J. Plant Biol. (2009) 52:462–465
DOI 10.1007/s12374-009-9059-1

ORIGINAL RESEARCH

Genetic Discrimination of *Catharanthus roseus* Cultivars by Pyrolysis Mass Spectrometry

Suk Weon Kim · Jong Hyun Kim · Jang R. Liu

Received: 5 March 2009 / Revised: 1 June 2009 / Accepted: 1 June 2009 / Published online: 28 July 2009

Abstract Pyrolysis mass spectrometry (PyMS) is a rapid, simple, high-resolution analytical method based on thermal degradation of complex materials in a vacuum. It is widely applied to the discrimination of closely related microbial strains. Leaf samples from eight cultivars ('Apricot Delight', 'Cooler Grape', 'Cooler Peppermint', 'Equator Grape', 'Equator Rose', 'Equator White', 'Equator White Eye', and 'Little Bright Eye') of *Catharanthus roseus* were subjected to PyMS for spectral fingerprinting. Discriminant analysis (DA) of PyMS data enabled us to assign these cultivars to discrete clusters. A hierarchical dendrogram based on DA provided a possible relationship among them that was in general agreement with a previously reported classification of the cultivars based on DNA fingerprints. Furthermore, those belonging to the same 'series' were grouped into a single cluster, which previously could not be achieved through similar approaches based on Fourier transform infrared spectroscopy or ^{1}H NMR data. Overall results suggest that chemical differences (i.e., in pyrolysate composition) among cultivars, as detected by mass spectrometry, reflect their genetic variation.

S. W. Kim
Biological Resources Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB),
111 Gwahangno, Yuseong-gu,
Daejeon 305-806, South Korea

J. H. Kim · J. R. Liu (✉)
Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB),
111 Gwahangno, Yuseong-gu,
Daejeon 305-806, South Korea
e-mail: jrliu@kribb.re.kr

Keywords Dendrogram · Discriminant analysis (DA) · Genetic relationships · Principal component analysis (PCA) · Pyrolysis mass spectrometry (PyMS)

Pyrolysis mass spectrometry (PyMS) is a high-resolution analytical method based on thermal degradation of complex materials in a vacuum. This procedure causes molecules to cleave at their weakest points to produce smaller, volatile fragments called pyrolysates (Irwin 1982), which are then further separated on the basis of their mass-to-charge ratio (*m/z*) via mass spectrometry. PyMS is rapid and simple to use, and requires only small amounts of a minimally prepared sample. It has been widely applied to the discrimination of closely related microbial strains (Freeman et al. 1994; Goodacre et al. 1994, 1996a).

PyMS has been performed in only a limited number of plant biology studies, including the discrimination of hybrid from non-hybrid seeds in Triticeae species (Valcarce et al. 1990) and for distinguishing among the seeds of four species from Begoniaceae and Campanulaceae at the genus, species, and varietal level (Goodacre et al. 1996b). We previously reported that discriminant analysis of PyMS data from leaf samples was able to separate higher plants even at the varietal level, where two naturally occurring varieties were included (Kim et al. 2004).

All of these studies demonstrate that PyMS is robust in chemotaxonomic classification, but it has remained to be determined if this approach is valid for representing an extensive number of plant cultivars of a given species. Here, we examined whether PyMS analysis could discriminate eight cultivars of *Catharanthus roseus* based on genetic relationships that had been determined by DNA-fingerprinting techniques, e.g., rapid amplification of polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP).

Materials and Methods

Eight cultivars of *C. roseus* (L.) G. Don ('Apricot Delight', 'Cooler Grape', 'Cooler Peppermint', 'Equator Grape', 'Equator Rose', 'Equator White', 'Equator White Eye', and 'Little Bright Eye') were subjected to PyMS analysis to compare the spectral fingerprints of various genotypes within the same species. All plants were reared in a growth chamber (25°C, approximately 70 μmol $m^{-2}s^{-1}$ from cool-white fluorescent lamps, 16-h photoperiod). Fully expanded leaves were excised at the flowering stage. Leaf discs (0.5 cm diameter) were punched with a cork borer from each plant and were homogenized with a small pestle in 1.5 ml Eppendorf tubes. After centrifugation at 800×g for 1 min, 5 μl of the supernatant was spotted on each metal foil sample carrier. These were run in triplicate using homogenized leaf discs from individual plants of each species. The supernatant samples were oven-dried at 50°C for 30 min before PyMS analysis.

We used a pyrolysis mass spectrometer RAPyD-400 (mass range 12 to 400 Da, ion counting to 25 MHz; SS Scientific Limited, England). Curie-point pyrolysis was conducted at 530°C for 3 s. PyMS data were normalized to the percentage total ion count to remove the influence of sample size. Principal component analysis (PCA) and discriminant analysis (DA) of normalized data were performed with the GENSTAT package. Three-dimensional principal component plots were generated to display groupings of species samples following PCA or DA. A hierarchical dendrogram was developed to show the relationship between plants from PCA or DA of the PyMS data, using the unweighted pair-group method with arithmetic mean (UPGMA) analysis from a multivariate statistical package (MVSP 3.13, Kovach Computing Services).

Results

Quantitative PyMS data were obtained for each sample (Fig. 1). PCA, an unsupervised clustering method requiring no knowledge of the data set, was performed with PyMS data and displayed in three-dimensional plots using the first three principal components (Fig. 2a). No triplicate samples from any cultivar were grouped into discrete clusters, indicating that PCA was unable to discern among cultivars. In contrast, DA, a supervised clustering method requiring a priori knowledge of replicates in the data set, recovered triplicate samples of 'Little Bright Eye' separately from the other cultivars; these were displayed in three-dimensional plots generated with the first three discriminant scores (Fig. 2b). This confirmed that DA could discretely discriminate each cultivar.

Our hierarchical dendrogram divided the eight cultivars into two major clusters (Fig. 2c). The first comprised two from the 'Cooler' series and 'Little Bright Eye', the second, 'Apricot Delight' and four from the 'Equator' series. Those two in the 'Cooler' series were further divided into individual cultivars. The second cluster was separated into two subclusters, the first having only 'Apricot Delight' and the second comprising the four 'Equator' cultivars. Those four were split into two sub-subclusters, one with 'Equator Grape' and 'Equator Rose', and the other with 'Equator White Eye' and 'Equator White', which then ended with individual cultivars.

In an earlier study (Kim et al. 2007a), we applied the two most popular DNA-fingerprinting methods—RAPD and AFLP—to evaluate genetic relationships among these eight cultivars and found slight differences between the two approaches. We had also previously discriminated those same cultivars through Fourier transform infrared spectroscopy (FTIR; Kim et al. 2007c) and ^{1}H NMR spectral fingerprints (Kim et al. 2007a). None of these DNA- or

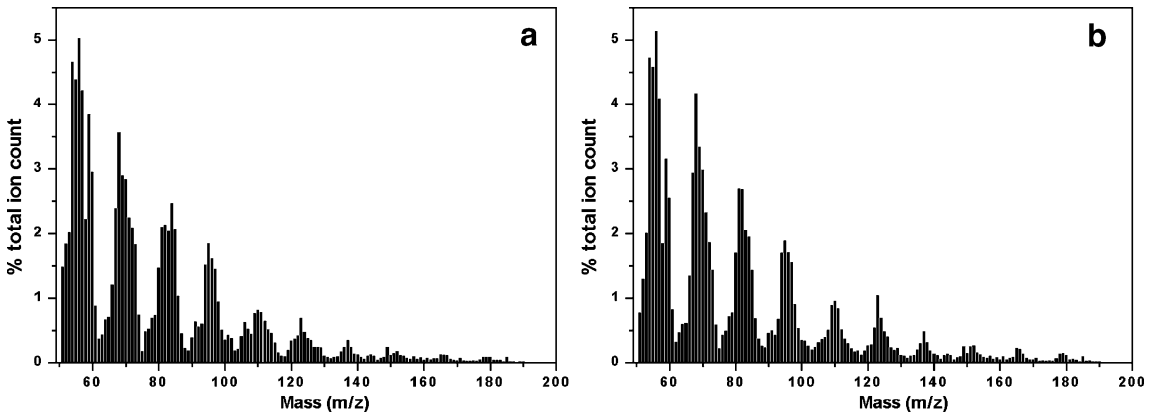

Fig. 1 Representative pyrolysis mass spectra of leaf samples from *Catharanthus roseus* 'Cooler Grape' (**a**) and 'Cooler Peppermint' (**b**). Fully expanded leaf discs were homogenized and, after centrifugation, supernatants were spotted on metal foil sample carriers. They were then oven-dried at 50°C for 30 min before PyMS analysis. Curie-point pyrolysis was performed at 530°C for 3 s

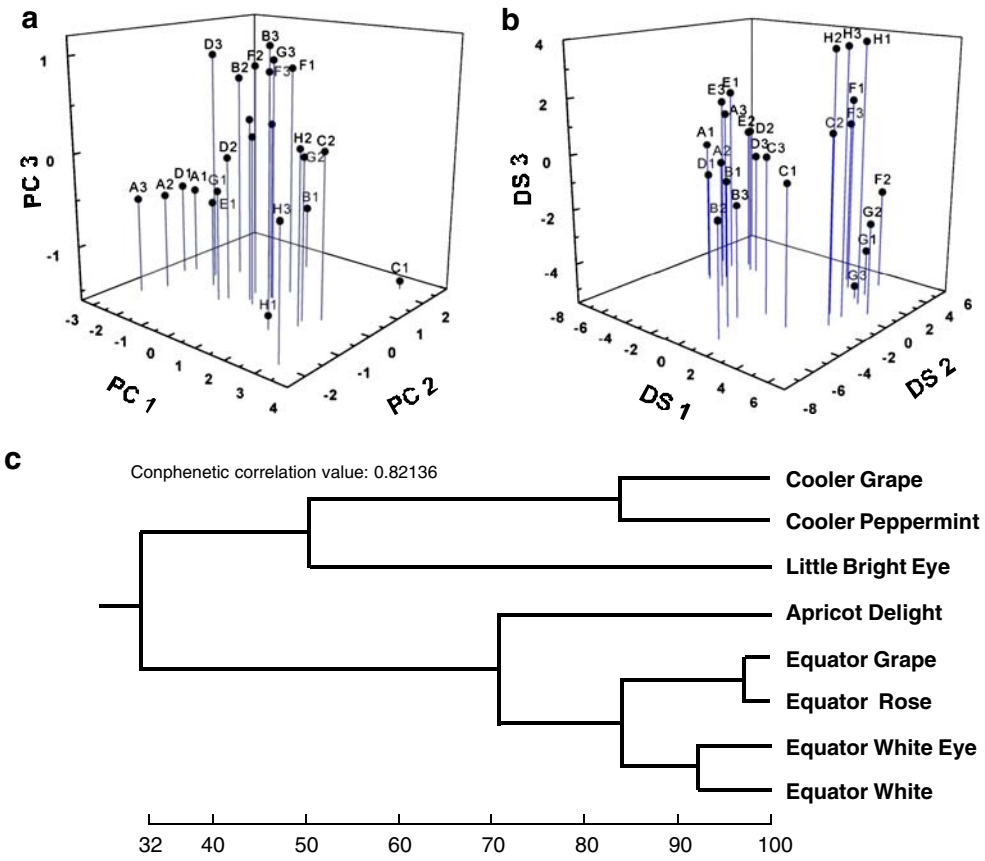

Fig. 2 Three-dimensional plots of principal components (**a**) and discriminant analysis plot (**b**) based on PyMS data analyzed by GENSTAT. First three components (PC 1, PC 2, and PC 3) are displayed, which respectively account for 65.7%, 17.6%, and 11.3% (94.6% total) of all variation. First three discriminant scores (DS 1, DS 2, and DS 3) are displayed, respectively, accounting for 51.6%, 24.0%, and 14.4% (90% total) of all variation. *Letters and numbers* indicate each cultivar and replicate of *Catharanthus roseus*: *A*, 'Apricot Delight'; *B*, 'Cooler Grape'; *C*, 'Cooler Peppermint'; *D*, 'Equator Grape'; *E*, 'Equator Rose'; *F*, 'Equator White'; *G*, 'Equator White Eye'; and *H*, 'Little Bright Eye'. **c** Hierarchical dendrogram based on PyMS data analyzed by GENSTAT depicting possible relationships among eight cultivars

spectral-fingerprinting methods revealed the same relationships among cultivars, although all were in general agreement with one another. Although information on the genetic backgrounds of these eight cultivars is not available, apparently, those within the same 'series' are derived from the same variety. Therefore, we can assume that cultivars within a given 'series' must be genetically closer than those across 'series'. For example, two in the 'Cooler' series ('Cooler Grape' and 'Cooler Peppermint') seem genetically closer to each other than to the four in the 'Equator' series ('Equator Grape', 'Equator Rose', 'Equator White', and 'Equator White Eye'). Our current PyMS spectral fingerprints have now led us to group discretely those cultivars from the same series into a single cluster, a result that could not be achieved by either DNA-fingerprinting (Kim et al. 2007a) or spectral-fingerprinting methods (Kim et al. 2007c).

To describe the relationships among cultivars based on their high-dimensional spectral variables, we reduced the complexity of those variables into three uncorrelated variables called principal components, which accounted for 96.5% of the total variation (Fig. 2a). PCA does not utilize a priori knowledge of the samples, thereby displaying their natural relationships. However, that procedure failed to recover identical replicate samples in discrete clusters. In contrast, DA processes a priori information, of which the samples are replicates. Therefore, we used it with three discriminant scores, which accounted for 90% of the total variation; this produced discrete clusters of identical replicate samples (Fig. 2b).

Discussion

PCA is one of the best known tools for obtaining an overview of multivariate data. It can be used to detect groupings and to evaluate correlations among variables by reducing data dimensionality while minimizing the loss of information. Here, PCA enabled us to depict the complicated variables comprising PyMS profiling data points in three PCs (Fig. 1), removing the remaining dimensions. The goal of this method is to find a new set of axes (PCA vectors) so that most of the data variability is reflected in the first few dimensions. PCs are independent and uncorrelated variables that explain observed variability; each is a linear combination of the original variables.

If data dimensionality is appropriately reduced for correlating the samples, the same sets of replicates should be grouped together. However, in this study, PCA failed to do so, whereas DA successfully overcame that problem. In contrast to PCA, which does not consider the difference between categories (i.e., distinct sets of replicates), DA

maximizes that difference between categories and also minimizes the difference within categories by identifying axes that best separate the categories.

We previously found that PyMS of leaf samples from higher plants provides high-resolution discrimination at the varietal level (Kim et al. 2004). In our current study, we demonstrated that an extensive number of cultivars from a given species can be separated via DA and that a hierarchical dendrogram based on DA can describe possible relationships among cultivars that are equivalent to the genetic relationships revealed by DNA-fingerprinting methods. We have also shown earlier that PyMS combined with multivariate analysis of the data allows us to distinguish embryogenic from non-embryogenic calli in various higher plants (Kim et al. 2006, 2007b).

PyMS analysis has been widely applied for discriminating closely related microbial strains (Freeman et al. 1994; Goodacre et al. 1994, 1996a). Goodacre et al. (1996b) pioneered the use of PyMS for taxonomic classification of higher plants. There, such data from four types of intact seeds enabled researchers to separate out those seeds at the species level when the spectral data were analyzed by DA. However, because of the high level of noise in that PyMS data, it was not possible to discriminate varieties belonging to the same species by conventional multivariate analysis such as PCA and DA. Instead, artificial neural networks were relied upon, which tends to be a black box without analytical basis. Therefore, it was suggested that leaves provide higher chemical complexity than do seeds, making the former tissue more suitable for PyMS-supported discrimination at the cultivar level in higher plants.

In conclusion, we have now shown that discriminant analysis of PyMS data from eight cultivars of *C. roseus* enables us to tell them apart based on their genetic differences. Our overall results imply that the pyrolysates detected from leaf samples by mass spectrometry reflected those genetic relationships. PyMS seemed to provide a higher resolution of genetic similarities than did FTIR or ^{1}H NMR. Therefore, this rapid, simple, high-resolution analytical method may be extensively applied for the discrimination of cultivars from various crops.

Acknowledgements This work was supported by a grant (ABC1000912) to SWK from the BioGreen 21 Program, funded by the Rural Development Administration; a grant to JRL from the Crop Functional Genomics Center of the 21st Century Frontier Research Program; a grant to JRL from the Korea Science and Engineering Foundation through the Plant Metabolism Research Center of the Kyung Hee University, funded by the Korea Ministry of Education, Science, and Technology; and a grant to JRL from the Marine Extreme Genome Research Center, funded by the Korean Ministry of Land, Transport, and Maritime Affairs.

References

Freeman R, Goodacre R, Sisson PR, Magee JG, Ward AC, Lightfoot NF (1994) Rapid identification of species within the *Mycobacterium tuberculosis* complex by artificial neural network analysis of PyMS data. J Med Microbiol 40:170–173

Goodacre R, Neal MJ, Kell DB, Greenham LW, Noble WC, Harvey RG (1994) Rapid identification using pyrolysis mass spectrometry and artificial neural networks of *Propionibacterium acnes* isolated from dogs. J Appl Bacteriol 76:124–134

Goodacre R, Hiom SJ, Cheeseman SL, Murdoch D, Weightman AJ, Wade WG (1996a) Identification and discrimination of oral asaccharolytic *Eubacterium* spp. by pyrolysis mass spectroscopy and artificial neural networks. Curr Microbiol 32:77–84

Goodacre R, Pygall J, Kell DB (1996b) Plant seed classification using pyrolysis mass spectrometry with unsupervised learning: the application of auto-associative and Kohonen artificial neural networks. Chemometr Intell Lab Syst 34:69–83

Irwin WJ (1982) Analytical pyrolysis: a comprehensive guide. Marcel Dekker, New York

Kim SW, Ban SH, Chung HJ, Choi DW, Choi PS, Yoo OJ, Liu JR (2004) Taxonomic discrimination of higher plants by pyrolysis mass spectrometry. Plant Cell Rep 22:519–522

Kim SW, Ban SH, Yoo OJ, Liu JR (2006) Identification of biomarkers for discriminating between embryogenic and nonembryogenic calluses of higher plants using pyrolysis gas chromatography-mass spectrometry and genetic programming. Biotechnol Bioprocess Eng 11:38–42

Kim SW, Ban SH, Jeong SC, Chung HJ, Ko SM, Yoo OJ, Liu JR (2007a) Genetic discrimination between *Catharanthus roseus* cultivars by metabolic fingerprinting using ^{1}H NMR spectra of aromatic compounds. Biotechnol Bioprocess Eng 12:646–652

Kim SW, Ban SH, Yoo OJ, Liu JR (2007b) Discrimination of higher plant calluses based on embryogenic capacity and taxonomic classification by pyrolysis mass spectrometry. Plant Biotechnol Rep 1:61–65

Kim SW, Cho SH, Chung HI, Liu JR (2007c) Genetic discrimination of *Catharanthus roseus* cultivars by multivariate analysis of Fourier transform infrared spectroscopy data. J Plant Biotechnol 34:271–275

Valcarce RV, Smith GG, Stevenson DN, Asay KH (1990) Chemometric analyses of Curie-point pyrolysis—mass spectral data for differentiating seeds of Triticeae grass species and hybrids. Chemometr Intell Lab Syst 9:95–105