Biochemical and Molecular Investigations of *Arabidopsis thaliana* Transformed with Genes of Rosmarinic Acid Biosynthesis

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Pharmazie der

Philipps-Universität Marburg

vorgelegt von

Agus Chahyadi

aus Lambuya/Indonesian

Marburg/Lahn 2017

Erstgutachter: Prof. Dr. Maike Petersen

Zweitgutachter: Prof. Dr. Alfred Batschauer

Eingereicht am 12.05.2017

Tag der mündlichen Prüfung am 28.06.2017

Hochschulkennziffer: 1180

List of abbreviations

2,4-D	2,4-dichlorophenoxyacetic acid
APS	Ammoniumpersulphate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
Caf	Caffeoyl
CaMV 35S	Cauliflower mosaic virus 35S promoter
CbRAS	*Coleus blumei*-RAS
CB2	*Coleus blumei*-B5 medium
CB(2)OH	CB2 medium without hormones
cDNA	complementary DNA
CoA	Coenzyme A
Cq	Quantification cycle
CTAB	Cetyltrimethylammonium bromide
CYP	Cytochrome P450
C4H	Cinnamate 4-hydroxylase
4CL	4-Coumarate coenzyme A ligase
DHPL	3,4-Hydroxyphenyllactic acid
DIECA	Diethyldithiocarbaminate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ε	Extinction coefficient
EIC	Extracted ion chromatography
EtOH	Ethanol
gDNA	Genomic deoxyribonucleic acid
HCT	Hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyltransferase
HPLC	High performance liquid chromatography
HPPD	Hydroxyphenylpyruvate dioxygenase
HPPR	Hydroxyphenylpyruvate reductase
HPR	Hydroxypyruvate reductase
Ic	Inhibitor cocktail
IES	Indole-3-acetic acid
KPi	Potassium phosphate buffer
LC-MS	Liquid chromatography – mass spectrometry
MeOH	Methanol
mRNA	Messenger ribonucleic acid
NAA	Naphthalene acetic acid
OD_{600}	Optical density at 600 nm

PAL	Phenylalanine ammonia-lyase
pC	4-coumaroyl
PCR	Polymerisation chain reaction
pHPP	4-Hydroxyphenylpyruvic acid
pHPL	4-Hydroxyphenyllactic acid
PLP	Pyridoxalphosphate
PMSF	Phenylmethylsulfonyl fluoride
PTGS	Post-transcriptional gene silencing
PTM	Post-translational modification
p.a.	Pro analysi
qPCR	Quantitative polymerase chain reaction
RA	Rosmarinic acid
RAS	Rosmarinic acid synthase
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TAIR	The *Arabidopsis* information source
TAT	Tyrosine aminotransferase
T-DNA	Transfer-DNA
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TEMED	N,N,N',N'-Tetramethylenediamine

Contents

1. Introduction

1.1 Rosmarinic acid

1.1.1 Occurrence

Rosmarinic acid (RA) is an ester of caffeic acid and 3,4-dihydroxyphenyllactic acid (DHPL), which was first isolated from rosemary (*Rosmarinus officinalis*, Lamiaceae) and then named accordingly (Scarpati and Oriente 1958). The occurrence of RA was found to be widely distributed not only in eudicotyledonous plants such as Boraginaceae and Lamiaceae, but also in certain families of monocotyledonous plants and even in lower plants, e.g. Blechnaceae (ferns) and Anthocerotaceae (hornworts), although only in a few species (Tab. 1; Petersen (2013)).

Table 1. Reported occurrence of RA in the plant kingdom (Petersen 2013)

Order	**Family**
Anthocerophyta (Hornworts)	
Anthocerotales	Anthocerotaceae
Ferns	
Polypodiales	Blechnaceae
Dennstaedtiales	Dennstaedtiaceae
Spermatophyta – Magnoliopsida	
Basal orders	
Chloranthales	Chloranthaceae
Monocotyledonous plants	
Alismatales	Araceae; Potamogetonaceae; Zosteraceae
Liliales	Melianthaceae
Asparagales	Iridaceae
Poales	Poaceae; Cyperaceae
Zingiberales	Cannaceae; Marantaceae
Eudicotyledonous plants	
Saxifragales	Crassulaceae
Celastrales	Celastraceae
Fabales	Fabaceae
Rosales	Rosaceae; Moraceae
Cucurbitales	Cucurbitaceae
Myrtales	Onagraceae; Myrtaceae
Sapindales	Sapindaceae

Malvales	Malvaceae (incl. former Sterculiaceae and Tiliaceae)
Brassicales	Brassicaceae
Caryophyllales	Portulacaceae; Amaranthaceae; Polygonaceae
Gentianales	Rubiaceae
Lamiales	Lamiaceae; Plantaginaceae; Acanthaceae; Scrophulariaceae; Linderniaceae
Solanales	Solanaceae
Without assigned order	Boraginaceae (incl. former Hydrophyllaceae)
Asterales	Asteraceae
Apiales	Apiaceae; Araliaceae
Dipsacales	Dipsacaceae

The presence of RA in plants is not consistent. Boraginaceae is the only family in which RA occurs throughout in all species, whereas in Lamiaceae, the occurrence is mainly restricted to the subfamily Nepetoideae (Petersen et al. 2009). In other plant taxa, the inconsistency is also reported to occur within an order, within a family and even within a genus, in which not all members contain RA in each level of taxa (Abdullah et al. 2008; Petersen et al. 2009; Petersen 2013).

1.1.2 Bioactivity

As a phenolic compound, RA is considered to be a plant defence compound against herbivores and pathogens (Petersen and Simmonds 2003) and a plant UV-protectant (Luis et al. 2007). It also possesses a number of interesting pharmacological activities such as antibacterial, antiviral, antioxidant and anti-inflammatory (Petersen and Simmonds 2003), and therefore contributes to the efficacy of several medicinal plants (e.g. *Salvia officinalis, Mentha x piperita, Thymus vulgaris, Melissa officinalis, Symphytum officinale*) within the families Lamiaceae and Boraginaceae (Petersen et al. 2009). In recent years, RA's pharmacological activities have been found to be extremely diverse. RA and its derivatives display potential usages as antiaging, antiallergic, anticancer, neuroprotective agents, and many more that have been reviewed in more detail by Bulgakov et al. (2012) and recently by Kim et al. (2015) and Amoah et al. (2016).

1.1.3 Biosynthesis

The biosynthetic pathway of RA in Lamiaceae species has been fully elucidated (Fig. 1; Weitzel and Petersen (2010)). The formation of RA utilises L-phenylalanine as a source of caffeic acid and L-tyrosine as a source of 3,4-dihydroxyphenyllactic acid (Ellis and Towers 1970). For the

formation of the caffeic acid precursor, the first step is generating *t*-cinnamic acid via enzymatic deamination of phenylalanine by phenylalanine ammonia-lyase (PAL), followed by hydroxylation at *para* position of the aromatic ring by a cytochrome P450 monooxygenase, cinnamate 4-hydroxylase (C4H), forming 4-coumaric acid and finally, activation of 4-coumaric acid to 4-coumaroyl-CoA by 4-coumarate coenzyme A ligase (4CL). The 3,4-dihydroxyphenyllactic acid part is generated from tyrosine which is transaminated by tyrosine aminotransferase (TAT) yielding 4-hydroxyphenylpyruvic acid (pHPP) which is then converted to either homogentisic acid, a precursor for tocopherols and plastoquinones needed in photosynthesis, or 4-hydroxyphenyllactic acid (pHPL) by hydroxyphenylpyruvate dioxygenase (HPPD) or hydroxyphenylpyruvate reductase (HPPR), respectively. 4-Coumaroyl-CoA and pHPL are condensed by rosmarinic acid synthase (RAS), a key enzyme in RA biosynthesis, to yield an ester 4-coumaroyl-4'-hydroxyphenylpyruvic acid. This intermediate is further hydroxylated consecutively in positions 3 and 3' of the aromatic rings by two cytochrome P450 monooxygenase activities, which finally generates caffeoyl-3',4'-dihydroxyphenyllactic acid (rosmarinic acid, RA). These seven or eight enzymes involved in RA biosynthesis have been well characterised and the respective genes isolated and analysed from various plants, e.g. *Coleus blumei* (Lamiaceae) (Petersen 1991, 1993; Kim et al. 2004; Berger et al. 2006; Eberle et al. 2009; Sander and Petersen 2011), *Melissa officinalis* (Lamiaceae) (Weitzel and Petersen 2010, 2011), *Salvia miltiorrhiza* (Lamiaceae) (Huang et al. 2008; Hu et al. 2009; Song and Wang 2009), *Anthoceros agrestis* (Anthocerotaceae) (Petersen 2003), and *Lithospermum erythrorhizon* (Boraginaceae) (Yamamura et al. 2001; Matsuno et al. 2002; Tsuruga et al. 2006).

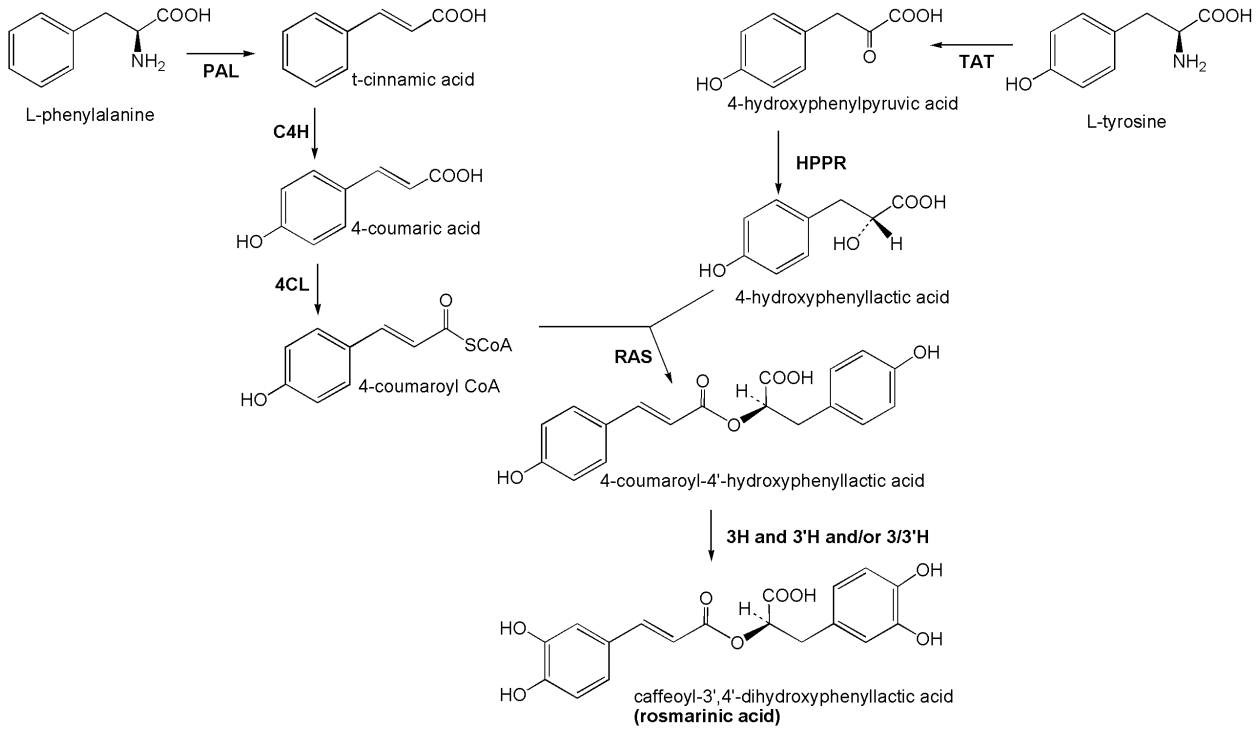


Figure 1. RA biosynthetic pathway (Weitzel and Petersen 2010)

Among these enzymes mentioned above, only two or three enzymes have been suspected to be specific enzymes for RA formation, namely RAS and the cytochrome P450s catalysing the 3- and 3'-hydroxylation (e.g. CYP98A14). The others, PAL, C4H, 4CL, TAT as well as HP(P)R, are well known as ubiquitous genes/enzymes that exist in higher plants and are involved in the phenylpropanoid pathway leading to the precursors for phenolic compounds, i.e. flavonoids, anthocyanins and lignin, and for the biosynthesis of tocopherols and plastoquinones (Petersen et al. 2009). In addition, HP(P)R might be involved in plant photorespiration (Peterhansel et al. 2010).

1.2 *Arabidopsis thaliana*

1.2.1 A model plant

For more than sixty years, *Arabidopsis thaliana* has been the subject of research and has gained increasing popularity as a model organism for the study of plant biology. *Arabidopsis* belongs to the mustard (Brassicaceae) family, which includes economically important crops such as cabbage, broccoli, and horseradish. Although closely related to these crops, *Arabidopsis* has no major agronomic significance. Despite this, its astonishing features such as short life cycle, small

habitus which limits the requirement for growth facilities, efficient reproduction through self-pollination, and relatively small and tractable genome that can be manipulated through genetic engineering more easily and rapidly than other plant genomes made *Arabidopsis* a prominent model organism in plant molecular genetics, development, physiology, and biochemistry (Meyerowitz 1987; Goodman et al. 1995; Koornneef and Meinke 2010)

In 1577, when the plant was first discovered in the Harz Mountains by Johannes Thal (1542-1583), a physician from Nordhausen, Thüringen, Germany, it was called *Pilosella siliquosa*. A number of name changes have been done since then. In 1753, Carl Linnaeus assigned names of *Pilosella siliquosa* minor as well as *Arabis thaliana* in honor of Thal. In 1841, the taxonomist Gustav Heynhold renamed *Arabis thaliana* as *Arabidopsis thaliana* (L.) Heynh (http://www.arabidopsis.com/main/tbl/chronology.htm). In this work, the use of the word "*Arabidopsis*" refers to *A. thaliana*.

Arabidopsis grows throughout in many regions and climates, ranging from high elevations in the tropics to the cold climate of northern Scandinavia, and including locations in Europe, Asia, Africa, Australia, and North America (Mitchell-Olds 2001). Its natural habitat is open free draining ground, such as sandy or gravelly soils (Anderson and Wilson 2002). *Arabidopsis* (Fig. 2) grows as a small ground-hugging rosette of about 5-10 cm diameter, from which a flowering stem is produced which can extend to a height of 20-70 cm. Flowers are 2-3 mm long, self-pollinate as the bud opens, but can also be crossed by applying pollen to the stigma surface, and produce several hundred seeds (0.5 mm in length) in slender fruits known as siliques. The siliques contain about 30-60 seeds each and shatter on ripening to allow the distribution of seed. Seedlings develop into the above-mentioned rosette plants. The entire life cycle, including seed germination, formation of a rosette plant, bolting of the main stem, flowering, and maturation of the first seeds, is completed in 6-8 weeks. Many different ecotypes have been collected from natural populations and are available for experimental analysis. The Columbia and Landsberg *erecta* ecotypes are commonly used as standards for genetic and molecular studies (Meinke et al. 1998; Anderson and Wilson 2002).

Figure 2. *Arabidopsis thaliana* at an early stage of flowering (Meinke et al. 1998)

The *Arabidopsis* genome is the first-ever plant genome that was completely sequenced. The *Arabidopsis* Genome Initiative (2000) stated five chromosomes with a total size of around 125-megabasepairs (Mbps) of the *Arabidopsis* genome, although other methods show a size of up to 157 Mbp (Bennett et al. 2003). The most up-to-date version of the *Arabidopsis* genome is preserved by the *Arabidopsis* Information Resource (TAIR) (https://www.arabidopsis.org/portals/genAnnotation/gene_structural_annotation/annotation_data.jsp). The TAIR10 release comprises 27,416 protein-encoding genes, 4827 pseudogenes or transposable element genes and 1359 non-coding RNAs (33,602 genes in all, 41,671 gene models). However, not all proteins and their functions have been assigned today.

1.2.2 Specialised metabolism in *Arabidopsis*

As the model plant, the enormous impact of *Arabidopsis* has advanced the current scientific knowledge in all aspects of plant biology. So far, *Arabidopsis* has also become a useful plant for studying specialised metabolism. In 2005, D´Auria and Gershenzon (2005) reported that more than

170 secondary metabolites have been identified in this species and they predicted many more compounds would be discovered. Five years later, the number has been increased to over 270 secondary metabolites (Soledade et al. 2010). Apparently, it will be growing in the future, since 103 compounds were found in root exudates of *Arabidopsis* and many of these compounds have not been reported before (Strehmel et al. 2014). As the most typical class of secondary metabolites contained in Brassicaceae and grouped according to their biosynthetic building blocks, the secondary metabolites of *Arabidopsis* are fatty acids and their derivatives (31 compounds), shikimate-derived metabolites (48, i.e. phenylpropanoids, benzenoids, flavonoids and their glycosides), isoprenoids (99, i.e. mono-, sesqui-, and triterpenes, and steroids), and alkaloids and sulphur metabolites (95, i.e. indole alkaloids, glucosinolates and their derivatives) (Soledade et al. 2010). According to the genome sequence of this species, there are 12 gene families encoding many classes of enzymes participating in specialised metabolism in *Arabidopsis*. A large number of them are enzymes involved in prominent reactions, such as hydroxylation (cytochrome P450 monooxygenases), glycosylation (glycosyltransferases), and acylation (acyltransferases) (D´Auria and Gershenzon 2005). However, the vast majority of these genes are so far uncharacterised and therefore revealing that a huge number of secondary metabolites are still waiting to be identified too.

In plants, acylation particularly plays a significant role to gain versatile properties of secondary metabolites. Most of the secondary metabolites generated from this reaction have a multitude of functions in ecological interaction (e.g. plant defence) (Landry et al. 1995; Treutter 2006), and plant growth and development (e.g. flowering, lignification) (Besseau et al. 2007; Grienenberger et al. 2009). The acylation includes reactions with various acids (donor compounds, e.g. acetate, malonate, hydroxycinnamate or (hydroxyl)benzoate) with oxygen-containing (e.g. shikimate, quinate, hydroxyphenyllactate) and nitrogen-containing acceptor substrates (e.g. choline or spermidine) to produce esters and amides, respectively. In *Arabidopsis*, prominent hydroxycinnamic acid derivatives are sinapic acid esters, e.g. sinapoylmalate, sinapoylcholine (Fig. 3) which are synthesised by serine carboxypeptidase-like acyltransferases with hydroxycinnamoylglucose esters as activated acyl donors (Milkowski and Strack 2004, 2010; Stehle et al. 2008, 2009).

According to D´Auria (2006) and Yu et al. (2009) 61/64 acyltransferases of the BAHD-superfamily, acyl CoA-utilising enzymes, are found in the genome of *Arabidopsis*. Members with known activities are involved in the biosynthesis of anthocyanins (D´Auria et al. 2007; Luo et al.

2007), aromatic suberin components (Gou et al. 2009), spermidine derivatives (Grienenberger et al. 2009; Luo et al. 2009), volatile fragrance compounds (D`Auria et al. 2002), and hydroxycinnamoylshikimate and -quinate (Besseau et al. 2007). The last mentioned enzyme, also known as hydroxycinnamoyl-coenzyme A shikimate:quinate hydroxycinnamoyltransferase (HCT), mainly uses shikimate as hydroxycinnamoyl acceptor and is involved in the formation of lignin monomers. Although it is able to form hydroxycinnamoylquinate esters, there is no report with regard to occurrence of chlorogenic or rosmarinic acid in *Arabidopsis*.

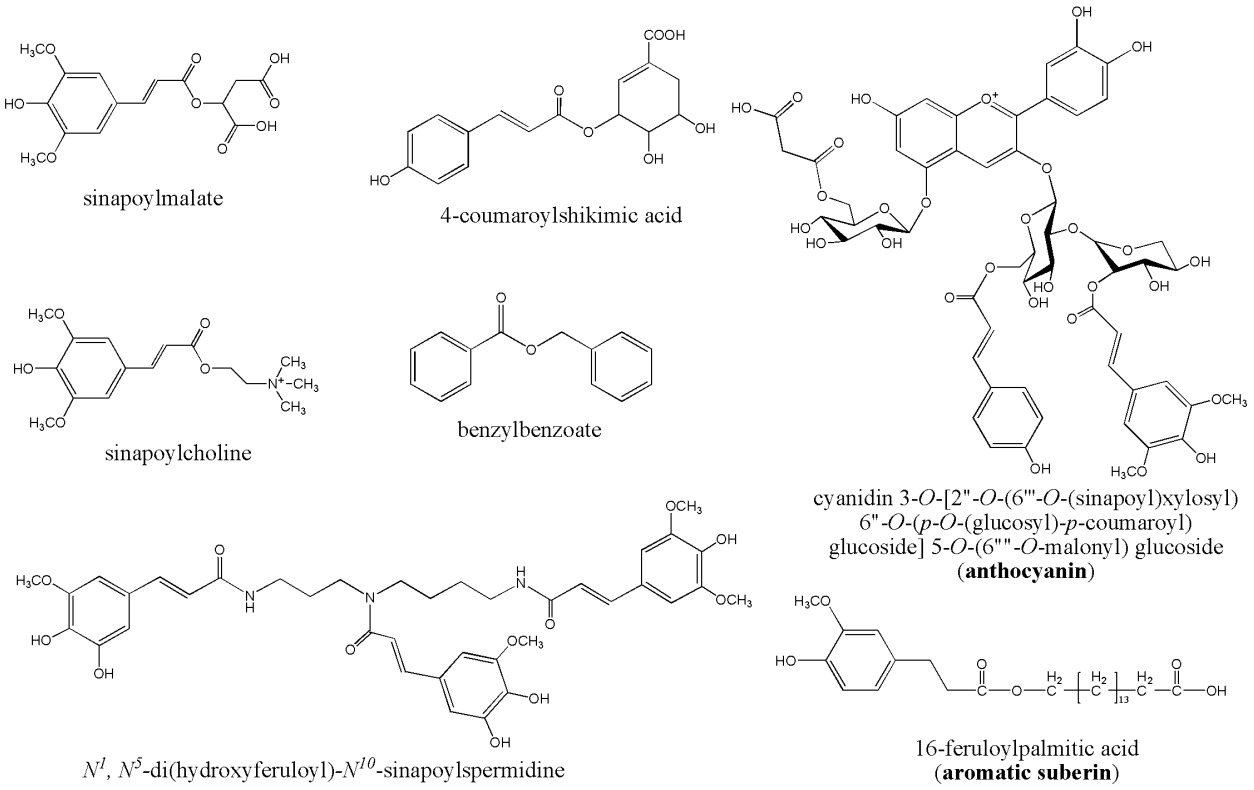


Figure 3. Hydroxycinnamic acid ester derivatives in *Arabidopsis* (Besseau et al. 2007; Luo et al. 2007; Gou et al. 2009; Matsuno et al. 2009; Milkowski and Strack 2010)

1.2.3 Presence of the rosmarinic acid biosynthetic pathway in *Arabidopsis*

Obviously, the biosynthetic pathway of RA (as shown in Fig. 1) is a mixed pathway of the two ubiquitous plant pathways, namely the phenylpropanoid and the tocopherols/plastoquinones pathway (Petersen et al. 2009). The phenylpropanoid pathway provides the important precursor 4-coumaroyl-CoA for synthesising phenolic compounds, e.g. flavonoids, anthocyanins and lignin. Therefore, this pathway should be present in all land plants. Furthermore, the pathway for the

biosynthesis of tocopherols and plastoquinones is essential for photosynthesis and thus should be present in all photosynthesising plants. Accordingly, both of pathways as well as their biosynthetic genes are present ubiquitously in higher plants and in *Arabidopsis* for sure, whereas the key steps of the RA pathway may be restricted to producing plants.

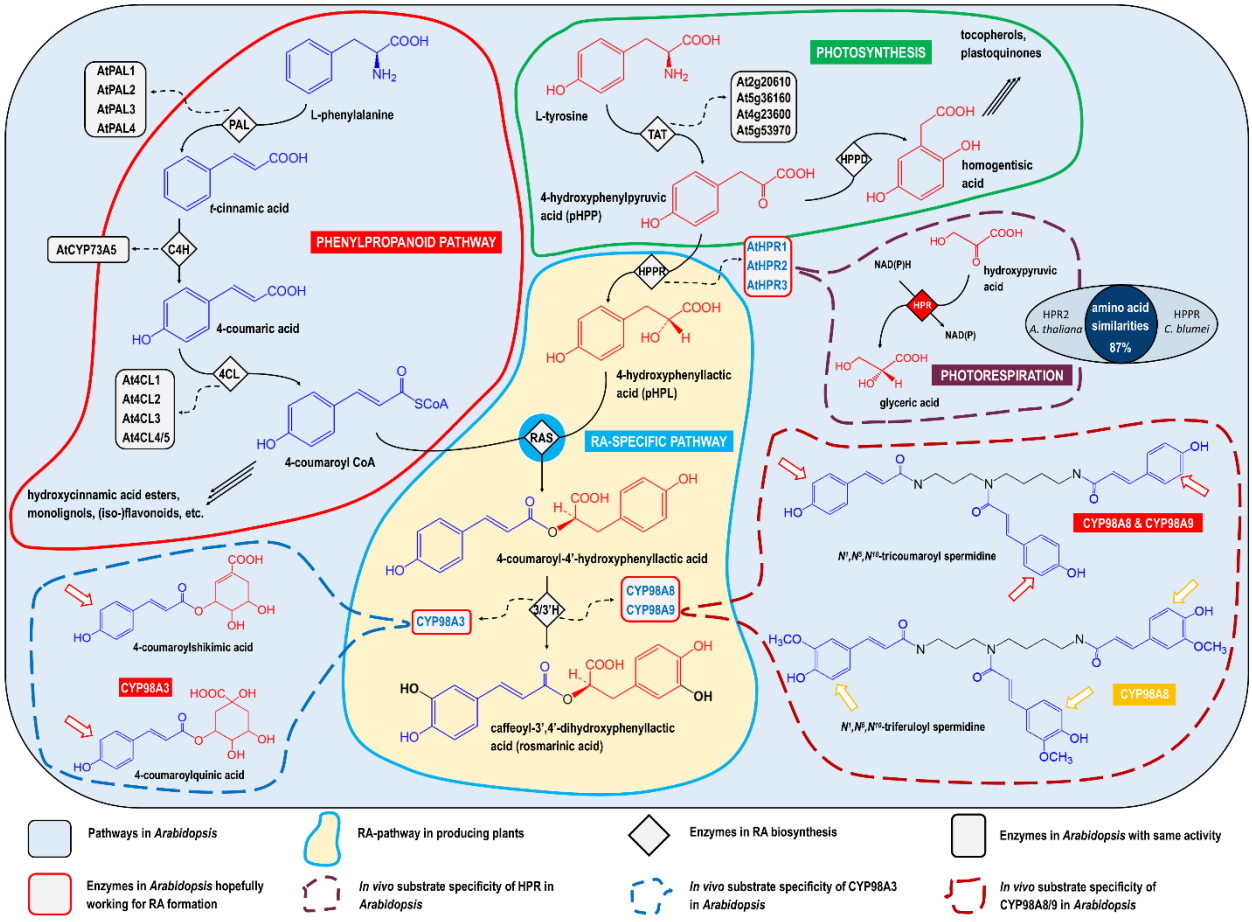


Figure 4. Presence of parts of the rosmarinic acid biosynthesis pathway in *Arabidopsis*

Regarding our investigation of genes and enzymes of the RA biosynthetic pathway in *Arabidopsis* (Fig. 4), at least four enzymes undoubtedly exist in this plant. They even have many isoforms, e.g. PAL, 4CL, and TAT, and were well characterised. The only missing enzymes are RAS, HPPR and cytochrome P450 monooxygenases for the *meta*-hydroxylation of both, the 4-coumaroyl and pHPL moieties. Thus, *Arabidopsis* - to our knowledge - cannot synthesise RA. However, recent investigations revealed that a few enzymes in *Arabidopsis* have high similarities with enzymes from RA-producing plants and so they might display similar activities (see Fig. 4). For instance, hydroxypyruvate reductases (HPRs), normally involved in photorespiration, can also

provide pHPL, although probably in low amounts. Other enzymes are cytochrome P450 monooxygenases being members of the CYP98A family that are also responsible for *meta*-hydroxylation of coumaric acid esters. These enzymes could probably accept ester products generated from RAS activity. The properties of these enzymes described above (PAL, C4H, 4CL, TAT, HPRs, CYP98A family) in *Arabidopsis* are described in detail below.

1.2.3.1 Phenylalanine ammonia-lyase (PAL)

PAL (EC 4.3.15) catalyses the non-oxidative deamination of phenylalanine to yield *trans*-cinnamic acid and ammonia which represents the first step of phenylpropanoid metabolism (Fig. 5; Cochrane et al. (2004)), a core pathway for the biosynthesis of a large class of aromatic plant natural products. PAL activity has been found in all higher plants analysed so far, and in some fungi and a few bacteria, but not in animals (Zhang and Liu 2015). As found in many plants, PAL of *Arabidopsis* is also encoded by a multi-gene family with presence of four PAL isoforms therein, AtPAL1 (At2g37040), AtPAL2 (At3g53260), AtPAL3 (At5g04230), and AtPAL4 (At3g10340). The four AtPAL isogenes are expressed ubiquitously in all organs of the plant, albeit they show different patterns on the level of expression. AtPAL1, 2, and 4 are expressed at high levels in the inflorescence stem, particularly during the plant developmental phase of active lignification (Raes et al. 2003). Two *Arabidopsis* PAL genes, AtPAL1 and AtPAL2, are structurally similar to PAL genes from other plants according to their conserved region (Wanner et al. 1995) and actively expressed in seedling development and in roots (Ohl et al. 1990; Wanner et al. 1995; Mizutani et al. 1997; Raes et al. 2003). AtPAL2 and AtPAL4 are also abundantly expressed in siliques and seeds (Mizutani et al. 1997; Raes et al. 2003). By contrast, AtPAL3 seems to be mostly expressed at a lower level and only at basal levels in stems (Mizutani et al. 1997; Raes et al. 2003; Rohde et al. 2004). Kinetic studies of these four AtPALs revealed that AtPAL1, 2, and 4 catalysed deamination of phenylalanine *in vitro* properly, whereas AtPAL3 exhibited only low activity (Cochrane et al. 2004).

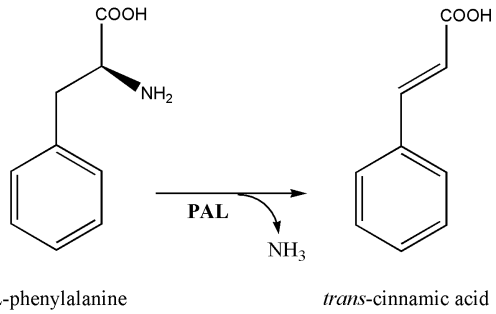
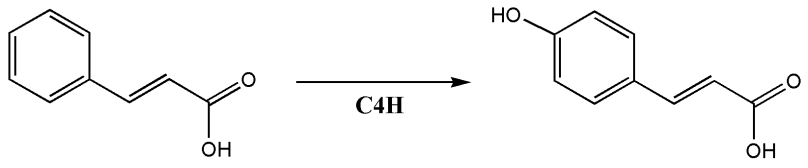


Figure 5. Catalytic activity of PAL, which catalyses the deamination of L-phenylalanine to *t*-cinnamic acid

On the basis of their high expression in the inflorescence stem, AtPAL1, 2, and 4 are believed to be closely associated with lignin biosynthesis (Raes et al. 2003; Rohde et al. 2004; Huang et al. 2010). Meanwhile, according to their phylogenetic relationship and the presence of specific promoter elements, AtPAL1 and AtPAL2 are proposed as the main PAL genes for the phenylpropanoid metabolism (Raes et al. 2003). Undoubtedly, their promoters are powerfully triggered during the lifetime of the plant either in early seedling development or in adult plants. In particular, the expression level in adult plants arises only in the vascular tissues of inflorescence stems, roots, and leaves, whereas in the meristematic tissues, i.e. root tip or shoot apical meristem the expression is low. In addition to their roles in plant development, AtPAL1 and AtPAL2 play a pivotal role in plant defence since their promoters are found to be stress-inducible (Ohl et al. 1990; Wong et al. 2012). The transcript level of AtPAL1 is also up-regulated under low temperatures and light presence which is correlated with the accumulation of anthocyanin in leaves and stems (Leyva et al. 1995). Furthermore, at low temperatures and shortages of nitrogen, the accumulation of some phenolic compounds, such as sinapic acid esters, flavonols and anthocyanins, is closely related to the high expression of AtPAL1 and AtPAL2 (Olsen et al. 2008). Moreover, the double knockout of *pal1* and *pal2* in *Arabidopsis* shows that the mutants accumulate phenylalanine and show a decreased level of their three major kaempferol glycosides, glycosylated vanillic acid, scopolin, and two feruloyl malates coupled to coniferyl alcohol (Rohde et al. 2004), as well as reduction in anthocyanin pigments in various tissues, and thus higher sensitivity to ultraviolet-B light (Huang et al. 2010). AtPAL4 only partially compensates the loss of AtPAL1 and AtPAL2 of the mutants (Rohde et al. 2004). Accordingly, it is clearly strengthened that AtPAL1 and AtPAL2 are the principal PAL genes for flavonoid biosynthesis in *Arabidopsis*.

1.2.3.2 Cinnamic acid 4-hydroxylase (C4H)

C4H (EC 1.14.13.11) is a cytochrome P450-dependent monooxygenase (P450; CYP) that catalyses the conversion of cinnamic acid through *para*- hydroxylation of the aromatic ring to 4-coumaric acid (Fig. 6; Bak et al. (2011)). In *Arabidopsis*, this enzyme exists as the only CYP73 family member and was designated CYP73A5. Catalysing the second step of the phenylpropanoid pathway, C4H is highly active in lignifying tissues which displays its role in supplying monolignols, monomers of lignin. Hence, relatively higher activities of CYP73A5 in inflorescence stems, roots, and siliques than in leaves and flowers reflect such a role (Bell-Lelong et al. 1997; Mizutani et al. 1997). An experiment using C4H promoter-driven β-glucuronidase (GUS) expression indicates that the expression of CYP73A5 is tissue-specific and wound-inducible. Strong GUS staining has been detected in the vascular tissue of stems and leaves as well as in wounded cells of mature leaves (Bell-Lelong et al. 1997). Light also effects a high expression of CYP73A5 (Bell-Lelong et al. 1997; Mizutani et al. 1997). This is closely coordinated with high expression of both PAL1 and 4CL since the C4H promoter region also shows all three *cis*-acting elements (boxes P, A, and L) conserved among the PAL and 4CL genes so far reported controlling their expression (Mizutani et al. 1997). Missense mutations of this enzyme in *Arabidopsis* reduce the levels of numerous classes of phenylpropanoid end-products (e.g. sinapoylmalate) and lignin deposition (Schilmiller et al. 2009). These *reduced epidermal fluorescence 3* (*ref3*) mutants also cause an alteration of lignin monomer content, in which the plant accumulates cinnamoylmalate, a novel hydroxycinnamic acid ester which is not found in the wild type. Such metabolite changes affect directly the developmental, structural, and reproductive phenotypes. For instance, the mutants have severely collapsed xylem vessels due to the reduction of the lignin content. They also fail to develop mature pollen and therefore are male-sterile. Moreover, the mutants are mainly dwarfed and show a loss of apical dominance. Taken together, these abnormalities demonstrate the pivotal role of C4H in physiological processes for plant developmental and survival (Ruegger and Chapple 2001; Schilmiller et al. 2009).

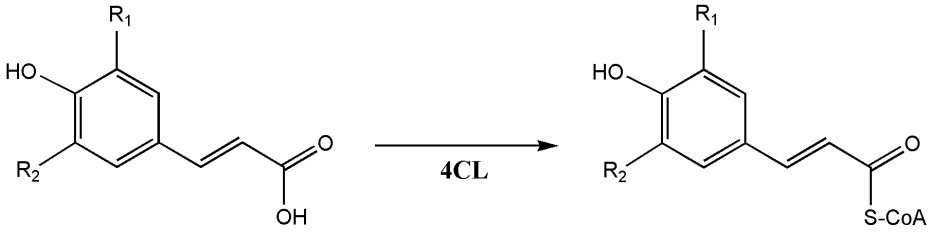


trans-cinnamic acid *para*-coumaric acid

Figure 6. The reaction catalysed by C4H

1.2.3.3 4-Coumarate:CoA-ligase (4CL)

4CL (EC 6.2.1.12) catalyses the formation of CoA esters of 4-coumaric acid, caffeic acid, ferulic acid, 5-hydroxyferulic acid, and sinapic acid (Fig. 7; Lee et al. (1997)). The genome of *Arabidopsis* encodes four 4CL isozymes, At4CL1-4. Each of them exhibits differential behaviour in expression, distinct substrate preferences and specificities (Lee et al. 1995; Ehlting et al. 1999; Hamberger and Hahlbrock 2004). At4CL1 is the first 4CL characterised in *Arabidopsis* and highly expressed in tissues undergoing lignification, i.e. young seedlings, bolting stems of mature plants, and siliques. In addition, its expression is inducible by a variety of stresses including pathogen infection, UV irradiation and wounding (Lee et al. 1995; Mizutani et al. 1997; Ehlting et al. 1999). Such patterns of expression are exhibited similarly by At4CL2 (Ehlting et al. 1999; Raes et al. 2003), an isozyme which is 83% identical to At4CL1. On the other hand, At4CL3, though this isozyme is 61% identical to At4CL1 (Costa et al. 2003), shows differential expression. The expression occurs mostly in siliques (Raes et al. 2003) and is not induced by pathogen attack but is induced by UV radiation. Those three isozymes display preferential activity towards 4-coumarate. However, according to their enzymatic properties, expression characteristics and evolutionary relationships, At4CL1 and At4CL2 are likely to participate in lignin biosynthesis, whereas At4CL3 might have a role in the biosynthetic pathway leading to flavonoids (Ehlting et al. 1999). In contrast, At4CL4 displays preferential activity towards sinapate and ferulate, not towards 4-coumarate, suggesting a different metabolic function in the production of additional phenolic compounds (Hamberger and Hahlbrock 2004).



4-coumaric acid	(R_1=R_2=H)	4-coumaroyl-CoA
caffeic acid	(R_1=H, R_2=OH)	caffeoyl-CoA
ferulic acid	(R_1=H, R_2=OCH_3)	feruloyl-CoA
5-hydroxyferulic acid	(R_1=OH, R_2=OCH_3)	5-hydroxyferuloyl-CoA
sinapic acid	(R_1=R_2=OCH_3)	sinapoyl-CoA

Figure 7. Diversity of substrates accepted by 4CL

A recent finding confirms that At4CL3 belongs to a different clade than the other At4CLs, and that At4CL1 and At4CL2 are more closely related to one another than to At4CL4. Furthermore, At4CL1 accounts for the majority (90%) of total 4CL activity and therefore significantly contributes to lignin biosynthesis, and either At4CL1 or At4CL2 is required for normal plant growth. In contrast, At4CL3 is expressed in a broad range of cell types and has acquired a distinct role in flavonoid metabolism. Both, At4CL3 and At4CL1 are essential and function redundantly in the biosynthesis of sinapoylmalate, the major hydroxycinnamoyl ester found in *Arabidopsis*. The last, At4CL4 has a limited expression profile and contributes modestly to lignin deposition, at least in the absence of At4CL1 (Li et al. 2015).

1.2.3.4 Tyrosine aminotransferase (TAT)

TAT (EC 2.6.1.5) catalyses the pyridoxalphosphate-dependent reaction from L-tyrosine to 4-hydroxyphenylpyruvate (pHPP) (Fig. 8). It is the first enzyme in the biosynthetic pathway leading via homogentisic acid to plastoquinone and tocopherols. Plastoquinone participates in the photosynthetic electron transport chain and plays a necessary role as an enzyme cofactor in carotenoid biosynthesis. In photosystems, carotenoids protect chlorophyll and photosynthetic membranes from oxidative degradation. Tocopherols function as radical scavengers in plants and thus protect the plant in a variety of different stress situations (Sandorf and Holländer-Czytko 2002; Grossmann et al. 2012).

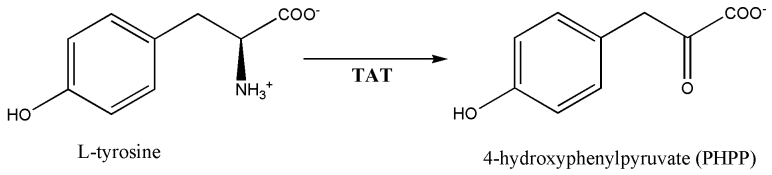


Figure 8. The reaction catalysed by TAT

In *Arabidopsis*, there are six annotated tyrosine aminotransferases (TATs), namely At2g24850, At4g23590, At4g28410, At4g28420, At4g3600, At5g36160, and At5g53970 (TAIR; www.*arabidopsis*.org). So far, only two of them have been identified to have TAT activity, i.e. At5g36160 (Prabhu and Hudson 2010) and At5g53970 (Grossmann et al. 2012; Riewe et al. 2012). The remaining four were computationally predicted. *In vitro,* the At5g36160 isozyme has the ability to convert L-tyrosine and L-phenylalanine*,* whereas At5g53970 converts L-tyrosine

preferentially. *In vivo*, both genes were also able to complement an *E. coli* mutant which is auxotrophic for L-tyrosine and L-phenylalanine (Prabhu and Hudson 2010; Riewe et al. 2012).

At5g53970 is the only gene that has been studied so far for its role and regulation as a TAT in *Arabidopsis*. Because of its role in tocopherol biosynthesis, the expression of At5g53970 is significantly elevated together with the increased level of α- and γ-tocopherol in leaves of aging *Arabidopsis*, supporting the role of this enzyme against, for instance, aging-induced oxidative stress. On the other hand, coronatine, a phytotoxin mimicking octadecanoids and leading to symptoms of senescence, moderately increased the expression of the gene and the level of tocopherols as well (Holländer-Czytko et al. 2005). Another study showed that the activity of this enzyme is inhibited by the herbicides cinmethylin and 5-benzyloxymethyl-1,2-isoxazolines derivatives (Grossmann et al. 2012). In the study of two independent *Arabidopsis* mutants, knocking out At5g53970, the total TAT activity in rosette leaves was reduced up to 50%, suggesting the major contribution of At5g53970 to total TAT activity. Furthermore, the mutation of At5g53970 did not affect the expression of other putative TATs and even the expression of At5g36160, the closest homologue of At5g53970. On the metabolic level, there was a specific increase in tyrosine levels and, as its consequence, the content of tocopherols was substantially reduced. This may show the role of At5g53970 as pivotal for tocopherol biosynthesis (Riewe et al. 2012).

1.2.3.5 Hydroxypyruvate reductase (HPR)

HPR (EC 1.1.1.29) is responsible in plant photorespiration for reducing hydroxypyruvate (HP) to glycerate (Fig. 9; Peterhansel et al. (2010)). Three HPRs are found in *Arabidopsis*, namely NADH-dependent peroxisomal HPR (HPR1), NADPH-dependent cytosolic HPR (HPR2), and NADPH-dependent chloroplast HPR (HPR3). Knock-out of a single HPR in *Arabidopsis* did not result in severe photorespiratory phenotypes. Meanwhile, either double or triple HPR knock-out showed the typical air-sensitivity and drastically reduced photosynthetic capacity of photorespiratory phenotypes. However, they were fully recovered in high-CO_2 conditions suggesting that those HPRs are involved in photorespiration. Of the three HPRs, HPR1 and HPR2 are the most active isoforms in *Arabidopsis*. HPR1 plays a pivotal role under moderate growth conditions whereas HPR2 utilises the excess of hydroxypyruvate leaking from peroxisomes under conditions of very high photorespiratory flow. On the other hand, HPR3 is only responsible for a small proportion of the HP-into-glycerate interconversion if HPR1 and HPR2 are present (Timm et al. 2008, 2011).

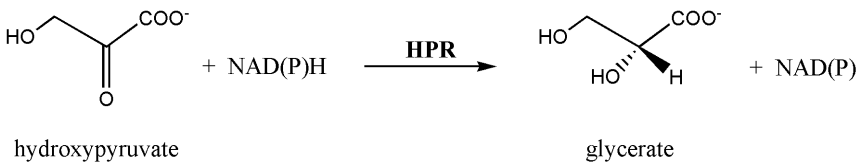


Fig 9. Reaction catalysed by HPR

An interesting hypothesis has been proposed that those photorespiratory enzymes might participate in secondary metabolism. One of them, HPR2, has high sequence similarities (87% on amino acid level) with HPPR from *C. blumei* (Petersen 2013). Heterologous expression of recombinant HPR2 from *Arabidopsis* in *E. coli* showed that this enzyme was capable of accepting hydroxyphenylpyruvate *in vitro*, although *Arabidopsis* cannot synthesise RA *in vivo* (Hücherig 2010).

1.2.3.6 Cytochrome P450-dependent monooxygenases (CYP98A family)

In the biosynthesis of caffeic acid esters, cytochrome P450 monooxygenases of the CYP98A family are responsible for the introduction of the 3-OH group into 4-coumaroyl moieties in ester linkages. In *Arabidopsis*, three cytochrome P450 monooxygenases from the CYP98A family have been identified. The first, CYP98A3 has as specific substrates 4-coumaroylshikimate and -quinate, whereas the other two, CYP98A8 and CYP98A89, have spermidine derivatives as specific substrates (Bak et al. 2011). According to ESTs of *Arabidopsis*, CYP98A3 is one of the constitutively expressed P450 genes. The expression of this gene takes place in all plant tissues. The highest expression occurs in stems and then in roots and siliques, whereas in leaves it is induced by wounding. Since its highest presence was found in lignifying tissues, CYP98A3 is allegedly related to lignin biosynthesis. It was anticipated, that CYP98A3 could be the second P450 enzyme in the phenylpropanoid pathway, also known as 4-coumaroyl ester 3'-hydroxylase (C3'H). This enzyme converts 4-coumaric acid to caffeic acid or the analogous reactions occurring at the level of the corresponding CoA thioesters. However, the yeast-expressed CYP98A3 was not found (or due to low activities) to metabolise free 4-coumaric acid or its glucose or CoA esters, 4-coumaraldehyde, or 4-coumaryl alcohol (Schoch et al. 2001; Nair et al. 2002; Franke et al. 2002b). Nevertheless, the catalytic activity of this enzyme displays high activity towards 5-*O*-shikimate and 5-*O*-*D*-quinate esters of *trans*-coumaric acid (Fig. 10). The shikimate ester was converted four times faster than the quinate derivative (Schoch et al. 2001). Of the two coumaroyl esters, two

products are generated, namely caffeoylshikimic acid and chlorogenic acid, respectively. None of these compounds, however, have been reported to be accumulated in *Arabidopsis*. Apparently, the shikimate esters are considered to be transient intermediates leading to the monolignols, by the fact that shikimate is the preferred 4-coumaroyl acceptor substrate of HCT over quinate and also followed by a higher turnover of CYP98A3 with the shikimate ester than with the quinate ester. As proposed, caffeoylshikimic acid could be then converted back into caffeoyl-CoA as precursor for further metabolic steps (Schoch et al. 2001; Bak et al. 2011; Fraser and Chapple 2011).

The role of CYP98A3 in lignin biosynthesis has been confirmed by characterising a number of *Arabidopsis* mutants displaying a reduced epidermal fluorescence (ref) phenotype. One of them, the *ref8* mutant, exhibited reduced levels of the blue-green fluorescence of its rosette leaves due to the absence of wild-type levels of sinapoylmalate. Instead of sinapate esters, the mutants accumulate 4-coumarate esters, showing the mutation has impaired C3'H activity (Franke et al. 2002b). This had influence on the composition of lignin in which the mutants deposited a 4-coumaroyl alcohol-derived lignin instead of S and G lignin found in normal plants. As result, the mutation leads to secondary phenotypes including vascular collapse, dwarf, and susceptibility to fungal attack, and hyperaccumulation of flavonoids. Therefore, CYP98A3 is essential as a 4-coumaroyl ester hydroxylase in the biosynthesis of soluble phenolics and lignin monomers that are responsible for normal plant growth (Franke et al. 2002a; Abdulrazzak et al. 2006; Li et al. 2010).

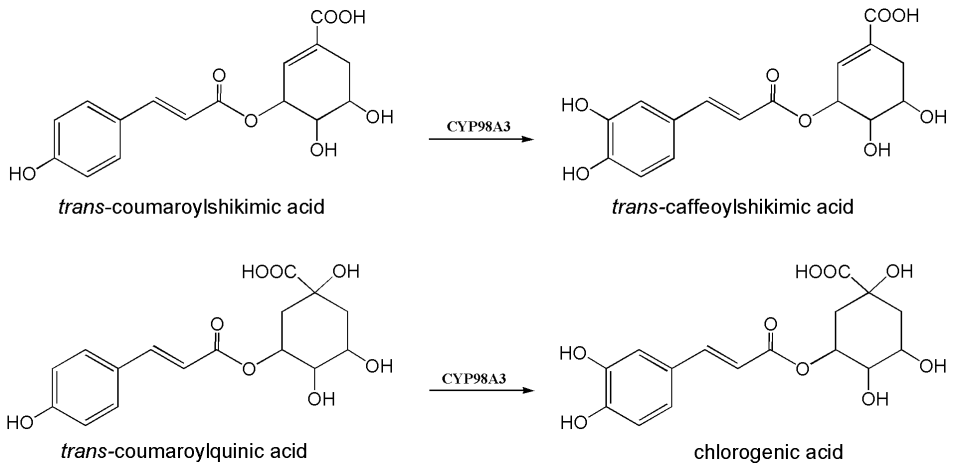


Fig. 10. Reactions catalysed by CYP98A3

The CYP98A8 and CYP98A9 genes are two chromosome 1-clustered duplications of an ancestor of CYP98A3 that occurred via mRNA-mediated transposition (retroposition) and

therefore are free of introns. Their proteins are 69% identical and each of them shares about 50% protein identity with CYP98A3 (Matsuno et al. 2009). Unlike CYP98A3, CYP98A8 and CYP98A9 are needed for synthesising specific phenolic compounds existing only in pollen, rather than in lignin biosynthesis. Analysis of *in silico* transcriptome and promoter:β-glucuronidase (GUS) transformed plant revealed a predominant expression of CYP98A8 and CYP98A9 in floral tissues and stamens, in particular tapetal cells and pollen (Ehlting et al. 2008; Matsuno et al. 2009). Metabolic profiling of wild type and mutant plants disclosed the contribution of both enzymes in the production of N^1,N^5-di(hydroxyferuloyl)-N^{10}-sinapoyl spermidine, a major constituent of the pollen coat. Both CYP98A8 and CYP98A9 share the ability to introduce 3-OH group into mono-, di-, and tri-coumaroyl spermidine (Fig. 11). Further, the second *meta*-hydroxylation of the methylated intermediate, N^1,N^5-N^{10}-triferuloyl spermidine, can be processed only by CYP98A8 (Matsuno et al. 2009). Suppression of enzymes in this pathway, either up- or downstream enzymes, leads to strong perturbations in pollen development and fertility (Fellenberg et al. 2009; Schilmiller et al. 2009).

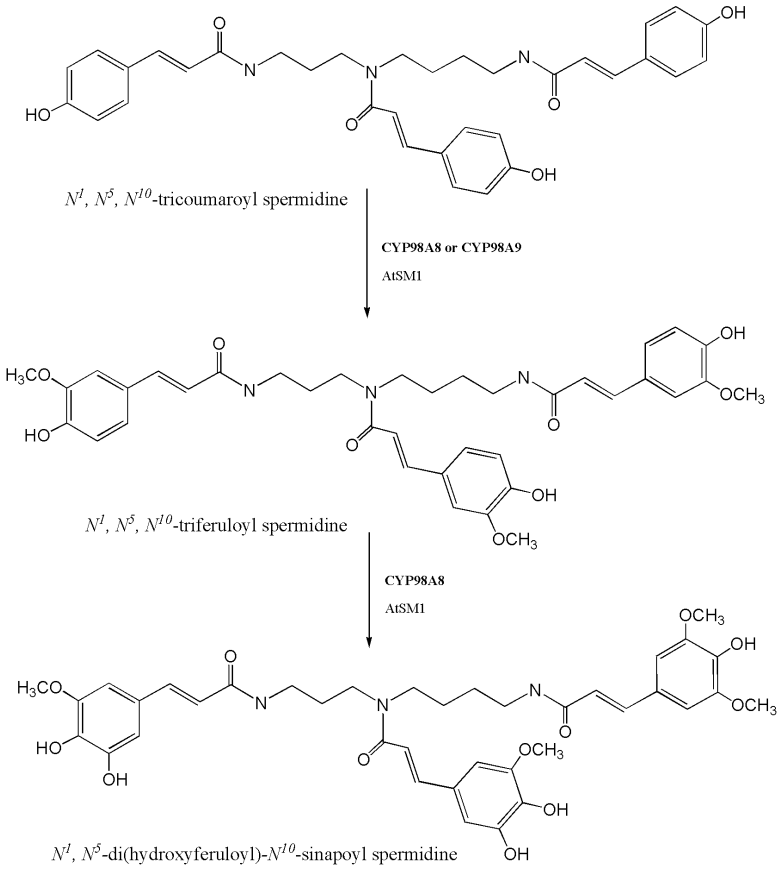


Fig. 11. Reactions catalysed by CYP98A8 and CYP98A9. AtSM1: *A. thaliana O*-methyltransferases

Although CYP98A3, CYP98A8 and CYP98A9 share the ability of *meta*-hydroxylation of coumaroyl esters, they differ in the acceptance of various coumaroyl esters. Neither CYP98A8 nor CYP98A9 has hydroxylase activities toward the shikimate/quinate esters of 4-coumaric acid (Schoch et al. 2001). In parallel, CYP98A3 was also found not to perform on N^1,N^5-N^{10}-triferuloyl spermidine, but a weak activity on N^1,N^5-N^{10}-tricoumaroyl spermidine, however, cannot be ruled out (Matsuno et al. 2009). An interesting result was shown for *in vitro* activities of the recombinant proteins of these three cytochrome P450s expressed in *Saccharomyces*, where they apparently have a broad substrate acceptance. They could also hydroxylate other coumaric acid-based compounds, e.g. umbelliferone, 4-coumaroyltyramine, and even resveratrol that has never been found in *Arabidopsis*. Unfortunately, none of these enzymes were found to hydroxylate 4-coumaroyl-pHPL or caffeoyl-pHPL (Morant et al. 2007). However, there is a slight possibility that the *in vivo* properties of these enzymes could be altered when plants are under stress conditions or when extraneous genes/compounds are present.

1.2.4 Engineering of secondary metabolite pathways in *Arabidopsis*

Arabidopsis has been the subject for studying specialised metabolism including regulation and modification of metabolic pathways to reveal their roles in the ecological interaction, plant growth and development, and even establishing new pathways that previously did not exist in *Arabidopsis*. A few works have shown how this plant successfully generated previous non-existent compounds after expressing transgenes.

The *Arabidopsis* plant has been engineered with FaNES1, a cytosolic dual linalool/nerolidol synthase from strawberry. The gene was constitutively expressed and directed to the plastids by the fusion with a plastid-targeting signal. As expected, both linalool and nerolidol which are found to be emitted only in flowers of wild type were emitted in rosette leaves of transgenic plants (Aharoni et al. 2003). This has also been achieved by expressing the gene with an inducible promoter (Yang et al. 2008). In addition, overexpressing this transgene has led the host plant to produce free, glycosylated, and hydroxylated derivatives of the expected product (Aharoni et al. 2003).

Arabidopsis is not known to have flavone biosynthesis, neither flavone compounds nor flavone synthase activity has been reported to exist in this plant. However, by overexpressing FNS-I, a flavone synthase I from parsley, the transgenic *Arabidopsis* can accumulate the flavone apigenin although only after feeding with naringenin (Yun et al. 2008).

In addition to establishing a new pathway, a few existing biosynthetic pathways in *Arabidopsis* have also been engineered. Overexpression of a bacterial pinoresinol reductase led the transgenic *Arabidopsis* to change its lignan composition (Tamura et al. 2014). In *Arabidopsis*, two UDP-dependent glycosyltransferases (UGTs), UGT72E2 and UGT72E3, are responsible for production of coniferin and syringin, respectively. Overexpression of a recombinant UGT72E2/3, generated by partial domain swapping strategy of the two genes, was able to improve syringin production in the plant (Chu et al. 2014). Moreover, overexpression of GmMYB12B2, the R2R3-MYB transcription factor from soybean, elevated the expression of several flavonoid biosynthetic genes and caused high accumulation of flavonoids in the transgenic *Arabidopsis* (Li et al. 2013).

Dhurrin, a cyanogenic glucoside derived from tyrosine, is a phytoanticipin found in *Sorghum bicolor*. It is known for its activity in plant defence against herbivores through releasing its side chain, in the form of hydrogen cyanide, as a result of β-glucosidase and α-hydroxynitrile lyase activities upon tissue damage. The pathway of dhurrin biosynthesis is constituted by three consecutive genes, namely two cytochrome P450s, CYP79A1 and CYP71E, and UGT85B1, a soluble UDPG-dependent glucosyltransferase (Fig. 12). Part or even the entire pathway for the synthesis of this compound has been transferred from *Sorghum bicolor* to *Arabidopsis*. By expressing CYP79A1 in *Arabidopsis*, the transgenic plants accumulate surprisingly large amounts of 4-hydroxybenzylglucosinolate, a novel glucosinolate, together with three phenolic glucosides in minor levels (Bak et al. 1999). The presence of 4-hydroxyphenylacetaldoxime in CYP79A1-engineered *Arabidopsis* has probably activated some endogenous enzymes of the glucosinolate pathway to metabolise the intermediate into 4-hydroxybenzylglucosinolate, since the pathway also uses oxime intermediates, although they are derived from different amino acids (Soledade et al. 2010). Obviously, there has been a metabolic crosstalk between the engineered and endogenous metabolic pathways (Bak et al. 1999). By introducing the subsequent enzyme CYP71E1, this enzyme can overcome the high activity of pre-existing enzymes from the glucosinolate pathway in metabolisation of the intermediate *p*-hydroxyphenylacetaldoxime. As a result, the formation of *p*-hydroxybenzylglucosinolate occurs only in very small amounts and, in addition, other phenolic glucosides are also increased. They are derived from 4-hydroxymandelonitrile, the intermediate product of CYP71E1, demonstrating the detoxification of xenobiotics in non-producing plants (Bak et al. 2000). However, expression the three genes of the pathway, CYP79A1, CYP71E, and UGT85B1, completely directed the metabolic flux towards the formation of high levels of the cyanogenic glucoside dhurrin (Tattersall et al. 2001), suggesting that the formation of dhurrin in

Arabidopsis needs the presence of UGT85B1, a *p*-hydroxymandelonitrile glucosyltransferase. So far, this is the only report of a new secondary metabolic pathway established in *Arabidopsis*.

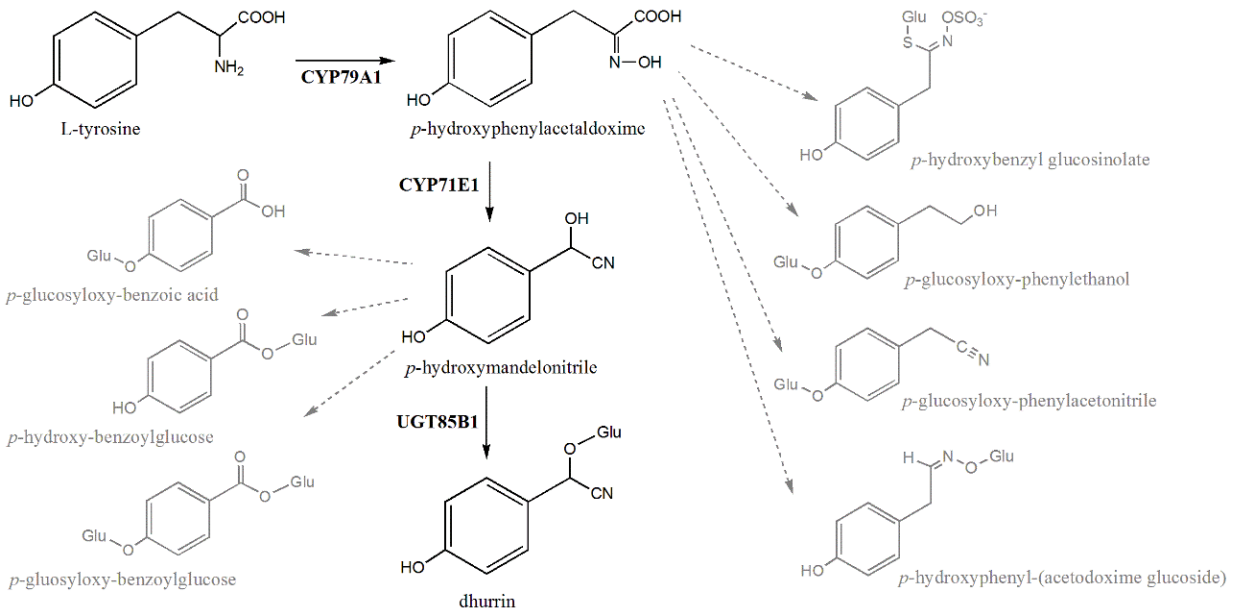


Fig. 12. Biosynthetic pathway of dhurrin (black) and unexpected products (grey) that were present after expressing CYP79A1 or CYP79A1 and CYP71E1 genes in *Arabidopsis*

Apparently, numerous unexpected products are sometimes present in *Arabidopsis* plants after expressing transgenes. Such phenomena are also found in prenyltransferase (PTase)-transgenic plants that do not normally produce prenylated flavonoids. Engineered *Arabidopsis* with *Sf*N8DT-1 cDNA, a plant flavonoid PTase from *Sophora flavescens* that only specifically transfers a dimethylallyl moiety to naringenin, unexpectedly produced 8-dimethylallyl kaempferol and dimethylallylated derivatives of apigenin and quercetin (Sasaki et al. 2008), while *Sf*N8DT-1-transgenic *Lotus japonicus* produced only 8-dimethylallyl naringenin, the specific product of *Sf*N8DT-1 (Sugiyama et al. 2011). Another interesting result has appeared in SCO7190-transgenic *L. japonicus* that produced dimethylallylapigenin and dimethylallyldaidzein after feeding with naringenin, although the SCO7190, a PTase from *Streptomyces*, can only catalyse the prenylation of dihydroxynaphthalenes, naringenin, and isosakuranetin but not daidzein or genistein (Sugiyama et al. 2011). These studies demonstrated that non-producing plants could be hosts for the production of desirable compounds by expressing key enzymes. In addition, the properties of targeted enzymes could be altered in a host species-dependent manner and the detoxification of xenobiotics might also occur, that eventually lead host plants to produce unexpected products.

1.3 Objectives

Older as well as recent investigations have shown that a number of enzymes involved in the formation of RA in plants are also active in other essential (primary) biosynthetic pathways and thus should be present in all higher plant species. Other enzymes of RA biosynthesis are closely related to other biosynthetic enzymes active in photorespiration (HPR2/HPPR) or the formation of monolignols (hydroxycinnamoyltransferases, CYP98A monooxygenases). It can thus be anticipated that the precursors necessary for ester formation by RAS are present also in plants that are free of RA. An idea is to use the model plant *A. thaliana* to investigate whether this plant species that is known to accumulate neither rosmarinic nor chlorogenic acid is able to form RA or RA-like esters after transformation with the cDNA for RAS. These investigations will furthermore shed light on the question whether transport of RA into the vacuole is a prerequisite for RA formation/accumulation.

In this project, suspension cultures of transgenic *Arabidopsis* carrying the RAS gene were established and analysed in detail:

- Activity and expression of RAS gene and ubiquitous genes: PAL, C4H, 4CL, TAT, HP(P)R
- Analysis of phenolic compounds, especially RA and RA-like esters, in transgenic and control cultures

To elaborate this investigation, the *ras*-transgenic *Arabidopsis* plants were also transformed with the HPPR cDNA, a further gene of RA-biosynthesis. The role of HPPR has been shown as a limiting point in RA biosynthesis. Overexpression of *hppr* significantly elevated the level of RA accumulation in both hairy roots of *C. blumei* (Hücherig and Petersen 2012) and *S. miltiorrhiza* (Xiao et al. 2011). Transformation of *ras*-transgenic *Arabidopsis* with HPPR cDNA was expected to lead to an increased formation of pHPL, which eventually might activate RAS leading to the esterification of pHPL with 4-coumaroyl-CoA. An *Agrobacterium rhizogenes*-mediated plant transformation was used to introduce *hppr* into the *ras*-transgenic *Arabidopsis* to generate hairy roots. Series of molecular and biochemical changes in the hairy roots were further investigated.

2. Materials and methods

2.1 Experimental material

Plant materials were transformed seeds of *A. thaliana* with RAS cDNA of *Coleus blumei* (CbRAS). The transformed seeds were obtained by using *Agrobacterium*-mediated plant transformation. The plant vector pFGC5941, *Agrobacterium tumefaciens* strain GV3101 pMP90 and *A. thaliana* cv. Columbia were given by Prof. Dr. Alfred Batschauer. Moreover, the floral dip transformation of plants was also performed in his laboratory (our unpublished results). Two different lines were generated from this transformation i.e. R-lines, plants transformed with pFGC5941-CbRAS and L-lines, plants transformed with empty vector (pFGC5941) as controls.

2.1.1 Establishment of sterile seedlings, callus and suspension cultures

Plant materials (seedlings, callus cultures, and suspension cultures) for further investigation were established under sterile conditions. The establishment of sterile plant material excludes an influence of bacteria or fungi on the results. It has been shown that the biosynthesis of RA works in undifferentiated plant cells and light-independently. Nevertheless, cultures incubated under light and in darkness were established, because photorespiration is light-dependent and thus the reduction of pHPP to pHPL might be more prominent under condition allowing photosynthesis.

Sterile seedlings were grown from the transformed seeds sterilised with the following method. Seeds were soaked in 70% ethanol for approximately 1 min. Then, they were transferred into a 10% commercial chlorine disinfectant solution for 20 min and subsequently washed three times with sterile distilled water. The sterilised seeds were then placed onto solid hormone-free CB2 (CB2OH) medium and incubated in permanent light at 25±1^{o}C. For callus induction, all parts of two-week-old seedling (leaf, stem, roots) were cut into small pieces and incubated on CB2-agar medium (Petersen and Alfermann 1988) in continuous light at 25±1^{o}C. Sub-culturing of callus was performed on the same medium in four-week intervals. Suspension cultures were established by transferring 5 g friable callus to 50 ml liquid CB2-medium in a 250 ml Erlenmeyer flask and shaking at 110 rpm. Suspension cultures were sub-cultivated weekly with 5 g wet cells per 50 ml CB2-medium.

2.1.2 Basic characterisation of sterile transformed plant materials

The sterile transformed plant material was analysed to ensure the absence of agrobacteria and the plasmid pFGC5941, and the presence of the stably integrated T-DNA and the RAS-cDNA. The absence of *A. tumefaciens* was tested by PCR-amplification of VirD2 (Haas et al. 1995), one of the virulence genes of agrobacteria that is not transferred into the host plant genome. The absence of the plasmid pFGC5941 was tested by PCR-amplification of the plasmid-encoded kanamycin resistance gene. The presence of the T-DNA was shown by PCR-amplification of the cauliflower mosaic virus (CaMV) 35S promoter. Finally, the presence of the RAS open reading frame was shown by PCR-amplification of the RAS-encoding sequence. The primer sequences and PCR protocol are described in Section 2.3.

2.2 Extraction of plant genomic DNA

2.2.1 Rapid DNA extraction from *Arabidopsis*

This method is the most suitable way of extracting DNA from a high number of plant samples, since it only needs a tiny part of plant tissues and is simple and quick. In this work, the method was used to rapidly provide a small amount of DNA during basic characterisation of sterile transformed *Arabidopsis* on the level of seedlings (Section 2.1.2).

A piece of a rosette leaf from a two week-old seedling was placed in a 1.5 ml centrifuge tube along with 200 µl extraction buffer (200 mM Tris/HCl pH 7.5, 25 mM EDTA, and 0.5% SDS) diluted 10-fold by TE buffer. Leaves were crushed with a plastic rod against the tube wall. The extracts were then centrifuged at 13000 rpm for 5 min. The supernatants were collected, stored and stable at -20 ^{o}C for at least one month. For PCR analysis, 1 µl of DNA solution was used as template for 25 µl PCR reaction volume (Kasajima et al. 2004).

2.2.2 Plant DNA extraction (CTAB method)

To obtain purified high molecular weight plant genomic DNA, a DNA extraction method based on the CTAB (cetyltrimethylammonium bromide)-method developed by Rogers and Bendich (1985) was adapted in this work. This method was used to isolate genomic DNA from callus and suspension cultures. Fresh or frozen cells were ground into a fine powder in a porcelain mortar by using liquid nitrogen. About 200 mg of ground cell material were placed in a 1.5 ml-centrifuge tube in liquid nitrogen. Then 300 µl of pre-heated 2X CTAB-buffer (65 ^{o}C) were added to the tube

followed by incubation at 65 ºC for 10 to 30 min with intermediate mixing twice. After incubation, the tube was chilled on ice for 5 to 30 min. For chloroform extraction, an equal volume of chloroform (300 µl) was added. After gentle but thorough mixing, the tube was centrifuged at 13000 rpm for 5 min. The top phase of the supernatant was transferred into a new 1.5 ml centrifuge tube containing 30 µl of pre-heated 10X CTAB buffer (65 ºC). A second chloroform extraction was performed with addition of 300 µl chloroform. After centrifugation, 250 µl of the top phase was transferred into a new 1.5 ml centrifuge tube containing 250 µl CTAB precipitation buffer. After thorough mixing, the tube was centrifuged at 13000 rpm for 10 min. A careful decantation was performed to remove the supernatant from the precipitate that adheres to the tube as an opaque thin film. The precipitate was redissolved in 200 µl high salt TE buffer, followed by adding 400 µl ethanol, mixing and incubation at -20 ºC for a minimum of 15 min. After centrifugation at 13000 rpm for 15 min, the precipitate was washed with 70% ethanol twice and dried at room temperature. Finally, the nucleic acid precipitate was dissolved in 20 to 50 µl TE buffer containing 100 µg/ml RNAse A and incubated at 37 to 50 ºC for 15-30 min to remove RNA that might be present. The DNA was stored at -20 ºC until use.

2.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis is a method for separating nucleic acids according to their molecular weight. It uses agarose gel at concentrations ranging between 1 and 3 percent which provide a wide variety of pore sizes. An electric current is applied to move the negatively charged DNA through the gel towards the positive electrode. Smaller DNA molecules generally run through the gel with a higher mobility than larger DNA fragments. After separation, the DNA fragments can be visualised with help of ethidium bromide, the most common method of DNA staining. The binding of ethidium bromide to DNA results a strong fluorescence under ultraviolet light (260-300 nm) where the DNA shows up on the gel as a band of orange fluorescence (Reece, 2004).

In this work, an agarose gel was made by suspending 0.7 to 1.5 g of dry agarose in 100 ml TBE buffer and boiling in a microwave until a clear solution was formed. The solution was poured into a suitable gel chamber containing a comb to form wells, after addition of 7 to 10 µl ethidium bromide (1%), and allowed to cool at room temperature for at least one hour to get a rigid gel. After the gel has set, the comb was removed and the gel was placed into an electrophoresis chamber containing TBE buffer. Samples of DNA (5 to 25 µl) were mixed with 5 µl 6X DNA loading dye and loaded into the wells. The gel was then subjected to a constant electric field (110 to 150 Volt)

for approximately 30 to 60 min. Evaluation of the gel was performed in a gel documentation system and the picture of the DNA bands was taken with the equipped camera.

2.2.4 Measurement of DNA concentration

The concentrations of DNA or RNA are rapidly measured with a photometer. Samples of DNA were diluted with a dilution factor 1:100 and measured at three different wavelengths, i.e. 230, 260 and 280 nm. Samples of DNA or RNA have high purity, as generally accepted, if the ratio A260/280 and A260/230 are in the range of 1.8-2.0 and 2.0-2.2, respectively.

2.3 Polymerase chain reaction (PCR)

2.3.1 General protocol for PCR

PCR was performed with GoTaq polymerase (Promega, Mannheim, Germany) following the manufacturer's protocol,

Reagents	Volume
500 ng DNA (or cDNA)	1.0 µl
5X GoTaq-buffer	5.0 µl
25 mM $MgCl_2$	3.0 µl
10 mM dNTPs	0.5 µl
10 mM *forward*-Primer	0.5 µl
10 mM *reverse*-Primer	0.5 µl
GoTaq-Polymerase	0.1 µl
H_2O	14.4 µl
total volume	25.0 µl

and the amplification program,

Temperature and time program	Cycle
94 °C, 120 s; X °C, 60 s; 70 °C, 90 s	1st
94 °C, 30 s; X °C, 60 s; 70 °C, 90 s	2nd to 39th
94 °C, 60 s; X °C, 60 s; 70 °C, 10 min	40th
6 °C, ∞	hold

where X corresponded to the optimum annealing temperature (T_a) of the primers. The PCR products were evaluated using electrophoresis on a 0.7-1.5 % agarose gel (Section 2.2.3).

2.3.2 Primers for PCR

All primers for PCR were purchased from Eurofins. The optimal primers were selected after nucleotide BLAST alignment (http://blast.ncbi.nlm.nih.gov/Blast.cgi). All primers and their properties are listed below.

Genes	Primer codes and sequences (5'→ 3')	T_a, °C	Expected product size (bp)
Cbras	ras-fw GCAAACGAGCACCACTTATCCGACGGCGTCGCCGCC	68	700
	ras-rv CACCGCCTTCCCCCACCCRAATCCGGCTGRTA		
Cbras (full length)	ras-fw ATGAAGATAGAAGTCAAAGACTCGAC	60	1290
	ras-rv TCAAATCTCATAAAACAACTTCTCAAATCTC		
CaMV 35S	35s-fw TCGACGAATTAATTCCAATCC	58	674
	35s-rv GATGCAATTAGTCCTGAATC		
kanR	kanR-fw GATACGGAAGGAATGTCTCC	57	609
	kanR-rv GCAGAAGGCAATGTCATACC		
virD2	virD2-fw ATGCCCGATCGAGCTCAAGT	50	224
	virD2-rv TCGTCTGGCTGACTTTCGTCATAA		
Cbhppr	hppr-fw ATGGAGGCGATTGGCGTTTTG	55	950
	hppr-rv TCAAACTACTGGAGTTAAGAGAGG		

rol A	rolA-fw CCGGACTAAACGTCGCCGGC rolA-rv GAACGTCCCGGTCGGGCTTG	60	200
rol B	rolB-fw TCTCACTCCAGCATGGAGCC rolB-rv GAGAGTCGCAGGGTTAGGTC	60	650
rol C	rolC-fw ATGCCTCACCAACTCACCAG rolC-rv GACAAGCAGCGATGAGCTAG	60	450
virC	virC-fw ATCATTTGTAGCGACT virC-rv AGCTCAAACCTGCTTC	55	700

2.4 Protein extraction and analysis

2.4.1 Plant protein extraction

Five grams of fresh cells were homogenised with 1 g Polyclar 10 and 5 ml 0.1 M KH_2PO_4/K_2HPO_4 (KPi) buffer pH 7.0 containing 1 mM dithiothreitol (DTT) with help of an UltraTurrax for 3 x 30 s with intermediate cooling on ice. Cell residues were sedimented by centrifugation at 10,000 x *g* for 20 min at 4 ^oC and the protein extract was decanted, used directly for enzyme assays or frozen at 20 ^oC until use.

2.4.2 Protein quantification

Protein concentrations were measured using the Bradford method with BSA as standard (Bradford 1976). The determination of protein concentrations in crude enzyme extracts was performed with duplicate samples. 2 ml Bradford solution was pipetted into a cuvette, followed by addition of 10 µl diluted protein extract (10 µl protein + 40 µl 0.1 M Tris/HCl pH 7.5 buffer) and shaking immediately. As blanks and standards, 10 µl buffer and 10 µl BSA (1 mg/ml) were used, respectively. After 15 minutes incubation, the absorption of the samples was measured at 595 nm.

2.4.3 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is the most widely used method for qualitatively analysing protein mixtures. It is particularly useful for monitoring protein purification, and because the method is based on the separation of proteins according to their size, the method can also be used to determine the relative molecular mass of proteins.

Proteins to be run on SDS-PAGE are fully denatured with help β-mercaptoethanol and SDS and boiling. They open each protein into a rod-shaped structure with a series of negatively charged SDS molecules along the polypeptide chain that swamp the native original charge of proteins. Once a current is applied, the negatively charged protein-SDS complexes pass through the gel towards the anode. However, as they pass through the gel the proteins separate, owing to the molecular sieving properties of the gel. The smaller the protein, the more easily it can pass through the pores of the gel, whereas large proteins are successively retarded by frictional resistance owing to the sieving effect of the gel. The separated proteins can be seen by staining with Coomassie Brilliant blue R-250 as bands. An approximate molecular mass can be determined by comparing to protein bands to those of protein markers with known molecular mass (Walker 2002).

A polyacrylamide gel consists of two parts, i.e. separating and stacking gel, that were set together in one gel cassette. The composition of both gels is described below

Reagents	Separating gel	Stacking gel
1.5 M Tris/HCl pH 8.8	1.25 ml	625 µl
H_2O	1.45 ml	1.4 ml
Rotiphorese® Gel 30 (30% acrylamide and 0.8% *bis*-acrylamide)	2.05 ml	375 µl
10% SDS	200 µl	100 µl
N,N,N',N'-tetramethylenediamine (TEMED)*	8 µl	5 µl
10% ammoniumpersulphate (APS)*	36 µl	20 µl

* TEMED and APS were added immediately to initiate the polymerisation reaction before gels being set into the cassette.

The separating gel mixture was transferred to the gel cassette with a pipet by running the solution carefully down one edge of the cassette until it reached a position 1 cm from the bottom of the comb that will form the loading wells. A small amount of n-butanol was added very carefully on top of the gel and left until the gel completely polymerised. This is to keep out air bubbles from the gel and ensure the gel sets with a smooth surface. As the separating gel has set, the overlaying n-butanol was poured off. Then, the stacking gel mixture was added to the gel cassette until it reached the cutaway edge. A well-forming comb was placed into the cassette and left until the gel set.

After the gel has completely set, the comb and the spacer from the bottom of the gel cassette were carefully removed and the gel cassette was assembled in the electrophoresis tank. The top

reservoir was filled with electrophoresis buffer and inspected for leaks. The bottom tank was also filled with electrophoresis buffer. The electrophoresis buffer contained 25 mM Tris/HCl pH 8.3, 192 mM L-glycine, and 0.1% SDS.

Samples of protein were prepared by mixing 5 µl protein sample with 20 µl Roti®-load buffer. The proteins were then denatured by heating in a water bath at 95 ºC for 5 min, chilled immediately in ice, and shortly centrifuged. The proteins and a protein marker (ROTI®-mark plus standard) were loaded into each well of the gel. The electrophoresis was then performed at a constant voltage of 150V for about 2 hours until the dye, bromophenol blue, reached the bottom of the gel.

To visualise the protein bands, the gel was stained with the staining solution for a minimum of 2 hours with gently shaking and then destained with destaining solution overnight. The destaining solution removed unbound background dye from the gel, leaving stained proteins visible as blue bands on a clear background. The composition of the staining and destaining solution are described as follows:

Reagents	Staining solution	Destaining solution
Methanol p.a.	150 ml	150 ml
H_2O	150 ml	150 ml
Acetic acid	30 ml	30 ml
Coomassie brilliant blue R-250	825 mg	-

2.5 Enzyme assays

2.5.1 PAL assay

Two methods are available to determine the phenylalanine ammonia-lyase (PAL) activity, either by using spectrophotometry or HPLC. Both methods are, however, based on the absorption maximum of *t*-cinnamic acid as a reaction product at 290 nm. The two methods are described in details below.

2.5.1.1 Spectrophotometric PAL assay

The spectrophotometric assay took place at 36 °C in a photometer equipped with a temperature control. The assay was performed in 500 µl reaction volume containing:

Reagents	Probe	Blank
Boric acid-borax buffer pH 8.8	400 µl	450 µl
0.1 M L-phenylalanine in boric acid-borate buffer	50 µl	-
5 mg/ml protein extract	50 µl	50 µl

The protein extract was added into the cuvette after 2 min pre-incubation. Measurement was then started 5 min after mixing and absorptions were recorded at 290 nm, every 60 s (cycle) for 15 min. With the help of the increase in extinction (ΔE) and the molar extinction coefficient of 10 $cm^2/\mu mol$, the activity was calculated by using the following formula:

$$Specific\ activity\ (\mu kat/kg) = \frac{\Delta E * V_r * 10^6}{c_p * V_p * d * t * \varepsilon}$$

in which, ΔE = extinction difference per minute (min^{-1})

V_r = reaction volume (ml)

c_p = protein concentration (mg/ml)

V_p = volume of protein (ml)

d = cuvette thickness (1 mm)

t = measuring time/cycle (s)

ε = extinction coefficient ($cm^2/\mu mol$)

2.5.1.2 PAL assay with HPLC

Assays for HPLC analysis consisted of:

Reagents	Volume
Boric acid-borax buffer pH 8.8	200 µl
0.1 M L-phenylalanine in boric acid-borate buffer	25 µl
5 mg/ml protein extract	25 µl

After an incubation for 10 min at 36 °C, the reaction was terminated by adding 20 µl 6 N HCl. The assays were extracted with 750 µl ethyl acetate twice. After evaporation of the solvent, the residues were re-dissolved in 50 µl 50% MeOH/0.01% H_3PO_4 and analysed by HPLC. The HPLC system was as follows:

Column:	Equisil ODS column (250 mm length, 4 mm diameter, 20 mm pre-column)
Solvent:	60% MeOH with 0.01% H_3PO_4
Flow rate:	1 ml/min
Wavelength:	290 nm
Standard:	25 µM *t*-cinnamic acid

2.5.2 C4H assay

Cinnamic acid 4-hydroxylase (C4H) is one of the cytochrome P450 enzymes which can be found in microsomes, fragments of the endoplasmic reticulum. Hence, to determine the activity of C4H, the protein should be isolated in form of the microsome fraction.

2.5.2.1 Isolation of microsomes

Suspension cells of *Arabidopsis* were filtered and fresh cells were homogenised in a pre-cooled mortar together with 1/5 of the FW of Polyclar 10 and an equivalent volume (1 ml/1 g FW) of buffer [0.1 M Tris–HCl pH 7.0, 1 mM DTT, 1 mM sodium diethyldithiocarbaminate (DIECA)]. The homogenate was centrifuged at 8,000 g for 20 min. The supernatant was adjusted to 50 mM $MgCl_2$ and stirred on ice for 20 min. After centrifugation at 48,000 g for 20 min the sediment containing aggregated membranes was re-suspended in buffer (as above) with the help of a Potter–Elvehjem glass homogenizer and adjusted to a final protein concentration of 0.83 mg/ml (= microsome suspension). Protein concentrations were determined according to Bradford (1976) using 1 mg/ml BSA as standard.

2.5.2.2 Determination of C4H activity

Standard assays were performed in 1.5 ml reaction tubes and contained:

Reagents	Volume
0.1 M Tris/HCl buffer pH 7.5	40 µl
0.1 M DTT	25 µl
5 mM *t*-cinnamic acid (in 50% MeOH)	10 µl
5 mM NADPH	25 µl
0.83 mg/ml microsome suspension	150 µl

The reaction was started by adding the microsome suspension and incubated in a water bath at 25 °C for 10 min. The reaction was terminated by adding 25 µl 6 N HCl, vigorously mixing, and cooling in ice. The assays were extracted twice with 750 µl ethyl acetate. After evaporation of the solvent, the products were re-dissolved in 100 µl 45% MeOH/0.01% H_3PO_4 and analysed with HPLC. The HPLC system was as follows:

Column: Equisil ODS column (250 mm length, 4 mm diameter, 20 mm pre-column)
Solvent: 45% MeOH with 0.01% H_3PO_4
Flow rate: 1 ml/min
Wavelength: 311 nm
Standard: 25 µM 4-coumaric acid

2.5.3 4CL assay

The activity of 4CL can be evaluated photometrically by measuring the formation of the respective coenzyme A (CoA) thioester at 311 nm (cinnamic acid), 333 nm (4-coumaric acid), 346 nm (caffeic acid), 345 nm (ferulic acid) and 352 nm (sinapic acid), respectively. In this experiment, caffeic acid was used as a substrate. The assay took place at 40 °C in a photometer equipped with temperature control. The reaction was performed in a quartz cuvette containing:

Reagents	Probe	Blank
0.1 M phosphate buffer pH 7.5	435 µl	435 µl
10 mM caffeic acid (in 50% MeOH)	25 µl	25 µl
125 mM ATP	10 µl	10 µl
125 mM $MgCl_2$	10 µl	10 µl
5 mg/ml protein extract	10 µl	10 µl
10 mM CoA	5 µl	-
H_2O	-	5 µl

The protein extract was added after pre-incubation for 5 minutes and the reaction was started with addition of CoA. Measurement was then started 5 min after mixing and absorptions were recorded at 346 nm, every 60 s (cycle) for 30 min. With the help of the extinction difference (ΔE) and the molar extinction coefficient of caffeoyl-CoA 18 cm^2/µmol, the activity was calculated using the formula at heading 2.2.1.1.

2.5.4 Assay of tyrosine aminotransferase (TAT)

In a 250 µl total reaction volume, the assay contained

Reagents	Volume
1 M Tris/HCl buffer pH 9.0	163 µl
2 mM pyridoxalphosphate (PLP)	10 µl
400 mM 2-oxoglutarate	20 µl
5 mg/ml protein extract	50 µl
150 mM L-tyrosine (in 0.5 M HCl)	7 µl

The reaction was started by addition of the substrate L-tyrosine, followed by incubation at 40 °C for 10 min. The reaction was terminated by addition 100 µl 6 N HCl, vigorous mixing and cooling in ice. The product 4-hydroxyphenylpyruvate (pHPP) was extracted twice with 750 µl ethyl acetate and the solvent was evaporated. The product was re-dissolved in 50 µl MeOH/0.01% H_3PO_4 and analysed with HPLC. The HPLC system was as follows:

Column: Equisil ODS column (250 mm length, 4 mm diameter, 20 mm pre-column)

Solvent: 35% MeOH with 0.01% H_3PO_4

Flow rate: 1 ml/min

Wavelength: 283 nm

Standard: 100 µM pHPP

2.5.5 HP(P)R assay

Hydroxyphenylpyruvate reductase (HP(P)R) catalyses the reduction of hydroxyl(phenyl)pyruvate to its corresponding lactate. The reaction also needs the equivalent amount of NAD(P)H as a co-substrate. Therefore, the activity can be determined by using either HPLC to quantify the corresponding 4-hydroxyphenyllactate, or a spectrophotometer to monitor the decrease of the co-substrate NAD(P)H.

2.5.5.1 HP(P)R assay with HPLC

Reduction of pHPP as substrate yields 4-hydroxyphenyllactate (pHPL) which can be measured quantitatively at 280 or 220 nm. The assay contained

Reagents	Volume
0.1 M KPi buffer pH 7.0	160 µl
1 mM ascorbic acid	10 µl
0.1 M DTT	10 µl
50 mM NADPH	10 µl
25 mM pHPP	10 µl
5 mg/ml protein extract	50 µl

After 10 min of incubation at 30 ºC, the assay was terminated by adding 50 µl 6 N HCl. The products were then extracted twice with ethyl acetate and dried at room temperature overnight. The products were re-dissolved in 30 µl MeOH/0.01% H_3PO_4 and analysed with HPLC. The HPLC system was as follows:

Column: Equisil ODS column (250 mm length, 4 mm diameter, 20 mm pre-column)

Solvent: 10% MeOH with 0.01% H_3PO_4

Flow rate: 1 ml/min

Wavelength: 280 nm

Standard: 100 µM pHPL

2.5.5.2 Spectrophotometric HP(P)R assay

This method was used based on the reduction of the NADPH concentration during the assay. The reaction was performed in a glass cuvette containing:

Reagents	Volume
0.1 M KPi buffer pH 7.0	250 µl
12.5 mM NADPH	40 µl
Substrate (10 mM lithium pyruvate or 25 mM pHPP)	10 µl
5 mg/ml protein extract	50 µl
H_2O	150 µl

The substrates were added 3 minutes after pre-incubation to start the reaction. The assay took place at 37 ºC in a photometer equipped with temperature control. The measurement was started 2 min after mixing and absorption values were recorded at 380 nm every 30 s (cycle) for 10 min.

2.5.6 RAS assay

The assay for rosmarinic acid synthase (RAS) activity was performed in a total volume of 125 µl containing

Reagents	Volume
0.1 M KPi buffer pH 7.0	50.5 µl
0.1 M DTT	12.5 µl
12.5 mM ascorbic acid	5 µl
100 mM pHPL (in 20% ethanol)	5 µl
2.5 mM caffeoyl-CoA	2 µl
5 mg/ml protein extract	50 µl

The assays were then incubated at 30 ºC for 10 min and stopped by addition of 25 µl 6 N HCl and cooling on ice. The assays were extracted twice with 750 µl ethyl acetate each and the solvent evaporated. Dry residues were dissolved in 100 µl 45% MeOH/0.01% H_3PO_4 and analysed with HPLC. The HPLC system was as follows:

Column:	Equisil ODS column (250 mm length, 4 mm diameter, 20 mm pre-column)
Solvent:	45% MeOH with 0.01% H_3PO_4
Flow rate:	1 ml/min
Wavelength:	333 nm
Standard:	25 µM caffeoyl-pHPL

2.6 Protease inhibitor treatment

Two protease inhibitors were used: phenylmethylsulfonyl fluoride (PMSF), an inhibitor of serine proteases, and Broad Range Inhibitor Cocktail (Ic), an inhibitor of a broad range of proteases (serine proteases, esterases, metalloproteases, cysteine proteases, and trypsin-like proteases). A stock solution of 0.1 M PMSF was prepared in 0.1 M Tris-HCl pH 7.0 (Gindro et al. 2012). Because of its limited water solubility, PMSF was dissolved in a small amount of ethanol before addition to the buffer. A stock solution of 100x Ic was prepared from one vial of Broad Range Inhibitor Cocktail following the manufacturer's instructions (Carl Roth). The 100x Ic stock solution contained 50 mM AEBSF (4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride), 15 µM aprotinin, 100 µM E-64, 50 mM EDTA, and 100 µM leupeptin.

For *in vivo* testing, 500 µl PMSF inhibitor was added to 50 ml 5-day-old cell suspension cultures to get a final concentration of 1 mM. Cells were harvested one day after treatment and directly prepared for the protein extraction as described in Section 2.4.1. For *in vitro* testing, PMSF and Ic were added during the protein extraction at a final concentration of 1 mM PMSF and 1x Ic, respectively. Each inhibitor was added separately before and after cell lysis.

2.7 Gene expression analysis

2.7.1 RNA isolation

Isolation of RNA from plant material was performed with a single step method by Chomczynski and Sacchi (1987). To prevent from any contaminating RNAses, all reagents,

reaction tubes and tips were autoclaved twice, whereas mortars and spatula were heated in an oven at 200 ºC for 2 hours. Fresh or frozen 4-day-old plant material was ground into a fine powder in a porcelain mortar in liquid nitrogen. An about 50 mg of pulverised cells were weighed into a 2 ml-centrifuge tube. 500 µl solution D (4 M guanidine thiocyanate; 25 mM Na-citrate pH 7.0; 0.5% (w/v) lauryl sarcosine) were added and incubated shortly at room temperature with often mixing. They were then mixed with 50 µl 2 M Na-acetate pH 4, 500 µl citrate buffer-saturated phenol, 100 µl pre-chilled chloroform, vortexed for 10 s, and incubated on ice for 15 min. After incubation, the mixtures were centrifuged at 12000 g, 4 ºC, for 15 min and the supernatants (400 µl) were transferred to new 1.5 ml-centrifuge tubes, followed by adding one volume (400 µl) isopropanol and incubation at -20 ºC for at least 15 min. RNA was precipitated by centrifugation at 12000 g, 4 ºC for 15 min. The precipitate was washed consecutively with 70% and absolute ethanol, followed by centrifugation at 7500 g for 5 min in between. The RNA precipitate was dried at 37 ºC for 5 min and dissolved in 20 µl H_2O at 50 ºC. The concentration and quality of the isolated RNA were determined with photometer and agarose gel electrophoresis, respectively. The RNA was then converted directly into cDNA and/or stored at -80 ºC until use.

2.7.2 cDNA synthesis

The isolated RNAs were transcribed into cDNA with qScript™ cDNA Supermix (Quanta Bioscience). The cDNA synthesis was performed in 20 µl reaction volume containing 4 µl 5X qScript cDNA Supermix and variable volumes of RNA template and RNAse/DNAse-free water to obtain 500 ng/µl as the final concentration of the RNA. The reaction was done in a thermal cycler following the reaction step 25 ºC 5 min; 42 ºC 30 min; 85 ºC 5 min; 4 ºC. The cDNA was stored at -20 ºC until use.

2.7.3 Real time quantitative PCR

The quantitative PCR (qPCR) was performed with the Perfecta SYBR® Green Supermix (Quanta Bioscience). The reaction took place in a 96-well microplate where each well contained 6.25 µl 2X Perfecta SYBR Green SuperMix, 0.25 µl of each forward and reverse primers, 0.75 µl H_2O and 5 µl cDNA. The reaction was performed in a real time PCR cycler (PIKOREAL, Thermo) following the reaction step 95 ºC 2 min, 50 cycles at 95 ºC 15 s; 50 ºC 30 s; 68 ºC 30 s (data acquisition), and determination of a melting curve at 50-95 ºC. Actin2 and 40sr16 were used as housekeeping genes.

2.7.4 Primers for real time PCR

All primers for quantitative PCR were designed through an online free software provided by Eurofins (https://ecom.mwgdna.com). Both, primers and PCR products were designed to have GC-contents between 50 to 60% and 20 to 80%, and melting temperatures (Tm) between 50 and 60°C and 70 and 95°C, respectively. PCR products were adjusted to have product sizes between 100 and 120 bp. The optimal primers were selected after nucleotide BLAST alignment (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using a data set of *A. thaliana*. All primers have a similar annealing temperature (T_a) at 50°C.

Genes	Primer codes and sequences (5' → 3')	Product size, bp
Atpal1	Atpal1-fw TCGCAGAGGTGATGAGTG	117
	Atpal1-rv AGCTTCCGTCGAGGATATG	
Atpal2	Atpal2-fw CGAAGAGTATCGTAGACCAGTC	112
	Atpal2-rv AGTCTCCGCTAACTCAACC	
Atpal3	Atpal3-fw CAACAGCATTGAGCGACC	115
	Atpal3-rv AGCTGCACCGTTCCTTTC	
Atpal4	Atpal4-fw CGATCCGTTGAACTGGAAC	113
	Atpal4-rv TCTCACCTCCTAACTTCACC	
Atc4h	Atc4h-fw AGGAACGAACTCGACACAG	119
	Atc4h-rv GGAATCGCCATTCTCAGAC	
At4cl1	At4cl1-fw GATGCTGACTCACAAGGGAC	106
	At4cl1-rv GGGCAAAACACAGAGTATGACG	
At4cl2	At4cl2-fw GTGCTAAGTTTCCTAACGCC	120
	At4cl2-rv CACATGCTCCTGACTTCAC	
At4cl3	At4cl3-fw ACCGATACATCCGTCGTC	119
	At4cl3-rv CAGTAAGTGTGGAGAGGGAG	

At4cl4	At4cl4-fw CGTCATCCTCTGTTTTCTCC	118
	At4cl4-rv CACCAGATTCAACTCGAACC	
Athct	Athct-fw GTCTACTTCTACAGACCCACC	116
	Athct-rv TCTCTCTTCAAGCGACCAG	
Atcyp98a3 (Atc3h3)	Atcyp98a3-fw GGACACGACAGCGATAACAG	112
	Atcyp98a3-rv GATCCGGTCAAGTCCAACC	
Atcyp98a8	Atcyp98a8-fw GTTCCTGGAGAGAGGATTGAC	108
	Atcyp98a8-rv AGGTATTGCTCTTGCGGC	
Atcyp98a9	Atcyp98a9-fw GACCACCAACGCGATTTC	120
	Atcyp98a9-rv GCTGTGATCCTAACCACACC	
Attat1 (At5g53970)	Attat1-fw CACCTTGAAGTTCGCTACGTCG	100
	Attat1-rv CCAAAGCAACCGTGTTTTCGTC	
Attat2 (At5g36160)	Attat2-fw GCTTTCGTACCGATCAAGCC	104
	Attat2-rv CTTCCTAGCGACAGGAACAC	
Attat3 (At4g23600)	Attat3-fw GCCTCCTTAGGCTCATATACC	106
	Attat3-rv CGGCTGTATTGCTGGTCTC	
Athpr1	Athpr1-fw CTCACATTGCTTCTGCTTCC	107
	Athpr1-rv GTTCGGGTCATGCCAAATC	
Athpr2	Athpr2-fw CCATTAAGCCTGATGTCGC	118
	Athpr2-rv CGGTCCACAATGTGTCTG	
Athpr3	Athpr3-fw TCTCCGTCGTATTCCTGCC	111
	Athpr3-rv CCAACTATCCCAACTCTCTTCC	
Atact2	Atact2-fw TGAGACCTTTAACTCTCCCG	105
	Atact2-rv ACACCATCACCAGAATCCAG	

At40s16	At40S16-fw TTCCAGCCAGAGATCCTCC	118
	At40S16-rv CCTGTGATGTGTGACCACC	
Cbras	Cbras-fw CTGGATGAGGACTACCTGC	109
	Cbras-rv TCACCCACAGATTAGGGC	
Cbact	Cbact-fw CTTATGTTGCCCTGGACTATGAGC	120
	Cbact-rv GGCAACGGAATCTCTCAGCTCC	

2.8 Transformation of *ras*-transgenic *Arabidopsis* with help of *Agrobacterium rhizogenes* carrying pART27-*hppr*

Overexpression of multiple genes of RAS biosynthesis, *hppr* + *ras*, in *Arabidopsis* was achieved through an *Agrobacterium*-mediated plant transformation. In this work, the *Agrobacterium* strain used was *Agrobacterium rhizogenes* LBA15834 carrying a plant expression vector, pART27, which has been inserted with the full length HPPR cDNA of *C. blumei* controlled by the CaMV 35S promoter (Hücherig 2010). Sterile seedlings of *ras*-transgenic *Arabidopsis* (R13 line) grown on CB2OH medium for 3 weeks were used as explants for the transformation. Leaves were cut from the plant and the leaf surface was infiltrated with pART27-HPPR-transformed *A. rhizogenes* suspension grown overnight under sterile condition. In addition, leaves were transformed with *A. rhizogenes* carrying the empty pART27 vector as control. The leaves were blotted on sterile filter paper to remove the excess of agrobacteria and transferred onto CB2OH agar plates. After incubation in the dark for 48-52 h, the transformed leaves were washed with 500 mg/l cefotaxime to remove the growing agrobacteria. The leaves were further kept on CB2OH agar containing 500 mg/l cefotaxime to prevent growth of agrobacteria. Single hairy roots from the leaves at an approximate length of 10 mm were cut and placed on CB2OH with 500 mg/l cefotaxime and 60 mg/l kanamycin for selection of positively transformed roots. The concentration of cefotaxime was reduced every week by transferring roots into new medium containing 400, 200, and 100 mg/l, consecutively. The agrobacteria-free hairy roots were transferred to 50 ml CB2OH medium and maintained by transferring 0.5 g root material into the fresh liquid medium every 21 days. The roots were cultivated on a rotary shaker at 110 rpm in the dark.

Some basic characterisation of sterile hairy roots were carried out by evaluating the presence of *rol A, rol B, rol C, hppr* and *ras* genes and the absence of *Agrobacteria* (*VirC* or *VirD2*

gene). The PCR protocols as well as the primers are described in Section 2.3. In addition, extraction of protein was performed as described in Section 2.4 and the protein was then tested for its HPPR and RAS activity as described in Section 2.5. RNA of hairy roots was extracted and reverse-transcribed into cDNA for the analysis of the expression of *hppr* and *ras* genes

2.9 Secondary metabolite analysis

2.9.1 Extraction

After separating cells from medium, the cells were frozen at -20 ^{o}C and lyophilised for 16 h. The lyophilised cells were ground to a fine powder. 1 g pulverised plant material was suspended in 20 ml 70% EtOH and mixed vigorously. The extraction was accelerated by sonication at 70 ^{o}C for 10 min twice with vigorous mixing in between, extracts were then centrifuged at 3000 *g* for 10 min. and concentrated by using a vacuum rotary evaporator until the volume of the extracts was reduced to 5 ml. 500 µl concentrated extract was dissolved in 500 µl 100% MeOH/0.6% CH_3COOH and centrifuged again at 16,000 g for 5 min before HPLC analysis.

2.9.2 HPLC

HPLC was performed as described in Section 2.2.6 but using 50% MeOH/0.3% CH_3COOH instead of 45% MeOH/0.01% H_3PO_4 as eluent. Extracts were analysed by comparing with standards including RA, caffeoyl-pHPL, 4-coumaroyl-pHPL and 4-coumaroyl-DHPL. Since the concentration of the expected products was very low and therefore could not be detected, the HPLC fraction from 3 to 10 minutes elution time was then collected and concentrated to remove the excess of methanol by using a vacuum rotary evaporator. The water-containing fractions were extracted twice with ethyl acetate and the ethyl acetate fractions were then evaporated to dryness. The dried residues were stored at -20 ^{o}C until use.

2.9.3 LC-MS

For mass spectrometric analysis the dried residues were re-dissolved in 100 µl solvent (50% MeOH/0.3% CH_3COOH), in which the final volume represented about 1 g of dry weight of cells per ml solvent. In addition, the solvent and authentic standards i.e. pHPL, RA, caffeoyl-pHPL, 4-coumaroyl-pHPL, and 4-coumaroyl-dHPL were included for the analysis.

Agilent HPLC series 1200 (Böblingen, Germany) alone and with a micrOTOF-Q III spectrometer with an ESI source (Bruker, Bremen, Germany) were used for analysis of the concentrated fraction. Analysis on LC-MS was carried out on a CS Multospher 120 RP 18 column (2 × 250mm, 5 µm) applying a linear gradient of 5–100 % acetonitrile in water, both containing 0.1% formic acid, in 40 min and a flow rate at 0.25 mL/min. The column was then washed with 100 % acetonitrile containing 0.1% formic acid for 5 min and equilibrated with 5 %, acetonitrile in water for 5 min. The parameters of the spectrometer were set as following: electrospray negative ion mode for ionization, capillary voltage with 4.5 kV, collision energy with 8.0 eV.

3. Results

3.1 Characterisation of *in vitro* plant material

Plant *in vitro* cultures of *Arabidopsis* either transformed with T-DNA carrying the RAS gene under the control of the CaMV 35S promoter or without additional genetic information as control have been successfully established. This is a suitable way of getting enough plant material and of ensuring the continuous supply of plant cells which have uniform traits. The genetic characterisation was performed to assess the genetic stability of transformed cells. Only transformed cells carrying pre-defined genes were continuously grown as cell suspension cultures. However, not all selected cells grew well as cell suspension cultures, some of them were unsuccessful to multiply or did not grow well. A few healthy growing cell lines were then characterised for their growth profiles, including the biomass and protein production.

3.1.1 Sterile seedlings, callus and cell suspension cultures

To characterise the alteration of the molecular and biochemical properties of the transformed *Arabidopsis*, sterile seedlings, callus and suspension cultures were cultured *in vitro* under sterile conditions. Transformed seeds from 21 R-lines and 15 L-lines were successfully germinated on CB2OH medium (Fig. 13). After two weeks of germination, five seedlings of each line were transferred onto CB2 medium for callus induction. In just three weeks, 105 seedlings of R-lines and 80 seedlings of L-lines have successfully developed into either friable or compact calli. However, after genetic characterisation, only 22 and 33 of the respective lines were continuously cultivated on the CB2 medium. Throughout several times of sub-culturing, the calli were re-characterised genetically to ensure the genetic stability of the transformed cells. The selected calli were transferred into the liquid CB2 medium to initiate cell suspension cultures. Nevertheless, in the end there were only 8 R-lines and 4 L-lines that have been established as cell suspension cultures, i.e. R2.4, R10.1, R13.1, R13.2, R15.3, R16.2, R17.6, L1.5, L7.2, L9.1, and L11.4.

Figure 13. Sterile seedling, callus and suspension cultures of the transformed *Arabidopsis*

3.1.2 Molecular profiles of sterile transformed plant materials

Molecular characterisation was performed to ensure that the selected transformed *Arabidopsis* is genetically stable which included the analysis of the presence of the stably integrated T-DNA, and the absence of agrobacteria and the plasmid pFGC5941. The PCR characterisation was performed by using the genomic DNA isolated from callus cultures as the template, the plasmid pFGC5941-CbRAS and water as the positive and negative controls, respectively. The vector map (Fig. 14) shows that the RAS-cDNA is part of the T-DNA from the plasmid pFGC5941-CbRAS. Accordingly, this gene was found to be present in nearly all R-lines. This was indicated by the presence of a 700 bp product of the RAS gene (partial) (Fig. 15a). The presence of the T-DNA in L-lines was confirmed by PCR amplification of the CaMV 35S promoter. The presence of the respective 674 bp product in some of the L-lines has been confirmed (Fig. 15b). These results show that the respective T-DNA has been stably integrated in several lines. Other lines that did not show any band were excluded for further experiments.

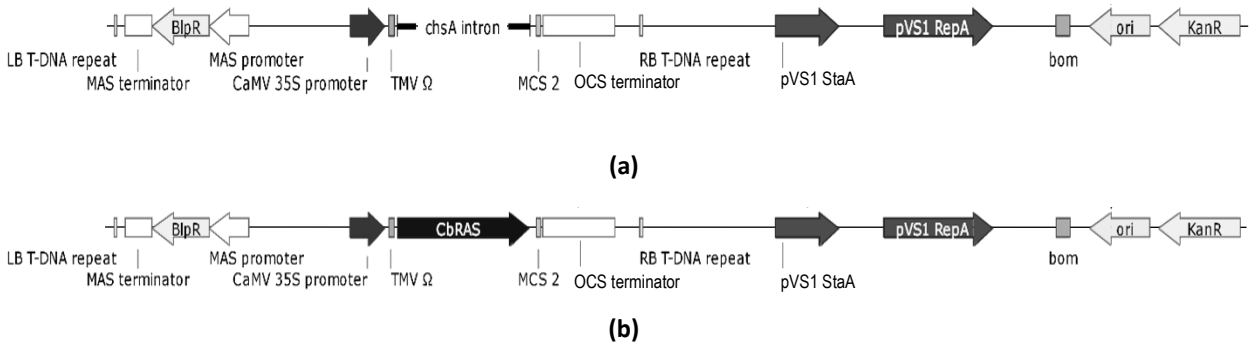


Figure 14. Plant vector constructs of pFGC5941 (a) and pFGC5941-CbRAS (b). LB T-DNA repeat: left border repeat from nopaline C58 T-DNA, MAS terminator: mannopine synthase terminator, BlpR: confers resistance to bialophos or phosphinothricine, MAS promoter: mannopine synthase promoter, CaMV 35S promoter: strong constitutive promoter from cauliflower mosaic virus, TMV Ω: translational enhancer from the tobacco mosaic virus 5'-leader sequence, chsA intron: chalcone synthase A intron from *Petunia hybrida,* MCS 2: multiple cloning site, OCS terminator: octopine synthase terminator, RB T-DNA repeat: right border repeat from nopaline C58 T-DNA, pVS1 StaA: stability protein from *Pseudomonas* plasmid pVS1, pVS1 RepA: replication protein from *Pseudomonas* plasmid pVS1, bom: basis of mobility region from pBR322, ori: high-copy-number ColE1/pMB1/pBR322/pUC origin of replication, KanR: confers resistance to kanamycin

The absence of agrobacteria was tested by using the presence of VirD2 gene, one of the virulence genes in the Ti-plasmid of *Agrobacterium*. Therefore, cells of *A. tumefaciens* GV3101 pMP90 were used directly as the positive control. The PCR amplification confirmed that only a few lines still contained the VirD2 gene and most of them are free from agrobacteria (Fig. 15c). Furthermore, the absence of the intact plasmid pFGC5941-CbRAS was tested by using the KanR gene (*nptII*) as indicator, since it is not part of the T-DNA (see Fig. 14a). However, amplified products of 609 bp were found to be present in all lines and in the positive control (pFGC5941-CbRAS) as well, but neither in the negative control (H_2O) nor in the wild type (Fig. 15d).

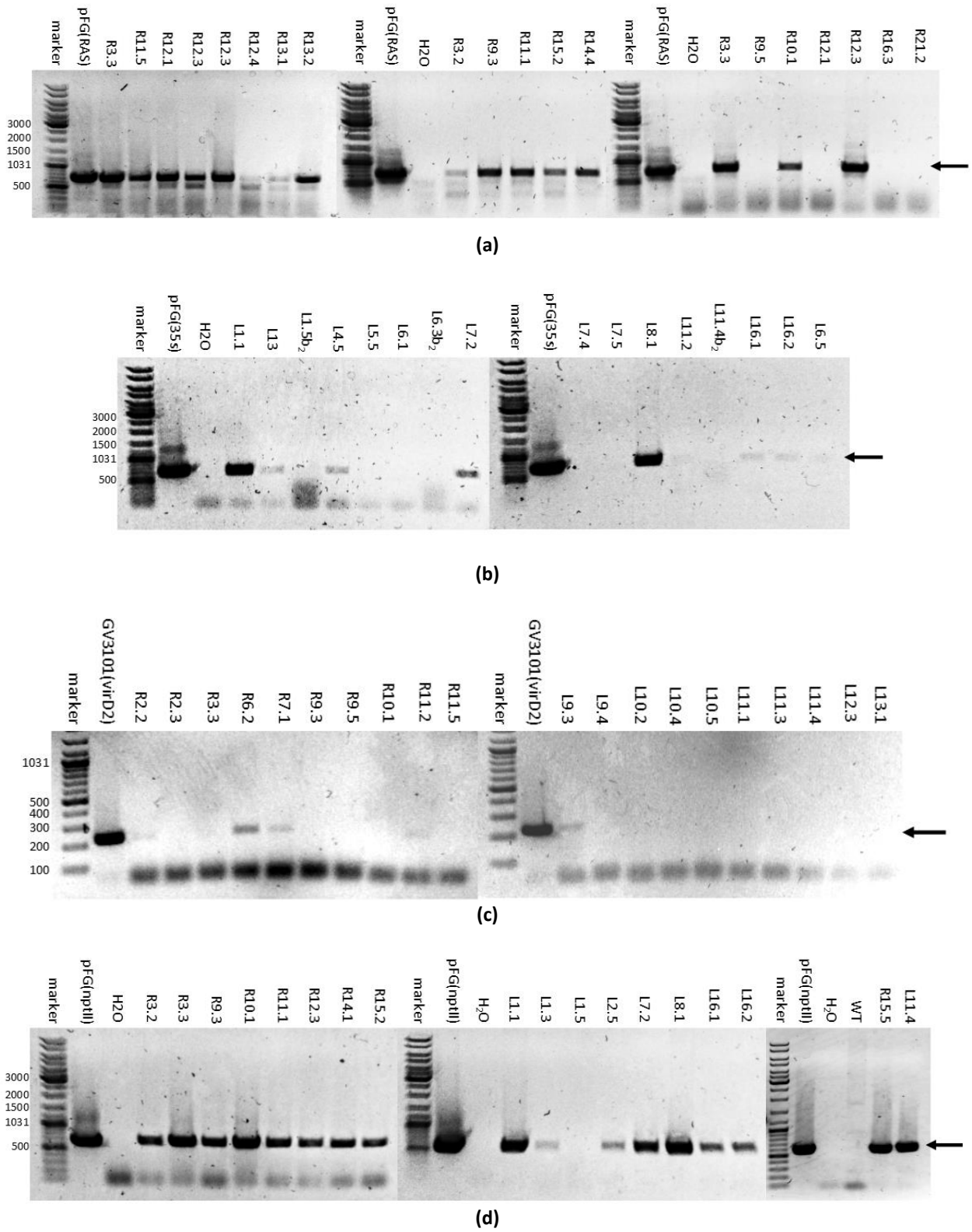
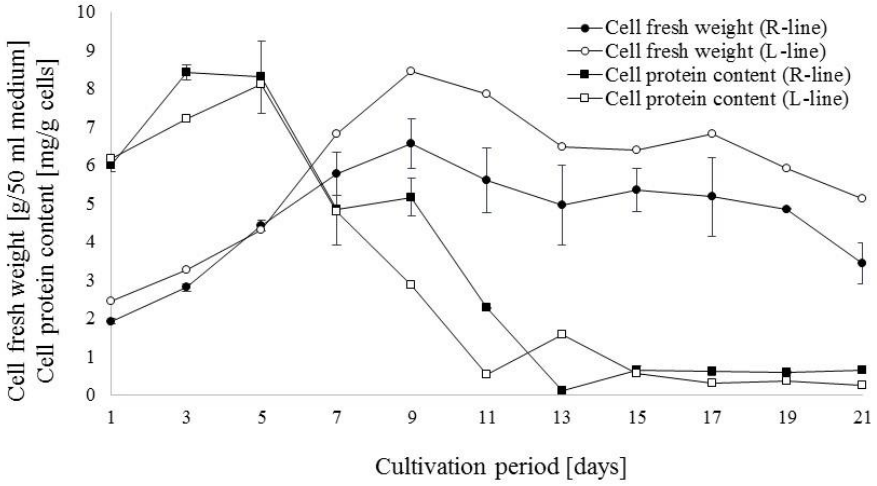


Figure 15. PCR amplification of partial sequences of RAS gene, *Cbras ~ 700 bp* (a), CaMV 35S promoter, *35s ~ 674 bp* (b), VirD2 gene, *virD2 ~ 224 bp* (c), and kanamycin resistance gene (*nptII*), *kanR ~ 609 bp* (d) with gDNA as the template. Marker 1 kb. Positive controls were pFGC5941-CbRAS (pFG) and *A. tumefaciens* (GV3101). Arrows indicate the expected amplification products.

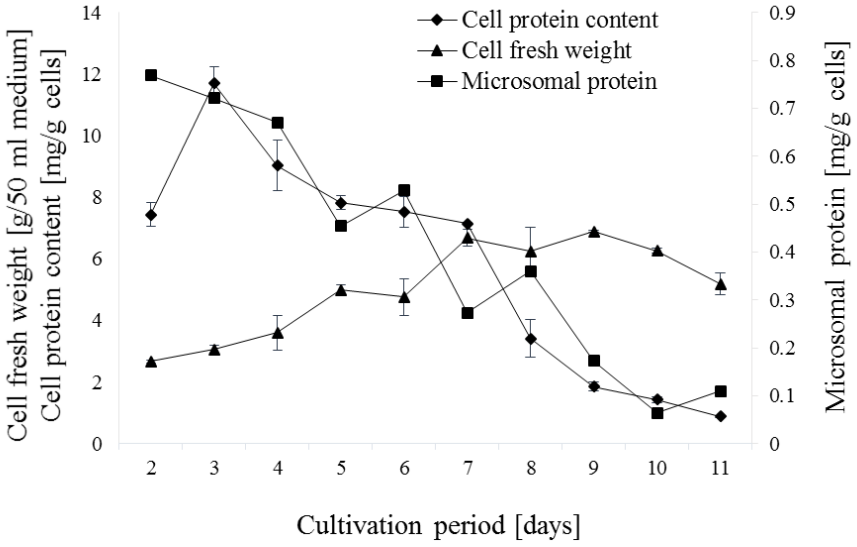
3.1.3 Growth profiles of cell suspension cultures

The established cell suspension cultures were characterised for their growth profiles. Figure 16a shows the comparison between two lines, R- and L-line which are represented by R13.2 and L11.4, respectively. The characterisation included the profiles of fresh weight and protein content of the cells during an observation period of three weeks with measurements every second day. The fresh weight of the cells started increasing rapidly from day 1 and reached the maximal value on day 9. They decreased slowly afterwards because cells underwent browning and lysis until the end of the observation period. No different pattern between R and L line has been observed in this regard, except a difference with respect to the total of cell biomass. Differences were observed with regard to the protein content, where the R-line reached the maximal value on day 3 while the L-line reached it on day 5. The protein content of both decreased significantly after day 5 and showed extremely low levels after day 11. These results suggest that the cultivation period should not exceed one week.

Since the R-line was the main object of the study, the same characterisation has been profoundly performed by using only R13.2 cell suspension cultures. The suspension cultures were cultivated for 11 days with the measurement of three flasks every day. Figure 16b shows that the growth profiles are quite similar. The fresh weight of cells increased from 2.5 g/flask (50 g/l) on day 2 to 7 g/flask (140 g/l) on day 9. The protein content also displayed such pattern, where the highest amount of protein was found to be 12 mg/g cells on day 3 but significantly decreased afterwards. The profile of the microsomal protein exhibited an interesting pattern. In contrast to the total protein, the content of the microsomal protein began to decrease gradually on day 2, even though the cells have already set to multiply on that day.



(a)



(b)

Figure 16. Growth profiles of the cell suspension cultures. (a) Fresh weight and protein contents of R13.2 and L11.4 during 21 days of the observation period and (b) fresh weight, protein contents, and microsomal protein of R13.2 during 11 days of the cultivation period. Data are shown as mean value ± standard deviation where each value is the average of three individual harvests, except for L-line and microsomal protein contents where the value is from one individual harvest.

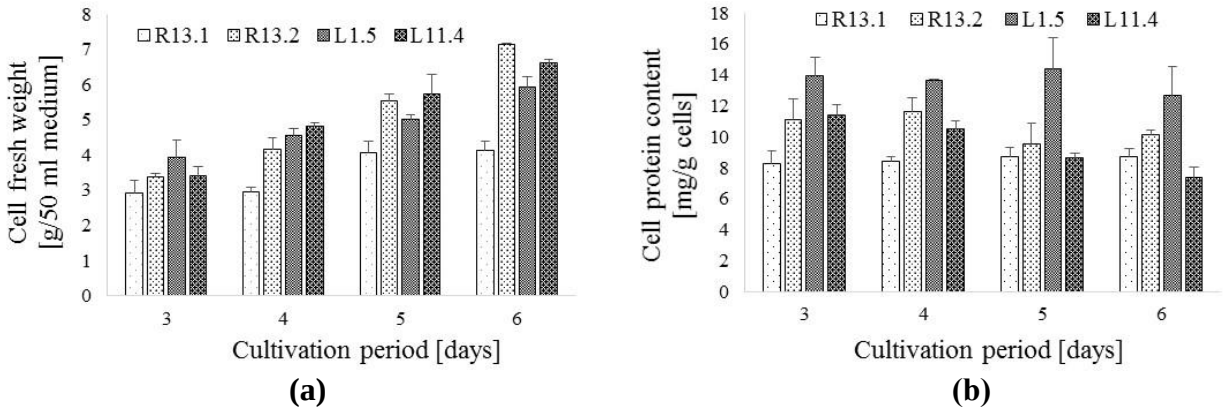


Figure 17. Growth profiles of the suspension cultures during the cultivation period. (a) fresh weight and (b) protein contents

Further analysis (see Fig 17) was performed with two cell suspension cultures of each line, R-line (R13.1, R13.2) and L-line (L1.5, L11.4). The profiles of fresh weight and protein contents were recorded during the logarithmic growth (3^{rd} – 6^{th} day). No change of the growth pattern was observed in all lines during the growth phase, they kept growing day after day. However, the difference lay on the type of lines and the individual level. The biomass was accumulated higher in L-lines than in R-lines (see Fig 16a and 17a), suggesting that the overexpression of foreign genes affected the growth of the cells. At the individual level, some differences were observed, even in these two individuals (R13.1, R13.2) that were established from the same line (R13). It shows that each individual has its own traits and therefore possesses a unique growth pattern. At the protein level, each individual displayed a different pattern as well (see Fig 17b). Two different lines, L11.4 and R13.2, had their maximum value on day 3 and 4 and then it decreased slowly on the following day (see also Fig 16b). In contrast, the other lines, L1.5 and R13.1, had no change in their protein contents. An interesting finding is that the overexpression of the RAS gene in the R-lines does not elevate the protein content as expected, even though the gene is expressed by a strong constitutive promoter. So, there might be some unknown regulation at the molecular level.

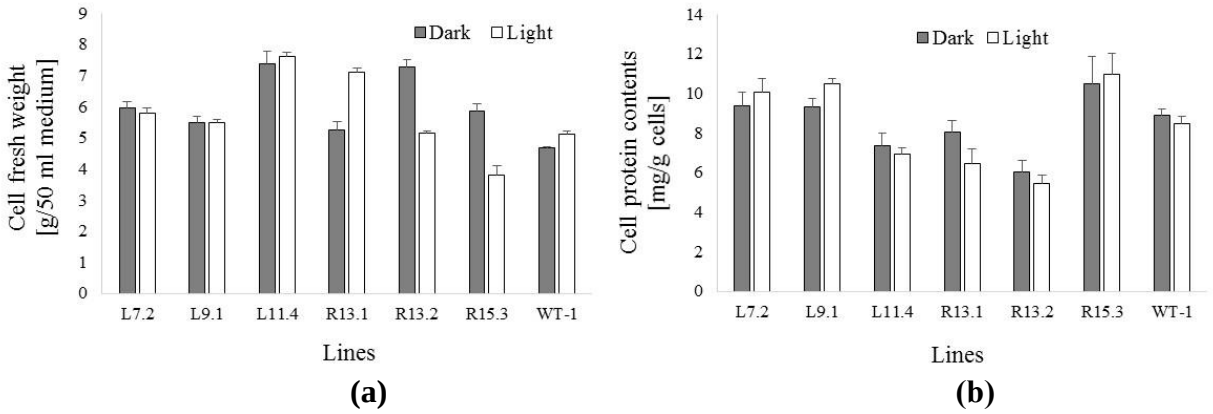


Figure 18. Effect of light on the growth profiles of cell suspension cultures. (a) Fresh weight and (b) protein contents

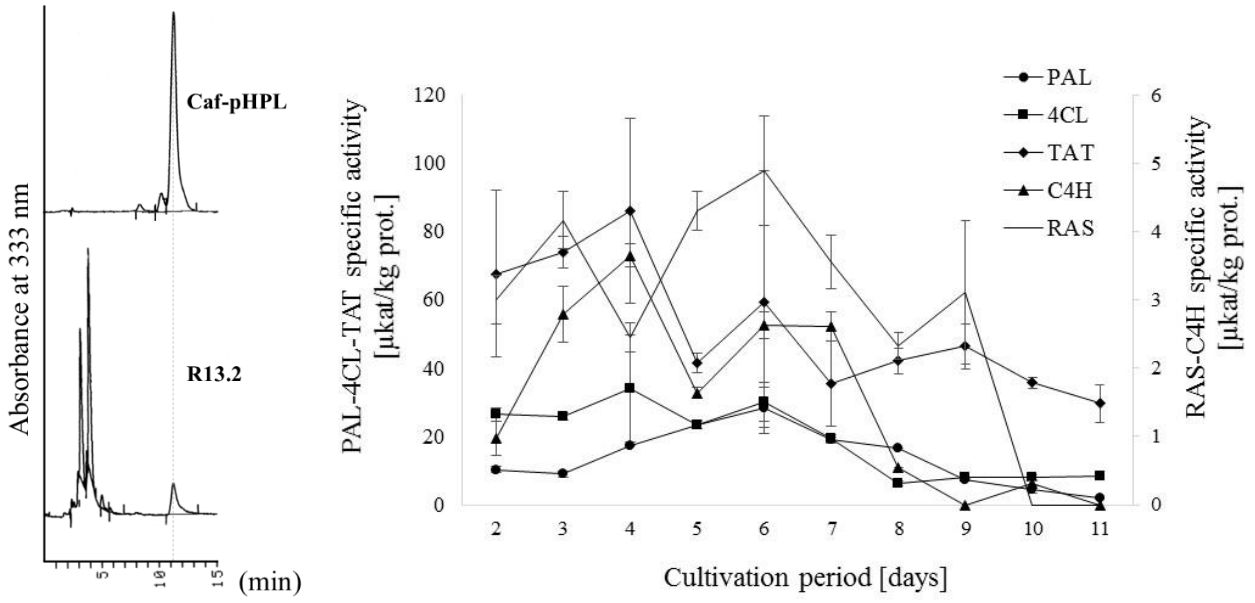
The effect of light on the cell growth was tested by continuously incubating cell suspension cultures in light and dark. Three L-lines (L7.2, L9.1, L11.4), three R-lines (R13.1, R13.2, R15.3), and one wild type (WT-1) have been characterised for their fresh weight and protein contents. Three flasks of each line from the two different conditions were collected on day 5 of the cultivation period. No significant difference of fresh weight was observed for the L-lines and wild type with regard to the presence of light (Fig. 18a). The light only affected R-lines but in the uneven way. In the dark, R13.2 and R15.3 had a relatively higher amount of biomass than in light. In contrast, the accumulation of R13.1 biomass was higher in continuous light. In addition, light was found to activate the photosynthetic activity since a few cell suspension lines were green cells, indicating the presence of chlorophyll. Likewise, the effect of light was found to be not influential on the protein content of the cell suspension cultures (Fig. 18b). These findings suggest that the cell growth and protein production are not severely affected by light.

3.2 Activities of RA-biosynthetic enzymes

3.2.1 Activities of PAL, C4H, 4CL, TAT and RAS during the growth phase

The R13.2 line was the first established cell suspension culture from R-line with RAS activity. The crude protein extract of R13.2 cell suspension culture showed the ability to catalyse the formation of caffeoyl-pHPL from the RAS subtrates caffeoyl-CoA and pHPL (Fig. 19). This revealed that the integrated RAS gene has been successfully expressed in the R13.2 line. As the representative of R-line, cell suspension cultures of R13.2 were then used to study the activity of RA biosynthetic enzymes. The enzymes included PAL, C4H, 4CL, TAT and RAS. Figure 24

displays the activity of each enzyme during the observation period of 11 days. Three enzymes, i.e. 4CL, C4H, and TAT, were found to be highly active at the beginning of the observation period where they reached their maximal activity on day 4 and decreased slowly afterwards. Unlike 4CL and C4H where the activity was almost depleted on day 9, the activity of TAT decreased more slowly and remained stably even until the end of the observation period. On the other hand, the two remaining enzymes i.e. PAL and RAS reached their optimum activities on day 6 of the observation period. The activities afterwards were reduced gradually and depleted at the end of the observation period.



(a) **(b)**

Figure 19. Activity of RA biosynthetic enzymes in R13.2 cell suspension cultures. (a) HPLC chromatogram of products of the RAS assay and (b) profiles of specific activities of RA biosynthetic enzymes.

3.2.2 HP(P)R activity

HP(P)R is the remaining enzyme in the upstream biosynthetic pathway of RA that is not included in Fig. 19b due to the failure in measuring an activity of HPR2 in protein extracts from the cell culture, although two methods have been used to assay the activity. The photometrical measurement showed that the amount of NADPH determined the decreased of absorption independent from the concentration of lithium pyruvate (Fig. 20a). This method failed to quantify the activity as shown in Fig. 20b, where no change of ΔE was observed in the reaction with or

without the presence of substrates. Also, the HPLC method failed to identify pHPL, the product yielded from the conversion of pHPP. As previously described in Weitzel (2009), the product was too difficult to be separated even with the very polar solvent, 10% MeOH with 0.01% H_3PO_4 (Fig. 20c). In addition, the conversion rate of the enzyme seemed to be quite low. Crude enzyme extracts from *Arabidopsis* and from the RA-producing plant, *Coleus blumei*, were tested to see their ability to convert pHPP to pHPL. As previously reported, the HPR2 of *Arabidopsis* has the ability to convert pHPP to pHPL as well. However, the crude enzyme from *C. blumei* also did not show any measurable activity. Presumably, the problem with the substrate affects the activity since pHPP forms keto-enol tautomerisation in aqueous solution.

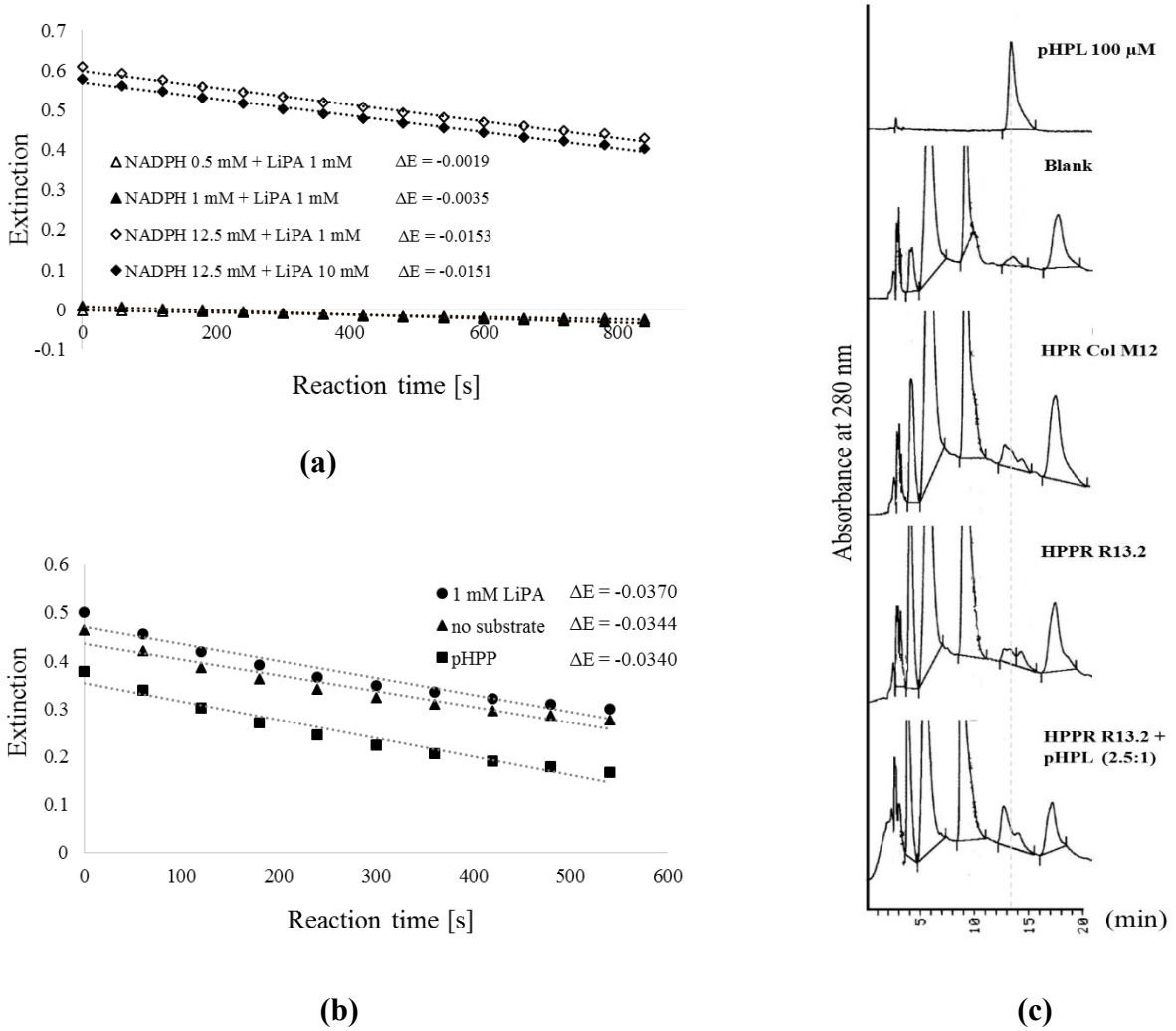


Figure 20. The HP(P)R activity assay. Spectrophotometric method, where the changes of extinction (ΔE) are displayed based on (a) the concentration of the substrate and co-substrate and (b) type of the substrate. (c) HPLC method showing the chromatograms of the authentic standard, blank, enzyme products from R13.2 and *C. blumei*, and enzyme products with the addition of pHPL standard.

3.2.3 Activities of PAL, C4H, 4CL, and TAT on cultivation days 3-6

Several established cell suspension cultures of R- and L-lines were then evaluated for their enzymatic activities. Two cell suspension cultures of each respective line were assayed for their PAL, 4CL, TAT and C4H activities from day 3 to day 6 of the cultivation period. Overall, the highest activity of PAL was found to be either on day 5 or day 6 (Fig. 21a). The 4CL activities of all lines remained stable during the observation period. Even though there were changes between days, the fluctuation was not significant (Fig. 21b). For TAT, the activities of all lines were found to be highly active either on day 4 or 5 of the cultivation period. No different pattern of the TAT activity was observed between R- and L-lines, except L11.4 that displayed higher activity than the others (Fig. 21c). Similar results were observed with respect to the C4H activity, where the highest activity was found on day 4 of the cultivation period, but no different pattern was displayed between R- and L-lines as well (Fig. 21d). These results suggest that overexpressing the RAS gene has no influence to genes encoding enzymes in the upstream biosynthetic pathway. The activity of the enzymes is found to be cell line-specific.

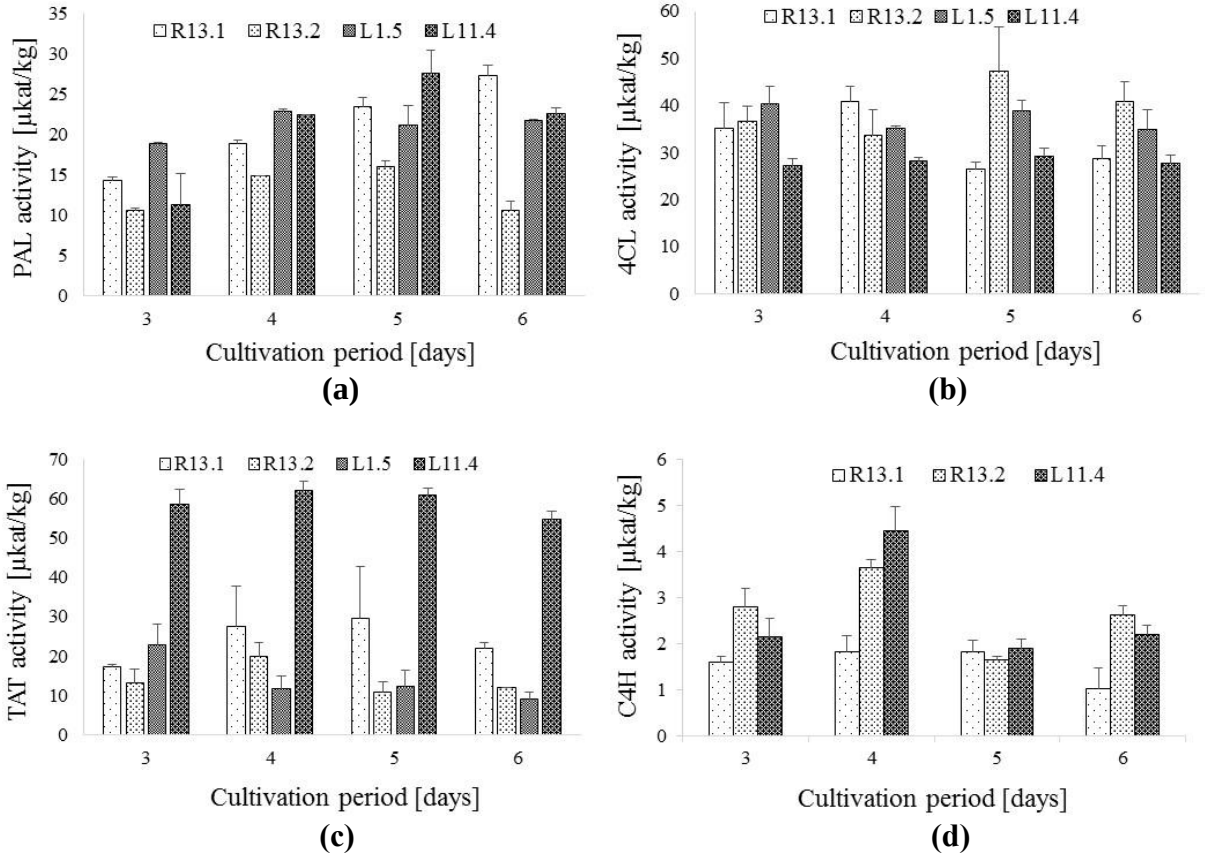


Figure 21. Profiles of specific activities of four enzymes from a two R- and two L-lines, (a) PAL, (b) 4CL, (c) TAT, and (d) C4H.

3.2.4 Influence of light on the activities of PAL and TAT

The presence of light has been reported to increase the activity of two enzymes, PAL and TAT. Therefore these two enzymes were evaluated for their activity by incubating cell suspension cultures under dark and light conditions during the cultivation period. In L-lines, the activity of PAL was found to be higher in the dark than in the light. In contrast, the wild type and all R-lines showed the opposite, where higher activity of PAL occurred in the presence of light (see Fig. 22a). Light showed no influence on the TAT activity in L-lines. However, the presence of light increased the TAT activity in all R-lines and the wild type (Fig. 22b).

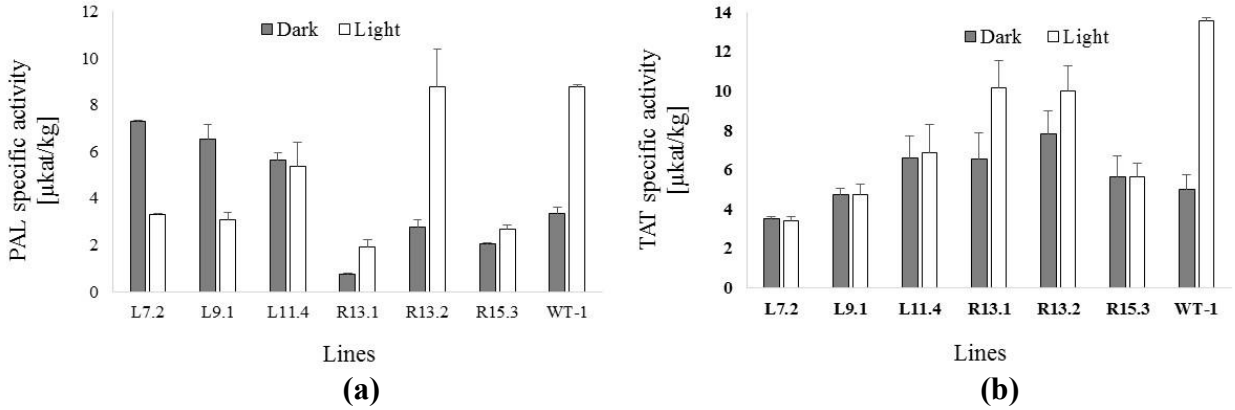


Figure 22. Effect of light on the specific activity of two enzymes, PAL (a) and TAT (b)

3.2.5 RAS activity: in depth study

Since this work has been focusing on the overexpression of RAS in RAS-transformed *Arabidopsis*, a comprehensive study has been performed to evaluate the established suspension cultures from R-lines with respect to RAS activity. These established suspension cultures were all R-lines that have been re-characterised on the callus-phase and known to have stably integrated T-DNA (RAS cDNA) and RAS activity as well, namely R2.4, R10.1, R11.1, R13.1, R13.2, R15.3, R16.2, and R17.6. However, after several times of sub-culturing, five of them i.e. R2.4, R10.1, R11.1, R16.2, and R17.6 did not show RAS activity anymore. A series of molecular identification experiments have been performed to answer this problem. The PCR result using gDNA from cell suspension cultures confirmed that the integrated RAS cDNA is still replicated in this phase, even in the R-lines that showed no RAS activity (Fig. 22). It was supposed that the cells underwent a post-transcriptional gene silencing (PTGS), a condition where cells failed to express a particular

gene on the transcriptional level. However, the PCR result of the cDNA transcribed from the total RNA of the cells showed that the mRNAs of RAS gene were present in all R-lines suggesting that PTGS was not a cause in this regard.

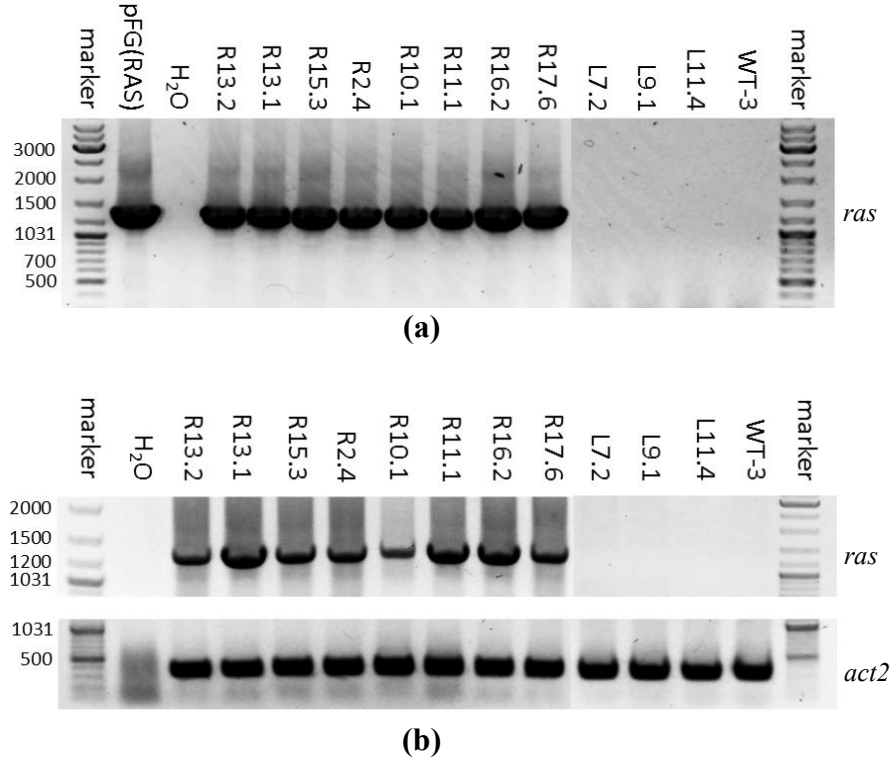


Figure 22. PCR amplification of RAS gene sequences from (a) gDNA and (b) cDNA isolated from the cell suspension cultures, where *Cbras ~ 1290 bp;* and actin, *Atact2 ~ 105 bp*

Two other possible reasons for the lack of RAS activity could be the presence of natural inhibitors or a mechanism called post-translational modification. The natural inhibitors might be present in *Arabidopsis* and therefore inactivate the enzyme. It was then proved by mixing the crude enzyme preparation of each R-line with the recombinant enzyme purified from the CbRAS-engineered *E. coli* (4:1, v/v). However, no change in the activity has been observed between the enzyme mixtures and the recombinant CbRAS alone.

Furthermore, SDS-PAGE has been performed to identify the presence of the RAS protein in the crude extracts. Here, protein extracts from two different separation methods, i.e. centrifugation and filtration, were investigated. Separation with a nylon filter was carried out to anticipate possible inclusion bodies of RAS protein that might occur in plant cells. Though the expression of RAS is controlled by a strong promoter, no thick band at an expected molecular mass of approximately 48 kDa, however, could be seen in R-lines but not in L-lines (Fig. 23). It was not

possible to detect a prominent band in crude protein extracts due to a multitude of proteins with similar molecular mass. Here, advanced protein analysis, e.g. by Western blot, would be necessary which could not be performed without antibodies against the RAS protein.

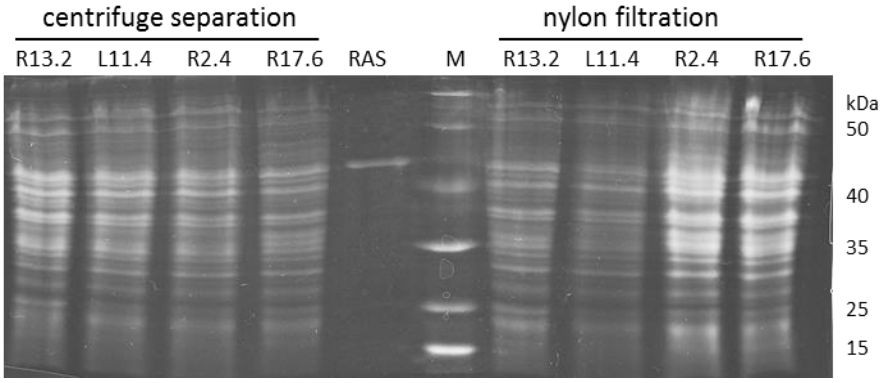


Figure 23. SDS-PAGE of crude protein extracts separated by centrifugation and filtration. CbRAS ~ 48 kDa.

To investigate a possible protein degradation, about 1 mM phenylmethylsulfonyl fluoride (PMSF), a serine protease inhibitor, was added to cell suspension cultures, one day prior the harvesting time (on day 5 of cultivation period). However, *in vivo* treatment with the inhibitor made the activity lower than in the untreated cells (Fig. 24a) because the inhibitor caused a severe damage to the cell growth (data not shown) and therefore resulted in decrease of the RAS activity. Furthermore, the effect of an inhibitor cocktail (Ic) with broad range protease inhibitory activity was tested *in vitro* by adding 1 mg/l Ic to the enzyme preparation directly after protein extraction. This inhibitor has a powerful activity by inhibiting many types of proteolytic enzymes. Fig. 24b shows that the addition of Ic prevents the loss of RAS activity in the enzyme preparations of lines R13.2 and R15.3. In the following week, these two inhibitors were then compared for their inhibitory activity on the enzyme preparations of all R-lines. As shown in Fig. 24c, Ic performed better than PMSF in preventing the loss of RAS activity after protein extraction. Since the plant proteases may only act after the breaking of the cell compartments, the addition of the inhibitors was performed prior and after cell lysis. The result shown in Fig. 24d suggested that the addition of the inhibitors prior to cell lysis saved the RAS activity even better than after cell lysis. Nevertheless, as shown in Fig. 24b and 24c, treatments with protease inhibitors did not recover RAS activity in all the R-lines that had already lost it. In addition, the loss of RAS activity was observed in the R13.2 and R15.3 lines one week after the above-mentioned experiments, because

RAS activity could no longer be detected and it has not recovered until now. Overall, the addition of protease inhibitors only prevented about 20% loss of RAS activity. These results suggest that the proteolytic degradation is not the causing factor of the post-translational modification in this regard. There might be other modification such as glycosylation, phosphorylation, nitrosylation, acylation, alkylation that actually need appropriate methods to be investigated.

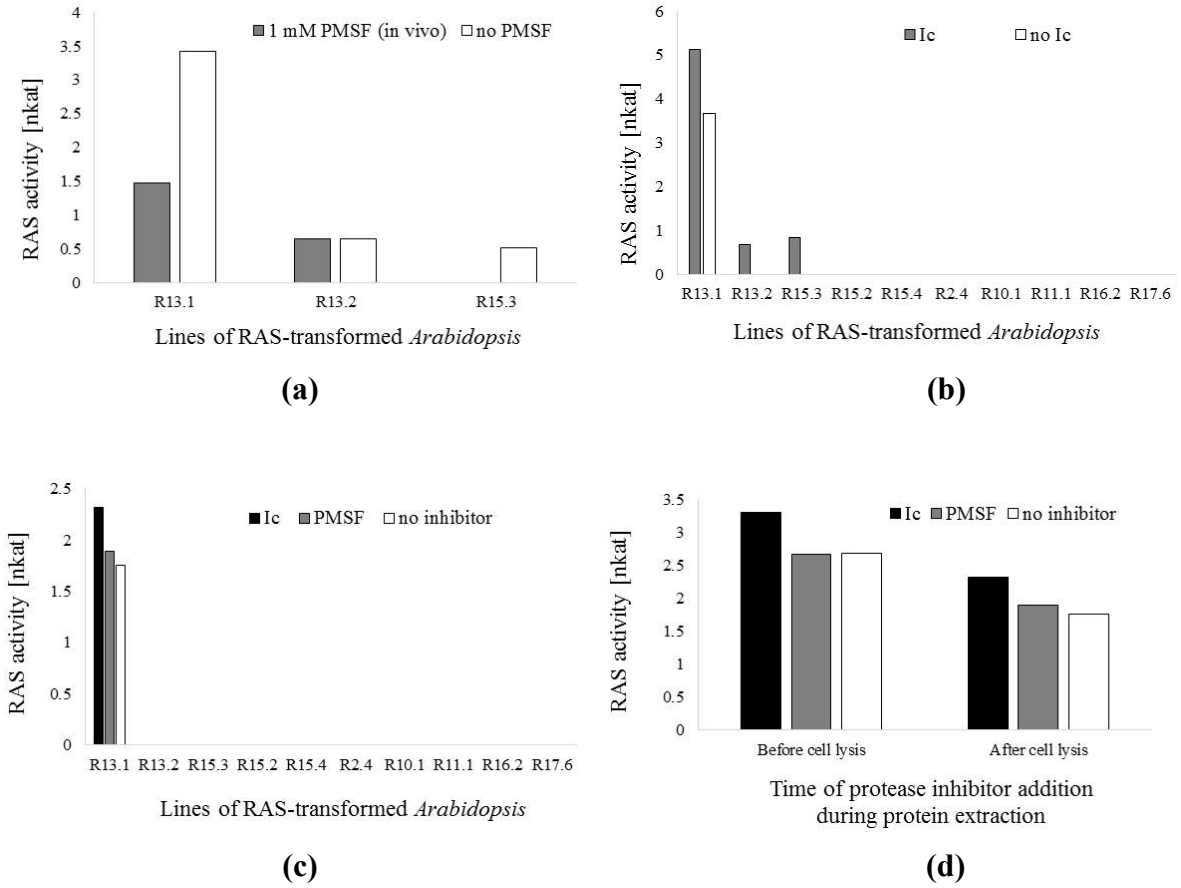


Figure 24. Effect of protease inhibitors on the RAS activity. (a) *in vivo* treatment with PMSF in the cell suspension cultures, (b) *in vitro* treatment with Ic in the protein preparation, (c) comparison of RAS activities after Ic and PMSF treatment and (d) RAS activities based on the addition of the inhibitors, prior or after cell lysis.

3.3 Expression of RA biosynthesis genes

As already described, many enzymes participating in the RA biosynthetic pathway have been confirmed to be present in *Arabidopsis* through the enzyme assays. In addition, the presence of RAS mRNA in the R-lines was positively confirmed by PCR of the cDNA, though as a qualitative result (see Fig. 22b). Furthermore, the quantitative real time PCR (qPCR) was carried

out to analyse the expression level of the respective genes. The expression of the RAS gene was analysed from the mRNA of cell suspension cultures harvested on day 5 of the cultivation period. The expression level of RAS was found to be higher in all R-lines than in the suspension cultures of *C. blumei* (Fig. 25), but the enzyme activity was higher in *C. blumei* than in *Arabidopsis*. Moreover, among the R-lines, only R13.1 had a relative expression change above 50-fold while the expression of others was far below 20-fold. This could be the reason why R13.1 was the only R-line displaying RAS activity, although it cannot explain why *C. blumei* showed higher RAS activity.

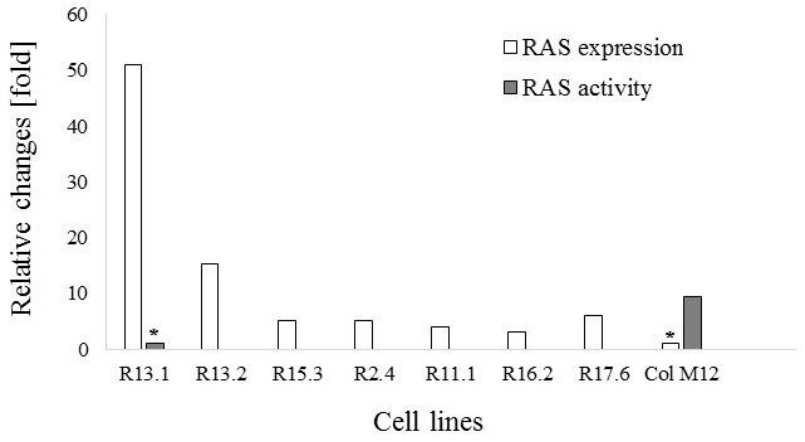


Figure 25. Relative changes of RAS gene expression and RAS activity in R-lines of *Arabidopsis* and *C. blumei* suspension cultures. *calibrator

The expression changes of genes involved in the RA biosynthetic pathway were analysed in three cell suspension cultures each of L-lines (L7.2, L9.1, L11.4) and R-lines (R13.1, R13.2, R15.3). The expression changes of every gene were based on the expression of the gene from the wild type (WT) suspension cultures. All mRNAs were isolated on day 4 of the cultivation period. The expression changes of PAL genes (four *pal* isoforms, *pal1-4*) are shown in Fig. 26b. The expression of *pal1* and *pal3* was found to be variable. The expression of *pal2* was suppressed in both lines except in R13.2 where the expression remains stable. On the other hand, the expression changes of *pal4* were found to be upregulated in both lines with no exception. This is the only *pal* isoform where the expression was found to be consistently upregulated.

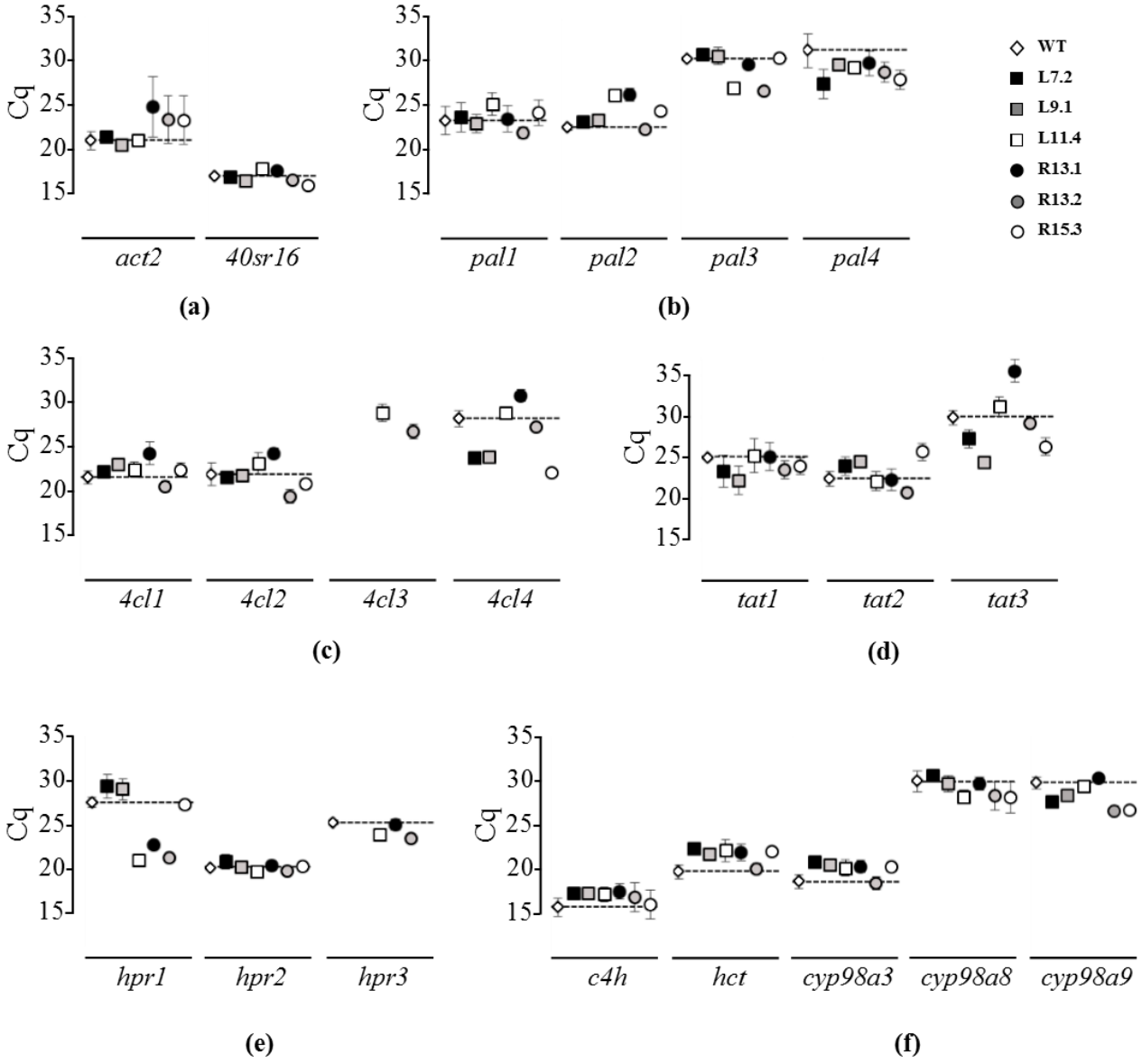


Figure 26. Transcript quantification of RA biosynthetic genes in *Arabidopsis* suspension cultures. Transcript levels of the housekeeping genes, ACT2 and 40SR16 (a), PALs (b), 4CLs (c), TATs (d), HPRs (e), C4H, HCT, CYP98A3, CYP98A8 and CYP98A9 (f) are given as Cq value. Data represent the average transcript level of two cDNA synthesis repeats (except *40sr16*), where each repeat was triplicate. The dotted line indicates the mean value of control wild type line.

The transcript levels of 4CL genes, which have four isoforms, *4cl1-4*, are depicted in Fig. 26b. The expression of *4cl1* was mostly suppressed in R- and L-lines except in R13.2. The transcript levels of *4cl2* and *4cl4* were found to be inconsistent, in which no expression pattern was observed between control and *ras*-transformed lines. Moreover, the *4cl3* isoform was found to be not expressed in most lines, except in L11.4 and R13.2.

Fig. 26d shows the transcript levels of three TAT isoforms, *tat1-3*. The *tat1* isoform was mostly upregulated except in L11.4 and R13.1 lines. The expression of *tat2* and *tat3* seemed to be randomly up or downregulated. For HPR genes (Fig. 26e), no expression changes of the *hpr2* isoform was observed in both lines. In addition, an inconsistent regulation was observed in the transcript levels of *hpr1* and *hpr2*. Therefore, there was no significant pattern in term of HPR gene expression.

The remaining genes, i.e. *c4h, hct, cyp98a3, cyp98a8*, and *cyp98a9*, are present as single genes and their transcript levels are shown in Fig. 26f. The expression of *c4h, hct,* and *cyp98a3* was found to be suppressed in L- and R-lines, although the transcript level of *cyp98a3* was relatively constant in the R13.2 line. The expression of *cyp98a8* was elevated in all lines, except in L7.2, where it was found to be slightly suppressed. The last gene, *cyp98a9* was overall upregulated in both lines, although its expression was slightly decreased in R13.1.

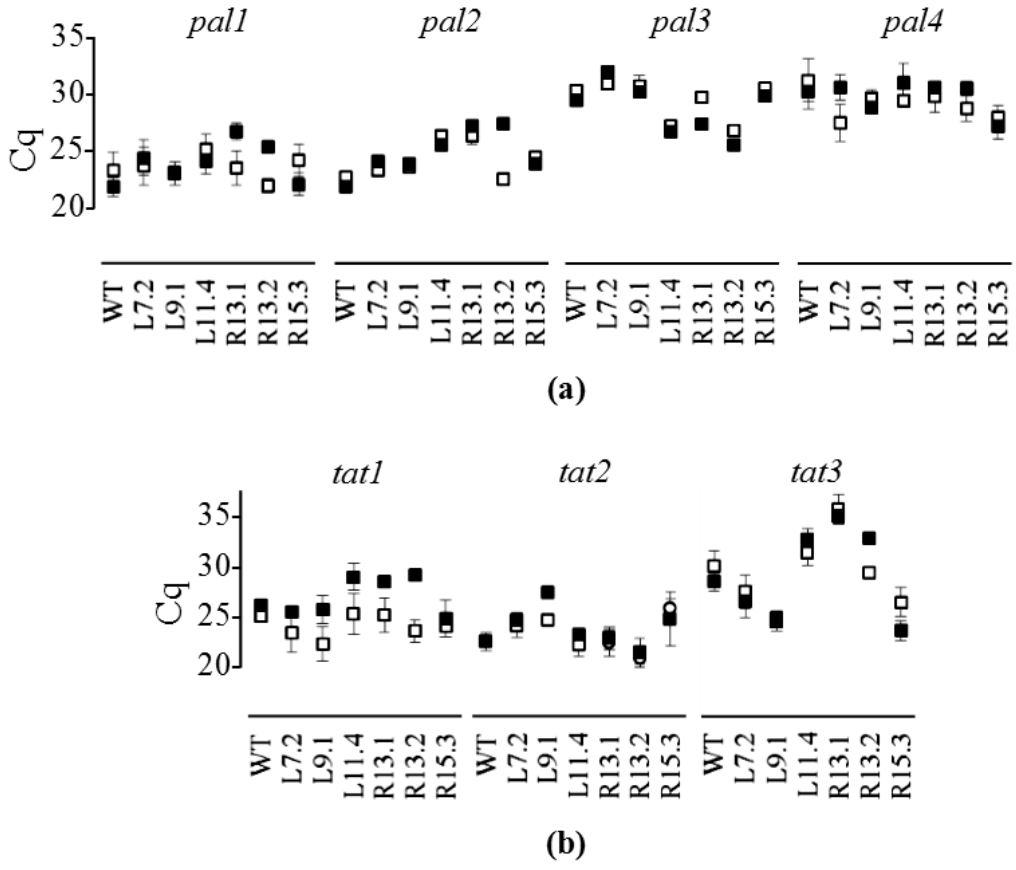


Figure 27. Effect of light on the transcript levels of two genes, *pal* (*pal1-4*) (a) and *tat* (*tat1-3*) (b). Data represent the average expression level of two cDNA synthesis repeats, where each repeat was triplicate. (■) dark, (□) light

The presence of light was reported to have an influence on the expression levels of PAL and TAT genes in *Arabidopsis* (Leyva et al. 1995; Sandorf and Holländer-Czytko 2002). Relative to the expression of genes in cells cultivated in the dark, however, all PAL isoforms were found to be suppressed completely in R15.3 and WT lines. Overall, the presence of light on the other lines has no effect on the transcript levels of all PAL isoforms since the expression was inconsistently regulated (see Fig. 27a). Moreover, the expression of *tat1* in all lines was found to be upregulated by the presence of light, whereas *tat2* was shown to be suppressed only in R15.3 line. The expression of *tat3* did not show a consistent pattern (Fig. 27b).

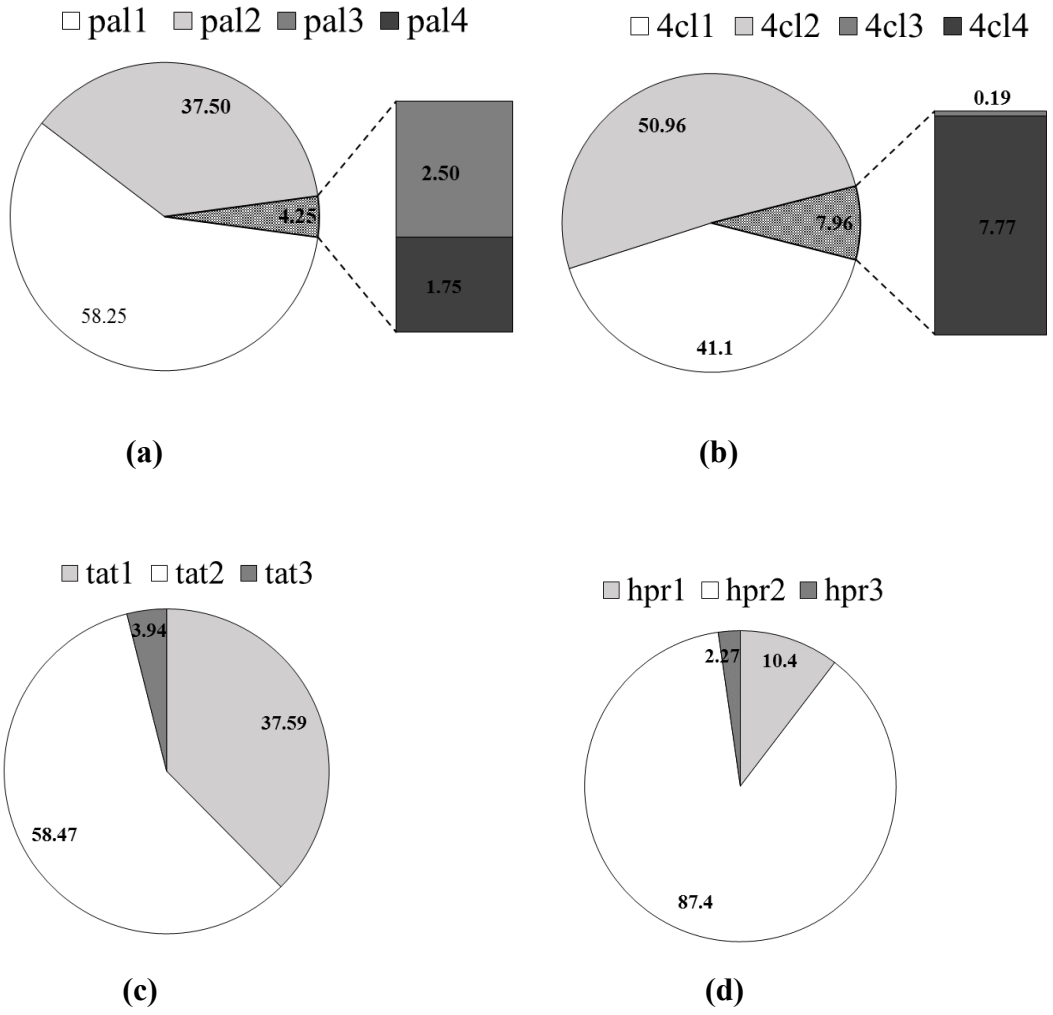


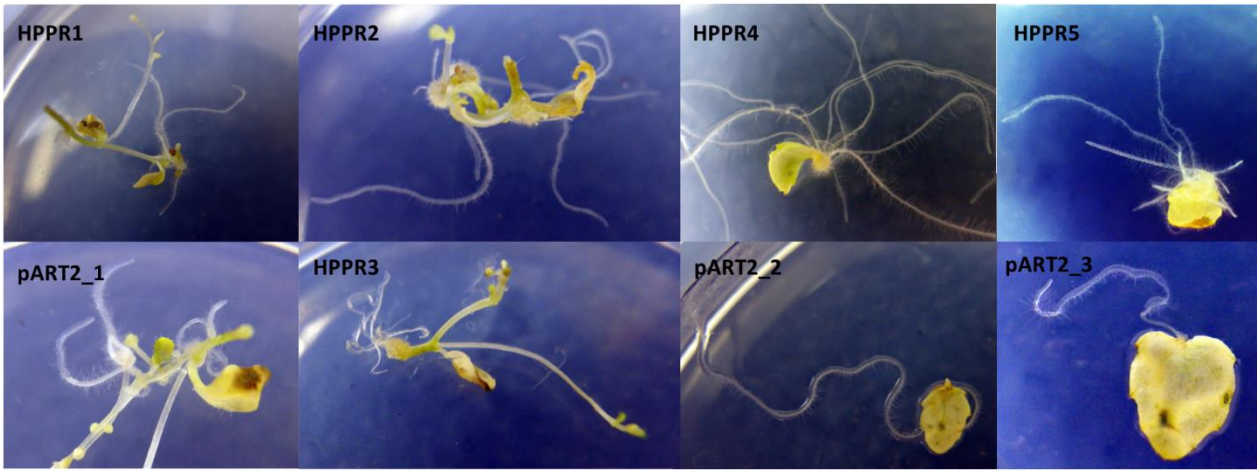
Figure 28. The relative expression between isoforms of a gene. (a) PAL genes, (b) 4CL genes, (c) TAT genes, and (d) HPR genes, where *pal4, 4cl3, tat3*, and *hpr3* were used as the calibrator the respective gene, respectively.

In addition to the expression changes of every gene, the relative expression between isoforms of a gene was analysed in order to identify which isoform is predominantly expressed in cell suspension cultures. For the expression of PAL genes, the highest Cq value that belongs to *pal4* was used as the calibrator. The *pal1* and *pal2* isoforms were found to be expressed at high levels, which is up to 95% of the total PAL expression. Of that, the expression of *pal1* took over 58% of the total expression followed by *pal2* at 37%. The expression of *pal3* was found at only 2.5% and the lowest, *pal4*, was expressed about 1.7% of the total expression (see Fig. 28a). Based on the total expression of 4Cl genes, *4cl1* and *4cl2* were identified to be predominantly expressed and amounted up to 90% of the total expression where the remaining was the expression of *4cl4* (Fig. 28b). The expression of *4cl3*, however, was barely detected (see also Fig. 26b) and therefore was used as the calibrator. Of the two principal isoforms, *4cl2* contributed a half of the total expression.

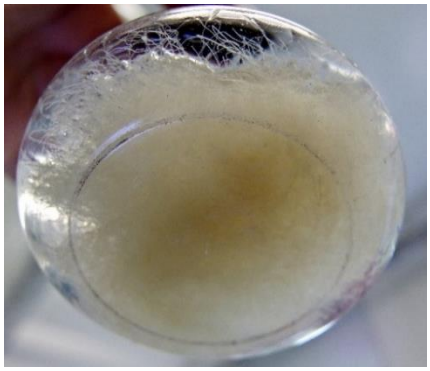
Of the TAT expression, *tat1* and *tat2* were the prominent isoforms that contribute to about 95% of the total expression. About 58% of the expression resulted from *tat2* and 37% resulted from expression of *tat1*. The expression of *tat3* contributed only about 4% of the total expression (Fig. 28c). In the cell suspension cultures, *hpr2* was dominantly expressed and contributed up to 87% of the total HPR expression. The expression of *hpr1* contributed only 10% of the total expression and a minor contribution resulted from expression of *hpr3* (see Fig. 28d).

3.4 The *hppr-ras*-transformed roots

In a few passages, hundreds of leaf explants from seedlings of the *Arabidopsis* R13 line have been transformed with *A. rhizogenes* harbouring the pART27 plasmid with the *hpp*r sequence under the control of the CaMV 35S promoter in order to generate hairy roots. However, only eight leaves were successfully able to generate roots, in which five of them were transformed with pART27-*hppr* and the rest were transformed with pART27 (an empty vector) as controls (see Fig 29a). After the establishment process of the transformed roots which included removing the excess of agrobacteria, selecting with antibiotic and establishing in liquid medium, only one transformed root resulted from HPPR2 explant that was able to multiply in the hormone-free CB2 medium (Fig 29b). The hairy roots were continuously maintained in liquid CBOH medium by sub-culturing every three weeks.



(a)



(b)

Figure 29. Transformed roots. (a) Induced roots from transformed explants, (b) established hairy roots, HPPR2 (*At*HR), in liquid CBOH medium.

Further analysis was performed to investigate the molecular characteristics of the hairy roots. The absence of agrobacteria was verified by PCR using a partial sequence of the *VirC* gene, one of the virulence genes in the Ri-plasmid but not a part of the T-DNA. PCR amplification of this gene resulted in many bands (Fig. 30a) suggesting that the primer is not specific. However, by comparison with the amplification product of 700 bp from the positive control suggested that the hairy roots were free from agrobacteria. This was more obvious when *virD2* was used as the marker gene instead, in which the band of VirD2 gene (224 bp) was found in the positive control but not in the hairy roots (Fig. 30b). The presence of genes that are responsible for hairy root induction, the *rol* genes (*rolA, rolB, rolC*), has been positively confirmed (see Fig 30a) suggesting that the transformed roots are fully hairy roots. Furthermore, PCR amplification of *hppr* showed the presence of the exact product size of 950 bp in both hairy root and the positive control (Fig 30b). It revealed that the T-DNA has been stably integrated in the genome of the hairy roots.

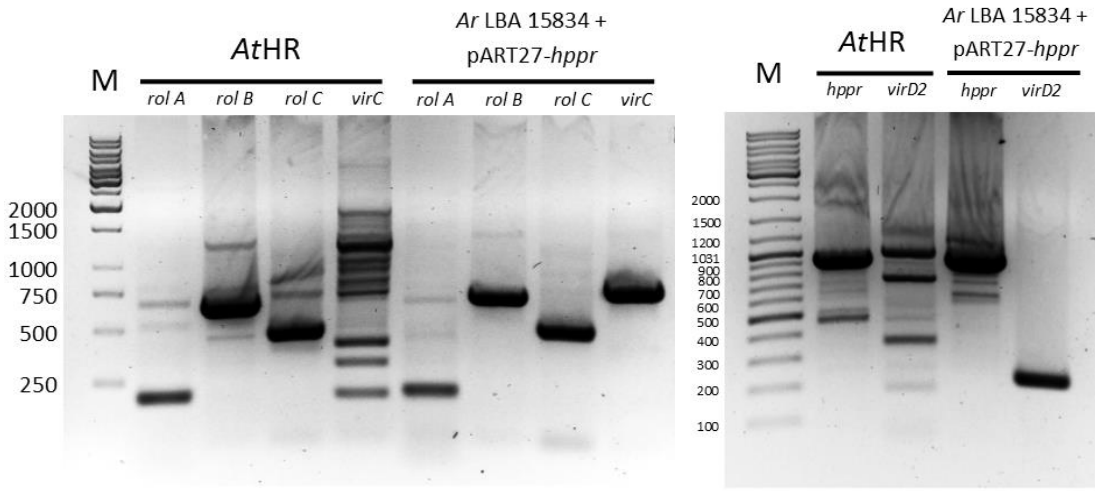


Figure 30. PCR confirmation of transformed hairy roots. (a) *rolA ~ 200 bp, rolB ~ 650 bp, rolC ~ 450 bp, virC ~ 700 bp.* (b) *hppr ~ 950 bp*, *virD2 ~ 224 bp.* PCR templates were gDNA of hairy roots (*At*HR) and *A. rhizogenes* LBA15834 carrying pART27-*hppr* as the positive control.

Enzyme activities were evaluated by using crude protein extracts from 3-week old hairy roots. The assays were only performed for two enzymes, RAS and HPPR. No product of RAS activity (Caf-pHPL) has been detected in hairy root as found in cell suspension cultures of the R13.1 line (Fig. 31a). Also the HPPR assay did not show any activity since no significant change in the reaction product (pHPL) was observed between the test and the control sample (see Fig. 31b). Analysis of the gene expression confirmed that the expression of *ras* was undetectable in the cDNA although this gene was shown to be integrated in the genome of the hairy roots (Fig. 32). This answered the question why no RAS activity has been detected. In contrast, the *hppr* gene was found to be successfully expressed in the hairy roots, although a measurable activity was not detectable.

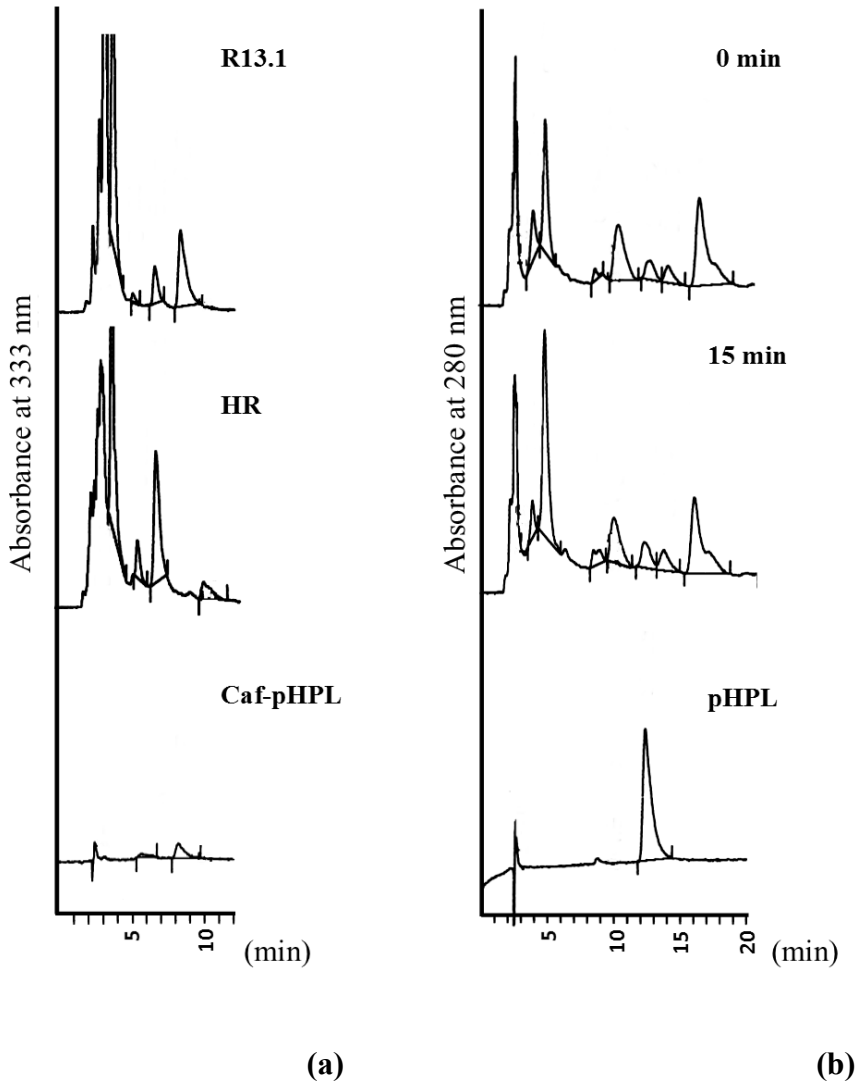


Figure 31. HPLC chromatograms of RAS (a) and HPPR (b) assays from protein extracts of the *hppr*-transformed hairy roots (HR)

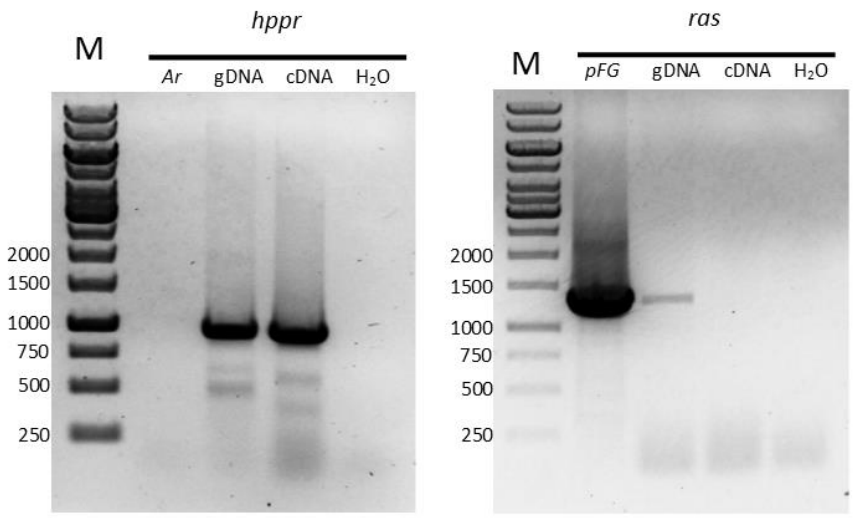


Figure 32. PCR amplification of *hppr* (*~950 bp*) and *ras* (*~ 1290 bp*) from gDNA and cDNA of the hairy roots

3.5 Metabolite accumulation in transformed *Arabidopsis*

The profile of metabolites in the transgenic *Arabidopsis* was preliminarily analysed by HPLC. They included RA and RA-like esters, i.e. caffeoyl-4'-hydroxyphenyllactic acid (Caf-pHPL), 4-coumaroyl-4'-hydroxyphenyllactic acid (pC-pHPL), and 4-coumaroyl-3',4'-dihydroxyphenyllactic acid (pC-DHPL) as authentic standards. Unfortunately, most of the metabolites were under the detection level, even with a high concentration of extract which represented 100 mg of lyophilised cells per ml (see Fig. 33). Based on the differing retention times of the authentic standards, the HPLC fraction potentially containing the esters was collected and concentrated for LC-MS-MS analysis.

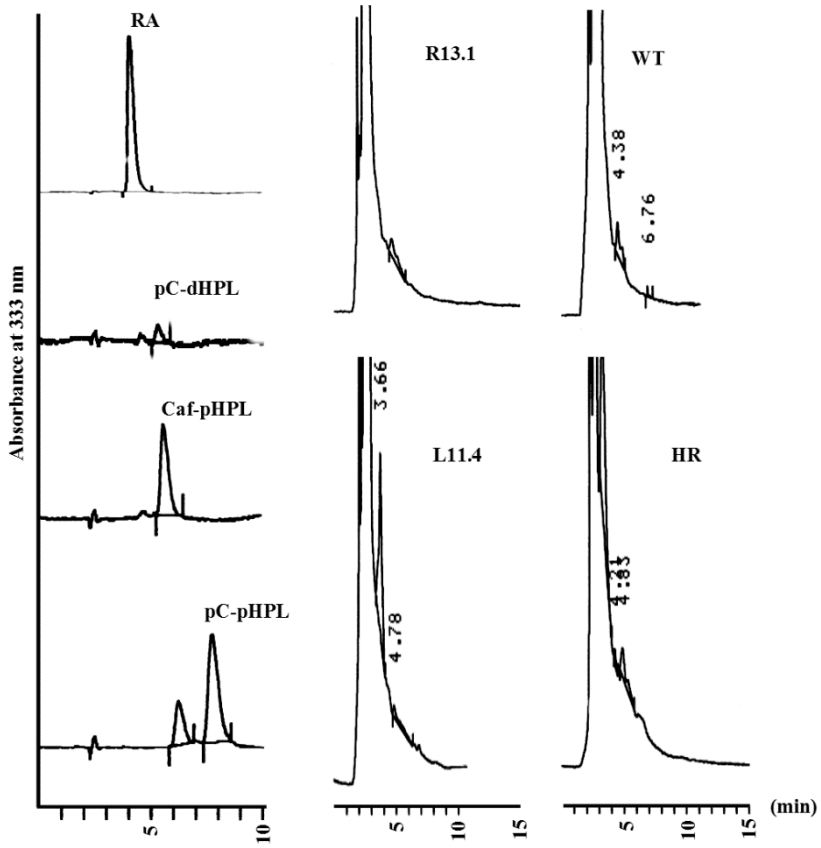


Figure 33. HPLC chromatograms of the authentic standards and extracts of the transformed (R, L, HR line) and wild type (WT) *Arabidopsis*

The ester products of RAS can be differentiated through their retention times, their masses and their fragmentation patterns. Tandem mass spectrometry fragments a compound into distinct ions and therefore is a powerful tool, particularly when analysing these esters from fractions that usually have compounds with similar molecular mass. In particular, it helped differing two esters, Caf-pHPL and pC-DHPL, with similar retention times and molecular mass as well (Petersen 1993, see also Fig. 34). Results depicted in Fig. 34 were the confirmation of the structure of these esters by tandem mass spectrometry. They were ionised in the negative, not in the positive ion mode as described in the previous report. However, the fragmentation pattern that resulted in both ion modes looks alike, where abundant fragment ions are structures on both sides of the oxygen atom of the ester bond. As can be found in the daughter ion spectra, abundant ions correspond to 4-coumaroyl (*m/z* 163), caffeoyl (*m/z* 161), pHPL (*m/z* 181), and DHPL (*m/z* 179 and 197). The occurrence of these ions allows a fast identification of the ester products of RAS in the ester fractions of transgenic *Arabidopsis*.

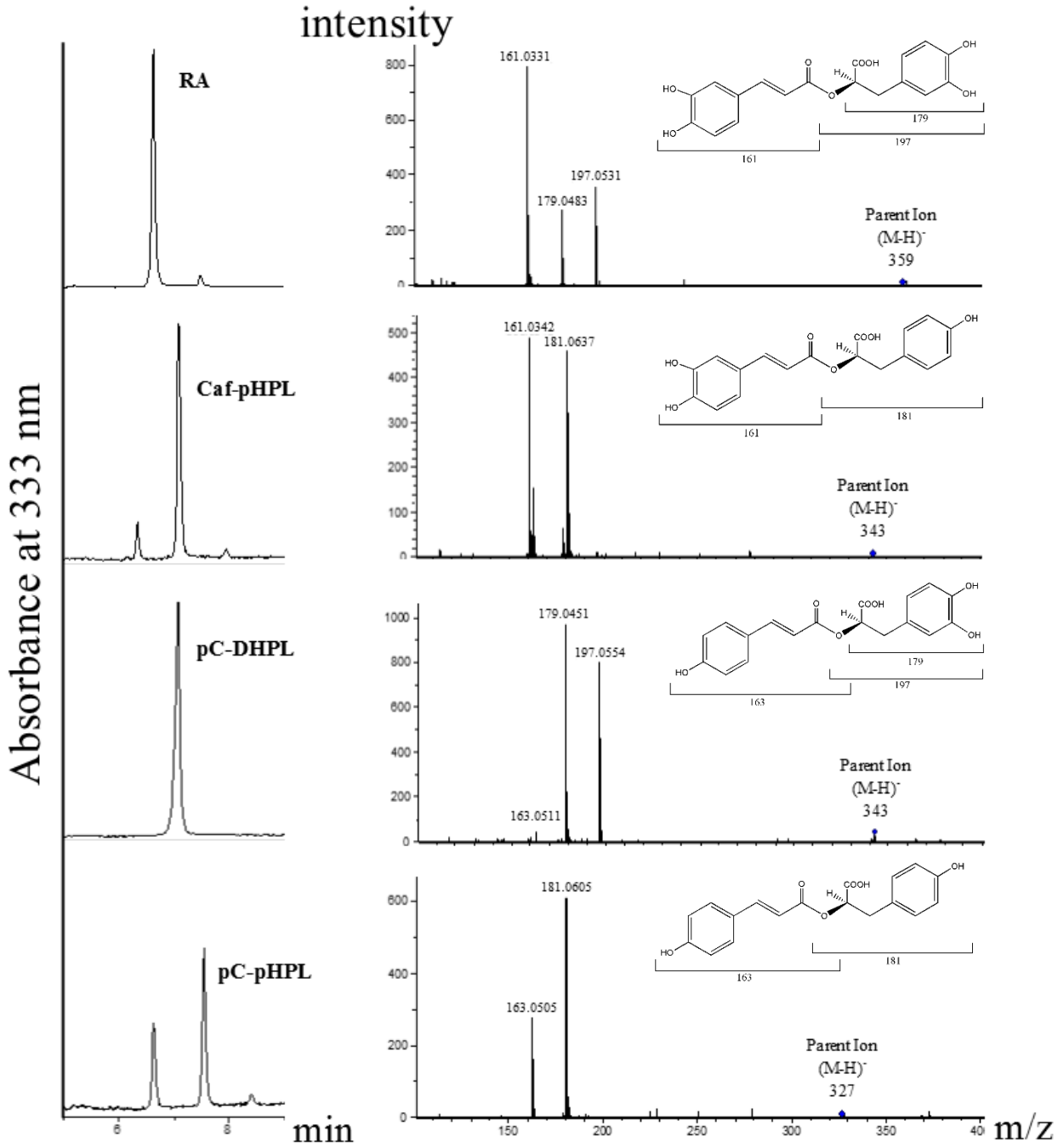


Figure 34. (a) HPLC chromatogram of the authentic standards and (b) daughter ion mass spectra of $[M-H]^-$ of RA (*m/z* 359), Caf-pHPL (*m/z* 343), pC-DHPL (*m/z* 343), and pC-pHPL (*m/z* 327). Fragments allowing identification are shown at the inserted structural formula.

For the mass spectrometry analysis, the fractions were collected from 6 individuals of R-lines (R13.1, R13.2, R13.3, R13.5, R13.8, R15.3), 3 individuals of L-lines (L7.2, L9.1, L11.4), one individual of hairy root and wild type as well. Each fraction represented 1 g lyophilised cells/ml, of which 20 µl was used for the analysis. Surprisingly, RA and Caf-pHPL were most probably found to occur in the RAS-transgenic *Arabidopsis*, even though they were found only in R13.1 and R13.2, respectively (Fig. 35). However, none of those RAS products have been shown to be present in other lines, though they do exhibit RAS activity (see Appendix). These results suggest that the expressed foreign RAS gene in *Arabidopsis* enabled the plant to synthesise the ester products as hypothesised. However, the crucial point was the failure to identify any detectable level of pHPL, even in the hairy roots where the HPPR gene has been overexpressed (see Appendix). In RA-producing plants, pHPL is also never identified as a free intermediate. It might be used directly for the subsequent reaction and therefore does not accumulate in cells to a measurable level. A metabolomic study with high-throughput analysis might be able to identify the occurrence of this intermediate in cells.

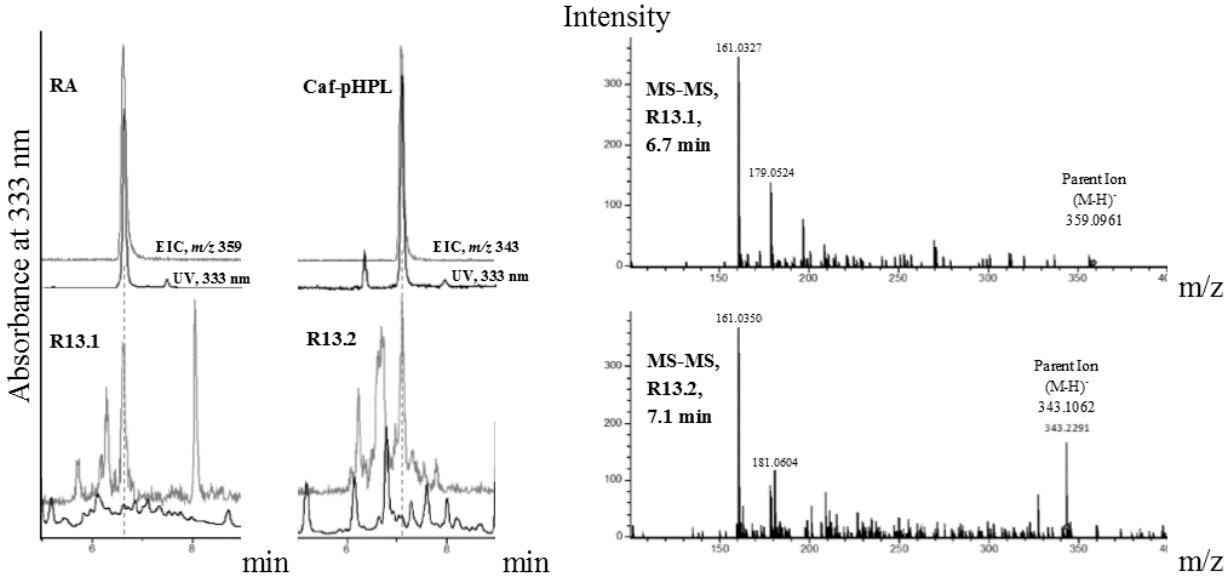


Figure 35. Metabolite analysis of the ester fraction of lines R13.1 and R13.2. The expected products were identified by aligning the HPLC (black) and EIC (grey) chromatograms between the fraction and authentic standard. RA and Caf-pHPL were detected in R13.1 and R13.2, respectively. They have been positively confirmed by comparing the daughter ion mass spectra (MS-MS) with the standards (see also Fig. 34b). EIC, extracted ion chromatography, where RA (*m/z* 359) and Caf-pHPL (*m/z* 343).

4. Discussion

4.1 Molecular characterisation and growth profiles of *in vitro* plant material

Plant *in vitro* cultures have been widely used to study many aspects of plant biology and biotechnology, including the use of plant cell cultures for the production of high-value specialised metabolites (Verpoorte et al. 1994, 2000, 2002; Rao and Ravishankar 2002; Matkowski 2008). One of them is rosmarinic acid (RA), a promising plant specialised compound that has shown many remarkable pharmacological activities (Amoah et al. 2016). Surprisingly, this compound can be produced to around 20% of the dry biomass by using plant cell factories (Petersen and Alfermann 1988; Döring and Petersen 2014).

From the evolutionary view, RA has attracted our interest since this phenolic ester is widespread in the plant kingdom. RA has been found to occur not only in higher plants but also in lower plants. However, there is an inconsistency since not all members in each level of plant taxa contain RA (Petersen 2013). A small approach was then taken in this work to reveal a small part of such phenomena by using the model plant *Arabidopsis thaliana*. The presence of general phenylpropanoid pathway genes as well as more specialised RA biosynthesis genes has been described in this plant, but the absence of one or more RA-specific genes causes the inability of the plant to synthesise RA. Hence, by introducing the specific genes, *ras* and *hppr,* into *Arabidopsis* will hopefully guide us into better understanding the evolution and the regulation of the RA biosynthetic pathway in plants. In this work, cell suspension cultures were chosen since the undifferentiated cells from various producing plants have been shown to be a suitable way for RA production (Bulgakov et al. 2012)

Seeds from T2 progeny of the transformed *Arabidopsis thaliana* were used to establish the experimental plant material. The plants consisted of two different lines, R- and L-lines, that were transformed with the vector (pFGC5941) containing the complete open reading frame for CbRAS and with the empty vector. Axenic cell suspension cultures were derived from callus established from aseptic seedling on CB2 medium. The medium used here was based on previous publications (Valvekens et al. 1988; Ozawa et al. 1998; Sugimoto and Meyerowitz 2013) and has been routinely used for the culture of various RA-producing plants in our laboratory, e.g. for *Coleus blumei* (Petersen 1991). The medium has been confirmed to be efficient for the cultivation of transformed *Arabidopsis*. In addition, the growth of the two transgenic lines in this medium was not distinguishable even not from the wild-type lines.

Molecular characterisation confirmed the successful integration of the T-DNA into the plant genome, in which the CbRAS gene is present in individuals of the R-lines and the CaMV 35S promoter sequence in the L-lines, respectively. Moreover the absence of the virD2 gene has asserted that the transformed plants are free from agrobacteria. This indicates that the PCR amplification product of the T-DNA resulted not from the complete plasmid of the agrobacteria but from the integrated T-DNA in the plant genome. However, the presence of the KanR gene in all individuals of both lines has been confusing us since this gene marker is located outside the T-DNA (see Fig. 14) and normally should not be transferred to the plant genome. Such phenomenon, however, has been described in several reports as the integration of the non-T-DNA binary vector backbone sequences. It is mostly caused by the properties of the border repeat sequences that determine the nicking process of the T-strands. In several reports, it was found to occur with percentages of more than 50% of transgenic *Arabidopsis* and tobacco plants (Kononov et al. 1997; De Buck et al. 2000).

In an early experiment, the growth profiles of cell suspension cultures, represented by R13.2 and L11.4, have been determined for their fresh weight and protein contents during the cultivation cycle. The logarithmic growth of the cell suspension cultures was during the first week of the observation period (see Fig. 16). This was similarly reported for cell suspension cultures of some RA-producing plants such as *C. blumei* (Gertlowski and Petersen 1993), *Melissa officinalis* (Weitzel and Petersen 2011), and *Glechoma hederacea* (Döring and Petersen 2014) which were cultured in the same medium. However, the cells of both lines kept growing until they reached the highest biomass on day 9. At this time point, the *Arabidopsis* cell cultures still contained green cells which was also observed by Hampp et al. (2012). In contrast, the total protein content was found to be at its highest level only during the logarithmic growth. More interestingly, cell suspension cultures of the R-line reached their highest protein content two days prior to the L-line. This may be due to the strong expression of the RAS gene that might change the flux of the protein synthesis in transformed plants at the early stages of cell division, where the nutrient supplies are readily available.

After characterising more individuals of each line, we observed that the overexpression of the RAS gene in *Arabidopsis* may reduce the cell culture growth but does not severely affect the cells since the growth reduction is only very low and no cell impairment has been observed. Overall, strong expression of the RAS gene did not increase the protein contents as compared to the control lines. Additional experiments showed that the growth profiles and protein contents are not influenced by permanent light compared to darkness.

4.2 Activities of RA biosynthesis enzymes

In the Section 3.2, the activities of RA-biosynthetic enzymes (PAL, C4H, 4CL, TAT, RAS) in crude protein extracts of cell suspension cultures were reported. The *ras*-transformed cell line R13.2 successfully synthesises functional RAS which is capable of catalysing the formation of the ester caffeoyl-pHPL from caffeoyl-CoA and pHPL in *in vitro* enzyme assays (see Fig. 19a). Similar successes have been documented in a few attempts of expressing transgenes in *Arabidopsis* that are related to specialised metabolism, for example, the expression of a bacterial pinoresinol reductase (Tamura et al. 2014) and a cytochrome P450, CYP79A1, involved in dhurrin biosynthesis (Bak et al. 1999).

Overall, the activities of RA biosynthetic enzymes in *Arabidopsis* increased in the first week of the observation period while the cells were actively multiplying. This suggests that the enzymes and their (specialised) products may have role in active cell growth. Similar results have been observed in the cell suspension cultures of RA-producing plants, e.g. *C. blumei* (Petersen 1991), *M. officinalis* (Weitzel and Petersen 2010, 2011), and *G. hederacea* (Döring and Petersen 2014). Unlike the other enzymes, the activities of RAS and TAT were quite stable after passing the day where the highest activity was achieved. For RAS, this may be based on the control of the strong constitutive promoter that keeps the gene continuously expressed as long as the cells are still growing. For the putative TAT, the activity seems to obviously highlight its role against the oxidative stress during the senescence process. Such a role has been observed in aging leaves of *Arabidopsis,* in which the expression level of one TAT isoform, *tat1* (*At5g53970*), as well as the level of tocopherols were elevated in the phase of senescence (Holländer-Czytko et al. 2005).

By expressing the CbRAS gene, the enzymes in the upstream pathway of RA might be triggered to produce more intermediate substrates (coumaroyl-CoA and pHPP), since these intermediates are also used by other metabolic pathways. Therefore, the presence of RAS might increase the demands of cells for these intermediates. However, by comparing more individuals of both lines, we found no denoted change with respect to the activity of these enzymes. The changes of the activity were observed to be cell-line specific, suggesting that the RAS overexpression has no influence on the activity of these enzymes. The activities of two essential enzymes in the upstream pathway, PAL and TAT, that provide the caffeic acid and pHPL parts, respectively, are known to be stimulated by light. One of the PAL genes in *Arabidopsis, pal1,* was found to be upregulated in the presence of light (Leyva et al. 1995). The influence of light on the activity of TAT cannot be ruled out since this enzyme is required for the formation of compounds essential for

photosynthesis (Riewe et al. 2012). The expression of the TAT gene was also found to be upregulated under UV-B radiation in *Salvia miltiorrhiza* (Huang et al. 2008). In the cell suspension cultures of *Arabidopsis*, the activity of both enzymes was randomly affected under dark and light conditions (see Fig. 22). Interestingly, the presence of light only increased the activity of the enzymes in *ras*-transgenic individuals and wild-type, which we still cannot explain at the moment.

Possibility of post-translational modification events

Attempts to overexpress a foreign gene in plants sometimes face many challenges as described in several reports. The constitutive expression of certain genes might be harmful to their host plants, affecting plant growth and development. In other cases, however, the host plants develop self-defences in various ways to overcome any change resulted from the integration of transgenes. In our case, we found that strong overexpression of the RAS gene in *Arabidopsis* did not make all *ras*-transformed lines having RAS activity. A few of them had exhibited the activity, but later on, the activity got lost. We have anticipated this case as a result of post-transcriptional gene silencing (PTGS), a mechanism developed by the host plants that results in the specific degradation of a population of either homologous or non-homologous RNAs. PTGS greatly reduces mRNA accumulation in plants but does not affect transcription. This mechanism is triggered more efficiently when strong promoters are used (Vaucheret et al. 2001), as observed in *ras*-overexpressing hairy roots of *C. blumei* that lead to co-suppression (Hücherig and Petersen 2012). This, however, has not been shown in our case since mRNAs of the RAS gene were found to be present in all individuals of *ras*-transformed lines as shown by the successful PCR-amplification of the derived cDNA (Fig. 22b). In addition, the presence of natural inhibitors, which might inhibit the RAS activity, however, was not the cause as well. This was demonstrated by mixing protein extracts from R-lines with the recombinant CbRAS purified from the CbRAS-engineered *E. coli.* However, no decrease of the activity has been observed between in the enzyme mixtures compared to the recombinant CbRAS alone (see Fig. 38 in the Appendix).

Since factors affecting the loss of RAS activity in *Arabidopsis* are not on the replication and transcription levels, we therefore looked into the post-translational regulation level. In general, during the post-translational process, the majority of newly synthesised plant peptides undergo additional modification that will finish their tertiary and quaternary structures to become functional proteins. The post-translational modification (PTM) has been shown to affect almost every aspect

of protein activity, including function, localisation, stability, and dynamic interactions with other molecules (Webster and Thomas 2012). Under certain circumstances, post-translational regulation was found to control the decrease of enzyme activity in a cell, in which the enzyme can be modified either covalently or by a degradation program (Hemmerlin 2013). It has long been known that plant proteases are involved in various processes, including senescence and defence responses (van der Hoorn et al. 2004). *In planta*, proteolysis has become an obstacle not only in the production of recombinant proteins (Benchabane et al. 2008; Pillay et al. 2014) but also in the natural production of putative proteins (Hemmerlin 2013; Zhang and Liu 2015; Kim et al. 2016). For instance, the putative flowering locus T protein (FT) was found to be post-translationally regulated in *Arabidopsis* plants. The FT protein levels gradually decreased to an almost undetectable level after the floral transition, although FT mRNA is still transcribed (Kim et al. 2016).

In *Arabidopsis*, more than 950 proteases have been documented which comprise 847 putative peptidases and 129 non-peptidase homologues. They can be distinguished as serine, cysteine, aspartic, and metalloproteases, where serine proteases are the largest class in this plant species (http://merops.sanger.ac.uk, Pillay et al. 2014). We therefore investigated cellular proteolysis events by using protease inhibitors (PMSF and inhibitor cocktail) with inhibitory activity against serine protease and broad range of proteases, respectively. However, both inhibitors only prevented around 20% loss of RAS activity during the protein extraction but did not recover the activity of cell lines whose activities had already been lost.

As a regulatory mechanism, proteolysis can be controlled by induction through specific metabolites (Hemmerlin 2013). For example, the decreased activity of HMG-CoA reductase in plants is thought to be post-translationally regulated by molecules such as phytochrome, sterols, abscisic acid, ubiquinone or 4-hydroxybenzoic acid, and by protease-mediated degradation (Russell and Davidson 1982; Korth et al. 2000). In yeast and mammalian cells, this enzyme is regulated by ubiquitination and proteasome-mediated degradation (Hemmerlin 2013). The ubiquitination mediates the protein degradation by the proteasome in eukaryotic systems. The degradation is initiated by modification of the protein target by covalent attachment of ubiquitin, a 76 amino acid protein, to lysine residues. Then, the protein is recognised, relocated into early endosomes and transported to lysosomes, and degraded by lysosomal proteases (Walsh et al. 2005; Hemmerlin 2013; Zhang and Liu 2015). Several phenylpropanoid and monolignol biosynthetic enzymes, including four PAL isoforms in *Arabidopsis,* are reported to be potentially degraded by proteasome-mediated proteolysis through ubiquitination (Zhang and Liu 2015). We therefore

performed a test of the RAS protein sequence with an online free software Ubpred (www.ubpred.org) to predict a possible ubiquitination. However, the result suggests that the ubiquitination may not be the cause in this case.

Our attempts to identify the presence of the RAS protein in the cell extracts were not successful due to the interference of putative plant proteins with similar molecular mass. We have thougth of overcoming this problem by gaining RAS-specific antibodies. Eventually, we faced a difficult challenge with the purification of heterologously expressed CbRAS from *E. coli* to high purity and yield. Therefore, we were not able to obtain a high purity recombinant CbRAS protein in sufficient amounts to raise polyclonal antibodies for Western blotting.

Despite our limitation to investigate the proteolysis events *in vivo*, we could speculate that the proteolytic degradation-related PTM is not the cause in this regard, though we cannot explain the gradual loss of activity in many cell lines despite the proven presence of transcripts. We also don't know the reason why the RAS activity can still be observed in other lines. Other possible mechanisms of PTMs such as protein processing and activation, protein localisation, protein-protein interactions, enzyme activity and kinetics (e.g. allosteric regulation) (Friso and van Wijk 2015) should be considered in future.

4.3 Expression of RA biosynthetic genes

4.3.1 Gene expression changes related to *ras* overexpression

The expression of a gene under control of a strong constitutive promoter is found to be frequently higher than under its own promoter. Some reported attempts of overexpression of RA biosynthetic genes in producing plants revealed that. The expression level of C4H, TAT and HPPR genes in hairy roots of *S. miltiorrhiza* under control of the CaMV 35S promoter are higher than in untransformed hairy roots (Xiao et al. 2011). Similar results are also observed in hairy roots of *C. blumei* transformed with the HPPR gene (Hücherig and Petersen 2012). In *tat*-transformed *Perilla frutescens*, the expression of TAT genes are 2-3 times higher than in the untransformed plants (Lu et al. 2013). One attempt of overexpression of the RAS gene in plants is the report of our group. The overexpression of the RAS gene did not lead to an increased expression level of the RAS gene in hairy roots of *C. blumei* as expected. On the contrary, it leads to the reduction of the RAS transcript level as well as the RA content in hairy roots, due to co-suppression effects or homology-dependent gene silencing (Hücherig and Petersen 2012).

In this work the RAS gene has been successfully overexpressed in *Arabidopsis*, a plant not known for RA production, for this time. The transgene was transcribed in all *ras*-transformed lines but varied among them (see Fig. 25). A similar result has been observed in *Arabidopsis* transformed with the bacterial pinoresinol reductase (Tamura et al. 2014), suggesting that the efficiency of overexpression might be affected by many factors such as the number and site(s) of T-DNA integration events or cellular responses of host plants or cell lines. Even though the expression of the RAS gene in transformed *Arabidopsis* cell cultures is much higher than in a (rather old) suspension culture of *C. blumei*, the activity is much lower and even undetectable in many *ras*-transformed lines. This result highlights the unknown cellular mechanisms which have already been discussed above.

Overexpression of the RAS gene in *Arabidopsis* changes the expression of genes related to the RA biosynthetic pathway. In general, the gene transcript levels fluctuated and showed line-specific differences. A comparison between the *ras*-transformed and the control lines resulted in no exact pattern for the expression of PALs, 4CLs, TATs, and HPRs. These are found to be inconsistently down- and/or upregulated in each line. C4H is the only gene in the RA pathway whose expression is downregulated in all lines. This could be the reason why the activity of C4H is below 5 µkat/kg which is much lower than for other enzymes (see Fig. 19b and 21d). The expression levels of other RA-related genes, HCT and CYP98A3, are downregulated, CYP98A8 is relatively constant in control lines but upregulated in *ras*-transformed lines, while CYP98A9 is upregulated in all lines. In the presence of light, the activity/expression of PAL and TAT is reported to be positively regulated (Leyva et al. 1995; Riewe et al. 2012). Our results, however, show that the expression levels are inconsistent and line specific.

There are many reports that the expression levels of genes in a biosynthetic pathway fluctuated in engineered plants. In hairy roots of *S. miltiorrhiza,* the transgenic manipulation variably affected the transcript level of other RA biosynthetic genes (Xiao et al. 2011). It has been known that PALs can be transcriptionally, post-translationally or metabolite feedback regulated (Zhang and Liu 2015). In *c4h*-knockout tobacco plants, PAL expression/activity was negatively regulated by the accumulation of its product *t*-cinnamic acid (Blount et al. 2000). In contrast, down-regulation of PAL by gene silencing did not reduce the activity of C4H in tobacco plants (Blount et al. 2000) and in PAL-suppressed *S. miltiorrhiza,* the transcript levels of C4H and 4CL2 were even found to be increased (Song and Wang 2011). This suggests that feedback regulation takes place at the entry point of phenylpropanoid pathway and operates in one direction only. In our

experiments, the decrease of C4H gene expression in all lines could be linked to the down-regulation of the PAL2 gene, one of the principal PALs. Although the expression of the PAL2 gene was not completely abolished, a significant increased expression of PAL4 was observed in all lines. A similar result has been reported by Rohde et al. (2004) where PAL4 partially compensated the loss of PAL1 and PAL2 in the double knockout *Arabidopsis* mutants.

The feedback regulation is not only influenced by the modulation from a direct pathway but also from branch pathways. For example, overexpression of TAT, HPPR, or TAT and HPPR genes in hairy roots of *S. miltiorrhiza* reduced the transcript levels of PAL (Xiao et al. 2011). Overexpression of a bacterial quercetin oxidoreductase in *Arabidopsis* reduced the transcript level of the C4H gene (Reuben et al. 2013). In *Arabidopsis*, the expression of HCT was reduced in the double knockout mutants of PAL1 and PAL2 (Rohde et al. 2004). In hairy roots of *C. blumei*, the expression of HCT remained stable irrespective of overexpression or suppression of the HPPR or RAS gene (Hücherig and Petersen 2012). In our experiments, non-RAS HCTs were down-regulated in both control and *ras*-transformed *Arabidopsis*. It seems that the regulation of gene expression is a more complex phenomenon in the overall phenylpropanoid network.

4.3.2 Expression profile of *Arabidopsis* genes with multiple isoforms in undifferentiated cells

The biosynthetic pathway studied here contains three genes with multiple isoforms, namely PAL, 4CL, and TAT, and HPR with three different enzymes that function with different co-substrates and in different cellular localisations. Regulation of their expression has been well studied during the growth and development phase of *Arabidopsis* plants including in organs and in specific tissues. However, no study has been reported with regard to their regulation in undifferentiated cells. For PAL genes, our result strengthens *pal1* and *pal2* as the principal PAL genes in *Arabidopsis*, where these two isoforms make up to 95% of the total expression. Interestingly, the expression of *pal3,* which is reported to have the lowest level in lignifying tissues (Mizutani et al. 1997; Raes et al. 2003; Rohde et al. 2004), shows higher level in undifferentiated cells than the expression of *pal4.* For 4CL genes, about 90% of 4CL expression resulted from the *4cl1* and *4cl2* isoforms, where *4cl2* contributed half of the total expression. This is in contrast to the expression of 4CL genes in lignifying tissues, where *4cl1* alone contributes to 90% of the total activity (Li et al. 2015). In contrast to the expression of *4cl3* and *4cl4* in tissues, *4cl4* contributes to the remaining expression, while *4cl3* is barely expressed in our undifferentiated *Arabidopsis* cells.

As reported, *tat1 (At5g53970)* could be the prominent TAT isoform since the *tat1* gene knockout reduced about 60% of the total TAT activity in *Arabidopsis* leaves. Even its closest homolog, *tat2 (At5g3610)*, was not able to compensate the loss of *tat1* function (Riewe et al. 2012). Here, we found that in undifferentiated cells, *tat2* was found to contribute more than a half of the total expression instead of *tat1* as the prominent isoform. This suggests that there might be different regulation patterns of TAT genes in different stages of plant development.

HPR genes encode plant photorespiratory enzymes that convert hydroxypyruvate (HP) to glycerate. Although they function similarly, the three HPRs have different properties. HPR1 uses NADH as the co-substrate and functions in peroxisomes, whereas HPR2 and HPR3 have NADPH as co-substrate but they function in different compartments of the cell, cytosol and chloroplast, respectively. In *Arabidopsis* plants, two isoforms (*hpr1* and *hpr2*) function as the prominent HPRs and work redundantly in certain conditions (Timm et al. 2008, 2011). However, in our undifferentiated *Arabidopsis* cells, the probably only have very low photosynthetic activity, the highest portion of HPR gene expression resulted from the expression of *hpr2,* not from *hpr1* expression. This is very interesting since *hpr2* is able to accept pHPP *in vitro* (Hücherig 2010) that we suppose could provide the pHPL substrate for the formation of RA-like esters.

4.4 *Hppr-ras*-transformed roots

Hairy roots with or without additional genes of interest are frequently used for studies of secondary metabolism and even for the commercial production of high valuable pharmaceuticals (Giri and Narasu 2000; Chandra and Chandra 2011). Hairy roots are produced from plant material after transformation with *Rhizobium rhizogenes* (syn. *Agrobacterium rhizogenes*), in which the transferred genetic information (T-DNA) from the root inducing (Ri) plasmid triggers root formation in infected plants (Spena et al. 1987). The responsible root inducing genes located in the T-DNA, *root loci (rol) A, B* and *C,* render plant cells more sensitive to endogenous auxin which eventually induces the formation of hairy roots (Spena et al. 1987; Schmülling et al. 1988). In addition, the expression of the *rol* genes is sometimes found to increase the amounts of secondary metabolites in hairy roots, for example, solasodine in *Solanum aviculare* (Kittipongpatana et al. 1998), tropane alkaloids in *Atropa belladonna* (Bonhomme et al. 2000), and RA and other antioxidants in *Ocimum basilicum* hairy root cultures (Srivastava et al. 2016). The *rol* genes are also able to induce the production of compounds that are usually synthesised only in specific organs

of intact plants, for example artemisinin which is produced only in glandular trichomes of *Artemisia annua* leaves and shoots (Mannan et al. 2008). The presence of these *rol*-genes, which are also found in our *hppr-ras*-transformed *Arabidopsis* roots, verifies successfully established hairy roots. In addition, specific traits of hairy roots such as high growth rate in hormone-free media were also observed.

Cultures of hairy roots have been shown to be a promising means for the production of RA *in vitro* and can be used to study the regulation of RA metabolism (Tada et al. 1996; Chen et al. 1999; Bais et al. 2002; Li et al. 2005; Lee et al. 2008; Bauer et al. 2009). Reports show that both overexpression and silencing of RA biosynthetic genes increase and decrease the RA content in hairy roots of producing plants, respectively (Xiao et al. 2011; Hücherig and Petersen 2012; Lu et al. 2013). Here, the double expression of *hppr* and *ras* genes in the same hairy root line was achieved after two individual *Agrobacterium* transformation steps. Our approach was not a classical co-transformation with two T-DNAs from different plasmids or one T-DNA with two transgenes. Moreover, the approach used here was not as often described where a single *Arabidopsis* plant is co-transformed at the same time during co-cultivation (De Buck et al. 2000; De Buck et al. 2009; Ghedira et al. 2013). Here, the already transformed *Arabidopsis* originating from a floral-dip transformation with *A. tumefaciens* carrying the CbRAS gene in its T-DNA was transformed again with *A. rhizogenes* carrying the CbHPPR gene. The efficiency of the transformation was very low since only one from hundreds of leaf explants successfully generated hairy roots. This result shows that the already transformed *Arabidopsis* may not be competent for the second transformation.

The hairy roots display genetic stability since the integration of RAS and HPPR genes is still observed after several times of subcultivation. However, RAS activity could not be detected and this is correlated to an undetectable level of *ras* transcript that might be caused by gene silencing. It should be stressed here that this is in contrast to the R13 line, which was used for HPPR-transformation, where most individuals from the R13 line exhibit detectable levels of *ras* transcript and in some cases show RAS activity. This suggests that the second transformation may be responsible in altering cellular regulatory mechanisms. Evaluation of the HPPR activity also did not show a measureable level, the same problem as already described in Section 3.2.2, even though the HPPR gene was strongly expressed. Thus, our expectation of an increased formation of RAS-ester products by performing the double expression of *hppr* and *ras* in *Arabidopsis* still remains unfulfilled.

4.5 Metabolite accumulation in transformed *Arabidopsis*

Metabolite contents were analysed with the help of liquid chromatography-tandem mass spectrometry. The identification was based on the pattern of ion fragments of four ester products from RAS activity, i.e. Caf-DHPL (RA), Caf-pHPL, pC-DHPL, and pC-pHPL. The positive ion mode did not result in fragmentation as reported in a previous work (Petersen and Metzger 1993), but the negative ion mode resulted in useful fragmentation patterns. In addition, the negative ion mode is common for fragmentation of other phenolic acid esters and phenylpropanoid derivatives (Innocenti et al. 2006; Hossain et al. 2010; Alberti-Dér 2013).

The four esters are fragmented only in two main structures located on both sides of the ester bond. Ions corresponding to the 4-coumaroyl (*m/z* 163) and caffeoyl (*m/z* 161) moieties are typical fragments that resulted from their esters (Hossain et al. 2010; Alberti-Dér 2013; Di et al. 2013). Different fragmentation is observed between pHPL and DHPL, the alcohol parts of the esters. The fragmentation results in one abundant ion of pHPL (*m/z* 181) and two abundant ions of DHPL (*m/z* 179 and 197). It seems that the different numbers of hydroxyl groups might affect the fragmentation pattern. Here, the fragmentation pattern of RA is found similar to other reports (Hossain et al. 2010; Di et al. 2013; Ma et al. 2013).

The introduction of the RAS gene into *Arabidopsis* obviously induced the putative formation of two ester products, RA and Caf-pHPL. While RA was only found in the R13.1 line, Caf-pHPL seemed to be present in R13.2 line. The amounts of these two products are extremely low which might be associated with the low RAS activity found in the transformed cell lines. The putative products are neither present in other lines where the RAS activity and expression are lower than in R13.1 and R13.2, nor in L-lines. This suggests that the products may only be detectable if the RAS activity is above a certain threshold to be capable of catalysing the formation of detectable product amounts *in vivo*. In addition, the result confirms the role of HPR2 in *Arabidopsis* as the enzyme which can also participate in secondary metabolism by providing pHPL. However, neither cell suspension cultures nor hairy roots overexpressing the HPPR gene showed a detectable level of pHPL. This was also observed in RA-producing plants, where pHPL occurs in very low to undetectable levels. In *C. blumei* suspension cultures, a measurable accumulation of radioactively labelled pHPL occurred only after inhibition of PAL, showing a correlation between the two parallel pathways of RA-precursor formation (Ellis et al. 1979).

4.6 New insights and questions

Our results give new insights and open up a few questions. Based on the report from Morant et al. (2007), CYP98A3, A8 and A9 are not able to *in vitro* catalyse the *meta*-hydroxylation of pC-pHPL, the direct ester product of RAS activity which we had anticipated could possibly occur in the *ras*-transformed *Arabidopsis* cell lines. So, could those cytochrome P450s be capable of performing this *meta*-oxidation *in vivo*? If yes – as could be anticipated from putatively finding RA in R13.1 - why did we still find Caf-pHPL in R13.2 instead of RA? Which CYP is capable of introducing the hydroxyl group in 3'-position of pHPL and finishing the formation of RA? And if none of them, then why cannot we find any detectable level of pC-pHPL?

The already established biosynthesis pathway in *Coleus blumei* shows that the first ester product of RAS is built from the mono-hydroxylated substrates, 4-coumaroyl-CoA and 4-hydroxyphenyllactic acid (pHPL). However, it could still be possible that caffeoyl-CoA is a suitable donor substrate, since *in vitro* tests show that heterologously expressed RAS prefers caffeoyl- over coumaroyl-CoA in the reaction saturated with pHPL (Sander and Petersen 2011). Under normal conditions, 4-coumaroyl-CoA is not used directly for the subsequent formation of monolignols, glucose esters, or other phenolic esters in *Arabidopsis*, but it should be converted first to caffeoyl-CoA. In addition, free coumaric acid or its glucose- or CoA-esters are not accepted by CYP98A3 (Schoch et al. 2006) and also not by CYP98A6 and CYP98A14 from the RA-producing plants, *Lithospermum erythrorhizon* (Matsuno et al. 2002) and *C. blumei* (Petersen 1997; Eberle et al. 2009), respectively. Hence, 4-coumaroyl-CoA is first converted to the shikimate ester, then hydroxylated in *meta*-position by CYP98A3, and then hydrolysed by caffeoylshikimate esterase (CSE) or transesterified by a HCT resulting in caffeic acid or caffeoyl-CoA, respectively (Schoch et al. 2006; Vanholme et al. 2013). Shikimate esters are always described as transient intermediates not accumulating in high amounts in plant tissues (Schoch et al. 2001). Therefore, if we stick to the early established RA biosynthetic pathway, we could speculate that the first intermediate pC-pHPL may occur only very shortly in living cells and be immediately turned over to the more hydroxylated forms.

The previous RA biosynthetic pathway with pC-pHPL as the first ester product of RAS is based on the K_m-values of RAS for differently hydroxylated substrates as determined in plant protein extracts (Petersen 1991) as well as the fact that CYP98A14 from *C. blumei* only specifically introduces 3- and 3'-hydroxyl groups in the respective 4-coumaroyl moiety of pC-DHPL and 4'-hydroxyphenyllactate of Caf-pHPL (Eberle et al. 2009). CYP98A14 seems to hydroxylate 4-

coumaroyl moiety preferentially since the incubation with pC-pHPL resulted in Caf-pHPL only. Neither RA nor pC-DHPL resulted from the reaction after a longer incubation (Eberle 2008). Among these RA-like esters, Caf-pHPL, also named isorinic acid, has been identified in detectable levels in methanolic extracts from suspension cells of *C. blumei* and *Anthoceros agrestis* and it has been reported as a natural constituent of *Lycopus europaeus* (Petersen 1991) or *Helicteres isora* (Satake et al. 1999). Caf-pHPL has been found as the main product besides RA in genetically modified *E. coli* expressing a combinatorial pathway for RA biosynthesis (Bloch and Schmidt-Dannert 2014). An alternative pathway for RA biosynthesis has been proposed by Di et al. (2013) in *Salvia miltiorrhiza* suggesting that the *meta*-hydroxylation of pHPL to DHPL occurs prior to the esterification. Accordingly, the product pC-DHPL undergoes a single hydroxylation. However, a biosynthetic pathway is sometimes not rigid for certain metabolites. They function as "metabolic grids" rather than as metabolic pathways, especially for pathways that are connected or channeled with other biosynthetic pathways. For further experiments, studies such as knockout of those cytochrome P450 genes involved in RA biosynthesis or knockout of the HCT/CSE genes, might give us more information regarding the biosynthetic pathway of RA.

Because of the protein-binding properties of RA, the *ras*-transformed cells of *Arabidopsis* might protect themselves from the toxicity of RA by sequestering it into vacuoles or release it out of the cells to the medium. It is known that RA is accumulated in vacuoles of producing plants, i.e. *C. blumei, Ocimum basilicum*, and *Anchusa officinalis* (Chaprin and Ellis 1984; Häusler et al. 1993). The uptake of RA into *Arabidopsis* vacuoles could be possible, considering that many phenolic compounds have been reported to be transported into the vacuole, e.g. anthocyanins (Saito et al. 2013) or saponarin (Frangne et al. 2002). The latter compound is not even synthesised in *Arabidopsis*. The uptake mechanism of RA into the vacuole of *Coleus* probably involves carriers since RA is relatively hydrophilic and present almost exclusively as an anion at the cytoplasmic pH (Häusler et al. 1993). Different transport mechanisms such as ATP-binding cassette (ABC) transporters or multidrug resistance-associated protein (MRP)-like ABC transporters may be involved (Klein et al. 2001; Frangne et al. 2002), although investigations by Häusler et al. (1993) show that the transport into vacuoles is saturable, time-, pH- and temperature-dependent but not dependent on ATP. In *Arabidopsis*, several MRPs have been identified to be involved in transport of glutathione conjugates (Klein and Roos 2009). Such transporter type might be involved in the vacuolar transport of RA in *Arabidopsis* as well. Our results showed that RA and Caf-pHPL were present only in cell extracts not in the medium. Unfortunately, we did not perform a stability test

of RA in the medium of the cell cultures. After feeding of RA into the habituated, cell-free medium of *C. blumei* cell cultures, it was undetectable after very short time periods (our unpublished work). Although we did not accomplish this study thoroughly, we could speculate that vacuoles are a prerequisite of cells to accumulate RA. Therefore, further experiments might be necessary to study the transport mechanism of RA in plant cells. This is very important, for example, when using plant cell culture technology as a means to commercially produce RA.

5. Summary

Rosmarinic acid (RA) is an ester of caffeic acid and 3,4-dihydroxyphenyllactic acid (DHPL). This compound is present in various higher plants and some lower plants. However, the occurrence is not consistent since not all members in each level of plant taxa where RA has been shown to be present contain RA. Recent investigations confirm that a number of enzymes involved in the formation of RA are present in all higher plant species and also active in other metabolic pathways. In *Arabidopsis thaliana* (Brassicaceae), the presence of certain RA biosynthesis genes has been described, but one or more RA specific genes are missing and thus the plant is unable to synthesise RA. Therefore, introducing the cDNA of CbRAS into *Arabidopsis* might prove the role of enzymes such as HPR2 (a photorespiratory enzyme) and cytochrome P450 monooxygenases from the CYP98A family (enzymes in the formation of monolignols) in providing the substrates and finishing the formation of RA or RA-like esters.

The transformed *Arabidopsis thaliana* was obtained via *Agrobacterium tumefaciens*-mediated transformation and has been successfully established as cell suspension cultures. The transformed cells were R-lines (plants transformed with the vector containing cDNA of CbRAS) and L-lines (control plants transformed with the empty vector). Molecular characterisation showed that the T-DNA parts, the RAS and CaMV 35S promoter sequences, have been stably integrated in the plant genome. However, the presence of the KanR gene (a non-T-DNA part) in all established cell suspension cultures revealed a phenomenon which is called the integration of the non-T-DNA binary vector backbone sequences. This has been reported in several plant transformation events.

The optimum growth of cell suspension cultures including the cell biomass and protein contents was observed in the first week of the cultivation period. Overexpression of the RAS gene slightly reduced the cell growth but did not impair the cells. In addition, strong expression of the RAS gene did not increase the protein content.

All the RA biosynthetic enzymes were found to be highly active during the logarithmic growth phase. Enzymes such as TAT, C4H and 4CL reached their maximum activity on day 4, while PAL and RAS on day 6. Overexpression of RAS gene had no influence on the activity of the enzymes in the upstream pathway of RA. Their activities were found to be cell line-specific.

The strong expression of the RAS gene did not make all *ras*-transformed lines having RAS activity. Some of them had exhibited the activity, but later on, the activity was lost. Gene expression analysis did not show any post-transcriptional gene silencing (PTGS) event. Therefore, a complex

regulation at the post-translational level might be involved.

The transcript level of the RAS gene varied among *ras*-transformed lines which might be caused by the number and site(s) of T-DNA integration events or cellular responses of the cell lines. The expression pattern of other RA-biosynthetic genes cannot be clearly distinguished between *ras*-transformed lines, control lines and the wild type. In general, gene transcripts fluctuated and were line-specific.

The attempt to overexpress two genes, CbHPPR and CbRAS, in *Arabidopsis* was achieved by generating hairy roots from *ras*-transformed lines with the help of *Agrobacterium rhizogenes* carrying the CbHPPR gene. Only one single hairy root line could be obtained containing both genes. However, the transformed roots exhibited neither RAS activity nor detectable levels of *ras* transcript; this might be caused by PTGS events. Also, the transformed roots did not show any significant change of the HPPR activity although high expression of the HPPR gene was clearly observed. Accordingly, the attempt to increase the formation of RAS-ester products by double expression of *hppr* and *ras* remains unsuccessful.

Analysis of metabolite contents by LC-MS showed that *ras*-transformed cells probably accumulated RA and Caf-pHPL, although the yields were extremely low. Thus, the presence of RAS in *Arabidopsis thaliana* demonstrates the dual role of HP(P)R in primary and secondary metabolism, by providing glycerate (in photorespiration) and pHPL (for RA biosynthesis), respectively. It also leads us into new perspectives regarding the role of CYP98As in *Arabidopsis*, the biosynthetic pathway of RA, and the role of vacuoles for RA accumulation. This, however, should be answered in future experiments.

6. Zusammenfassung

Rosmarinsäure (RA) ist ein Ester von Kaffeesäure und 3,4-Dihydroxyphenylmilchsäure (DHPL). Dieser Naturstoff kommt in verschiedenen höheren Pflanzen und einigen niedrigeren Pflanzen vor. Allerdings ist das Vorkommen nicht konsistent, da nicht alle Mitglieder in jeder Taxonebene, in der RA gezeigt worden ist, RA enthalten. Neuere Untersuchungen bestätigen, dass einige Enzyme, die an der Bildung von RA beteiligt sind, in allen höheren Pflanzenarten vorhanden und auch in anderen Stoffwechselwegen aktiv sind. In *Arabidopsis thaliana* (Brassicaceae) wurde die Anwesenheit bestimmter RA-Biosynthese-Gene beschrieben, aber ein oder mehrere RA-spezifische Gene fehlen und somit ist die Pflanze nicht in der Lage, RA zu synthetisieren. Daher könnte das Einbringen der cDNA von CbRAS in *Arabidopsis* zeigen, ob Enzyme wie HPR2 (ein photorespiratorisches Enzym) und Cytochrom P450 Monooxygenasen aus der CYP98A-Familie (Enzyme bei der Bildung von Monolignolen) an der Bereitstellung der Substrate und der Bildung von RA oder RA-ähnlichen Estern beteiligt sein können.

CbRAS-transformierte *Arabidopsis thaliana* wurden durch *Agrobacterium tumefaciens*-vermittelte Transformation gewonnen und erfolgreich als Zellsuspensionskulturen etabliert. Es wurden R-Linien (Pflanzen, die mit dem Vektor transformiert wurden, der cDNA von CbRAS enthielt) und L-Linien (Kontrollpflanzen, die mit dem leeren Vektor transformiert wurden) gewonnen. Die molekulare Charakterisierung zeigte, dass die T-DNA-Teile (RAS- und CaMV-35S-Promotorsequenzen) stabil in das Pflanzengenom integriert wurden. Die Anwesenheit des KanR-Gens (eines Nicht-T-DNA-Teils) in allen etablierten Zellsuspensionskulturen zeigte jedoch ein Phänomen, das auf die Integration der Nicht-T-DNA-Sequenzen des Vektors beruht. Dies wurde schon mehrfach in der Literatur beschrieben.

Optimales Wachstum der Zellsuspensionskulturen, bestimmt anhand der Zellbiomasse und des Proteingehalts, wurde in der ersten Woche der Kultivierungsperiode beobachtet. Die Überexpression des RAS-Gens verringerte das Zellwachstum leicht, beeinträchtigte aber die Zellen nicht. Die starke Expression des RAS-Gens erhöhte nicht den Proteingehalt.

Alle RA-Biosynthese-Enzyme hatten ihre höchste Aktivität während der logarithmischen Wachstumsphase. Enzyme wie TAT, C4H und 4CL erreichten ihre maximale Aktivität am 4. Tag, PAL und RAS am 6. Tag. Die Überexpression des RAS-Gens hatte keinen Einfluss auf die Aktivität der im RA-Biosyntheseweg vorgeschalteten Enzyme. Ihre Aktivitäten wurden als zelllinienspezifisch eingestuft.

Die starke Expression des RAS-Gens führte nicht in allen *ras*-transformierten Linien zu einer messbaren RAS-Aktivität. Einige Linien hatten anfangs nachweisbare RAS-Aktivität, die jedoch mit zunehmender Passagenanzahl verloren ging. Die Genexpressionsanalyse zeigte, dass posttranskriptionelles Gen Silencing (PTGS) nicht die Ursache hierfür war. Daher könnte eine komplexe Regulation auf der posttranslationalen Ebene beteiligt sein.

Die RAS-Transkriptmenge variierte unter den *ras*-transformierten Linien. Dies könnte durch die Anzahl und die Stelle(n) von T-DNA-Integrationsereignissen oder zellulären Reaktionen der Zelllinien verursacht werden. Das Expressionsmuster anderer an der RA-Biosynthese beteiligter Gene lässt sich nicht klar zwischen *ras*-transformierten Linien, Kontrolllinien und dem Wildtyp unterscheiden. Im Allgemeinen sind die Transkriptlevel schwankend und linienspezifisch.

Der Versuch, zwei Gene, CbHPPR und CbRAS, gleichzeitig in *Arabidopsis* zu überexprimieren, wurde durch die Erzeugung von „Hairy Root“ aus *ras*-transformierten Linien mit Hilfe von *Agrobacterium rhizogenes* erreicht, das das CbHPPR-Gen tragen. Es konnte nur eine einzige „Hairy Root“-Linie erhalten werden, die beide Gene überführte. Die transformierten Wurzeln zeigten jedoch weder RAS-Aktivität noch nachweisbare Mengen an *ras*-Transkript. Dies könnte durch PTGS-Ereignisse verursacht werden. Außerdem zeigten die transformierten Wurzeln keine signifikante Veränderung der HPPR-Aktivität, obwohl das HPPR-Gen stark exprimiert wurde. Dementsprechend bleibt der Versuch erfolglos, die Bildung von RAS-Esterprodukten durch doppelte Expression von *hppr* und *ras* zu erhöhen.

Die phenolischen Inhaltsstoffe wurden über LC-MS analysiert. Hierbei zeigte sich, dass *ras*-transformierte Zellen mit hoher Wahrscheinlichkeit RA und Caf-pHPL akkumulierten, obwohl die Ausbeute extrem niedrig war. Dies ist ein Hinweis darauf, dass bei Anwesenheit von RAS in *Arabidopsis thaliana* HP(P)R eine doppelte Rolle im Primär- und Sekundärmetabolismus hat, indem sie Glycerat (in der Photorespiration) und pHPL (für die RA-Biosynthese) lieferte. Daraus ergeben sich neue Fragen in Bezug auf die Rolle von CYP98As in *Arabidopsis*, den Biosyntheseweg der RA und die Rolle der Vakuolen für die RA-Akkumulation. Dies sollte jedoch in zukünftigen Experimenten geklärt werden.

7. References

Abdullah Y, Schneider B, Petersen M (2008) Occurrence of rosmarinic acid, chlorogenic acid and rutin in Marantaceae species. Phytochem Lett 1:199–203.

Abdulrazzak N, Pollet B, Ehlting J, et al (2006) A coumaroyl-ester-3-hydroxylase insertion mutant reveals the existence of nonredundant meta-hydroxylation pathways and essential roles for phenolic precursors in cell expansion and plant growth. Plant Physiol 140:30–48.

Aharoni A, Giri AP, Deuerlein S, et al (2003) Terpenoid metabolism in wild-type and transgenic *Arabidopsis* plants. Plant Cell 15:2866–2884.

Alberti-Dér Á (2013) LC-ESI-MS/MS methods in profiling of flavonoid glycosides and phenolic acids in traditional medicinal plants: *Sempervivum tectorum* L. and *Corylus avellana* L. Dissertation, Semmelweis University Budapest.

Amoah SKS, Sandjo LP, Kratz JM, Biavatti MW (2016) Rosmarinic acid - pharmaceutical and clinical aspects. Planta Med 82:388–406.

Anderson M, Wilson F (2002) *Arabidopsis*: a practical approach. Oxford University Press.

Bais HP, Walker TS, Schweizer HP, Vivanco JM (2002) Root specific elicitation and antimicrobial activity of rosmarinic acid in hairy root cultures of *Ocimum basilicum*. Plant Physiol Biochem 40:983–995.

Bak S, Beisson F, Bishop G, et al (2011) Cytochromes P450. In: The *Arabidopsis* Book. The American Society of Plant Biologists, p e0144.

Bak S, Olsen CE, Halkier BA, Møller BL (2000) Transgenic tobacco and *Arabidopsis* plants expressing the two multifunctional *Sorghum* cytochrome P450 enzymes, CYP79A1 and CYP71E1, are cyanogenic and accumulate metabolites derived from intermediates in dhurrin biosynthesis. Plant Physiol 123:1437–1448.

Bak S, Olsen CE, Petersen BL, et al (1999) Metabolic engineering of *p*-hydroxybenzylglucosinolate in *Arabidopsis* by expression of the cyanogenic CYP79A1 from *Sorghum bicolor*. Plant J 20:663–671.

Bauer N, Kiseljak D, Jelaska S (2009) The effect of yeast extract and methyl jasmonate on rosmarinic acid accumulation in *Coleus blumei* hairy roots. Biol Plant 53:650–656.

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997) Cinnamate-4-hydroxylase expression in *Arabidopsis* (regulation in response to development and the environment). Plant Physiol 113:729–738.

Benchabane M, Goulet C, Rivard D, et al (2008) Preventing unintended proteolysis in plant protein biofactories. Plant Biotechnol J 6:633–648.

Bennett MD, Leitch IJ, Price HJ, Johnston JS (2003) Comparisons with *Caenorhabditis* (100 Mb) and *Drosophila* (175 Mb) using flow cytometry show genome size in *Arabidopsis* to be 157 Mb and thus 25 % larger than the *Arabidopsis* genome initiative estimate of 125 Mb. Ann Bot 91:547–557.

Berger A, Meinhard J, Petersen M (2006) Rosmarinic acid synthase is a new member of the superfamily of BAHD acyltransferases. Planta 224:1503–1510.

Besseau S, Hoffmann L, Geoffroy P, et al (2007) Flavonoid accumulation in *Arabidopsis* repressed in lignin synthesis affects auxin transport and plant growth. Plant Cell 19:148–162.

Bloch SE, Schmidt-Dannert C (2014) Construction of a chimeric biosynthetic pathway for the *de novo* biosynthesis of rosmarinic acid in *Escherichia coli*. Chembiochem 15:2393–2401.

Blount JW, Korth KL, Masoud SA, et al (2000) Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway. Plant Physiol 122:107–116.

Bonhomme V, Laurain-Mattar D, Lacoux J, et al (2000) Tropane alkaloid production by hairy roots of *Atropa belladonna* obtained after transformation with *Agrobacterium rhizogenes* 15834 and *Agrobacterium tumefaciens* containing rol A, B, C genes only. J Biotechnol 81:151–158.

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254.

Bulgakov VP, Inyushkina YV, Fedoreyev SA (2012) Rosmarinic acid and its derivatives: biotechnology and applications. Crit Rev Biotechnol 32:203–217.

Chandra S, Chandra R (2011) Engineering secondary metabolite production in hairy roots. Phytochem Rev 10:371–395.

Chen H, Chen F, Zhang YL, Song JY (1999) Production of rosmarinic acid and lithospermic acid B in Ti transformed *Salvia miltiorrhiza* cell suspension cultures. Process Biochem 34:777–784.

Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159.

Chu Y, Kwon T, Nam J (2014) Enzymatic and metabolic engineering for efficient production of syringin, sinapyl alcohol 4-*O*-glucoside, in *Arabidopsis thaliana*. Phytochemistry 102:55–63.

Cochrane FC, Davin LB, Lewis NG (2004) The *Arabidopsis* phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65:1557–1564.

Costa MA, Collins RE, Anterola AM, et al (2003) An *in silico* assessment of gene function and organization of the phenylpropanoid pathway metabolic networks in *Arabidopsis thaliana* and limitations thereof. Phytochemistry 64:1097–1112.

D`Auria JC, Chen F, Pichersky E (2002) Characterization of an acyltransferase capable of synthesizing benzylbenzoate and other volatile esters in flowers and damaged leaves of *Clarkia breweri*. Plant Physiol 130:466–476.

D´Auria JC (2006) Acyltransferases in plants : a good time to be BAHD. Curr Opin Plant Biol 9:331–340.

D´Auria JC, Gershenzon J (2005) The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. Curr Opin Plant Biol 8:308–316.

D´Auria JC, Reichelt M, Luck K (2007) Identification and characterization of the BAHD acyltransferase malonyl-CoA: anthocyanidin 5-*O*-glucoside-6"-*O*-malonyltransferase (At5MAT) in *Arabidopsis thaliana*. FEBS Lett 581:872–878.

De Buck S, De Wilde C, van Montagu M, Depicker A (2000) T-DNA vector backbone sequences are frequently integrated into the genome of transgenic plants obtained by *Agrobacterium*-mediated transformation. Mol Breed 6:459–468.

De Buck S, Podevin N, Nolf J, et al (2009) The T-DNA integration pattern in *Arabidopsis* transformants is highly determined by the transformed target cell. Plant J 60:134–145.

Di P, Zhang L, Chen J, et al (2013) 13C tracer reveals phenolic acids biosynthesis in hairy root cultures of *Salvia miltiorrhiza*. ACS Chem Biol 8:1537–1548.

Döring AS, Petersen M (2014) Production of caffeic, chlorogenic and rosmarinic acids in plants and suspension cultures of *Glechoma hederacea*. Phytochem Lett 10:cxi–cxvii.

Eberle D (2008) Klonierung und funktionelle Expression von CYP98A14 und der NADPH:Cytochrom P450 Reduktase aus *Coleus blumei*. Dissertation, Philipps-Universität Marburg.

Eberle D, Ullmann P, Werck-Reichhart D, Petersen M (2009) cDNA cloning and functional characterisation of CYP98A14 and NADPH:cytochrome P450 reductase from *Coleus blumei* involved in rosmarinic acid biosynthesis. Plant Mol Biol 69:239–253.

Ehlting J, Büttner D, Wang Q, et al (1999) Three 4-coumarate:coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. Plant J 19:9–20.

Ehlting J, Sauveplane V, Olry A, et al (2008) An extensive (co-)expression analysis tool for the cytochrome P450 superfamily in *Arabidopsis thaliana*. BMC Plant Biol 8:47.

Ellis BE, Remmen S, Goeree G (1979) Interactions between parallel pathways during biosynthesis of rosmarinic acid in cell suspension cultures of *Coleus blumei*. Planta 147:163–167.

Ellis BE, Towers GHN (1970) Biogenesis of rosmarinic acid in *Mentha*. Biochem J 118:291–297.

Fellenberg C, Böttcher C, Vogt T (2009) Phenylpropanoid polyamine conjugate biosynthesis in *Arabidopsis thaliana* flower buds. Phytochemistry 70:1392–1400.

Frangne N, Eggmann T, Koblischke C, et al (2002) Flavone glucoside uptake into barley mesophyll and *Arabidopsis* cell culture vacuoles. Energization occurs by H^+-antiport and ATP-binding cassette-type mechanisms. Plant Physiol 128:726–733.

Franke R, Hemm MR, Denault JW, et al (2002a) Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of *Arabidopsis*. Plant J 30:47–59.

Franke R, Humphreys JM, Hemm MR, et al (2002b) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J 30:33–45.

Fraser CM, Chapple C (2011) The phenylpropanoid pathway in *Arabidopsis*. In: The *Arabidopsis* Book. The American Society of Plant Biologists, p e0152

Friso G, van Wijk KJ (2015) Posttranslational protein modifications in plant metabolism. Plant Physiol 169:1469–1487.

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res 50:151–158.

Gertlowski C, Petersen M (1993) Influence of the carbon source on growth and rosmarinic acid production in suspension cultures of *Coleus blumei*. Plant Cell Tissue Organ Cult 34:183–190.

Ghedira R, De Buck S, Van Ex F, et al (2013) T-DNA transfer and T-DNA integration efficiencies upon *Arabidopsis thaliana* root explant cocultivation and floral dip transformation. Planta 238:1025–1037.

Gindro K, Berger V, Godard S, et al (2012) Protease inhibitors decrease the resistance of Vitaceae to *Plasmopara viticola*. Plant Physiol Biochem 60:74–80.

Giri A, Narasu ML (2000) Transgenic hairy roots: recent trends and applications. Biotechnol Adv 18:1–22.

Goodman HM, Ecker JR, Dean C (1995) The genome of *Arabidopsis thaliana*. Proc Natl Acad Sci USA 92:10831–10835.

Gou J, Yu X, Liu C (2009) A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in *Arabidopsis*. Proc Natl Acad Sci USA 106:18855–18860.

Grienenberger E, Besseau S, Geoffroy P, et al (2009) A BAHD acyltransferase is expressed in the tapetum of *Arabidopsis* anthers and is involved in the synthesis of hydroxycinnamoyl spermidines. Plant J 58:246–259.

Grossmann K, Hutzler J, Tresch S, et al (2012) On the mode of action of the herbicides cinmethylin and 5-benzyloxymethyl-1,2-isoxazolines: putative inhibitors of plant tyrosine aminotransferase. Pest Manag Sci 68:482–492.

Haas JH, Moore LW, Ream W, Manulis S (1995) Universal PCR primers for detection of phytopathogenic *Agrobacterium* strains. Appl Environ Microbiol 61:2879–2884.

Hamberger B, Hahlbrock K (2004) The 4-coumarate:CoA ligase gene family in *Arabidopsis thaliana* comprises one rare, sinapate-activating and three commonly occurring isoenzymes. Proc Natl Acad Sci USA 101:2209–2214.

Hampp C, Richter A, Osorio S, et al (2012) Establishment of a photoautotrophic cell suspension culture of *Arabidopsis thaliana* for photosynthetic, metabolic, and signaling studies. Mol Plant 5:524–527.

Häusler E, Petersen M, Alfermann A (1993) Isolation of protoplasts and vacuoles from cell suspension cultures of *Coleus blumei* Benth. Plant Cell Rep 12:510–512.

Hemmerlin A (2013) Post-translational events and modifications regulating plant enzymes involved in isoprenoid precursor biosynthesis. Plant Sci 203–204:41–54.

Holländer-Czytko H, Grabowski J, Sandorf I, et al (2005) Tocopherol content and activities of tyrosine aminotransferase and cystine lyase in *Arabidopsis* under stress conditions. J Plant Physiol 162:767–770.

Hossain MB, Rai DK, Brunton NP, et al (2010) Characterization of phenolic composition in Lamiaceae spices by LC-ESI-MS/MS. J Agric Food Chem 58:10576–10581.

http://www.arabidopsis.com/main/tbl/chronology.htm. Chronology of *Arabidopsis*. Accessed 1 Aug 2014.

https://www.arabidopsis.org/. The *Arabidopsis* information resource. Accessed 7 Oct 2015.

https://www.arabidopsis.org/portals/genAnnotation/gene_structural_annotation/annotation_data.jsp. TAIR - Genome annotation. Accessed 6 Oct 2015.

Hu YS, Zhang L, Di P, Chen WS (2009) Cloning and induction of phenylalanine ammonia-lyase gene from *Salvia miltiorrhiza* and its effect on hydrophilic phenolic acids levels. Chin J Nat Med 7:449–457.

Huang BB, Yi B, Duan YB, et al (2008) Characterization and expression profiling of tyrosine aminotransferase gene from *Salvia miltiorrhiza* (Dan-shen) in rosmarinic acid biosynthesis pathway. Mol Biol Rep 35:601–612.

Huang J, Gu M, Lai Z, et al (2010) Functional analysis of the *Arabidopsis* PAL gene family in plant growth, development, and response to environmental stress. Plant Physiol 153:1526–1538.

Hücherig S (2010) RNAi-Untersuchungen und Überexpression von Genen der Rosmarinsäurebiosynthese. Dissertation, Philipps-Universität Marburg.

Hücherig S, Petersen M (2012) RNAi suppression and overexpression studies of hydroxyphenylpyruvate reductase (HPPR) and rosmarinic acid synthase (RAS) genes related to rosmarinic acid biosynthesis in hairy root cultures of *Coleus blumei*. Plant Cell Tissue Organ Cult 113:375–385.

Innocenti M, la Marca G, Malvagia S, et al (2006) Electrospray ionisation tandem mass spectrometric investigation of phenylpropanoids and secoiridoids from solid olive residue. Rapid Commun Mass Spectrom 20:2013–2022.

Kasajima I, Ide Y, Ohkama-Ohtsu N, et al (2004) A protocol for rapid DNA extraction from *Arabidopsis thaliana* for PCR analysis. Plant Mol Biol Report 22:49–52.

Kim GD, Park YS, Jin YH, Park CS (2015) Production and applications of rosmarinic acid and structurally related compounds. Appl Microbiol Biotechnol 99:2083–2092.

Kim KH, Janiak V, Petersen M (2004) Purification, cloning and functional expression of hydroxyphenylpyruvate reductase involved in rosmarinic acid biosynthesis in cell cultures of *Coleus blumei*. Plant Mol Biol 54:311–323.

Kim SJ, Hong SM, Yoo SJ, et al (2016) Post-translational regulation of FLOWERING LOCUS T protein in *Arabidopsis*. Mol Plant 9:308–311.

Klein M, Martinoia E, Hoffmann-Thoma G, Weissenböck G (2001) The ABC-like vacuolar transporter for rye mesophyll flavone glucuronides is not species-specific. Phytochemistry 56:153–159.

Klein M, Roos W (2009) Handling dangerous molecules: transport and compartmentation of plant natural products. In: Plant-derived Natural Products. Springer US, New York, NY, pp 229–267

Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector "backbone" sequences into the tobacco genome: evidence for multiple complex patterns of integration. Plant J 11:945–957.

Koornneef M, Meinke D (2010) The development of *Arabidopsis* as a model plant. Plant J 61:909–921.

Korth KL, Jaggard DAW, Dixon RA (2000) Developmental and light-regulated post-translational control of 3-hydroxy-3-methylglutaryl-CoA reductase levels in potato. Plant J 23:507–516.

Landry LG, Chapple CC, Last RL (1995) *Arabidopsis* mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiol 109:1159–1166.

Lee D, Ellard M, Wanner LA, et al (1995) The *Arabidopsis thaliana* 4-coumarate:CoA ligase (4CL) gene: stress and developmentally regulated expression and nucleotide sequence of its cDNA. Plant Mol Biol 28:871–884.

Lee D, Meyer K, Chapple C, Douglas CJ (1997) Antisense suppression of 4-coumarate:coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. Plant Cell 9:1985–1998.

Lee SY, Xu H, Kim YK, Park SU (2008) Rosmarinic acid production in hairy root cultures of *Agastache rugosa* Kuntze. World J Microbiol Biotechnol 24:969–972.

Leyva A, Jarillo JA, Salinas J, Martinez-Zapater JM (1995) Low temperature induces the accumulation of phenylalanine ammonia-lyase and chalcone synthase mRNAs of *Arabidopsis thaliana* in a light-dependent manner. Plant Physiol 108:39–46.

Li W, Koike K, Asada Y, et al (2005) Rosmarinic acid production by *Coleus forskohlii* hairy root cultures. Plant Cell Tissue Organ Cult 80:151–155.

Li X, Bonawitz ND, Weng JK, Chapple C (2010) The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. Plant Cell 22:1620–1632.

Li XW, Li JW, Zhai Y, et al (2013) A R2R3-MYB transcription factor, GmMYB12B2, affects the expression levels of flavonoid biosynthesis genes encoding key enzymes in transgenic *Arabidopsis* plants. Gene 532:72–79.

Li Y, Kim JI, Pysh L, Chapple C (2015) Four isoforms of *Arabidopsis* 4-coumarate:CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. Plant Physiol 169:2409–2421.

Lu X, Hao L, Wang F, et al (2013) Molecular cloning and overexpression of the tyrosine aminotransferase (TAT) gene leads to increased rosmarinic acid yield in *Perilla frutescens*. Plant Cell, Tissue Organ Cult 115:69–83.

Luis JC, Pérez RM, González FV (2007) UV-B radiation effects on foliar concentrations of rosmarinic and carnosic acids in rosemary plants. Food Chem 101:1211–1215.

Luo J, Fuell C, Parr A, et al (2009) A novel polyamine acyltransferase responsible for the accumulation of spermidine conjugates in *Arabidopsis* seed. 21:318–333.

Luo J, Nishiyama Y, Fuell C, et al (2007) Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. Plant J 50:678–695.

Ma B, Wang Y, Zhang Q, et al (2013) Simultaneous determination of oridonin, ponicidin and rosmarinic acid from herba *Isodi rubescentis* extract by LC-MS-MS in rat plasma. J Chromatogr Sci 51:910–918.

Mannan A, Shaheen N, Arshad W, et al (2008) Hairy roots induction and artemisinin analysis in *Artemisia dubia* and *Artemisia indica*. African J Biotechnol 7:3288–3292.

Matkowski A (2008) Plant *in vitro* culture for the production of antioxidants - A review. Biotechnol Adv 26:548–560.

Matsuno M, Compagnon V, Schoch GA, et al (2009) Evolution of a novel phenolic pathway for pollen development. Science 325:1688–1692.

Matsuno M, Nagatsu A, Ogihara Y, et al (2002) CYP98A6 from *Lithospermum erythrorhizon* encodes 4-coumaroyl-4′-hydroxyphenyllactic acid 3-hydroxylase involved in rosmarinic acid biosynthesis. FEBS Lett 514:219–224.

Meinke DW, Cherry JM, Dean C, et al (1998) *Arabidopsis thaliana*: a model plant for genome analysis. Science 282:662–682.

Meyerowitz EM (1987) *Arabidopsis thaliana*. Annu Rev Genet 21:93–111.

Milkowski C, Strack D (2010) Sinapate esters in brassicaceous plants: biochemistry, molecular biology, evolution and metabolic engineering. Planta 232:19–35.

Milkowski C, Strack D (2004) Serine carboxypeptidase-like acyltransferases. Phytochemistry 65:517–524.

Mitchell-Olds T (2001) *Arabidopsis thaliana* and its wild relatives: a model system for ecology and evolution. Trends Ecol Evol 16:693–700.

Mizutani M, Ohta D, Sato R (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from *Arabidopsis* and its expression manner in planta. Plant Physiol 113:755–763.

Morant M, Schoch GA, Ullmann P, et al (2007) Catalytic activity, duplication and evolution of the CYP98 cytochrome P450 family in wheat. Plant Mol Biol 63:1–19.

Nair RB, Xia Q, Kartha CJ, et al (2002) *Arabidopsis* CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. Plant Physiol 130:210–220.

Ohl S, Hedrick SA, Chory J, Lamb CJ (1990) Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. Plant Cell 2:837–848.

Olsen KM, Lea US, Slimestad R, et al (2008) Differential expression of four *Arabidopsis* PAL genes; PAL1 and PAL2 have functional specialization in abiotic environmental-triggered flavonoid synthesis. J Plant Physiol 165:1491–1499.

Ozawa S, Yasutani I, Fukuda H, et al (1998) Organogenic responses in tissue culture of srd mutants of *Arabidopsis thaliana*. Development 125:135–142.

Peterhansel C, Horst I, Niessen M, et al (2010) Photorespiration. In: The *Arabidopsis* Book. The American Society of Plant Biologists, p e0130.

Petersen M (2013) Rosmarinic acid: new aspects. Phytochem Rev 12:207–227.

Petersen M (1991) Characterization of rosmarinic acid synthase from cell cultures of *Coleus blumei*. Phytochemistry 30:2877–2881.

Petersen M (1993) Purification of rosmarinic acid synthase from cell cultures of *Coleus blumei* Benth. Planta 191:18–22.

Petersen M (2003) Cinnamic acid 4-hydroxylase from cell cultures of the hornwort *Anthoceros agrestis*. Planta 217:96–101.

Petersen M (1997) Cytochrome P450-dependent hydroxylation in the biosynthesis of rosmarinic acid in *Coleus*. Phytochemistry 45:1165–1172.

Petersen M, Abdullah Y, Benner J, et al (2009) Evolution of rosmarinic acid biosynthesis. Phytochemistry 70:1663–1679.

Petersen M, Alfermann AW (1988) Two new enzymes of rosmarinic acid biosynthesis from cell cultures of *Coleus blumei*: hydroxyphenylpyruvate reductase and rosmarinic acid synthase. Z Naturforsch 43c:501–504.

Petersen M, Metzger JW (1993) Identification of the reaction products of rosmarinic acid synthase from cell cultures of *Coleus blumei* by ion spray mass spectrometry and tandem mass spectrometry. Phytochem Anal 4:131–134.

Petersen M, Simmonds MSJ (2003) Rosmarinic acid. Phytochemistry 62:121–125.

Pillay P, Schlüter U, van Wyk S, et al (2014) Proteolysis of recombinant proteins in bioengineered plant cells. Bioengineered 5:15–20.

Prabhu PR, Hudson AO (2010) Identification and partial characterization of an L-tyrosine aminotransferase (TAT) from *Arabidopsis thaliana*. Biochem Res Int: Article ID 549572.

Raes J, Rohde A, Christensen JH, et al (2003) Genome-wide characterization of the lignification toolbox in *Arabidopsis*. Plant Physiol 133:1051–1071.

Rao SR, Ravishankar GA (2002) Plant cell cultures: chemical factories of secondary metabolites. Biotechnol Adv 20:101–153.

Reuben S, Rai A, Pillai BVS, et al (2013) A bacterial quercetin oxidoreductase QuoA-mediated perturbation in the phenylpropanoid metabolic network increases lignification with a concomitant decrease in phenolamides in *Arabidopsis*. J Exp Bot 64:5183–5194.

Riewe D, Koohi M, Lisec J, et al (2012) A tyrosine aminotransferase involved in tocopherol synthesis in *Arabidopsis*. Plant J 71:850–859.

Rogers S, Bendich A (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Plant Mol Biol 5:69–76.

Rohde A, Morreel K, Ralph J, et al (2004) Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell 16:2749–2771.

Ruegger M, Chapple C (2001) Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. Genetics 159:1741–1749.

Russell DW, Davidson H (1982) Regulation of cytosolic HMG-CoA reductase activity in pea seedlings: Contrasting responses to different hormones, and hormone-product interaction, suggest hormonal modulation of activity. Biochem Biophys Res Commun 104:1537–1543.

Sander M, Petersen M (2011) Distinct substrate specificities and unusual substrate flexibilities of two hydroxycinnamoyltransferases, rosmarinic acid synthase and hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase, from *Coleus blumei* Benth. Planta 233:1157–1171.

Sandorf I, Holländer-Czytko H (2002) Jasmonate is involved in the induction of tyrosine aminotransferase and tocopherol biosynthesis in *Arabidopsis thaliana*. Planta 216:173–179.

Sasaki K, Mito K, Ohara K, et al (2008) Cloning and characterization of naringenin 8-prenyltransferase, a flavonoid-specific prenyltransferase of *Sophora flavescens*. Plant Physiol 146:1075–1084.

Satake T, Kamiya K, Saiki Y, et al (1999) Studies on the constituents of fruits of *Helicteres isora* L. Chem Pharm Bull (Tokyo) 47:1444–1447.

Scarpati M, Oriente G (1958) Isolamento e costituzione dell'acido rosmarinico (dal rosmarinus off.). Ric Sci 28:2329–2333.

Schilmiller AL, Stout J, Weng JK, et al (2009) Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in *Arabidopsis*. Plant J 60:771–782.

Schmülling T, Schell J, Spena A (1988) Single genes from *Agrobacterium rhizogenes* influence plant development. EMBO J 7:2621–2629.

Schoch GA, Goepfert S, Morant M, et al (2001) CYP98A3 from *Arabidopsis thaliana* is a 3′-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem 276:36566–36574.

Schoch GA, Morant M, Abdulrazzak N, et al (2006) The meta-hydroxylation step in the phenylpropanoid pathway: a new level of complexity in the pathway and its regulation. Environ Chem Lett 4:127–136.

Singleton P, Keyser H, Sande E (2002) Development and evaluation of liquid inoculants: inoculants and nitrogen fixation of legumes. In: ACIAR Proceeding. Vietnam, p 109e.

Soledade M, Pedras C, Zheng Q (2010) The chemistry of *Arabidopsis thaliana*. In: Liu HW, Mander L (eds) Comprehensive Natural Products II. Elsevier, Oxford, p 1297–1315.

Song J, Wang Z (2011) RNAi-mediated suppression of the phenylalanine ammonia-lyase gene in *Salvia miltiorrhiza* causes abnormal phenotypes and a reduction in rosmarinic acid biosynthesis. J Plant Res 124:183–192.

Song J, Wang ZZ (2009) Molecular cloning, expression and characterization of a phenylalanine ammonia-lyase gene (SmPAL1) from *Salvia miltiorrhiza*. Mol Biol Rep 36:939–952.

Spena A, Schmülling T, Koncz C, Schell JS (1987) Independent and synergistic activity of rol A, B and C loci in stimulating abnormal growth in plants. EMBO J 6:3891–3899.

Srivastava S, Conlan XA, Adholeya A, Cahill DM (2016) Elite hairy roots of *Ocimum basilicum* as a new source of rosmarinic acid and antioxidants. Plant Cell Tissue Organ Cult 126:19–32.

Stehle F, Brandt W, Stubbs MT, et al (2009) Sinapoyltransferases in the light of molecular evolution. Phytochemistry 70:1652–1662.

Stehle F, Stubbs MT, Strack D, Milkowski C (2008) Heterologous expression of a serine carboxypeptidase-like acyltransferase and characterization of the kinetic mechanism. FEBS J 275:775–787.

Strehmel N, Böttcher C, Schmidt S, Scheel D (2014) Profiling of secondary metabolites in root exudates of *Arabidopsis thaliana*. Phytochemistry 108:35–46.

Sugimoto K, Meyerowitz EM (2013) Regeneration in *Arabidopsis* tissue culture. In: De Smet I (ed) Plant organogenesis: methods and protocols, methods in molecular biology. Springer Science+Business Media, New York, p 265–275.

Sugiyama A, Linley PJ, Sasaki K, et al (2011) Metabolic engineering for the production of prenylated polyphenols in transgenic legume plants using bacterial and plant prenyltransferases. Metab Eng 13:629–637.

Tada H, Murakami Y, Omoto T, et al (1996) Rosmarinic acid and related phenolics in hairy root cultures of *Ocimum basilicum*. Phytochemistry 42:431–434.

Tamura M, Tsuji Y, Kusunose T, et al (2014) Successful expression of a novel bacterial gene for pinoresinol reductase and its effect on lignan biosynthesis in transgenic *Arabidopsis thaliana*. Appl Microbiol Biotechnol 98:8165–8177.

Tattersall DB, Bak S, Jones PR, et al (2001) Resistance to an herbivore through engineered cyanogenic glucoside synthesis. Science 293:1826–1828.

Timm S, Florian A, Jahnke K, et al (2011) The hydroxypyruvate-reducing system in *Arabidopsis*: multiple enzymes for the same end. Plant Physiol 155:694–705.

Timm S, Nunes-Nesi A, Pärnik T, et al (2008) A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in *Arabidopsis*. Plant Cell 20:2848–2859.

Treutter D (2006) Significance of flavonoids in plant resistance: a review. Environ Chem Lett 4:147–157.

Tsuruga A, Terasaka K, Kamiya K, et al (2006) Elicitor-induced activity of isorinic acid 3'-hydroxylase, an enzyme catalyzing the final step of rosmarinic acid biosynthesis in *Lithospermum erythrorhizon* cell suspension cultures. Plant Biotechnol 23:297–301.

Valvekens D, Van Montagu M, Van Lijsebettens M (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. Proc Natl Acad Sci USA 85:5536–5540.

van der Hoorn RAL, Leeuwenburgh MA, Bogyo M, et al (2004) Activity profiling of papain-like cysteine proteases in plants. Plant Physiol 135:1170–1178.

Vanholme R, Cesarino I, Rataj K, et al (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. Science 341:1103–1106.

Vaucheret H, Béclin C, Fagard M (2001) Post-transcriptional gene silencing in plants. J Cell Sci 114:3083–3091.

Verpoorte R, Contin A, Memelink J (2002) Biotechnology for the production of plant secondary metabolites. Phytochem Rev 1:13–25.

Verpoorte R, van der Heijden R, Hoge JH, ten Hoopen HJ (1994) Plant cell biotechnology for the production of secondary metabolites. Pure Appl Chem 66:2307–2310.

Verpoorte R, van der Heijden R, Memelink J (2000) Engineering the plant cell factory for secondary metabolite production. Transgenic Res 9:323–343.

Walker JM (2002) SDS polyacrylamide gel electrophoresis of proteins. In: Walker JM (ed) The protein protocols handbook, 2nd ed. Humana Press, Totowa-New Jersey.

Walsh CT, Garneau-Tsodikova S, Gatto GJ (2005) Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chemie Int Ed 44:7342–7372.

Wanner LA, Li G, Ware D, et al (1995) The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. Plant Mol Biol 27:327–338.

Webster DE, Thomas MC (2012) Post-translational modification of plant-made foreign proteins; glycosylation and beyond. Biotechnol Adv 30:410–418.

Weitzel C, Petersen M (2010) Enzymes of phenylpropanoid metabolism in the important medicinal plant *Melissa officinalis* L. Planta 232:731–742.

Weitzel C, Petersen M (2011) Cloning and characterisation of rosmarinic acid synthase from *Melissa officinalis* L. Phytochemistry 72:572–578.

Wong JH, Namasivayam P, Abdullah MP (2012) The PAL2 promoter activities in relation to structural development and adaptation in *Arabidopsis thaliana*. Planta 235:267–277.

Xiao Y, Zhang L, Gao SH, et al (2011) The *c4h, tat, hppr* and *hppd* genes prompted engineering of rosmarinic acid biosynthetic pathway in *Salvia miltiorrhiza* hairy root cultures. PLoS One 6:e29713.

Yamamura Y, Ogihara Y, Mizukami H (2001) Cinnamic acid 4-hydroxylase from *Lithospermum erythrorhizon*: cDNA cloning and gene expression. Plant Cell Rep 20:655–662.

Yang L, Mercke P, van Loon JJA, et al (2008) Expression in *Arabidopsis* of a strawberry linalool synthase gene under the control of the inducible potato PI2 promoter. Agric Sci China 7:521–534.

Yu XH, Gou JY, Liu CJ (2009) BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. Plant Mol Biol 70:421–442.

Yun CS, Yamamoto T, Nozawa A, Tozawa Y (2008) Expression of parsley flavone synthase I establishes the flavone biosynthetic pathway in *Arabidopsis thaliana*. Biosci Biotechnol Biochem 72:968–973.

Zhang X, Liu CJ (2015) Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids. Mol Plant 8:17–27.

8. Appendix

8.1 Materials and instruments

8.1.1 Materials

8.1.1.1 Chemical materials

Acetic acid, p.a.	Roth
Acrylamide, Rotiphorese Gel 30	Roth
Adenosine 5`-triphosphate, Na-salt (ATP)	Biomol
Agar	C.E. Roeper
Agarose	Biozym
Ammoniumperoxidisulphate (APS)	Merck
Ammonium sulphate	Roth
Ampicillin	Roth
L-(+)-Ascorbic acid	Roth
Boric acid, p.a.	Merck
n-Butanol	Merck
Caffeic acid	Alfa Aesar
Caffeoyl-pHPL	Institute collection (synthesised)
Calcium chloride dihydrate, p.a.	Merck
Cefotaxime	Mannheim Fresenius Kabi
Cetyltrimethylammoniumbromide (CTAB)	Roth
Chloroform, p.a.	Roth
Cinnamic acid	Duchefa
Cobalt(II) chloride	Merck
Coenzyme A trilithium salt	Applichem,
Coomassie Brilliant Blue G-250	Fluka
Coomassie Brilliant Blue R-250	Fluka
Copper(II)sulphate-pentahydrate	Fluka
Diethyldithiocarbaminate (DIECA)	Fluka
4-Coumaric acid	Alfa Aesar
D-Glyceric acid	Fluka
2,4-Dichlorophenoxyacetic acid	Duchefa
1,4-Dithiothreitol (DTT)	Roth
EDTA, Na-salt	Roth
Ethanol, p.a.	Roth
Ethyl acetate	Roth
DL-4-Hydroxyphenyllactate (pHPL)	Sigma-Aldrich
4-Hydroxyphenylpyruvate (pHPP)	Sigma-Aldrich
Ethidium bromide, 1%	Applichem,
Glycin	Merck
Glyoxylate	Fluka
Guanidinthiocyanate	Roth

Indole-3-acetic acid (IES)	Duchefa
Imidazole	Roth
Isopropyl-β-thiocyanate (IPTG)	Roth
Iron(II)sulphate heptahydrate	Roth
Isopropanol	Roth
Kanamycin	Roth
Kaliumdihydrogen phosphate	Roth
di-Kaliumhydrogen phosphate	Roth
Kalium iodide	Merck
Kalium nitrate	Roth
Kinetin	Duchefa
Laurylsarcosine	Sigma-Aldrich
L-Phenylalanine	Roth
L-Tyrosine	Fluka
Magnesium chloride hexahydrate	Roth
Magnesium sulphate heptahydrate	Roth
Mangan(II) sulphate	Duchefa
Methanol	Fisher Scientific
myo-Inositol	Roth
Naphthaleneacetic acid (NAA)	Duchefa
Nicotine amide adenine dinucleotide, reduced (NADH)	Biomol
Nicotine amide adenine dinucleotide phosphate, oxidised (NADP)	Biomol
Nicotine amide adenine dinucleotide phosphate, reduced (NADPH)	Roth
Sodium acetate trihydrate p.a.	Roth
Sodium tetraborate decahydrate	Roth
Sodium chloride, p.a.	Roth
Sodium citrate	Roth
Sodium dihydrogen phosphate	Duchefa
Sodium hydroxide, p.a.	Roth
Sodium molybdate dihydrate	Fluka
Nickel(II) sulphate heptahydrate	Roth
Nicotinic acid	Duchefa
NZ-amines A	Sigma-Aldrich
2-Oxoglutarate, Na-salt	Merck
PD10-column	GE Healthcare
Peptone	Roth
Phenol (citrate buffer-saturated)	Sigma
o-Phosphoric acid, 85%	Sigma-Aldrich
Phenylmethylsulfonyl fluoride (PMSF)	Serva
Polyclar 10	ISP
Protease Inhibitor Cocktail® Broad Range	Roth
Pyridoxal phosphate	Sigma-Aldrich
Pyridoxine-HCl	Duchefa

Pyruvate	Fluka
Rosmarinic acid	Institute collection
Sodium dodecylsulphate (SDS) Pellets	Roth
Spectinomycin	AppliChem
Streptomycin Sulphate	Strepto-Hefa
Sucrose	Aldi
Tetramethylethylenediamine (TEMED)	Roth
Thiamine dichloride	Roth
Trishydroxymethylaminomethan (Tris)	Roth
Triton X-100.	Sigma
Tween 20 pure	Serva
Yeast extract	Roth
β-Hydroxypyruvate	Fluka
Zinc sulphate heptahydrate p.a.	Duchefa

8.1.1.2 Kits for molecular biology

dNTP Mix	Fermentas
GeneRuler™ DNA Ladder Mix	Fermentas
His-select® HF Nickel Affinity Gel	Sigma-Aldrich
Perfecta SYBR® Green Supermix	Quanta Bioscience
qScript™ cDNA Supermix	Quanta Bioscience
RevertAid™ First Strand cDNA Synthesis Kit	Fermentas
Roti® mark	Roth
GoTaq-DNA-Polymerase with buffer (5 U/µl)	Promega
RNase A, 90%	Fluka
Bovine Serum Albumin Fraction V (BSA)	Roth

8.1.2 Instruments

Autoclave	Systec VX-150, Systec GmbH Labor-Systemtechnik, Advantage-Lab AL 02-02-100
Water purifier	Milli RX20 Millipore, Stak Pure
Ice machine	Scotsman AF100
Gel Documentation System	FAS-Digi
Gel Electrophoresis System for Agarose	Agagel Mini Biometra, Biomed Anaytik GmbH, VWR Scientific
Gel Electrophoresis System for SDS-PAGE	Mini-Protean® 3 Cell, BioRad
HPLC	L-6000 Pump Merck-Hitachi D-2500 Chromato-Integrator, Merck-Hitachi
PCR-Cycler	Mastercycler Gradient, Eppendorf
Real-Time PCR System	Pikoreal 96, Thermo Scientific

pH-Meter	Accumet Basic, Fischer Scientific
Ultrasonic bath	Sonorex Super RK 510H
Plant Cell Homogeniser	Ultra Turrax T25 Basic, IKA
Rotary Vacuum Evaporator	Rotovapor RE 120, Büchi
Ultrasonic processor	UP200s, Dr. Hielscher GmbH
Rotary shaker for plant cell culture	Certomat® BS-1, B. Braun Biotech International HT (small), Infors AG RS-306, Infors AG
Gel Shaker	Duomax 1030
Shaker Incubator	Gallenkamp
Magnetic stirrer	MR3001 und MR1000, Heidolph
Microwave	Bosch
Thermomixer comfort	Eppendorf
Centrifuge	Eppendorf Centrifuge 5415D, Eppendorf Biofuge 17RS Heraeus Sepatech SIGMA 3K30S B. Braun Biotech Sorvall® RC 50 plus DuPont Mini centrifuge
Spectrophotometer	Specord 200 plus, analytikjena Bio Photometer, Eppendorf AG
Analytical Balance	Mettler H64, Mettler-Toledo GmbH Kern EW, Kern & Sohn GmbH PT310 Sartorius, Sartorius AG
Portable UV-lamp	Bachofer GmbH
Sterile Cabinet	Class 100, Gelaire, Gelman Instrument Opera
Transilluminator	TFX -35, Vilber Lourmat
Vacuum Freeze Dryer	Christ L1, B.Braun Biotech
Water bath	Thermomix® Me, B.Braun
Vortex	Vortex-Genie 2, Scientific Industries

8.2 Media, reagents and buffers

8.2.1 Media

8.2.1.1 CBOH and CB2 medium

Both, CBOH and CB2 media are B5-based media (Gamborg et al. 1968; Petersen and Alfermann 1988). CBOH does not contain hormones and NZ-Amine whereas CB2 does.

Composition	1l of Stock Solution (100X)	Volume for 1 l medium
CB-Macroelements		
KNO_3	250 g	10 ml
$MgSO_4$ x 7 H_2O	25 g	10 ml
NaH_2PO_4 x H_2O	17.2 g	10 ml
$CaCl_2$ x 2 H_2O	15 g	10 ml
$(NH_4)_2SO_4$	13.4 g	10 m
$FeSO_4$ x 7 H_2O with	2.56 g	10 ml
NaEDTA x 2 H_2O	3.43 g	
CB Microelements		10 ml
H_3BO_3	200 mg	
$ZnSO_4$ x7 H_2O	300 mg	
$MnSO_4$ x H_2O	100 mg	
KI	75 mg	
Na_2MoO_4 x 2 H_2O	25 mg	
$CuSO_4$ x $5H_2O$	25 mg	
$CoCl_2$ x $6H_2O$	25 mg	
CB-Vitamins		1 ml
Thiamine dichloride	1 g	
Pyridoxine-HCl	100 mg	
Nicotinic acid	100 mg	
Hormones		
Kinetin	20 mg	10 ml
IES	50 mg	10 ml
2,4-D	200 mg	10 ml
NAA	50 mg	10 ml
Myo-inositol		100 mg
Sucrose		20 g
NZ-Amine		2 g

Adjust pH to 5.5 with 0.5 N HCl

For solid medium, 0.8% (w/v) agar was added

8.2.1.2 LB medium

Preparation of 1 l LB medium (Bertani, 1951)

10 g Tryptone
5 g Yeast extract
10 g NaCl

Adjust pH to 7.0 with 0.5 N NaOH

For solid medium, 1% (w/v) agar was added.

8.2.1.3 YMB medium

Preparation of 1 l YMB medium (Singleton et al. 2002)

0.5 g KH_2PO_4
0.2 g $MgSO_4$ x $7H_2O$
0.1 g NaCl
10 g Mannitol
0.5 g Yeast Extract

Adjust pH to 7 with 0.5 N HCl

Adjust to 1 l with distilled water

For solid medium, 1% (w/v) agar was added.

8.2.2 Reagents and buffers

8.2.2.1 Bradford solution

Preparation of 1 l Bradford solution (Bradford 1976)

100 mg Coomassie Brilliant Blue G250
50 ml 95% ethanol p.a.
100 ml 85% ortho-phosphoric acid

Adjust to 1 l with distilled water and filter twice before use

8.2.2.2 CTAB buffers

2X CTAB-buffer		10X CTAB-buffer	
CTAB	2% (w/v)	CTAB	10% (w/v)
Tris/HCl, pH 8.0	100 mM	NaCl	0.7 M
EDTA, pH 8.0	29 mM		
NaCl	1.4 M		
Polyvinylpyrrolidone (MW 40000)	1% (w/v)		

High Salt TE buffer		CTAB-precipitation-buffer	
Tris/HCl, pH 8.0	10 mM	CTAB	1% (w/v)
EDTA, pH 8.0	1 mM	Tris/HCl, pH 8.0	50 mM
NaCl	1 M	EDTA, pH 8.0	10 mM

8.2.2.3 TBE buffer

Tris	45 mM
Boric acid	45 mM
Na-EDTA	1.5 mM

8.2.2.4 TE buffer

Tris/HCl, pH 7.5	10 mM
EDTA	1 mM

8.2.2.5 Boric acid/borax buffer

Preparation of 500 ml boric acid/borax buffer pH 8.8

0.078 mol/l H_3BO_3 (MW 61.83 g/mol) = 2.4114 g
0.03 mol/l $Na_2B_4O_7$ (MW 381.37 g/mol) = 5.7206 g
0.02 mol/ NaCl (MW 58.5 g/mol) = 0.585 g
Adjust to 0.5 l with distilled water

8.2.2.6 Potassium phosphate buffer

Preparation of 1 L of 0.1 M potassium phosphate (KPi) buffer

Solution A: 500 ml of 1 M K_2HPO_4 (MW 174.18 g/mol) = 87.09 g

Solution B: 500 ml of 1 M KH_2PO_4 (MW 136.09 g/mol) = 68.045 g

pH	Volume of solution A [ml]	Volume of solution B [ml]
7.0	61.5	38.5
7.5	83.4	16.6

The combined 1 M stock solution was diluted to 1 l with distilled water.

8.2.2.7 Tris/HCl buffer

Preparation of 100 ml of 1 M Tris/HCl

1 M Tris (MW 121.14 g/mol) = 12.114 g
Adjust to 100 ml with distilled water
Adjust pH to 7.5 or pH 9.0 with HCl

8.3 HPLC chromatogram of enzyme assays' products

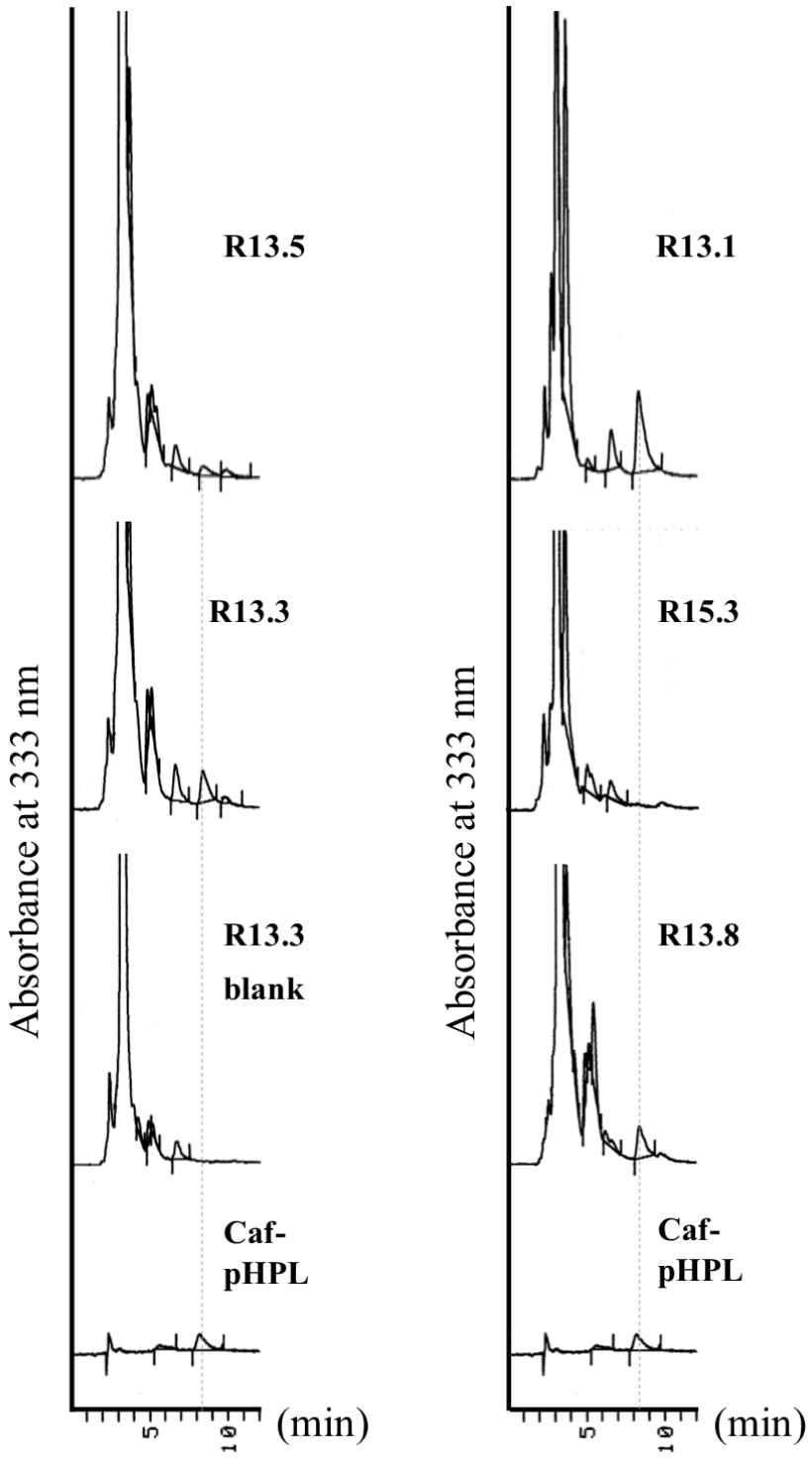


Figure 36. HPLC chromatogram of RAS assay products from the established cell suspension cultures of R-lines. The assay was performed for 15 min and 0 min for blank. Caf-pHPL was used as standard.

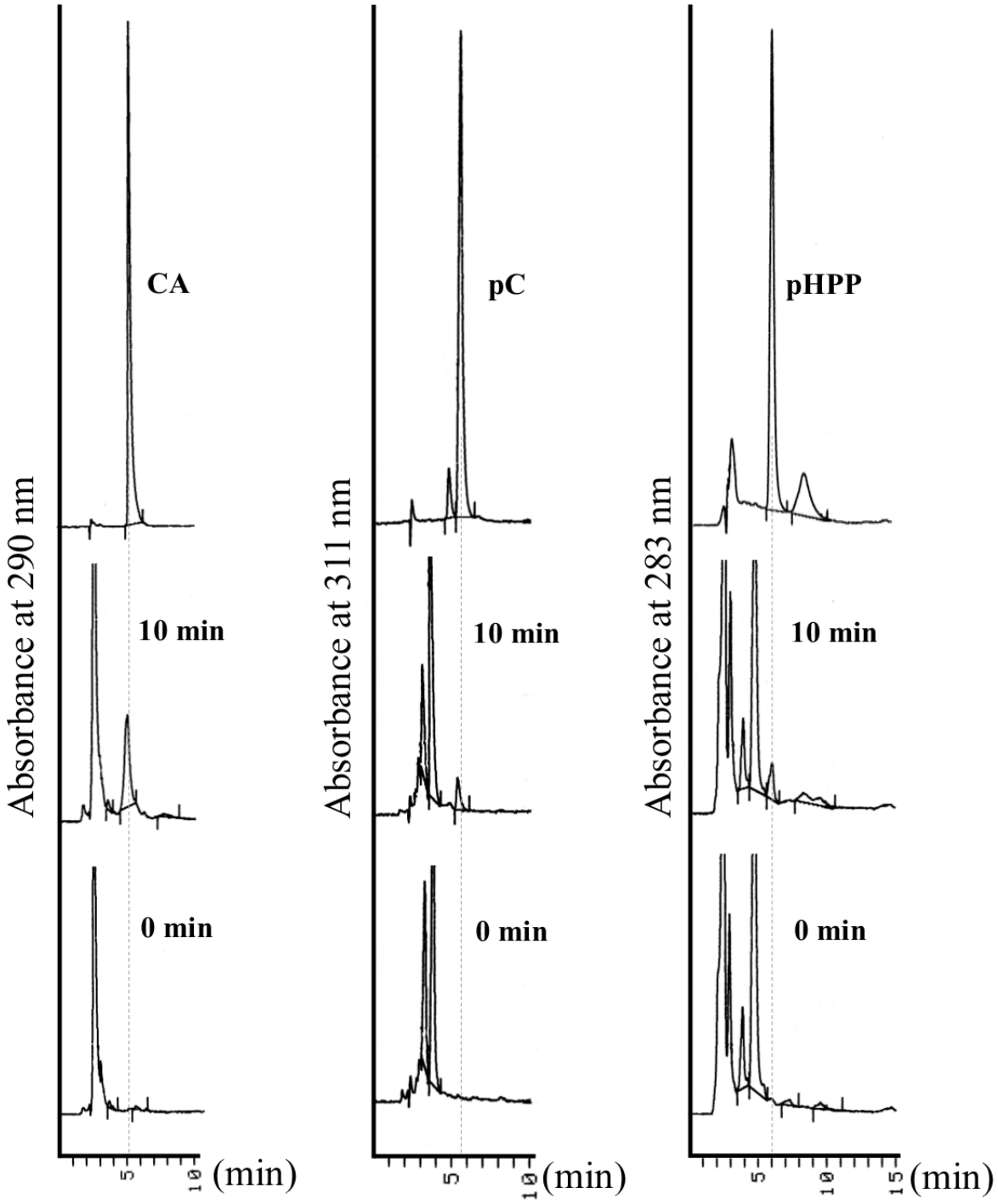


Figure 37. HPLC chromatogram of products from PAL (left), C4H (middle), and TAT (right) assays. Standards of CA (*trans*-cinnamic acid), pC (4-coumaric acid), and pHPP (4-hydroxyphenylpyruvic acid) were used at a concentration of 25, 25, and 100 μM, respectively.

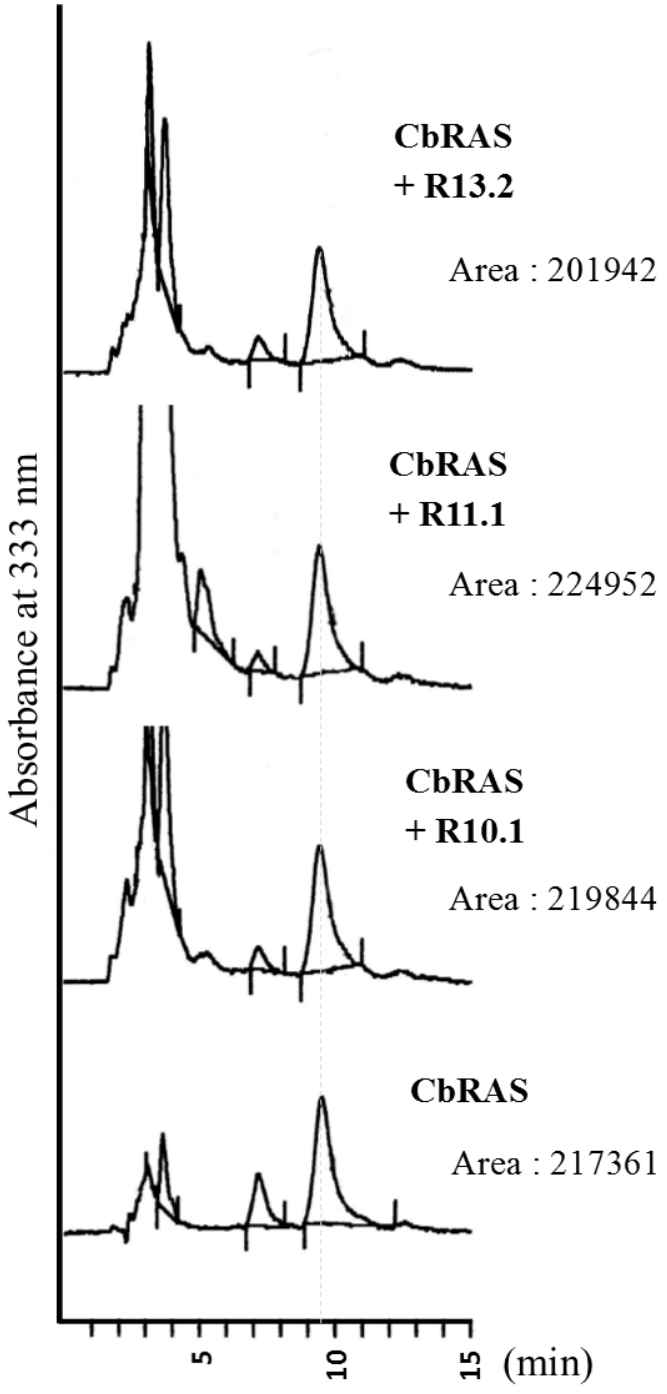


Figure 38. HPLC chromatogram of RAS assay products from the protein mixtures of the recombinant CbRAS and the protein extract of established cell suspension cultures of R-lines. The protein mixtures were incubated 10 min prior to assays. Area values show the quantitative different of the peak intensity of the product (Caf-pHPL) between protein mixtures.

8.4 HPLC and EIC profiles of metabolite contents

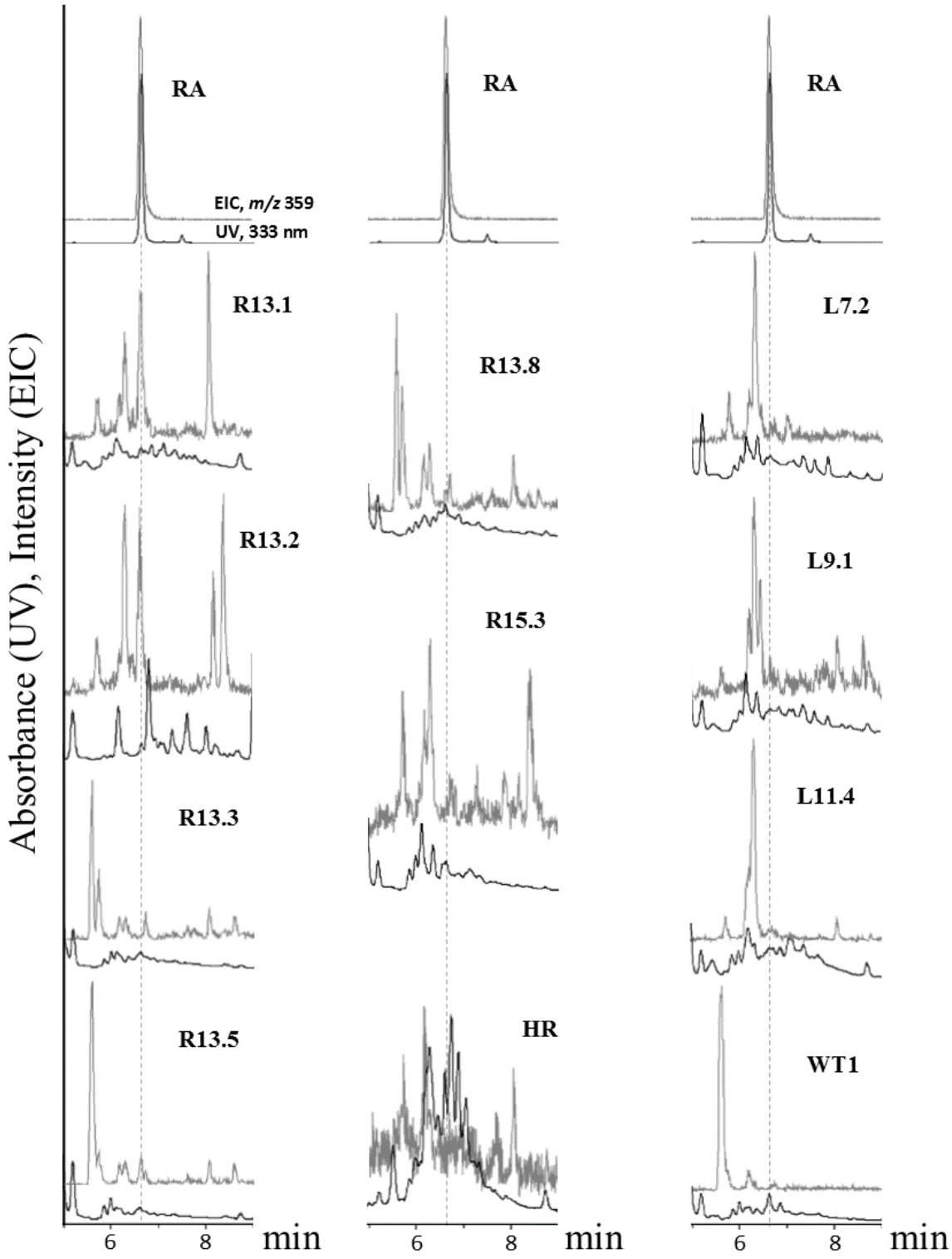


Figure 39. HPLC (black) and extracted ion chromatogram (EIC, grey) profiles of the RA content of all established cell lines.

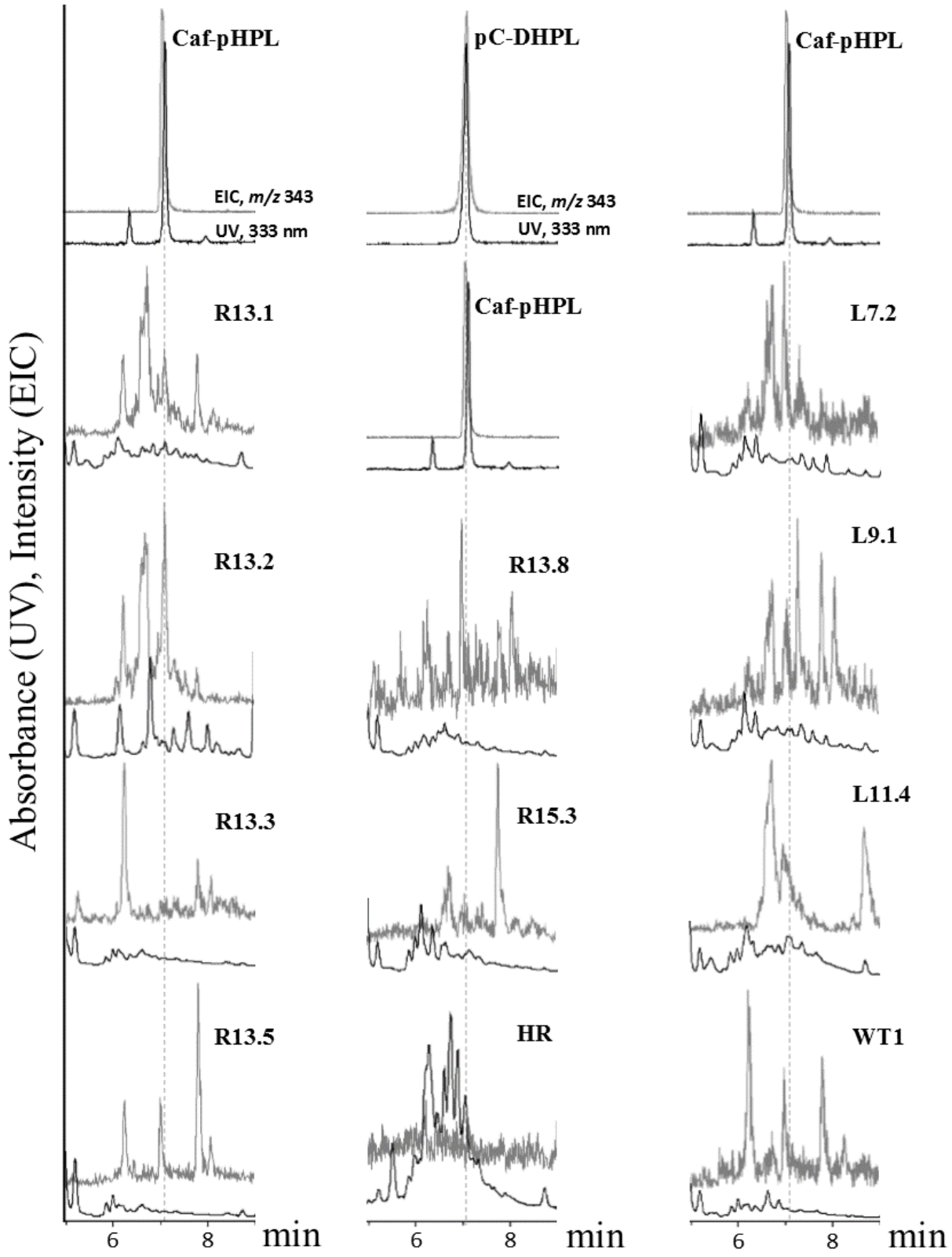


Figure 40. HPLC (black) and extracted ion chromatogram (EIC, grey) profiles of the Caf-pHPL and pC-DHPL contents from all established cell lines.

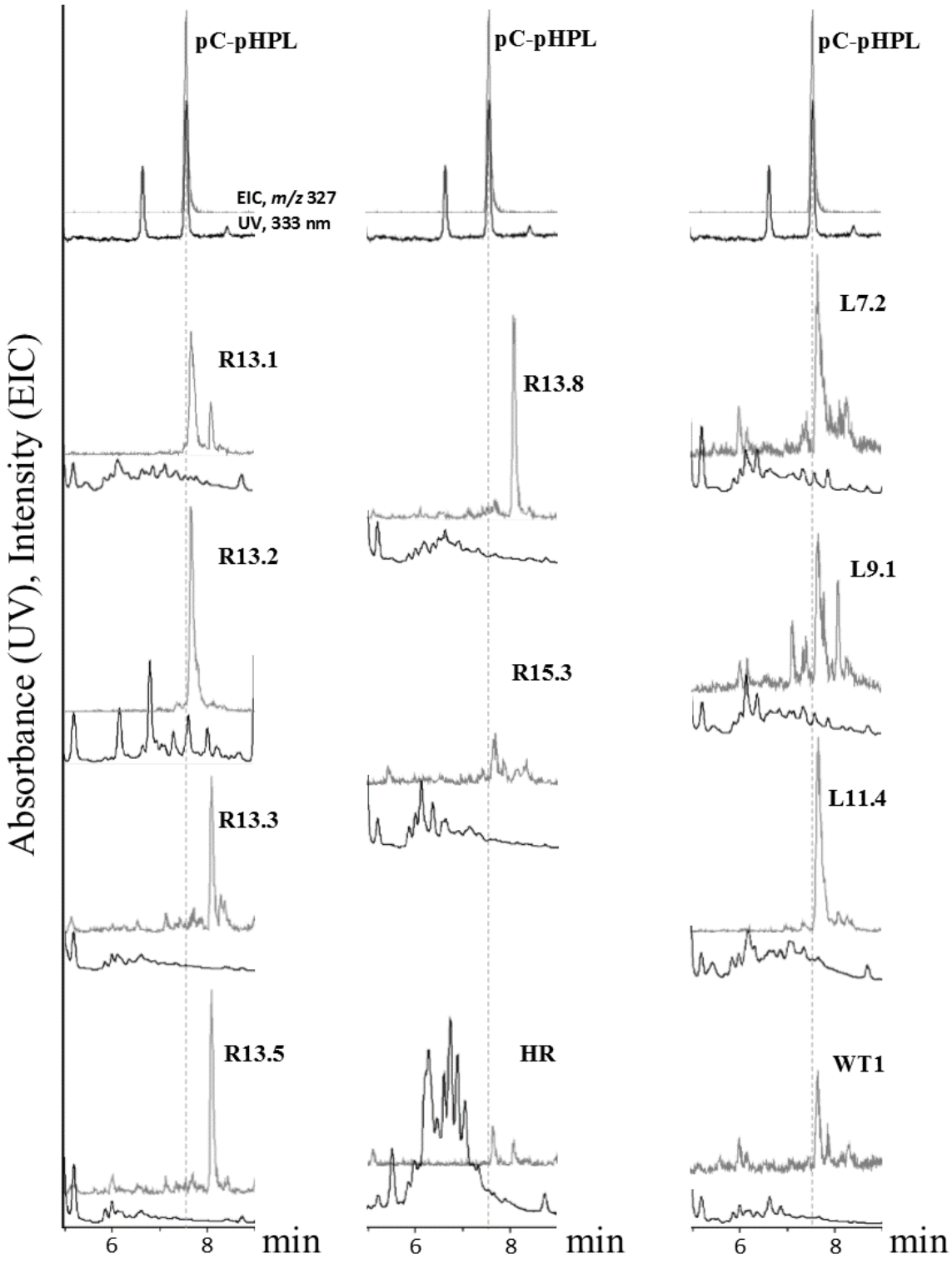


Figure 41. HPLC (black) and extracted ion chromatogram (EIC, grey) profiles of the pC-pHPL contents from all established cell lines.

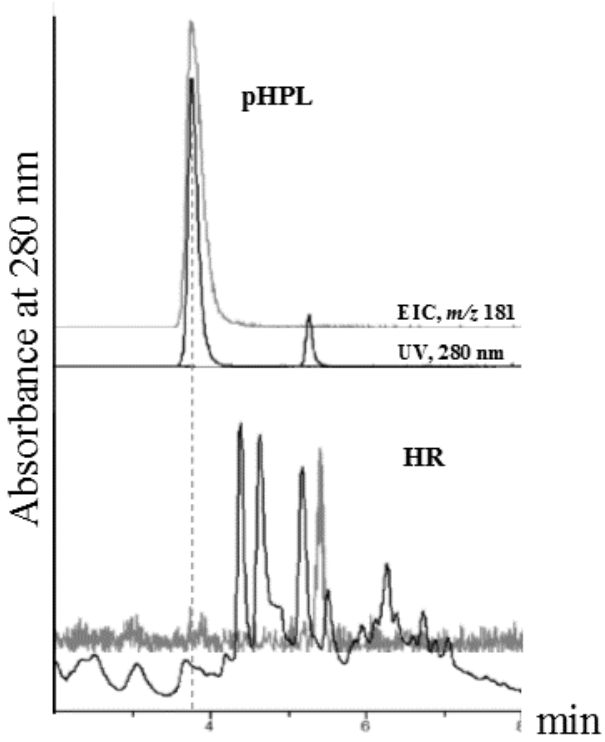


Figure 42. HPLC (black) and extracted ion chromatogram (EIC, grey) profiles of the pHPL contents from the established hairy root cultures.

ERKLÄRUNG

Ich versichere, dass ich meine Dissertation

„Biochemical and molecular investigations of *Arabidopsis thaliana* transformed with genes of rosmarinic acid biosynthesis“

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen sind Zitate als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 12.05.2017

Agus Chahyadi

Curriculum vitae

Personal data

Full name:	Agus Chahyadi
Place of birth:	Lambuya
Date of birth	August 6th, 1986
Nationality:	Indonesian

Educational background

2001-2004:	Senior high school, SMAN 4 Kendari, Indonesia.
2004-2008:	BSc in chemistry, Faculty of mathematics and natural sciences, Universitas Halu Oleo (UHO), Kendari-Indonesia.
2009-2011:	Master of science, School of Pharmacy, Institut Teknologi Bandung (ITB), Bandung-Indonesia.
2014-2017:	Doctor in natural sciences, Institute for Pharmaceutical Biology and Biotechnology, Philipps-Universität Marburg, Germany.
	Financial support by The Directorate General of Higher Education (DGHE), Ministry of Research, Technology and Higher Education of the Republic of Indonesia.

Employment record

2008-2009:	Academic assistant at the Department of Chemistry, Universitas Halu Oleo (UHO), Kendari-Indonesia.
2011-2013:	Research assistant at the Division of Pharmaceutical Biology, School of Pharmacy, Institut Teknologi Bandung (ITB), Bandung-Indonesia.
since 2013:	Academic assistant at the Faculty of Pharmacy, Universitas Halu Oleo (UHO), Kendari-Indonesia.
since 2016:	Founder of two startups (Wegronomy and Filosofood), Indonesia

Marburg, 12.05.2017

Agus Chahyadi

Publications

Chahyadi A , Petersen M (2014) Can *Arabidopsis thaliana* synthesize rosmarinic acid? 1st Eduard Strasburger-Workshop 2014 of the German Botanical Society, Seminar/Workshop der Sektion „Pflanzliche Naturstoffe“ der Deutschen Botanischen Gesellschaft, Nürnberg. (Oral)

Chahyadi A, Petersen M (2014) Engineering of *Arabidopsis thaliana* with rosmarinic acid biosynthesis genes. International Autumn School on Synthetic Biology “BIOLOGY FEAT. ENGINEERING HEIDELBERG, JÜLLICH, MARBURG”, November 17th – 28th, 2014. (Poster)

Chahyadi A, Petersen M (2015) Engineering of *Arabidopsis thaliana* with rosmarinic acid synthase gene. Botanikertagung 2015: from molecules to the field, Freising. (Poster; rewarded with a poster prize)

Chahyadi A, Petersen M (2016) Engineering of *Arabidopsis thaliana* with rosmarinic acid biosynthesis genes. Meeting of the section “Natural Products”, Deutsche Botanische Gesellschaft, Meisdorf. (Oral)

Acknowledgements

Alhamdulillah, finally this Dissertation is finished. I would like to express my gratitude to all of you, especially from the members of the Pharmaceutical for Biology and Biotechnology, Philipps-Universität Marburg. Your support does not only include the scientific and academic aspects, but also the social life during the years of my stay in the lovely small city, Marburg, „klein aber schön". I will never forget you and this nice city.

I truly acknowledge the Directorate General of Higher Education (DGHE) of Indonesia for the financial support during my doctoral study. I also never forget one of the scholarships from DGHE to enrol German course in Indonesia two years before I came here. It really helped me to adapt the language and the culture quickly.

I wish to thank my supervisor Prof. Dr. Maike Petersen, who accepted me to do research for my doctoral degree. I always remember our first acquaintance and engagement during your visit in Bandung. It was the first time I practiced my German, even though my German was bad but you let me keep talking. Finally, we ended up here. I will never forget your friendliness, patience and guidance during my PhD period. You have always tried to help me to solve the problem I had. „Du bist meine Mutter hier". I will miss you when I go back to home.

I would like to express my sincere appreciation to Prof. Dr. Alfred Batschauer as my co-supervisor. Parts of this research will not be done without your contribution. To Prof. Dr. Jens Kockskämper and Prof. Cornelia M. Keck, I am greatly thankful for your participation as auditors when I defended my dissertation. Also, I thank to Dr. Katja Backhaus, who wanted to be the protocol during my defence.

Many thanks, I want to address to my colleagues, Elke Bauerbach, Dr. Soheil Pezehski, Jennifer Robinson, Lennart Poppe, Victoria Werner, Julia Wohl, Thang Son Ta, Tobias Busch, Marie-Claire Badouin and Olga Haag. I am really grateful working together with all of you. I also thank to the Prof. Li's group for togetherness and a nice atmosphere.

I address my warmest gratitude to my Indonesian colleagues in several cities in Germany. It was nice to have you all here. Our brotherhood makes this country feels like home.

My best gratitude goes to my parents, my brothers and my sisters for support in many ways, praying and a lot of other things you spent to help me and to strengthen me.

Finally I come to address my appreciation to my wife, Nur Sulviyana. It is hard to find words to express how lucky I am to have you. You are stronger than I imagined. You are always supporting me and you are always praying for my sake. Yes, we were separated 6,844 miles away and I am not even able to say sorry. I hope, it was the best decision that we made at the moment. I believe, every struggle requires sacrifice. Someday we will see a bright future for us.

I am sure that some people are missing to be mentioned. I would like to thank all of you.

Thank you very much!
Vielen Dank!
Terimakasih banyak!

Agus Chahyadi