome (MT3.5; (Young et al., 2011)) with TopHat version 2.0.3 (Trapnell et al., 2009) and counting of the uniquely mapped reads and calculation of the FPKM values with Cufflinks version 1.3.0 (Trapnell et al., 2010) were performed using default parameters as described (Pollier et al., 2013).

Acknowledgments

Azra Gholami, Nathan De Geyter, Jacob Pollier and Tessa Moses are indebted to the Iranian Ministry for Health and Medical Education, Agency for Innovation by Science and Technology FWO and the VIB International PhD Program, respectively for predoctoral fellowships.

References

Arimura, G., Ozawa, R., Horiuchi, J., Nishioka, T., and Takabayashi, J. (2001). Plant-plant interactions mediated by volatiles emitted from plants infested by spider mites. Biochemical Systematics and Ecology **29,** 1049-1061.

Arimura, G.I., Ozawa, R., and Maffei, M.E. (2011). Recent Advances in Plant Early Signaling in Response to Herbivory. International Journal of Molecular Sciences **12,** 3723-3739.

Bate, N.J., and Rothstein, S.J. (1998). C-6-volatiles derived from the lipoxygenase pathway induce a subset of defense-related genes. Plant Journal **16,** 561-569.

Boatright, J., Negre, F., Chen, X., Kish, C.M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D., and Dudareva, N. (2004). Understanding *in vivo* benzenoid metabolism in petunia petal tissue. Plant Physiology **135,** 1993-2011.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C., and Barker, D.G. (2001). *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Molecular Plant-Microbe Interactions **14,** 695-700.

Bouwmeester, H.J., Gershenzon, J., Konings, M.C., and Croteau, R. (1998). Biosynthesis of the monoterpenes limonene and carvone in the fruit of caraway. I. Demonstration of enzyme activities and their changes with development. Plant Physiology **117,** 901-912.

Broeckling, C.D., Huhman, D.V., Farag, M.A., Smith, J.T., May, G.D., Mendes, P., Dixon, R.A., and Sumner, L.W. (2005). Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. Journal of Experimental Botany **56,** 323-336.

Chiang, Y.-H., Zubo, Y.O., Tapken, W., Kim, H.J., Lavanway, A.M., Howard, L., Pilon, M., Kieber, J.J., and Schaller, G.E. (2012). Functional characterization of the GATA transcription factors GNC and CGA1 reveals their key role in chloroplast development, growth, and division in Arabidopsis. Plant Physiology **160,** 332-348.

D'Auria, J.C., Pichersky, E., Schaub, A., Hansel, A., and Gershenzon, J. (2007). Characterization of a BAHD acyltransferase responsible for producing the green leaf volatile (*Z*)-3-hexen-1-yl acetate in *Arabidopsis thaliana*. The Plant journal : for cell and molecular biology **49,** 194-207.

Dudareva, N., and Pichersky, E. (2000). Biochemical and molecular genetic aspects of floral scents. Plant Physiology **122,** 627-633.

Dudareva, N., Negre, F., Nagegowda, D.A., and Orlova, I. (2006). Plant volatiles: recent advances and future perspectives. Critical Reviews in Plant Sciences **25,** 417-440.

Feussner, I., and Wasternack, C. (2002). The lipoxygenase pathway. Annual Review of Plant Biology **53,** 275-297.

Gershenzon, J., McConkey, M.E., and Croteau, R.B. (2000). Regulation of monoterpene accumulation in leaves of peppermint. Plant Physiology **122,** 205-214.

Hartl, M., Giri, A.P., Kaur, H., and Baldwin, I.T. (2010). Serine protease inhibitors specifically defend *solanum nigrum* against generalist herbivores but do not influence plant growth and development. Plant Cell **22,** 4158-4175.

Hirao, T., Okazawa, A., Harada, K., Kobayashi, A., Muranaka, T., and Hirata, K. (2012). Green leaf volatiles enhance methyl jasmonate response in Arabidopsis. Journal of Bioscience and Bioengineering **114,** 540-545.

Kanamaru, K., and Tanaka, K. (2004). Roles of chloroplast RNA polymerase sigma factors in chloroplast development and stress response in higher plants. Bioscience, Biotechnology, and Biochemistry **68,** 2215-2223.

Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY$^{((TM))}$ vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science **7,** 193-195.

Kishimoto, K., Matsui, K., Ozawa, R., and Takabayashi, J. (2005). Volatile C6-aldehydes and allo-ocimene activate defense genes and induce resistance against *Botrytis cinerea* in *Arabidopsis thaliana*. Plant and Cell Physiology **46,** 1093-1102.

Kotera, E., Tasaka, M., and Shikanai, T. (2005). A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature **433,** 326-330.

Liavonchanka, A., and Feussner, N. (2006). Lipoxygenases: Occurrence, functions and catalysis. Journal of Plant Physiology **163,** 348-357.

Manavella, P.A., Dezar, C.A., Bonaventure, G., Baldwin, I.T., and Chan, R.L. (2008). *HAHB4*, a sunflower HD-Zip protein, integrates signals from the jasmonic acid and ethylene pathways during wounding and biotic stress responses. Plant Journal **56,** 376-388.

Matsui, K. (2006). Green leaf volatiles: hydroperoxide lyase pathway of oxylipin metabolism. Current Opinion in Plant Biology **9,** 274-280.

Naoumkina, M.A., He, X.Z., and Dixon, R.A. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. BMC Plant Biology **8,** 132-146.

Newman, J.D., and Chappell, J. (1999). Isoprenoid biosynthesis in plants: carbon partitioning within the cytoplasmic pathway. Critical Reviews In Biochemistry And Molecular Biology **34,** 95-106.

Oh, Y., Baldwin, I.T., and Galis, I. (2012). NaJAZh regulates a subset of defense responses against herbivores and spontaneous leaf necrosis in *Nicotiana attenuata* plants. Plant Physiology **159,** 769-788.

Pichersky, E., Noel, J.P., and Dudareva, N. (2006). Biosynthesis of plant volatiles: Nature's diversity and ingenuity. Science **311,** 808-811.

Pollier, J. (2011). Combinatorial Biochemistry of Triterpene Saponins in Plants (Faculty Of Sciences VIB Department Of Plant Systems Biology Ghent University).

Pollier, J., Rombauts, S., and Goossens, A. (2013). Analysis of RNA-Seq data with TopHat and Cufflinks for genome-wide expression analysis of jasmonate-treated plants and plant cultures. In In Jasmonate Signaling: Methods and Protocols. Methods in Molecular Biology (Humana Press), pp. in press.

Rodriguez-Concepcion, M., and Boronat, A. (2002). Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology **130,** 1079-1089.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A., and Barkan, A. (2006). A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 Pre-mRNA. Plant Cell **18,** 2650-2663.

Schuurink, R.C., Haring, M.A., and Clark, D.G. (2006). Regulation of volatile benzenoid biosynthesis in petunia flowers. Trends in Plant Science **11,** 20-25.

Seo, H.S., Song, J.T., Cheong, J.J., Lee, Y.H., Lee, Y.W., Hwang, I., Lee, J.S., and Choi, Y.D. (2001). Jasmonic acid carboxyl methyltransferase: A key enzyme for jasmonate-regulated plant responses. Proceedings of the National Academy of Sciences of the United States of America **98,** 4788-4793.

Shao, H., Dixon, R.A., and Wang, X. (2007). Crystal structure of vestitone reductase from alfalfa (*Medicago sativa* L.). Journal of Molecular Biology **369,** 265-276.

Shiojiri, K., Kishimoto, K., Ozawa, R., Kugimiya, S., Urashimo, S., Arimura, G., Horiuchi, J., Nishioka, T., Matsui, K., and Takabayashi, J. (2006). Changing green leaf volatile biosynthesis in plants: An approach for improving plant resistance against both herbivores and pathogens. Proceedings of the National Academy of Sciences of the United States of America **103,** 16672-16676.

Song, M.S., Kim, D.G., and Lee, S.H. (2005). Isolation and characterization of a jasmonic acid carboxyl methyltransferase gene from hot pepper (*Capsicum annuum* L.). Journal of Plant Biology **48,** 292-297.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25,** 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology **28,** 511-515.

Troxler, R.F., Zhang, F., Hu, J., and Bogorad, L. (1994). Evidence that sigma-factors are components of chloroplast RNA-polymerase. Plant Physiology **104,** 753-759.

Turlings, T.C., Loughrin, J.H., McCall, P.J., Rose, U.S., Lewis, W.J., and Tumlinson, J.H. (1995). How caterpillar-damaged plants protect themselves by attracting parasitic wasps. Proceedings of the National Academy of Sciences of the United States of America **92,** 4169-4174.

Underwood, B.A., Vanderhaeghen, R., Whitford, R., Town, C.D., and Hilson, P. (2006). Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations. Plant Biotechnology Journal **4,** 317-324.

Verdonk, J.C., de Vos, C.H.R., Verhoeven, H.A., Haring, M.A., van Tunen, A.J., and Schuurink, R.C. (2003). Regulation of floral scent production in petunia revealed by targeted metabolomics. Phytochemistry **62,** 997-1008.

Xie, D.-Y., Jackson, L.A., Cooper, J.D., Ferreira, D., and Paiva, N.L. (2004). Molecular and Biochemical Analysis of Two cDNA Clones Encoding Dihydroflavonol-4-Reductase from Medicago truncatula. Plant Physiology **134,** 979-994.

Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., and al., e. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature **480,** 520-524.

Zavala, J.A., Patankar, A.G., Gase, K., Hui, D.Q., and Baldwin, I.T. (2004). Manipulation of endogenous trypsin proteinase inhibitor production in Nicotiana attenuata demonstrates their function as antiherbivore defenses. Plant Physiology **134,** 1181-1190.

6

Functional characterization of small signalling peptides that potentially steer secondary metabolite biosynthesis in *Medicago truncatula*

Azra Gholami, Jacob Pollier, Alain Goossens

Author contribution:

Generation of transgenic lines, qRT-PCR transcript profiling, data interpretation

Abstract

Plants are an enormously rich source of peptides that can function as signaling molecules governing several aspects of plant growth and development. Several signaling peptides have been identified to control secondary metabolism in plants. Recently, Taximin, a plant specific peptide has been identified in Taxus baccata that seems to be a conserved regulator of plant metabolism. In this study, we explored the effects of ectopically expressed Taximin on secondary metabolism in M. truncatula hairy root cultures. The results indicated that Taximin can modulate saponin production in M. truncatula hairy roots. In addition, we describe the cloning, characterization and functional analysis of the Taximin homologs in the model legume M. truncatula. Five close homologs of the Taximin gene were identified in M. truncatula. Gain- or loss function studies suggested that MtTAX genes might be involved in the control of (secondary) metabolism in M. truncatula.

1 Introduction

The inducible defense mechanisms in plants are activated upon perception of extra- or intracellular signals by receptors on the surface of the plasma membrane or endomembrane or in the nucleus. Subsequently, a signal transduction network is orchestrated leading to the activation or *de novo* biosynthesis of transcription factors which regulate the expression of the genes involved in the defense mechanisms such as biosynthesis of plant secondary metabolites. Several defense regulators, including jasmonic acid (JA), salicylic acid (SA) and ethylene trigger the induction of the defense mechanisms in plants (Odjakova and Hadjiivanova, 2001; Zhao et al., 2005). It has become evident that many secretory and nonsecretory small peptides also function as signaling molecules in plants. Small peptides are protein molecules with a molecular mass of up to 10 kDa that can be involved in various aspects of plant growth regulation, including callus growth, meristem organization, root growth, leaf-shape regulation, nodule development, organ abscission, self-incompatibility and defense responses (Matsubayashi and Sakagami, 2006; Butenko et al., 2009; Marshall et al., 2011). Furthermore, some plant peptides are involved in the regulation of secondary metabolism in plants. For example, systemin induces volatile organic compounds in tomato and phytosulfokine (PSK-α) was shown to stimulate production and accumulation of tropane alkaloids in hairy root cultures of *Atropa belladonna* (Sasaki and Takahashi, 2002) and taxol in *Taxus canadensis* cell cultures when added

simultaneously with Methyl jasmonate (MeJA) (Kim et al., 2006). Therefore, they can be used as an elicitor to enhance the production of metabolites in plant species.

Recently, Taximin, a 73 amino acid peptide, was identified through genome-wide transcript profiling of MeJA-elicited cell cultures of *Taxus baccata.* Taximin was shown to be capable of activating secondary metabolism genes in tobacco and causing an increase in nicotine production in tobacco hairy roots (Onrubia Ibáñez, 2012). Taximin constitutes of a 27 amino-acid signaling peptide and a 46 amino-acid mature peptide containing 6 cysteine residues. Therefore, it has a common feature of the Cys-rich proteins including small size (less than 160 amino acid residues), a conserved N-terminal signal peptide and a C-terminal mature peptide with conserved cysteine residues (Marshall et al., 2011). Considering the small size and lack of homology to known domains, Taximin has been suggested as a Cys-rich pathogenesis-related peptide involved in MeJA- inducible defense systems in plants (Onrubia Ibáñez, 2012). Taximin is a plant specific peptide and seems to be conserved across the plant kingdom. The mature peptide part of Taximin presents high homology to predicted proteins from several plant species including *Vitis vinifera*, *Arabidopsis*, *Poplus trichocarpa*, *Ricinus communis* and *Glycine max* (Onrubia Ibáñez, 2012).

In this study, the impact(s) of ectopic expression of *Taximin* on secondary metabolism in *M. truncatula* have been explored. Higher levels of saponin glycosides were detected in *Taximin* overexpressing transgenic lines as compared to control lines. In addition, we identified five close homologs of the *Taximin* gene in *M. truncatula* and assessed their role(s) in secondary metabolism by gain-or loss-of-function analyses. The results suggest that *Taximin* and *MtTAX* genes can modulate secondary metabolite biosynthesis in *M. truncatula*. However, their exact mechanism of action is not determined yet.

2 Results

2.1 Identification of *Medicago* Taximin-like sequences and phylogenetic analysis

T. baccata sequence for the Taximin protein were retrieved from the previous study (Onrubia Ibáñez, 2012) and used for mining plant genome sequences using the BLASTP program. Two similar sequences for *Arabidopsis* (At2g20562 and At2g311090) and three in *Glycine max* (soybean) (ACU17275, ACU14314 and ACU14776) were identified but no hit was obtained from the *M. truncatula* genome.

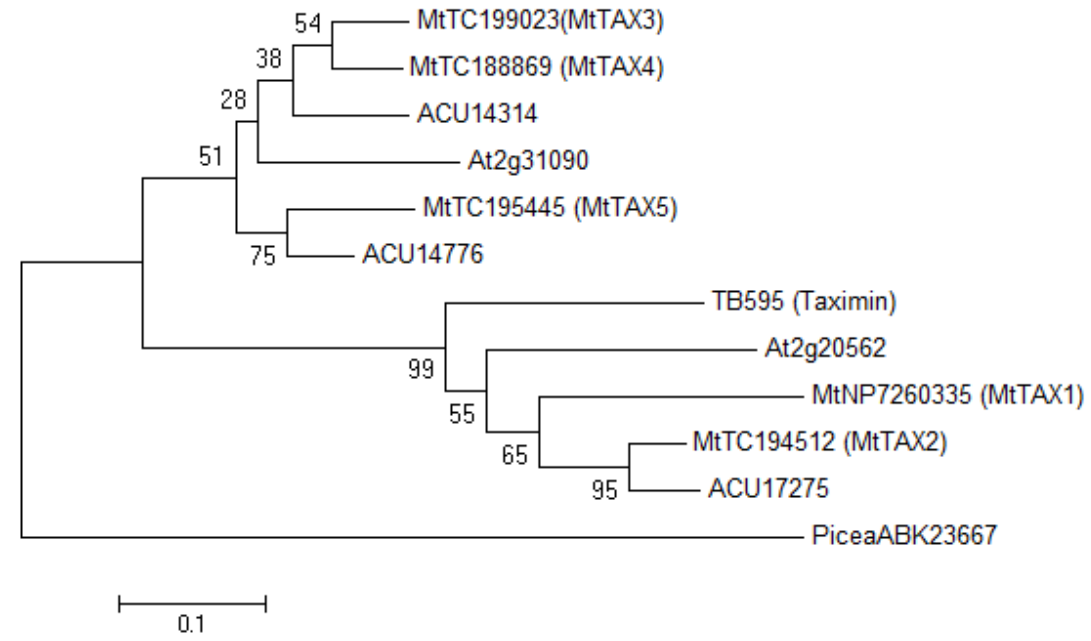
Then we used the hits from soybean to BLAST against the *M. truncatula* Database in the DFCI *Medicago* Gene Index (http://compbio.dfci.harvard.edu/cgibin/tgi/gimain.pl?gudb=medicago) for the available EST contigs with the highest similarity. Five unique ESTs for the *Taximin* homolog genes could be identified, including NP7260335, TC194512, TC199023, TC188869, TC195445, and named *MtTAX1* to *MtTAX5,* respectively.

Phylogenetic analysis with the amino acid sequences of Taximin-like protein in *M. truncatula* showed two subclades (Fig. 1A). The first subclade contains three of the characterized *MtTAX* genes including TC199023 (*MtTAX3*), TC188869 (*MtTAX4*) and TC195445 (*MtTAX5*) together with two genes from soybean and one homologous gene from *Arabidopsis thaliana* (Fig. 1A).

The second subclade contains the two other *MtTAX* genes, NP7260335 (*MtTAX1*) and TC194512 (*MtTAX2*), together with one homologous gene from soybean and *Arabidopsis* each and *Taximin* from *T. baccata* (*TB595*).

The amino acid identity of Taximin with its closest *M. truncatula* homologs, NP7260335 (MtTAX1) and TC194512 (MtTAX2) is 79.7% and 82.1%, respectively. Taximin shares 70.1%, 70.1% and 70.0% amino acid similarity with TC199023 (MtTAX3), TC188869 (MtTAX4) and TC195445 (MtTAX5), respectively (Fig. 1B). Similar to Taximin, the MtTAX peptides from *M. truncatula* have a small size (80 amino acids or less), a C-terminal mature peptide with the conserved six cysteine residues and an N-terminal signal peptide, altogether representing the common structural feature of Cys-rich proteins.

A



B

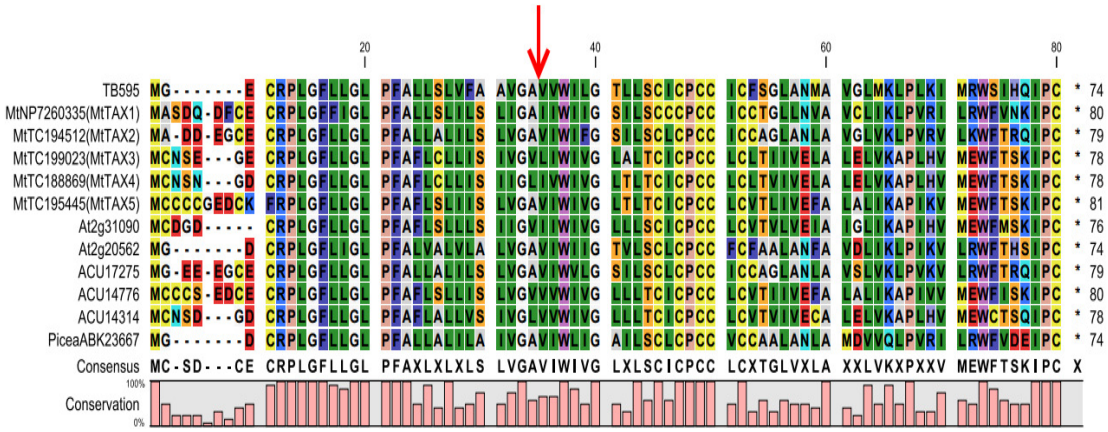
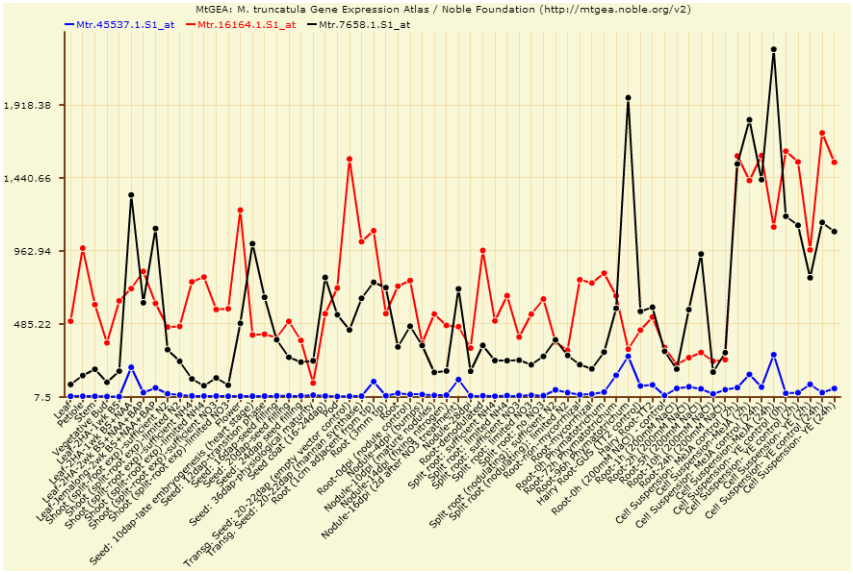


Figure 1. A) Phylogenetic analysis of Taximin and the identified peptides with the highest similarity. B) Alignment of the amino acid sequences of Taximin and its homologs from *Arabidopsis*, *Glycine max*, *Picea abies* and *M. truncatula*. The red arrow shows the cleave site of the signaling peptide.

A comparative *in silico* analysis of the expression of *MtTAX* genes using the MtGEA tool (http://mtgea.noble.org) revealed that except for *MtTAX3* all identified *MtTAX* genes are inducible by MeJA in *M. truncatula* cell suspensions (Fig. 2 A, B). In addition, it was shown that *MtTAX3* and *MtTAX4* are repressed by yeast elicitor (YE) elicitation, while the other *MtTAX* genes are induced by YE elicitation (Fig. 2A & B).

A



B

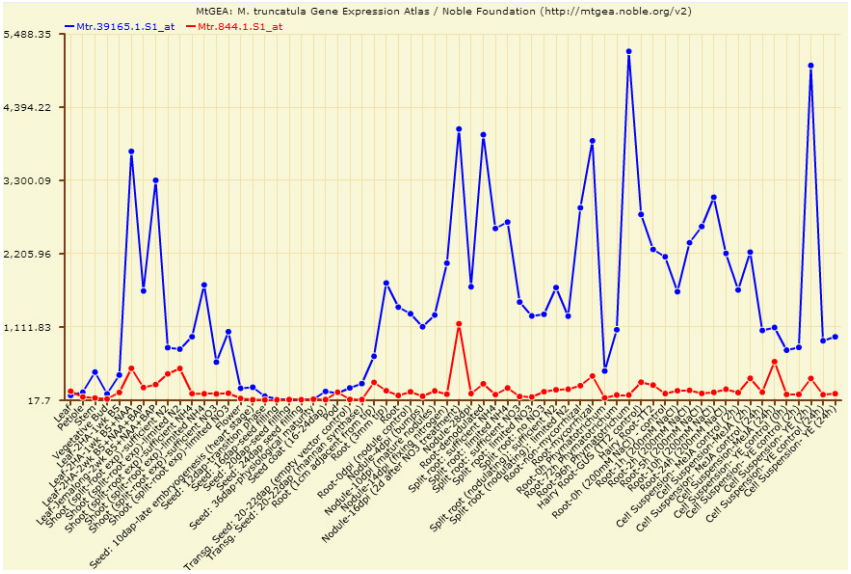


Figure 2. Screenshot of the *in silico* expression analysis of *M. truncatula TAX* genes belonging to A) the first subclade of phylogenetic tree *MtTAX3* (red), *MtTAX4* (black), *MtTAX5* (blue); and B) the second subclade; *MtTAX1* (red); *MtTAX2* (blue) obtained with the MtGEA tool.

2.2 Gain- and loss-of-function of the *MtTAX* genes is associated with prominent phenotypic changes in *M. truncatula* hairy roots

For functional characterization of the *MtTAX* genes, a reverse genetics screen was launched. Gain- and loss-of-function of the candidate genes was carried out by overexpression of the full-length open reading frame (FL- ORF) or of a hairpin RNA-mediated interference (hpRNAi) construct, respectively, under constitutive CaMV35S promoter in *M. truncatula* hairy roots. Moreover, we generated double, triple and quintuple hpRNAi constructs in which the paralogs of each phylogenetic subclade (in

double and triple knockdown) or all five identified *MtTAX* genes (quintuple knockdown) were targeted. In addition, the overexpression construct harboring *T. baccata Taximin* was generated as well (Table 1). The overexpression and downregulation of the *MtTAX* genes in the transgenic hairy roots was confirmed via qRT-PCR analyses and three independent transgenic lines were selected for each construct (Fig. 3A). No knockdown line was obtained for *MtTAX3*, *MtTAX4* and *MtTAX5*, neither after transformation with single hpRNAi constructs, nor after transformation with triple knockdown construct. QRT-PCR analysis of the MtTAX1^{KD}, 2^{KD} and MtTAXQ^{KD} hairy roots showed that from all five *MtTAX* genes only the transcript levels of *MtTAX2* were consistently and slightly decreased (Fig. 3 B & C).

Table 1. List of the constructs designed for the reverse genetics screen of *Taximin* like genes in *M. truncatula*

Construct ID	**Constructs**
Taximin (TB595)	pK7WG2D-TB595
Control	pK7WG2-Venus
MtTAX1 ^{OE}	pK7WG2D-NP7260335
MtTAX2 ^{OE}	pK7WG2D-TC194512
MtTAX3 ^{OE}	pK7WG2D-TC199023
MtTAX4 ^{OE}	pK7WG2D-TC188869
MtTAX5 ^{OE}	pK7WG2D-TC195445
MtTAX1 ^{KD}	pK7GWIWG2D-NP7260335
MtTAX2 ^{KD}	pK7GWIWG2D-TC194512
MtTAX3 ^{KD}	pK7GWIWG2D-TC199023
MtTAX4 ^{KD}	pK7GWIWG2D- TC188869
MtTAX5 ^{KD}	pK7GWIWG2D-TC195445
MtTAX1^{KD}, 2^{KD}	Double: pK7GWIWD2D-TC194512-NP7260335
MtTAX-3,-4,-5^{KD}	Triple: pK7GWIWD2D-TC199023- TC188869-TC195445
MtTAX1-5^{KD} (MtTAX.Q^{KD})	Quintuple:pK7GWIWD2D-TC194512-NP7260335-TC199023- TC195445- TC188869

A

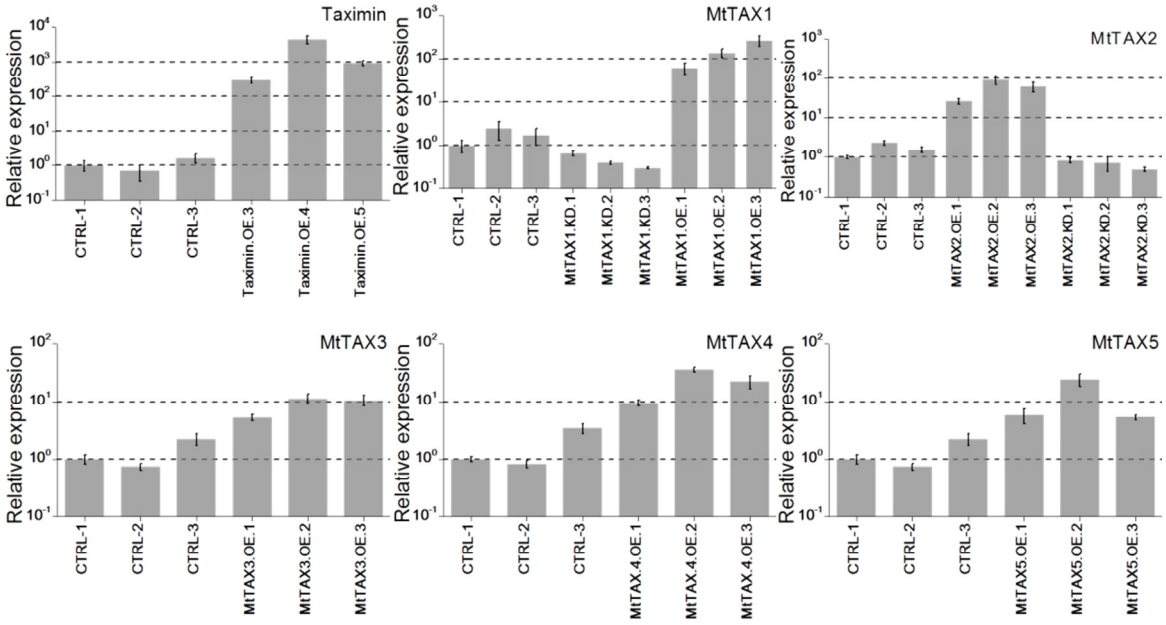


Figure 3A. Expression analyses of the *Taximin* and *MtTAX* genes in their corresponding transgenic hairy root lines. CTRL, control lines; OE, overexpression lines, KD, knockdown lines. Y-axis represents relative expression ratio as compared to the control line 1 (CTRL-1). Error bars represent standard error of mean (SEM) of three technical repeats.

B

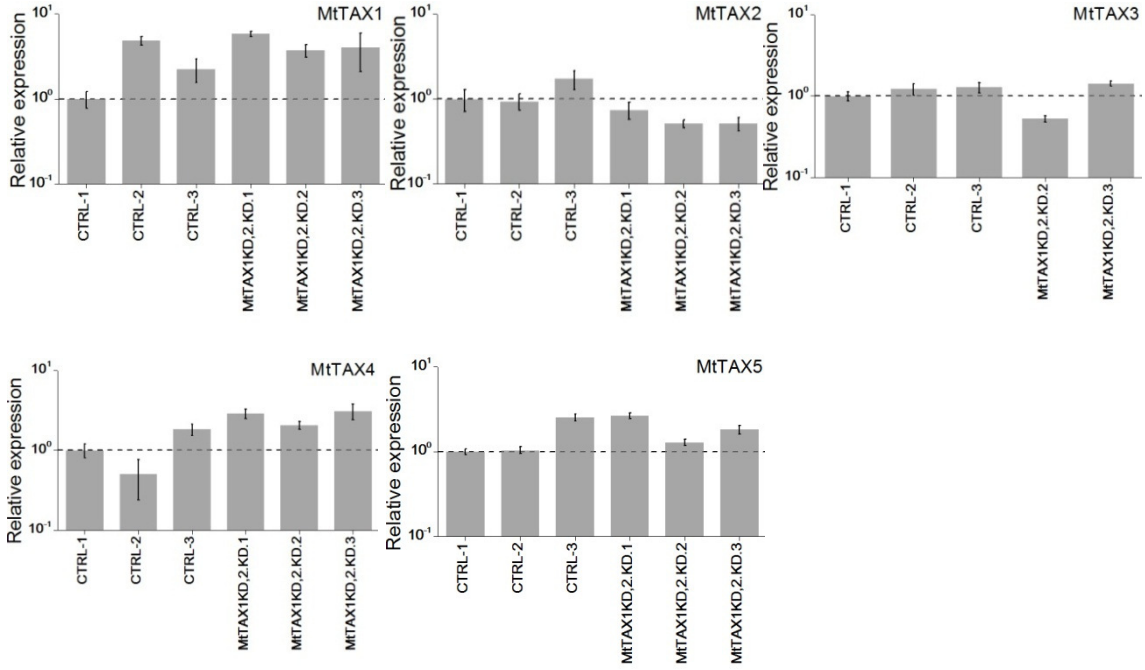


Figure 3B. Expression analyses of the *MtTAX* genes in *M. truncatula* hairy roots transformed with the double knockdown construct (MtTAX1^{KD}, 2^{KD}). CTRL, control lines; KD, knockdown lines. Y-axis represents relative expression ratio as compared to the control line 1 (CTRL-1). Error bars represent SEM of three technical repeats.

C

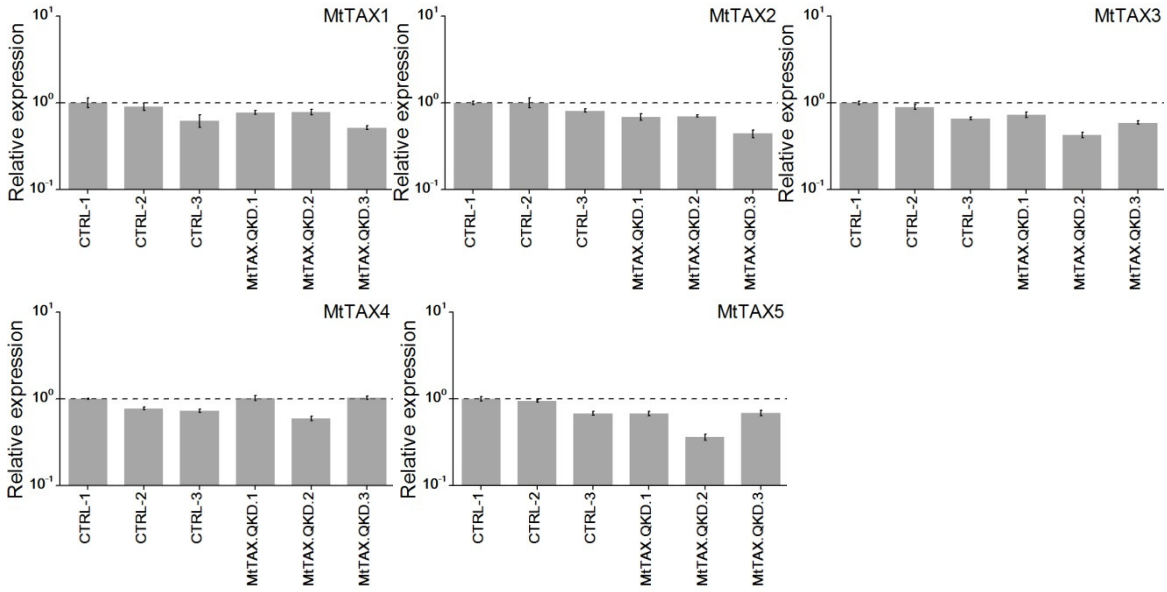


Figure 3C. Expression analyses of the *MtTAX* genes in *M. truncatula* hairy roots transformed with the quintuple knockdown construct (MtTAXQ^{KD}). CTRL, control lines; KD, knockdown lines. Y-axis represents relative expression ratio as compared to the control line 1 (CTRL-1). Error bars represent SEM of three technical repeats.

Overexpression of the *MtTAX4* gene in transgenic hairy roots caused the production of thick, slowly grown and fragile roots that possess no root hairs. The other overexpression constructs developed roots with no obvious phenotype when compared with the control lines. The hairy roots transformed with *MtTAX1* and *MtTAX2* hpRNAi constructs as well as double knockdown (MtTAX1^{KD},2^{KD}) and quintuple knockdown (MtTAXQ^{KD}) constructs produced dissociated and very short roots growing in clumps (Fig. 4). The transgenic lines transformed with the triple hpRNAi construct (MtTAX-3,-4,-5^{KD}) produced roots with severe growth defect, thick, fragile, and with no root hairs.

Based on the observed phenotypes and the results obtained from the expression analysis of the *MtTAX* genes; MtTAX2^{KD}, MtTAXQ^{KD}, MtTAX1^{OE}, MtTAX2^{OE}, MtTAX4^{OE} and *Taximin* overexpressing lines were selected for more investigations.

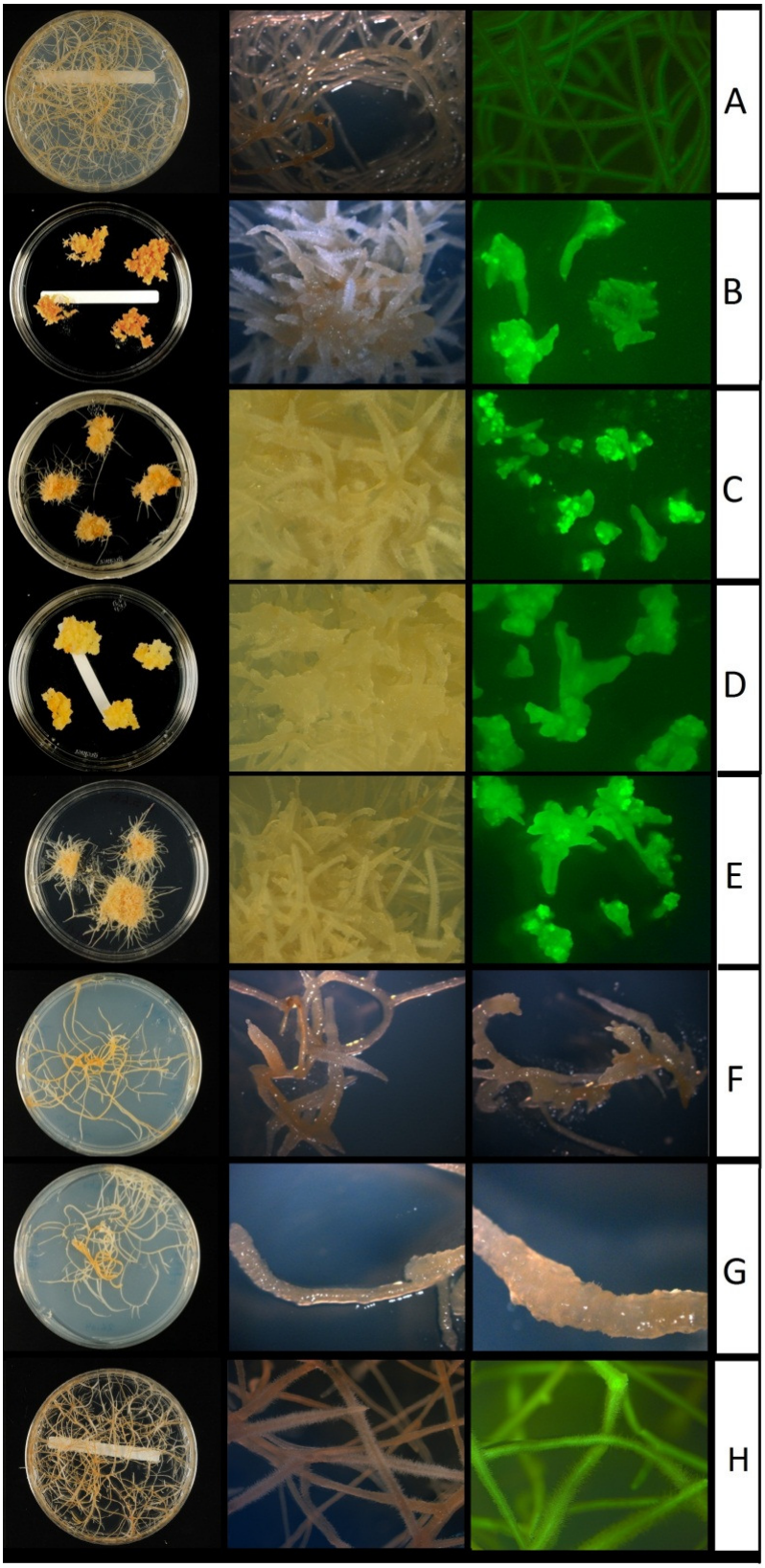


Figure 4. Phenotypic features of the *MtTAX* transgenic lines. Roots of A) control, B) MtTAX1^{KD}, C) MtTAX2^{KD} D) double knockdown [MtTAX1^{KD}, 2^{KD}] E) quintuple knockdown [MtTAXQ^{KD}] F) triple knockdown [MtTAX3 ^{KD},4 ^{KD},5^{KD}] and G) MtTAX4^{OE} H) Taximin ^{OE} lines grown on solid medium. KD: knockdown lines, OE: overexpression lines.

2.3 Ectopic expression of *Taximin* in *M. truncatula* hairy roots enhanced accumulation of glycosylated triterpene saponins

Considering the induction of alkaloid biosynthesis in tobacco plants that ectopically express *Taximin* (Onrubia Ibáñez, 2012), we assessed whether *M. truncatula* secondary metabolism was also affected by the ectopic expression of *Taximin*. We also checked the impact of *Medicago MtTAX* genes on the secondary metabolite production in transgenic roots. We performed metabolite profiling using LC-ESI-FT-ICRMS on *Taximin* expressing roots, MtTAX2^{OE} and MtTAX2^{KD} transgenic hairy roots and compared their metabolite accumulation with that of the control roots. Neither MtTAX2^{OE} nor MtTAX2^{KD} lines showed any significant differences in their metabolite profiles compared to control lines. In contrast, comparative analysis of the root extracts of *Taximin*-expressing and control hairy root yielded a total of 7320 *m/z* peaks, containing 195 peaks which were significantly changed ($p<10^{-5}$) (Fig. 5A). Among the differentially present peaks, the peaks corresponding to triglycosylated saponins were more represented among the up-regulated peaks. These findings suggest that ectopic expression of *Taximin* in *M. truncatula* hairy root cultures can stimulate the triterpene biosynthetic pathway in *M. truncatula*. Some compounds with the higher accumulation level in Taximin^{OE} lines than control lines are listed in Table 2. Because of relatively high variability between samples however, this assay will be repeated in near future to confirm these preliminary observations.

Table 2. Identified triterpene saponins with increased accumulation level in Taximin^{OE} hairy roots of *M. truncatula* compared to control lines.

Compounds	Fold of changes
Hex-Hex-HexA-Aglycone A	3.61
Hex-Hex-HexA-Hederagenin	3.39
Hex-Hex-HexA-Soyasapogenol E	4.40
Malonyl-Hex-Hex-HexA-Hederagenin	2.53
Malonyl-Hex-Hex-HexA-Aglycone A	2.45
Malonyl-Hex-HexA-Bayogenin	3.99
3-Glc-28-Glc-Medicagenic acid	3.51

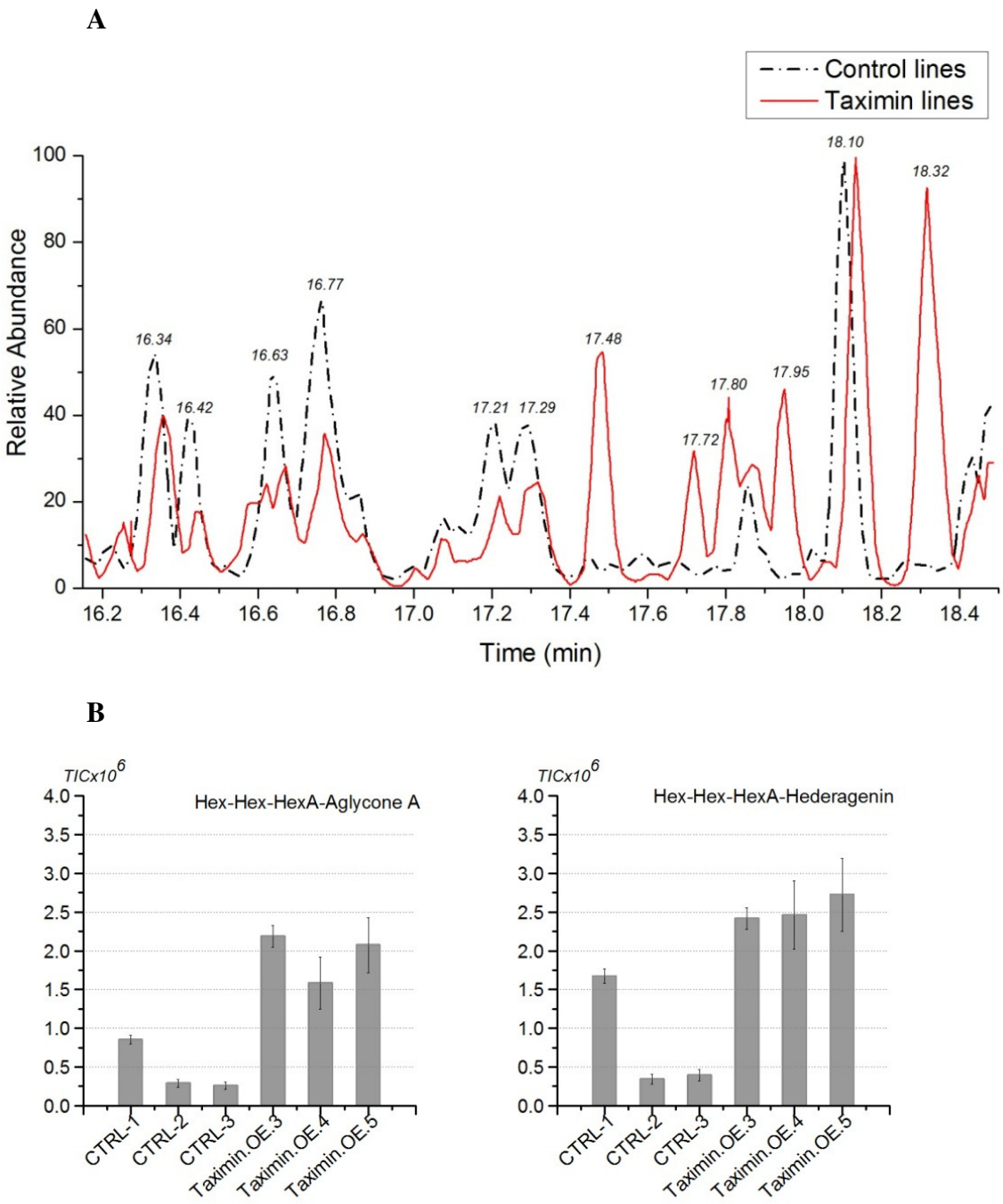


Figure 5 Metabolite profiling of *Taximin*-overexpressing hairy roots. (A) Detail of the full MS scan of a control line (black) and a *Taximin*-overexpressing line (red). CTRL, control lines; OE, overexpression lines. The number indicated on each peak shows the retention time of the corresponding compound. The peak at t_R 17.48 and 17.80 represent Hex-Hex-HexA-Aglycone A and Hex-Hex-HexA-Hederagenin, respectively. (B) Average Total Ion Current (TIC) of the main masses corresponding to Hex-Hex-HexA-Aglycone A and Hex-Hex-HexA-Hederagenin. Error bars represent SEM of five technical repeats.

2.4 Altered expression of *MtTAX* genes affects the expression of biosynthesis genes involved in saponin and flavonoid biosynthesis pathways

Since the expression of the *Taximin* gene from *T. baccata* can activate secondary metabolite biosynthesis genes from *T. baccata* (yew), *Nicotiana tabacum* or *Catharanthus roseus* (Onrubia Ibáñez, 2012), we also explored whether loss- or gain-

of- function of the *MtTAX* genes can alter the expression of secondary metabolite biosynthesis genes in *M. truncatula*. Therefore, the transcript levels of the genes involved in saponin and (iso)flavonoid biosynthesis were measured in transgenic hairy roots expressing *Taximin*, MtTAX1^{OE}, MtTAX2^{OE}, MtTAX4^{OE}, MtTAX2^{KD} and MtTAXQ^{KD} lines. QRT-PCR analysis revealed that despite some fluctuations, expression of none of the (iso)flavonoid biosynthesis genes, nor of the saponin biosynthesis genes, was markedly changed in the lines ectopically expressed *Taximin* (Fig. 6) or *MtTAX* genes (Fig. 7).

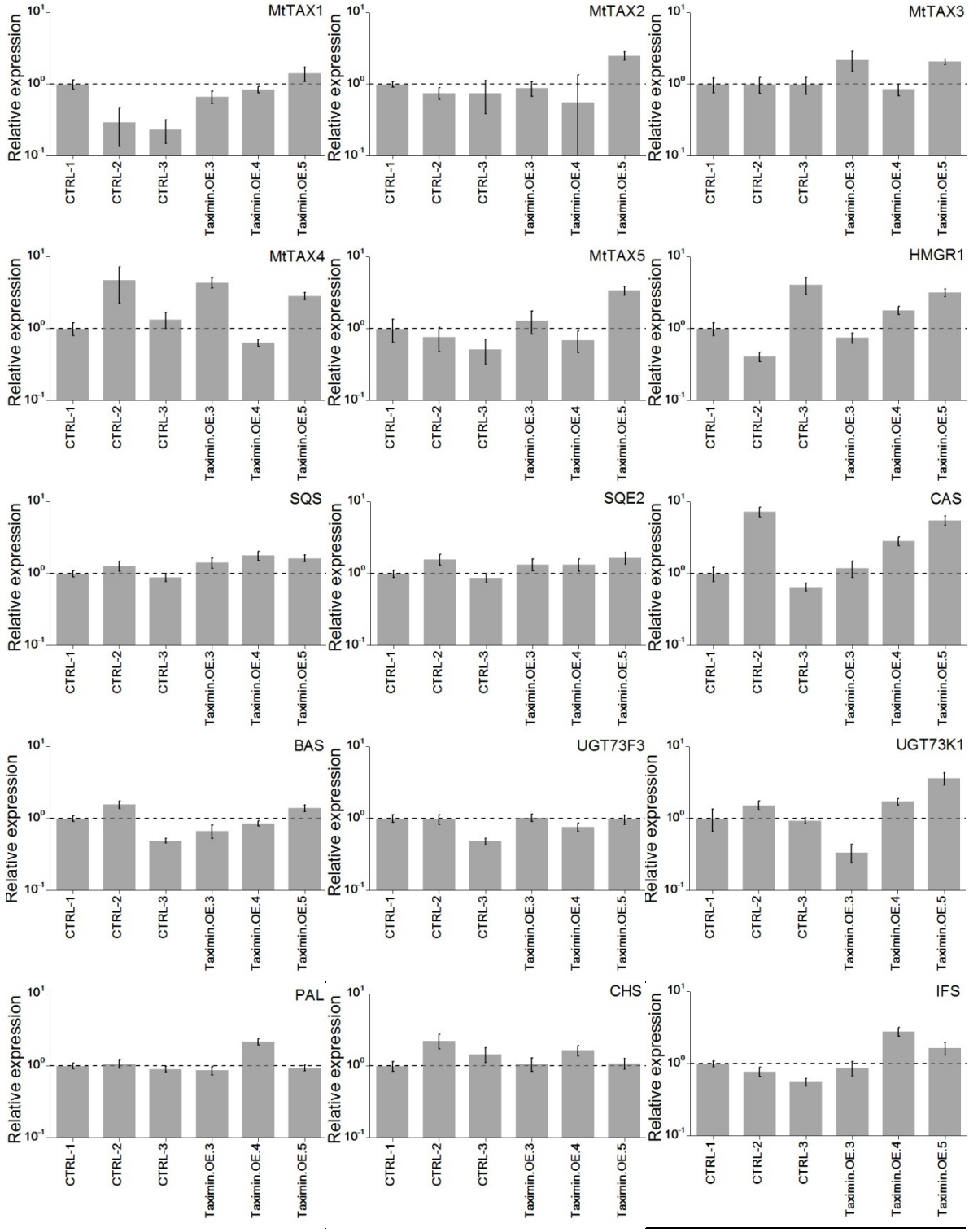


Figure 6. QRT-PCR analysis of biosynthesis genes involved in secondary metabolite in *Taximin*-expressing hairy roots. CTRL, control lines; OE, overexpression lines. Y-axis represents relative expression ratio relative to the CTRL-1 which is set to 1. Error bars represent SEM of three technical repeats.

Figure 7. QRT-PCR analysis of biosynthesis genes involved in secondary metabolite in MtTAX1^{OE}, MtTAX2^{OE} and MtTAX4^{OE} lines. CTRL, control lines; OE, overexpression lines. Y-axis represents relative expression ratio relative to the CTRL-1 which is set to 1. Error bars represent SEM of three technical repeats.

Conversely, the results obtained from the knockdown lines, MtTAX2^{KD} and MtTAXQ^{KD} showed that the steady-state levels of the genes involved in triterpene saponin biosynthesis were decreased (Figs. 8A and 9A). In MtTAX2^{KD} lines, the steady-state levels of the genes downstream of squalene synthase (SQS) in the pathway toward biosynthesis of triterpene saponin including, *SQS*, *squalene epoxide isoform 2* (*SQE2*) and triterpene saponin biosynthesis genes including *β-amyrin synthase* and two saponin-specific glycosyltransferases, *UGT73F3* (Naoumkina et al., 2010) and *UGT73K1* (Achnine et al., 2005) were markedly decreased (Fig. 8A). Similarly, the expression of almost all of the known triterpene saponin biosynthetic genes was strongly downregulated in the MtTAXQ^{KD} lines. *BAS* and *UGT73K1* transcript levels were reduced 3- to 8-fold and the expression of the *CYP93E2* gene, which is the *M. truncatula* gene encoding the β-amyrin 24-hydroxylase (Fukushima et al., 2011), was reduced up to 50-fold (Fig. 9A). In addition, the steady-state levels of *HMGR1* were slightly decreased, the transcript levels of *SQE1* were slightly increased and the *SQS* and *SQE2* expression levels did not change in the MtTAXQ^{KD} lines (Fig. 9A).

The steady state levels of *cycloartenol cyclase* (*CAS*) transcripts corresponding to the key enzyme in sterol biosynthesis which shares the same biosynthetic pathway with triterpene saponins up to the cyclization of squalene epoxide, remains unchanged in both MtTAX2^{KD} and MtTAXQ^{KD} lines (Figs. 8A and 9A).

Among the genes involved in flavonoid biosynthesis the steady state level of *chalcone synthase* (*CHS*) transcripts was slightly decreased in both knockdown lines (Figs. 8B and 9B). Whereas the expression of *UGT78G1* (Modolo et al., 2009), *UGT72L1* (Pang et al., 2008) and *UGT85H2* (Li et al., 2007) that are involved in flavonoid glycosylation did not change in either knockdown lines (Figs. 8B and 9B).

Taken together, overexpression of *MtTAX* genes did not alter saponin gene expression whereas silencing of *MtTAX* genes had a negative effect on saponin gene expression. Surprisingly, this does not correlate with the metabolomics data in which only a positive effect of *Taximin* overexpression on saponin accumulation was observed.

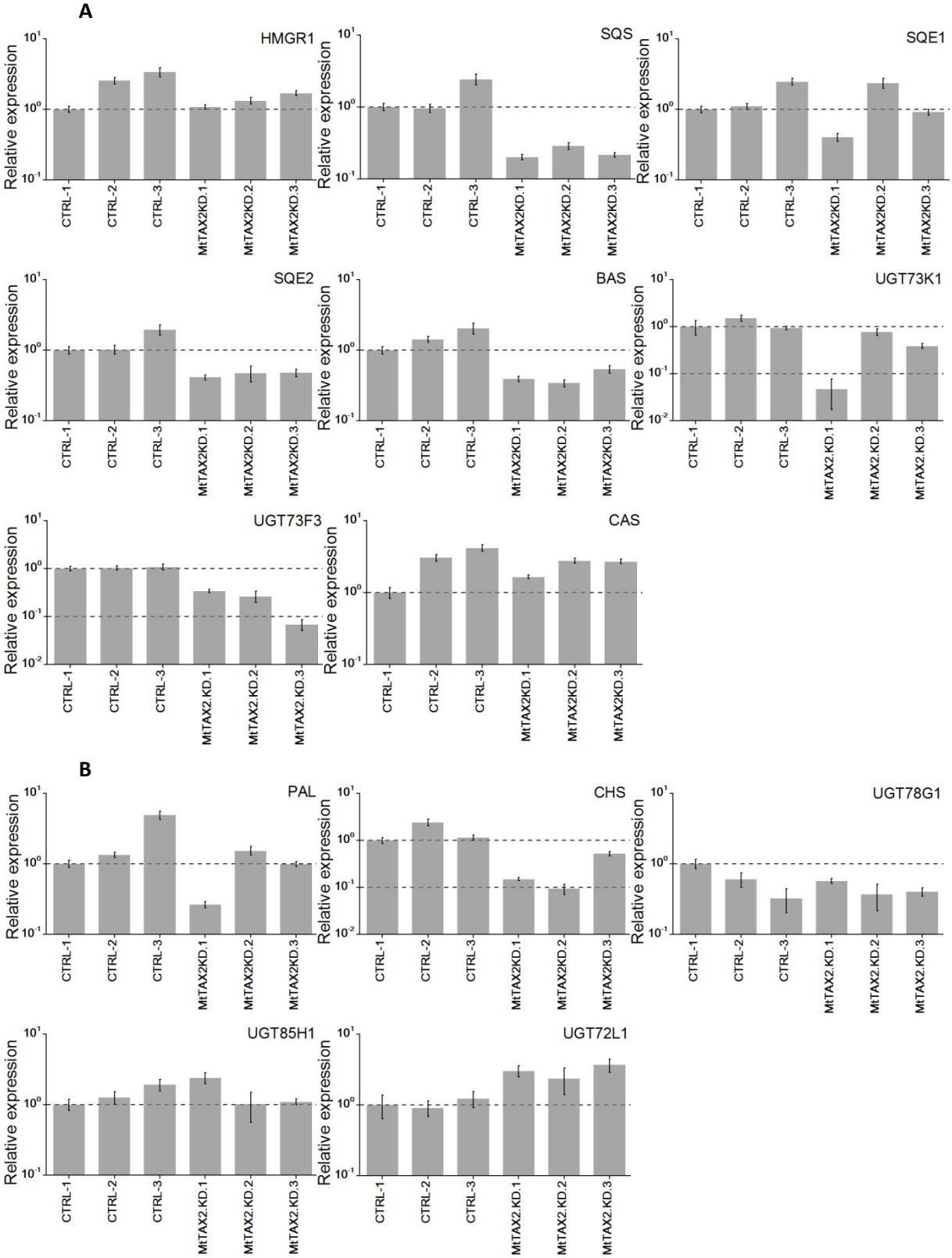


Figure 8. QRT-PCR analysis of biosynthesis genes involved in A) the triterpene pathway toward biosynthesis of triterpene saponins and sterols; B) the flavonoid biosynthesis pathway in *MtTAX2* knockdown ($MtTAX2^{KD}$) *M. truncatula* hairy roots. Y-axis represents relative expression ratio relative to the CTRL-1 which is set to 1. Error bars represent SEM of three technical repeats.

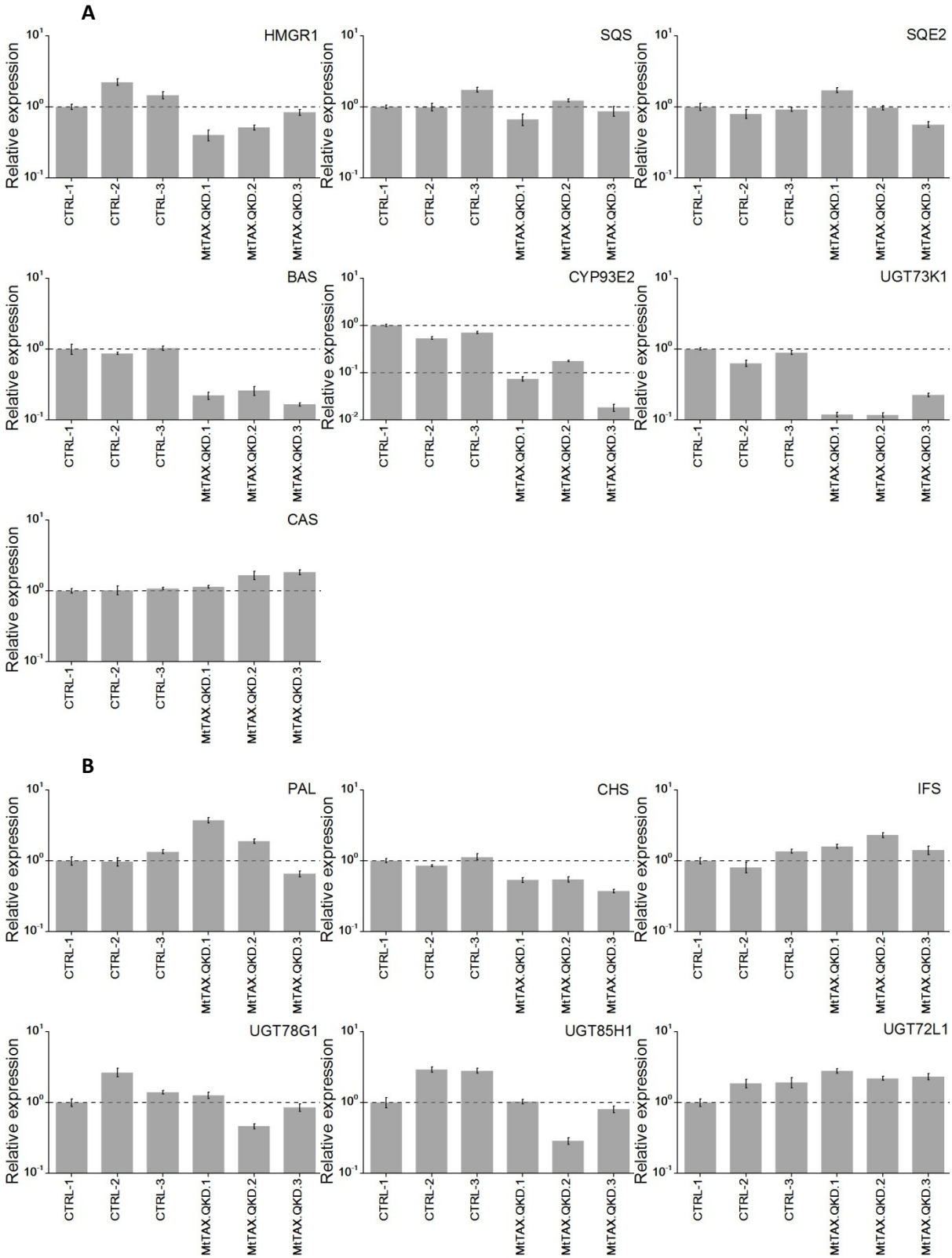


Figure 9. QRT-PCR analysis of biosynthesis genes involved in A) the triterpene pathway toward biosynthesis of triterpene saponins and sterols; B) the flavonoid biosynthesis pathway in quintuple knockdown (MtTAXQ^{KD}) *M. truncatula* hairy roots. Y-axis represents relative expression ratio relative to the CTRL-1 which is set to 1. Error bars represent SEM of three technical repeats.

2.5 RNA-Seq analysis

To obtain a genome-wide overview on the genes modulated in the transgenic *MtTAX* lines, we performed RNA sequencing transcript profiling (RNA-Seq analysis) on MtTAX1^{OE}, the closest homolog of *Taximin* in the phylogenetic tree, and MtTAXQ^{KD} lines. The transcript abundances were measured in fragments per kilobase of exon per million fragments mapped (FPKM) values of the different genes that are present on the genome. Gene annotations were retrieved using NCBI.

We found *MtTAX1*, *MtTAX3* and *MtTAX5* in the RNA-Seq data list, but *MtTAX2* and *MtTAX4* were not present on the *M. truncatula* reference genome and thereby, we were not able to detect them in the RNA-Seq list. As shown in Table 3, *MtTAX1* is highly upregulated in MtTAX1^{OE} lines and is downregulated about 2-fold in MtTAXQ^{KD}, while *MtTAX5* level did not change in any lines and *MtTAX3* transcript levels were about 5-fold higher in MtTAXQ^{KD} lines than in the control lines (Table 3). Surprisingly, these results are not in agreement with the results obtained from qRT-PCR analysis of the MTTAXQ^{KD} line, which showed that the steady-state of *MtTAX2* transcript levels were slightly decreased while the transcript levels of other *MtTAX* genes remain unchanged (Fig. 3B). It should be noted that the *M. truncatula*, genome annotation is currently being reassessed by the *Medicago* genome consortium. When an updated curated genome will be available; we will reanalyze our RNA-Seq data. Meanwhile, we continue the analysis and interpretation with the available version, which may include errors on generated artifact in our interpretations.

From the expressed genes in the list, we selected the genes with ≥5 fold induction (36 in MtTAXQ^{KD} and 2 MtTAX1^{OE} lines) and reduction (16 and 3 in MtTAXQ^{KD} and MtTAX1^{OE} lines, respectively) in their transcript levels (Tables 4, 5 and 6).

Tables 4 and 5 represent the genes with 5-fold induction and reduction in MtTAXQ^{KD} lines, respectively. The genes were grouped based on their putative annotation.

Table 3. Transcript levels of *MtTAX* genes in MtTAX1^{OE} and MtTAXQ^{KD} hairy root cultures (n=3).

Gene_id	Annotation	FPKM- normalized expression counts				
		Control mean ± SD	MtTAX1^{OE} mean ± SD	MtTAXQ^{KD} mean ± SD	Fold of induction in MtTAX1^{OE}	Fold of induction in MtTAXQ^{KD}
MTR_4g078840	MtTAX1: NP260335	12.49±10.3	966.7±196.9	6.900±4.25	77.38	0.552
MTR_4g079680	MtTAX3: TC199023	19.71±3.89	15.98±1.77	91.92±4.01	0.812	4.664
MTR_2g015540	MtTAX5: TC195445	97.55± 4.69	121.2±32.06	134.2±13.98	1.242	1.376

Table 4. Genes significantly induced in the MtTAXQ^{KD} hairy root cultures (n=3).

Gene_id	Annotation	FPKM- normalized expression counts		
		Control mean ± SD	MtTAXQ^{KD} mean ± SD	Fold of induction
Gene with a putative role in cell wall biosynthesis				
MTR_4g122640	Peroxidase	4.22±0.66	25.02±10.13	5.93
MTR_2g044880	Pectinesterase	0.46±0.155	3.28±1.36	7.21
MTR_2g038860	Xyloglucan endotransglucosylase/hydrolase	1.15±1.238	10.91±1.46	9.58
MTR_7g087960	Alpha-L-arabinofuranosidase	0.22±0.25	2.85±1.47	12.72
MTR_7g087880	Alpha-L-arabinofuranosidase	1.00±1.55	15.56±2.80	15.54
MTR_7g087890	Alpha-L-arabinofuranosidase	0.58±0.72	10.09±2.80	17.34
Genes involved in stress response				
MTR_8g074330	Endochitinase	33.46±12.17	190±48.17	5.67
MTR_1g007030	Serine protease-like protein	0.25±0.306	6.60±4.86	26.4
MTR_8g079400	Cc-nbs-lrr resistance protein	5.77±1.3	37.18±15.56	6.44
MTR_7g070220	Dehydration responsive element binding protein	0.22±0.17	3.41±0.702	15.5
MTR_5g010250	Glutamate-cysteine ligase (oxidative stress)	0.303±0.3	2.62±2.34	8.66
MTR_4g081380	Thioredoxin-like protein (oxidative stress)	78.27±57.2	421±35.22	5.38

Signal Transduction				
MTR_3g084390	Two-component response regulator ARR12	0	2.64±0.414	-
MTR_4g092570	Leucine-rich repeat receptor-like protein kinase	0.62±0.513	5.57±2.559	8.97
Primary metabolism genes				
MTR_7g090140	Protein WAX2	0.41±0.309	2.10±0.852	5.15
MTR_5g090070	Alanine glyoxylate aminotransferase	44.60±26.25	240±55.64	5.39
MTR_2g045280	Cytidine deaminase	241±149	1540±395	6.37
MTR_3g036100	Cytokinin dehydrogenase	0.54±0.432	3.51±1.622	6.45
MTR_3g102370	L-asparaginase	2.18±0.65	15.76±2.42	7.23
MTR_4g118350	Pyruvate orthophosphate dikinase	0.63±0.41	8.60±7.219	13.7
Genes with no known function				
MTR_4g062160	Hypothetical protein	1.04±0.86	9.33±1.84	8.98
MTR_7g072260	Hypothetical protein	0.67±0.63	4.91±2.33	7.30
MTR_1g098470	Hypothetical protein	3.99±0.28	23.73±10.6	5.95
MTR_8g092240	Hypothetical protein	0.96±0.79	5.09±1.60	5.32
MTR_8g094560	Hypothetical protein	0	14.05±7.03	-
MTR_8g037760	Hypothetical protein	0	10.61±8.28	-
MTR_4g014400	Hypothetical protein	0	6.41±1.01	-
MTR_4g119460	Hypothetical protein	0	3.24±2.55	-
MTR_3g052400	Hypothetical protein	0	3.12±1.66	-
MTR_2g007860	Hypothetical protein	0	1.86±0.29	-
MTR_3g014420	Hypothetical protein	0	1.88±1.77	-
MTR_7g086040	MtN20 protein	0	3.16±1.30	-
Transporters				
MTR_2g005130	Peptide transporter PTR1	1.54±0.538	9.78±0.54	6.38
MTR_2g026160	Inositol transporter	0.40±0.315	3.98±1.054	9.95
Others				

MTR_5g063970	Zinc finger MYM-type protein	0.107±0.092	1.98±1.27	18.42

Table 5. Genes significantly reduced in the MtTAXQ^{KD} hairy root cultures (n=3).

Gene_id	Annotation	FPKM- normalized expression counts		
		Control mean ± SD	MtTAXQ^{KD} mean ± SD	Fold of induction
Gene with a putative role in cell wall biosynthesis				
MTR_4g091010	AM3 (cell wall macromolecule catabolic process)	40.1±11.7	6.99±0.09	0.174
MTR_3g064510	Expansin	1.35±1.31	0	0
MTR_7g023730	Polygalacturonase inhibitor protein	0.508±0.26	0.05±0.07	0.091
MTR_2g008226	UDP-glucosyltransferase	6.41±0.20	0.83±0.03	0.129
MTR_2g008225	Glucosyltransferase	6.35±2.48	1.07±0.47	0.168
Secondary metabolite biosynthesis genes				
MTR_4g031800	UGT73K1	101±16.3	21.97±0.16	0.218
MTR_4g005190	Beta-amyrin synthase	32.12±4.4	7.21±1.64	0.224
MTR_4g022290	Anthranilate N-benzoyltransferase	8.06±1.59	2.01±0.47	0.248
MTR_7g084300	Chalcone synthase	5.52±4.50	0.36±0.16	0.066
Genes with no known function				
MTR_3g016250	Hypothetical protein	2.53±0.62	0	0
MTR_7g074750	Hypothetical protein	1.70±1.64	0.073±0.11	0.043
MTR_4g039740	Hypothetical protein	7.32±2.86	0.67±0.54	0.091
MTR_6g010810	Hypothetical protein	1.81±0.74	0.22±0.08	0.121
MTR_4g059670	Hypothetical protein	64.12±11.9	11.36±2.82	0.177
Others				
MTR_6g013200	Lamin-like protein	1.43±0.57	0.13±0.19	0.089
MTR_1g087230	Multidrug resistance protein ABC transporter family protein	2.90±0.68	0.50±0.39	0.171

Table 6. Genes significantly altered in the MtTAX1^{OE} hairy root cultures sorted according to fold of induction (n=3).

		FPKM- normalized expression counts		
Gene_id	**Annotation**	**Control mean ± SD**	**MtTAX1OE mean ± SD**	**Fold of induction**
MTR_4g078840	MtTAX1: NP260335	12.49±11.55	966±197	77.38
Up-regulated genes				
MTR_3g117160	Defensin	0	1.67±0.37	-
MTR_4g059760	Hypothetical protein	0	4.22±1.46	-
Down-regulated genes				
MTR_5g061610	Hypothetical protein	10.11±7.41	0	0
MTR_4g108470	ATP synthase subunit a chloroplastic	2.08±2.3	0	0
MTR_4g018820	Jasmonate *O*-methyltransferase	31.32±16.23	3.41±0.85	0.11

The results obtained from RNA-Seq analysis revealed that several proteins involved in cell wall function, defense and stress response and some genes encoding proteins with a putative role in signal transduction were upregulated in MtTAXQ^{KD} lines (Table 4). Among the downregulated genes in MtTAXQ^{KD} lines, we found known genes encoding some key enzymes of secondary metabolite pathways. These include the genes involved in flavonoid biosynthesis (*chalcone synthase* and *anthranilate N-benzoyltransferase*) and saponin biosynthesis (*Beta-amyrin synthase* and *UGT73K1*) (Table 5), which is in accordance with the results obtained from qRT-PCR analysis (Fig. 9). Expression of several UGTs was decreased in MtTAXQ^{KD} (Table 7), among which the saponin specific UGT, *UGT73K1* (Achnine et al., 2005). Therefore, we specifically looked for other hitherto known UGTs that are functionally characterized as the UGTs involved in glycosylation of flavonoid and saponin compounds in *M. truncatula*. The data showed that transcript levels of the saponin specific UGTs including *UGT73F3*, *UGT73K1* and *UGT71G1* were decreased; whereas the transcript levels of flavonoid-specific UGTs including *UGT85H2*, *UGT78G1* and *UGT72L1* did not change in MtTAXQ^{KD} lines (Table 7). These results confirmed the results obtained from qRT-PCR analysis of MtTAXQ^{KD} and MtTAX2^{KD} hairy root cultures (Figs. 8 and 9).

Table 6 shows the genes with the altered transcript levels in MtTAX1^{OE} lines. We found only five genes with a more than 5-fold alteration in transcript levels in MtTAX1^{OE} lines (2 upregulated and 3 downregulated genes). *Defensin* and *jasmonate O-methyltransferase* (*JMT*) were among the genes with altered transcript levels in MtTAX1^{OE} lines. The former is a member of a large gene family encoding putative Cys-rich defense protein (Hanks et al., 2005) and the latter is involved in MeJA biosynthesis.

Next, we looked for genes that were altered in opposite directions in MtTAX1^{OE} and MtTAXQ^{KD}. Twenty four genes were oppositely altered (Table 8). Again, some known genes encoding key enzymes of secondary metabolite pathways were presented in this list. These include *chalcone synthase*, *flavanone 3-hydroxylase* involved in flavonoid biosynthetic pathway, *anthranilate N-benzoyltransferase* gene encoding the enzyme catalyzing the first committed reaction of phytoalexin biosynthesis, and phytoene synthase that catalyzes the first committed step toward carotenoid biosynthesis (Table 8).

Table 7. The UGTs mentioned in this study (n=3).

Gene_id	Annotation	FPKM- normalized expression counts				
		Control mean ± SD	MtTAX1^{OE} mean ± SD	MtTAXQ^{KD} mean ± SD	Fold of induction in MtTAX1^{OE}	Fold of induction in MtTAXQ^{KD}
UGTs involved in saponin glycosylation						
MTR_5g070090	UGT71G1	19.74±5.48	16.86±4.62	11.39±1.5	0.85	0.58
MTR_4g031800	UGT73K1	100±16.3	67.60±5.06	21.97±0.16	0.67	0.22
MTR_2g035020	UGT73F3	40.70±5.45	41.37±3.51	14.55±1.17	1.02	0.36
UGTs involved in flavonoid glycosylation						
MTR_6g014290	UGT85H2	0.25±0.17	0.201±0.15	0.10±0.09	0.80	0.39
MTR_4g128670	UGT78G1	22.81±4.76	22.78±3.91	25.36±4.92	0.99	1.11
MTR_8g009070	UGT72L1	27.39±1.85	27.00±2.16	32.62±3.89	0.95	1.19
Other UGTs with altered transcript levels						
MTR_2g008226	UDP-glucosyltransferase	6.41±0.20	7.91±1.87	0.83±0.03	1.23	0.13
MTR_2g008225	Glucosyltransferase	6.35±2.48	14.97±7.53	1.07±0.47	2.36	0.17
MTR_7g102450	Cytokinin-*O*-glucosyltransferase	0.84±0.23	0.24±0.082	2.36±0.26	2.91	0.29
MTR_5g035580	*N*-hydroxythioamide *S*-beta-glucosyltransferase	1.350±1.01	5.44±2.21	0.33±0.10	4.03	0.24

Table 8. Genes with transcript levels altered in opposite directions in MtTAX1^{OE} and MtTAXQ^{KD} hairy roots (n=3).

		FPKM- normalized expression counts				
Gene_id	**Annotation**	**Control mean ± SD**	**MtTAX1^{OE} mean ± SD**	**MtTAXQ^{KD} mean ± SD**	**MtTAX1^{OE} fold of induction**	**MTTAXQ^{KD} fold of induction**
MTR_4g078840	MtTAX1:NP260335	12.49±10.3	966.7±196.9	6.90±4.25	77.38	0.55
Gene with a putative role in cell wall biosynthesis						
MTR_2g089140	Xyloglucan endotransglucosylase/hydrolase	172±0.59	82.24±0.072	366±1.41	0.478	2.13
Genes involved in secondary metabolism						
MTR_7g084300	Chalcone synthase	5.52±4.50	18.558±16.5	0.37±0.17	3.36	0.07
MTR_4g022290	Anthranilate N- benzoyltransferase	8.08±1.58	17.23±5.82	2.01±0.48	2.13	0.25
MTR_5g076620	Phytoene synthase	8.98±0.70	24.97±10.04	3.22±0.85	2.78	0.36
MTR_8g075830	Flavanone 3-hydroxylase	2.50±0.53	5.74±1.60	1.27±0.22	2.30	0.51
Transcription factors						
MTR_5g014520	BHLH transcription factor	2.90±0.21	7.21±1.19	0.61±0.10	2.49	0.21
MTR_1g106430	BTB/POZ domain-containing protein	46.10±8.05	128±35.2	24.19±2.52	2.77	0.53
Genes involved in primary metabolism						
MTR_8g031850	Carbonyl reductase	9.73±5.42	30.20±10.92	1.26±0.94	3.10	0.13
MTR_4g017200	Pyruvate decarboxylase isozyme	30.40±3.66	164±58.4	10.79±2.23	5.41	0.36
Genes with no known function						
MTR_6g006210	Hypothetical protein	2.45±1.07	6.54±0.47	0.42±0.65	2.67	0.17
MTR_8g071920	Hypothetical protein	13.58±2.4	31.95±14.27	3.60±0.94	2.35	0.27

MTR_1g090630	Hypothetical protein	4.39±1.82	14.39±6.47	2.18±0.54	3.28	0.50
MTR_7g073390	Hypothetical protein	1.67±1.2	0.40±0.20	3.80±1.02	0.24	2.29
MTR_7g073410	Hypothetical protein	1.67±1.2	0.40±0.20	3.80±1.026	0.24	2.29
Transporters						
MTR_1g084720	ABC transporter family protein	44.1±14.8	119 ±53.41	23.46±1.33	2.70	0.53
MTR_3g099700	Ripening regulated protein DDTFR18 (antiporter activity)	1.61±0.28	0.41±0.071	3.30±1.41	0.26	2.05
Glycosyltransferases						
MTR_5g035580	*N*-hydroxythioamide *S*-beta-glucosyltransferase	1.35±1.01	5.44±2.21	0.33±0.10	4.03	0.24
MTR_7g102450	Cytokinin-*O*-glucosyltransferase	0.84±0.23	0.24±0.08	2.36±0.26	2.91	0.30
Others						
MTR_3g108710	Cryptochrome DASH (Photolyase)	2.34±1.12	6.17±1.51	0.30±0.12	2.64	0.127
MTR_5g074680	*O*-methyltransferase	0.84±0.23	0.24±0.082	2.36±0.26	2.38	0.13
MTR_5g033520	Alliin lyase	0.25±0.05	0	1.10±0.51	0.29	2.81
MTR_5g020250	2-hydroxy-6-oxononadienedioate hydrolase	4.74±1.07	13.81±6.31	1.389±0.21	0	4.481

3 Discussion

Taximin was previously reported as a MeJA-inducible small peptide encoding gene in *T. baccata* cell cultures. *Taximin* modulates the expression of genes involved in secondary metabolism and consequently secondary metabolite production (Onrubia Ibáñez, 2012). In this study, we explored the effect of ectopically expressed *Taximin* on the secondary metabolism of *M. truncatula* hairy roots. In addition, we describe the cloning and characterization of *Taximin* homologs, the *MtTAX* genes, from *M. truncatula* and investigated their role in the regulation of *M. truncatula* metabolism.

Five *Taximin* homolog genes were identified in *M. truncatula* and named as *MtTAX1* to *MtTAX5*. According to the phylogenetic study these genes fell into two subclades, the first subclade constitutes of *MtTAX-3, -4, -5*, whereas *MtTAX1* and *-2* fell in the second subclade, to which also *Taximin* belongs. Gain- or loss function studies in hairy roots showed that downregulation of the *MtTAX* genes and overexpression of *MtTAX4* was often associated with growth defect. Overexpression of the other *MtTAX* genes or *Taximin* did not affect general growth and morphology of *M. truncatula* hairy roots.

RNA-Seq analysis indicated that several genes involved in cell wall modification were up- and down-regulated in the MtTAXQ^{KD} plants (Table 4 and 5), which may be related with the mutant phenotype. In addition, several defense- and stress-related genes were coordinately induced in MtTAXQ^{KD} hairy root cultures (Table 4).

The plant cell wall is a complex structure composed of polysaccharides, phenolic compounds, and proteins. The wall proteins are involved in at least three categories of cell wall functions. The cell wall modifying proteins contribute to modification of matrix polysaccharides (Darley et al., 2001). Structural proteins involved in wall architecture (Cassab, 1998), and defense proteins help in responses to biotic and abiotic stresses (Baluška et al., 2003; Qin et al., 2003).

Cell wall modifying proteins include various endoglycanases that may cleave the backbone of matrix polysaccharides; glycosidases that may remove side chains, and therefore allowing greater interactions between polysaccharide backbones; glycosyltransferases that may cut polysaccharides and ligate them together, esterases that remove methyl groups from pectins and cleave ester linkages between polysaccharide chains, and peroxidases that may form or break phenolic linkages in the wall (Cosgrove, 2001; Jamet et al., 2006). All these enzymes cause wall-loosening

required for plant cell expansion. Transcription analysis of MtTAXQ^{KD} hairy roots showed that the steady-state levels of the gene transcripts corresponding to the saponin-specific UGTs, *UGT73K1*, *UGT73F3* and *UGT71G1* were decreased in MtTAXQ^{KD} lines, whereas the flavonoid specific UGTs including *UGT85H2*, *UGT78G1* and *UGT72L1* did not alter in these knockdown lines (Figs. 8 and 9, Table 7).

Saponins are able to form a complex with sterols and change membrane permeability. since monodesmosides saponins have generally a stronger membrane activity than bidesmosides (Tava and Avato, 2006), it is possible that sugar side chains on the aglycone structures have a role on the effects of saponin on cell membranes. Taking into account the structural resemblance of membrane sterols, it can be speculated that triterpene saponins can interfere with membrane-related processes within cells (Osbourn et al., 2011). For instance, it is reported that accumulation of incompletely glycosylated triterpene saponin in oat mutants results in disruption of membrane trafficking that is associated with growth and developmental defects in plant (Mylona et al., 2008). Therefore, it is possible that accumulation of less glycosylated saponin in *MtTAX* knockdown lines, due to downregulation of saponin specific glycosyltransferases, interferes with cell membrane trafficking and probably cause defects in the cell wall biogenesis and integrity. Changes in cell wall integrity can act as a signal for the activation of defense and stress responses (Vorwerk et al., 2004). However, further metabolite and microscopic evidences are needed to explore these relationships and better establish whether it is truly causal.

Higher alkaloid production in tobacco hairy roots expressing the *Taximin* gene was reported previously (Onrubia Ibáñez, 2012). Here we provide preliminary data that *Taximin*-expressing *M. truncatula* hairy roots may accumulate higher levels of glycosylated saponins than control hairy roots (Fig. 5). These results suggest that *Taximin* may modulate accumulation of saponin glycosides in *M. truncatula* hairy roots. However no changes have been observed on the steady state levels of the saponin biosynthesis gene transcript and therefore, the mechanism of this potential regulatory role is still unclear.

Conversely, qRT-PCR analysis of the cDNA obtained from MtTAX2^{KD} and MtTAXQ^{KD} lines showed decreases in the steady state levels of the transcripts corresponding to genes involved in saponin biosynthesis (Figs. 8A and 9A). In addition, RNA-Seq analysis of the MtTAX1^{OE} and MtTAXQ^{KD} lines showed that the

transcript levels of several genes involved in different secondary metabolite biosynthesis pathways including triterpene saponin, flavonoid, phytoalexin and carotenoid biosynthesis pathways were increased in MtTAX1^{OE} and decreased in MtTAXQ^{KD} hairy roots. However, metabolite profiling of MtTAX2^{OE} and MtTAX2^{KD} lines revealed no changes in their metabolite levels. Therefore, although the results indicated that *Taximin* and its *MtTAX* homologs may have a role in modulating secondary metabolism in *M. truncatula*, the possible mechanism(s) of such regulatory functions are still unknown. It was previously suggested that Taximin is a pathogenesis-related Cys-rich peptide. Members of CRP can be either produced constitutively as a first line defense in cell surface of nutrient-rich organs like flowers and seeds, or induced by several defense or stress-signaling pathways such as the jasmonate/ethylene signaling pathway (Silverstein et al., 2007). Here, *in silico* analysis revealed that *MtTAX1* and *MtTAX2* are induced in MeJA elicited cell cultures of *M. truncatula*. Collectively, based on these results and according to the high structural similarity of MtTAX peptides with Taximin (Fig. 1), it is tempting to speculate that *Medicago TAX* genes encode Cys-rich peptides involved in plant defense systems as well. However, further analyses are clearly needed to allow determining how the *MtTAX* genes connect with plant secondary metabolism.

4 Material and methods

4.1 Generation of DNA constructs

The FL-ORF sequences of *MtTAX* genes were derived from the *Medicago truncatula* Gene Index database (http://compbio.dfci.harvard.edu/tgi/), PCR-amplified and cloned into the entry vector pDONR221 by Gateway^{TM} recombination. To obtain entry clones with stop codons, Gateway primers were designed according to Underwood et al. (Underwood et al., 2006). All entry constructs were sequence-verified. For the overexpression experiments, Gateway recombination was carried out with the pK7WG2D binary vector (Karimi et al., 2002). For hpRNAi, specific fragments of the *MtTAX* genes were PCR-amplified and by Gateway^{TM} recombination cloned into the binary vector pK7GWIWG2D(II) (Karimi et al., 2002). To generate double, triple and quintuple hpRNAi constructs we used fusion PCR. For the double knockdown construct containing MtTAX1RNAi and MtTAX2RNAi, two parallel PCR amplifications were performed on pDONR221- MtTAX1RNAi and pDONR221-MtTAX2RNAi constructs using the primer pairs TAX025 + TAX027 and TAX020 +

TAX026 (Table 9), respectively. The gel purified PCR fragments were then used as a template to generate a fused fragment using the primer pair TAX020+TAX025 (Table 9). Accordingly, to create triple knockdown constructs containing MtTAX3RNAi, MtTAX4RNAi and MtTAX5RNAi, three parallel PCR amplifications were carried out on pDONR221- MtTAX3RNAi, pDONR221- MtTAX4RNAi and pDONR221- MtTAX5RNAi using the primer pairs TAX016 + TAX028, TAX031 + TAX023 and TAX029 + TAX030 (Table 9), respectively. Further, a PCR fusion amplification was performed on gel purified PCR fragments using the primer pairs TAX016 + TAX023 (Table 9). The quintuple knockdown construct was generated using the gel purified double and triple knockdown constructs via two PCR amplification reactions. First two parallel PCR amplifications were performed on double and triple knockdown constructs using the primer pairs TAX020 + TAX032 and TAX033 + TAX023 (Table 9), respectively. Then, the PCR amplification was performed on the gel purified first PCR product using the primer pairs TAX020 + TAX023 (Table 9). The fused PCR fragments were then cloned into the binary vector pK7GWIWG2D(II) by Gateway™ recombination (Karimi et al., 2002). The resulting clones were transformed into the *A. rhizogenes* strain LBA 9402/12 for generation of hairy roots.

Table 9. Primers used in this study.

Name	Sequence	Description
Generation of overexpression constructs		
TAX001	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTGCAATTCAGAAGGTGAATGC	Fw: AttB1- MtTAX3
TAX002	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAGCAGGGGATCTTAGAGGTG	Rv: AttB2- MtTAX3
TAX003	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTGTTGCTGTTGTGGTGAAG	Fw: AttB1- MtTAX5
TAX004	GGGGACCACTTTGTACAAGAAAGCTGGGTACTAACATGGAATCTTGGATGTG	Rv: AttB2- MtTAX5
TAX005	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCAGATGATGAAGGTTGTG	Fw: AttB1- MtTAX2
TAX006	GGGGACCACTTTGTACAAGAAAGCTGGGTACTAACAAGGAATTTGTCGAG	Rv: AttB2- MtTAX2
TAX007	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGTGTAATTCAAATGGTGATTGC	Fw: AttB1- MtTAX4
TAX008	GGGGACCACTTTGTACAAGAAAGCTGGGTATCAACATGGAATCTTAGAGG	Rv: AttB2- MtTAX4
TAX009	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGCAAGTGACCAAGACTTTTG	Fw: AttB1- MtTAX1
TAX010	GGGGACCACTTTGTACAAGAAAGCTGGGTACTAACAAGGAATTTTGTTAAC	Rv: AttB2- MtTAX1
Generation of RNAi constructs		
TAX016	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGATCCCCTGCTGAAGATAG	Fw:AttB1- MtTAX3
TAX017	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTAATAAACAAATTAATTCAC	Rv:AttB2- MtTAX3
TAX018	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGGTTATGGAGTGGTTCAC	Fw:AttB1- MtTAX5
TAX019	GGGGACCACTTTGTACAAGAAAGCTGGGTACTAATAGTTTAAACTAATGC	Rv:AttB2- MtTAX5
TAX020	GGGGACAAGTTTGTACAAAAAAGCAGGCTTACTCGACAAATTCCTTGTTAG	Fw:AttB1- MtTAX2
TAX021	GGGGACCACTTTGTACAAGAAAGCTGGGTAGAAAGAATCAAATATTCACC	Rv:AttB2- MtTAX2
TAX022	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAAGATTCCATGTTGATGTTAG	Fw:AttB1- MtTAX4
TAX023	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTTCTCATAAACAATATCAG	Rv:AttB2- MtTAX4
TAX024	GGGGACAAGTTTGTACAAAAAAGCAGGCTTATGTTTGATAAAGCTTCCAG	Fw:AttB1- MtTAX1
TAX025	GGGGACCACTTTGTACAAGAAAGCTGGGTAGTTTCAATTGAGAAATTCCC	Rv:AttB2- MtTAX1
Generation of fused RNAi constructs		
TAX026	CTGGAAGCTTTATCAAACAGAAAGAATCAAATATTCACC	Rv for Double KD
TAX027	GGTGAATATTTGATTCTTTCTGTTTGATAAAGCTTCCAG	Fw for Double KD
TAX028	GTGAACCACTCCATAACCTGTAATAAACAAATTAATTC	Rv1 for Triple KD
TAX029	GAATTAATTTGTTTATTACAGGTTATGGAGTGGTTCAC	Fw1 for Triple KD
TAX030	CTAACATCAACATGGAATCTCTAATAGTTTAAACTAATGC	Rv2 for triple KD
TAX031	GCATTAGTTTAAACTATTAGAGATTCCATGTTGATGTTAG	Fw2 for triple KD
TAX032	CTATCTTCAGCAGGGGATCTGTTTCAATTGAGAAATTCCC	Rv for quintuple KD
TAX033	GGGAATTTCTCAATTGAAACAGATCCCCTGCTGAAGATAG	Fw for quintuple KD

4.2 Generation and cultivation of transgenic *M. truncatula* hairy roots

Protocol for *A. rhizogenes*-mediated transformation of *M. truncatula* (ecotype Jemalong J5) hairy roots was adapted from Boisson-Dernier *et al.* (Boisson-Dernier et al., 2001) with modifications and performed as described in chapter 4 page 130-1.

4.3 Transcript profiling

For quantitative Real Time PCR (qRT-PCR), RNA from were isolated with the RNeasy mini kit (Qiagen) according to manufacturer instruction. Quality control and quantification were performed with a Nanodrop spectrometer (Isogen, Hackensack, NJ). cDNA were prepared using SuperscriptTM II Reverse Transcriptase (Invitrogen). Primers were designed with Beacon Designer version 4.0 (Premier Biosoft International, Palo Alto, CA, USA). qRT-PCR was carried out as described in chapter 4 page 131.

4.4 Metabolite extractions

M. truncatula hairy roots were harvested in five biological repeats of three independent transgenic lines per transgene construct and washed with purified water under vacuum filtration. Metabolite extraction was performed as described in chapter 4, page 131.

4.5 LC ESI FT-ICR MS

For reversed-phase LC, an Acquity UPLC BEH C18 column (150 x 2.1 mm, 1.7 µm; Waters, Milford, MA) was mounted on an ultra-high-performance LC system consisting of a Accela pump (Thermo Electron Corporation, Waltham, MA, USA) and Accela autosampler (Thermo Electron Corporation). The Accela LC system was hyphenated to a LTQ FT Ultra (Thermo Electron Corporation) via an electrospray ionization source. The following gradient was run using water:MeCN (99:1, v:v) acidified with 0.1% (v:v) HOAc (solvent A) and MeCN:water (99:1, v:v) acidified with 0.1% (v:v) HOAc (solvent B): time 0 min, 5% B; 30 min, 55% B; 35 min, 100% B. The loop size, flow, and column temperature were 25 µL, 300 µL/min and 80°C, respectively. Full loop injection was applied. Negative ionization was obtained with the following parameter values: capillary temperature 150°C, sheath gas 25 (arbitrary units), aux. gas 3 (arbitrary units), and spray voltage 4.5 kV. Full FT-MS spectra between m/z 120– 1400 were recorded at a resolution of 100,000. For identification, full MS spectra were interchanged with a dependent MS^2 scan event in which the most abundant ion in the previous full MS scan was fragmented, and two dependent MS^3 scan events in which the two most abundant daughter ions were fragmented. The collision energy was set at 35%. Elucidation of the MS^n spectra was according to Pollier et al (Pollier et al., 2011) for the saponins and Morreel et al. (Morreel et al., 2006; Morreel et al., 2010) for the flavonoids and (neo-)lignans. The resulting chromatograms were integrated and aligned with the XCMS package (Smith et al., 2006) in R version 2.6.1. with the following parameter values: xcmsSet (fwhm=8, max=300, snthresh=5, mzdiff=0.5), group(bw=8, max=300), rector (method=loess, family=symmetric). A second grouping was done with the same parameter values. Due to in- source fragmentation, multiple m/z peaks for each compound were often observed. The number of compounds was estimated with “peak groups” consisting of m/z peaks with the same retention time (window, x s) that were correlated (Pearson; threshold, 0.85) across all control samples.

4.6 Phylogenetic analysis

Protein sequences were aligned with ClustalW and the phylogenetic tree was generated with the MEGA 5.0.1 software (Tamura et al., 2011), by the Neighbor-Joining method. Bootstrapping was performed with 10,000 replicates and the evolutionary distances were computed using the Poisson correction method. All positions containing gaps and missing data were eliminated from the data set (complete deletion option).

4.7 RNA-Seq

Total RNA from three independent MtTAX1^{OE}, MtTAXQ^{KD} and control lines was extracted with the RNeasy mini kit (Qiagen) and cDNA prepared with SuperScriptTM II Reverse Transcriptase (Invitrogen). RNA samples were sent to "Genomics core" (KU, Leuven, Belgium) for mRNA purification, cDNA library construction and Illumina Hiseq 2000-based RNA sequencing with Solexa technology.

Read mapping on the *Medicago* genome (MT3.5; (Young et al., 2011)) with TopHat version 2.0.3 (Trapnell et al., 2009) and counting of the uniquely mapped reads and calculation of the FPKM values with Cufflinks version 1.3.0 (Trapnell et al., 2010) were performed using default parameters as described (Pollier et al., 2013).

References

Achnine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., and Dixon, R.A. (2005). Genomics-based selection and functional characterization of triterpene glycosyltransferases from the model legume *Medicago truncatula*. Plant Journal **41,** 875-887.

Baluška, F., Šamaj, J., Wojtaszek, P., Volkmann, D., and Menzel, D. (2003). Cytoskeleton-plasma membrane-cell wall continuum in plants. Plant Physiology **133,** 482-491.

Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C., and Barker, D.G. (2001). *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Molecular Plant-Microbe Interactions **14,** 695-700.

Butenko, M.A., Vie, A.K., Brembu, T., Aalen, R.B., and Bones, A.M. (2009). Plant peptides in signalling: looking for new partners. Trends in Plant Science **14,** 255-263.

Cassab, G.I. (1998). Plant cell wall proteins. Annual Review of Plant Physiology and Plant Molecular Biology **49,** 281-309.

Cosgrove, D.J. (2001). Wall structure and wall loosening. a look backwards and forwards. Plant Physiology **125,** 131-134.

Darley, C., Forrester, A., and McQueen-Mason, S. (2001). The molecular basis of plant cell wall extension. Plant Molecular Biolology **47,** 179-195.

Fukushima, E.O., Seki, H., Ohyama, K., Ono, E., Umemoto, N., Mizutani, M., Saito, K., and Muranaka, T. (2011). CYP716A Subfamily Members are Multifunctional Oxidases in Triterpenoid Biosynthesis. Plant and Cell Physiology **52,** 2050-2061.

Hanks, J.N., Snyder, A.K., Graham, M.A., Shah, R.K., Blaylock, L.A., Harrison, M.J., and Shah, D.M. (2005). Defensin gene family in *Medicago truncatula:* structure, expression and induction by signal molecules. Plant Molecular Biology **58,** 385-399.

Jamet, E., Canut, H., Boudart, G., and Pont-Lezica, R.F. (2006). Cell wall proteins: a new insight through proteomics. Trends in Plant Science **11,** 33-39.

Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY$^{((TM))}$ vectors for *Agrobacterium*-mediated plant transformation. Trends in Plant Science **7,** 193-195.

Kim, B.J., Gibson, D.M., and Shuler, M.L. (2006). Effect of the plant peptide regulator, phytosulfokine-alpha on the growth and taxol production from *Taxus sp.* suspension cultures. Biotechnology and Bioengineering **95,** 8-14.

Li, L.N., Modolo, L.V., Escamilia-Trevino, L.L., Achnine, L., Dixon, R.A., and Wang, X.Q. (2007). Crystal structure of *Medicago truncatula* UGT85H2 - Insights into the structural basis of a multifunctional (Iso)flavonoid glycosyltransferase. Journal of Molecular Biology **370,** 951-963.

Marshall, E., Costa, L.M., and Gutierrez-Marcos, J. (2011). Cysteine-Rich Peptides (CRPs) mediate diverse aspects of cell–cell communication in plant reproduction and development. Journal of Experimental Botany **62,** 1677-1686.

Matsubayashi, Y., and Sakagami, Y. (2006). Peptide hormones in plants. Annual Review of Plant Biology **57,** 649-674.

Modolo, L.V., Li, L.N., Pan, H.Y., Blount, J.W., Dixon, R.A., and Wang, X.Q. (2009). Crystal structures of glycosyltransferase *UGT78G1* reveal the molecular basis for glycosylation and deglycosylation of (iso)flavonoids. Journal of Molecular Biology **392,** 1292-1302.

Morreel, K., Goeminne, G., Storme, V., Sterck, L., Ralph, J., Coppieters, W., Breyne, P., Steenackers, M., Georges, M., Messens, E., and Boerjan, W. (2006). Genetical metabolomics of flavonoid biosynthesis in Populus: a case study. Plant Journal **47,** 224-237.

Morreel, K., Kim, H., Lu, F.C., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inze, D., Messens, E., Ralph, J., and Boerjan, W. (2010). Mass spectrometry-based fragmentation as an identification tool in lignomics. Analytical Chemistry **82,** 8095-8105.

Mylona, P., Owatworakit, A., Papadopoulou, K., Jenner, H., Qin, B., Findlay, K., Hill, L., Qi, X., Bakht, S., Melton, R., and Osbourn, A. (2008). *Sad3* and *Sad4* are required for saponin biosynthesis and root development in oat. Plant Cell **20,** 201-212.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y.H., Sumner, L.W., and Dixon, R.A. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell **22,** 850-866.

Odjakova, M., and Hadjiivanova, C. (2001). The complexity of pathogen defense in plants. Bulgarian Journal of Plant Physiology **27,** 101-109.

Onrubia Ibáñez, M. (2012). A molecular approach to taxol biosynthesis. PhD thesis In experimental and health sciences (Universitat PUMPEU FABRA).

Osbourn, A., Goss, R.J.M., and Field, R.A. (2011). The saponins - polar isoprenoids with important and diverse biological activities. Natural Product Reports **28,** 1261-1268.

Pang, Y.Z., Peel, G.J., Sharma, S.B., Tang, Y.H., and Dixon, R.A. (2008). A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of *Medicago truncatula*. Proceedings of the National Academy of Sciences of the United States of America **105,** 14210-14215.

Pollier, J., Rombauts, S., and Goossens, A. (2013). Analysis of RNA-Seq data with TopHat and Cufflinks for genome-wide expression analysis of jasmonate-treated plants and plant

cultures. In In Jasmonate Signaling: Methods and Protocols. Methods in Molecular Biology (Humana Press), pp. in press.

Pollier, J., Morreel, K., Geelen, D., and Goossens, A. (2011). Metabolite profiling of triterpene saponins in *Medicago truncatula* hairy roots by liquid chromatography fourier transform ion cyclotron resonance mass spectrometry. Journal of Natural Products **74,** 1462-1476.

Qin, Q., Bergmann, C.W., Rose, J.K.C., Saladie, M., Kolli, V.S.K., Albersheim, P., Darvill, A.G., and York, W.S. (2003). Characterization of a tomato protein that inhibits a xyloglucan-specific endoglucanase. The Plant Journal **34,** 327-338.

Sasaki, K., and Takahashi, T. (2002). A flavonoid from *Brassica rapa* flower as the UV-absorbing nectar guide. Phytochemistry **61,** 339-343.

Silverstein, K.A.T., Moskal, W.A., Wu, H.C., Underwood, B.A., Graham, M.A., Town, C.D., and VandenBosch, K.A. (2007). Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants. The Plant Journal **51,** 262-280.

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

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution **10,** 2731–2739.

Tava, A., and Avato, P. (2006). Chemical and biological activity of triterpene saponins from *Medicago* species. Natural Product Communications **1,** 1159-1180.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics **25,** 1105-1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology **28,** 511-515.

Underwood, B.A., Vanderhaeghen, R., Whitford, R., Town, C.D., and Hilson, P. (2006). Simultaneous high-throughput recombinational cloning of open reading frames in closed and open configurations. Plant Biotechnology Journal **4,** 317-324.

Vorwerk, S., Somerville, S., and Somerville, C. (2004). The role of plant cell wall polysaccharide composition in disease resistance. Trends in Plant Science **9,** 203-209.

Young, N.D., Debelle, F., Oldroyd, G.E.D., Geurts, R., and al., e. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature **480,** 520-524.

Zhao, J., Davis, L.C., and Verpoorte, R. (2005). Elicitor signal transduction leading to production of plant secondary metabolites. Biotechnology Advances **23,** 283-333.

7

Conclusion and perspectives

Plant metabolites are produced through complex processes including multiple enzymatic steps, branched pathways and regulation by a number of functionally redundant transcription factors (TFs). Engineering secondary metabolite pathways depends upon exhaustive knowledge of the whole biosynthetic pathway and a detailed insight into the regulatory networks controlling the biosynthesis process.

Generally, two types of regulatory networks are involved in differential gene expression: transcription regulatory networks and post-transcription regulatory networks. Transcription regulatory networks contain two types of nodes: regulatory TFs and their target DNA sequences. Such networks are composed of protein–DNA interactions between TFs and their target genes. Protein-DNA interactions can be identified via two complementary strategies: TF-centered and gene-centered approaches. In the former approach the DNA sequences that interact with a TF or set of TFs of interest are identified, whereas the later comprises the methods that lead to the identification of the TFs that interact with a regulatory DNA sequence or set of DNA sequences of interest (e.g. gene promoters) (Walhout, 2006).

Today, TFs are well-known as regulators of various plant functions and many of them are characterized as key regulators of metabolic pathways (De Geyter et al., 2012) and therefore their manipulation would be an effective approach for controlling of plant metabolite biosynthesis, both quantitatively and qualitatively.

The major goal of this study was to identify regulators of the biosynthesis of saponins and isoflavonoids, the two major secondary metabolite classes in *Medicago truncatula*. Plant saponins are subject of many metabolic engineering efforts to enhance their production in cultured cells and corresponding plants. Currently our knowledge on saponin biogenetic pathways is still linear and insufficient. On the contrary, isoflavonoid biosynthesis is fairly well characterized. Nonetheless there is still an incomplete view on the regulation of this class of metabolites as well. Studies for identification of genes involved in secondary metabolism in *M. truncatula* using the techniques of genome-wide profiling have been initiated in the last few years (Suzuki et al., 2002; Broeckling et al., 2005; Naoumkina et al., 2008; Farag et al., 2009; Naoumkina et al., 2010).

Our research was launched based on comprehensive combined transcript and metabolite profiling studies of MeJA- elicited *M. truncatula* cell cultures that establish a substantial collection of *M. truncatula* genes potentially involved in secondary metabolism (Pollier, 2011). We used a reverse genetics approach and a combination of transcript analysis and metabolite profiling for further characterization of the selected genes (chapter 4). Out of 6

selected MeJA-responsive genes with the putative regulatory role, the only regulatory gene with some effects on secondary metabolite biosynthesis genes was *Mt148* that caused lower expression levels of some triterpene saponin and (iso) flavonoid genes in hairy roots overexpressing this regulatory gene. However, stable overexpression of *Mt148* did not lead to any changes in metabolite profiles of hairy roots.

It is possible that coordinated expression of a combination of other regulatory genes is necessary for manipulation of secondary metabolite biosynthesis, especially when taking into account that the Mt148 protein is a member of the large Ccr4-Not regulatory complex involved in several aspect of gene expression regulation (Collart and Panasenko, 2012). Hence, studies on protein-protein interactions could be a way to get more insight into the regulatory function of this gene. Indeed, tandem affinity purification platform is being established in *M. truncatula* hairy roots and used to map complex involved in JA-signaling, including *Mt148* (PhD project of Nathan De Geyter).

In chapter 5, another early MeJA upregulated *M. truncatula* gene, a MYB TF *Mt061*, was chosen for further characterization. Overexpression of this TF in hairy roots could be associated with the production and emission of green leaf volatile (GLV) compounds and higher accumulation of isoflavonoids and pterocarpans. To identify differential gene expression of $Mt061^{OE}$ lines, and to get an insight into the regulatory mechanism of *Mt061*, we performed RNA-Seq transcript profiling on the transgenic lines.

RNA-Seq is the first sequencing-based method that allows the entire transcriptome to be surveyed in a very high-throughput and quantitative manner (Wang et al., 2009). With RNA- Seq technique it is possible to quantify differentially expressed genes and identify novel transcribed regions and alternative splice events with high levels of accuracy and specificity (Wang et al., 2009). The RNA- Seq approach was shown to have relatively little variation between technical replicates (Marioni et al., 2008). During the data analysis, the transcriptome sequencing reads are usually mapped to the reference genome sequences or transcriptome databases if they are available. However, during our analysis it appeared that the *M. truncatula* genome is not still fully sequenced, though has been recently claimed that it is sequenced up to 94 % (Young et al., 2011). Hence, we mapped the RNA-Seq data from short reads onto the *M. truncatula* reference genome (chapters 5 and 6) (Young et al., 2011). Unexpectedly, the results obtained from the RNA-Seq revealed that several genes are still missing in reference genome and that annotation needs to be improved. For example, HPL2, a key enzyme in production of GLV

compounds, many known saponin biosynthesis genes or two *MtTAX* genes (*MtTAX2* and *MtTAX5*) are not present on the reference genome yet.

De novo transcriptome assembly is an alternative method that has been developed to produce a genome-scale transcription map for those species without genomic reference sequence (Chen et al., 2011). Therefore, to investigate the genes that are missing from the *M. truncatula* reference genomes due to the incompleteness of reference sequences, *de novo* assembly is a suitable approach.

In addition, in further attempts to map the regulatory network of *Mt061*, determination of the binding ligands via protein-protein and protein-DNA interactions would be the most logical next step. The steady state level of *Mt061* was not affected by MeJA elicitation of *M. truncatula* hairy roots and the overexpression of *Mt061* led to the upregulation of several GLV and MeJA biosynthesis genes. Therefore, exploring the transcriptional and metabolite changes upon MeJA elicitation of transgenic lines might be helpful to gain a better understanding of its incorporation in stress signaling network.

In chapter 6, we showed that heterologous expression of the *Taximin* gene from *Taxus baccata* in hairy root of *M. truncatula* led to enhanced accumulation of saponin glycosides. Taximin is a new Cys-rich peptide identified through cDNA-AFLP studies of MeJA elicited cell cultures of *T. baccata*. It was previously shown that this peptide is activated by MeJA and can modulate alkaloid and taxol biosynthesis in tobacco and *Taxus baccata,* respectively (Onrubia Ibáñez, 2012). However, qRT-PCR analysis did not show any changes in the expression levels of saponin biosynthesis genes in the *M. truncatula* hairy roots expressing this gene, hence more comprehensive genome-wide analysis and protein-protein interaction might be needed to unravel the mechanism of action of this peptide. Furthermore neither overexpression nor downregulation of *Taximin* homolog genes in *M. truncatula* (*MtTAX*) had any effect on the metabolite profiles of *M. truncatula* hairy roots, whereas downregulation of some *MtTAX* genes led to decreased expression of some saponin biosynthesis genes. Besides, RNA-Seq analysis showed that several biosynthesis genes involved in other secondary metabolite biosynthesis pathways were downregulated in *MtTAX* knockdown lines. Therefore, it is proposed that the peptides from this class might have a role in signaling pathways leading to the accumulation of secondary metabolism.

Although further studies about the activity of the *M. truncatula* taximin homologs (the *MtTAX*) on the enzymes involved in the secondary metabolism are needed, the results obtained in this study suggested that *MtTAX* genes are likely involved in the regulation of

the production of triterpene saponin biosynthesis pathway. It could thus be interesting to check the interaction of the *MtTAX* peptides with the genes involved in downstream steps of the triterpene saponin biosynthesis pathway. In addition, preliminary RNA-Seq data indicate that investigation of their role in other metabolic pathways such as the carotenoid and phenylpropanoid biosynthesis pathways is also worthwhile.

Looking back to the initial goal of this project, identifying regulators of secondary metabolite biosynthesis in the model legume *Medicago truncatula*, no regulatory elements controlling saponin biosynthesis have been identified through this study.

The gene-centered protein–DNA interaction mapping methods, such as high-throughput yeast one-hybrid (Y1H) assays may facilitate the identification of regulatory elements that can bind to *cis*-regulatory sequence and promoters of saponin biosynthesis genes.

In conclusion, in this study we have showed that the integration of metabolomics and transcriptomics can provide a powerful approach to better understand gene-to-metabolite networks for identifying the function of unknown genes and for the further characterization of key regulatory elements and metabolites involved in plant adaptation to environmental signals. However, the results of this study have revealed that transcript and metabolite interaction is often quite complex. Intermediate steps between transcription and metabolite production such as post-translational modification, regulation or modulation of the expression of the biosynthetic and regulatory genes by metabolite levels may be responsible for the non-linear and complex transcriptome-metabolome relationship. In light of the modern functional genomics technologies which are becoming more and more available, more progress on dissecting of a signal transduction pathways, transcriptional regulatory networks, and transcript to metabolite events can be envisaged.

References

Broeckling, C.D., Huhman, D.V., Farag, M.A., Smith, J.T., May, G.D., Mendes, P., Dixon, R.A., and Sumner, L.W. (2005). Metabolic profiling of *Medicago truncatula* cell cultures reveals the effects of biotic and abiotic elicitors on metabolism. Journal of Experimental Botany **56,** 323-336.

Chen, G., Yin, K., Wang, C., and Shi, T. (2011). *De novo* transcriptome assembly of RNA-Seq reads with different strategies. Science China Life Sciences **54,** 1129-1133.

Collart, M.A., and Panasenko, O.O. (2012). The Ccr4-Not complex. Gene **492,** 42-53.

De Geyter, N., Gholami, A., Goormachtig, S., and Goossens, A. (2012). Transcriptional machineries in jasmonate-elicited plant secondary metabolism. Trends in Plant Science **17,** 349-359.

Farag, M.A., Deavours, B.E., de Fatima, A., Naoumkina, M., Dixon, R.A., and Sumner, L.W. (2009). Integrated metabolite and transcript profiling identify a biosynthetic mechanism for hispidol in *Medicago truncatula* cell cultures. Plant Physiology **151,** 1096-1113.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008). RNA-Seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Research **18,** 1509-1517.

Naoumkina, M.A., He, X.Z., and Dixon, R.A. (2008). Elicitor-induced transcription factors for metabolic reprogramming of secondary metabolism in *Medicago truncatula*. BMC Plant Biology **8**.

Naoumkina, M.A., Modolo, L.V., Huhman, D.V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L.W., and Dixon, R.A. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. Plant Cell **22,** 850-866.

Onrubia Ibáñez, M. (2012). A molecular approach to taxol biosynthesis. PhD thesis in experimental and health sciences (Universitat PUMPEU FABRA).

Pollier, J. (2011). Combinatorial Biochemistry of Triterpene Saponins in Plants (Faculty of Sciences VIB Department of Plant Systems Biology Ghent University).

Suzuki, H., Achnine, L., Xu, R., Matsuda, S.P.T., and Dixon, R.A. (2002). A genomics approach to the early stages of triterpene saponin biosynthesis in *Medicago truncatula*. The Plant Journal **32,** 1033-1048.

Walhout, A.J.M. (2006). Unraveling transcription regulatory networks by protein–DNA and protein–protein interaction mapping. Genome Research **16,** 1445-1454.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics **10,** 57-63.

Young, N.D., Debelle, F., Oldroyd, G.E.D., and et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature **480,** 520-524.

Summary

Plant secondary metabolites are an extremely huge group of natural compounds involved in plant responses to biotic and abiotic stress signals. From a human point of view, many of these bioactive compounds have wonderful pharmaceutical, industrial and agricultural applications. The induction of secondary metabolism by several biotic and abiotic stresses is often mediated by regulatory signaling molecules, such as jasmonate (JA), ethylene, salicylic acid, and their derivatives. JAs are the most important molecules for induction of secondary metabolism and have been found to have a conserved role in induction of the biosynthesis of a variety of secondary metabolites. Several studies revealed that JA orchestrates comprehensive transcriptional reprogramming leading to the shifts in metabolic fluxes. Transcription factors (TFs) are characterized as the key components of signaling pathways toward the onset of secondary metabolites biosynthesis. Many TFs were characterized as jasmonate-modulated regulators involved in the regulation of the biosynthesis of secondary metabolites (chapter 2).

In this thesis we explored putative regulatory genes potentially involved in the biosynthesis of secondary metabolites in the model legume, barrel medic (*Medicago truncatula*). *Medicago* species produce a variety of bioactive natural products that possess promising pharmaceutical and agricultural benefits (chapter 3).

To identify putative regulators of secondary metabolism in *M. truncatula* a reverse-genetics screen was performed on genes identified from a genome-wide cDNA-AFLP transcript profiling on MeJA elicited cell cultures of *M. truncatula* previously. Two regulatory proteins were identified by assessing the effect of gain-of- function on secondary metabolite biosynthesis in *M. truncatula* hairy roots (chapters 4 and 5). *Mt148* encodes a CCR4-associate factor1 (Caf1) protein presumably involved in mRNA metabolism and post- transcriptional gene regulation. Overexpression of *Mt148* led to the downregulation of some secondary metabolite genes in *M. truncatula* hairy roots via a mechanism that is still unknown (chapter 4). We also identified *Mt061*, which encodes an R2R3-type MYB family transcription factor, as a candidate regulator of green leaf volatile (GLV) biosynthesis in *Medicago truncatula*. Overexpression of *Mt061* in *M. truncatula* hairy roots caused the production and

emission of GLV compounds. In addition, the enhanced transcript levels of several defense genes and phenolic biosynthesis genes and higher levels of phytoalexin compounds detected in *Mt061*-overexpressing lines suggest that *Mt061* might be involved in the overall regulation of defense response processes in the plant (chapter 5).

Among other components involved the signaling pathways, several signaling peptides have been identified to control secondary metabolism in plants. During this thesis we also explored the effects of ectopic expression of a small signaling peptide, *Taximin*, from *T. baccata* on secondary metabolism in *M. truncatula* hairy roots (chapter 6). The results showed that *Taximin* can modulate saponin production in *M. truncatula* hairy roots. Furthermore cloning, characterization and functional analysis of the homologs of the *Taximin* gene from *M. truncatula*, the *MtTAX* genes of which we identified, suggested that they may have a general role in the regulation of secondary metabolite biosynthesis in *M. truncatula*.

ACKNOWLEDGEMENTS

Looking back, I am very glad for everything I have received over the last four years of my life. Since I have moved to the beautiful city of Gent in Belgium for higher education, without help, support and encouragement from several people it was hardly possible for me to succeed in my doctoral work. Here is a small tribute to all who have contributed to my experience.

First of all, I would like to express my greatest gratitude to my supervisor Prof. Alain Goossens for providing me with the opportunity to complete my PhD thesis. Thanks for all of his guidance, support and encouragement that he has given me throughout my PhD research work.

I am deeply indebted to my committee members Prof. Wout Boerjan, Prof. Sofie Goormachtig, Prof. Bartel Vanholme, Prof. Jan Van Bocxlaer, Prof. Rosa Cusido, Prof. Tom Beeckman and Dr. Jacob Pollier for their time and effort in reviewing this work. Their critical suggestions were valuable help in the improvement of this thesis.

I would also like to extend my gratitude towards my lab mates, Amparo, Laurens, Andres, Janine, Jacob, Tessa, Alex, Rebecca, Robin, Sofie, Nathan, Astrid, Jonas, Gino, Michele, Michiel, Sandra, Jan and Philipp for a friendly and cooperative atmosphere that they provided for me.

I have made many friends along the way. They have helped me directly or indirectly to complete my Ph.D. Special thanks to Mansour, Wilson, Nino, Delphine, Kristof, Jackie, Hendrik, Annick Bleys and Karel for their valuable help that I received in the lab.

Finally, I must acknowledge my husband and best friend, Ali, without whose love, encouragement and editing assistance, I would not have finished this thesis. I dedicated my thesis to my dearest Ali.