Metabolic fingerprinting of wild type and transgenic tobacco plants by 1H NMR and multivariate analysis technique

Choi, H.-K.; Choi, Y.H.; Verberne, M.; Lefeber, A.W.M.; Erkelens, C.; Verpoorte, R.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Phytochemistry 65 (2004) 857–864
PHYTOCHEMISTRY
www.elsevier.com/locate/phytochem

Metabolic fingerprinting of wild type and transgenic tobacco plants by ^{1}H NMR and multivariate analysis technique

Hyung-Kyoon Choi[a,b], Young Hae Choi[a], Marianne Verberne[a], Alfons W.M. Lefeber[c], Cornelis Erkelens[c], Robert Verpoorte[a,*]

[a]*Department of Pharmacognosy, Section of Metabolomics, Institute of Biology, Leiden University, PO Box 9502, 2300 RA Leiden, The Netherlands*
[b]*College of Pharmacy, Chung-Ang University, Seoul 156-756, South Korea*
[c]*Division of NMR, Institute of Chemistry, Gorlaeus Laboratories, PO Box 9502, 2300 RA Leiden, The Netherlands*

Received 16 September 2003; received in revised form 22 January 2004

Abstract

The metabolomic analysis of wild type and constitutive salicylic acid producing tobacco plants (CSA tobacco, *Nicotiana tabacum* 'Samsun' NN) plants overexpressing salicylate biosynthetic genes was carried out by ^{1}H NMR spectrometry and multivariate analysis techniques. The principle component analysis (PCA) of the ^{1}H NMR spectra showed a clear discrimination between those samples by PC1 and PC2. The discrimination of non-inoculated, TMV-virus inoculated, and systemic leaves or veins could also be obtained by PCA analysis. Major peaks in ^{1}H NMR spectra contributing to the discrimination were assigned as those of chlorogenic acid, malic acid, and sugars. This method allows an efficient differentiation between wild type and transgenic plants without any pre-purification steps.

Keywords: Metabolomic analysis; Tobacco; *Nicotiana tabacum*; ^{1}H NMR; Multivariate analysis; Salicylic acid

1. Introduction

In the post genomic era, it is crucial to perform unbiased metabolic analyses and eventually define the biochemical functions of plant primary and secondary metabolic pathways (Tretheway et al., 1999). The term 'metabolome' has been used to describe the observable chemical profile or fingerprint of the metabolites in whole tissues (Ott et al., 2003).

In metabolite profiling, it is preferable to use a wide spectrum of chemical analysis techniques, which are rapid, reproducible, and stable in time while needing only a very basic simple sample preparation. NMR is one of the techniques that potentially meets those demands. On the other hand, the NMR has disadvantages such as relatively low sensitivity and the necessity of expensive instrument to obtain greater resolution and separation of chemical shifts. In spite of its some disadvantages a number of techniques have been devised to develop NMR spectroscopy as a fingerprinting tool for the interpretation and quality assessment of industrial and natural products and multivariate or pattern recognition techniques such as the well-known principal component analysis (PCA) have been specifically designed to analyze complex data sets (Sumner et al., 2003). There were many reports to use NMR and multivariate analysis technique to show differences between samples and elucidate biomarkers in the filed of food science (Vogels et al., 1996; Belton et al., 1998; Gall et al., 2001; Brescia et al., 2002; Charlton et al., 2002; Duarte et al., 2002). In addition, Gavaghan et al. (2000) reported that the NMR-based metabolomic approach could be applied to differentiate between genetic strains in mouse lines. The analysis of the consequences of the genetic manipulation and strain differentiation in strains of yeast was also made with ^{1}H NMR (Raamsdonk et al., 2001). Investigation of metabolite profiling of rat urine using ^{1}H NMR and multivariate analysis was performed also (Bollard et al., 2001). Ward et al. (2003) reported that the various

* Corresponding author. Tel.: +31-71-527-4510; fax: +31-71-527-4511.
E-mail address: verpoort@chem.leidenuniv.nl (R. Verpoorte).

0031-9422/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.phytochem.2004.01.019

ecotypes of *Arabidopsis thaliana* could be distinguished using ^{1}H NMR and multivariate analysis.

The introduction of the *entC* gene from *Escherichia coli* and the *pmsB* gene from *Pseudomonas fluorencens*, encoding two enzymes involved in the bacterial biosynthetic pathway of salicylic acid (SA), into tobacco (*Nicotiana tabacum* 'Samsun' NN) resulted in constitutive salicylic acid producing plants (CSA) (Verberne et al., 2000). Salicylic acid (SA) is known as an important signal compound in systemic acquired resistance (SAR) observed in tobacco after infection of leaves with tobacco mosaic virus (TMV) (Raskin, 1992). Indeed as expected, CSA-tobacco plants showed increased resistance against viral infection (Verberne et al., 2000). Metabolite profiling of the CSA plants using various chromatographic methods was conducted in the fields of phytoalexins (Nugroho et al., 2002a), alkaloids (Nugroho et al., 2002b), and flavonoids (Nugroho et al., 2002c). However, those trials did not use general profiling but a limited and targeted approach on a specific group of compounds and no clear differentiation between wild type and CSA plants was found except for the phenylpropanoid products which were found to be different for CSA and wild type plants (Halim et al., 2003).

In this report, we describe a simple and efficient method to discriminate between wild type and CSA tobacco and to fingerprint their metabolomic response to TMV infection using ^{1}H NMR spectroscopy method coupled with multivariate analysis. In addition, the revelation of the major components contributing the discrimination is performed. This should lead to the identification of metabolic pathways connected with the defense response against TMV-infection.

2. Results and discussion

2.1. Visual inspection of ^{1}H NMR spectra and assignments of the compounds

Little difference was observed between the spectra of the $CHCl_3$ extracts of the various samples (data not shown). The ^{1}H NMR spectra of the aqueous extracts for the wild type and CSA plant are shown in Fig. 1. The phenylpropanoid signals were relatively higher in the wild type leaves than in the CSA-line #16 leaves. In the aromatic region (δ 6.0–8.0), the signals were smaller compared with those in the aliphatic or sugar region. The signals of the main aromatic compound in the extract were assigned to chlorogenic acid (Fig. 2). The ^{1}H NMR spectrum is in accordance with a phenylpropanoid, showing the characteristic signals due to two *trans* olefinic protons (1H each, *d*, J = 15.9 Hz at δ 7.68, H-7′ and δ 6.42, *d*, J = 15.9 Hz, H-8′). In addition, three aromatic protons at δ 7.24 (1H, *s*), δ 7.15 (1H, *d*, J = 8.5 Hz), and δ 6.96 (1H, *d*, J = 8.5 Hz) correspond to H-2′, H-6′, and H-5′ of the aromatic ring of chlorogenic acid, respectively. Other signals were detected close to those of chlorogenic acid. They are downfield shifted ca. 0.05 ppm from the chlorogenic acid signals and assumed to be those of chlorogenic acid derivatives such as 4-*O*-caffeoylquinic acid or 5-*O*-caffeoylquinic acid because of the same coupling constants and correlation pattern in the ^{1}H–^{1}H COSY spectrum.

In the region of δ 2.0–5.0, there were big differences between wild type and CSA plants. Compared to the wild type leaves, the signals at δ 2.47 (*dd*, J = 15.6, 3.2 Hz), 2.69 (*dd*, J = 15.6, 9.5 Hz), 4.32 (*dd*, J = 9.4, 3.2 Hz) were greatly increased in the CSA samples. These signals were assigned to be H-3a (close to H-2), H-3b (close to C-OH), and H-2 of malic acid, respectively. In addition to these signals, signals from threonine at δ 1.34 (*d*, J = 7.5 Hz), alanine at δ 1.48 (*d*, J = 7.3 Hz), glycine at δ 3.56 (*s*), fructose at δ 4.22 (*d*, J = 8.7 Hz), β-glucose at δ 4.64 (*d*, J = 7.9 Hz), α-glucose at δ 5.24 (*d*, J = 3.9 Hz), sucrose at δ 5.41 (*d*, J = 3.9 Hz), and fumaric acid at δ 6.53 (*s*) were assigned on the basis of the comparison with the chemical shifts of standard compounds and 2D-NMR using ^{1}H–^{1}H COSY (correlation spectroscopy), HMQC (heteronuclear multiple quantum coherence), and HMBC (heteronuclear multiple bond coherence).

Although there were clear visual differences between the spectra, for nonbiased interpretation of the results, the samples were analyzed using PCA. Subsequently, further investigations were made to assign the peaks.

2.2. PCA analysis

Principle component analysis (PCA) is an unsupervised clustering method requiring no knowledge of the data set and acts to reduce the dimensionality of multivariate data while preserving most of the variance within it (Eriksson et al., 2001). In applying PCA, no large differences were observed between the spectra of the $CHCl_3$ fractions of each sample, and all the spots in the PCA gathered in the same region (data not shown). Therefore the focus was placed on the results obtained from the aqueous fraction. As seen in Fig. 3, there is a clear discrimination possible between the samples of wild type and CSA plants, both in leaves and veins. This separation took place in the first two principal components which cumulatively account for 89.6% of variation. The separation between the wild type and CSA plants was easily achieved in both of the scores of principal component 1 (PC1) and principal component 2 (PC2). For the wild type and CSA leaves, the non-inoculated (WNL and CNL leaves) and systemic leaves (WSL and CSL leaves), which are not infected directly by tobacco mosaic virus (TMV) but the nearby leaves of directly infected leaves by TMV, were located close to each other, while the location of the plots of inoculated

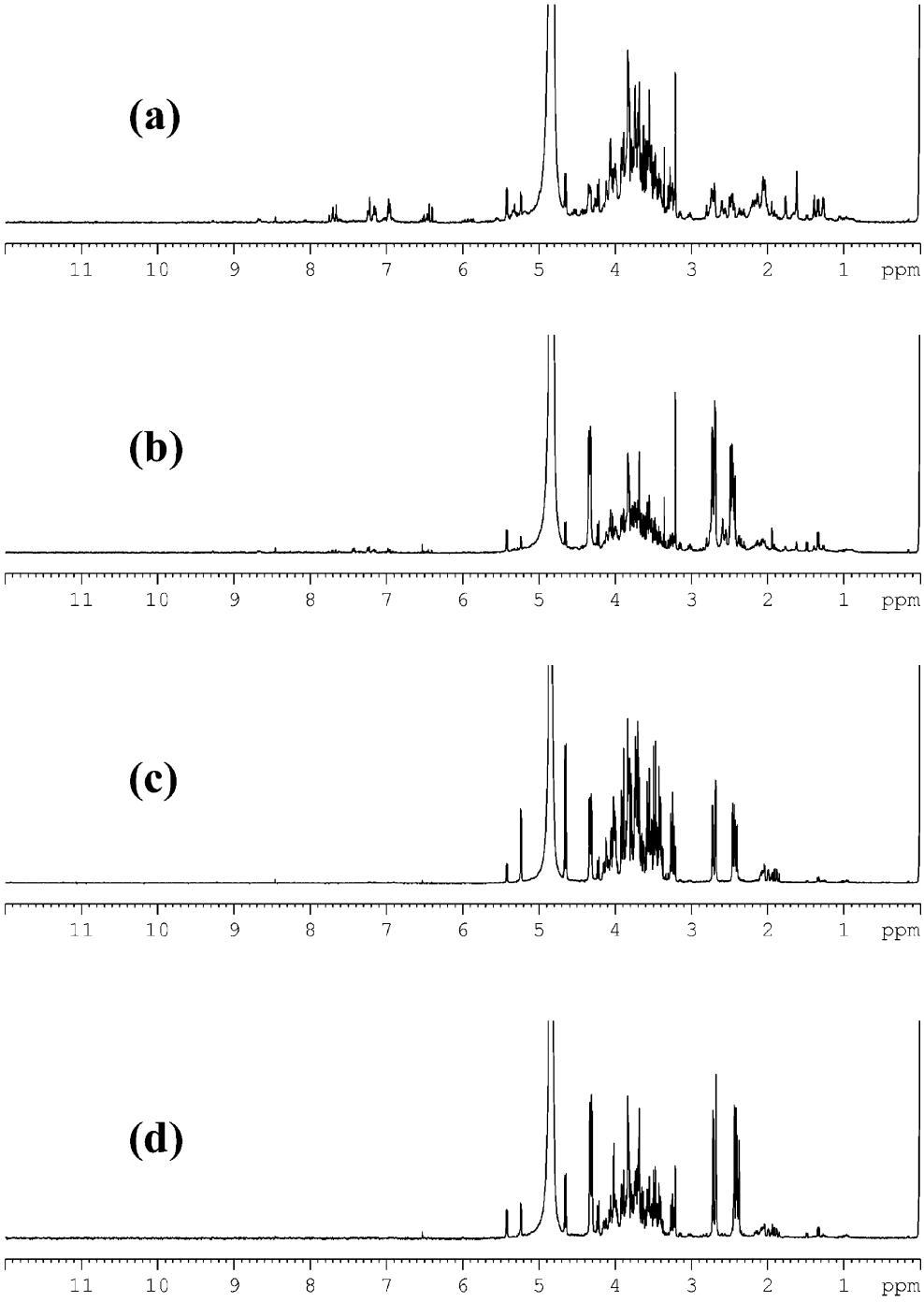

Fig. 1. ^{1}H NMR spectra for aqueous fractions of samples (a) leaves of wild type plants, (b) leaves of CSA plants, (c) veins of wild type plants, (d) veins of CSA plants.

leaves (WIL and CIL leaves) were quite well separated from them (Fig. 3).

In the samples of veins, there was also a clear difference between wild type and CSA veins, and the discrimination could be achieved in both PC1 and PC2 (Fig. 3). The PC1 and PC2 of inoculated veins (WIL and CIL veins) were well separated from those of non-inoculated (WNL and CNL veins) and systemic veins (WSL and CSL veins).

In the comparison of the metabolomic profiles of leaves and veins, the discrimination between leaves and veins in the wild type plants (the leaves and veins of WNL, WIL, and WSL) was found to be considerably larger than between CSA plants (the leaves and veins of CNL, CIL, and CSL) (Fig. 3). In particular, the leaves and veins of TMV-virus inoculated CSA plants (CSL leaves and veins) showed more similar metabolomic

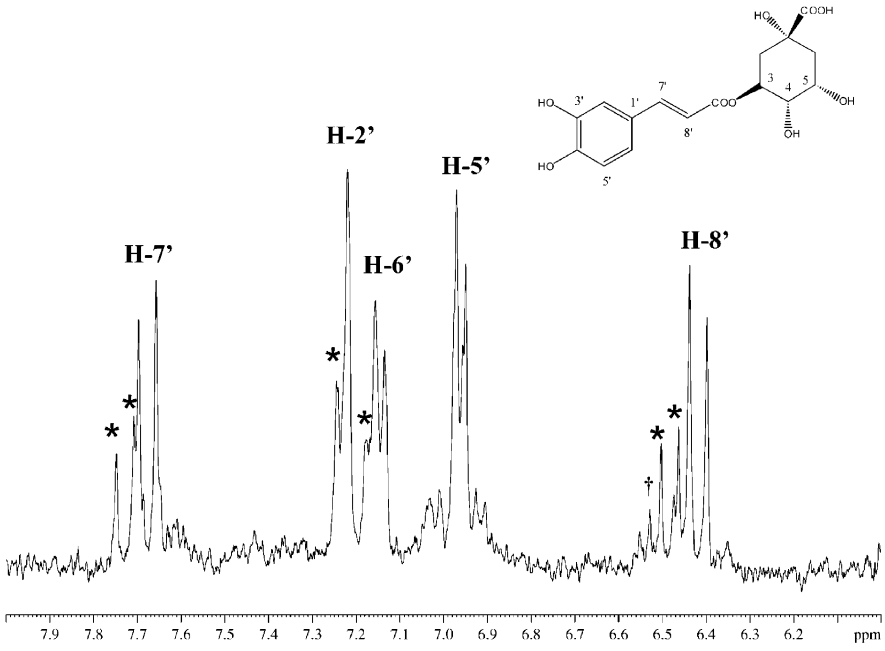

Fig. 2. ^{1}H NMR spectra for aqueous fractions of samples leaves of wild type plants in the rage of δ 6.0 – 8.0. * Possible signals of chlorogenic acid derivatives; †: fumaric acid.

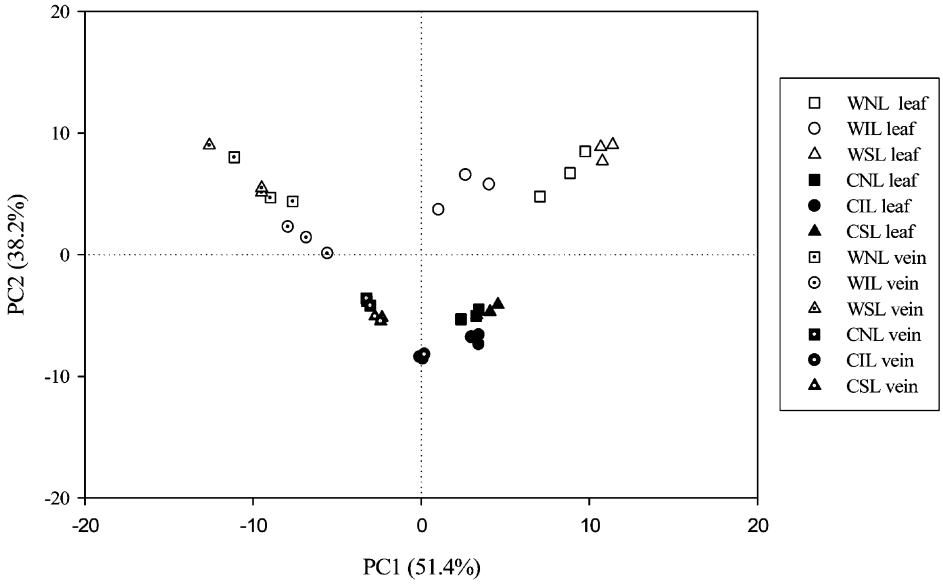

Fig. 3. PC1 scores vs. PC2 scores of aqueous fraction for wild type and CSA-line #16 leaves and veins following PCA analysis. WNL: wild type non-inoculated leaf, WIL: wild type inoculated leaf, WSL: wild type systemic leaf, CNL: CSA non-inoculated leaf, CIL: CSA inoculated leaf, CSL: CSA systemic leaf, WNV: wild type non-inoculated vein, WIV: wild type inoculated vein, WSV: wild type systemic vein, CNV: CSA non-inoculated vein, CIV: CSA inoculated vein, CSV: CSA systemic vein.

profiles when compared with the others. Actually, the inoculation with TMV-virus seems to have reduced the difference between the metabolomic profiles to become closer. In every experiment evaluated in this study of the leaves and veins of wild type and CSA plants, the inoculated ones always showed similar PC1 and PC2 values (Fig. 3).

It was possible to determine variable importance by analyzing the correlation of each variable with PC1 and PC2 scores. The loading plot of all ^{1}H NMR signals evaluated is shown in Fig. 4. The discrimination between wild type and CSA plants both in leaves and veins occurred predominantly in PC2. It showed that the wild type non-inoculated (WNL leaves) and systemic

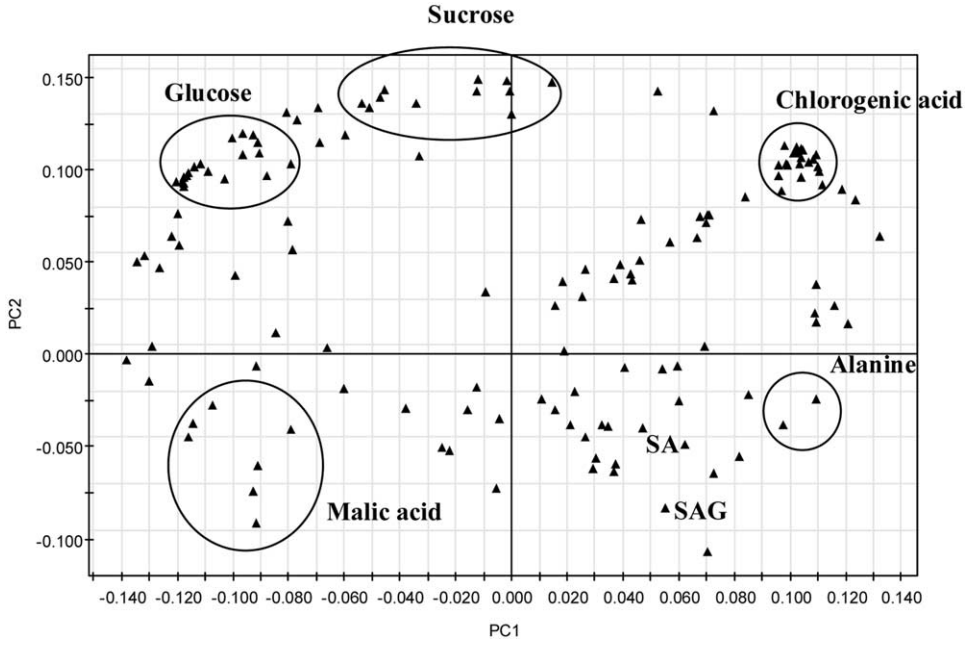

Fig. 4. Loading plot of all ^{1}H NMR signals. The variables for salicylic acid (SA) and salicylic acid glucoside (SAG) were obtained from HPLC result (see Section 4).

leaves (WSL leaves) contain a much higher level of chlorogenic acid. Glucose is the discriminating component for wild type non-inoculated (WNL veins) and systemic veins (WSL veins). In CSA plants, however, alanine and malic acid are relatively higher. As an example, the intensities of ^{1}H NMR signals of chlorogenic acid (δ 7.66) and malic acid (δ 2.72) are shown for all plant samples evaluated in Fig. 5. Chlorogenic acid could clearly distinguish wild type non-inoculated and systemic leaves but malic acid is predominant compounds in CSA leaves.

Chlorogenic acid is mainly present in the wild type leaves. It was already reported that the suppression of chlorogenic acid accumulation in CSA plants might be due to the channeling of chorismate into isochorismate and away from the phenylpropanoid pathway (Nugroho et al., 2002c). The decrease of the glucose, fructose and sucrose levels in CSA plants indicates the rapid utilization of those compounds for the production of other compounds, or a lower production of carbohydrates. Considering that the CSA plants were a bit shorter than the wild type plants, the sugars might be used for rendering resistance to the stress conditions caused by constitutively produced salicylic acid (SA) rather than for growth. These observations are in accordance with a report that exogenous treatment of tobacco plants with SA decreased the growth, transpiration, chlorophyll content of their leaves, and the size

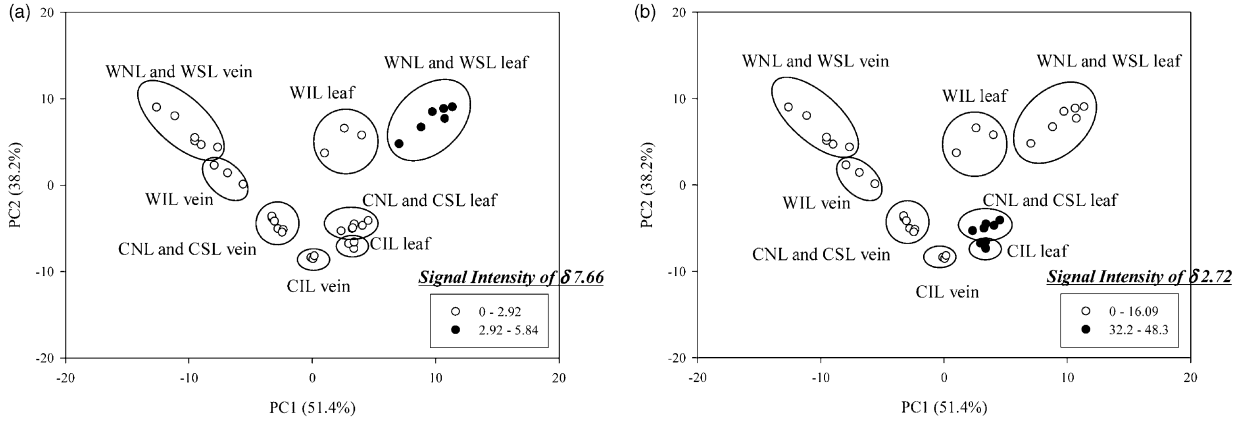

Fig. 5. Effect of chlorogenic acid (δ 7.66) and malic acid (δ 2.72) on the differentiation of aqueous fraction of tobacco plants on the plot of PC1 and PC2 scores. WNL: wild type non-inoculated leaf, WIL: wild type inoculated leaf, WSL: wild type systemic leaf, CNL: CSA non-inoculated leaf, CIL: CSA inoculated leaf, CSL: CSA systemic leaf, WNV: wild type non-inoculated vein, WIV: wild type inoculated vein, WSV: wild type systemic vein, CNV: CSA non-inoculated vein, CIV: CSA inoculated vein, CSV: CSA systemic vein.

Table 1
Salicylic acid (SA) and salicylic glucoside (SAG) accumulation (μg/g FW) in leaves and veins of wild type and CSA line #16 tobacco plants, with or without inoculation with TMV

		Non-inoculated	Inoculated	Systemic
Wild type				
Leaf	SA	0.00	1.12	0.05
	SAG	0.04	9.14	0.16
Vein	SA	0.00	0.05	0.00
	SAG	0.01	0.33	0.01
CSA #16				
Leaf	SA	0.33	1.39	0.43
	SAG	33.21	32.79	33.23
Vein	SA	0.04	0.08	0.08
	SAG	1.16	1.31	1.00

of the stomatal openings (Roggero and Pennazio, 1988). SA is also known as a compound inducing a temperature rise in flowers of *Arum* plants (Raskin et al., 1987). In addition, Pierpoint suggested that the SA might inhibit photosynthesis and inhibit the reduction of toxic nitrite in the presence of light condition, hence inducing a general stress to plants (Pierpoint, 1994). An increased level of malic acid in CSA plant implies that there might be the inhibition of photosynthesis.

Table 1 shows the summary of the salicylic acid (SA) and salicylic acid glucoside (SAG) accumulation in various samples in this study. Most of the salicylic acid was present in the SAG form, and higher contents of SA and SAG were detected in the CSA plants as reported previously (Verberne et al., 2002). To confirm the separation between wild type and CSA plants by SA and SAG, their analytical results were plotted together with ^{1}H NMR signals after principal component analysis. SA and SAG are correlated with the ^{1}H NMR signals such as δ 1.34, 2.10, 2.34, 2.36, 2.58, 3.00. These signals are assumed to be due to plant organic acids or amino acids but it is difficult to elucidate the signals because they are very minor signals.

3. Conclusion

The work has proven that it is possible to discriminate between wild type and transgenic CSA tobacco plants by multivariate analysis of ^{1}H NMR spectra of crude extracts of the plant materials. The major compounds contributing to the discrimination were chlorogenic acid, malic acid, glucose and sucrose. Because genetic engineering is widely used for major crop plants, investigations on the possible changes in the metabolic profiles have to be made, and the consequences of any difference for safety of the use of those plants as food or feed have to be considered. The technology described in this report can be applied for the detection of differences in metabolite profiles between wild type and transgenic plants, and identification of the effects of the genetic changes made, as in this case the effect of introduction of a new pathway of SA on plant metabolism. The method using the NMR and multivariate analysis used in this report are relatively simple and efficient one and it would be easily applicable for metabolomic fingerprinting of various sources such as transgenic tissues or environmentally stressed samples.

4. Experimental

4.1. Tobacco plants

Seeds of T1 CSA-line #16 of *Nicotiana tabacum* 'Samsun' NN were germinated on Murashige and Skoog (1962) agar medium containing kanamycin (100 mg/l). The kanamycin resistant seedlings and non-transgenic tobacco seedlings were transferred into soil and grown in the greenhouse to the 4–5 leaves stage (about 2 months old). Eight plants were used for TMV inoculation. The leaves of four plants were harvested and frozen in liquid nitrogen. The veins of the frozen leaves were carefully removed from the mesophyll using forceps in liquid nitrogen. The leaves and the veins were homogenized and stored at −80 °C. The veins were cleaned in liquid nitrogen separating the remaining mesophyll cells with a razor blade.

4.2. Tobacco mosaic virus (TMV) inoculation

Eight plants of both wild type and CSA-line #16 were inoculated with TMV. Three leaves per plant were dusted with carborundum powder prior to virus inoculation. A TMV stock solution was diluted to 1 μg/ml in water and each dusted leaf was inoculated with 100 μl of this solution. Five days after inoculation, the inoculated leaves of four plants were harvested, frozen in liquid nitrogen and stored at −80 °C. Ten days after inoculation, three systemic leaves directly above the inoculated leaves of the remaining four plants were harvested in the same way. The four plants per sample were pooled, veins and mesophyll were separated as described above, homogenized, and used for further analysis.

4.3. Solvents and chemicals

First grade acetic acid cyclohexane, and ethyl acetate were purchased from J.T. Baker (Deventer, The Netherlands); chloroform and methanol were from Merck Biosolve Ltd. (Valkenswaard, The Netherlands). $CDCl_3$ (99.96%) and D_2O (99.00%) were purchased from Cambridge Isotope Laboratories Inc (Miami, FL, USA). and NaOD was purchased from Cortec (Paris,

France). Salicylic acid was purchased from Sigma (St. Louis, MO, USA) and 3,4-DHBA from Fluka (Buchs, Switzerland). Ethanol and sodium acetate were obtained from Merck (Darmstadt, Germany).

4.4. Quantitation of salicylic acid (SA) and salicylic acid glucoside (SAG)

The extraction of salicylic acid and salicylic acid glucoside was performed using the methods as reported previously (Verberne et al., 2002). Salicylic acid and salicylic acid glucoside were quantified using a Phenomenex column, type LUNA 3 μ C18. The HPLC system consisted of a 2150 HPLC pump from LKB, and a Rheodyne 7010 injector with a 100 μl loop. Twenty microlitres were injected into the HPLC system. The eluent consisted of 0.2 M sodium acetate (pH 5.5): methanol (9:1), the flow rate was 0.8 ml min^{-1}. Salicylic acid levels were determined before (SA) and after acid hydrolysis (SAG). Detection was performed with a Shimadzu RF-10Axl spectrofluorometric detector, using an emission and excitation wavelength of 407 and 305 nm, respectively.

4.5. Extraction for plant materials

The plant materials were obtained from three plants cultivated in three different times, pooled, and ground in liquid nitrogen in a mortar with a pestle. Three hundred milligrams of ground material were transferred into a centrifuge tube. Five millilitres of a 50% water–methanol mixture and 5 ml chloroform were added to the tube followed by vortexing for 30 s and sonication for 1 min. The material was then centrifuged at 3000 rpm for 20 min. The extraction was performed twice. The aqueous and organic fractions were transferred separately into a 25-ml round bottom flask and dried with a rotary vacuum evaporator.

4.6. NMR measurements

KH_2PO_4 was added to D_2O as a buffering agent. The pH of the D_2O for NMR measurements was adjusted to 6.0 using a 1 N NaOD solution. All spectra were recorded on a Brüker AV-400 NMR spectrometer operating at a proton NMR frequency of 400.13 MHz. For each sample, 128 scans were recorded with the following parameters: 0.126 Hz/point, pulse width (PW) = 4.0 μs (30°), and relaxation delay (RD) = 1.0 s. FIDs were Fourier transformed with LB = 1.0 HZ, GB = 0, and PC = 1.0. For quantitative analysis, peak integral was used. The spectra were referenced to trimethyl silane propionic acid sodium salt (TSP) at 0.00 ppm for aqueous fractions and for $CHCl_3$ fractions to residual solvent at 7.26 ppm. Hexamethyl disilane (HMDS, 0.01%, v/v) for $CDCl_3$ and TSP, 0.01%, w/v were used for internal standard. The whole peak intensities in every 0.02 ppm in ^1H NMR spectra in the range of δ −0.30–12.0 were used as variables.

4.7. Data analysis

The ^1H NMR spectra were automatically reduced to ASCII files using AMIX (v. 3.7, Brüker Biospin). Spectral intensities were scaled to HMDS for $CHCl_3$ extract and TSP for aqueous extract, and reduced to integrated regions of equal width (0.02 ppm) corresponding to the region of δ −0.30 to 10.00. The region of δ 4.6 to 5.8 was excluded from the analysis because of residual signal of water. Principle component analysis (PCA) were performed with the SIMCA-P software (Umetrics, Umeå, Sweden).

Acknowledgements

The post-doctoral Fellowship Program of Korea Science and Engineering Foundation (KOSEF) is gratefully acknowledged for their support of H.-K. Choi and Y.H. Choi. The financial support of the van Leersumfonds (KNAW) is also gratefully acknowledged.

References

Belton, P.S., Colquhoun, I.J., Kemsley, E.K., Delgadillo, I., Roma, P., Dennis, M.J., Sharman, M., Holmes, E., Nicholson, J.K., Spraul, 1998. Application of chemometrics to the ^1H NMR spectra of apple juices: discrimination between apple varieties. Food Chem. 61, 207–213.

Bollard, M.E., Holmes, E., Lindon, J.C., Mitchell, S.C., Branstetter, D., Zhang, W., Nicholson, J.K., 2001. Investigations into biochemical changes due to diurnal variation and estrus cycle in female rats using high resolution ^1H NMR spectroscopy of urine and pattern recognition. Anal. Biochem. 295, 194–292.

Brescia, M.A., Caldarola, V., de Giglio, A., Benedetti, D., Fanizzi, F.P., Sacco, A., 2002. Characterization of the geographical origin of Italian red wine based on traditional and nuclear magnetic resonance spectrometric determinations. Analytica Chimica Acta 458, 117–186.

Charlton, A.J., Farrington, H.H., Brereton, P., 2002. Application of ^1H NMR and multivariate statistics for screening complex mixtures: quality control and authenticity of instant coffee. J. Agric. Food. Chem. 50, 3098–3103.

Duarte, I., Barros, A., Belton, P., Righelato, R., Spraul, M., Humpfer, E., Gil, A.M., 2002. High resolution nuclear magnetic resonance spectroscopy and multivariate analysis for the characterization of beer. J. Agric. Food. Chem. 50, 2475–2481.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wold, S., 2001. Multi-and Megavariate Data Analysis. Umetrics Academy, Umeå, Sweden.

Gall, G.L., Puaud, M., Colquhoun, I.J., 2001. Discrimination between orange juice and pulp wash by ^1H nuclear magnetic resonance spectroscopy: identification of marker compounds. J. Agric. Food. Chem. 49, 580–588.

Gavaghan, C.L., Holmes, E., Lenz, E., Wilson, I.D., Nicholson, J., 2000. A NMR based metabonomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk:ApfCD mouse. FEBS Lett. 484, 169–174.

Halim, V.A., Verberne, M.C., Verpoorte, R., 2003. Differential metabolic profiling by HPLC-PDA-MS of wild type and transgenic

tobacco plants constitutively producing salicylic acid. J. Liq. Chromatogr. Rel. Technol. 26, 369–383.

Murashige, T., Skoog, F., 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol. Plant. 15, 473–497.

Nugroho, L.H., Peltenburg-Looman, A.M.G., Verberne, M.C., Verpoorte, R., 2002a. Is accumulation of sesquiterpenoid phytoalexins induced in tobacco plants constitutively producing salicylic acid? Plant Sci. 162, 989–993.

Nugroho, L.H., Peltenburg-Looman, A.M.G., de Vos, H., Verberne, M.C., Verpoorte, R., 2002b. Nicotine and related alkaloids accumulation in constitutive salicylic acid producing tobacco plants. Plant Sci. 162, 575–581.

Nugroho, L.H., Verberne, M.C., Verpoorte, R., 2002c. Activities of enzymes involved in the phenylpropanoid pathway in constitutively salicylic acid-producing tobacco plants. Plant Physiol. Biochem. 40, 755–760.

Ott, K.-H., Aranibar, N., Singh, B., Stockton, Q.W., 2003. Metabonomic classifies pathway affected by bioactive compounds. Artificial neural network classification of NMR spectra of plant extracts. Phytochem. 62, 971–985.

Pierpoint, W.S., 1994. Salicylic acid and its derivatives in plants: medicine, metabolites and messenger molecules. Adv. Bot. Res. 20, 164–235.

Raamsdonk, L.M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M.C., Berden, J.A., Brindle, K.M., Kell, D.B., Rowland, J.J., Westerhoff, H.V., van Dam, K., Oliver, S.G., 2001. A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. Nat. Biotechnol. 19, 45–50.

Raskin, I., 1992. Salicylate, a new plant hormone. Plant Physiol. 99, 799–803.

Raskin, I., Ehmann, A., Melander, W.R., Meeuse, B.J.D., 1987. Salicylic acid-a natural inducer of heat production in *Arum* lilies. Science 237, 1545–1556.

Roggero, P., Pennazio, S., 1988. Effects of salicylate on systemic invasion of tobacco plants by various viruses. J. Phytopathol. 123, 207–216.

Sumner, L.W., Mandes, P., Dixon, R.A., 2003. Plant metabolomics: large-scale phytochemistry in the functional genomics era. Phytochem. 62, 817–836.

Trethewey, R.N., Krotzky, A.J., Willmitzer, L., 1999. Metabolic profiling: a Rosetta stone for genomics? Curr. Opin. Plant Biol. 2, 83–85.

Verberne, M.C., Verpoorte, R., Bol, J.F., Mercado-Blanco, J., Linthorst, H.J.M., 2000. Overproduction of salicylic acid in plants by bacterial transgenes enhances pathogen resistance. Nat. Biotechnol. 18, 779–783.

Verberne, M.C., Brouwer, N., Delbianco, F., Linthorst, H.J.M., Bol, J.F., Verpoorte, R., 2002. Method for the extraction of the volatile compound salicylic acid from tobacco leaf material. Phytochem. Anal. 13, 45–50.

Vogels, J., Terwel, L., Tas, A.C., van den Berg, F., Dukel, F., van der Greef, J., 1996. Detection of adulteration in orange juices by a new screening method using proton NMR spectroscopy in combination with pattern recognition techniques. J. Agric. Food. Chem. 44, 175–180.

Ward, J.L., Harris, C., Lewis, J., Beale, M.H., 2003. Assessment of ^{1}H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. Phytochem. 62, 949–958.