

Original Article

Metabolite profiles of rice cultivars containing bacterial blight-resistant genes are distinctive from susceptible rice

Jiao Wu^{1,2†}, Haichuan Yu^{2†}, Haofu Dai¹, Wenli Mei¹, Xin Huang³, Shuifang Zhu^{3*}, and Ming Peng^{1*}

¹Institute of Tropic Bioscience and Biotechnology, Chinese Academy of Tropic Agricultural Sciences, Haikou 571101, China

²Xinxiang Medical University, Xinxiang 453002, China

³Institute of Animal and Plant Quarantine, Chinese Academy of Inspection and Quarantine, Beijing 100029, China

[†]These authors contributed equally to this work.

*Correspondence address. Tel: +86-898-66890981; Fax: +86-898-66980978; E-mail: mmpeng_2000@yahoo.com (M.P.)/

Tel: +86-10-64896608; Fax: +86-10-64930934; E-mail: zhushf@netchina.com.cn (S.Z.)

The metabolic changes of bacterial blight-resistant line C418/*Xa23* generated by molecular marker-assisted selection ($n = 12$), transgenic variety C418-*Xa21* generated by using the *Agrobacterium*-mediated system ($n = 12$), and progenitor cultivar C418 ($n = 12$) were monitored using gas chromatography/mass spectrometry. The validation, discrimination, and establishment of correlative relationships between metabolite signals were performed by cluster analysis, principal component analysis, and partial least squares-discriminant analysis. Significant and unintended changes were observed in 154 components in C418/*Xa23* and 48 components in C418-*Xa21* compared with C418 ($P < 0.05$, Fold change > 2.0). The most significant decreases detected ($P < 0.001$) in both C418/*Xa23* and C418-*Xa21* were in three amino acids: glycine, tyrosine, and alanine, and four identified metabolites: malic acid, ferulic acid, succinic acid, and glycerol. Linoleic acid was increased specifically in C418/*Xa23* which was derived from traditional breeding. This line, possessing a distinctive metabolite profile as a positive control, shows more differences vs. the parental than the transgenic line. Only succinic acid that falls outside the boundaries of natural variability between the two non-transgenic varieties C418 and C418/*Xa23* should be further investigated with respect to safety or nutritional impact.

Keywords metabolomics; rice; gas chromatography/mass spectrometry; genetically modified

Received: February 27, 2012

Accepted: April 9, 2012

Introduction

Bacterial blight (BB), which is caused by *Xanthomonas oryzae* pv. *oryzae* (*Xoo*), is one of the most destructive

diseases in both inbred and hybrid rice throughout the world [1]. More than 30 BB resistance genes or loci against *Xoo* have been identified in rice so far [2]. However, it has been difficult to use these genes to improve the resistance of the parent for the purpose of hybrid improvement [3]. The cloned and completely dominant gene *Xa21* [4] is widely used for resistance breeding, but its resistance only shows at the later tillering stage [5]. Recently, the near-isogenic line CBB23 with *Xa23* on chromosome 11 was developed and used in breeding programs [6]. Although it is not cloned, the dominant *Xa23* gene displays the broadest spectrum and highest resistance to BB at all growth stages.

C418, an elite restorer line of hybrid rice, has a number of hybrids widely used in rice production [7]. However, C418 is susceptible to BB, introducing *R* genes into this variety is a direct and convenient way [8]. To improve the BB resistance of C418, the transgenic variety C418-*Xa21* was produced through an *Agrobacterium*-mediated system [9,10]. Introgression line C418/*Xa23* was generated by means of molecular marker-assisted selection (MAS) and recurrent backcrossing [6,11,12]. Polymerase chain reaction, Southern blot, and northern blot analyses demonstrated that the foreign genes in the two genetically modified varieties were inherited and expressed steadily through generations. Field demonstration also indicated that they had increased resistance to BB [9,12,13]. However, there is concern that genetic engineering may introduce unforeseen traits into crops, which may cause undesirable metabolites [14]. As the most important crop in the world, the safety of genetically modified rice must be evaluated strictly prior to large-scale production.

Metabolomics enables the parallel assessment of the levels of a broad range of metabolites [15]. Metabolomic technologies, which are typically based on nuclear magnetic resonance analysis, gas chromatography/mass

spectrometry (GC/MS), liquid chromatography/mass spectrometry, and capillary electrophoresis/mass spectrometry, have been successfully applied to various research fields, including plant genotype discrimination based on the multivariate analysis of complex biological profiles [16,17]. GC/MS is of relatively low cost and provides high separation efficiency to resolve complex biological mixtures [18]. The mass spectral and retention time index (MSRI) library and other available libraries meet the expressed demand within the metabolomics community [19–21] and enable the use of GC/MS in more fields than other analysis techniques [16,17,22–25]. Coupled with pattern recognition, GC/MS is a good approach for metabolite profiling of plant genotypes [17,26].

Previous studies have focused on the comparative safety assessment for biotech crops [27,28], but little attention has been paid to new rice varieties generated by MAS techniques. In this study, the comprehensive quantitative metabolic profiles of C418/*Xa23*, C418/*Xa21*, and their genetic background C418 were monitored using GC/MS. Cluster analysis (CA) and principal component analysis (PCA) were used to overview the distribution of all 36 brown rice samples, classify the three rice genotypes, and find the differences between complex samples. Furthermore, we employed the supervised partial least squares-discrimination analysis (PLS-DA) method to construct a visual model with good predictive power. We evaluated the metabolomic differences between genetically modified (GM) variants and wild cultivars using metabolite profiles based on an available Q_MSRI library (GC/MS libraries: www.csbdb.mpimgolm.mpg.de/gmd.html), and then investigated whether it is possible to discriminate between the three rice genotypes using metabolomic variations. In addition, key potential biomarkers with the greatest influence on the discrimination between sample classes were discovered.

Materials and Methods

Plant materials

The plant materials used in the study included three *japonia* (*Oryza sativa* ssp. *indica*) cultivars: C418/*Xa21*, a transgenic restorer line containing *Xa21* [9]; C418/*Xa23*, a new BB-resistant line containing *Xa23* by molecular MAS [12]; ‘C418’, the restorer line for many elite hybrids widely grown in China [7]. These cultivars were obtained from the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (Table 1). Seeds were surface sterilized with 10% H₂O₂ (v/v) for 10 min, rinsed thoroughly with distilled water, and germinated on moist filter paper for 3 days in an incubator at 25°C. After germination, the seeds were then transferred to a net floating on 0.5 mM CaCl₂ in a plastic container. After 6 days, the seedlings

Table 1 Sample information

Sample	Description	Methods of transformation
C418- <i>Xa21</i>	<i>Japonica</i> restorer C418 transformed with <i>Xa21</i> gene containing the marker gene <i>hpt</i>	<i>Agrobacterium</i> -mediated system
C418	Wild comparator	
C418/ <i>Xa23</i>	<i>Japonica</i> restorer C418 transformed with <i>Xa23</i> gene	Polymerase chain reaction-based MAS system

were transferred to a 10-L plastic pot in a growth chamber at a 400-μM photons/m²·s and 28°C/18°C under a 14/10 h light/dark regime. The nutrient solution for *O. sativa* was half-strength Kimura B solution containing the macronutrients including 0.18 mM (NH₄)₂SO₄, 0.27 mM MgSO₄·7H₂O, 0.09 mM KNO₃, 0.18 mM Ca(NO₃)₂·4H₂O, and 0.09 mM KH₂PO₄ and the micronutrients including 20 μM NaEDTAFe·3H₂O, 6.7 μM MnCl₂·4H₂O, 9.4 μM H₃BO₃, 0.015 μM (NH₄)₆Mo₇O₂₄·4H₂O, 0.15 μM ZnSO₄·7H₂O, and 0.16 μM CuSO₄·5H₂O. The nutrient solution was prepared with purified water, aerated daily, and renewed every 3 days. The pH of the solution was adjusted daily to 5.6 with diluted HCl and/or NaOH.

Seeds were harvested separately for each cultivar after 40 days, starting from the day on which the first panicle of rice was observed. The seeds were dried at 30°C for 3 days, threshed from the panicles manually, and then collected. All seeds in the husks were stored at 4°C under dark conditions until analysis.

Chemicals

All chemicals were purchased from Sigma-Aldrich Chemie GmbH (Deisenhofen, Germany). The methoxyamination reagent methoxyamine hydrochloride (Sigma, Munich, Germany) was dissolved at 40 mg/ml in pure pyridine (Merck, Darmstadt, Germany). The per-silylation reagent was 1 ml vials of *N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide (MSTFA; Macherey & Nagel, Düren, Germany).

Extraction and derivatization of the samples

Extraction of the plant samples was performed as previously described [8] with the exception that different volumes of solvents were used according to the fresh weight of the samples and different aliquots of the upper phase were taken for derivatization. Briefly, 100 seeds were selected for each cultivar according to average weight and length. After separating the husks from the seeds, the brown rice seeds were bulked and crushed using a Retsch mixer mill

MM400 at a frequency of 25 Hz^{-1} for 1 min at 4°C . Then, 1 ml of extraction buffer [isopropanol/acetonitrile/water (3:3:2, v:v:v), -20°C] was added to 100 mg of the obtained powder. The samples were vortexed vigorously for 10 s and then shaken for 15 min at 4°C . After centrifugation, a 150- μl aliquot of the supernatant was drawn and transferred into a glass insert vial. The extracts were evaporated to dryness in an SPD111 V-230 SpeedVac concentrator from ThermoSavant (Thermo Electron Corporation, Waltham, USA) at room temperature for 6 h.

Derivatization was performed as described previously [8]. In summary, 2 μl of a C8-C40 n-alkanes mixture was used to convert retention times to retention indexes. Carbonyl groups were protected by 10 μl of a solution of 40 mg/ml methoxyamine in pyridine at 30°C for 90 min. Ninety microliters of MSTFA + 1% Trimethylchlorosilane was added for trimethylsilylation of acidic protons at 37°C for 30 min. All the derivatized samples were analyzed randomly within 24 h.

GC-MS analysis and operating conditions

The reaction mixture was transferred to a 2-ml clear glass autosampler vial with microinsert (Agilent, Santa Clara, USA). Sample analysis was performed essentially as described [29]. Briefly, samples (volume of 1 μl) were injected with splitless using a hot-needle technique. The GC-MS system consisted of a 7683B autosampler, a GC 7890A gas chromatograph and a 5975C quadrupole mass spectrometer (Agilent, Atlanta, USA). Gas chromatography was performed on a 30-m capillary column Rtx-5Sil MS with a 0.25-mm inner diameter, integrated guard column, and 0.25- μm film (Restek GmbH, Bad Homburg, Germany). The injection temperature was 230°C . The transfer line and ion source temperatures were 250°C and 200°C , respectively. The carrier gas was helium set at a constant flow rate of 1.0 ml/min. The temperature program starts isothermal for 5 min at 70°C followed by a $5^{\circ}\text{C}/\text{min}$ ramp to 350°C and a final 5-min heating at 330°C . Cooling is performed as fast as instrument specifications allow. The system was then temperature equilibrated for 1 min at 70°C prior to injection of the next sample. Mass spectra were recorded at two scans per second with a 50–400- m/z scanning range. The acceleration voltage was turned on after a solvent delay of 9.5 min.

Mass spectral data processing

All mass spectral data were processed by Automated Mass Spectrometry Deconvolution and Identification System (AMDIS) software version 2.66 from National Institute of Standards and Technology (NIST) (Gaithersburg, USA) and exported in Elu and Fin format to Mass Profiler Professional (MPP) software version 2.0 (Agilent Technologies, Palo Alto, USA). Metabolites were unambiguously assigned by

the AMDIS, using retention index and mass spectrum as the two most important identification criteria. Automatic peak detection and mass spectrum deconvolution were performed using a peak width set to 2.0 s. Peaks with signal-to-noise (S/N) value lower than 10 were rejected. The S/N values were based on the masses chosen by the software for quantification. Data, including baseline correction, peak deconvolution, and peak annotation for GC-quad-MS were processed. All data pretreatment procedures, such as data normalization, baseline correction, chromatogram alignment, and the subsequent data treatments, were performed using custom scripts [30] for multivariate statistical analysis of metabolite phenotype clustering. To obtain accurate peak areas for the deconvoluted components, the unique masses for each component were specified and the samples were reprocessed. The resolved MS spectra obtained from the custom scripts were matched with reference mass spectra using the commercial NIST 08 mass spectral library and the Q-MSRI library in the Golm Metabolome Database (GMD) at CSB.DB [31,32]. The extracted mass spectra were finally identified or annotated according to their retention index and comparison with the reference mass spectra in the libraries.

Statistics

MPP software version 2.0 was used for retention time and mass abundance adjustments. Multiple data sets were aligned and normalized to correct data for retention time and response drift before they were compared. The workflow followed these steps: (i) alignment of features across all samples, (2) ‘per mass’ normalization to the median, and (3) filtering. All independent biological replicate samples per variety were analyzed as one variety after the ion intensity for each molecular ion was averaged across all biological replicates. The relative frequency of detection was varied and compared. The highest quality features analyzed were those that were only present in all replicates, in at least one or more groups. The features were then analyzed by volcano plots to reveal both fold changes and statistical significance ($P < 0.05$).

Multivariate statistical analysis was performed by unsupervised PCA, cluster analysis (CA), and supervised PLS-DA with mean centering and scaling to visually demonstrate the variance in the metabolic phenotypes based on the discriminating features from the analysis of variance. Hierarchical cluster analysis of metabolites was performed to reveal association between replicate biological samples within a group based on the similarity of their mass abundance profiles. For univariate statistics, data were Log(2)-transformed to reduce the weight of outliers and to render the data sets more normally distributed. Two-tailed Student’s *t*-tests were performed with significance thresholds of $P < 0.05$ and < 0.001 . Ratios were calculated from antilogged class averages. PCA was carried out on unit

variance scaled data in unsupervised mode for dimension reduction purposes. The metabolic differences between the transgenic rice and non-transgenic counterparts were visualized three-dimensionally [33], and simple comparison of peak area means was performed by *t*-test. A difference of $P < 0.05$ was regarded to be significant. Daily quality controls were used as Fiehn's method [34].

Results

Physical characteristics of genetically modified and non-transgenic rice seeds

Physical characteristics, such as the shape, uniformity and translucence, of grains are crucial aspects of grain quality [35,36]. The characteristics of C418/Xa23, C418-Xa21, and C418 were compared in Fig. 1. The length, width, thickness, and weight of a single grain of brown rice from each rice genotype were measured. Among the three rice samples, no marked difference in seed appearance was observed. The weights of the two genetically modified rice seeds were significantly decreased compared with C418.

The number of detected peaks increased linearly with class frequency threshold reduction

Eight-fold replicate extraction tests were performed to determine the technical precision of the method. For the identified metabolites, an overall median precision of 17%

relative standard deviation (RSD) was determined, validating that the extraction and analysis protocol was suitable for GC-Quad-MS-based metabolite profiling [37]. The technical reproducibility was determined as 10% RSD for the 27 more abundant compounds and 23% RSD for the 80 less abundant metabolites. Thirty-six brown rice samples of each rice genotype were pretreated and analyzed. There were 12 samples from C418, C418-Xa21, and C418/Xa23. GC-Quad data were deconvoluted by AMDIS and constrained by Q-MSRI GMDs. The number of detected peaks increases proportionally to lowering the threshold of positive detection for an entity in at least one condition, e.g. present in 12 of 12 samples replicates rice grain extracts per one or more classes (100% threshold) or lower thresholds (50% and 75%) (Table 2). Generally, the reduction in entities with higher threshold stringency improves the confidence in subsequent statistical investigation by limiting the dimensionality and reducing the noise in the data [8]. Therefore, we calculated the relationship between the class frequency threshold (percent relative frequency of entities detected across replicates in at least one group) in liquid injection GC-Quad analyses and the number of reported entities. The number of differential peaks at >2-fold change is shown for various comparisons with decreasing class frequency thresholds for peak detections. Parentheses indicate the number of compounds having both greater than 2-fold change and $P < 0.05$ (Table 3). We

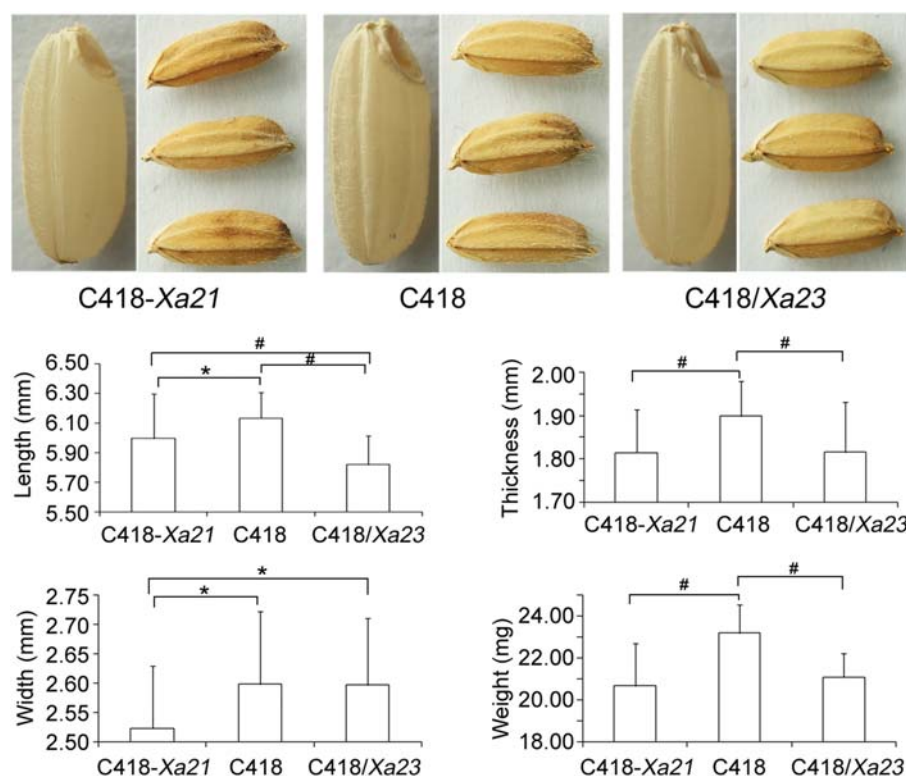


Figure 1 Comparative morphology, length, width, thickness, and weight of bacterial blight-resistant rice lines and non-transgenic rice single grain. * $P < 0.05$, # $P < 0.01$.

found that the number of entities roughly increased linearly with decreasing class frequency threshold.

In total, 609 unfiltered entities were detected in GC-EI-Quad ion mode. Upon filtering, the entities were reduced to 467 (50% threshold) or 180 (100% threshold). GC-Quad-MS data of structurally uncharacterized metabolites can be classified into identified and structurally annotated metabolites via retention index and mass spectral matching to reference libraries. In a manner similar to the total number of metabolic signals, the number of identified metabolites linearly increased from 83 to 164 non-redundant metabolites with decreasing relative frequency threshold levels. This finding confirmed that setting the positive relative frequency threshold at 50% was reasonable for comprehensive metabolomics. The fact that only around one-third of all peaks were structurally assigned is typical for unbiased metabolomics screening, suggesting that functional annotation of plant genes would greatly benefit from improved algorithms and libraries used for compound identification [8].

Next, detailed pair-wise comparisons were performed between the different groups that included fold change and significance analysis based on unpaired *t*-tests. The results were summarized in Table 3. The number of significant metabolic differences ($P < 0.05$) between C418/*Xa23* and C418 was clearly more than those between C418/*Xa2* and C418. This observation was especially evident for those significant differences that were >2 -fold.

Genetically modified variants and their progenitor line C418 differ in metabolic phenotypes

The representative total ion current chromatograms were obtained from the total isopropanol/acetonitrile/water extract of each rice genotype after silylation with MSTFA [Fig. 2(A–C)]. High-level sucrose [Retention time (RT) 40.3 min] and raffinose (RT 48.7 min) were measured as a common feature in all three rough rice genotypes.

Table 2 Application of different stringency filters to the processing of raw spectrometry data files significantly affects interpretation of global metabolite profiling analyses

Class frequency threshold (%)	GC-Quad total	GC-Quad unknowns	GC-Quad identified
50	467	303	107
75	334	209	95
100	180	97	62

GC-Quad data were deconvoluted by AMDIS and constrained by Q-MSRI Golm metabolome databases. The number of detected peaks increases proportionally to lowering the threshold of positive detection for an entity in at least one condition, e.g. present in 12 of 12 samples replicates rice grain extracts per one or more classes (100% threshold) or lower thresholds (50% and 75%).

As part of our exploratory phase, we performed cluster analysis based on 3 K-means, 3×4 Self-Organizing Map (SOM), and hierarchical grouping of the samples into clusters based on similarity of their metabolite abundance profiles. Figure 2(D–F) present heat maps that contain all detected metabolites in all samples by GC-Quad. The heat maps reveal good clustering of sample replicates. Interestingly, the samples from C418 were included in C418-*Xa21* by using HCA. The samples from C418 and C418-*Xa21* were clustered together, but not crossed by 3×4 SOM, indicating an analogical metabolite profile between rice cultivar C418 and its *Xa21* transgenic variant different from C418/*Xa23*.

As shown in the PCA score plot [Fig. 3(A)], the distribution of the samples in the score plot with the first two principal components accounted for 87%. Each point represents a particular sample. Rice samples of the C418/*Xa23* variety were both discriminated mainly along principal component 1 (60.2% of explained total variance of all variables) from the C418-*Xa21* and C418, indicating a distinctive metabolite profile. Thus, PCA analysis poorly distinguished C418-*Xa21* from C418.

After the distribution of samples in the variable space by HCA and PCA was overviewed, the supervised partial least squares-discriminant analysis (PLS-DA) method was used to establish a visual model for the discrimination of GM varieties and wild-type controls [Fig. 3(B)]. Direct visualization of the significance and magnitude of effects in the rice grain comparison was provided by *t*-score plots of PLS-DA and volcano plots for each pair-wise variety. As shown in Fig. 4(A–C), autoscaling not only centered the data set but also made the standard variation variable have the same units [18]. The red points in the volcano plots, based on the differential abundance between pair-wise conditions using fold change and *P* values for visualization [Fig. 4(A)], represent 48 significant differences between the rice grains of C418-*Xa21* and C418 [Fold change (FC) > 2.0 and $P < 0.05$]. A total of 154 components

Table 3 Univariate analysis of variance statistics for C418-*Xa21* and C418/*Xa23* compared with C418

Class frequency threshold (%)	C418- <i>Xa21</i> versus C418	C418/ <i>Xa23</i> versus C418
GC/Quad 100	137 (16)	179 (66)
GC/Quad 75	260 (29)	307 (119)
GC/Quad 50	397 (64)	442 (163)

The number of differential peaks with >2 -fold change is shown for various comparisons with decreasing class frequency thresholds for peak detections. Parentheses indicate the number of compounds having both >2 -fold change and $P < 0.05$.

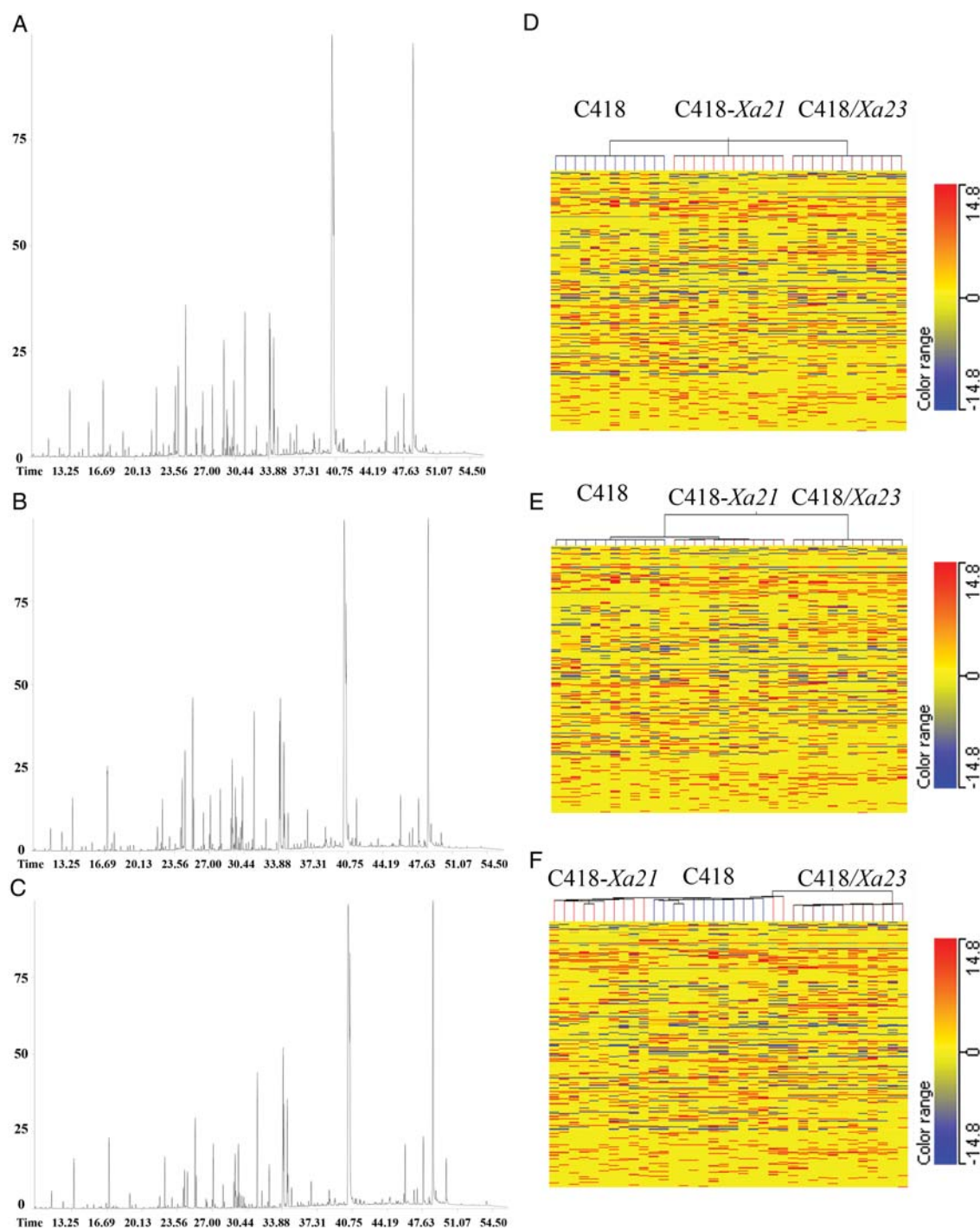


Figure 2 Cluster analysis of GC-MS revealed a more distinct profile of the new rice variety C418/Xa23 compared with rice cultivar C418 and its transgenic variety C418-Xa21 (A–C) Typical GC/MS total ion current chromatograms for the total isopropanol/acetone/water extract of the three rice genotype samples C418-Xa21(A), C418 (B), and C418/Xa23 (C) after silylation with MSTFA. (D–F) Cluster analysis of GC-MS data for 12 biological replicates for the three rice genotypes based on 3 K-means, 3 × 4 SOM, and hierarchical. The mass trees (dendrogram) reveal the relationships between the different conditions based on their abundance levels with each rectangle or cell representing an accurate mass and colored by its abundance intensities on a normalized scale from –14.8 (low) to 14.8 (high). Masses with similar abundance patterns are placed closer to each other. The mass tree branches at the top of the figure are colored by each condition, and show the correlations for the main clades or branches.

varied between C418/Xa23 and C418 [Fig. 4(B); FC > 2.0 and $P < 0.05$], and 112 varied between C418/Xa23 and C418-Xa21 [Fig. 4(C)]. Furthermore, a total of 77 components varied between C418/Xa23 and C418 (FC > 2.0 and

$P < 0.001$), and 13 varied between C418-Xa21 and C418. Alanine, glycerol, glycine, malic acid, succinic acid, ferulic acid, and tyrosine were all present at significantly lower levels ($P < 0.001$) in C418/Xa23, whereas linoleic acid

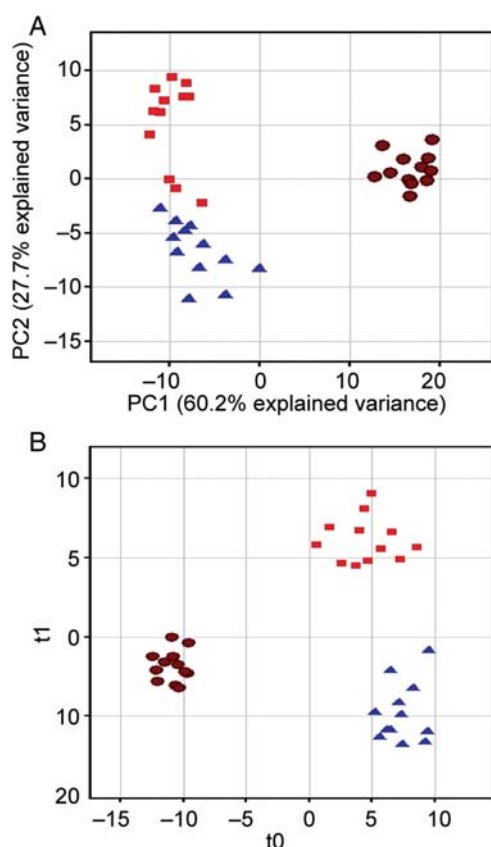


Figure 3 PCA and PLS-DA of GC-MS data showed inherent brown rice metabolism difference among the three genotypes (A) 2D-PCA scores plot represent the discrimination among two GM variants, C418-*Xa21* and C418/*Xa23*, and their counterpart. 2D-PCA score plots consisting of the PC1 and PC2 are shown. (B) PLS-DA scores plot display the diversity among C418-*Xa21*, C418/*Xa23*, and their counterpart. The squares (red), triangles (blue), and circles (brown) indicate C418-*Xa21* ($n = 12$), C418/*Xa23* ($n = 12$), and C418 ($n = 12$), respectively.

was present at considerably higher levels ($P < 0.001$) than in C418 (Fig. 5). Although most significant differences are still inside between C418 and C418/*Xa23*, succinic acid that falls outside the boundaries of natural variability between the two non-transgenic varieties should be further investigated with respect to safety or nutritional impact.

Discussion

We also noticed that linoleic acid increased to 29% and 25% in transgenic rice *O. sativa* L. ssp. *indica* with resistance to rice blast and *O. sativa* L. Kefeng No. 6 with insect resistance, respectively, compared with their non-transgenic counterparts, which means that the transgenic effect was positive on the fatty acid compositions [27]. Interestingly, we observed that plant height and leaf length of C418/*Xa23* were significantly larger than C418 and its transgenic variant C418-*Xa21* throughout the vegetative

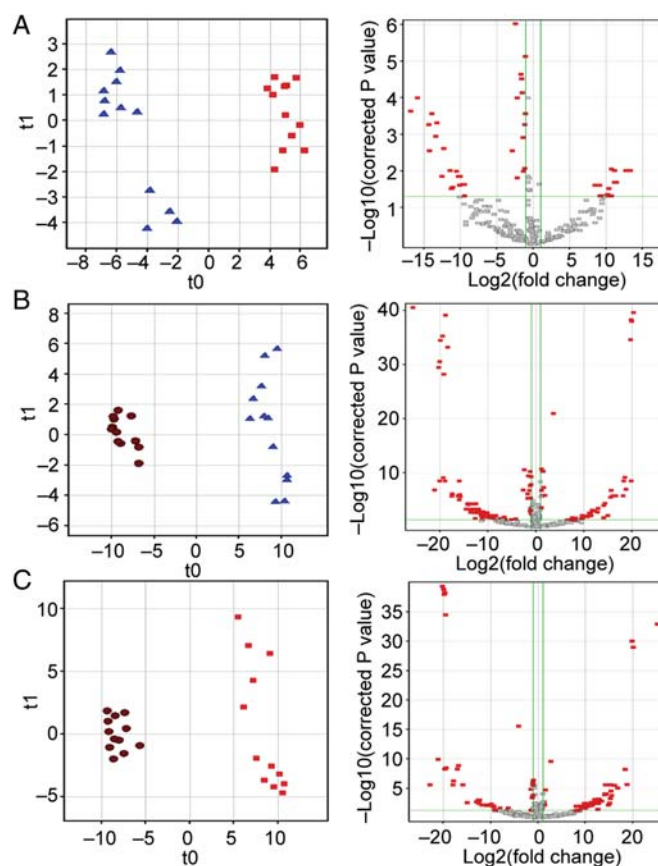


Figure 4 PLS-DA score plots (left) and volcano plots (right) of metabolomic comparison (A) C418-*Xa21* versus C418, (B) C418/*Xa23* versus C418, and (C) C418/*Xa23* versus C418-*Xa21*. Significant difference with $P < 0.05$ and >2 -fold intensity ratios (50% threshold) are shown as red rectangle (gray rectangle means no significant difference). The squares (red), triangles (blue), and circles (brown) indicate C418-*Xa21* ($n = 12$), C418/*Xa23* ($n = 12$), and C418 ($n = 12$), respectively.

stage. In the absence of ammonium salt as a nitrogen source, only C418/*Xa23* grew normally (data not shown). Simultaneously, up-regulation of linoleic acid only occurred in C418/*Xa23*, which might result in promoting the vegetative stage through the activation of phosphoenolpyruvate carboxylase (PEPC) [38], which is consistent with recent results that the genus *oryza* has a unique mechanism for providing organic acids for ammonium assimilation that involves a chloroplastic PEPC, and that this route is crucial for growth with ammonium [39].

In addition, Sana *et al.* [8] have shown that there is inherent difference in leaf metabolism and that central carbon catabolism is reduced in correlation with gene and metabolite expression in both *Xa21* and TP309 rice genotypes. This was interpreted as a pleiotropic effect of primary genetic alteration but not found in pathways involved in resistance to bacterial infection. Zhou *et al.* [28] reported that sucrose, mannitol, and glutamic acid

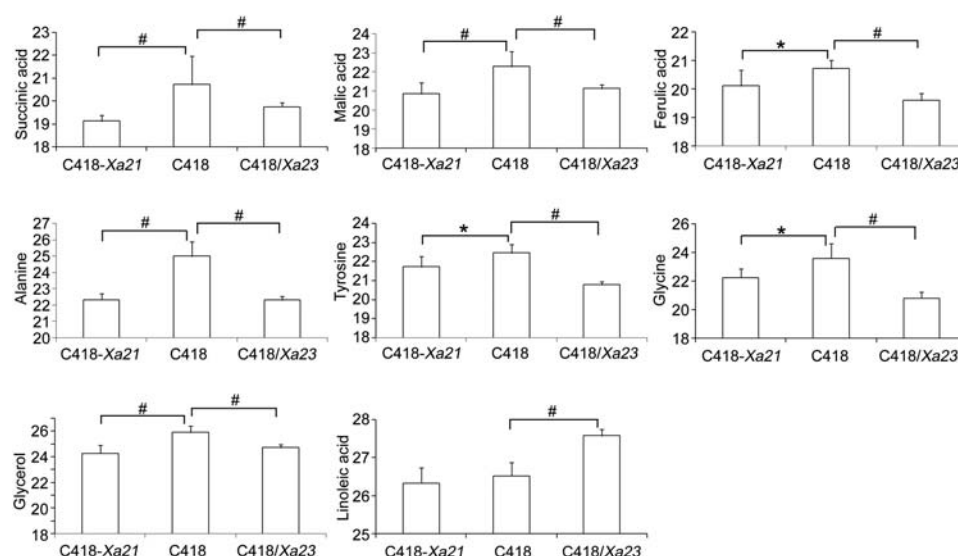


Figure 5 One-way analysis of variance of Log 2 (GC-Quad data) of metabolic differences in two bacterial blight-resistant rice variants compared with their counterpart C418 plant. Error bars represent one standard deviation. * $P < 0.05$; # $P < 0.001$. $n = 12$ for each rice genotype.

were widely affected by insertion of the *cryIac* and *sck* genes. Jiao *et al.* [27] showed that glycine decreased 23%–74% in three transgenic rice seeds compared with their counterparts and surmised that the nutritional value of transgenic rice was decreased. In this study, amino acids (glycine, alanine, and tyrosine), organic acids (succinic acid, malic acid, and ferulic acid), and glycerol were the common, novel discriminatory metabolites decreased in two genetically modified rice grains (**Fig. 5**). The reason for unintended compositional alterations and unintended changes of physical characteristic in GM rice compared with the progenitor line might be related to the genetic transformation and introduced foreign gene or its product, which may interfere with the metabolic pathway by interacting with enzymes and bringing about an accumulation or disappearance of metabolites in the host cells [27]. Recently, Chen *et al.* [40] identified a new class of sugar transporters SWEETs and discovered that the rice homologues SWEET11 and SWEET14 are specifically exploited by bacterial pathogens for virulence by means of mediating glucose transport. The mutagenesis in a unique *fbab* (fructose-bisphosphate aldolase) gene of *Xanthomonas oryzae* pv. *oryzicola* led the pathogen unable to use pyruvate and malate and reduced extracellular polysaccharide production, i.e. carbohydrates play critical roles in rice pathogen *X. oryzae* pv. *oryzicola* [41]. Furthermore, in our study, the common difference in physical characteristics and metabolic profile variation of both BB resistant rice seeds may be induced by the defense response owing to foreign gene expression. Although our analyses enable differentiation of metabolites of two BB resistant rice genotypes and their susceptible counterpart, the coherent interpretation

of metabolite profiles and BB defense phenotypes needs to be elucidated by additional studies because of the technological limitations.

The food safety of GM crop plants is a controversial topic worldwide. Unbiased, discovery-based metabolomics analyses yield novel insights into the response to diverse and foreign gene manipulation. Defining the composition of biotech products is a key step in safety assessment, especially in the context of substantial equivalence and comparative risk assessment [42]. Brown rice metabolome analysis using GC/MS confirmed that the levels of various metabolites were altered in GM variants compared with their parental counterpart. Our results indicated that the C418/Xa23 variety possessed a distinct metabolite fingerprint compared with C418-Xa21 and C418, i.e. the transgenic variant C418-Xa21 in this study appear substantially equivalent to traditional cultivar C418 apart from the same targeted changes in two resistant rice varieties.

Molecular MAS has increasingly become an important measure to modify and select rice varieties. In fact, large variation in the metabolite profiles among conventional cultivars has been found, and those differences were never sought as desired traits in traditional breeding programs [43]. Since new hybrid rice varieties generated by MAS or conventional breeding techniques, for example C418/Xa23, have to be positive control when the risk of transgenic rice C418-Xa21 is assessed before popularizing in subsequent studies. We suggest that only significant differences that fall outside the boundaries of natural variability among a range of equivalent cultivars and not solely the parental line should need further investigation with respect to safety or nutritional impact.

Acknowledgements

The authors thank Prof. Dr Wen-Xue Zhai (Institute of Genetics and Developmental Biology, Chinese Academy of Sciences) for kindly providing the seed materials, and Dr Fuquan Yang, Dr Wei Chen, and Dr Xiaorong Ran for their useful interpretation of MPP software.

Funding

This study was supported grants from the Basic Research on Breed Improvement of the Important Tropical Crop (Cassava) (2010CB126600) and the Genetically Modified Organisms Breeding Major Projects (2008ZX08012001).

References

- Mew TW. Current status and future prospects of research on bacterial blight of rice. *Ann Rev Phytopathol* 1987, 25: 359–382.
- Zhang Q. Highlights in identification and application of resistance genes to bacterial blight. *Chin J Rice Sci* 2005, 19: 453–459.
- Chen S, Lin XH, Xu CG and Zhang Q. Improvement of bacterial blight resistance of ‘Minghui 63’, an elite restorer line of hybrid rice, by molecular marker-assisted selection. *Crop Sci* 2000, 40: 239–244.
- Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T and Gardner J, *et al.* A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* 1995, 270: 1804–1806.
- Khush GS, Bacalango E and Ogawa T. A new gene for resistance to bacterial blight from *O. longistaminata*. *Rice Genet Newslett* 1990, 7: 121–122.
- Zhang Q, Wang CL, Zhao KJ, Yang WC, Qiao F, Zhou YL and Liu GC. Development of near-isogenic line CBB23 with a new resistance gene to bacterial blight in rice and its application. *Chin J Rice Sci* 2002, 16: 206–210.
- Yang ZY, Zhang ZX, Wei YL, Zhao YC and Gao Y. Breeding and characteristics of japonica type wide compatibility line C418. *Hybrid Rice* 1998, 13: 31–32.
- Sana TR, Fischer S, Wohlgemuth G, Katrekar A, Jung KH, Ronald PC and Fiehn O. Metabolomic and transcriptomic analysis of the rice response to the bacterial blight pathogen *Xanthomonas oryzae* pv. *oryzae*. *Metabolomics* 2010, 6: 451–465.
- Zhai W, Chen C, Zhu X, Chen X, Zhang D, Li X and Zhu L. Analysis of T-DNA- *Xa21* loci and bacterial blight resistance effects of the transgene *Xa21* in transgenic rice. *Theor Appl Genet* 2004, 109: 534–542.
- Bai H, Li LY and Liu GZ. Advances of rice bacterial blight disease resistance gene *Xa21*. *Yi Chuan* 2006, 28: 745–753.
- Zheng JT, Tu SH, Zhang JF, Zhao KJ, Zhang SJ and Xie HA. Breeding of restorer lines of hybrid rice with bacterial blight resistance gene *Xa23* by using marker-assisted selection. *Chin J Rice Sci* 2009, 23: 437–439.
- Xia Z, Zhao X, Fan H, Jin L, Gao L, Luo Y and Zhai W. Improving of bacterial blight resistance of hybrid rice parents through marker assisted selection of the *Xa23* gene. *Mol Plant Breeding* 2010, 8: 652–656.
- Li XB, Yi CD, Zhai WX, Yang ZY and Zhu LH. A genetically modified japonica restorer line, C418-*Xa21*, and its hybrid rice with bacterial blight resistance. *Sheng Wu Gong Cheng Xue Bao* 2001, 17: 380–384.
- Kok EJ and Kuiper HA. Comparative safety assessment for biotech crops. *Trends Biotechnol* 2003, 21: 439–444.
- Fernie AR and Schauer N. Metabolomics-assisted breeding: a viable option for crop improvement? *Trends Genet* 2009, 25: 39–48.
- Taylor J, King RD, Altmann T and Fiehn O. Application of metabolomics to plant genotype discrimination using statistics and machine learning. *Bioinformatics* 2002, 18(Suppl. 2): S241–248.
- Fiehn O. Metabolomics—the link between genotypes and phenotypes. *Plant Mol Biol* 2002, 48: 155–171.
- Yi LZ, He J, Liang YZ, Yuan DL and Chau FT. Plasma fatty acid metabolic profiling and biomarkers of type 2 diabetes mellitus based on GC/MS and PLS-LDA. *FEBS Lett* 2006, 580: 6837–6845.
- Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J and Nikolau BJ, *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci* 2004, 9: 418–425.
- Fernie AR, Trethewey RN, Krotzky AJ and Willmitzer L. Metabolite profiling: from diagnostics to systems biology. *Nat Rev Mol Cell Biol* 2004, 5: 763–769.
- Schauer N, Steinhauser D, Strelkov S, Schomburg D, Allison G, Moritz T and Lundgren K, *et al.* GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett* 2005, 579: 1332–1337.
- Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN and Willmitzer L. Metabolite profiling for plant functional genomics. *Nat Biotechnol* 2000, 18: 1157–1161.
- Roessner-Tunali U, Hegemann B, Lytovchenko A, Carrari F, Bruedigam C, Granot D and Fernie AR. Metabolic profiling of transgenic tomato plants overexpressing hexokinase reveals that the influence of hexose phosphorylation diminishes during fruit development. *Plant Physiol* 2003, 133: 84–99.
- Strelkov S, von Elstermann M and Schomburg D. Comprehensive analysis of metabolites in *Corynebacterium glutamicum* by gas chromatography/mass spectrometry. *Biol Chem* 2004, 385: 853–861.
- Lee SH, Woo HM, Jung BH, Lee J, Kwon OS, Pyo HS and Choi MH, *et al.* Metabolomic approach to evaluate the toxicological effects of nonylphenol with rat urine. *Anal Chem* 2007, 79: 6102–6110.
- Skogerson K, Harrigan GG, Reynolds TL, Halls SC, Ruebels M, Iandolo A and Pandravada A, *et al.* Impact of genetics and environment on the metabolite composition of maize grain. *J Agric Food Chem* 2010, 58: 3600–3610.
- Jiao Z, Si XX, Li GK, Zhang ZM and Xu XP. Unintended compositional changes in transgenic rice seeds (*Oryza sativa* L.) studied by spectral and chromatographic analysis coupled with chemometrics methods. *J Agric Food Chem* 2010, 58: 1746–1754.
- Zhou J, Ma C, Xu H, Yuan K, Lu X, Zhu Z and Wu Y, *et al.* Metabolic profiling of transgenic rice with cryIAC and scd genes: an evaluation of unintended effects at metabolic level by using GC-FID and GC-MS. *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, 877: 725–732.
- Lisec J, Schauer N, Kopka J, Willmitzer L and Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat Protoc* 2006, 1: 387–396.
- Jonsson P, Johansson AI, Gullberg J, Trygg J, Jiye A, Grung B and Marklund S, *et al.* High-throughput data analysis for detecting and identifying differences between samples in GC/MS-based metabolomic analyses. *Anal Chem* 2005, 77: 5635–5642.
- Steinhauser D, Usadel B, Luedemann A, Thimm O and Kopka J. CSB.DB: a comprehensive systems-biology database. *Bioinformatics* 2004, 20: 3647–3651.
- Kopka J, Schauer N, Krueger S, Birkemeyer C, Usadel B, Bergmuller E and Dormann P, *et al.* GMD@CSB.DB: the Golm Metabolome Database. *Bioinformatics* 2005, 21: 1635–1638.
- Hirai MY, Yano M, Goodenowe DB, Kanaya S, Kimura T, Awazuha M and Arita M, *et al.* Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* 2004, 101: 10205–10210.
- Lee do Y and Fiehn O. High quality metabolomic data for *Chlamydomonas reinhardtii*. *Plant Methods* 2008, 4: 7.

- 35 Ge G, Zheng X, Wu J, Ye Z and Shi C. Analysis of the conditional correlations from different genetic systems between the protein content and the appearance quality traits of Indica rice. *J Genet Genomics* 2007, 34: 129–137.
- 36 Fitzgerald MA, McCouch SR and Hall RD. Not just a grain of rice: the quest for quality. *Trends Plant Sci* 2009, 14: 133–139.
- 37 Fiehn O, Wohlgemuth G, Scholz M, Kind T, Lee do Y, Lu Y and Moon S, *et al.* Quality control for plant metabolomics: reporting MSI-compliant studies. *Plant J* 2008, 53: 691–704.
- 38 Kameshita I, Tokushige M, Izui K and Katsuki H. Phosphoenolpyruvate carboxylase of *Escherichia coli*. Affinity labeling with bromopyruvate. *J Biochem* 1979, 86: 1251–1257.
- 39 Masumoto C, Miyazawa S, Ohkawa H, Fukuda T, Taniguchi Y, Murayama S and Kusano M, *et al.* Phosphoenolpyruvate carboxylase intrinsically located in the chloroplast of rice plays a crucial role in ammonium assimilation. *Proc Natl Acad Sci USA* 2010, 107: 5226–5231.
- 40 Chen LQ, Hou BH, Lalonde S, Takanaga H, Hartung ML, Qu XQ and Guo WJ, *et al.* Sugar transporters for intercellular exchange and nutrition of pathogens. *Nature* 2010, 468: 527–532.
- 41 Guo W, Zou LF, Li YR, Cui YP, Ji ZY, Cai LL and Zou HS, *et al.* Fructose-bisphosphate aldolase exhibits functional roles between carbon metabolism and the hrp system in rice pathogen *Xanthomonas oryzae* pv. *oryzicola*. *PLoS One* 2012, 7: e31855.
- 42 Reynolds TL, Nemeth MA, Glenn KC, Ridley WP and Astwood JD. Natural variability of metabolites in maize grain: differences due to genetic background. *J Agric Food Chem* 2005, 53: 10061–10067.
- 43 Catchpole GS, Beckmann M, Enot DP, Mondhe M, Zywicki B, Taylor J and Hardy N, *et al.* Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. *Proc Natl Acad Sci USA* 2005, 102: 14458–14462.