Metabolomic characteristics of Catharanthus roseus plants in time and space

Qifang, P.

Chapter 4

Investigation of the chemomarkers correlated with flower color in different organs of *Catharanthus roseus* using NMR-based metabolomics

Qifang Pan [1, 2, 4], Yuntao Dai [1, 4], Tri Rini Nuringtyas [1], Natali Rianika Mustafa [1, 3], Anna Elisabeth Schulte [3], Robert Verpoorte [1] and Young Hae Choi [1]

[1] Natural Products Laboratory, Institute of Biology, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

[2] Plant Biotechnology Research Center, SJTU-Cornell Institute of Sustainable Agriculture and Biotechnology, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai 200240, PR China

[3] ExPlant Technologies B.V., Einsteinweg 55, 2333 CC Leiden, The Netherlands

[4] Co-first authors

Abstract

Flower color is a complex phenomenon involving a wide range of secondary metabolites, e.g. phenolics and carotenoids as well as co-pigments. Biosynthesis of those metabolites occurs through complex pathways and is finally expressed mostly in flowers. It was of interest, thus, to investigate the correlation between flower color and the metabolic profiles of the plant to find chemomarkers in other organs (leaf, stem, and root), which might facilitate prediction of the flower color of a plant. To investigate the metabolic profiles of leaves, stems, roots and flowers of *Catharanthus roseus* (Pacific cherry red) with four flower colors (orange, pink, purple, and red), ^{1}H-NMR and multivariate data analysis were used to characterize the metabolites in the organs. The results showed that flower color is characterized by a special pattern of metabolites such as anthocyanins, flavonoids, organic acids, and sugars; the leaves, stems, and roots also presented metabolic differences according to the flower color. Plants with orange flowers featured a relatively high level of kaempferol analogues in all organs except roots; those flowering red showed a high level of malic acid, fumaric acid and asparagine in both flowers and leaves; those flowering purple and pink presented a high level of sucrose, glucose and 2,3-dihydroxybenzoic acid. Quercetin analogues are highest in the flowers and leaves of purple flower plants. Compounds in different plant parts related to specific flower colors must be a result of similar biosynthetic pathways and interactions between these pathways. Most importantly it seems possible to predict the flower color through profiling the metabolites in leaves, stems, or roots, which may be a helpful tool for plant breeding.

Introduction

Flowers, as reproductive organs, have developed a great diversity of colors. The chemistry of color is particularly associated to the presence of a wide range of metabolites. Flower pigments fall into three major categories: flavonoids, carotenoids, and betalains. It is possible to find both flavonoids and carotenoids in the same flower (Miller *et al.* 2011). Flavonoids derive from the phenylpropanoid biosynthetic pathway (Mustafa and Verpoorte, 2007) and are responsible for the widest variety of flower colors. They include flavonols, flavones, isoflavonoids and anthocyanins. Anthocyanins are responsible for the red, pink, mauve, purple, violet, and blue color of most flowers. Anthocyanins can adopt different resonance structures depending on the pH, their bonding to co-pigments, such as organic acids or phenylpropanoids, and molecular stacking with metals, all of which result in the wide range of colors in flowers (Robinson *et al.*, 1938; Yoshida *et al.*, 2003; 2009; Malien-Aubert *et al.*, 2001; Mulder-Krieger and Verpoorte, 1994).

The flower is an important phenotype of plants and its color may be correlated with the metabolic activities in the whole plant, as exemplified by the color change of *Brunfelsia calycina* flowers from purple to white with increased levels of 19 volatiles, and a further 17 primary and secondary metabolites (Bar-Akiva *et al.*, 2010). Moreover, precursors of pigments, such as flavonols (quercetin and kaempferol), are known to accumulate not only in flowers but also in other organs, including leaves, stems and seeds (Pereira *et al.*, 2009; Ferreres *et al.*, 2008). Recent research on the modification of flower color has focused mostly on the genetic engineering of the flavonoid and carotenoid biosynthetic pathways. However, due to the complexity of pigment chemistry and metabolism in plants, the new colors generated in cultivars did not result in stable phenotypes (Tanaka *et al.*, 2010). It is thus clear, that a more comprehensive insight into the metabolome of different plant parts is necessary to understand how pigment chemistry correlates with biosynthetic pathway in other plant parts.

In this study, we used *Catharanthus roseus* with four different flower colors (orange, pink, purple and red) as a model plant because it has a quite broad range of metabolites e.g. terpenoids, indole alkaloids, flavonoids, and phenylpropanoids. *Catharanthus roseus* is grown as an ornamental plant for gardens and parks due to its rich variety of colors and a long-blooming season. Each cultivar of *C. roseus* can give a variety of flower colors. Chemical analysis of the flower pigments revealed that petunidin, malvidin and hirsutidin are all present in two different glycosidic forms and that the total and relative amounts of the pigments varies between the different colored flowers (Milo *et al.*, 1985).

Catharantus roseus plants produce a great diversity of valuable-bioactive terpenoid indole alkaloids (TIAs), phenolic compounds, such as flavonoid glucosides (Chung *et al.*, 2009; Brun *et al.*, 1999; Pereira *et al.*, 2009; Ferreres *et al.*, 2008) and anthocyanins (Filippini *et al.*, 2003; Toki *et al.*, 2008) via the

shikimate pathway. After the intermediate chorismate, the pathways separate, and alkaloids follow the tryptophan route, whereas the flavonoids and anthocyanins follow the phenylpropanoid pathway (Mustafa and Verpoorte, 2007). Both groups of secondary metabolites are stored in vacuoles. Interestingly, a higher yield of ajmalicine was found in roots of *C. roseus* plants with red-eyed flowers than those with pink and white flowers (Weissenberg, 1988). However, no further studies were made of the correlation between flower color and metabolites profiles in *C. roseus.*

In this context a comprehensive profiling method is required to investigate the metabolic networks related to plant flower color. Metabolomics undoubtedly is such a method, since it is a powerful tool that provides quantitative and qualitative information of a wide range of primary and secondary metabolites in plants, and may help to reveal the biochemical status of an organism. NMR-based metabolomics, in particular, has been widely used in metabolomic studies in diverse fields of plant physiology and classification (Choi *et al.*, 2004; Kim *et al.*, 2010a).

In this study, NMR profiling techniques followed by multivariate data analysis were applied to investigate the metabolic difference of *C. roseus* with four flower colors to find marker metabolites in non-flower organs characteristic for each flower color. This metabolomics study was also implemented to obtain a comprehensive insight of the total metabolic network in *C. roseus* plants and the interrelationships between the occurrences of diverse metabolites in different organs of *C. roseus.*

Methods and materials

Plant materials

Four different flower-colored *Catharanthus roseus* purchased from Intratuin Pijnacker (Postbus 1016, 1700BA, Heerhugowaard, The Netherlands), belonging to the Pacific series i.e. Pacifica XP Apricot (orange, 1), Pacifica cherry red (red, 2), Pacifica XP Blush (pink, 3), and Pacifica Orchid Halo (purple, 4) were used in this study (Supplementary Fig. 1). At least five biological replicates of each cultivar were collected in September 2009 during the blooming season of *C. roseus*. Flowers, leaves, stems, and roots of the plants were collected separately as independent samples. Two types of leaf samples were collected according to their position in the plants: young leaves in the upper 3 layers from the top and other leaves from the lower 3 layer from the bottom. Samples of fresh root, leaf, flower and stem were ground into powder with liquid nitrogen and freeze-dried in 2 days.

Extraction and NMR measurements

The methods of extraction and NMR analysis used in this study were those described previously (Kim *et al.,* 2010b). A dried sample of 50 mg was transferred to a 2 ml-micro tube to which 1.5 ml of methanol-d_4 (750 μL) and D_2O (750 μL) (KH_2PO_4 buffer, pH 6.0), containing 0.01% (w/w) TMSP-d_4 (trimethylsilyl propionic acid sodium salt), were added. The mixture was vortexed for 1 minute, ultrasonicated for 30 minutes, and centrifuged for 20 minutes at 13,000 rpm at room temperature. An aliquot of 800 μL of the supernatant from the mixture was transferred to a new Eppendorf tube for a second centrifugation at 13,000 rpm for 5 minutes, after which 750 μl of the supernatant was transferred to a 5 mm-NMR tube and used for the ^{1}H-NMR analysis. Both deuterated methanol and water were purchased from Cambridge Isotope (Andover, MA, USA).

The ^{1}H-NMR spectra were recorded using a Bruker DMX 600 spectrometer (Bruker, Karlsruhe, Germany). For each sample, 128 scans were recorded with the following parameters: 0.167 Hz/point, pulse width (PW) = 4.0 μs, acquisition time (AQ) = 3.17 s, demiscans (DS) = 4, and relaxation delay (RD) = 1.5 s. Free induction decays (FIDs) were Fourier transformed with line broadening (LB) = 0.3 Hz. Manual phase adjustment and baseline correction were applied prior to integration of target regions for quantitative analysis. The parameters for two-dimensional NMR spectra such as *J*-resolved, correlated spectroscopy (COSY), heteronuclear single quantum coherence (HSQC), and heteronuclear multiple bond correlation (HMBC) were those described previously (Kim *et al.,* 2010b).

Data Analysis

^{1}H-NMR spectra were automatically binned by AMIX software (v.3.7, Biospin, Bruker). Spectral intensities were scaled to total intensity and the region of δ 0.30 - δ 10.00 was reduced to integrated regions of 0.04 ppm each. The regions of δ 4.7 - δ 5.0 and δ 3.28 - δ 3.34 were excluded from the analysis because of the residual water and methanol signal, respectively. SIMCA-P$^+$ software (v.12.0, Umetrics, Umeå, Sweden) was used for principal component analysis (PCA) and partial least square (PLS) modeling with Pareto-scaled data of binned ^{1}H-NMR. All experiments were conducted with five replicates. Statistical analysis was performed using one way analysis of variance (ANOVA) followed by Duncan's Multiple Range (DMRT) test. The values are mean ± SD for three samples in each group. P values ≤ 0.05 were considered as significant. The ANOVA for all the data was performed by SPSS (version 14.0, Chicago, IL, USA).

Results and Discussion

Assignment of the signals in ^{1}H-NMR spectra

^{1}H-NMR spectra of the organs of *C. roseus* such as root, stem, upper leaf, lower leaf, and flower show very different signals (Fig. 1). Even upper and lower leaves are found to have quite different metabolites both in the aliphatic and aromatic regions of the ^{1}H-NMR. *Catharanthus roseus* plants with the flowers of the four colors, i.e. orange, red, pink and purple exhibited obvious differences in the proton spectra (Fig. 2). ^{1}H-NMR has been widely used in metabolomics and related fields of plant science. Nevertheless, limitations in resolution and signal overlapping in ^{1}H-NMR spectra require the combination of different NMR technologies, such as *J*-resolved, COSY, HSQC, or HMBC etc. Thus, the use of 2D NMR extends the capability of NMR for use in metabolomics. This allowed a number of primary and secondary metabolites to be identified based on information from the ^{1}H-NMR, *J*-resolved, COSY and HSQC, and HMBC spectra and from previous publications (Table 1) (Choi *et al.*, 2004; Yang *et al.*, 2009; Kim *et al.*, 2010a).

Signals were assigned to metabolites with clear difference in the area of amino acids and aliphatic compounds (δ 0.5 - δ 3.1), the area of sugars (δ 3.5 - δ 5.5), and the area of aromatic/phenolic compounds (δ 6.0 – δ 8.5) (Fig. 1 and 2). The signals of the main aromatic compounds in the flower and leaf extracts were assigned to quercetin-3-*O*-glucoside, chlorogenic acid, 4-*O*-caffeoylquinic acid, secologanin, and vindoline but they were not detected in the spectra of stem and root extracts. The ^{1}H-NMR spectra of flowers showed more signals of phenolic compounds, especially those of flavonoids. The signals of H-2' and H-6' from two kaempferol analogues were observed at δ 8.04 (d, J = 7.9 Hz) and δ 8.08 (d, J = 7.9 Hz), respectively in the spectrum of the orange flowers, both of which were correlated with the signals at δ 7.00 (d, J = 7.9 Hz) of H-3' and H-5' in the ^{1}H-^{1}H COSY spectrum. Another signal at δ 6.50 of H-8 (d, J = 2.1 Hz) was correlated with one at δ 6.30 of H-6 (d, J = 2.1 Hz). In the NMR spectra of other organs, however, one of the kaempferol resonances at δ 8.04 was found at very low intensity, almost undetectable. In the ^{1}H-NMR spectrum of flowers, the quercetin signal at δ 7.65 of H-6' correlated with a signal at δ 6.99 of H-5' and a signal at δ 7.88 of H-2' in the ^{1}H-^{1}H COSY spectrum. Three aromatic protons at δ 7.52 (dd), δ 7.26 (dd), and δ 6.83 (t) which were observed in the spectra of purple and pink flowers correspond to H-6, H-4, and H-5 of the aromatic ring of 2,3-dihydroxybenzoic acid (DHBA). The spectra of stems and roots show two additional signals at δ 7.08 and δ 5.30 that did not appear in flowers and leaves spectra. Based on reference compounds, these signals were assigned to the H-3 and H-1 of loganic acid (Choi, *et al.*, 2004). The signals corresponding to vindolinine and catharanthine were detected in the spectra of all organs (Table 1).

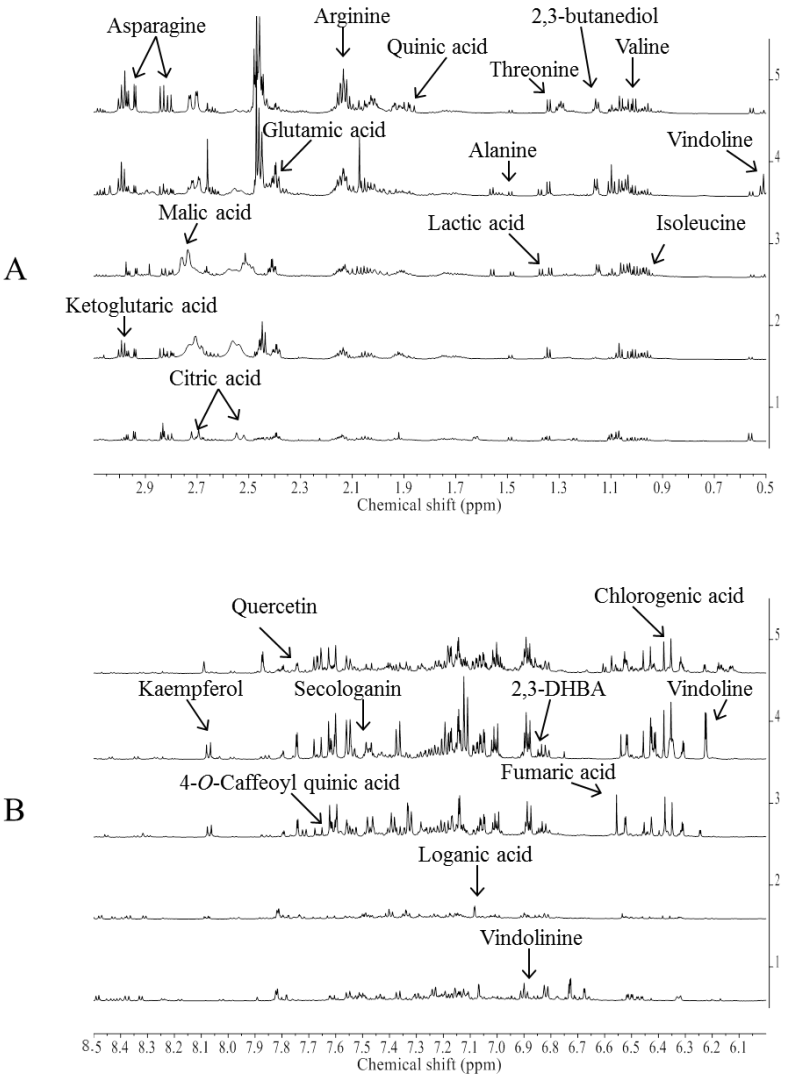


Fig. 1 ^{1}H-NMR spectra of extracts of different organs of red-colored *Catharanthus roseus* with the assigned metabolites: 1: root; 2: stem; 3 lower leaf; 4 upper leaf; 5 flower. A: aliphatic amino acids in the range of δ 0.5 – δ 3.1. B: aromatic metabolites in the range of δ 6.0 – δ 8.5.

In the aliphatic region, the high signal intensities facilitated the identification of amino acids and organic acids. This part of the spectra showed significant differences for the different plant parts. The signals of isoleucine at δ 0.96, leucine at δ 0.97 and δ 0.99, valine at δ 1.01 and δ 1.06, threonine at δ

1.34, glutamic acid at δ 2.39, asparagine at δ 2.82 and δ 2.96 were higher in the spectra of flowers and leaves than that of stems and roots. Some organic acids, such as quinic acid (δ 1.88 and δ 1.92), malic acid (δ 2.72) and ketoglutaric acid (δ 2.99), were also more intense in flower and leaf spectra. Stems exhibited a strong signal corresponding to citric acid at δ 2.52. Resonances of 2,3-butanediol observed at δ 1.15 correlated with the protons at δ 3.51 in ^{1}H-^{1}H COSY spectra and the carbon at δ 72.5 in HSQC spectra. In the sugar part, flower spectra showed extremely strong signals of glucose at δ 5.19 and δ 4.58 as compared to the spectra of other organs. Signal of sucrose at δ 5.41 were observed in all parts of *C. roseus*.

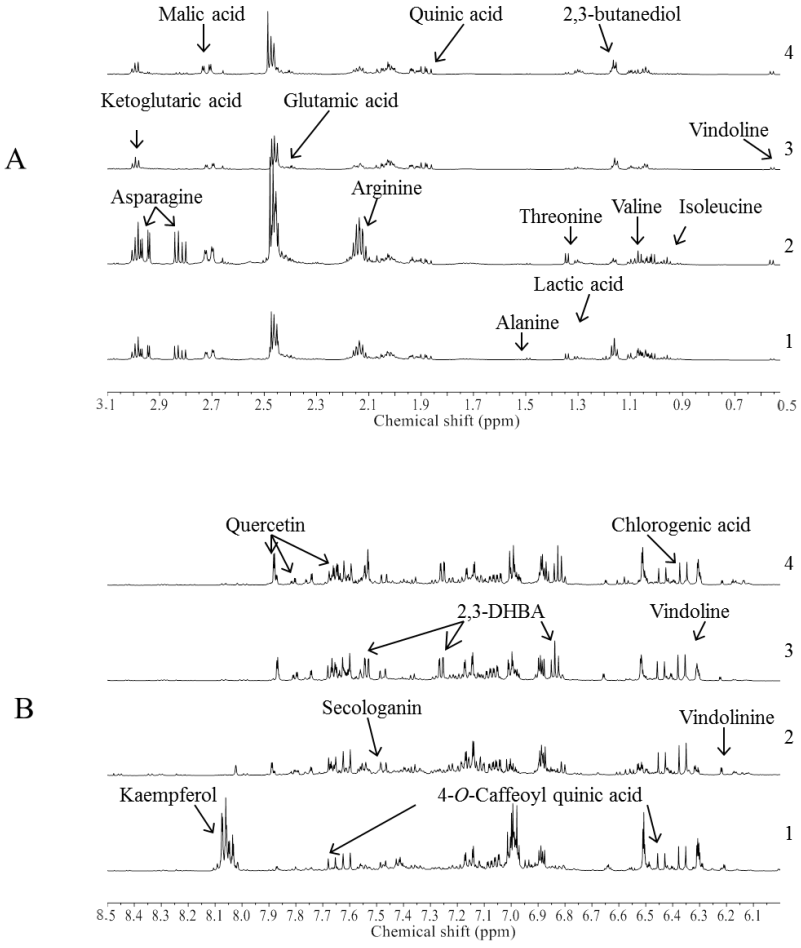


Fig. 2 ^{1}H-NMR spectra of extracts of flowers of *Catharanthus roseus* with four colors: 1: orange; 2: red; 3 pink; 4 purple. A: aliphatic amino acids in the range of δ 0.5 – δ 3.1. B: aromatic metabolites in the range of δ 6.0 – δ 8.5.

Table 1. ^{1}H-NMR chemical shift (d) and coupling constants (Hz) of identified metabolites based on ^{1}H-NMR, J-resolved, COSY, HSQC, HMBC, and reference compounds.

No	Compounds	Chemical shifts (δ) (ppm)
1	Isoleucine	0.96 (t, *J* = 7.4), 1.03 (d, *J* = 6.8)
2	Leucine	0.97 (d, *J* = 6.3), 0.99 (d, *J* = 6.3)
3	Valine	1.01 (d, *J* = 7.0), 1.06 (d, *J* = 7.0), 2.28 (m)
4	Threonine	1.34 (d, *J* = 6.6)
5	Alanine	1.48 (d, *J* = 7.2)
6	Arginine	1.70 (m), 1.90 (m)
7	Glutamic acid	2.04 (m), 2.12 (m), 2.39 (m)
8	Glutamine	2.15 (m), 2.48 (m)
9	Aspartic acid	2.64 (dd), 3.83 (dd)
10	Asparagine	2.82 (dd, *J* = 16.9, 8.2), 2.96 (dd, *J* = 16.9, 3.9)
11	Serine	3.78 (dd, *J* = 6.2, 3.7), 3.92 (dd, *J* = 12, 6.2), 3.98 (dd, *J* = 12, 3.7)
12	2,3-butanediol	1.15 (d, *J* = 6.4), 3.50 (m)
13	EtOH	1.19 (t, *J* = 7.1)
14	Quinic acid	1.88 (dd), 1.92 (m)
15	Lactic acid	1.37 (d, *J* = 7.2)
16	Acetic acid	1.94 (s)
17	Malic acid	2.45 (dd, *J* = 15.6, 7.2), 2.72 (dd, *J* = 15.6, 3.9)
18	Citric acid	2.44 (d, *J* = 15.6), 2.71 (d, *J* = 15.6)
19	Ketoglutaric acid	2.99 (t, *J* = 7.5)
20	Succinic acid	2.51 (s)
21	Oxalacetic acid	3.65 (s)
22	Fumaric acid	6.55 (s)
23	Sucrose	5.41 (d, *J* = 3.8)
24	*α*-glucose	5.19 (d, *J* = 3.7)
25	*β*-glucose	4.58 (d, *J* = 7.9)
26	Choline	3.21 (s)
27	Chlorogenic acid	7.61 (d, *J* = 15.9), 7.14 (d, *J* = 2.1), 7.05 (dd, *J* = 8.4, 2.1) 6.36 (d, *J* = 15.9)
28	2,3-DHBA	6.83 (t, *J* = 8.1), 7.26 (dd, *J* = 8.1, 1.5), 7.52 (dd, *J* = 8.1, 1.5)
29	Quercetin analogues	7.88 (d, *J* = 2.1), 7.65 (dd, *J* = 8.5, 2.1), 6.99 (d, *J* = 9.3)
30	Kaempferol analogues	6.30 (d, *J* = 2.1), 6.51 (d, *J* = 2.1), 7.00 (d, *J* = 9.3), 8.04 (d, *J* = 9.3)

31	Kaempferol analogues	6.30 (d, *J* = 2.1), 6.51 (d, *J* = 2.1), 6.98 (d, *J* = 9.3), 8.08 (d, *J* = 9.3)
32	Loganic acid	7.06 (d, *J* = 1.1), 5.30 (d, *J* = 3.2), 4.72 (d, *J* = 8.0), 1.07 (d, *J* = 6.9).
33	Secologanin	7.46 (d, *J* = 0.9), 7.48 (d, *J* = 0.9), 7.56 (d, *J* = 1.5), 9.65 (q)
34	Catharanthine	1.10 (t, *J* = 7.3), 7.36 (d, *J* = 8), 7.55 (d, *J* = 8)
35	Vindoline	0.51 (t, *J* = 7.4), 2.00 (s), 2.66 (s), 5.93 (m), 6.22 (d, *J* = 2.3), 7.11 (d, *J* = 8)
36	Vindolinine	1.08 (d, *J* = 5.6), 5.92 (m), 6.51 (dd, *J* = 9.9, 3.1), 6.82 (d, J = 7.9), 7.24 (d, *J* = 7.3)
37	4-*O*-Caffeoyl quinic acid	7.67 (d, *J* = 15.9), 7.17 (d, *J* = 2.1), 7.08 (dd, *J* = 8.3, 2.1), 2.1), 6.44 (d, *J* = 15.9), 2.09 (m)

Multivariate data analysis to find the chemo-marker compounds for different flower color

The ^{1}H-NMR data of all flower samples were subjected to different multivariate data analysis methods for the metabolic discrimination of different colored flowers and the investigation of the metabolites associated to these flower colors. Principal component analysis (PCA) is usually used first to analyze data in an unbiased and unsupervised way, providing an insight into the differences or similarities among the samples. The principal components (PC) can be presented graphically as a score plot. The corresponding loading plot shows the metabolites responsible for the separation on the score plot. To get more specific information on the common or discriminating metabolites in different groups of samples, a supervised multivariate analysis method such as partial least square discriminant analysis (PLS-DA) can be further applied to the separation. The PLS-DA model was validated by the permutation method through 20 applications.

Based on ^{1}H-NMR data, the PLS-DA score plot of flower samples showed all four cultivars of different flower-colored plants to be separated into three groups by PLS components 1 and 2: orange, red and pink/purple (Fig. 3A). From the PLS-DA loadings scatter plot of PLS component 1 and 2 (Fig. 3B), the intense signals of kaempferol analogues, 2,3-butanediol, leucine/isoleucine, and ethanol indicated high levels of these compounds in the orange flower group. Purple and pink flowers were characterized by the signals of quercetin (δ 7.88), 2,3-DHBA, glucose, sucrose, and choline. In the case of red flowers, signals of chlorogenic acid, 4-*O*-caffeoyl quinic acid, malic acid, citric acid, fumaric acid, quinic acid, glutamine and vindoline were found to be higher than for other colors.

Flower color is determined by a combination of pigments and co-pigments, which are all secondary metabolites, stored in cell vacuoles of floral tissues (Tanaka *et al.*, 2010). The accumulation of secondary metabolites, such as flavonoids, carotenoids and betalains, is responsible for flower color development (Tanaka *et al.*, 2010). Flavones and/or flavonols can form a weak complex with anthocyanidins (Seigler, 1995). The flavonols kaempferol and quercetin have been identified in *C. roseus* flowers (Forsyth and Simmonds, 1957). The biosynthesis of these compounds is catalyzed by flavonol synthase (FLS) from the dihydroflavonols substrate, which can also be reduced to the precursors of anthocyanins through the action of dihydroflavonol 4-reductase (DFR) (Nishihara and Nakatsuka, 2010). Competition between FLS and DFR thus affects flower color (Tanaka *et al.*, 2010).

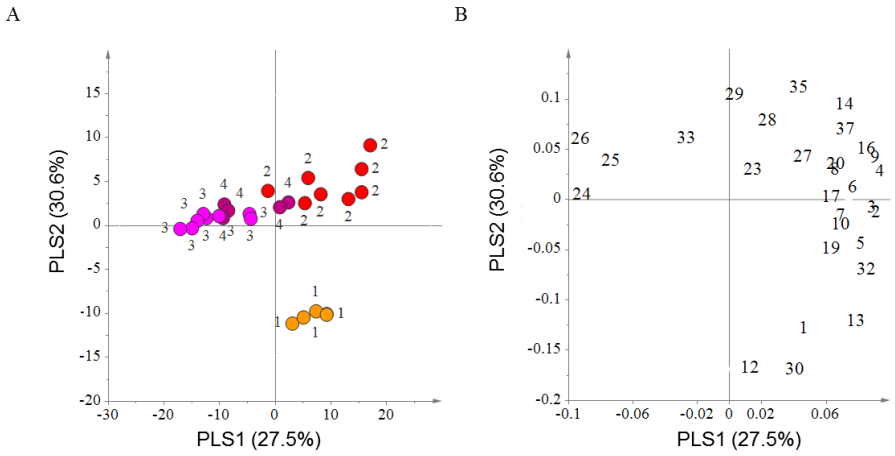
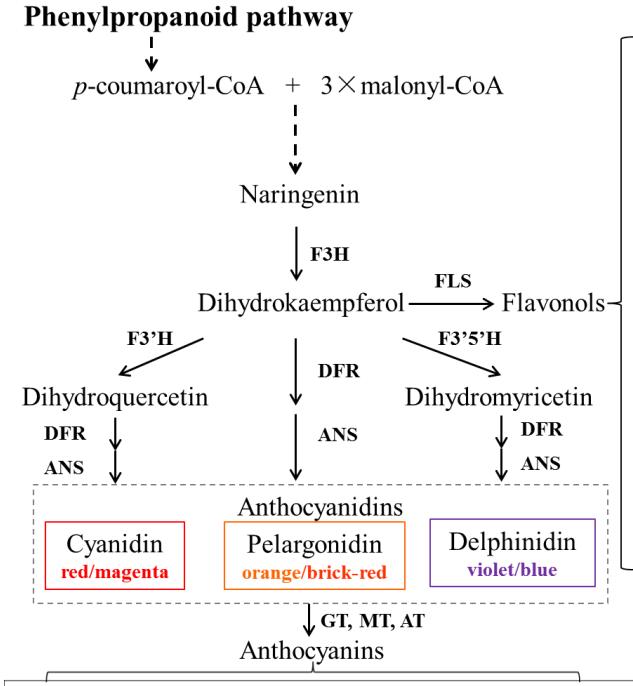


Fig. 3 PLS-DA results of flowers with four colors from *C. roseus* plants. 1: orange; 2: red; 3 pink; 4 purple. A: score plot. B: loadings plot. Numbers in loadings plot are the same as those of Table 1 of metabolites.

Our data showed a high level of kaempferol in orange flowers but a low level of quercetin, indicating that dihydroflavonols may be mainly converted into kaempferol as a co-pigment in the orange color flowers (Fig. 4). Conversely, quercetin was more abundant in purple, pink and red flowers. Different pigment/co-pigment compositions have proved to be responsible for the basic color differences between white, cream, pink, mauve and purple flowers (Markham and Ofman, 1993). Except quercetin, purple flowers featured a high level of 2,3-DHBA, which is the iron-binding component of enterobactin (Devireddy *et al.*, 2010).



Flavonols in *C. roseus*	Parts
Kaempferol-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)-7-*O*-hexoside	Seed, leaf
Kaempferol-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)	Seed, petal, stem, leaf
Kaempferol-3-*O*-(2,6-di-*O*-rhamnosyl-glucoside)	Seed, petal
Kaempferol-3-*O*-(6-*O*-rhamnosyl-galactoside)	Seed, petal
Kaempferol-3-*O*-(6-*O*-rhamnosyl-glucoside)	Petal
Quercetin-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)	Seed, petal, stem, leaf
Isorhamnetin-3-*O*-(2,6-di-*O*-rhamnosyl-glucoside)	Seed
Isorhamnetin-3-*O*-(6-*O*-rhamnosyl-galactoside)	Petal
Isorhamnetin-3-*O*-(6-*O*-rhamnosyl-glucoside)	Seed, petal
3',4'-di-*O*-methylquercetin-7-*O*-D-glucopyranoside	Hairy root
4'-*O*-methylkaempferol-3-*O*-D-glucopyranoside	Hairy root
3',4'-di-*O*-methylbutin-7-*O*-D-glucopyranoside	Hairy root
4'-*O*-methylbutin-7-*O*-D-glucopyranoside	Hairy root
syringetin-3-*O*-robinobioside	Stem
Kaempferol-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)	Stem
Quercetin-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)	Stem
Isorhamnetin-3-*O*-(2,6-di-*O*-rhamnosyl-galactoside)	Stem

Anthocyanins in *C. roseus*	Color	Parts
Hirsutidin 3-*O*-glucoside	red	Flower, cell culture
Hirsutidin 3-*O*-(6-*O*-*p*-coumaroyl) glucoside	red	Flower, cell culture
Malvidin 3-*O*-glucoside	red	Flower, cell culture
Malvidin 3-*O*-(6-*O*-*p*-coumaroyl) glucoside	red	Flower, cell culture
Petunidin 3-*O*-glucoside	red	Flower, cell culture
Petunidin 3-*O*-(6-*O*-*p*-coumaroyl) glucoside	red	Flower, cell culture
Rosinidin 3-*O*-[6-*O*-(a-rhamnopyranosyl)-b-galactopyranoside]	orange-red	Flower
7-*O*-methylcyanidin 3-*O*-[6-*O*-(a-rhamnopyranosyl)-b-galactopyranoside]	orange-red	Flower

Fig. 4 Scheme of flavonols and anthocyanins identified in *Catharanthus roseus*.

Red flowers exhibited higher intensity signals of organic acids than other colored flowers. It suggests that flower-colors vary because anthocyanins change their color according to the pH (Willstatter, 1915): in acidic media, the pigments are red, but turn blue in alkaline pH. Petal color changes in morning glory (*Ipomoea tricolor* cv. Heavenly Blue) from red to blue during the flower-opening period, is due to an unusual increase in vacuolar pH (pHv) from 6.6 to 7.7 in colored epidermal cells (Yoshida *et al.*, 2009).

Multivariate data analysis to find the specific metabolites correlated with flower color in different organs

In order to explore the correlation between expressed colors and metabolic profiles of other organs, ^{1}H-NMR spectrum of young leaf, old leaf, stem, and root sample were also investigated by multivariate data analysis. The PLS-DA score plot showed that the separation of young (upper) leaf samples was consistent with that of flower samples. The young leaves of red and orange flowers were clearly separated while those of pink and purple flowers were still overlapping (Fig. 5A). Also, the samples of old (lower) leaves showed the similar pattern as young leaves and flowers in the PLS-DA score plot (Fig. 5B). In the case of stems, samples from all four flower-color plants were better separated than other organs (Fig. 5C). Root samples were analyzed by PLS-DA and classified into four groups as in the case of stem samples (Fig. 5D).

From the loadings scatter plot, characteristic signals were found responsible for the separation and assigned to certain compounds (Fig. 6). Based on ^{1}H-NMR spectra, the mean peak areas of the identified compounds were quantified relative to the signal of TMSP as internal standard. The compounds were thus relatively quantified in flowers, leaves, stems and roots of the different flower-colored *C. roseus*. These data were then subjected to ANOVA to confirm the participation of different metabolites in the discrimination of organs of the different flower-colored plants. Results obtained with a statistical significance (Alpha < 0.05) were in agreement with those obtained from multivariate data analysis. Table 2 shows that each organ has its own specific metabolic characteristics.

Each of the flower color groups was dominated by a series of specific metabolites. Orange colored-flowers *C. roseus* plants exhibited a relatively high level of kaempferol in all organs except roots. Its stems, young and old leaves also showed a significantly higher level of secologanin and threonine as compared to the other three flower-colored plants. A relatively higher level of malic acid and fumaric acid occurred in the flowers and young leaves of red-colored *C. roseus* but no difference was observed for old leaves, stems and roots. The flowers, young leaves, old leaves and stems of orange and red flower-colored plants had significantly higher levels of several amino acids, such as leucine, isoleucine, threonine and asparagine if compared to purple and

pink flower-colored plants. The latter are characterized by a significantly higher level of glucose, sucrose and 2,3-DHBA not only in their flowers but also in their young leaves, old leaves and stems. Quercetin content was significantly higher in the purple flowers of *C. roseus*, but showed no difference in leaves and stems of all flower color plants. Red and purple flower-color plants also exhibited a higher level of vindoline in their flowers and young leaves. Roots of purple and pink-colored flower plants had a higher content of loganic acid than orange and red samples.

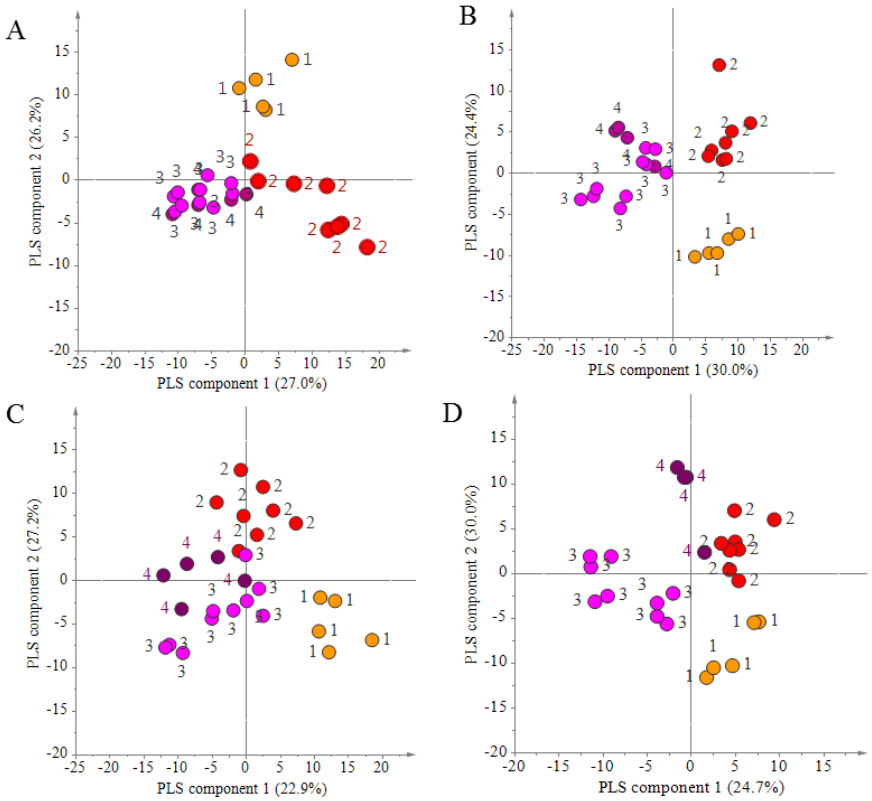


Fig. 5 PLS-DA score plots of young leaf (upper leaf, A), old leaf (lower leaf, B), stem (C), root (D). 1: orange; 2: red; 3 pink; 4 purple.

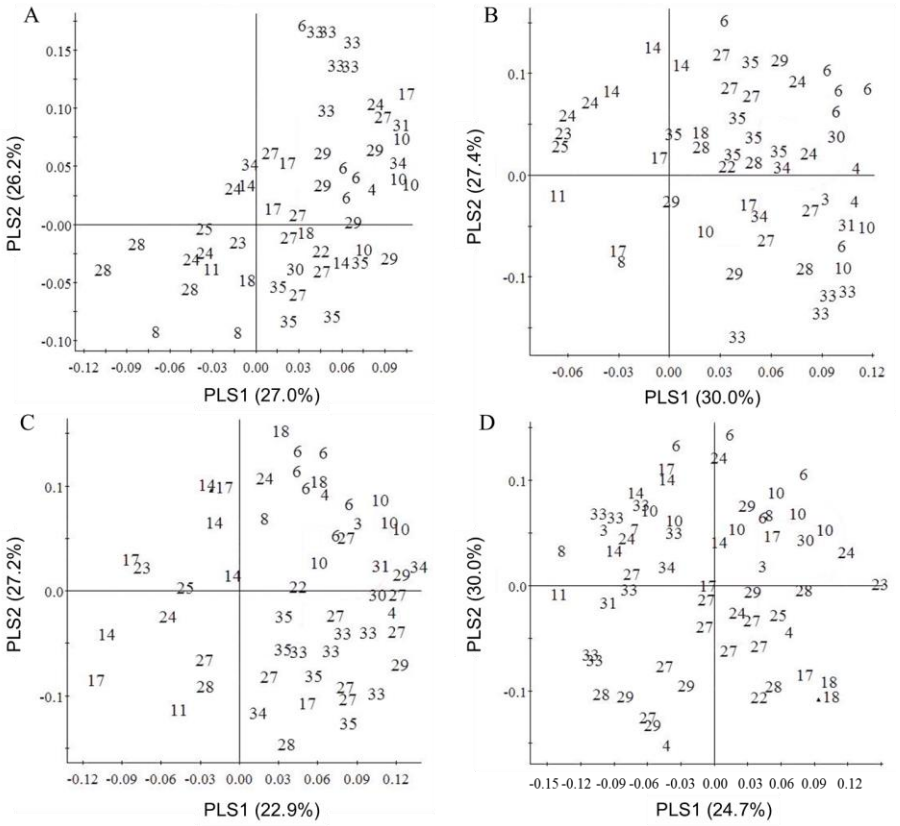


Fig. 6 PLS-DA loading plots of young leaf (upper leaf, A), old leaf (lower leaf, B), stem (C), root (D). 1: orange; 2: red; 3 pink; 4 purple.

The metabolic correlation among the different organs of different flower-colored plants showed that, as suspected, grouping was not limited to the flowers themselves but also extended to their leaves, stems and roots. A young leaf is physically the closest organ to the flower. Kaempferol derivatives are the markers which are high not only in the flowers but also in the young leaves of orange flower-bearing *C. roseus* plants. Purple and pink-flowered plants were characterized by the presence of 2,3-DHBA and this was also found in their leaves and stems. Red-flowered plants accumulated more acidic metabolites in both flowers and young leaves. The stem, which connects all organs, had no clear markers related to flower colors. Though the root is furthest away from flowers in fact there is a difference in the metabolomes of roots of the different flower-colored plants. For example, roots showed a relatively higher level of loganic acid, which was particularly significant in purple and pink-flower plants. In leaves and stems a significantly higher level of secologanin was detected in

orange-flowering plants but in the case of roots highest levels were detected in purple and pink-flowered plants. Secologanin, a monoterpenoid, derives from the plastidial MEP pathway and acts as a precursor of terpenoid indole alkaloids in *C. roseus*. The MEP pathway also leads to the biosynthesis of carotenoids, which are important compounds in orange and yellow color in flowers (Zhu *et al.*, 2010).

Table 2. Difference of marker metabolites among orange (**1**), red (**2**), pink (**3**), and purple (**4**) colors in different organs based on ^{1}H-NMR intensities.

Compounds	**Flower**				**Upper leaf**				**Lower leaf**				**Stem**				**Root**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Leucine&Isoleucine	+	+			+	+						+	+	+				+		+
Threonine						+			+	+			+					+		
Alanine	+	+							+	+			+		+					
Asparagine		+				+			+	+				+				+		+
2,3-butanediol	+																			
Quinic acid																				
Lactic acid					+				+	+			+				+		+	
Malic acid		+				+			+						+					+
Fumaric acid		+			+	+														+
Sucrose			+	+				+			+	+			+	+				+
Glucose			+	+				+			+	+			+	+			+	+
2,3-DHBA			+	+				+				+			+	+				
Quercetin-3-O-glucoside			+	+																
Kaempferol	+				+				+	+			+	+						
Loganic acid																			+	+
Secologanin					+				+				+		+				+	+
Vindoline		+		+		+		+				+	+		+					

^{1}+: significantly high (Alpha<0.05 by ANOVA).

Pigments and nectar are produced in flowers as signals to attract pollinating organisms as a strategy of increasing seed dispersion for survival (Miller *et al.*, 2011). Our data showed that flowers contain a higher level of defense-related compounds, such as 2,3-DHBA and 2,3-butanediol as compared to other organs. According to previous studies, 2,3-DHBA is formed through

the isochorismate pathway in *C. roseus* after elicitation (Budi Muljono *et al.*, 2002) and a recent report revealed that the accumulation of 2,3-DHBA depends on EDS1 (Enhanced Disease Susceptibility) in resistance responses and during the ageing of plants (Bartsch *et al.*, 2010). This compound has been reported to cause induced systemic resistance (ISR) of plants, which might play a role in the triggering of the production of other metabolites (Ryu *et al.*, 2004). As the location for plant communication with other organisms, flowers may contain more compounds related to the defense system to protect itself.

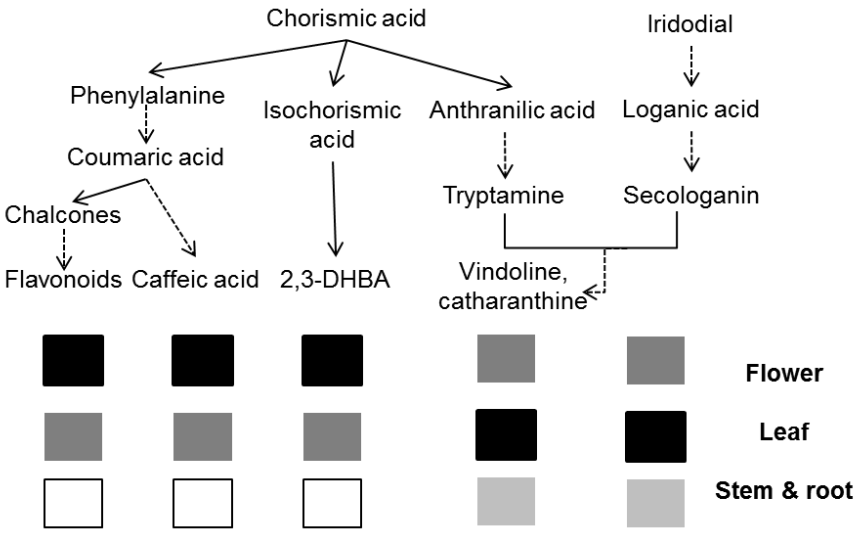


Fig. 7 Scheme of the phenylalanine pathway, isochorismic pathway, indole pathway and iridoid pathway with their intermediates. The levels of intermediates were determined by the signals integration based on ^{1}H-NMR spectra and tested by ANOVA as in table 2 (■ high level; ■ low level; ■ very low level; □ undetectable). Solid arrows represent one-step reactions; broken arrows represent multiple or uncharacterized reactions.

This study also shows the distribution of metabolites in different pathways in four organs. Three pathway branches from the chorismic acid (isochorismic pathway, phenylalanine pathway and indole pathway) and iridoid pathway are clearly involved in the different metabolic patterns observed in the four types of plants studied. The plants also differed in their primary metabolism. The data from NMR showed that the metabolites related to the phenylalanine pathway (caffeoylquinic acid, flavonols and anthocyanins) and the isochorismic pathway (2,3-DHBA) were present in higher levels while the metabolites related to the terpenoid indole alkaloid (secologanin, vindoline and catharanthine) were present in lower amounts in flowers. The situation in leaves was opposite (Fig.

7). This indicates that the phenylalanine and isochorismic pathways are competitive with the indole alkaloid pathway for various physiological functions of *C. roseus* plants.

Conclusion

This study provides the first comprehensive comparative analysis of the primary and secondary metabolite composition of flowers, leaves, stems and roots of the Pacific series of *C. roseus* with four flower colors (orange, red, purple and pink). The results revealed clear differences in the metabolome of all four organs of the different flower colored *C. roseus*. The four organs also showed specific metabolic profiles that correlate with their flower color, which might be used to predict the color the flower already in small plantlets long before flowering. NMR-based metabolomics could thus be a very useful tool for plant breeding, but first further studies wih other genotypes with different colors are needed to confirm this conclusion

References

Bar-Akiva A, Ovadia R, Rogachev I, Bar-Or C, Bar E, *et al.* (2010) Metabolic networking in *Brunfelsia calycina* petals after flower opening. *J Exp Bot* 61: 1393-1403.

Bartsch M, Bednarek P, Vivancos PD, Schneider B, von Roepenack-Lahaye E, Foyer CH, Kombrink E, Scheel D, Parker JE. 2010. Accumulation of isochorismate-derived 2,3-dihydroxybenzoic 3-*O*- -D-xyloside in *Arabidopsis* resistance to pathogens and ageing of leaves. *J Biol Chem* 285: 25654-2565.

Brun G, Dijoux MG, David B, Mariotte AM. 1999. A new flavonol glycoside from *Catharanthus roseus*. *Phytochemistry* 50: 167-169.

Budi Muljono RA, Scheffer JJC, Verpoorte R. 2002. Isochorismate is an intermediate in 2,3-dihydroxybenzoic acid biosynthesis in *Catharanthus roseus* cell cultures. *Plant Physiol Biochem* 40: 231-234.

Choi YH, Tapias EC, Kim HK, Lefeber AW, Erkelens C, Verhoeven JT, Brzin J, Zel J, Verpoorte R. 2004. Metabolic discrimination of *Catharanthus roseus* leaves infected by phytoplasma using ^{1}H-NMR spectroscopy and multivariate data analysis. *Plant Physiol* 135: 2398-2410.

Chung IM, Ahmad A, Ali M, Lee OK, Kim MY, Kim JH, Yoon DY, Peebles CA, San KY. 2009. Flavonoid glucosides from the hairy roots of *Catharanthus roseus*. *J Nat Prod* 72: 613-620.

Devireddy L, Hart D, Goetz D, Green M. 2010. A mammalian siderophore synthesized by an enzyme with a bacterial homolog involved in enterobactin production. *Cell* 141: 1006-1017.

Ferreres F, Pereira DM, Valentão P, Andrade PB, Seabra RM, Sottomayor M. 2008. New phenolic compounds and antioxidant potential of *Catharanthus roseus*. *J Agric Food Chem* 56: 9967-9974.

Filippini R, Caniato R, Piovan A, Cappelletti EM. 2003. Production of anthocyanins by *Catharanthus roseus*. *Fitoterapia* 74: 62-67.

Forsyth WGC, Simmonds NW. 1957. Anthocyanidins of *Lochnera rosea*. *Nature* 180: 247.

Jaleel CA, Gopi R, Panneerselvam R, 2008. Growth and photosynthetic pigments responses of two varieties of *Catharanthus roseus* to triadimefon treatment. *C R Biologies* 331: 272-277.

Kim HK, Choi YH, Verpoorte R. 2010. NMR based metabolomic analysis of plants. *Nat Protoc* 5: 536-548.

Kim HK, Saifullah, Khan S, Wilson EG, Kricun SD, Meissner A, Goraler S, Deelder AM, Choi YH, Verpoorte R. 2010. Metabolic classification of South American Ilex species by NMR-based metabolomics. *Phytochemistry* 71: 773-784.

Knobloch KH, Bast G, Berlin J. 1982. Medium and light induced formation of serpentine and anthocyanins in cell suspension cultures of *Catharanthus roseus*. *Phytochemistry* 21: 591-594.

Malien-Aubert C, Dangles O, Amiot MJ. 2001. Color stability of commercial anthocyanin- based extracts in relation to the phenolic composition. Protective effects by intra and intermolecular copigmentation. *J Agric Food Chem* 49: 170-176.

Markham KR, Ofman DJ. 1993. Lisianthus flavonoid pigments and factors influencing their expression in flower colour. *Phytochemistry* 34: 679-685.

Miller R, Owens SJ, Rørslett B. 2009. Plants and colour: Flowers and pollination. *Opt Laser Technol* 43: 282-294.

Mulder-Krieger Th, and Verpoorte R. 1994. Anthocyanins as flower pigments. Possibilities for flower colour modification. Kluwer Academic Publishers, Dordrecht, 1994, 154 pp

Mustafa NR, Verpoorte R. 2007. Phenolic compounds in *Catharanthus roseus*. *Phytochem Rev* 6: 243-258.

Nishihara M, Nakatsuka T. 2011. Genetic engineering of flavonoid pigments to modify flower color in floricultural plants. *Biotechnol Lett* 33: 433-441.

Pereira DM, Ferreres F, Oliveira J, Valentão P, Andrade PB, Sottomayor M. 2009. Targeted metabolite analysis of *Catharanthus roseus* and its biological potential. *Food Chem Toxicol* 47: 1349-1354.

Piovan A, Filippini R, Favretto D. 1998. Characterization of the anthocyanins of *Catharanthus roseus* (L.) G. Don *in vivo* and *in vitro* by electrospray ionization ion trap mass spectrometry. *Rapid Commun Mass Spectrom* 12: 361-367.

Robinson GM, Robinson R. 1931. A survey of anthocyanins. I. *Biochem* XXV: 1687-1705.

Roytrakul S, Verpoorte R. 2007. Role of vacuolar transporter proteins in plant secondary metabolism: *Catharanthus roseus* cell culture. *Phytochem Rev* 6: 383-396.

Ryu CM, Farag MA, Hu CH, Reddy MS, Kloepper JW, Paré PW. 2004. Bacterial volatiles induce systemic resistance in Arabidopsis. *Plant Physiol* 134: 1017-1026.

Seigler DS. 1995. *Plant secondary metabolism*. Kluwer Academic Publishers: London.

Tanaka Y, Brugliera F, Kalc G, Senior M, Dyson B, Nakamura N, Katsumoto Y, Chandler S. 2010. Flower color modification by engineering of the flavonoid biosynthetic pathway: practical perspectives. *Biosci Biotechnol Biochem* 74: 1760-1769.

Tanaka Y, Ohmiya A. 2008. Seeing is believing: engineering anthocyanin and carotenoid biosynthetic pathways. *Curr Opin Biotechnol* 19: 190-197.

Toki K, Saito N, Irie Y, Tatsuzawa F, Shigihara A, Honda T. 2008. 7-*O*-Methylated anthocyanidin glycosides from *Catharanthus roseus*. *Phytochemistry* 69: 1215-1219.

Weissenberg M. Levy A, Schaeffler I, Levy EC.1988. High-performance liquid chromatographic analysis of the ajmalicine distribution in roots of

Catharanthus roseus lines with different flower colours. *J Chromatogr A* 452: 485-490.

Wilkinson M, Sweeney JG, Lacobucci GA. 1977. High performance liquid chromatography of anthocyanins. *J Chromatogr* 132: 349-351.

Willstätter R & Mallison H. 1915. Untersuchungen über die Anthocyane. X. Über Variationen der Blütenfarben. *Justus Liebigs Ann Chem* 408: 147-162.

Yang SO, Kim SH, Kim Y, Kim HS, Chun YJ, Choi HK. 2009. Metabolic discrimination of *Catharanthus roseus* calli according to their relative locations using ^{1}H-NMR and principal component analysis. *Biosci Biotechnol Biochem* 73: 2032-2036.

Yoshida K, Mori M, Kondo T. 2009. Blue flower color development by anthocyanins: from chemical structure to cell physiology. *Nat Prod Rep* 26: 857-964.

Yoshida K, Toyama-Kato Y, Kameda K, Kondo T. 2003. Sepal color variation of Hydrangea macrophylla and vacuolar pH measured with a proton-selective Microelectrode. *Plant Cell Physiol* 44: 262-268.

Zhu C, Bai C, Sanahuja G, Yuan D, Farré G, Naqvi S, Shi L, Capell T, Christou P. 2010. The regulation of carotenoid pigmentation in flowers. *Arch Biochem Biophys* 504: 132-14