Analytical Methods

Discrimination of cabbage (Brassica rapa ssp. pekinensis) cultivars grown in different geographical areas using $^1$H NMR-based metabolomics

Jahan Kim$^{a,1}$, Youngae Jung$^{a,1}$, Byeongyeol Song$^{b,c}$, Yeon-Sik Bong$^b$, Do Hyun Ryu$^d$, Kwang-Sik Lee$^{b,c,*}$, Geum-Sook Hwang$^{a,b,c,*}$

$^a$Integrated Metabolomics Research Group, Seoul Center, Korea Basic Science Institute, Seoul 136-713, Republic of Korea
$^b$Division of Earth and Environmental Science, Ochang Center, Korea Basic Science Institute, Chungbuk 363-883, Republic of Korea
$^c$Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 305-764, Republic of Korea
$^d$Department of Chemistry, Sungkyunkwan University, Suwon 440-746, Republic of Korea

A R T I C L E   I N F O

Article history:
Received 8 February 2012
Received in revised form 5 October 2012
Accepted 6 October 2012
Available online 17 October 2012

Keywords:
Cabbage
Metabolomics
Geographical origin
Cultivars
NMR
Multivariate statistical analysis

A B S T R A C T

Cabbage (Brassica rapa ssp. pekinensis) is one of the most popular foods in Asia and is widely cultivated in many countries for the production of lightly fermented vegetables. In this study, metabolomic analysis was performed to distinguish two cultivars of cabbage grown in different geographical areas, Korea and China, using $^1$H NMR spectroscopy coupled with multivariate statistical analysis. Principal component analysis (PCA) showed clear discrimination between extracts of cabbage grown in Korea and China for two different cultivars (Chunmyeong and Chunjung). The major biochemicals (metabolites) that contributed to discrimination between cabbages grown in the two regions were 4-aminobutyrate (GABA), acetate, asparagine, leucine, isoleucine, O-phosphohydroxyl, phenylacetate, phenylalanine, succinate, sucrose, tyrosine, and valine. These results suggest that the levels of the major metabolites that differ significantly between cabbages grown in these two areas were influenced by environmental factors such as climate and geology. Our study demonstrates that $^1$H NMR based on metabolomics, coupled with multivariate statistics, can be applied to identify the regions of cultivation of various cabbage cultivars.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Cabbage (Brassica rapa ssp. pekinensis) is one of the most popular foods in Asia. It is commonly used for the production of lightly fermented vegetables and is widely cultivated in many countries, especially in Korea and China. Brassica as a food source has a high content of antioxidant metabolites such as phenolics, flavonoids, phenylpropanoids, and glucosinolates that are beneficial for health and nutrition (Onyilagha et al., 2003; Vallejo, Gil-Izquierdo, Perez-Vicente, & Garcia-Viguera, 2004). Glucosinolates are thought to help Brassica plants survive under various environmental conditions and stresses (Halkier & Du, 1997; Loivamaki, Holopainen, & Nerg, 2004). Although this plant has become an essential food in many countries, there is no information to determine the origin or quality of cabbage. As consumers’ concerns about the origin and safety of food have increased, accurate and reliable methods to determine food quality and geographical origin are demanded.

Biochemical (i.e., metabolite) composition is a good criterion for the determination of food quality and origin. NMR-based global metabolic profiling, coupled with multivariate analysis, has been used in studies of different kinds of foods, such as meat (Jung et al., 2010; Shintu, Caldarelli, & Franke, 2007), honey (Donarski, Jones, Harrison, Driffield, & Charlton, 2010; Schievano, Peggion, & Mammi, 2009), wine (Son et al., 2008; 2009), and other plants (Consonni, Cagliani, Stocchero, & Porretta, 2009; Jung et al., 2011; Kim et al., 2011; Lee et al., 2009; Tarachiwin, Katoh, Ute, & Fukusaki, 2008). Discrimination based on quantitative differences of metabolites in food has significant advantages.

$^1$H NMR is well suited for metabolite profiling because it allows simultaneous detection of a diverse group of secondary metabolites in addition to abundant primary metabolites. The use of NMR also simplifies sample preparation and decreases the time required for analysis. Principal component analysis (PCA) is one of the most used tools in chemometric methods and is used to visualise samples present in an n-dimensional space of variables to minimise dimensionality and successively maximise the variance of data (Kemsley, 1996; Sumner, Mendes, & Dixon, 2003).

Several metabolic profiles of cabbage and other Brassica species have been reported in recent years. B. rapa leaves were analysed to differentiate various cultivars and plant ages using NMR.
spectroscopy (Abdel-Farid, Kim, Choi, & Verpoorte, 2007). Differences in metabolite profiles between control and methyl ester-treated B. rapa leaves were detected using J-resolved 2D NMR spectra for identifying the individual phenyl propanoids in B. rapa leaves (Liang, Choi, Kim, Linthorst, & Verpoorte, 2006a; 2006b). However, metabolite profiles that can identify the geographic origin of cabbage have not been reported. In the present study, we used $^1$H NMR-based global metabolite profiling to distinguish the origins of two cabbage cultivars grown in China and Korea.

2. Materials and methods

2.1. Sample preparation

Cabbages (B. rapa ssp. pekinensis) were cultivated in Zhanjiangou (China, 20°51′N, 114°59′E) and Pyeongchang (Korea, 37°26′N, 128°17′E). Two different cultivars, Chunmyeong (CM) and Chunjung (CJ), were collected on 6–8 August 2010 and 6 July 2010 in China and Korea, respectively. A total of 40 cabbages (10 cabbage samples for each group) were collected according to the cultivar and cultivation area: Chunmyeong-China (CMC), Chunjung-China (CJC), Chunmyeong-Korea (CMK), and Chunjung-Korea (CJK). Additionally 282 cabbages that were not identified their cultivars were cultivated in China and Korea.

3. Results and discussion

3.1. $^1$H NMR spectra of cabbage extracts

Representative 600 MHz $^1$H NMR spectra of cabbage samples cultivated in China and Korea are shown in Fig. 1. The most dominant part in each spectrum is the carbohydrate region (3.0–4.5 ppm), except for the spectrum from the CMC sample. NMR spectra of the two cultivars grown in Korea (CMK and CJK) showed relatively low-intensity peaks in aromatic regions. Spectral resonances of metabolites were assigned according to the literature (Bong et al., 2012) and then stored at −80°C until NMR analysis. From each sample, about 50 mg of ground cabbage was put into a 1.5-mL tube containing 2.8 mm zirconium oxide beads and homogenised twice at 5000 rpm with 500 μL methanol (d4), 400 μL 0.2 M sodium phosphate buffer (pH 7), and 100 μL 5 mM DSS (3-(trimethylsilyl)-1-propanesulphonic acid sodium salt) in D2O for 15 s using a Precellys 24 tissue grinder (Bertin Technologies, AmpèreMontigny-le-Bretonneux, France). After homogenisation, the mixture was vortexed vigorously for 1 min. The extracts were sonicated for 20 min, followed by a 10-min centrifugation (16,609 g) at room temperature.

2.2. NMR spectroscopy

The $^1$H NMR spectra were acquired on a VNMRS-600 MHz NMR spectrometer (Agilent Technologies Inc., Santa Clara, CA, USA) using a triple resonance 5-mm HCN salt-tolerant cold probe. $^1$H NMR spectra were acquired using the NOESY-PRESAT pulse sequence, which was applied to suppress the residual water signal. A total of 64 scans were collected into 32 K data points using a spectral width of 9615.4 Hz with a relaxation delay of 2.0 s, an acquisition time of 4.00 s, and a mixing time of 100 ms. A 0.5 Hz line-broadening function was applied to all spectra for Fourier transformation (FT), phasing, and baseline correction. Signal assignment for representative samples was facilitated by acquisition of two-dimensional (2D) total correlation spectroscopy (TOCSY), heteronuclear multiple bond correlation (HMBC), heteronuclear single quantum correlation (HSQC), spiking experiments, and comparisons to literature.

2.3. Chemometric analysis

All $^1$H NMR spectra were phased and baseline-corrected using Chenomx 6.0 (Chenomx Inc., Canada). A total of 1670 complex data points were acquired with a spectral width of 0.005 ppm. The regions corresponding to the solvent (4.7–5.1 ppm for water and 3.29–3.33 ppm for methanol) were excluded, and the remaining spectral regions were divided into 0.005-ppm bins over a range of δ 0.7–9.5 ppm. Spectral data were normalised to the total spectral area. The binning files were aligned using MATLAB (R2008a, The Mathworks Inc., Matick, MA, USA). The resulting data were imported into SIMCA (SIMCA-P+ 12, Umetrics AB, Umea, Sweden), in which PCA was initially performed to examine intrinsic variation in the data set and obtain an overview of variation among groups. All variables were Pareto-scaled for multivariate analysis. The quality of the models was described by $R^2$ and $Q^2$ values. $R^2$ is defined as the proportion of variance in the data explained by the model and indicates goodness-of-fit, and $Q^2$ is defined as the proportion of variance in the data predicted by the model and indicates predictability. Nonparametric two-tailed Mann–Whitney analysis was applied to metabolite concentration data (Wilcoxon, 1945) obtained with Chenomx Inc. software using SPSS 12.0 for Windows (SPSS, Chicago, IL, USA).

3.1. $^1$H NMR spectra of cabbage extracts

Representative 600 MHz $^1$H NMR spectra of cabbage samples cultivated in China and Korea are shown in Fig. 1. The most dominant part in each spectrum is the carbohydrate region (3.0–4.5 ppm), except for the spectrum from the CMC sample. NMR spectra of the two cultivars grown in Korea (CMK and CJK) showed relatively low-intensity peaks in aromatic regions. Spectral resonances of metabolites were assigned according to the literature and the 600-MHz library from Chenomx NMR suite version 6.1, total correlation spectroscopy (2D $^1$H–$^1$H TOCSY), heteronuclear multiple bond correlations (2D $^1$H–$^1$C HMBC), and spiking experiments. A total of 29 cabbage metabolites were identified through $^1$H NMR analysis: 4-aminobutyrate (GABA), acetate, alanine, asparagine, aspartate, citrate, formate, fructose, fumarate, glucose, glutamate, glutamine, isoleucine, lactate, leucine, malate, methylsuccinate, O-phosphocholine, phenylacetate, phenylalanine, proline, succinate, sucrose, threonine, tryptophan, tyrosine, and valine. The chemical shifts of identified metabolites are described in Supplementary Table S1.

Visual inspection of the $^1$H NMR spectra revealed marked differences in the metabolites alanine, lactate, threonine, leucine, isoleucine, glucose, sucrose, fructose, tyrosine, phenylalanine, and phenylacetate among the cabbage sample groups. The concentrations of some metabolites can be changed during sample collection because cabbage can be decomposed quickly in hot weather. A preparatory experiment was conducted to investigate such change in concentration of metabolites in sample collection, such as transport and packaging of samples. Through this experiment, it was found that the concentrations of alanine, lactate, and threonine were significantly different among the samples according to their packaging. Hence, the spectral resonances corresponding to these three metabolites were excluded in this analysis.

3.2. Pattern recognition analysis of $^1$H NMR spectra

PCA is an unsupervised classification method requiring no a priori knowledge of the data set and acts to reduce the dimensionality of multivariate data while preserving most of the variance within it. PCA score plots were used to determine whether the metabolic fingerprints of cabbage samples were sufficiently unique to identify metabolic markers for the different cultivation regions. The PCA score plots derived from the NMR spectra of two different cabbage cultivars grown in the two countries are shown in Fig. 2. Cabbage samples from four different groups were clustered according to cultivation country and cultivar. The PCA model revealed good clustering of samples with high goodness-of-fit and predictability, as indicated by $R^2$ and $Q^2$ values of 0.76 and 0.66, respectively. Each
Fig. 1. Representative $^1$H NMR spectra of the two cabbage cultivars grown in two different regions. (A) Chunmyeong-China (CMC); (B) Chunmyeong-Korea (CMK); (C) Chunjung-China (CJC); (D) Chunjung-Korea (CJK).

Fig. 2. 2D (A) and 3D (B) PCA score plots derived from $^1$H NMR spectra of cabbage samples ($R^2 = 0.76$, $Q^2 = 0.66$). Symbols represent each sample according to country and cultivar: CMC, red; CJC, black; CMK, green; CJK, blue. Abbreviations are as in Fig. 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
The cabbage group was separated mostly by the first PCA axis. The difference between the CMC and CJC groups, two cultivars from China, was much higher than that of the CMK and CJK groups from Korea.

The heatmap data combined with hierarchical clustering analysis of $^1$H NMR spectra from cabbage extracts show that the 40 samples could be divided into two branches (Fig. 3). The length of the horizontal lines in the dendrogram is a measure of similarity, with shorter lines indicating more common features. Similarity assessment for clustering was based on the Pearson correlation coefficient and the average linkage method. The colours of the

![Fig. 3. Heatmap combined with hierarchical clustering analysis of $^1$H NMR spectra of 40 cabbage samples from four different groups. Abbreviations are as in Fig. 1. The x-axis represents the NMR spectral bin.](image)

![Fig. 4. PCA scores (A and C) and corresponding loading plots (B and D) derived from $^1$H NMR spectra of cabbages grown in China (open squares) and Korea (filled circles) for the Chunmyeong (A and B, $R^2 = 0.79$, $Q^2 = 0.69$) and Chunjung (C and D, $R^2 = 0.69$, $Q^2 = 0.51$) cultivars.](image)
Heatmap represent unit variance-scaled peak intensities from spectra, so that red is high intensity and green is low intensity.

All of the cabbage samples were divided into two major branches that consist of samples grown in the two countries. As seen in the length of the horizontal line, CMK and CJK were more similar than CMC and CJC. The Korean cabbage samples showed better clustering than the Chinese samples. There is no particular clustering around a cultivar across the major branch. The peak intensity patterns of the samples are nearly separated by the boundary between the two branches. In the upper part, the peak intensities are high in the phenolic compounds (7–8 ppm) and aliphatic regions. In contrast, at the aromatic and sugar (3–4 ppm) regions, high intensity is found in the bottom part. Considering the overall hierarchical clustering analysis results, the cabbage extracts are better divided into cultivation regions than into cultivars.

3.3. Discrimination between cabbages grown in different regions

To provide a comparative interpretation and visualisation of the metabolic differences in cabbage according to the different growing areas, PCA was applied to the NMR spectral data set for each cultivar (Fig. 4). The PCA score plots showed fairly clear differences between China- and Korea-grown samples of both cultivars (Fig. 4A and C), indicating that cabbage metabolites are strongly influenced by the environmental conditions of each cultivation area. The differentiation between CMC and CMK is characterised by the first principal component in the plot, with $R^2$ and $Q^2$ values of 0.79 and 0.69, respectively (Fig. 4A). PCA modelling between CJC and CJK was also significant, with an $R^2$ value of 0.69 and a $Q^2$ value of 0.51 (Fig. 4C).

The loading plot of cabbage extracts from CMC and CMK showed that CMC extracts contained higher levels of GABA, acetate, aspartate, citrate, formate, fumarate, isoleucine, leucine, phenylacetate, phenylalanine, succinate, tryptophan, tyrosine, and valine than CMK samples. Lower levels of asparagine, fructose, glucose, glutamine, malate, O-phosphocholine, and sucrose were observed in the CMC group (Fig. 4B).

The loading plot of cabbage extracts from CJC and CJK showed that the separation was dominated by increases in GABA, acetate, fumarate, isoleucine, leucine, phenylacetate, phenylalanine, succinate, tryptophan, tyrosine, and valine in the CJC group (Fig. 4D). Higher levels of asparagine, O-phosphocholine, and sucrose were found in Korean cabbages than in China-grown cabbages (Fig. 4D).

To further extend our understanding of the dependence of the cabbage metabolic patterns on geographical origin, the 26 previously assigned metabolites were quantified using the 600-MHz library from Chenomx NMR Suite version 6.0 (Jung et al., 2011; Lee et al., 2009). And then PCA was applied (Supplementary Fig. 5). Quantification of significant metabolites identified from extracts of four different cabbage cultivars. Abbreviations are as in Fig. 1. The significant metabolites were selected using Mann–Whitney tests ($p < 0.05$).
This PCA model obtained from targeted profiling gave slightly better predictability than that from the global spectral analysis (for CM, \(R^2 = 0.86, Q^2 = 0.64\); for CJ, \(R^2 = 0.79, Q^2 = 0.57\)), indicating good discrimination between geographic origins. This result demonstrates that the loading plot and statistical values of the CM cultivar showed better discrimination between the two different countries along PC1 than that of the CJ cultivar.

Most of the significant cabbage metabolites responsible for the observed differences between China and Korea are presented in Fig. 6. To qualify the extent and distribution of significant differences, p-values using Mann–Whitney tests were examined. Fig. 5 summarizes the statistically significant (\(p < 0.05\)) metabolites in both cultivar groups. Differentiation between cabbages grown in China and Korea was caused by higher levels of the metabolites GABA, acetate, isoleucine, leucine, phenylacetate, phenylalanine, succinate, tryptophan, tyrosine, and valine, and lower levels of asparagine, glutamate, guanosine, proline, and sucrose in cabbage from China, compared to cabbage from Korea, regardless of the cultivar.

3.4. Validation of discrimination

In our model, we used cabbage samples having authentic cultivars (Chunmyeong and Chunjung) and origins (the Zhanjiakou of China and the Pyeongchang of Korea) to minimise the variations of metabolic patterns between groups. To validate our PCA model and identified metabolites for discrimination between geographical origins, PCA analysis was also applied on the spectra of the \(^1\text{H} NMR\) data of the additional 282 cabbage samples collected from China and Korea (Fig. 6A). These cabbage samples were composed of several different cultivars grown in various regions around each country, PCA score plot showed differentiation between the Koran and Chinese cabbage samples despite some overlaid samples (\(R^2 = 0.84, Q^2 = 0.74\)). Korean cabbages showed a fairly good clustering whereas Chinese cabbages exhibited a large variation within the group. This variation in Chinese cabbage is likely due to the longer distance among region in China (200–800 km) than in Korea (<200 km). The major contributors to the discrimination among origins were similar with identified metabolites mentioned above, namely, amino acids (leucine, isoleucine, valine, phenylalanine, phenylacetate, tryptophan, tyrosine etc.), sugars (glucose, sucrose, fructose, etc.), succinate, and 4-aminobutyrate (Fig. 6B). It may be considered that our PCA model was well applied to discriminate the geographical origins of cabbage extracts and these metabolites make a criterion between origins including the case of having different cultivars or being cultivated in various regions around each country.

3.5. Characteristics of biomarkers

The factors responsible for this discrimination between cabbages from different countries are likely complex and could be related to environmental factors, such as climate, soil type, and fertilisation. Zhanjiakou in China is at a higher altitude (1325 m) and higher latitude (40°51’N) than Pyeongchang in Korea (559 m, 37°26’N), resulting in more cooler weather conditions; the annual average temperatures in these areas were 8.7 °C and 11.2 °C and the annual rainfall was 576 mm and 1340 mm, respectively. Cabbage samples from China were grown at a higher altitude and in a cool and dry climate, whereas cabbages from Korea were grown with high rainfall and high temperatures. Zhanjiakou shows a wide distribution of volcanicogenic sedimentary rocks, whereas Pyeongchang mainly consists of granitic rocks and clastic sedimentary rocks. These distinct climatic and environmental factors are known to have considerable effects on metabolite production.

Water stress is known to increase sugar content, including cabbage. A high level of sugars was found in water-stressed seedlings as compared to control samples (Jahangir, Abdel-Farid, Kim, Choi, & Verpoorte, 2009; Sasaki, Ichimura, Okada, & Oda, 1998). The influence of growth temperature on free sugar in spinach leaf tissue has been studied, and an increase in sucrose and its primary biosynthetic enzyme have been found to occur in low temperatures (Guy, Huber, & Huber, 1992). GABA (4-aminobutyrate) is a widely distributed, non-protein amino acid that is found in high amounts in many plants (Narayan & Nair, 1990). In animals, GABA occurs at high levels in the brain and plays a major role in neurotransmission (Bouche & Fromm, 2004). Generally, GABA is rapidly produced in response to biotic and abiotic stresses (Kinnerseley & Turano, 2000; Shelp, Bown, & McLean, 1999; Snedden & Fromm, 1999). Moreover, chitosan fertiliser treatment affects cabbage growth and GABA content (Oh, Seo, Choi, Han, & Choi, 2000). In higher plants, GABA is synthesized primarily through the \(H^+\)-consuming \(\alpha\)-decarboxylation of \(\gamma\)-glutamate in a reaction catalysed by L-glutamate (Shelp et al., 1999). In this study, glutamate, precursors of GABA, showed higher level in Korean cabbage compared to Chinese cabbage, whereas GABA exhibited higher level in Chinese samples.

Levels of inosine and guanosine were higher in cabbage extracts from Korea in our study. Inosine, guanosine, and adenosine are associated with purine and pyrimidine metabolism, and participate in many biochemical processes in plants. However, adenosine was not detected in our samples. Guanosine and inosine are salvaged by kinases and/or nucleoside phosphotransferase. The activation and regulation of this pathway are related to salt stress, water stress, and iron deficiency (Itai et al., 2000; Kandpai & Appaji,
Phenolics are synthesized through a number of different pathways, but the shikimic acid pathway is most important in plants. Phenylalanine, tyrosine, and tryptophan are synthesized through this pathway (Mann, 1978). Moreover, glucosinolates, which are an important Brassica metabolite, are synthesized from tryptophan and amino acids such as phenylalanine, leucine, and valine (Chen & Andreasson, 2001; Halkier & Du, 1997; Stoewsand, 1995). It is known that the glucosinolate content of Brassica is affected by seasonal and environmental factors (Abdel-Farid et al., 2007). In this study, significant differences were found in the levels of aromatic compounds in cabbage cultivars from China and Korea. In particular, concentrations of the phenylalazines such as phenylalanine, phenylacetic acid, tyrosine, and tryptophan were much higher in cabbages from China than in Korean cabbages. Glucosinolates were not detected, but their precursors such as leucine and valine had also higher levels in China cabbages than in Korean cabbages.

4. Conclusion

This study demonstrates that 1H NMR-based metabolomic fingerprinting is a useful tool for distinguishing the origins of cabbage samples and that its use in combination with chemometric analysis improves sample classification. A complementary approach using both non-targeted and targeted metabolite profiling was employed to discriminate the cultivation origins of cabbage and identify potential biomarkers. Our results suggest that PCA clustering and hierarchical analysis of cabbage extracts depend on the region of origin rather than the cultivar. The levels of metabolites in cabbages were clearly different between China and Korea, according to chemometric analysis using PCA. Several metabolites were identified as candidate biomarkers that could be used to easily differentiate the origin of the cabbages. High levels of sugars (sucrose, glucose, and fructose) and metabolites related to purine and pyrimidine metabolism (inosine, guanosine) were found in Korean cabbages. However, levels of GABA and phenolic acid in samples from China were much higher than those from Korea. Many factors affect the growth and metabolism of plants, but water stress, temperature, and fertiliser treatment may lead to differences in the metabolism of the studied cabbages, because the two locations in this study have very different climates and growing conditions. The result of this study demonstrated our PCA model works well to discriminate the geographical origins of different cabbage cultivars through the validation for a large number of samples. This metabolite profiling, coupled with chemometric analysis, could be useful to discriminate between different regions of cultivation using biomarkers to distinguish between the cultivation environments in the two countries.

Acknowledgement

This work was supported by an IPET (Korea Institute of Planning and Evaluation for Technology of Food, Agriculture, Forestry and Fisheries) grant and partly by a grant of the Korea Basic Science Institute (T32730) and NRF-2012R1A6A1040282 (Priority Research Centers Program).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2012.10.012.

References


1984; Stasolla, Katahira, Thorpe, & Ashihara, 2003; Weretilnyk et al., 2001).