

¹H NMR 스펙트럼 데이터의 다변량 통계분석에 의한 벼 품종의 구분 및 주요 당 화합물의 정량분석

김석원¹, 구분초¹, 김종현², 유장렬^{2*}

¹생물자원센터, ²식물유전체연구센터, 한국생명공학연구원

Metabolic Discrimination of Rice Cultivars and Relative Quantification of Major Sugar Compounds Using ¹H NMR Spectroscopy Combined by Multivariate Statistical Analysis

Suk Weon Kim¹, Bon Cho Koo¹, Jonghyun Kim² and Jang Ryol Liu²

¹Biological Resources Center and ²Plant Genome Research Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Daejeon 305-333, Korea

ABSTRACT Discrimination of 5 rice cultivars (Sangjubyeo <SJ>, Dongjinbyeo <DJ>, Simbaekbyeo <SB>, Hwamanbyeo <HM>, and Simbaek-hetero <SH>) using metabolic profiling was carried out. Whole cell extracts from each cultivar were subjected to ¹H NMR spectroscopy. When spectral data were analyzed by principal component analysis, 5 cultivars were clustered into 3 groups: SJ, DJ + SB, and HM + SH. The cultivars showed great difference in carbohydrate region of ¹H NMR spectra, suggesting that qualitative and quantitative differences in carbohydrate compounds play a major role in discrimination of the cultivars. In addition, it was readily possible to determine relative quantification of major carbohydrates including sucrose, glucose, maltose from spectral data of the cultivars. SJ showed 2 to 4 times higher content of maltose than the other rice cultivars. Overall results indicate that metabolic discrimination of rice cultivars using ¹H NMR spectroscopy combined by multivariate statistical analysis can be used for rapid discrimination of numerous rice cultivars and simple quantitative analysis system of major carbohydrate compounds in rice grains.

서 론

최근 동식물 및 미생물을 포함하여 대사산물의 질적 양적 변화를 조사함으로써 대사경로의 조절 및 생화학적 대사 경로를 이해하고자 metabolome 연구가 활발하게 이루어지고 있다. 특히 식물분야에서는 표현형이 수반되지 않는 silent mutant로부터 metabolic profiling을 통해서 phenotype을 인지할 수 있을 것으로 보고 관련 유전자의 기능 연구체계를 접목함으로써 본격적인 연구의 단초가 마련되었다 (Fiehn et al. 2000, Raamsdonk et al. 2001).

식물세포로부터 세포내외의 대사산물의 변화를 추적 조사하기 위하여 HPLC, GC, LC/GC-MS 등 여러 분석 장비들이 활용되고 있지만 세포내 모든 대사산물의 질적 양적 변화를 정확히 조사할 수 있는 단일 분석 장비는 아직 없는 실정이다 (Dunn et al. 2005). 또한 이들 분석 장비의 경우 물질 분석을 위한 시료의 추출과정이 복잡하고 여러 분획과정을 거치므로 많은 노동력과 시간이 소요된다. 아울러 세포내 전체 대사산물은 추출방법 및 추출용매에 따라 정량 및 정성 분석 결과의 변화가 이루어진다 (Maharjan and Ferenci 2003, Villas-Boas et al. 2005). 이에 반하여 ¹H-NMR spectroscopy는 혈청이나 소변과 같은 시료의 경우 간편한 과정을 통해 whole cell extracts의 분석이 이루어지고 있으며 복잡한

*Corresponding author Tel 042-860-4430 Fax 042-860-4608
E-mail: jrliu@kribb.re.kr

overlapped signals 형태로 화합물에 대한 정보를 제공한다 (Gavaghan et al. 2000). 이와 같은 복잡한 화학적 정보는 다변량 통계분석 기법들을 활용함으로써 overlapping multiple signal profiles로부터 유용한 생물학적 정보를 유추 및 해석할 수 있다. ^1H NMR spectroscopy는 수소 이온을 포함하는 모든 화합물에 대한 화학적 정보를 제공한다. 따라서 생물의 whole cell extracts는 대사산물의 양적 질적인 변화에 대한 화학적인 정보를 제공하며 whole cell extracts의 ^1H NMR 스펙트럼 조사를 통하여 *Arabidopsis* (Ward et al. 2003) 및 일일초 (Kim et al. 2006) 등에서 품종의 구분이 보고 된 바 있다.

벼의 품종구분은 초기에는 isozyme 패턴 (Glaszmann 1987, Eun et al. 1990)이 활용되었으나 최근에는 RAPD (Fukuoka et al. 1992, Yu and Nguyen, 1994, Mackill 1995), RLGS (Kawase 1994), RFLP (Ghareyazie et al. 1995), microsatellite DNA polymorphism (Yang et al. 1994), AFLP (Subudhi et al. 1998) 등 다양한 유전자 지문분석 방법이 개발되어 활용되고 있다. 그러나 국내에 재배중인 벼 품종은 내병충성 및 다수성과 함께 양질미 우선으로 인하여 동일교배 조합에서 품종이 육성되었거나, 한가지 품종이 여러 교배조합에 이용되어 유전적으로 매우 단순화되어 있다. 따라서 이들 자포니카 벼 품종들은 RFLP 및 RAPD 등 유전자 지문방법에 의한 품종구분은 어려움이 있다. 또한 최근 수입 쌀의 시판을 앞두고 우리 쌀과 수입 쌀을 구별할 수 있는 보다 간편하고 빠른 품종 식별 방법의 개발이 요구되고 있는 실정이다.

따라서 본 연구에서는 건조된 5 종류의 벼 품종의 낱알을 사용하여 ^1H NMR 스펙트럼 조사에 의한 신속한 벼 품종 구분이 가능한 지 여부와 아울러 ^1H NMR 스펙트럼데이터로부터 주요 당 화합물의 상대적인 정량분석이 가능한 지 여부를 조사함으로써 벼의 신속한 품종구분 체계로 활용하고자 하였다.

재료 및 방법

벼 종자 시료 확보

건조된 5 종류의 벼 품종 (상주, 화만, 심백, 심백 hetero, 동진벼, 포항공대 안진홍 교수 제공)으로부터 각 품종별로 종자 5 개를 무작위로 선발하여 막자 사발에 넣고 마쇄하였다. 각 품종별로 3개의 반복시료를 준비하였으며 마쇄된 가루는 Eppendorf tube (1.5 ml)로 옮긴 다음 -70°C 초저온 냉동고에 보관한 다음 ^1H -NMR 분석을 위해 사용하였다.

^1H NMR 스펙트럼 조사

^1H NMR 분석을 위해 각각의 벼 품종별로 마쇄된 낱알 가

루를 각각 15 mg을 취하여 Eppendorf tube (1.5 ml)로 옮긴 다음 0.005% (w/v) TSP-d₄:sodium salt of trimethylsilylpropionic acid 첨가된 1 ml의 deuterated solvent mixture (v/v, D₂O:CD₃OD, 80:20)를 첨가하여 잘 혼합하였다. Eppendorf tube를 50°C 수조에 넣고 10분간 반응시킨 후 whole cell extract를 13,000 rpm에서 5분간 상온에서 원심분리한 다음 상층액 750 μl 를 5 mm NMR tube로 옮겨 각각 ^1H NMR 스펙트럼을 조사하였다. ^1H -NMR 스펙트럼은 Varian Unity 500 NMR spectrometer를 사용하여 조사하였으며 수분 peak는 presaturation pulse를 통해 제거하였다. 각 스펙트럼은 60회 반복 측정하여 평균스펙트럼을 확보하였으며 시료에 첨가된 TSP-d₄를 이용하여 ^1H NMR chemical shifts를 0.00 ppm으로 조정하였다.

^1H NMR 데이터의 다변량통계분석

각각의 벼 품종으로부터 얻어진 ^1H NMR 스펙트럼은 ADVASP Lite (short version of ADVASP package, Umatek Inc.)를 이용하여 ASCII 파일로 전환하였으며 각 스펙트럼의 세기는 internal standard로 첨가한 TSP-d₄의 peak 세기를 기준으로 각 시료의 peak scale을 조정하였다. 각각의 ^1H NMR 스펙트럼의 data 간격은 12.0 to 0 ppm으로 조정하였다. 이와 같이 가공된 ^1H NMR 스펙트럼 데이터는 Matlab (version 6.5)에 도입한 다음 NIPALS algorithm (Wold 1966)에 따라 PCA (principal component analysis) 및 PCR (principal component regression) 분석을 수행하였다. 또한 NMR 스펙트럼데이터는 각각 total metabolites (0-12 ppm), carbohydrates (3-6 ppm), aliphatics (1-3 ppm), 그리고 aromatics (6-9 ppm) region으로 구분하여 각각 PCA 분석을 수행하였다.

^1H NMR 스펙트럼 데이터로부터 Peak Assignment 및 당 성분의 정량분석

^1H NMR 스펙트럼으로부터 주요 당 성분 (fructose, glucose, sucrose, maltose) 및 아미노산 (threonine, alanine, glutamine, glutamate, asparagines, phenylalanine)들의 peak assignment를 위하여 각각의 순수화합물의 농도를 deuterated solvent mixture (D₂O:CD₃OD, 80:20)를 이용하여 1 mg/ml로 조정한 다음 상기의 ^1H NMR 스펙트럼조사와 동일한 조건으로 각각 ^1H NMR 스펙트럼을 조사하였다. 벼 낱알 시료의 ^1H NMR 스펙트럼상에 존재하는 proton signal과 순수 화합물의 chemical shift를 상호 비교하여 각각의 화합물의 존재를 확인하였다. 벼 시료에서 주요 당 성분 및 아미노산의 상대적인 정량분석을 위해 먼저 deuterated solvent mixture로 각 순수 화합물의 농도를 각각 0, 0.1, 0.5, 1, 5 mg/ml로 조정하여 각각의 ^1H NMR 스펙트럼 데이터를 확보하였다. 각 순수 화

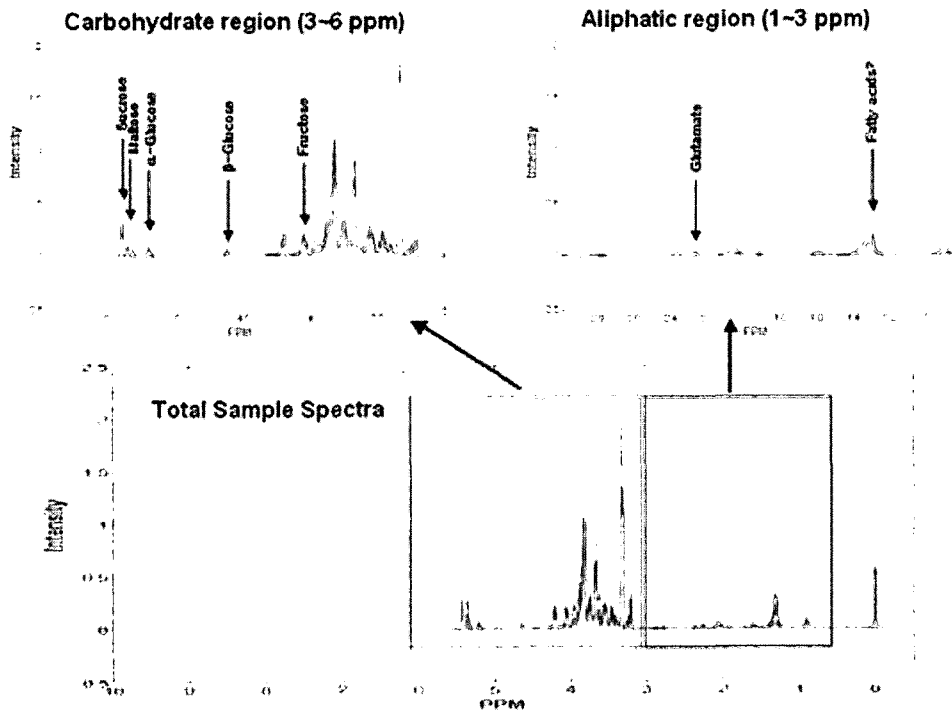


Figure 1. Representative ¹H NMR spectrum and peaks assignment from rice grains of 5 cultivars.

합물의 ¹H NMR 스펙트럼데이터로부터 주요 peaks의 chemical shifts 범위를 결정한 다음 chemical shifts 범위내의 proton signal area를 적분하였다. 각 순수 화합물의 농도별 적분 값과 실제 농도 사이의 선형회귀분석은 상용 프로그램인 Origin (version 6.0)를 사용하여 수행하였다. 각각의 순수 화합물의 농도 예측 회귀모델을 이용하여 벼 시료에서 각 화합물의 상대적인 함량을 비교 분석하였다.

결과 및 고찰

¹H NMR 스펙트럼 데이터의 PCA 분석에 의한 품종구분

각 벼 품종의 ¹H NMR 스펙트럼에 나타나는 주요 당 성분 (fructose, glucose, sucrose, maltose) 및 아미노산 (threonine, alanine, glutamine, glutamate, asparagines, phenylalanine)들의 peak assignment를 위해 순수 화합물의 chemical shifts를 비교 분석하였다. 그 결과 aliphatic region에서 glutamate (2.05, 2.35 ppm) 그리고 carbohydrates region에서 fructose (4.05 ppm), α-glucose (5.15 ppm), β-glucose (4.53 ppm), maltose (5.38 ppm) 그리고 sucrose (5.39 ppm)의 peak를 확인하였다 (Fig. 1). 각 벼 품종의 grain으로부터 얻어진 ¹H NMR 스펙트럼을 보면 carbohydrate region (3-6 ppm)에서는 sucrose, glucose, maltose 그리고 aliphatic region(1-3 ppm)에서는 glutamate 가 주요

peak로 관찰되었으며 양적인 변화가 이루어짐을 알 수 있었다 (Fig. 1). 그러나 ¹H NMR 스펙트럼의 aromatic region (6-9 ppm)에서는 주요 peak들이 관찰되지 않았다 (Fig. 1). 이 결과는 벼 낱알내의 대사산물은 carbohydrate 성분이 주요 성분이며 그리고 소량의 aliphatic 아미노산 성분들이 존재하며 aromatic 아미노산 성분들은 상대적으로 거의 존재하지 않음을 의미하는 것으로 벼 낱알내의 모든 대사산물에 대한 분석이 이루어져 상호 비교가 이루어진다면 매우 흥미로운 것으로 사료된다.

5개의 벼 품종 ¹H NMR 스펙트럼 데이터의 total region으로부터 PCA 분석을 수행한 결과 벼 품종은 크게 3개의 그룹으로 구분이 이루어졌다 (Fig. 2). PCA plot상에서 분석에 사용된 각 품종 별 3개의 반복구가 심백벼를 제외하고는 일정한 영역을 차지하며 분포가 이루어지는 것으로 보아 PCA 분석 결과는 재현성있게 이루어진 것으로 판단되었다. ¹H NMR 스펙트럼데이터의 carbohydrate region 으로부터 PCA 분석을 수행한 결과는 total region의 결과와 유사한 패턴을 보였다. 특히 상주벼의 경우는 독립된 그룹으로 구분이 이루어졌으며 다른 4 품종의 벼는 동진벼와 심백벼, 그리고 화만벼와 심백 hetero 품종이 각각 하나의 그룹으로 구분이 이루어졌다 (Fig. 2). 벼의 품종구분은 isozymes 패턴 (Glaszmann 1987, Eun et al. 1990), RAPD (Fukuoka et al. 1992, Yu and Nguyen 1994, Mackill 1995), RLGs (Kawase 1994), RFLP (Ghareyazie et al.

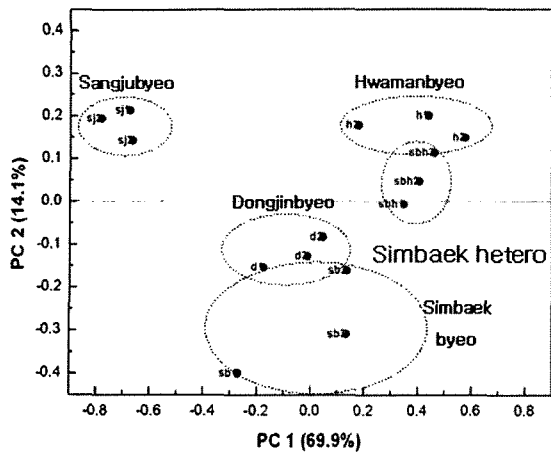


Figure 2. PCA score plots from total region of ¹H NMR spectral data. Each symbol on the PCA plot represents three replicates.

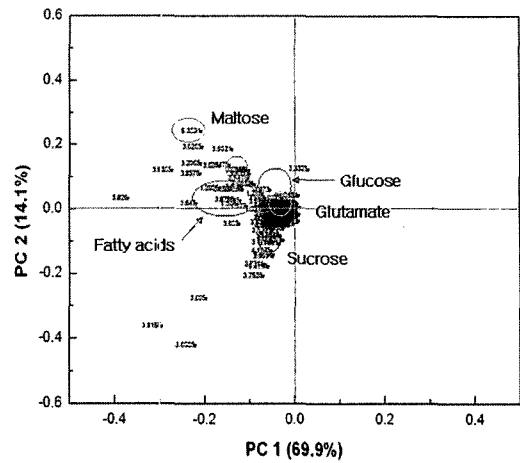


Figure 3. PC score loading plot from total region of ¹H NMR spectral data. Chemical compounds represent important variables for discrimination of 5 rice cultivars.

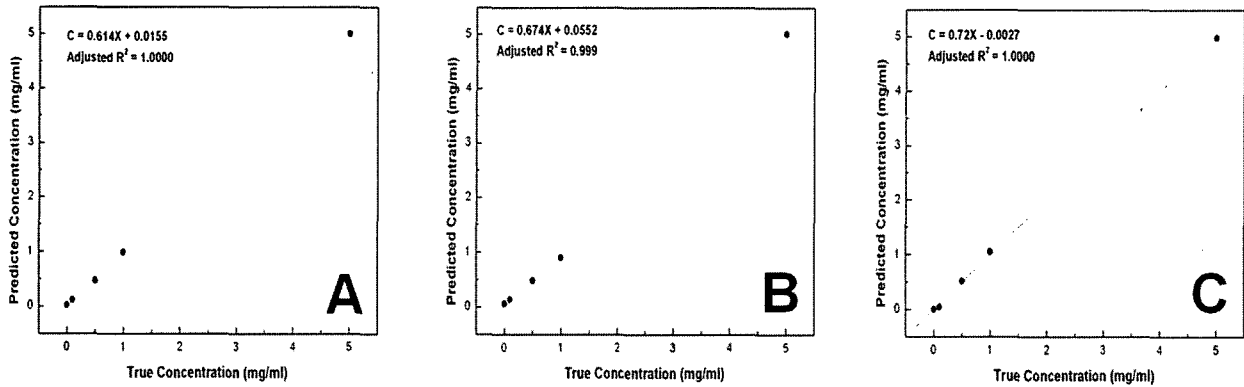


Figure 4. Calibration curve of sucrose (A), glucose (B), and maltose (C) from ¹H NMR spectrum. Regression analysis of sucrose was performed using linear fit tool of Origin software (version 6.0).

1995), microsatellite DNA polymorphism (Yang et al. 1994), AFLP (Subudhi et al. 1998) 등 다양한 유전자 지문분석 방법이 개발되어 활용되고 있다. 따라서 대사산물의 지문분석법과 유전자 지문분석법에 의한 벼 품종의 유연관계 결과와 상호비교는 매우 흥미로운 것으로 사료된다. 그러나 본 연구결과에서는 제한된 소수의 품종시료만을 대상으로 대사산물 지문분석이 이루어져 벼 품종의 계통분류학적 유연관계 규명은 이루어지지 않았다.

PCA 분석을 통한 벼 품종 구분에 중요한 역할을 하는 화합물의 종류를 조사하기 위하여 score loading plot를 조사한 결과 ¹H NMR 스펙트럼상에서 주요 화합물로 확인된 sucrose, glucose, maltose 등이 PCA 분석에서 중요한 역할을 차지하고 있는 화합물임을 알 수 있었다 (Fig. 3). 이 결과는 벼의 품질 구분에서 당 화합물들이 중요한 역할을 하고 있는 사실에 비추어 볼 때 매우 흥미로운 결과라 사료된다. 즉 벼의 품종 구분에 유전학적 마커는 물론 대사산물 마커 역시 중요한 역할을 할 것으로 기대된다.

¹H NMR Peak Assignment와 주요 당 화합물 정량분석

각 벼 품종의 ¹H NMR 스펙트럼에 나타나는 주요 당 화합물 (fructose, glucose, sucrose, maltose) 및 아미노산들의 정량 분석을 위해 각 순수 화합물의 선형회귀분석을 실시하였다. Sucrose (Fig. 4A), glucose (Fig. 4B), 그리고 maltose (Fig. 4C)의 회귀분석 결과 농도별 실측값과 각 화합물의 peak area 적분값 사이에는 매우 높은 상관관계를 보여 각 화합물의 정량 분석이 가능하였다. PCA plot (Fig. 2)상에서 5개 벼 품종 중에서 상대적으로 다른 그룹을 형성하였던 상주벼의 경우 다른 벼 품종에 비해 sucrose 및 glucose 함량은 큰 차이가 없었으나 maltose 함량이 타 품종에 비해 약 2.4배 높음을 알 수 있었다 (Fig. 6). 상주벼가 타품종에 비해 maltose 함량이 많은지 여부는 더 많은 품종들을 대상으로 대사산물 분석이 이루어진다면 보다 확실한 결과를 얻을 수 있을 것으로 사료된다.

따라서 본 연구에서 확립한 벼 종자의 whole cell extracts로부터 ¹H NMR 스펙트럼을 이용한 metabolic profiling 방법을

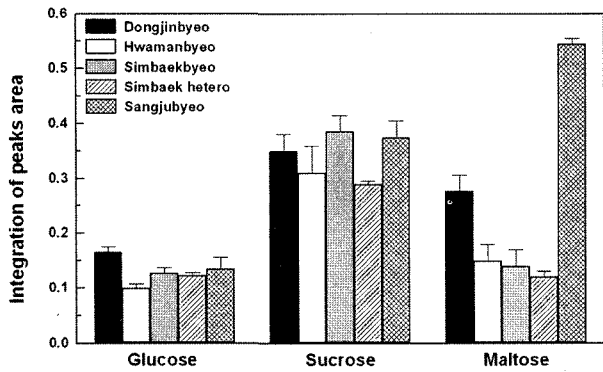


Figure 5. Comparison of cellular carbon metabolites contents. Each metabolite was quantified by integration of peak area using reference chemicals subjected to ¹H NMR spectroscopy. Each treatment consisted of three replicates. Vertical bars represent SD.

기반으로 더 많은 품종을 대상으로 표본 분석 데이터가 구축되어 활용이 이루어진다면 벼 품종구분에 사용될 수 있는 대사산물 수준에서 마커 탐색이 가능할 것으로 사료된다. 아울러 다양한 carbohydrates 및 alphatic 아미노산 성분들의 품종간 비교를 위한 간편한 정량분석 체계로 활용이 가능할 것으로 기대된다.

적 요

건조된 벼 5 품종의 whole cell extracts로부터 ¹H NMR 스펙트럼 조사를 통해 다변량 통계분석법을 활용하여 벼 종자의 품종 구분이 가능함을 조사하였다. ¹H NMR 스펙트럼 데이터에 기초한 PCA 분석 결과 크게 3개의 그룹으로 구분이 이루어졌다. 즉, 상주벼가 나머지 4 품종의 벼와 크게 다르게 구분이 이루어졌으며 동진벼와 심백벼, 그리고 화만벼와 심백hetero 품종이 각각 하나의 소그룹으로 구분이 이루어졌다. 스펙트럼 영역에 있어서는 carbohydrate region이 품종에 따라 크게 달라지는 것으로 보아 탄수화물의 정량정성적 차이가 metabolic profiling에 의한 품종 구분에 중요한 역할을 하는 것으로 추론된다. 또한 ¹H NMR 스펙트럼 데이터에 기초하여 주요 당 화합물 (sucrose, glucose, maltose 등)의 상대적인 정량분석을 조사한 결과 상주벼의 경우 다른 벼 품종에 비해 sucrose 및 glucose 함량은 큰 차이가 없었으나 maltose 함량이 타 품종에 비해 약 2-4배 높음을 알 수 있었다. 따라서 본 연구에서 확립한 벼 종자의 whole cell extracts로부터 ¹H NMR 스펙트럼을 이용한 metabolic profiling 방법은 다양한 벼 종자의 신속한 품종구분은 물론 주요 carbohydrates의 간편한 정량분석 체계로 활용이 가능할 것으로 예상된다.

사 사

실험용 벼 종자를 제공해주신 포항공대 안진홍 교수님께 감사드립니다. 본 논문은 농촌진흥청 바이오그린21 사업, 과학기술부의 21세기 프론티어 프로그램 작물유전체기능연구사업단, 과학재단 SRC의 경희대 식물대사연구센터, 한국생명공학연구원 기관고유사업의 연구비 지원으로 이루어졌다.

인용문헌

Dunn WB, Bailey NJC, and Johnson HE (2005) Measuring the metabolome: current analytical technologies. *Analyst* 130: 606-625

Eun MY, Kim YK, Cho YG, Kim YW, Chung TY, Choi HC (1990) Classification of Korean native rice cultivars by isozyme variations. *Korean J Breed* 21: 293-299

Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L (2000) Metabolite profiling for plant functional genomics. *Nat Biotechnol* 18: 1157-1161

Fukuoka S, Hosaka K, Kamijima O (1992) Use of random amplified polymorphic DNAs (RAPDs) for identification of rice accessions. *Japan J Genet* 67: 247-252

Gavaghan CL, Holmes E, Lenz E, Wilson ID, Nicholson JK (2000) An NMR-based metabolomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk:ApfCD mouse. *FEBS Lett* 484: 169-174

Ghareyazie B, Huang N, Second G, Bennett J, Khush GS (1995) Classification of rice germplasm. I. Analysis using ALP and PCR-based RFLP. *Theor Appl Genet* 91: 218-227

Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. *Theor Appl Genet* 74: 21-30

Kawase M (1994) Application of the restriction landmark genomic scanning (RLGS) methods to rice cultivars as a new fingerprinting technique. *Theor Appl Genet* 89: 861-864

Kim SW, Ban SH, Jeong SC, Chung HJ, Ko S, Yoo OJ, Liu JR (2006) Genetic discrimination between *Catharanthus roseus* cultivars by metabolite fingerprinting using ¹H-NMR spectra of aromatic compounds. *Plant Cell Rep* (accepted)

Mackill DJ (1995) Classifying japonica rice cultivars with RAPD markers. *Crop Sci* 35: 889-894

Maharjan RP, Ferenci T (2003) Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*. *Anal Biochem* 313: 145-154

Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, Westerhoff HV, van Dam K, Oliver SG (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotech* 19: 45-50

- Subudhi PK, Nandi S, Casal C, Virmani SS, Huang N (1998) Classification of rice germplasm: III. High-resolution fingerprinting of cytoplasmic genetic male-sterile (CMS) lines with AFLP. *Theor Appl Genet* 96: 941-949
- Villas-Boas SG, Hojer-Pedersen J, Akesson M, Smedsgaard J, Nielsen J (2005) Global metabolite analysis of yeast: evaluation of sample preparation methods. *Yeast* 22: 1155-1169
- Ward JL, Harris CH, Lewis J, Beale MH (2003) Assessment of ¹H NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. *Phytochemistry* 62: 949-957
- Wold H (1966) Estimation of principal components and related models by iterative least squares. In: Krishnaiah KR (ed), *Multivariate Analysis*, Academic Press, New York, pp. 391-420
- Yang GP, Saghai-Marooof MA, Xu CG, Zhang Q, Biyashev RM (1994) Comparative analysis of microsatellite DNA polymorphism in landrace and cultivars of rice. *Mol Gen Genet* 245: 187-194
- Yu LX, Nguyen HT (1994). Genetic variation detected with RAPD markers among upland and lowland rice cultivars (*Oryza sativa* L) *Theor Appl Genet* 87: 668-672

(접수일자 2006년 12월 3일, 수리일자 2006년 12월 11일)