Journal of Cereal Science 73 (2017) 165–173

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Polymeric hydrophilic interaction liquid chromatography coupled with Orbitrap mass spectrometry and chemometric analysis for untargeted metabolite profiling of natural rice variants

Muhammad Irfan Akram [a], Isabel M. Vincent [b], Amna Jabbar Siddiqui [a], Syed Ghulam Musharraf [a, *]

[a] H.E.J. Research Institute of Chemistry, International Center for Chemical and Biological Sciences, University of Karachi, Karachi, 75270, Pakistan
[b] Wellcome Centre For Molecular Parasitology, Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom

ARTICLE INFO

Article history:
Received 29 September 2016
Received in revised form
29 December 2016
Accepted 2 January 2017
Available online 3 January 2017

Keywords:
Rice
LC-MS
Metabolomics
HILIC Orbitrap

ABSTRACT

GC-MS based metabolomics of genetically modified rice varieties are widely reported but involves tedious derivatization processes. Broad-scale LC-MS-based metabolomics are lacking or limited to targeted analyses. In this study, metabolite diversity in natural rice variants including basmati and non-basmati rice varieties was investigated using polymeric hydrophilic interaction chromatography coupled to an Orbitrap mass analyzer. Identification was performed using exact mass and standard retention times. A total of 329 metabolites were selected using IDEOM (Identification of metabolites) software. Statistical tests were used to filter out top 50 rice varieties discriminating metabolites. Several up- or down-regulated metabolites were observed in rice varieties and can be used as biomarker compounds to differentiate between rice varieties. The developed method and generated data can be used for the comparison of genetically modified rice varieties with natural ones, an important requirement for the implementation of comparator-based GMO risk assessment procedures demanded by national and international regulations.

1. Introduction

Rice (*Oryza sativa* L.) is one of the most important cereal crops in the world, grown on more than 11 percent of the world's cultivable area (usriceproducers.com). Aromatic rice is consumed predominantly in the world because of its aroma and texture. The main driving force behind rice research is to breed rice varieties that have important characteristics such as drought tolerance, pest resistance and an improved nutritional profile. Metabolite profiling provides the most direct measure of the phenotypic state of a biological system (Breitling et al., 2008); therefore, metabolite profiles may be closely related to the important traits of rice such as yield, nutrient content and defense mechanisms(Hu et al., 2014; Matsuda et al., 2012).

It can be difficult to analyze all the metabolites present in a system with a single analytical technique, because they are very diverse in chemical nature. Highly accurate and robust analytical platforms have recently been developed to fulfill the requirements of metabolomics analysis(Breitling et al., 2006). Two analytical platforms, NMR (Nuclear Magnetic Resonance) and MS (Mass Spectrometry) are widely used to study metabolites but each has its own limitations and advantageous(Zhu et al., 2013). Due to its sensitivity, specificity and broader applicability, mass spectrometry has a great scope in metabolite analysis(Dettmer et al., 2007; Hollywood et al., 2006; Want et al., 2007). Mass spectrometers are usually coupled with chromatography techniques such as GC (Gas Chromatography) or LC (Liquid Chromatography) to analyze complex samples. GC-MS based metabolomics has become very common because of the availability of the NIST library and Fiehn metabolomics database, which allow rapid metabolite identification(Kind et al., 2009). LC-MS based metabolomics is gaining popularity due to the ability to analyze polar compounds without

* Corresponding author.
E-mail addresses: irfan.akram@iccs.edu (M.I. Akram), isabel.vincent@glasgow.ac.uk (I.M. Vincent), jabbar_siddiqui08@yahoo.com (A.J. Siddiqui), musharraf1977@yahoo.com (S.G. Musharraf).

http://dx.doi.org/10.1016/j.jcs.2017.01.001
0733-5210/

derivatization(Cubbon et al., 2010). Reversed Phase LC-MS (RP-LC-MS) has been applied extensively for the analysis of small polar molecules but now the hydrophilic interaction liquid chromatography (HILIC) is gaining popularity in combination with high resolution mass analyzers (Urban and Jandera, 2012). Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) and Orbitrap mass spectrometry are approaches used for metabolomics as these analyzers allow the analysis of metabolites (Cubbon et al., 2010; Dunn et al., 2011) with mass accuracies better than 1 ppm (Makarov et al., 2006).

Although the LC-MS related approaches have been extensively applied in food chemistry (Sasaki et al., 2014), a few researches addressed the use of this technique in rice metabolome (Pareja et al., 2012). Wei Chen et al., have reported an LCMS method to compare leaf metabolites of transgenic rice varieties with their natural counter parts (Chen et al., 2013). Similarly Chang, Y.et al., have reported the metabolites of genetically modified rice by using LC-Quadruple-Time-of-Flight Mass Spectrometry (Q-ToF-MS). However, the procedure used to identify the metabolites was solely based on exact molecular mass searches through databases/literature without providing any experimental proofs using authentic standards (Chang et al., 2012). Different studies of untargeted/targeted GC-MS analysis of rice metabolome utilizing solvent extraction followed by silylation have been reported (Frank et al., 2012; Shu et al., 2008; Zhou et al., 2009). NIST library is the major identification source for most of the approaches (Zhang et al., 2013) dealing with metabolite identification based on EI (Electron Ionization) spectra matching.

In this present work, a robust method is used to develop a discrimination model between different rice varieties by studying the metabolite profiles. In this respect, non-target approach followed in HILIC-LC-Orbitrap platform with a simple linear discrimination technique was used to classify the samples with different rice varieties. Data analysis was done using XCMS (Smith et al., 2006), mzMatch (Scheltema et al., 2011) and IDEOM (Creek et al., 2012). Both, positive and negative ESI (Electrospray Ionization) modes (rapid switching) were used to expand the metabolite coverage. We report metabolites with two levels of identification as published by Sumner et al., 2007 (Sumner et al., 2007). Level 1 identifications (represented by one star) are based on the mass and retention time match of a metabolite to its standard and level 2 identifications (represented by two stars) are putatively annotated based on the exact molecular mass and predicted retention times of metabolites based on an algorithm using the retention times from the authentic standards (Creek et al., 2011).

Table 1
Rice varieties and their codes.

No.	Name	Group	Code
01	Basmati 370	Basmati	F1
02	Basmati 385	Basmati	F2
03	Super Basmati	Basmati	F3
04	Shaheen Basmati	Basmati	F4
05	Basmati 515	Basmati	F6
06	Basmati 2000	Basmati	F7
07	MG Basmati	Basmati	F8
08	Basmati 198	Basmati	F9
09	GA 5015	Non-Basmati	F5
10	Super Fine	Non-Basmati	R1
11	PK 386	Non-Basmati	R2
12	KSK 282	Non-Basmati	R3
13	IRRI 9	Non-Basmati	R4
14	Super	Non-Basmati	R5
15	Supra	Non-Basmati	R6
16	IR 6	Non-Basmati	C1
17	KS 282	Non-Basmati	C2
18	KSK 133	Non-Basmati	C3
19	NIAB IR 9	Non-Basmati	C4
20	Shahkar	Non-Basmati	C5
21	JP 5	Non-Basmati	C6
22	Sawat 1	Non-Basmati	C7
23	Dilrosh 97	Non-Basmati	C8
24	Fakhr-e-Malakand	Non-Basmati	C9

2. Experimental

2.1. Rice samples and chemicals

Totally, 24 rice varieties were used in which donated by National Agricultural Research Center (NARC) of Islamabad (18 rice varieties) and Agricultural Research Council (PARC) in University of Karachi (6 rice varieties). These samples with corresponding information of groups are listed in Table 1. All solvents were purchased from Tedia (Tedia way, Fairfield, USA).

Table 2
List of top 50 Metabolites from basmati group after ANOVA test.

No.	RT	Compounds	Formula	P value
1	12.42	Leucine[b]	$C_6H_{13}NO_2$	4.17E-21
2	8.831	erythro-4-Hydroxyglutamate[b]	$C_5H_9NO_5$	2.59E-20
3	8.033	HEPES (Contaminant)	$C_8H_{18}N_2O_4S$	1.18E-18
4	8.062	3-Dehydroxycarnitine[b]	$C_7H_{15}NO_2$	2.85E-16
5	4.659	RT-4.659	$C_8H_8O_2$	6.72E-15
6	8.321	Dihydrobiopterin[b]	$C_9H_{13}N_5O_3$	7.10E-15
7	7.046	2-Amino-6-oxoheptanedioate[b]	$C_7H_{11}NO_5$	2.35E-14
8	6.656	4-Hydroxyphenylacetaldehyde oxime[b]	$C_8H_9NO_2$	1.09E-13
9	7.532	5-Methylthio-D-ribose[b]	$C_6H_{12}O_4S$	4.90E-13
10	8.054	RT-8.054	$C_9H_{19}NO_4$	7.70E-13
11	6.726	3-Methyldioxyindole[b]	$C_9H_9NO_2$	1.34E-12
12	7.976	3-Mercaptolactate[b]	$C_3H_6O_3S$	1.66E-12
13	7.489	2-Aminoadipate[a]	$C_6H_{11}NO_4$	1.89E-12
14	4.504	Gentisate aldehyde[b]	$C_7H_6O_3$	2.70E-12
15	7.295	N-Acetyl-beta-alanine[b]	$C_5H_9NO_3$	3.46E-12
16	8.678	Inosine [a]	$C_{10}H_{12}N_4O_5$	5.22E-12
17	9.143	Hercynine[b]	$C_9H_{15}N_3O_2$	7.15E-12
18	9.251	N(pi)-Methyl-L-histidine[a]	$C_7H_{11}N_3O_2$	7.19E-12
19	9.599	4-Trimethylammoniobutanoate[b]	$C_7H_{15}NO_2$	7.45E-12
20	8.26	N6-Acetyl-L-lysine[b]	$C_8H_{16}N_2O_3$	1.22E-11
21	8.436	Ribose[b]	$C_5H_{10}O_5$	1.68E-11
22	7.139	Histidinal[b]	$C_6H_9N_3O$	2.46E-11
23	8.41	N2-(D-1-Carboxyethyl)-L-lysine[b]	$C_9H_{18}N_2O_4$	2.63E-11
24	9.997	sn-glycero-3-Phosphocholine[a]	$C_8H_{20}NO_6P$	4.65E-11
25	8.132	Adenine[a]	$C_5H_5N_5$	6.96E-11
26	7.922	5,6-Dihydrothymine[b]	$C_5H_8N_2O_2$	1.21E-10
27	7.452	RT-7.452	$C_7H_{10}O_4$	1.94E-10
28	15.71	Choline[a]	$C_5H_{13}NO$	1.98E-10
29	8.481	Methyloxaloacetate[b]	$C_5H_6O_5$	2.68E-10
30	9.528	5-Acetamidopentanoate[b]	$C_7H_{13}NO_3$	2.87E-10
31	10.79	3-Ketosucrose[b]	$C_{12}H_{20}O_{11}$	3.33E-10
32	7.911	5-Hydroxy-L-tryptophan[b]	$C_{11}H_{12}N_2O_3$	4.42E-10
33	8.307	5-Phosphomevalonate[b]	$C_6H_{13}O_7P$	8.77E-10
34	4.89	3,4-Dihydroxyphenylacetaldehyde[b]	$C_8H_8O_3$	9.87E-10
35	8.433	N-Formimino-L-glutamate[b]	$C_6H_{10}N_2O_4$	1.22E-09
36	13.15	Acetylcholine[a]	$C_7H_{15}NO_2$	1.51E-09
37	9.032	1-Aminocyclopropane-1-carboxylate[a]	$C_4H_7NO_2$	1.60E-09
38	17.22	Thiamin[b]	$C_{12}H_{16}N_4OS$	1.99E-09
39	15.54	Homoarginine[b]	$C_7H_{16}N_4O_2$	2.01E-09
40	8.186	4-Hydroxybenzoate[b]	$C_7H_6O_3$	2.14E-09
41	6.757	Nicotinamide[a]	$C_6H_6N_2O$	2.24E-09
42	9.112	Tryptophan[a]	$C_{11}H_{12}N_2O_2$	2.28E-09
43	10.45	Sucrose[b]	$C_{12}H_{22}O_{11}$	2.53E-09
44	8.783	Leucine[a]	$C_6H_{13}NO_2$	2.65E-09
45	8.512	2-Amino-3-oxobutanoic acid[b]	$C_4H_7NO_3$	2.94E-09
46	6.781	5-Hydroxyindoleacetate[b]	$C_{10}H_9NO_3$	3.41E-09
47	8.676	Sedoheptulose[b]	$C_7H_{14}O_7$	5.74E-09
48	6.585	isoglutamine[b]	$C_6H_{12}N_2O_2$	5.85E-09
49	7.133	O-Succinyl-homoserine[b]	$C_8H_{13}NO_6$	6.06E-09
50	8.373	4-Hydroxy-4-methylglutamate[b]	$C_6H_{11}NO_5$	6.73E-09

[a] Level I identified metabolites.
[b] Level II identified metabolites according to MSI (metabolomics standards initiative).