

Progress in NMR-based metabolomics of *Catharanthus roseus*

Qifang PAN, Jingya ZHAO, Yuliang WANG, Kexuan TANG (✉)

Joint International Research Laboratory of Metabolic & Developmental Sciences, Key Laboratory of Urban Agriculture (South), Ministry of Agriculture, Plant Biotechnology Research Center, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract Metabolomics has been rapidly developed as an important field in plant sciences and natural products chemistry. As the only natural source for a diversity of monoterpene indole alkaloids (MIAs), especially the low-abundance antitumor agents vinblastine and vincristine, *Catharanthus roseus* is highly valued and has been studied extensively as a model for medicinal plants improvement. Due to multistep enzymatic biosynthesis and complex regulation, genetic modification in the MIA pathway has resulted in complicated changes of both secondary and primary metabolism in *C. roseus*, affecting not only the MIA pathway but also other pathways. Research at the metabolic level is necessary to increase knowledge on the genetic regulation of the whole metabolic network connected to MIA biosynthesis. Nuclear magnetic resonance (NMR) is a very suitable and powerful complementary technique for the identification and quantification of metabolites in the plant matrix. NMR-based metabolomics has been used in studies of *C. roseus* for pathway elucidation, understanding stress responses, classification among different cultivars, safety and quality controls of transgenic plants, cross talk between pathways, and diversion of carbon fluxes, with the aim of fully unravelling MIA biosynthesis, its regulation and the function of the alkaloids in the plant from a systems biology point of view.

Keywords *Catharanthus roseus*, monoterpene indole alkaloids, NMR, metabolomics

1 Introduction

In the postgenomic era, metabolomics is the latest tool for functional genomics^[1]. As a fast growing powerful technology, metabolomics is useful for phenotyping and diagnostic analyses of plants, and is rapidly becoming a key tool in functional annotation of genes and in the comprehensive understanding of the cellular response to various biological conditions^[2]. Metabolomics covers metabolic profiling, fingerprinting, footprinting and metabolic flux analysis, which can be used to measure the effect of developmental stage, environment, daily and seasonal changes and stress on the plant metabolome, as well as to assess the natural variance in metabolite content between individual plants. Also, metabolomics is a powerful approach with great potential for the improvement of the compositional quality of crops, for characterizing and identifying cultivars, chemotaxonomy and quality control of plant products (e.g., food, medicinal plants). Moreover, metabolomics in combination with transcriptomics has become a major tool for providing extensive information for use in functional genomics. A good illustration of this is the elucidation of the iridoid pathway in *Catharanthus roseus*^[3].

Catharanthus roseus (Madagascar periwinkle), belonging to the Apocynaceae family, is a medical plant of great pharmaceutical interest for its capacity to biosynthesize a great variety of MIAs (>130), which have a high economic value due to their wide spectrum of pharmaceutical applications. Besides the most well-known bisindole alkaloids (vinblastine and vincristine), *C. roseus* also produces ajmalicine used as an antihypertensive and serpentine used as a sedative. The trace amounts of the dimeric alkaloids in *C. roseus* and the difficulty in extracting and purifying them explain the high costs of these MIAs. Although total chemical synthesis of these complex alkaloids is of academic interest, this is not likely

to be applied commercially due to the low yields. However, the dimerization reaction coupling vindoline and catharanthine, which in the plant is catalyzed by a peroxidase, has been mimicked chemically and is now used to couple the much more abundant monomers. The *in vitro* production systems using plant cell cultures or hairy roots of *C. roseus* have been developed but failed to synthesize vindoline, one of the precursors needed for producing the bisindole alkaloids. To develop novel sources of these compounds, e.g., by synthetic biology, requires thorough knowledge of all genes, enzymes, and intermediates in the MIA biosynthetic pathway and the underlying regulation mechanisms. In the past decades there have been extensive efforts and in-depth studies on MIA biosynthesis in *C. roseus* by numerous groups of researchers across the globe. Nowadays the “omics” tools, such as genomics, transcriptomics, proteomics and especially metabolomics, can provide us with enormous amounts of information about the genes, enzymes, transcription factors, intermediates, pathways, and compartmentation of MIA biosynthesis in *C. roseus* cell cultures, hairy roots and plants. These tools will be very helpful in order to clarify some unresolved parts of the iridoid pathway, catharanthine biosynthesis, transport, and the signal-transduction and regulation of the pathway via transcription factors such as the octadecanoid-responsive *Catharanthus* AP2-domain (ORCA). Metabolomic analysis combined with other tools is being used to rapidly narrow down candidate genes for further functional investigation of their biochemical roles^[4].

Nuclear magnetic resonance spectroscopy (NMR) is a very powerful method that allows the simultaneous detection of diverse groups of secondary metabolites (such as flavonoids, alkaloids and terpenoids) besides the abundant primary metabolites (such as sugars, organic acids and amino acids)^[5]. The non-selectiveness of NMR makes it an ideal tool for unbiased plant metabolomics studies. NMR is also a very useful technique for structure elucidation using various 2D NMR measurements without further fractionation of the extract. As signals are proportional to their molar concentration in an NMR spectrum, it is possible to make a direct comparison of concentrations of all compounds without the need for calibration curves for each individual compound. Moreover, the analysis time is short and the sample preparation is simple and rapid, enabling the analysis of large numbers of samples per hour. These are major advantages of NMR compared with the more sensitive methods as MS, LC-MS and GC-MS, which all suffer from lack of absolute quantitative data. NMR application in metabolomics, however, is limited by its sensitivity and signal overlap, which are greatly improved by the recent developments of NMR hardware and two-dimensional NMR^[5]. These two approaches offer a choice of data quality: a large number of metabolites with only relative quantitation for each

compound, or smaller number of metabolites with full quantitation. However, one should keep in mind that in all present metabolomics methods, the visible metabolome is determined, or perhaps better to say limited, by the method of extraction. Only soluble compounds might be visible, and poorly soluble compounds will always be present up to saturation levels, thus not showing any variation above that level, even when large differences in accumulation may occur in the plant. As the first choice for metabolomics, NMR-spectroscopy has been applied to plant metabolomic studies of *C. roseus* for many purposes (Table 1), i.e., identification of novel metabolites, elucidation of metabolic pathways, metabolic responses to stress, metabolic characterization and classification, and metabolic flux analysis. Thus dozens of primary and secondary metabolites have been identified in cell cultures, hairy roots and plants of *C. roseus*^[16].

The latest developments in the studies of the biosynthesis of MIA and its regulation in *C. roseus* by metabolomic studies are reviewed in the present update.

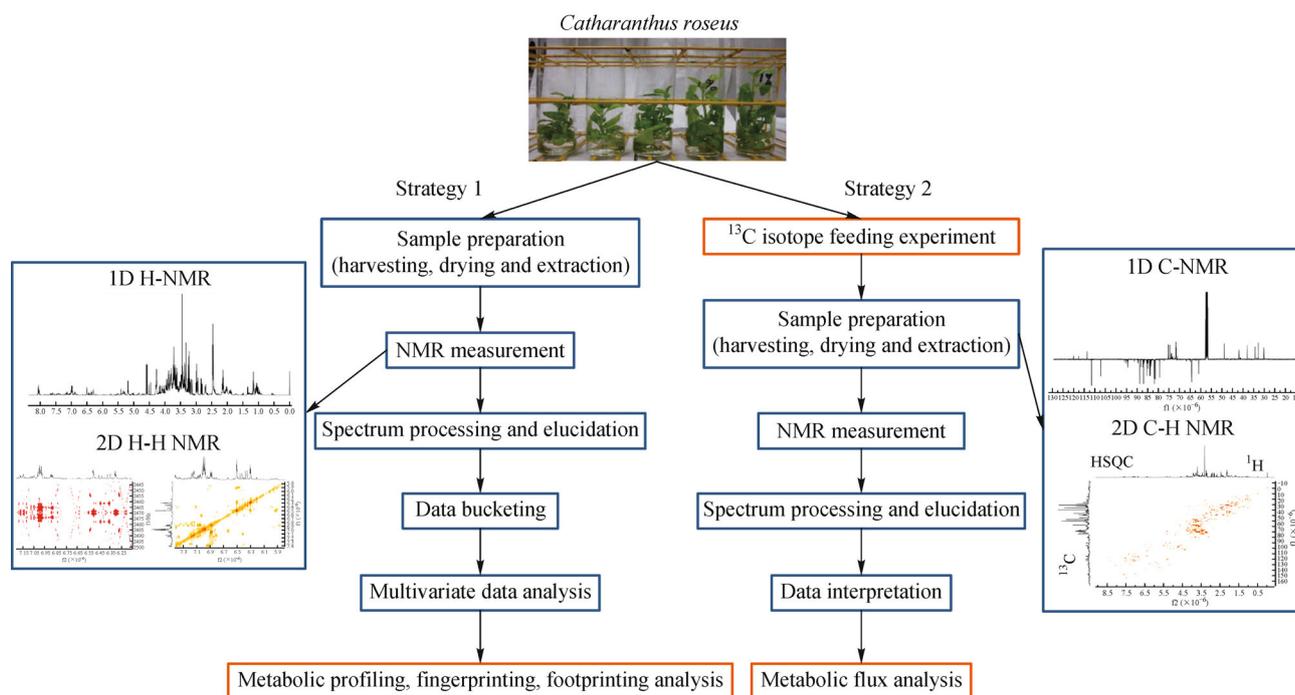
2 Strategies for NMR-based metabolomic study of *Catharanthus roseus*

2.1 Metabolite profiling, fingerprinting and footprinting using ¹H-NMR combined with multivariate data analysis

Metabolite profiling, fingerprinting and footprinting are commonly used as efficient methods in metabolomics studies (Fig. 1). Profiling aims to analyze quantitatively sets of metabolites in a selected biochemical pathway, or a specific class of compounds. Fingerprinting uses high throughput qualitative screening of the metabolite content of an organism or tissue with the primary aim of sample comparison and discrimination analysis. Footprinting is the fingerprinting analysis of metabolites that are excreted by cells to the culture medium. In all these approaches, ¹H-NMR is sufficient to generate metabolomics data of a sample within a relatively short time (5–10 for 64–128 scans)^[5] (Fig. 2). To accurately assign peaks to metabolites, 2D NMR experiments, such as *J*-resolved and COSY, are generally used in conjunction with compound identification in plant crude mixture/extracts^[17]. After metabolite identification, multivariate data analysis of ¹H-NMR spectra assists in recognizing patterns and finding discriminating signals^[18] (Fig. 3). Multivariate or pattern recognition techniques such as the well-described unsupervised principal component analysis (PCA), hierarchical cluster analysis (HCA), supervised partial least squares-discriminant analysis (PLS-DA) and orthogonal projections to latent structures (OPLS) are useful tools to analyze complex data sets^[1]. PCA and PLS-DA are widely applied regression methods used to reduce the multidimensionality of the

Table 1 Different types of NMR applied to identification of metabolites in *Catharanthus roseus*

Materials	Identified metabolites	NMR application	Publication
Stem	Syringetin-3- <i>O</i> -robinobioside	$^1\text{H-NMR}$, $^{13}\text{C-NMR}$	[6]
Leaf	Glutamates, glucose, polyphenols, succinic acid, sucrose, chlorogenic acid, loganic acid, secologanin, and vindoline	$^1\text{H-NMR}$, <i>J</i> -resolved, COSY, HMBC	[7]
Hairy root	Lysine, proline, tyrosine, hydroxyacetone, levulinic acid	HSQC, TOCSY	[8]
Flower	Anthocyanidin-1-galactose, anthocyanidin-1-rhamnose, anthocyanidin-2-galactose, anthocyanidin-2-rhamnose	$^1\text{H-NMR}$, $^{13}\text{C-NMR}$	[9]
Root	Flavonoid glucosides	$^{13}\text{C-NMR}$, COSY, HETCOR, HSQC	[10]
Calli	Valine, lactic, threonine, alanine, arginine, acetic acid, glutamic acid, malic acid, succinic acid, citric acid, asparagine, 2-oxoglutaric acid, choline, oxalacetic acid, sucrose, lactose, glucose, fumaric acid, phenylalanine, tryptophan, formic acid	$^1\text{H-NMR}$	[11]
Cell cultures	Leucine, isoleucine, valine, threonine, alanine, glutamic acid, aspartic acid, glucose, sucrose, fumaric acid, phenylalanine, salicylic acid, gentisic acid glucoside (2,5-dihydroxybenzoic acid glucoside)	$^1\text{H-NMR}$, <i>J</i> -resolved, COSY	[12]
Cell cultures	Salicylic acid, 2,3-DHBA	$^{13}\text{C-NMR}$, HSQC	[13]
Flower, leaf, stem, root	Isoleucine, leucine, valine, threonine, alanine, arginine, glutamic acid, glutamine, aspartic acid, asparagine, serine, 2,3-butanediol, quinic acid, lactic acid, acetic acid, malic acid, citric acid, ketoglutaric acid, succinic acid, oxalacetic acid, fumaric acid, sucrose, α -glucose, β -glucose, choline, chlorogenic acid, 2,3-DHBA, quercetin analogs, kaemferol analogs, loganic acid, secologanin, catharanthine, vindoline, vindolinine, 4- <i>O</i> -caffeoyl quinic acid	$^1\text{H-NMR}$, <i>J</i> -resolved, COSY	[14,15]

**Fig. 1** Scheme of NMR-based metabolomics analysis in *Catharanthus roseus*

metabolomics data. They provide an excellent platform to study, for example, the stress response in plants; quality

control and authentication of medicinal plants (like *Artemisia annua*, *Angelica acutiloba*, and *Panax*

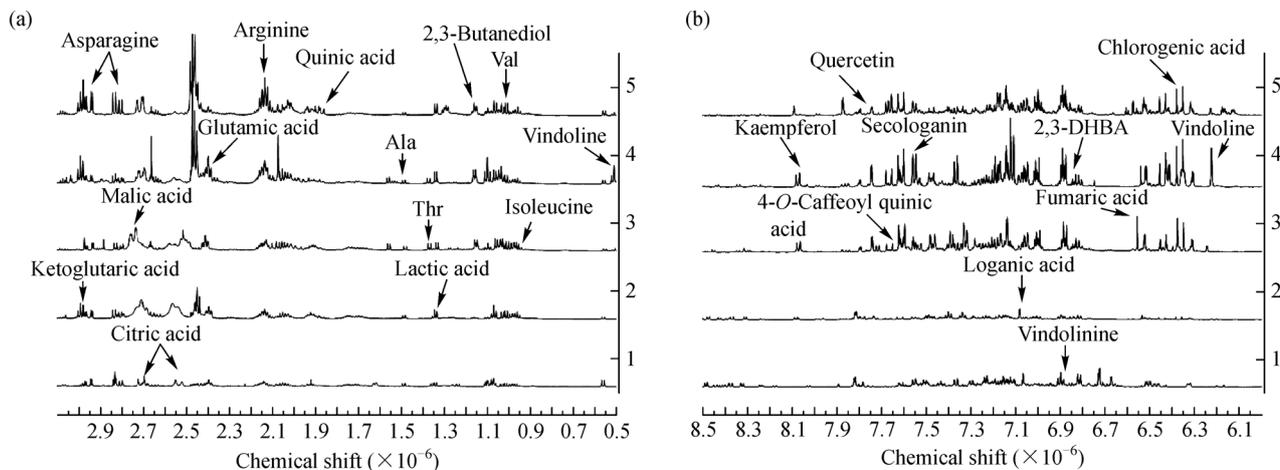


Fig. 2 ¹H-NMR spectra of extracts of different organs of *Catharanthus roseus* with the assigned metabolites. 1, root; 2, stem; 3, lower leaf; 4, upper leaf; 5, flower. (a) Aliphatic amino acids in the range of δ 0.5–3.1; (b) aromatic metabolites in the range of δ 6.0–8.5.

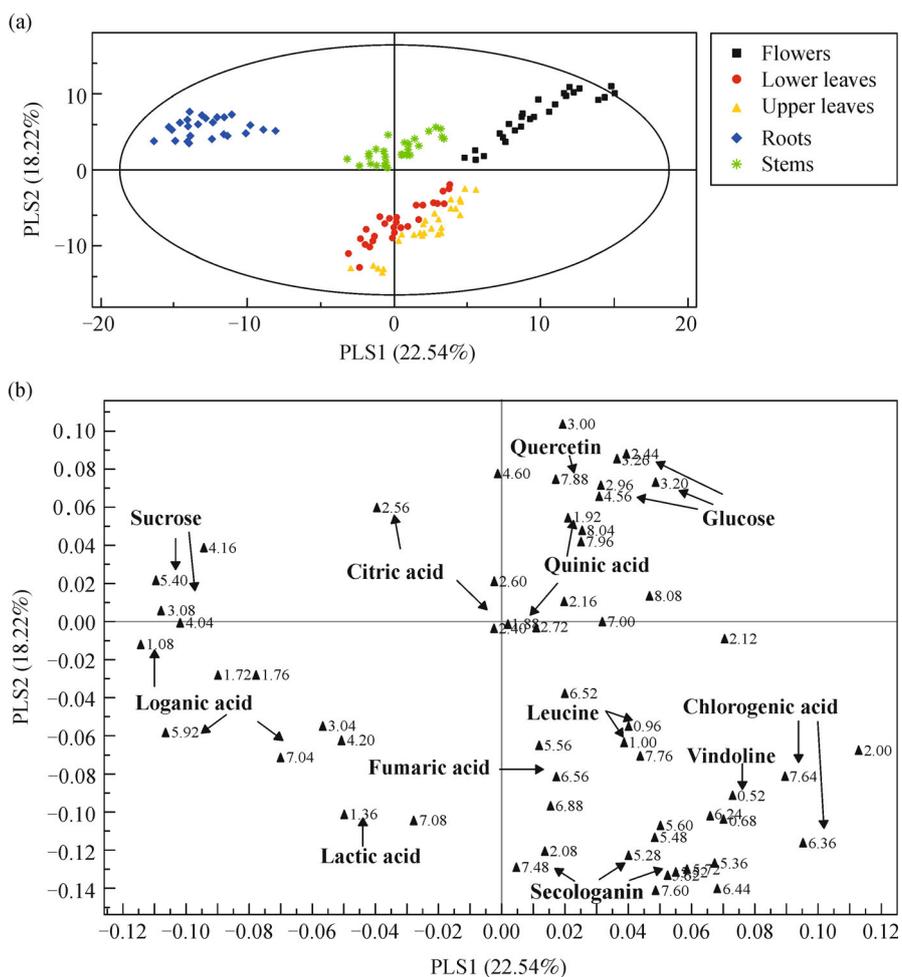


Fig. 3 ¹H-NMR investigation and multivariate data analysis of different organs in *C. roseus*. (a) PLS-DA score plot, in which blue color represents root, green color represents stem, red color represents old leaf, orange color represents young leaf, and black color represents flower; (b) PLS-DA loading plot.

notoginseng), classification of different plant species genotypes or ecotypes, identification of biomarkers for disease diagnosis, or identification of bioactive compounds in plants. The unsupervised methods show the maximum separation between all samples in two or three dimensions. If this does not result in any clear grouping of the samples, supervised methods can be applied to reveal possible characteristic differences between certain defined classes. In fact metabolomics has become the tool for systems biology, studying the response of the whole system, rather than a reductionist approach in which only a few parameters are measured. The metabolome of a plant, in fact, is the sum of a mixture of the metabolomes of different tissues and even different cells. One may thus speak about a macrometabolome and a micrometabolome, and even nanometabolome if one considers the role of cellular compartments in the cellular metabolism. The localization and distribution of MIA biosynthesis in different cells and cellular compartments is clearly important and a metabolomic analysis of a single cell, or at least single cell type, would be of great value in understanding the biosynthesis. The first experiments in this direction have already been made, a clear example being the analysis of epidermis cells^[19,20].

2.2 Metabolic flux analysis based on ¹³C labeling experiment

One of the problems of metabolomics is that it is like a two-dimensional picture, it measures the amounts of compounds present at a certain time point, but it does not tell anything about the turnover. A major compound can be a stored product or part of a very active metabolic

pathway, only measuring the dynamics of the system can give the answer that means measuring the flux through pathways, making a film, rather than a picture.

Metabolic flux analysis (MFA), the quantification of all intracellular fluxes in an organism, is thus an important cornerstone of metabolic engineering and systems biology (Fig. 1). Each flux reflects the function of a specific pathway within the network. As all biological activity is related to metabolic activity, it is these fluxes that deliver the phenotype of an organism^[21]. Flux measurements complement transcriptomic, proteomic, and metabolomic technologies in defining phenotypes, and provide a useful complementary parameter for the system-wide characterization of metabolic networks. MFA of different phenotypes in plants can provide valuable information, which facilitates the selection of metabolic engineering targets, elucidation of metabolic pathways, and construction of metabolic models^[22]. Metabolic flux analysis is usually carried out using ¹³C-NMR and ¹H-¹³C HSQC NMR analysis in experiments where ¹³C isotope-labeled compounds are added to the organisms studied. The NMR approach enables the determination not only of the overall percentage of labeling but also of the site of incorporation (Fig. 4). The data are analyzed and interpreted by mathematical models and software like ¹³C-FLUXTM and 4F^[23]. The ¹³C isotope is widely used since it is not radioactive and NMR analysis enables determination of the precise site of the label in a molecule. Natural abundance of ¹³C is 1.1%, so already a labeling of 1.1% will lead to a doubling of the percentage of the labeled carbons and consequently to a clear increase of the signal concerned. ¹³C labeling experiments and (¹³CLE)-based MFA have been applied to *C. roseus* for pathway elucidation, finding

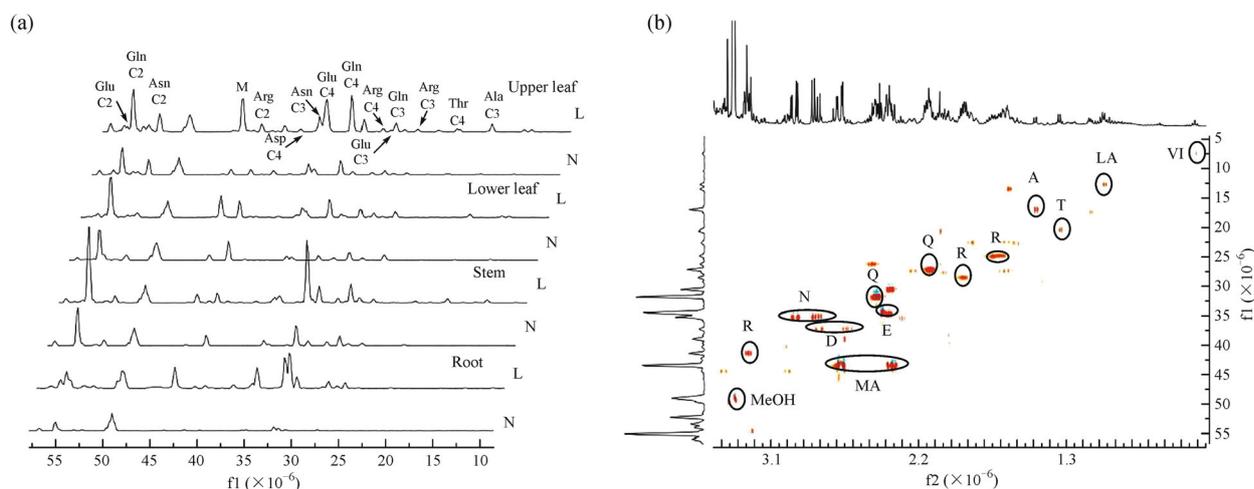


Fig. 4 ¹³C dimension of HSQC spectra of amino acids (δ 10×10^{-6} – 55×10^{-6}) in different organs (a) and 2-D [¹³C, ¹H] HSQC spectrum displaying amino acid resonances in leaves (b) of *Catharanthus roseus* after feeding [^{1-¹³C}] glucose. (a) L, labeled samples; N, non-labeled samples; M, malate; C2/3/4, the position of carbon in metabolites; (b) A, alanine; D, aspartate; E, glutamate; T, threonine; N, asparagine; Q, glutamine; R, arginine; LA, loganic acid; MA, malate; VI, vindoline; MeOH, methanol.

cross links between pathways, and flux quantification in central carbon metabolism.

3 NMR-based metabolomic applications to *Catharanthus roseus* plants, cells and hairy roots

3.1 Monitoring the response to stress of biotic or abiotic origin of *Catharanthus roseus*

The response of plants to elicitors, hormones, stress, wounding, herbivory or infection causes change in a wide range of metabolites^[2]. To understand the stress response in *C. roseus*, a comprehensive metabolomic profiling of leaves infected by 10 types of phytoplasmas was carried out using one-dimensional and two-dimensional NMR spectroscopy followed by PCA^[7]. The results showed that the major factors for discriminating phytoplasma-infected *C. roseus* leaves from healthy ones were increases of metabolites related to the biosynthetic pathways of phenylpropanoids and MIAs, chlorogenic acid, loganic acid, secologanin, and vindoline. Furthermore, a greater abundance of glycine, glucose, polyphenols, succinic acid, and sucrose was detected in the phytoplasma-infected leaves. ¹H-NMR-based metabolomics analysis can differentiate between inner and outer calli of *C. roseus*, monitoring with various elicitors in solid-state cultures^[11]. The cells with different localization patterns in the calli treated with silver nitrate and methyl jasmonate could be separated in PCA score plots. The levels of valine, threonine, alanine, asparagine, phenylalanine, tryptophan, choline, lactose, lactic acid, acetic acid, malic acid, succinic acid, citric acid, fumaric acid and formic acid were found to be higher in the inner callus than in the outer callus, whereas 2-oxoglutaric acid, oxalacetic acid, sucrose and glucose dominated in the outer callus. The effect of salicylic acid (SA) on the metabolic profile of *C. roseus* cell cultures in a time course (0, 6, 12, 24, 48 and 72 h after treatment) was studied using ¹H-NMR spectroscopy and PCA^[13]. Adding 25 μmol of sodium SA into 100 mL of 5 days old cell cultures altered the metabolome compared with the non-treated cells. A dynamic change in amino acids, phenylpropanoids, and tryptamine was found in cells at 48 h after SA treatment. Additionally, 2,5-dihydroxybenzoic-5-*O*-glucoside was detected only in SA-treated cells^[13].

3.2 Classification and characterization of different phenotypes or genotypes of *Catharanthus roseus*

Classification and characterization of plant products are now major areas of interest. Using NMR-based metabolomic approaches, it has been possible to discriminate different species of plants and identify biomarkers suitable for their discrimination^[18]. Based on the ¹H-NMR

metabolite profiles, eight cultivars of *C. roseus* plants could be discriminated genetically^[24]. Hierarchical dendrograms based on the NMR data for regions of *C. roseus* aromatic compounds were in general agreement with the genetic relationships determined by standard DNA fingerprinting methods. According to the signal assignment of the ¹H-NMR spectra, secologanin and polyphenols contributed most to the discrimination between cultivars^[24]. In another study, ¹H-NMR and multivariate data analysis were used to characterize the metabolites and investigate the metabolic profiles of leaves, stems, roots and flowers of *C. roseus* with four flower colors (orange, pink, purple and red)^[14]. The results showed that flower color is characterized by a special pattern of metabolites such as anthocyanins, flavonoids, organic acids and sugars. Not only the flowers but also the leaves, stems, and roots showed metabolic differences correlating with flower color. Most importantly, it seems possible to predict flower color through profiling the metabolites in leaves, stems, or roots, which may be a helpful tool for plant breeding.

3.3 Discrimination of wild/transgenic plants for quality and safety control

Transgenic crops are widespread in the agro-economy world but are also highly contentious because of a risk of possible unintended effects or unpredictable changes in plants. Metabolomic profiling via NMR can make broad and deep assessments of food quality and content^[25]. To better understand the effect of genetic engineering on plant metabolism, transgenic *C. roseus* plants overexpressing ORCA3 alone (OR lines), or co-overexpressing G10H and ORCA3 (GO lines) were investigated by metabolomics^[15]. ¹H-NMR-based metabolomics confirmed the higher accumulation of monomeric indole alkaloids (strictosidine, vindoline, catharanthine and ajmalicine) in OR and GO lines. Moreover, multivariate data analysis of ¹H-NMR spectra showed a clear separation between transgenic and control lines, which was determined by a change of amino acids, organic acids, sugars and phenylpropanoids levels in both OR and GO lines compared to the controls. The results indicate that enhancement of MIA biosynthesis by ORCA3 and G10H overexpression might also affect other metabolic pathways in the metabolism of *C. roseus* plants (Fig. 5).

3.4 Elucidation of the MIA Pathway based on feeding experiment

In *C. roseus*, mevalonate was at first considered to be the exclusive precursor of isopentenyl diphosphate in the biosynthesis of secologanin. However, later research indicated that the alternative MEP pathway might be involved. A feeding experiment using [1-¹³C] glucose to *C. roseus* cell cultures followed by analysis of its incorporation into secologanin using ¹³C NMR spectroscopy was performed. The data on the sites of incorporation of the ¹³C

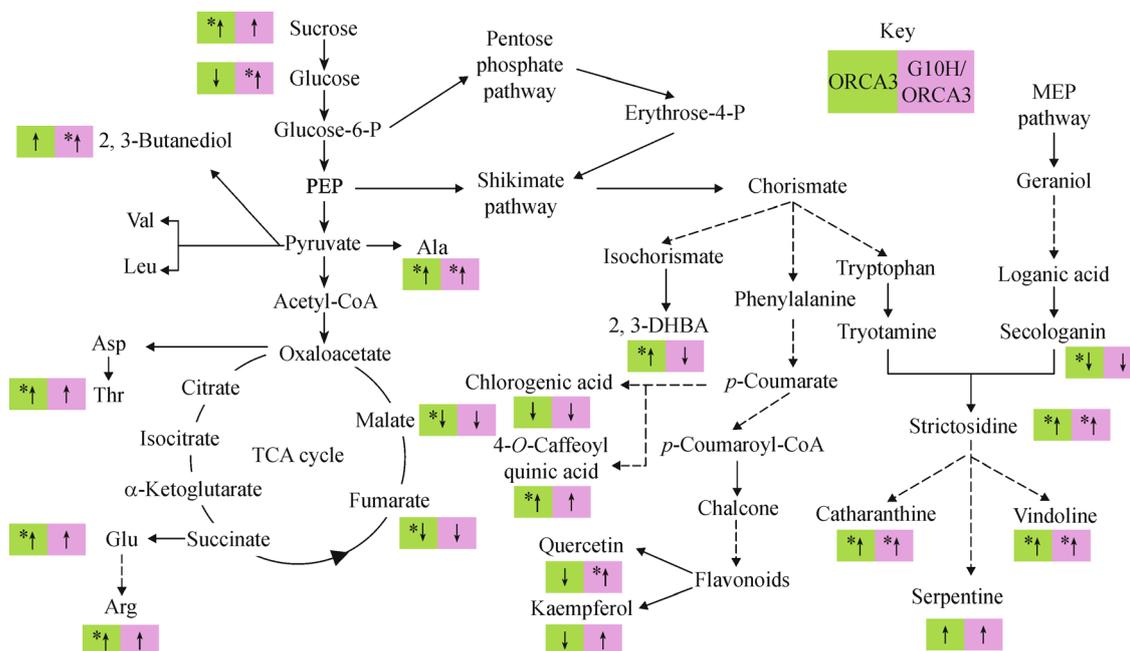


Fig. 5 Schematic effects of ORCA3, or G10H, overexpression on the metabolism of *Catharanthus roseus* plants based on NMR spectrum. Green box shows ORCA3 overexpression (the OR lines) and pink box is for G10H and ORCA3 co-overexpression (the GO lines). The up arrow in the box represents the increase of metabolite content. The down arrow in the box represents the decrease of metabolite content. Arrows with star in the box represent significant difference ($P < 0.05$ by ANOVA) of metabolite content compared with the controls^[15].

label showed that the MEP pathway and not the mevalonate pathway was the major route for secologanin biosynthesis^[26]. The biosynthetic pathways of SA and 2,3-DHBA were studied using a similar feeding-NMR method. The data led to the conclusion that the isochorismate pathway is responsible for the biosynthesis of both compounds, presenting the first full chemical evidence for the isochorismate pathway for the biosynthesis of SA as an important signal molecule in plants^[12,27].

3.5 Analysis of metabolic network fluxes in *Catharanthus roseus*

To assess quantitatively the crosstalk between the MEP and the mevalonate pathways, $[2-^{13}\text{C}_1]$ mevalonolactone or $[\text{U}-^{13}\text{C}_6]$ glucose were supplied to *C. roseus* cell cultures grown in light or dark^[28]. The incorporation of exogenous $[2-^{13}\text{C}_1]$ mevalonolactone into the DMAPP and IPP precursors of sitosterol and lutein were 48% and 7% respectively. With $[\text{U}-^{13}\text{C}_6]$ glucose as precursor, at least 95% of sitosterol precursors were obtained from the mevalonate pathway, whereas phytol appeared to be biosynthesized via the deoxyxylulose phosphate pathway (about 60%) as well via the mevalonate pathway (about 40%).

Hairy roots of *C. roseus*, a pharmaceutically significant system for production of plant compounds and an important metabolic engineering target, were used as a model system in the study of CLE-based MFA. $[\text{U}-^{13}\text{C}_6]$

glucose was fed to the hairy roots of *C. roseus* to investigate its elemental and biomolecular composition, in which the abundances of lipids, lignin, cellulose, hemicellulose, starch, protein, proteinogenic amino acids, mineral ash, and moisture were quantified^[29]. Moreover, 12 biomass synthetic fluxes relating to the metabolic map of the plant system of *C. roseus* hairy roots were precisely calculated. The results highlighted the flux of carbon from β -glucose consumed in the hairy roots into various products, which could enable the design of metabolic engineering strategies to divert carbon to the economically attractive MIAs^[29]. The application of bondomers (isomers of a metabolite differing in the connectivity of their C–C bonds) was introduced to MFA study as a computationally alternative to the isotopomer concept in *C. roseus*^[8]. Hairy roots were cultured on (5% w/w $[\text{U}-^{13}\text{C}_6]$, 95% w/w naturally abundant) sucrose. HSQC and COSY spectra of the hydrolyzed aqueous extract were acquired from the hairy roots. Analysis of these spectra yielded a data set of 116 bondomers of beta glucans and proteinogenic amino acids from the hairy roots. Fluxes were evaluated from the bondomer data by using comprehensive bondomer balancing, most of which were identified with precision in a three-compartment model of central carbon metabolism. Pentose phosphate pathways were observed to occur in parallel in the cytosol and plastids with significantly different fluxes. The fluxes between phosphoenolpyruvate and oxaloacetate in the cytosol and between malate and pyruvate in the mitochon-

dria were relatively high (60.1 ± 2.5 mol per 100 mol sucrose uptake, or 22.5 ± 0.5 mol per 100 mol mitochondrial pyruvate dehydrogenase flux).

The development of a comprehensive flux analysis tool for the plant system of *C. roseus* is expected to be valuable in assessing the metabolic impact of genetic or environmental changes and in controlling the fluxes to targeted metabolites during biosynthesis.

3.6 Combination with other omics tools for gene discovery and engineering MIA pathway

Integration of multiple omics with the analysis of various traits of plants is used to predict gene functions and characterize the complex interaction and coordination of plant metabolic networks in biological processes from a system biological point of view^[30]. Combination of non-targeted approaches, such as transcriptomics and metabolomics, can reveal potential gene-to-metabolite networks^[31], filter out candidate genes for certain metabolic pathways^[32], and suggest gene functions by overexpression^[33]. The integration of omics approaches can help to reveal the organization of the whole system and thus to identify interesting targets for further studies. A comprehensive profiling analysis of *C. roseus* was performed by combining genome-wide transcript profiling of cDNA-amplified fragment-length polymorphism with metabolic profiling of elicited *C. roseus* cell cultures to yield a collection of known and previously undescribed transcript tags and metabolites associated with MIAs^[34]. Previously undescribed gene-to-gene and gene-to-metabolite networks were drawn up by searching for correlations between the expression profiles of 417 gene tags and the accumulation profiles of 178 metabolite peaks. These networks revealed that the different branches of MIA biosynthesis and various other metabolic pathways are subject to different hormonal regulation. These networks also served to identify a select number of genes and metabolites likely to be involved in the MIA biosynthesis. So, the combination of multiple omics tools should contribute greatly to identification of key regulatory steps and characterization of the pathway interactions in various processes, aiming at elucidating the systemic coordination and communication among plant metabolic network. Metabolomics combined with other tools are being used to rapidly narrow down candidate genes for functional expression studies and discovery of their biochemical roles^[4].

4 The future prospects

More research is needed to discover the missing structural genes, enzymes and intermediates of MIA biosynthesis in *C. roseus*, as well as genes involved in the regulatory pathway. This knowledge is needed to develop genetically

modified plants, plant cells or microorganisms for the commercial production of the very valuable dimeric alkaloids. So far, the genetic modification of the plant, plant cell cultures or microorganisms has not led to the desired economically feasible production of MIA. In fact, it seems that the pathway is more complex than just a series of enzyme catalyzed steps. Moreover, the MIA pathway does not exist independently from the total metabolic network of the plant, but crosslinks and interacts with other branching pathways, which means it is part of a complex matrix, which raises the question how much of the total carbon flux in the plant can be channeled into MIA biosynthesis. To eventually solve all these problems a systems biology approach is required, which means that all omics will be needed to identify the missing links in the MIA biosynthetic pathway, and map the dynamics of the system. Identification of the regulatory genes might be more difficult as in different species the regulation of the biosynthesis steps can be different. Even in a single plant the regulation will be different between different MIA producing tissues, and between single cells dealing with different parts of the pathway. The single cell approach will thus be a major tool for unraveling pathways and also for studying the regulation and the physiological role of the alkaloids in the plant.

On their own, genetic/molecular tools are not sufficient to figure out the regulation landscape of MIA biosynthesis. Metabolomics, as a powerful technique to reveal changes in metabolic fluxes, is the ultimate level of post-genomic analysis and can facilitate a deeper insight into the function of genes, pathways and single cells through a systems biology approach. The combination of metabolomics with other omics will speed up the elucidation of the MIA pathway and lead to breakthroughs in overcoming the bottlenecks in the production of MIAs in *C. roseus*. Moreover, NMR-based metabolomics in conjunction with genetics strategies could aid in gene annotation and identification of candidate genes for biotechnology and/or breeding strategies in crops, to further facilitate crop improvement.

Acknowledgements This work was funded by National High-tech R&D Program of China (2011AA100605), Shanghai Key Discipline Cultivation and Construction Project (Horticulture), and Shanghai Jiao Tong University Agri-Engineering Program.

Compliance with ethics guidelines Qifang Pan, Jingya Zhao, Yuliang Wang, and Kexuan Tang declare that they have no conflict of interest or financial conflicts to disclose.

This article does not contain any studies with human or animal subjects performed by any of the authors.

References

1. Sumner L W, Mendes P, Dixon R A. Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochem-*

- istry, 2003, **62**(6): 817–836
2. Schauer N, Fernie A R. Plant metabolomics: towards biological function and mechanism. *Trends in Plant Science*, 2006, **11**(10): 508–516
 3. Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Pollier J, Woittiez L, van der Krol S, Lugan R, Ilc T, Verpoorte R, Oksman-Caldentey K M, Martinoia E, Bouwmeester H, Goossens A, Memelink J, Werck-Reichhart D. The seco-iridoid pathway from *Catharanthus roseus*. *Nature Communications*, 2014, **5**: 3606
 4. De Luca V, Salim V, Thamm A, Masada S A, Yu F. Making iridoids/secoiridoids and monoterpenoid indole alkaloids: progress on pathway elucidation. *Current Opinion in Plant Biology*, 2014, **19**: 35–42
 5. Kim H K, Choi Y H, Verpoorte R. NMR-based metabolomic analysis of plants. *Nature Protocols*, 2010, **5**(3): 536–549
 6. Brun G, Dijoux M G, David B, Mariotte A M. A new flavonol glycoside from *Catharanthus roseus*. *Phytochemistry*, 1999, **50**(1): 167–169
 7. Choi Y H, Tapias E C, Kim H K, Lefeber A W M, Erkelens C, Verhoeven J T J, Brzin J, Zel J, Verpoorte R. Metabolic discrimination of *Catharanthus roseus* leaves infected by phytoplasma using ¹H-NMR spectroscopy and multivariate data analysis. *Plant Physiology*, 2004, **135**(4): 2398–2410
 8. Sriram G, Fulton D B, Shanks J V. Flux quantification in central carbon metabolism of *Catharanthus roseus* hairy roots by ¹³C labeling and comprehensive bondomer balancing. *Phytochemistry*, 2007, **68**(16–18): 2243–2257
 9. Toki K, Saito N, Irie Y, Tatsuzawa F, Shigihara A, Honda T. 7-*O*-Methylated anthocyanidin glycosides from *Catharanthus roseus*. *Phytochemistry*, 2008, **69**(5): 1215–1219
 10. Chung I M, Ahmad A, Ali M, Lee O K, Kim M Y, Kim J H, Yoon D Y, Peebles C A M, San K Y. Flavonoid glucosides from the hairy roots of *Catharanthus roseus*. *Journal of Natural Products*, 2009, **72**(4): 613–620
 11. Yang S O, Kim S H, Kim Y, Kim H S, Chun Y J, Choi H K. Metabolic discrimination of *Catharanthus roseus* calli according to their relative locations using ¹H-NMR and principal component analysis. *Bioscience, Biotechnology, and Biochemistry*, 2009, **73**(9): 2032–2036
 12. Mustafa N R, Kim H K, Choi Y H, Verpoorte R. Metabolic changes of salicylic acid-elicited *Catharanthus roseus* cell suspension cultures monitored by NMR-based metabolomics. *Biotechnology Letters*, 2009, **31**(12): 1967–1974
 13. Mustafa N R, Kim H K, Choi Y H, Erkelens C, Lefeber A W M, Spijksma G, van der Heijden R, Verpoorte R. Biosynthesis of salicylic acid in fungus elicited *Catharanthus roseus* cells. *Phytochemistry*, 2009, **70**(4): 532–539
 14. Pan Q, Dai Y, Nuringtyas T R, Mustafa N R, Schulte A E, Verpoorte R, Choi Y H. Investigation of the chemomarkers correlated with flower colour in different organs of *Catharanthus roseus* using NMR-based metabolomics. *Phytochemical Analysis*, 2014, **25**(1): 66–74
 15. Pan Q, Wang Q, Yuan F, Xing S, Zhao J, Choi Y H, Verpoorte R, Tian Y, Wang G, Tang K. Overexpression of ORCA3 and G10H in *Catharanthus roseus* plants regulated alkaloid biosynthesis and metabolism revealed by NMR-metabolomics. *PLoS ONE*, 2012, **7**(8): e43038
 16. Pan Q, Mustafa N R, Tang K, Choi Y H, Verpoorte R. Monoterpenoid indole alkaloids biosynthesis and its regulation in *Catharanthus roseus*: a literature review from genes to metabolites. *Phytochemistry Reviews*, 2015, DOI: 10.1007/s11101-015-9406-4 (first published online)
 17. Mahrous E A, Farag M A. Two dimensional NMR spectroscopic approaches for exploring plant metabolome: a review. *Journal of Advanced Research*, 2015, **6**(1): 3–15
 18. van der Kooy F, Maltese F, Choi Y H, Kim H K, Verpoorte R. Quality control of herbal material and phytopharmaceuticals with MS and NMR based metabolic fingerprinting. *Planta Medica*, 2009, **75**(7): 763–775
 19. Murata J, Luca V D. Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-*O*-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *Plant Journal*, 2005, **44**(4): 581–594
 20. Murata J, Roepke J, Gordon H, De Luca V. The leaf epidermome of *Catharanthus roseus* reveals its biochemical specialization. *Plant Cell*, 2008, **20**(3): 524–542
 21. Ratcliffe R G, Shachar-Hill Y. Revealing metabolic phenotypes in plants: inputs from NMR analysis. *Biological Reviews of the Cambridge Philosophical Society*, 2005, **80**(1): 27–43
 22. Stephanopoulos G, Stafford D E. Metabolic engineering: a new frontier of chemical reaction engineering. *Chemical Engineering Science*, 2002, **57**(14): 2595–2602
 23. Ratcliffe R G, Shachar-Hill Y. Measuring multiple fluxes through plant metabolic networks. *The Plant Journal*, 2006, **45**(4): 490–511
 24. Kim S W, Ban S H, Jeong S C, Chung H J, Ko S M, Yoo O J, Liu J R. Genetic discrimination between *Catharanthus roseus* cultivars by metabolic fingerprinting using ¹H NMR spectra of aromatic compounds. *Biotechnology and Bioprocess Engineering*, 2007, **12**(6): 646–652
 25. Hoekenga O A. Using metabolomics to estimate unintended effects in transgenic crop plants: problems, promises, and opportunities. *Journal of Biomolecular Techniques*, 2008, **19**(3): 159–166
 26. Contin A, van der Heijden R, Lefeber A W M, Verpoorte R. The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture. *FEBS Letters*, 1998, **434**(3): 413–416
 27. Muljono R A B, Scheffer J J C, Verpoorte R. Isochorismate is an intermediate in 2, 3-dihydroxybenzoic acid biosynthesis in *Catharanthus roseus* cell cultures. *Plant Physiology and Biochemistry*, 2002, **40**(3): 231–234
 28. Schuhr C A, Radykewicz T, Sagner S, Latzel C, Zenk M H, Arigoni D, Bacher A, Rohdich F, Eisenreich W. Quantitative assessment of crosstalk between the two isoprenoid biosynthesis pathways in plants by NMR spectroscopy. *Phytochemistry Reviews*, 2003, **2**(1–2): 3–16
 29. Sriram G, González-Rivera O, Shanks J V. Determination of biomass composition of *Catharanthus roseus* hairy roots for metabolic flux analysis. *Biotechnology Progress*, 2006, **22**(6): 1659–1663
 30. Fukushima A, Kusano M, Redestig H, Arita M, Saito K. Integrated

- omics approaches in plant systems biology. *Current Opinion in Chemical Biology*, 2009, **13**(5–6): 532–538
31. Urbanczyk-Wochniak E, Luedemann A, Kopka J, Selbig J, Roessner-Tunali U, Willmitzer L, Fernie A R. Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO Reports*, 2003, **4**(10): 989–993
32. Okazaki Y, Shimojima M, Sawada Y, Toyooka K, Narisawa T, Mochida K, Tanaka H, Matsuda F, Hirai A, Hirai M Y, Ohta H, Saito K. A chloroplastic UDP-glucose pyrophosphorylase from *Arabidopsis* is the committed enzyme for the first step of sulfolipid biosynthesis. *Plant Cell*, 2009, **21**(3): 892–909
33. Lackman P, González-Guzmán M, Tilleman S, Carqueijeiro I, Pérez A C, Moses T, Seo M, Kanno Y, Häkkinen S T, Van Montagu M C, Thevelein J M, Maaheimo H, Oksman-Caldentey K M, Rodriguez P L, Rischer H, Goossens A. Jasmonate signaling involves the abscisic acid receptor PYL4 to regulate metabolic reprogramming in *Arabidopsis* and tobacco. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, **108**(14): 5891–5896
34. Rischer H, Orešič M, Seppänen-Laakso T, Katajamaa M, Lammer-tyn F, Ardiles-Diaz W, Van Montagu M C, Inzé D, Oksman-Caldentey K M, Goossens A. Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, **103**(14): 5614–5619