Metabolic changes in Arabidopsis thaliana plants overexperssing chalcone synthase

Dao, T.H.H.

Chapter 6

Metabolic profiling of *Arabidopsis thaliana* transformed with a heterologous *chs* cDNA from *Cannabis sativa*

T.T.H. Dao[1, 2], H.K. Kim[1], H.J.M. Linthorst [3], Y. H. Choi[1] and R. Verpoorte[1]

[1] Section Metabolomics, Institute of Biology, Leiden University, Leiden, The Netherlands
[2] Traditional Pharmacy Department, Hanoi Pharmacy University, Hanoi, Vietnam
[3] Section Plant Cell Physiology, Institute of Biology, Leiden University, Leiden, The Netherlands

Abstract

Nuclear Magnetic Resonance (NMR) and Multivariate Data Analysis (PDA) are important analytical tools for macroscopic profiling of metabolomes. This study describes the use of this approach to measure the metabolome of transgenic *Arabidopsis thaliana* plants with a high expression level of a heterologous chalcone synthase gene. Five transgenic *Arabidopsis* lines were analyzed in this study. The Partial least square-Discriminant Analysis (PLS-DA) showed a very good separation between the transgenic plants and controls. This analysis indicated that the level of sugars, flavonoids and phenylpropanoids are higher in the CHS transgenic plants than in control plants. These results show that chalcone synthase overexpression affects both plant secondary metabolism and primary metabolism.

Keywords: chalcone synthase, flavonoids, phenylpropanoids, *Arabidopsis*, NMR, multivariate data analysis.

6.1. Introduction

We can look at what genes are being expressed, and what proteins are present, but what are the end products in the form of cellular functions? Metabolomics attempts to answer this question. Linking functional metabolomic information to mRNA and protein expression data makes it possible to visualize the functional genomic repertoire of an organism. Metabolomics is becoming a widely used technology to evaluate global metabolite levels. In the context of functional genomics, the non-targeted profiling of metabolites in biological samples is now regarded as a viable counterpart to protein and transcript profiling technologies.

Metabolomics is the study of all the metabolites of a biological sample. Several analytical tools such as gas chromatography (GC), liquid chromatography (LC), mass spectrometry (MS) and nuclear magnetic resonance (NMR) have been used to profile the metabolome. None of these is able to give a comprehensive view of the complete metabolome. Issues as dynamic range, polarity, volatility, and stability are in fact limiting factors in obtaining a complete picture of all metabolites. A combination of the various methods should be used to have the most comprehensive view on the metabolome. However, with each single analytical method a wealth of information about the metabolome of an organism can already be obtained under certain conditions.

Nuclear magnetic resonance spectroscopy (NMR) has been considered as a powerful platform in metabolomics because of its ease of sample preparation, short time of analysis and as the only method that allows direct absolute quantitation of all metabolites. Also with ^{1}H-NMR analysis we can detect and do structure elucidation of various metabolites in the sample. However, most metabolites have many signals in their ^{1}H-NMR spectra, thus overlapping may cause a problem in identifying individual metabolites. 2D NMR or chromatographic separation can be used to overcome this problem.

Transgenic or mutant plants in combination with metabolomics provide an excellent means to look at changes in metabolic networks through the specific perturbation of a gene of interest. Plant extracts are very complex in composition and, if many samples are examined, it is difficult to make a meaningful comparison of large numbers of spectra or chromatograms ‘by eye.’ Thus multivariate statistical methods can be

extremely useful, as they are able to compress data into a more easily managable form, allowing the visualiation of the relation between samples.

In this study we used NMR spectroscopy and multivariate data analysis to monitor metabolome changes in transgenic CHS overexpressing *Arabidopsis thaliana* Col 0. Various metabolites were detected and their NMR signals were assigned. Finally, the differences between metabolites levels in transgenic plants and controls were determined.

The first committed step in the biosynthesis of flavonoids is catalysed by the enzyme chalcone synthase (CHS), resulting in a yellow coloured chalcone. In the majority of plants chalcones are not end-products, but intermediates in the pathway proceeding with several further enzymatic steps to other classes of flavonoids, such as flavanones, dihydroflavonols and finally to the anthocyanins, the major water-soluble pigments in flowers and fruits. Other flavonoid classes (i.e. isoflavones, aurones, flavones, pro-anthocyanidins (PA) and flavonols) represent side branches of the flavonoid pathway and are derived from intermediates in the anthocyanin formation. Little is known about the effect of CHS in the total plant metabolomic network, e.g. leading to lignan- and phenylpropanoid derivatives.

Prior to the present study, we transformed a *chs*-cDNA gene from *Cannabis sativa* into *A. thaliana.* The introduction of the first gene in the flavonoid biosynthesis pathway may alter directly or indirectly the level of other metabolites present in *A. thaliana*. This study aimed at the identification of metabolomic pathways that are affected by the overexpression of the *CHS* gene in this plant. ^{1}H-NMR based metabolomics coupled with multivariate data analysis was applied to distinguish between control and CHS transgenic *A. thaliana* plants (ACS).

6.2. Materials and methods

6.2.1. Growth of plant materials

Arabidopsis thaliana ecotype Columbia (Col.0) seeds were obtained from the section Plant Cell Physiology (IBL, Leiden Universiy, The Netherlands) and 5 transgenic ACS lines (ACS 1, 2, 14, 20, 21) were generated as described in **Chapter 3**. Seeds were surface sterilized by incubation for 1 min in 70% ethanol, 15 min in 2.5% sodium hypochlorite and rinsed with sterile water five times. Surface sterilized seeds were

grown on a plate containing half MS medium with 0.6% agar supplemented with 20 mg/L hygromycin for selection of transgenic plants. After keeping 3 days at 4°C in the dark, the seeds were germinated at 23°C under long day condition (16/8 h light/dark cycle) in a growth chamber. All the plant tissues were harvested when the seedlings are 10-days old. Immediately after harvesting, the material was frozen in liquid nitrogen and kept at −80° C until used.

6.2.2. Solvents and chemicals

D_2O (99%) and CH_3OH-d_4 (99.8%) were obtained from Cambridge Isotope Laboratories Inc (Miami, FL, USA). NaOD was purchased from Cortec (Paris, France). Potassium dihydrogen phosphate and trimethylsilane propionic acid sodium salt (TSP) were purchased from Merck (Darmstadt, Germany). As buffering agent 1.232 g KH_2PO_4 and 10 mg TSP (internal standard) were added to 100 g D_2O. Finally, the pH of the solution was adjusted to 6.0 using 1 M NaOD.

6.2.3. Extraction of plant materials

Plants were ground in liquid nitrogen and pooled before subjected to freeze-drying. Twenty-five milligrams of freeze-dried material were transferred to a micro-centrifuge tube before adding 600 µl of CH_3OH-d_4. The mixture was vortexed for 2 min and sonicated for 20 min, followed by centrifugation at 13,000 rpm for 5 min at room temperature. Five hundred microliters of the supernatant were then transferred into 2 ml micro-centrifuge tubes and 250 microliters of KH_2PO_4 buffer, pH 6.0, containing 0.1% trimethyl silyl propionic acid sodium salt (w/v) were added. The mixture was left for 30 min in 4°C and followed by centrifugation at 6000 rpm for 5 min at room temperature. Seven hundred microliters of the supernatant were then transferred into 5 mm NMR tubes for analysis.

6.2.4. NMR measurement

^{1}H-NMR, 2D J-resolved, ^{1}H-^{1}H correlated spectroscopy (COSY), and heteronuclear multiple bonds coherence (HMBC) spectra were recorded at 25 ºC on a 600 MHz Bruker AV 600 spectrometer equipped with cryo-probe operating at a proton NMR frequency of 600.13 MHz. CH_3OH-d_4 was used as the internal lock. Each ^{1}H-NMR

spectrum consisted of 128 scans requiring 10 min acquisition time with the following parameters: 0.25 Hz/point, pulse width (PW) = 30° (10.8 μsec), and relaxation delay (RD) = 1.5 sec. A presaturation sequence was used to suppress the residual H_2O signal with low power selective irradiation at the H_2O frequency during the recycle delay. FIDs were Fourier transformed with LB = 0.3 Hz and the spectra were zerofilled to 32 K points. The resulting spectra were manually phased and baseline corrected, and calibrated to TSP at 0.0 ppm, using Topspin (version 2.1, Bruker).

6.2.5. Data analysis

The spectral intensities were reduced to integrated regions of an equal width of 0.04 ppm (buckets or bins) corresponding to the region of δ 0.3 - δ 10.0. The regions of δ 4.8-δ 4.9 and δ 3.28- 3.40 were excluded from the analysis because of the residual signal of water and CH_3OH. Principal Compnent Analysis (PCA), Partial least square-Discriminant Analysis (PLS-DA) and Hierarchical Clustering Analysis (HCA) were performed with the SIMCA-P software (v. 12.0, Umetrics, Umeå, Sweden).

6.3. Results and discussion

Seedlings of Arabidopsis Col. 0 wild type (WT), transgenic plants with empty vector (E) and 5 CHS transgenic plant lines (ACS1, 2, 14, 20, 21) were subjected to metabolomic analysis by ^{1}H-NMR and 2D-NMR spectroscopy. For the assignments of metabolites in the ^{1}H-NMR spectra is referred to **Chapter 5.** ^{1}H-NMR spectra of ACS 2, ACS 20 and WT are shown in **Figure 6.1**. In the aromatic area, three kaempferol glycosides, quercetin rhamnose, 4 malate conjugated cinamic acid derivatives (hydroxyferulic, caffeic, coumaric, and sinapoyl), sinapoyl glucoside and formic acid were found. Visually in ^{1}H-NMR spectra we can see that all the identified phenolic compounds are significantly higher in ACS 20 compared to ACS 2 which has the same level of phenolic compounds as WT. This is in accordance with the results reported in **Chapter 4** showing that ACS 2 has similar CHS activity as WT.

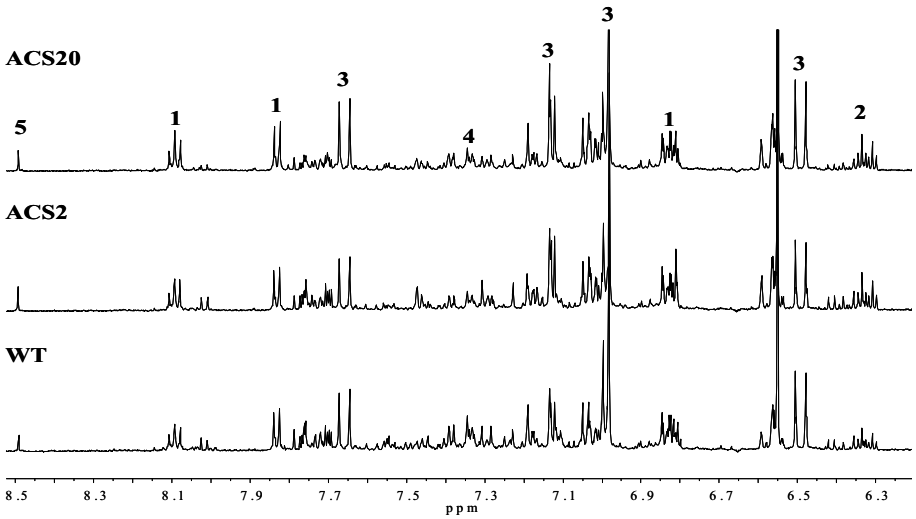

Figure 6.1. Aromatic regional ^{1}H-NMR spectra of Arabidopsis Col.0 (WT) and CHS transgenic Arabidopsis (ACS2, ACS20). 1. Keampferol glycosides (Kaempferol 3,7-*O*- dirhamnopyranoside, Kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*- rhamnopyranoside, Kaempferol 3-*O*-glucopyranoside-7-rhamnopyranoside), 2. *trans*-phenylpropanoids (caffeoyl malate, coumaroyl malate, hydroxyferuoyl malate), 3. *trans*-synapoyl, 4. Quercetin derevative, 5. Formic acid.

In order to distinguish the samples based on the ^{1}H-NMR spectra, multivariate data analysis was applied. Firstly, principal component analysis (PCA) was used to reduce the data set from the ^{1}H-NMR spectra to fewer components enabling to group samples or to do pattern recognition. To remove the biological variation of the samples of CHS transgenic plants and non CHS transgenic plants, a supervised multivariate data analysis, termed Partial Least Square-Discriminant Analysis (PLS-DA), was employed. After the PLS-DA analysis, Hierachical Clustering Analysis (HCA) based on PLS-DA was performed. The results show that E samples are very close to WT and all the samples of transgenic line ACS 2 and two samples of transgenic line ACS 21 are separated from other transgenic lines (**Figure 6.2**). This result fits with the genetic data of the transgenic plants which are reported in **Chapter 3**. The transgenic ACS 2 and ACS 21 have 5 copies of the *CHS* transgene. Since the E group shows similar metabolites profiles as WT group, it was not further studied.

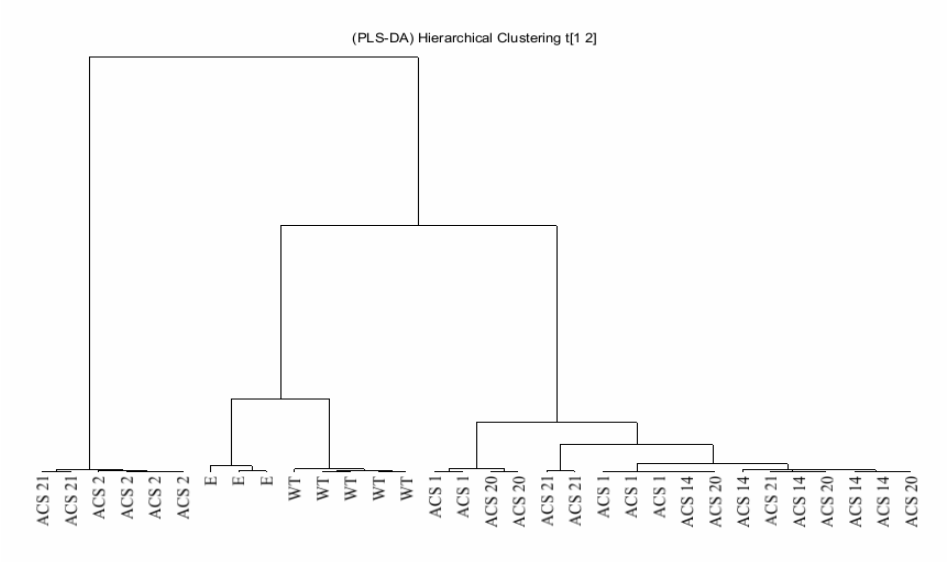

Figure 6.2. Hierarchical Clustering Analysis (HCA) based on PLS-DA of transgenic and control *Arabidopsis* plants. E: transgenic Arabidopsis with empty vector, WT: wild type *A. thaliana* Col.0; ACS 1….ACS 21: CHS transgenic *Arabidopsis* plants

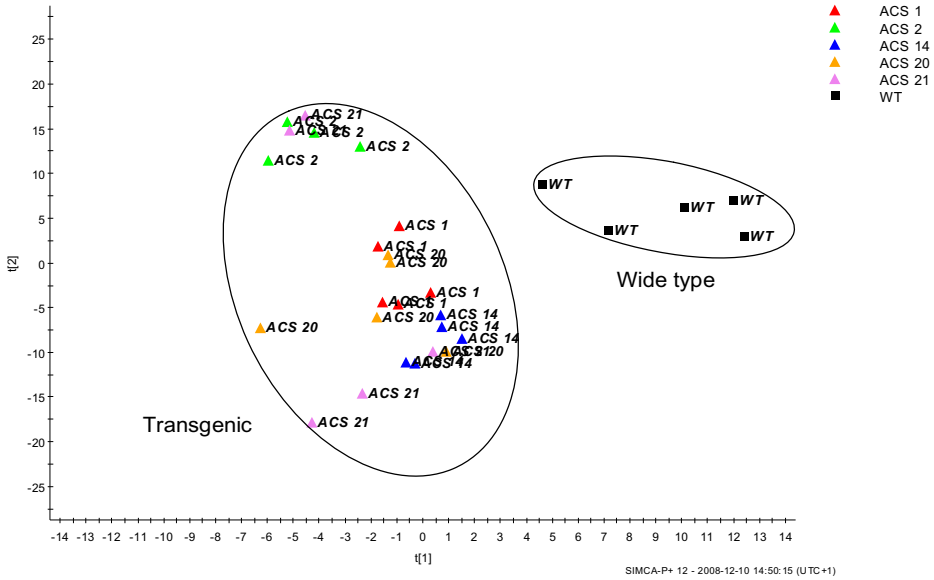

Figure 6.3. The PLS-DA score scatter plot of CHS transgenic (ACS 1-21) and wild type (WT) *Arabidopsis* Col. 0 plants.

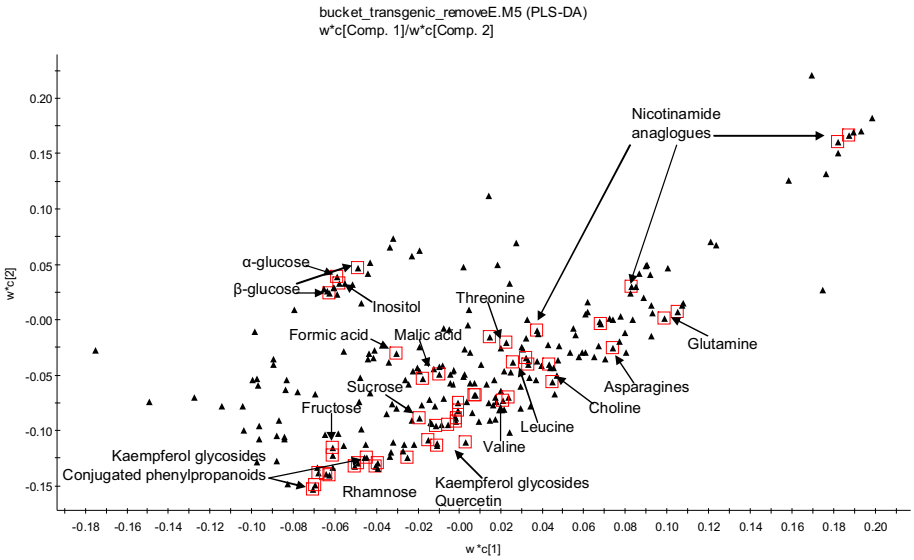

Figure 6.4. The PLS-DA loading scatter plot of transgenic (ACS 1-21) and wild type (WT) *Arabidopsis* Col. 0

When PLS-DA was applied, all the transgenic lines and WT plants could be clearly observed as separate clusters in the PLS-DA score scatter plot in PC 2 (**Figure 6.3**). Most of the transgenic plants are located on the negative side of the PLS PC 2 axis while all WT plants are on the positive side. To find out which metabolites contribute to the discrimination between the transgenic plants, a PLS-DA loading scatter plot (w*c) was analysed as shown in **Figure 6.4**. Positive values of w*c[1] are seen for the variable associated with WT plants whereas negative values are associated with transgenic plants. Examination of the loading plot shows that the metabolites strongly contributing to the separation of the WT plants were amino acids such as threonine (δ 1.32), alanine (δ 1.48), leucine (δ 0.96), asparagine (δ 2.8 and 2.97) and glutamine (δ 2.12 and 2.48). For the transgenic plants, the responsible loading plots correspond to the signals of sugars (α-glucose δ 5.18, β-glucose δ 4.58, rhamnose δ 5.62, sucrose δ 5.4), flavonoids (kaempferol 3-*O*-glucopyranoside 7-*O*-rhamnopyranoside δ 8.09, kaempferol 3,7-*O*-dirhamnopyranoside δ 7.83, kaempferol 3-O-[rhamnosyl (1–2) D-glucoside]-7-O-rhamnopyranoside δ 8.11, quercetine moiety δ 7.72), phenylpropanoids (*trans*-sinapoyl malate δ 6.48 and 7.66, *trans*-sinapoyl glucose δ 6.49 and 7.77, *trans*-5-hydroxyferuoyl malate, *trans*-caffeoyl malate, *trans*-coumaroyl malate δ 6.32 - 6.37 and 7.54 - 7.59).

These results indicate that sugars, flavonoids and phenylpropanoids are accumulated at higher levels in the transgenic plants than in WT plants, which contain higher levels of amino acid and nicotinamide analogues. It suggests that the presence of the heterogonous *CHS* gene in the transgenic plants influences the flavonoid pathway directly and indirectly, resulting in upregulation of sugar and phenylpropanoid metabolism. A simplified biosynthetic network including the compounds discussed in this chapter is shown in **Figure 6.4**. The phenylpropanoid pathways and secondary metabolic pathways are presented in **Figure 6.5** and **Figure 6.6.** Our results showed that the heterologous *CHS* gene may influence the expression of other genes so a micro array analysis could help to obtain a more complete picture of the effect of expression the *CHS* transgene on plant metabolism.

In plants, sugars have dual functions as nutrition and as important signal molecules [Rolland *et al.,* 2002]. There are several studies indicating that sugars upregulate the flavonoid and anthocyanin biosynthesis pathway strongly. The *CHS* gene derived from petunia (*Petunia hybrida*) petals expressed in transgenic *Arabidopsis* leaves was induced by sugars and *Arabidopsis* grown on a sucrose containing medium showed high levels of anthocyanins [Tsukaya *et al.,* 1991; Ohto *et al.,* 2001]. Here we also found a correlation between high expression of *CHS* and sugars, though now the effect seems inversed, the CHS expression causing increase of sugars.

The increase of flavonoid and phenylpropanoid levels in the transgenic plants suggests that CHS transgenic plants could be a way to achieve crop plants with a higher level of these phenolic compounds. They play an important role in plant development and a plant's interaction with the environment [Rasmussen and Dixon, 1999]. It is reported that flavonoids have antimicrobial activity, and can protect plants from herbivory. Their function as UV protectant in plants is well-known. Moreover important medicinal and nutritional values are described for flavonoids such as antioxidant activity, anti-inflammatory activity and anti-tumor activity [Harborne and William, 2000; Dixon, 2000]

A comparision of metabolites in wild type *Arabidopsis* and 5 CHS transgenic plants (ACS 1, 2, 14, 20, 21) based on the loading scatter plot are presented in **Table 6.1.**

Table 6.1. The metabolites in wild type *Arabidopsis* and transgenics

Compounds	WT	ACS1	ACS2	ACS14	ACS20	ACS21
Amino/organic acids						
Threonine	+	–	–	–	–	–
Alanine	+	–	–	–	–	–
Glutamine	+	–	–	–	–	–
Valine	+	–	–	–	–	–
Leucine	+	–	–	–	–	–
Asparagine	+	–	–	–	–	–
Formic acid	–	+	+	+	+	+
Sugars						
β-glucose	–	+	+	+	+	+
α-glucose	–	+	+	+	+	+
Rhamnose	–	+	+	+	+	+
Fuctose	–	+	+	+	+	+
Sucrose	–	+	+	+	+	+
Phenylpropanoids/Flavonoids						
Kaemferol 3-*O*-glucopyranoside-7-rhamnopyranoside	–	+ +	+	+	+ +	+
Kaemferol 3,7-*O*-dirhamnopyranoside	–	+ +	+	+	+ +	+
kaempferol 3-O- rhamnosyl (1–2) D-glucoside]-7-O-rhamnoside	–	+ +	+	+	+ +	+
Quercetine direvatives	–	+ +	+	+	+ +	+
trans-5-hydroxyferuoyl malate	–	+	+	+	+	+
trans-caffeoyl malate	–	++	+	+	++	+
trans-coumaroyl malate	–	++	+	+	++	+
trans-sinapoyl glucoside	–	+ +	+	+	+ +	+
trans-sinapoyl malate	–	+ +	+	+	+ +	+
Other compounds						
Choline	+	–	–	–	–	–
Nicotinamide analogue	+	–	–	–	–	–
Inositol	–	+	+	+	+	+

Note: – : lower, + : higher, ++: much higher

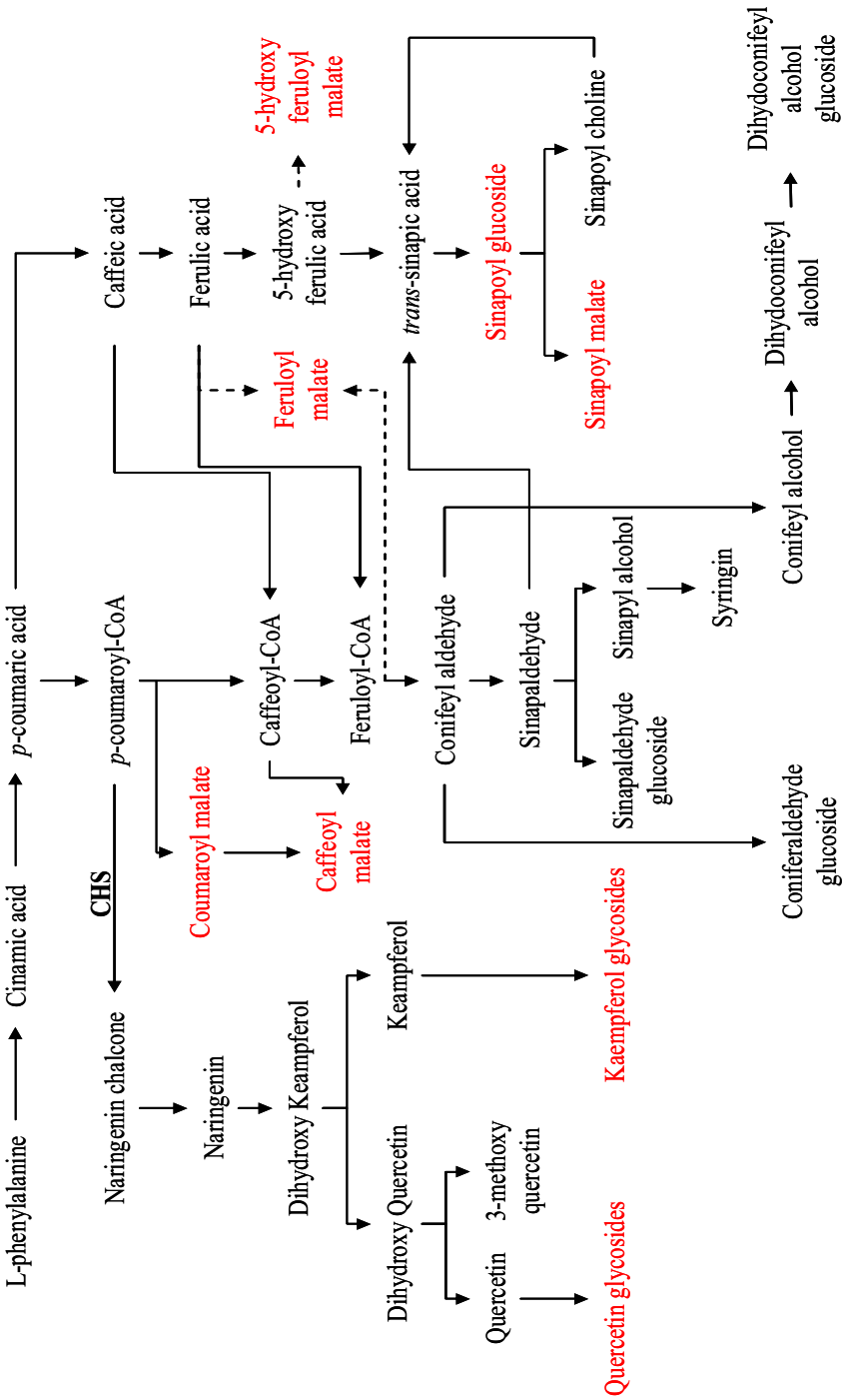

Figure 6.5. Phenylpropanoids biosynthesis pathways in *Arabidopsis* (Reference: *AraCyc* pathway; Sullivan, 2009)
The compounds with red letters are increased in transgenic CHS plants. ——→ known pathway; - - - - - → unknown pathway.

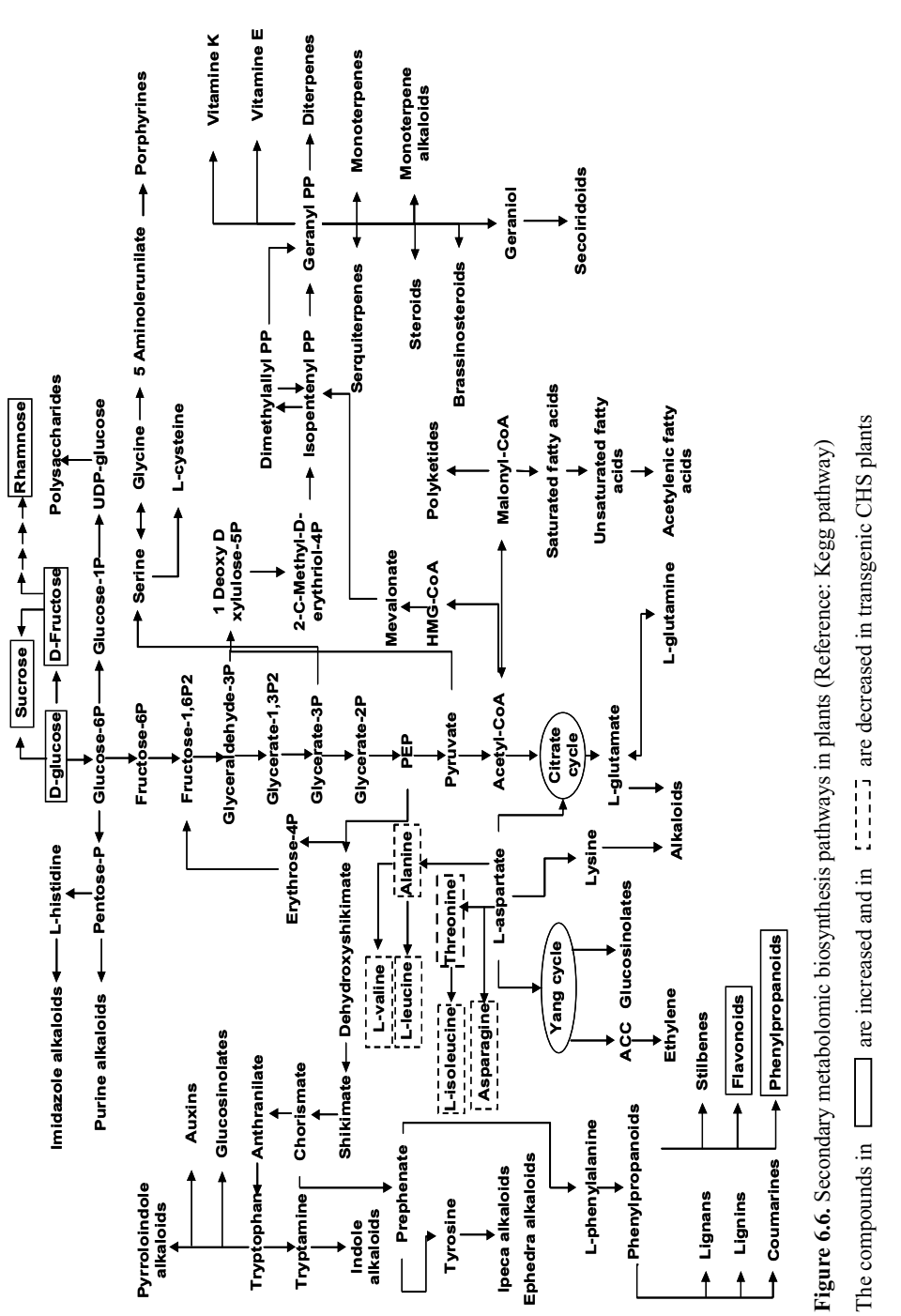

Figure 6.6. Secondary metabolomic biosynthesis pathways in plants (Reference: Kegg pathway)

The compounds in ▭ are increased and in ▭ are decreased in transgenic CHS plants

6.4. Conclusion

This study shows that a specific extraction method focusing on phenolic compounds such as flavonoids and phenylpropanoids and using NMR-based metabolomics coupled with multivariate data analysis is able to visualize metabolome changes in CHS transgenic plants. This analytical method allows identification of a broad range of primary and secondary metabolites in crude samples without any purification steps. The introduction of the heterogonous *CHS* gene in *A. thaliana* has influence on the whole plant metabolism, and not only on the pathway where the transgenic protein ic expressed. The results also show a potential future for CHS transgenic plants, as modifying the flavonoid biosynthesis pathway could improve nutritional or medicinal value of a plant.