Metabolomics reveals the effects of producing region and varieties on substance variation in characteristic rice in Yunnan (#101490)

First submission

Guidance from your Editor

Please submit by **1 Aug 2024** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for guidance.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files
Download and review all files from the materials page.

4 Figure file(s)
2 Table file(s)
1 Raw data file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- **Impact and novelty is not assessed.** Meaningful replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Metabolomics reveals the effects of producing region and varieties on substance variation in characteristic rice in Yunnan

Xuheng Nie [1], **Shuiyan Yang** [1], **Ying Guo** [1], **Xin Wang** [1], **Yunman Wen** [1], **Chao Liu** [Corresp., 2], **Fuying Liu** [1]

[1] Yunnan Provincial Academy of Food and Oil Sciences, Kunming, China

[2] College of Biological Resource and Food Engineering, Yunnan Engineering Research Center of Fruit Wine, Qujing Normal University, Qujing, China

Corresponding Author: Chao Liu
Email address: liuchao_80@163.com

The characteristic rice grown in Yunnan Province has good taste and is favored by consumers, but their metabolite composition is not clear. In this study, the metabolic profile of different rice planted in various producing regions was evaluated. A total of 1005 metabolites were identified, including nucleotides and their derivatives, amino acids and their derivatives, alkaloids, organic acids, phenolic acids, lignans and coumarins, lipids, terpenoids, quinones, flavones, tannins, and others. Procucing region and varieties can be clearly distinguished on the PCA diagram. Differential metabolites accumulated in the MSD502 vs. MSR88(138)/ LHHG(234)/ LHR88(188) comparison groups. The results in this study provide scientific information for the origin tracing and variety differentiation of raw rice materials.

Metabolomics reveals the effects of producing region and varieties on substance variation in characteristic rice in Yunnan

Xuheng Nie[1], Shuiyan Yang[1], Ying Guo, Xin Wang[1], Yunman Wen[1], Chao Liu*[2], Fuying Liu*[1]

1 Yunnan Provincial Academy of Food and Oil Sciences, Kunming, 650033, China

2 College of Biological Resource and Food Engineering, Yunnan Engineering Research Center of Fruit Wine, Qujing Normal University, Qujing, Yunnan, 655011, China

*Corresponding author：Chao Liu, E-mail: Liuchao_80@163.com

Fuying Liu, E-mail: 372243690@qq.com

Xuheng Nie and Shuiyan Yang should be regarded as joint first authors.

ABSTRACT

The characteristic rice grown in Yunnan Province has good taste and is favored by consumers, but their metabolite composition is not clear. In this study, the metabolic profile of different rice planted in various producing regions was evaluated. A total of 1005 metabolites were identified, including nucleotides and their derivatives, amino acids and their derivatives, alkaloids, organic acids, phenolic acids, lignans and coumarins, lipids, terpenoids, quinones, flavones, tannins, and others. Procucing region and varieties can be clearly distinguished on the PCA diagram. Differential metabolites accumulated in the MSD502 vs. MSR88(138)/ LHHG(234)/ LHR88(188) comparison groups. The results in this study provide scientific information for the origin tracing and variety differentiation of raw rice materials.

Keywords: Rice, Producing region, Differential metabolites, Metabolic pathways

INTRODUCTION

Rice is the most important staple food crop in Yunnan Province, which produces 40% of the total grain output with 25% of the cultivated land area, and 67% of the population of the province take rice as the staple food [1, [2]. The annual planting area is about 1 million square hectares. Most of the rice farming areas in Yunnan Plateau range from 1200 m to 2400 m above sea level, which is one of the special rice farming areas with the most biodiversity in China and even in the world. Yunnan Province has a variety of terrain and climate, which can adapt to different types of plateau characteristic rice varieties, but also has a unique rice planting area, which has an independent flavor grown in these areas. Due to the special natural ecological conditions, Yunnan has its own unique rice system, that is, special rice system. According to statistics, more than 30 quality rice varieties have formed a certain scale of planting and production in Yunnan, which is one of the quality rice farming areas with the most industrialization development prospects in China.

In recent years, quality grain project in China is closely related to the prosperity of food industry, the increase of farmers' income and the improvement of enterprises' efficiency. With the advantage of rich plateau characteristic food resources (such as, soft rice, red rice, purple rice), Yunnan accelerates the development of modern agriculture, meets the needs of consumers from the rigid demand type to the quality type, and starts to pay attention to the food quality. However, the food safety situation is changing, and economic profit-driven adulteration, shoddy goods, counterfeit products and origin violations are common, which not only damage the reputation of quality products, but also seriously harm the interests of consumers. In the face of such illegal

behavior, a variety of new analytical techniques are needed for grain quality and adulteration identification, and metabolomics may be an important technical means to solve this problem [3].

Starting from the overall level, metabolomics mainly studies the differences that the whole biological system responds to changes in certain conditions under different behavioral conditions, and then reflects the overall changes through the monitoring of biomarkers of differences [4]. It is widely used in nutrition science [5], food adulteration quality identification, food origin traceability [6], disease diagnosis [7], toxicology [8], plant metabolism and response mechanism [9], safety of food [10] (such as rice [11], potato [12], tomato [13]) and so on. The metabolome can more accurately reflect the nutrient composition and content of plant cells, so it plays an important role in evaluating food quality [14, [15]. In this study, 15 rice samples (divided into 5 groups) were used as research materials, and the metabolites of rice were detected using an extensively employed method in the field of metabolomics, known as ultra high performance liquid chromatography-tandem mass spectrometry(UPLC-MS/MS), and their differences were compared and analyzed. By examining various metabolites, variations in phenolic acids, alkaloids, flavonoids, lipids, amino acids and their derivatives were identified. The aim of this study was to provide possible solutions to the problems of identifying quality rice varieties, origin tracing and adulteration in Yunnan.

MATERIALS AND METHODS

Study materials

Fifteen rice samples from Lianghe County and Mangshi City of Dehong in Yunnan Province were collected and divided into 5 groups for metabolic study, namely, Mangshi Diantun 502 (MSD502), Lianghe Soft Rice 88 (LHR88), Mangshi Soft 88 (MSR88), Lianghe Honggu (LHHG) and mixed samples, with 3 biological replicates in each group. These rice were harvested in September 2020 and stored at -20℃.

Preparation and extraction of rice samples

The specimens were subject to freeze-drying under vacuum conditions using the lyophilizer (Scientz-100F), followed by pulverization into a fine powder using the grinder (MM 400, Retsch) at a frequency of 30Hz for a duration of 1.5 minutes. A quantity of 100 mg powder was precisely weighed and dissolved in 1.2 mL of methanol solution containing 70%, with vortexing performed every half an hour for a period of 30 seconds each time, totaling six cycles. Subsequently, the sample was refrigerated overnight at a temperature of 4℃. The centrifuge operated at a speed of 12000 *rpm* for ten minutes to separate the supernatant from the solid residue. After centrifugation, filtration through microporous membranes with pore size measuring at 0.22 μm was conducted on the obtained sample solution. Finally, storage in designated vials ensued prior to UPLC-MS/MS analysis.

Parameters for UPLC

The UPLC-ESI-MS/MS system utilized for the analysis of the sample extracts consisted of a SHIMADZU Nexera X2 UPLC and an Applied Biosystems 4500 Q TRAP mass spectrometer. The analytical conditions involved the use of an Agilent SB-C18 column (100×2.1 mm, 1.8 μm).

A gradient program was employed for sample processing, starting with a composition of 95% A and 5% B. Over a period of 9 minutes, there was a gradual transition to a composition of 5% A and 95% B, which was maintained for another minute. Subsequently, within just over one minute, the composition was adjusted to be predominantly composed of 95% A and only 5% B, which remained constant for approximately three minutes. The flow rate used during analysis was set at a value equivalent to injecting 0.35 ml per minute into the system. To maintain optimal temperature conditions, the column oven temperature was maintained at a steady state of 40 °C throughout the analysis process while ensuring that each injection contained precisely four microliters.

Parameters for MS/MS

The acquisition of LIT and triple quadrupole (QQQ) scans was performed using an AB4500 Q TRAP UPLC/MS/MS System, which is a triple quadrupole-linear ion trap mass spectrometer. The system was equipped with an ESI Turbo Ion-Spray interface and operated in both positive and negative ion mode. Control of the instrument was carried out using Analyst 1.6.3 software developed by AB Sciex. The operational parameters for the ESI source were as follows: the temperature of the ion source was set at 550℃; positive ion mode, the ion spray voltage (IS) was 5500 V, while in negative ion mode it was -4500 V; gas I (GSI), gas II (GSII), and curtain gas (CUR) pressures were adjusted to 50, 60, and 25.0 psi respectively; collision-activated dissociation (CAD) was set to high. The instrument underwent tuning and mass calibration using polypropylene glycol solutions with concentrations of 10 μmol/L and 100 μmol/L in QQQ and LIT modes respectively. MRM experiments were conducted for QQQ scans with medium collision gas (nitrogen). DP and CE optimization were performed for individual MRM

transitions by adjusting DP and CE values accordingly. A specific set of MRM transitions were monitored during each period based on the eluted metabolites within that period.

Analysis of Metabolite concentration with multiple variables

Utilizing a self-constructed metabolite database and relevant mass spectrometry database from Metware Biotechnology Co., Ltd (Wuhan, China), the detection of MRM metabolites was conducted through multi-peak mapping. The triple quadrupole technique was employed to screen characteristic ions for each substance, and subsequently, the detector measured the signal intensity of these ions.

This study employed multivariate statistical analysis, utilizing R (http://www.r-project.org/) for conducting principal component analysis (PCA) and Cluster Analysis (CA) on five distinct sample groups. Orthogonal partial least squares discriminant analysis (OPLS-DA) and OPLS-DA S-plot were utilized to identify orthogonal metabolites between the two samples. Variable importance in project (VIP) values, one-dimensional statistical P-values, and differential multiples were used to identify different metabolites. A cluster system was employed to screen the diverse metabolites in the samples, which were subsequently subjected to correlated pathway analysis using the KEGG (Kyoto Encyclopedia of Genes and Genomes) database website.

Analysis of KEGG annotations and enrichment

The KEGG database is utilized to annotate various metabolites and link them to the KEGG pathways database. Metabolite concentration analysis (MSEA) was conducted on pathways where significant regulation of metabolites occurred. The significance of these findings was assessed using the P value derived from a hypergeometric test.

148

149 **RESULTS**

150

151 **Metabolic composition profiling**

152

153 Metabolome analysis was performed on 15 rice samples. According to the self-constructed
154 metabolite database and relevant mas spectrometry database, a total of 1005 metabolites
155 belonging to 12 different classes were successfully identified (Table 1), The metabolites included
156 63 Nucleotides and their derivatives，110 Amino acids and their derivatives，87 alkaloids, 80
157 Organic acids，137 Phenolic acids，21 Lignans and coumarins，172 lipids，23 terpenoids，5
158 quinones，184 flavonoids，6 tannins，and 117 Others metabolites.

159

160 **Multivariate analysis of metabolite profiles**

161

162 Multivariate statistical analysis was conducted to evaluate the variations in metabolic profiles
163 across various rice samples. The rice samples exhibited distinct categorization into 5 groups,
164 with noticeable segregation between the groups, suggesting substantial variations in metabolite
165 composition among these five groups of rice samples. Furthermore, the three biological
166 replicates demonstrated close clustering together (Fig. 2A). Through conducting PCA analysis
167 on a set of rice samples, we were able to assess the extent of variation both within and between
168 the five groups of rice samples. The contribution rate of PC1 was 27.32%, and that of PC2 was
169 18.08%. The separation between the five groups of rice samples was obvious in the two-
170 dimensional diagram, which can generally reflect the differences across the five distinct sample
171 categories. The close clustering of the three biological replicates for each sample suggests that

the test exhibited high repeatability and reliability, thereby ensuring reliable quality control analysis (Fig. 2B).

Analysis of differential metabolites

The OPLS-DA method was used to screen for variables that contributed to the differences between the five groups. In this research, the OPLS-DA model was utilized to analyze and compare metabolite composition in rice samples, aiming to identify variations between different groups, LHR88 and LHHG ($R^2X = 0.743$, $R^2Y = 1$, $Q^2 = 0.975$), MSD502 and LHHG ($R^2X = 0.682$, $R^2Y = 1$, $Q^2 = 0.967$), MSD502 and LHR88 ($R^2X = 0.613$, $R^2Y = 1$, $Q^2 = 0.944$), MSD502 and MSR88 ($R^2X = 0.628$, $R^2Y = 1$, $Q^2 = 0.918$), MSR88 and LHR88($R^2X = 0.613$, $R^2Y = 1$, $Q^2 = 0.954$) (Fig. 2). All comparison groups had Q^2 values of more than 0.9, indicating that the model was stable. The OPLS-DA score map showed that the five groups of rice samples were well separated in pairs, suggesting that the metabolic phenotypes of the five groups of rice samples were significantly different.

According to the screening criteria, FC ≥ 2 to ≤ 0.5 and VIP ≥ 1, of differential metabolites, there was a significant difference in the expression levels of metabolites between LHR88 and LHHG, with 228 metabolites showing down-regulation and 99 metabolites showing up-regulation. The identified down-regulated and up-regulated metabolites mainly belonged to categories such as lipids, flavonoids, phenolic acids, amino acids, and their derivatives (Table 2). Terpenoids, lignans and coumarins were the least up-regulated, while quinones and blends were the least down-regulated. There was a significant decrease in the levels of 174 metabolites and an increase in the levels of 60 metabolites when comparing MSD502 to LHHG (Table 2). The most down-regulated and up-regulated metabolites were flavonoids, amino acids and their derivatives,

phenolic acids, etc. The least down-regulated metabolites were quinones and tannins, while the least up-regulated metabolites were terpenoids, other metabolites, nucleotides and their derivatives, lipins and coumarins. There was a significant difference in the expression levels of 79 metabolites down-regulated and 109 metabolites up-regulated when comparing MSD502 to LHR88(Table 2). The most down-regulated metabolites were flavonoids and phenolic acids, while the most up-regulated metabolites were lipids and alkaloids. Quinones, blends, lignans and coumarins were least down-regulated and up-regulated.Between MSD502 vs MSR88, there were 31 down-regulated metabolites and 107 up-regulated metabolites(Table 2). The most down-regulated metabolites were amino acids and their derivatives, while the most up-regulated metabolites were flavonoids and lipids. Quinones, blends, lignans and coumarins were least down-regulated and up-regulated. There was a significant decrease in the levels of 139 metabolites and an increase in the levels of 53 metabolites when comparing MSR88 to LHR88 (Table 2). The most down-regulated metabolites were flavonoids, phenolic acids, etc., and the most up-regulated metabolites were amino acids and their derivatives, alkaloids, lipids, etc. Terpenoids, quinones and blends were least down-regulated and up-regulated.

We analyzed the pathway enrichment of metabolites through KEGG database. The pathways differential metabolites remarkably enriched in the group of MSD502 vs. MSR88 were “linoleic acid metabolism”, “cutin, suberine and wax biosynthesis”, “citrate cycle (TCA cycle)”, “carbon metabolism”, “arginine biosynthesis” and “alanine, aspartate and glutamate metabolism” (P<0.05)(Fig.3B). The pathways differential metabolites remarkably enriched in the group of MSD502 vs. LHHG were “starch and sucrose metabolism”, “propanoate metabolism”, “lysine degradation”, “glutathione metabolism” and “amino sugar and nucleotide sugar metabolism” (P<0.05)(Fig.3C). The pathways differential metabolites remarkably enriched in the group of MSD502 vs. LHR88 were “riboflavin metabolism”, “pentose and glucuronate interconversions”,

219 “linoleic acid metabolism”, “biosynthesis of unsaturated fatty acids” and “alpha-linolenic acid
220 metabolism” (P<0.05)(Fig.3D).
221 In this research, a total of 406 distinct metabolites were identified in these groups. By comparing
222 the different groups using a Venn diagram (Fig.3A), only 16 unique metabolites (MSD502 vs.
223 MSR88/ LHHG/ LHR88) were found to be common among them. These shared metabolites
224 include flavonoids, organic acids, amino acids and their derivatives, alkaloids, lipids, phenolic
225 acids, terpenoids, and others. The findings indicate significant variations in the metabolic
226 profiles of four rice varieties obtained from diverse sources.
227 We compared the variation ratios of metabolite quantitative information within each group and
228 subsequently processed these variations (log2FC). The first 20 changes in each group (10 up-
229 regulated and 10 down-regulated) were differentially expressed metabolic components, as shown
230 in Figure 4. Ten significantly upregulated metabolites in the group of LHR88 vs. LHHG
231 contained three tannins (procyanidin B1, procyanidin B3, arecatannin A2), three flavonoids
232 (catechin, phloretin-4'-O-glucoside (Trilobatin), epicatechin gallate), three phenolic acide
233 (protocatechuic Acid Methyl Ester, 4-(3,4,5-Trihydroxybenzoxy)benzoic acid, salicylic acid) and
234 one organic acids(Jasmonic acid,)
235 (Figure 4A). There are eight metabolites with differences greater than 10 times. The
236 significantly down-regulated metabolites in the group of LHR88 vs. LHHG included four
237 alkaloids (indirubin, feruloylcholine glucoside, hexadecylsphingosine, dhurrin), four flavonoids
238 (quercetin-5-O-β-D-glucoside, wogonin (5,7-Dihydroxy-8-Methoxyflavone) , feruloylcholine,
239 genkwanin (Apigenin 7-methyl ether)), one phenolic acids (p-Coumaroylmalic acid) and one
240 lipids (7-O-Methylnaringenin) (Figure 4A). The differences of these metabolites were more than
241 10 times.

Ten significantly upregulated metabolites in the group of MSD502 vs. LHHG contained three tannins (procyanidin B1, procyanidin B3, arecatannin A2), five flavonoids (catechin, phloretin-4'-O-glucoside (Trilobatin), epicatechin gallate, naringenin-4'-O-glucoside, Diosmetin-6-C-glucoside), two phenolic acids (protocatechuic Acid Methyl Ester, 4-(3,4,5-Trihydroxybenzoxy)benzoic acid)(Figure 4B). There are seven metabolites with differences greater than 10 times. The significantly down-regulated metabolites in the group of MSD502 vs. LHHG included seven flavonoids (5,4'-Dihydroxy-7-methoxyflavanone (Sakuranetin), dihydroquercetin(Taxifolin), quercetin-3-O-galactoside (Hyperin), quercetin-5-O-β-D-glucoside, wogonin (5,7-Dihydroxy-8-Methoxyflavone), genkwanin (Apigenin 7-methyl ether), 7-O-Methylnaringenin), one nucleotides and derivatives (uric acid), one alkaloids (dhurrin) and one phenolic acids (p-Coumaroylmalic acid)(Figure 4B). There are seven metabolites with differences greater than 10 times.

Ten significantly upregulated metabolites in the group of MSD502 vs. LHR88 included four alkaloids (feruloylcholine, feruloylcholine glucoside, indirubin, sinapine), three lipids (13(S)-HODE;13(S)-Hydroxyoctadeca-9Z,11E-dienoic acid, 9S-Hydroxy-10E,12Z-octadecadienoic acid, 12,13-Epoxy-9-Octadecenoic Acid), one flavonoids (kaempferol-7-O-glucoside), one phenolic acids (3'-Hydroxy-3,4,5-trimethoxybibenzyl) and one others (machilusolide D) (Figure 4C). There are three metabolites with differences greater than 10 times. The significantly down-regulated metabolites in the group of MSD502 vs. LHR88 included one lipids (1-α-Linolenoyl-glycerol-3-O-glucoside), four phenolic acids (arbutin, gallacetophenone, 1-O-p-Coumaroyl-β-D-glucose, N-Acetyl-L-leucine), one organolic acids (2-Hydroxy-3-phenylpropanoic acid), two animo acids and derivatives (vanillic acid-4-O-glucoside, L-Theanine), one flavonoids (quercetin-3-O-rutinoside (Rutin)) and one organic acids (jasmonic acid) (Figure 4C).

Ten significantly upregulated metabolites in the group of MSD502 vs. MSR88 comprised four alkaloids (feruloylcholine, feruloylcholine glucoside, indirubin, sinapine), four flavonoids (catechin, phloretin-4'-O-glucoside (Trilobatin), kaempferol-7-O-glucoside, epicatechin gallate), one amino acids and derivatives (nicotinuric acid) and one lipids (lysoPC 20:5) (Figure 4D). There are five metabolites with differences greater than 10 times. The significantly down-regulated metabolites in the group of MSD502 vs. MSR88 included two flavonoids (3-Hydroxy-4',5,7-Trimethoxyflavanone, 7-Methoxyisoflavone), three amino acids and derivatives (glutathione reduced form, L-Theanine, N-Acetyl-L-leucine), four nucleotides and derivatives (Uridine 5'-diphospho-D-glucose, inosine 5'-monophosphate, adenosine 5'-monophosphate, uridine 5'-monophosphate) and one others (2-(2'-hydroxypropyl)-5-methyl-7-hydroxychromone) (Figure 4D).

Ten significantly upregulated metabolites in the group of MSR88 vs. LHR88 contained three others (2-(2'-hydroxypropyl)-5-methyl-7-hydroxychromone, ribitol, D-Arabitol), one flavonoids (7-Methoxyisoflavone), four alkaloids (indirubin, 2-Amino-4,5-dihydro-1H-imidazole-4-acetic acid, 1,4-Dihydro-1-Methyl-4-oxo-3-pyridinecarboxamide, stachydrine) and two amino acids and derivatives (D-Proline betaine, 1-Methylpiperidine-2-carboxylic acid) (Figure 4E). The significantly down-regulated metabolites in the group of MSR88 vs. LHRR88 contained two phenolic acids (vanillic acid-4-O-glucoside, 5-O-Feruloyl quinic acid glucoside,), seven flavonoids (yuanhuanin, rhoifolin, 4'-O-Glucosylvitexin, "Vitexin-2""-O-galactoside", saponarin(Isovitexin-7-O-glucoside), isosaponarin(Isovitexin-4'-O-glucoside), quercetin-3-O-rutinoside (Rutin)) and one organic acids (jasmonic acid)(Figure 4E).

DISCUSSION

Metabolites identified in rice samples

In the study, we conducted a widely-targeted metabolomics analysis of 15 rice samples and provided their comprehensive metabolic profile. Previous studies used near infrared spectroscopy and chemometrics to study wheat flour doping and tea origin information, which are relatively easy to achieve, but it is difficult to achieve the qualitative and quantitative components of plant cell metabolism. The metabolome can more accurately reflect the nutrient composition and content of plant cells, so it plays an important role in evaluating food quality [16, [17].

Based on the results of metabolomics, a total of 1005 metabolites were identified by qualitative and quantitative analysis based on ion pair information of compounds in rice samples. Compared with 732 metabolites identified in colored rice [18] and 672 metabolites identified in millet whole grains [19], we identified a wider variety of metabolites in rice. The metabolites included 63 Nucleotides and their derivatives，110 Amino acids and their derivatives，87 alkaloids, 80 Organic acids，137 Phenolic acids，21 Lignans and coumarins，172 lipids，23 terpenoids，5 quinones，184 flavonoids，6 tannins，and 117 Others metabolites (Table 1). Most of compounds were flavonoids，lipids and phenolic acids. A low percentage of compounds were tannins，quinones. Most of differential metabolites were flavonoids, lipids, and phenolic acids. However, Lignans and coumarins, quinines, and tannins showed a small change not only in different cultivars but also in each treatment (Table 2). Among the different metabolites of the five groups of rice samples, most substances were flavonoids, phenolic acids, lipids, amino acids and their derivatives. LHHG belongs to colored rice, and its phenolic and flavonoid compounds are more highly expressed than white rice. The results show that the metabolites that cause the difference between colored rice and non-colored rice are quite different. Flavonoids exhibit superior antioxidant and anti-inflammatory characteristics [20, [21, [22] and are important as

313 they contributed to people's health via gut microbiota regulation [23, [24]. As an important
314 component of plant phenolic compounds [25], they have antiviral, anti-free radical, anti-
315 oxidation, reducing lipid peroxidation, anti-inflammation, anti-aging, anti-cancer and prevention
316 of cardiovascular diseases [9]. The researchers [26] found that phenolic compounds, including
317 their subflavonoids, are widely found in grains, legumes and nuts. Previous study [27] have
318 demonstrated that flavonoids are mainly present in cotyledon and bran, but not in ground grains.
319 In addition, different metabolites such as lipids, amino acids and their derivatives, organic acids
320 were also more in colored and non-colored rice. In non-colored rice, MSD502 showed that
321 significant differences in different metabolites (such as lipids, phenolic acids, alkaloids and
322 flavonoids) are different from other rice varieties.

323 **Differential metabolites between MSD502 and MSR88, LHHG, LHR88.**

324 In present study, a total of 406 distinct metabolites were identified in these groups, only 16
325 unique metabolites (MSD502 vs. MSR88/ LHHG/ LHR88) were found to be common among
326 them (Fig.3A). The pathways of different metabolites in different comparison groups were
327 significant differences. For example, the pathways differential metabolites remarkably enriched
328 in the group of MSD502 vs. MSR88 were "linoleic acid metabolism", "cutin, suberine and wax
329 biosynthesis", "citrate cycle (TCA cycle)", "carbon metabolism", "arginine biosynthesis" and
330 "alanine, aspartate and glutamate metabolism" (P<0.05)(Fig.3B). The pathways differential
331 metabolites remarkably enriched in the group of MSD502 vs. LHHG were "starch and sucrose
332 metabolism", "propanoate metabolism", "lysine degradation", "glutathione metabolism" and
333 "amino sugar and nucleotide sugar metabolism" (P<0.05)(Fig.3C). The reasons for these
334 different results may be different ecological environment and climate characteristics. Therefore,
335 it is of great significance to promote the planting of plateau characteristic rice in different places
336 to meet the nutritional and flavor needs of different consumers.

To further analyze the characteristic metabolites between different treatment groups, ten significantly upregulated metabolites and ten significantly downregulated metabolites were chosen. The major upregulated metabolites in the group of LHR88 vs. LHHG were tannins, flavonoids, phenolic acids and organic acids, and the major downregulated metabolites were flavonoids, phenolic acids, alkaloids and lipids. The major upregulated metabolites in the group of MSD502 vs. LHHG were tannins, flavonoids and phenolic acids, and the major downregulated metabolites were nucleotides and derivatives, flavonoids, alkaloids and phenolic acids. The major upregulated metabolites in the group of MSD502 vs. LHR88 were alkaloids, flavonoids, lipids, phenolic acids and others, and the major downregulated metabolites were lipids, Phenolic acids, organic acids, amino acids and derivatives, flavonoids. The major upregulated metabolites in the group of MSD502 vs. MSR88 were alkaloids, flavonoids, amino acids and derivatives, flavonoids, lipids, and the major downregulated metabolites were flavonoids, amino acids and derivatives, nucleotides and derivatives, flavonoids, others. The major upregulated metabolites in the group of MSR88 vs. LHR88 were flavonoids, alkaloids, amino acids and derivatives, others, and the major downregulated metabolites were phenolic acids, flavonoids, phenolic acids, organic acids and others. These findings shown that notable variances in the metabolites among distinct rice varieties originating from identical production regions or within the same variety across diverse production areas are common which is consistent with the findings of previous study [28]. It can be seen that compared with non-colored rice, the characteristic flavone in LHHG is proanthocyanidins [29] and does not contain anthocyanidins [30]. The study [31] have shown that proanthocyanidins mainly exist in rice bran. These results are of great significance for preventing the adulteration of the characteristic rice in the plateau and distinguishing the planting areas.

CONCLUSION

In this study, a total of 1005 metabolites belonging to 12 different classes were detected in the analysis of 15 rice samples. These classes encompassed nucleotides and their derivatives, amino acids and their derivatives, alkaloids, organic acids, phenolic acids, lignans and coumarins, lipids, terpenoids, quinones, flavonoids, tannins as well as other categories. The pathways and enrichment of differential metabolites in the MSD502 vs. MSR88/ LHHG/ LHR88 comparison groups are significantly different, and only 16 unique metabolites were found to be common among them. The present work is helpful to understand the metabolite composition and functional compounds of different varieties of rice, and provides a basis for the breeders to screen out nutrition-rich and functional quality rice varieties through comparative evaluation. Different varieties of rice planted in different producing area can be clearly separated on PCA plot, which provides a good idea for identifying the adulteration problem of quality rice varieties in Yunnan Province.Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

This research has been confirmed for publication in the journal.

Competing interests

The authors have no conflicts of interest.

Funding

This work was supported by the National Natural Science Foundation of China (32060710).

References

1. Deng A, Yang C, Chen Q, Shi Y, Chen Y, Congying LI, Shilin XU, Mao G, Guiy-Ong LI, Xia Q. Effects of Different Fertilization Methods on Yield and Photosynthetic Material Production of Direct Seeding Rice at Different Densities. China Rice. 2017; 23:123-129
2. Deng A, Yang C, Luo J, Guiyong LI, Zhu H, Xia Q, Yaping PU, Shuqing MA. Present Situation of Rice Production in Yunnan Province and Cuntermeasures for Green Development. China Rice. 2019; 25
3. Deng ML, Xia YP, Zhao B, Da-Wei TU, Yang XS. Application of Food Omics Technology in Risk Control About Food Safety. Food and Fermentation Technology. 2015; 51:95-99
4. Hui-Run T, Yu-Lan W. Metabonomics: a Revolution in Progress. ProgBiochem Biophys. 2006; 33:401-417
5. Tebani A, Bekri S. Paving the Way to Precision Nutrition Through Metabolomics. Frontiers in Nutrition. 2019; 6. https://doi.10.3389/fnut.2019.00041.
6. Hao J, Jiang J, Mao T, Sun XD. Research progress of metabonomics in food safety risk monitoring. Journal of Food Safety & Quality. 2017; 8
7. GA Nagana Gowda SZ, Haiwei Gu, Vincent Asiago, Narasimhamurthy Shanaiah, Daniel Raftery. Metabolomics-based methods for early disease diagnostics. Expert Review of Molecular Diagnostics. 2008; 8:617-633. https://doi.10.1586/14737159.8.5.617.
8. Ramirez T, Daneshian M, Kamp H, Bois FY, Clench MR, Coen M, Donley B, Fischer SM, Ekman DR, Fabian E. t4 report : metabolomics in toxicology and preclinical research. Altex. 2013; 30:209-225. https://doi.10.14573/altex.2013.2.209.
9. Xia LI, Yi-Bo XU, Sun J, Qian Y, Shang XJ. Extraction of Total Flavonoids from Brown Rice and Its Scavenging Capacity on Hydroxyl Free Radicals. Journal of Jilin Agricultural University. 2014; 36:347-351
10. Development OfEC-oa. Safety evaluation of foods derived through modern biotechnology: concepts and principles [M]. Paris: Organization for Economic. 1993;
11. Chang Y, Zhao C, Zhu Z, Wu Z, Zhou J, Zhao Y, Lu X, Xu G. Metabolic profiling based on LC/MS to evaluate unintended effects of transgenic rice with cry1Ac and sck genes. Plant Mol Biol. 2012; 78:p.477-487. https://doi.10.1007/s11103-012-9876-3.
12. Catchpole GS, Beckmann M, Enot DP, Mondhe M, Zywicki B, Taylor J, Hardy N, Smith A, King RD, Kell DB. Hierarchical metabolomics demonstrates substantial compositional similarity between genetically modified and conventional potato crops. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102:14458-14462. https://doi.10.1073/pnas.0503955102.

13. Le Gall Gl, Dupont MS, Mellon FA, Davis AL, Collins GJ, Verhoeyen ME, Colquhoun IJ. Characterization and Content of Flavonoid Glycosides in Genetically Modified Tomato (Lycopersicon esculentum) Fruits. Journal of Agricultural & Food Chemistry. 2003; 51:2438-2446. https://doi.10.1021/jf025995e.
14. MichaelR.LaFrano. Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. Metabolomics. 2009; 5:435-458. https://doi.10.1007/s11306-009-0168-0.
15. Fiehn O. Metabolomics – the link between genotypes and phenotypes. Plant Mol Biol. 2002; 48:155-171. https://doi.10.1023/A:1013713905833.
16. Yi-Ru XU, Cui-Ling L, Xiao-Rong S, Sheng-Nan WU, Xiu-Li D. Study on Determination of wheat flour quality and doping identification method based on spectrum technology NIR-MIR. Food Science. 2024;
17. Ana, Marcos, Andrew, Fisher, Gerry, Rea, Steve, J., Hill. Preliminary study using trace element concentrations and a chemometrics approach to determine the geographical origin of tea. Journal of Analytical Atomic Spectrometry. 1998; 13:521-525. https://doi.10.1039/A708658J.
18. Zhang L, Cui D, Ma X, Han B, Han L. Comparative analysis of rice reveals insights into the mechanism of colored rice via widely targeted metabolomics. Food Chem. 2023; 399:133926. https://doi.10.1016/j.foodchem.2022.133926.
19. Li W, Wen L, Chen Z, Zhang Z, Guo Y. Study on metabolic variation in whole grains of four proso millet varieties reveals metabolites important for antioxidant properties and quality traits. Food Chem. 2021; 357:129791. https://doi.10.1016/j.foodchem.2021.129791.
20. Alves GH, Ferreira CD, Vivian PG, Fernandes Monks JL, Elias MC, Vanier NL, De Oliveira M. The revisited levels of free and bound phenolics in rice: Effects of the extraction procedure. Food Chem. 2016; 208:116-123. https://doi.10.1016/j.foodchem.2016.03.107.
21. Min B, Gu L, Mcclung AM, Bergman CJ, Chen MH. Free and bound total phenolic concentrations, antioxidant capacities, and profiles of proanthocyanidins and anthocyanins in whole grain rice (Oryza sativa L.) of different bran colours. Food Chem. 2012; 133:715–722. https://doi.10.1016/j.foodchem.2012.01.079.
22. Min B, Mcclung AM, Chen MH. Phytochemicals and Antioxidant Capacities in Rice Brans of Different Color. J Food Sci. 2011; 76:C117-C126. https://doi.10.1111/j.1750-3841.2010.01929.x.
23. X.-Y. L, Meng L, Shen L, H.-F. J. Regulation of gut microbiota by vitamin C, vitamin E and β-carotene. Food Res Int. 2023. https://doi.10.1016/j.foodres.2023.112749.

24. Esfandiar Z, Hosseini-Esfahani F, Mirmiran P, Azizi F. Higher dietary flavonol and isoflavonoid intakes are associated with lower incidence of type 2 diabetes. International Journal for Vitamin and Nutrition Research. 2023:-. https://doi.10.1024/0300-9831/a000782.
25. Ren CY, Lu SW, Hong B, Zhang YL, Guan LJ, Li B, Huang WG, Lu WH. Non-targeted Metabolomic Analysis of Brown Rice and White Rice. Food science. 2022; 43:183-190.(In chinese)
26. Kris-Etherton PM, Hecker KD, Bonanome A, Coval SM, Etherton TD. Bioactive compounds in foods: their role in the prevention of cardiovascular disease and cancer. Am J Med. 2002; 113:71-88. https://doi.10.1016/S0002-9343(01)00995-0.
27. Yang Z, Nakabayashi R, Mori T, Takamatsu S, Kitanaka S, Saito K. Metabolome Analysis of Oryza sativa (Rice) Using Liquid Chromatography-Mass Spectrometry for Characterizing Organ Specificity of Flavonoids with Anti-inflammatory and Anti-oxidant Activity. Chemical and Pharmaceutical Bulletin. 2016; 64:952-956. https://doi.10.1248/cpb.c16-00180.
28. Tianxin FU, Yuchao F, Liyuan Z, Xue LI, Changyuan W. Metabonomics Study on Rice from Different Geographical Areas Based on Gas Chromatography-Mass Spectrometry Food Science (In Chinese). 2019; 40:176-180. https://doi.10.7506/spkx1002-6630-20180621-412.
29. Xya B, Ty A, Qqa B, Yd A, Js C, Xl A, Yl A, Hz A, Zz A, Ning YA. Comparison of the contents of phenolic compounds including flavonoids and antioxidant activity of rice (Oryza sativa) and Chinese wild rice (Zizania latifolia). Food Chem. 2020. https://doi.10.1016/j.foodchem.2020.128600.
30. Finocchiaro F, Ferrari B, Gianinetti A. A study of biodiversity of flavonoid content in the rice caryopsis evidencing simultaneous accumulation of anthocyanins and proanthocyanidins in a black-grained genotype. Journal of Cereal Science. 2010; 51:28-34. https://doi.10.1016/j.jcs.2009.09.003.
31. Shao Y, Bao J. Polyphenols in whole rice grain: Genetic diversity and health benefits. Food Chem. 2015; 180:86-97. https://doi.10.1016/j.foodchem.2015.02.027.

Table 1(on next page)

Overview of annotated metabolites

1

2 Table 1. Overview of annotated metabolites.

Class	Number	Percentage (%)
Nucleotides and their derivatives	63	6.27
Amino acids and their derivatives	110	10.95
alkaloid	87	8.66
Organic acid	80	7.96
Phenolic acids	137	13.63
Lignans and coumarins	21	2.09
lipid	172	17.11
terpenoids	23	2.29
quinones	5	0.50
flavonoids	184	18.31
tannins	6	0.60
Others	117	11.64
Total	1005	100

3

4

Table 2(on next page)

Metabolites significantly changed in different treatments

Table 2. Metabolites significantly changed in different treatments.

Class		LHR88 vs LHHG	MSD502 vs LHHG	MSD502 vs LHR88	MSD502 vs MSR88	MSR88 vs LHR88
Nucleotides and their derivatives	down	13	11	4	5	2
	up	8	2	0	1	3
Amino acids and their derivatives	down	30	24	2	11	2
	up	3	5	11	6	14
alkaloid	down	23	13	4	4	1
	up	10	7	18	8	12
Organic acid	down	23	18	6	3	7
	up	8	3	5	6	2
Phenolic acids	down	16	19	21	2	28
	up	19	13	7	11	1
Lignans and coumarins	down	4	4	2	0	6
	up	2	2	0	1	1
lipid	down	50	5	14	1	7
	up	20	6	49	32	10
terpenoids	down	8	5	1	1	1
	up	0	0	5	2	0
quinones	down	0	1	0	0	2
	up	2	2	0	2	0
flavonoids	down	34	53	25	2	77
	up	21	16	2	33	4
tannins	down	0	0	0	0	0
	up	3	3	0	0	0
Others	down	27	21	0	2	6
	up	3	1	12	5	6
Total	down	228	174	79	31	139
	up	99	60	109	107	53

Figure 1

Heatmap (A) and PCA plot (B) illustrating the distribution of metabolites across different treatments in rice examples

In A, each sample is represented by a column, while each metabolite is represented by a row. The colors green and red are used to indicate low and high abundance, respectively. In B, PC1 and 2 demonstrate strong cohesion within groups and effective separation among the rice accessions.

Fig. 1. Heatmap (A) and PCA plot (B) illustrating the distribution of metabolites across different treatments in rice examples. In A, each sample is represented by a column, while each metabolite is represented by a row. The colors green and red are used to indicate low and high abundance, respectively. In B, PC1 and 2 demonstrate strong cohesion within groups and effective separation among the rice accessions.

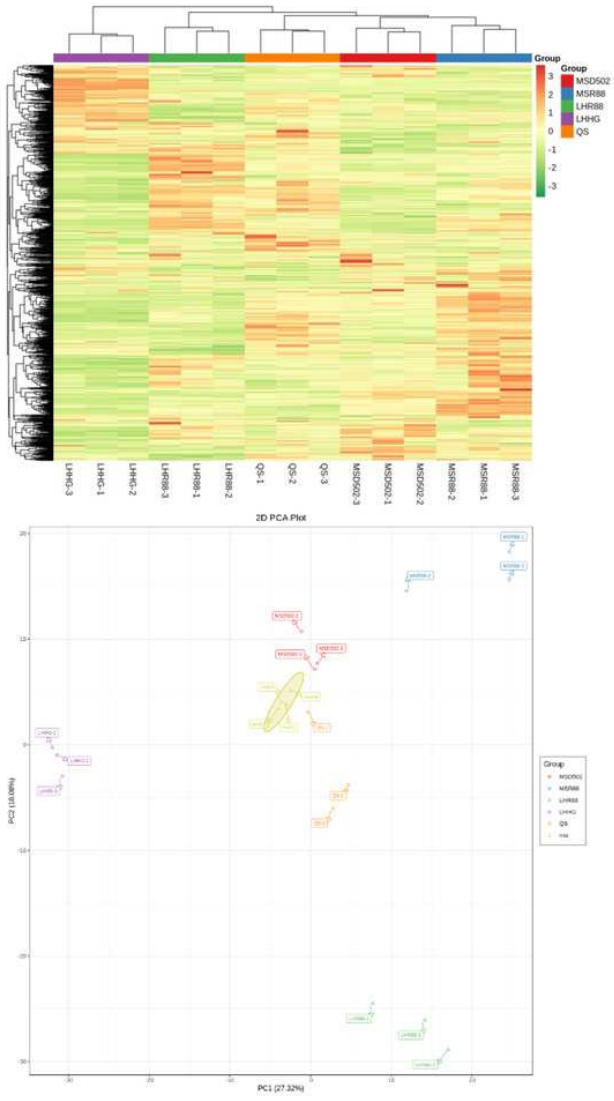


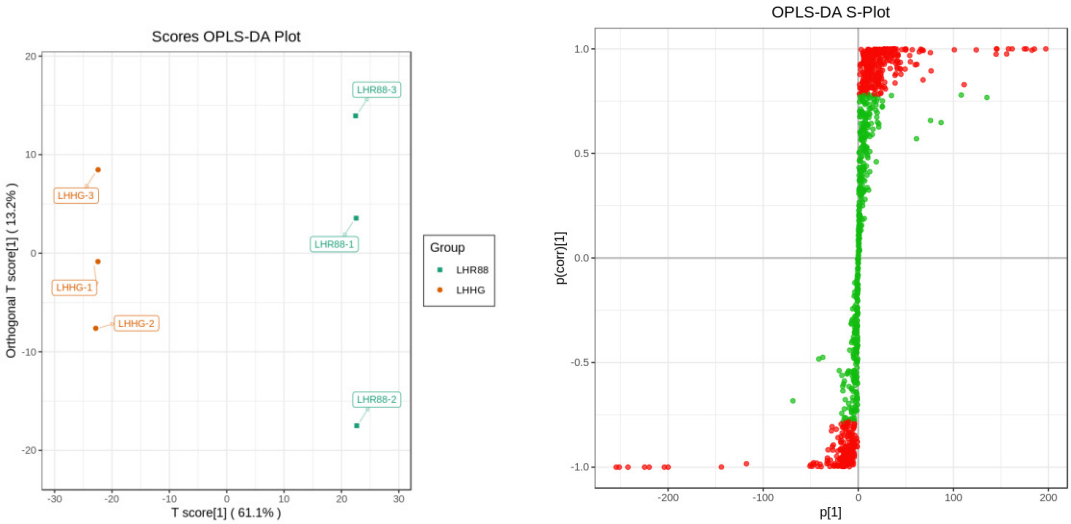
Figure 2(on next page)

The OPLS-DA score chart displays the comparison of 5 groups of rice, namely LHR88 vs LHHG, MSD502 vs LHHG, MSD502 vs LHR88, MSD502 vs MSR88, and MSR88 vs LHR88

A and E represent the OPLS-DA model diagrams for these comparisons.

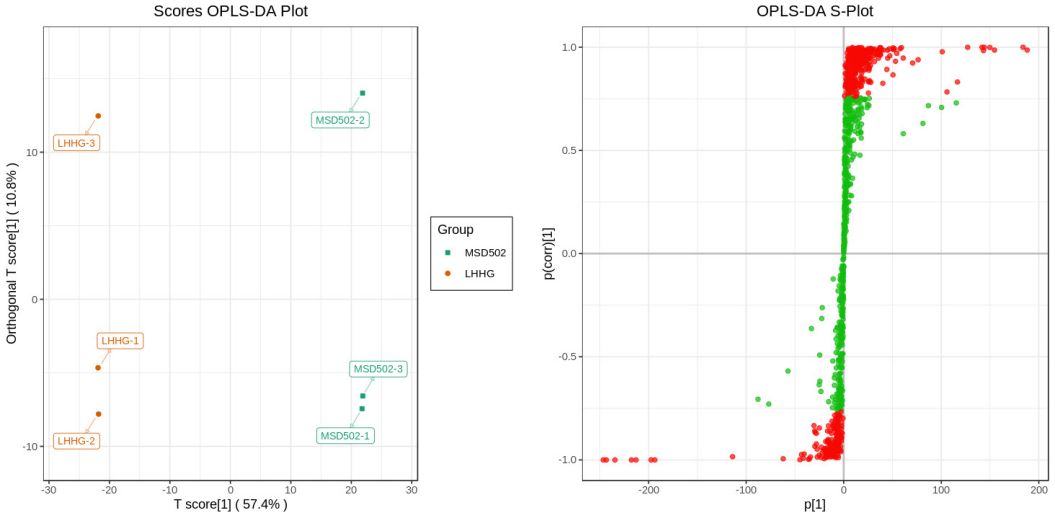
Fig. 2. The OPLS-DA score chart displays the comparison of 5 groups of rice, namely LHR88 vs LHHG, MSD502 vs LHHG, MSD502 vs LHR88, MSD502 vs MSR88, and MSR88 vs LHR88. A and E represent the OPLS-DA model diagrams for these comparisons.

A R 2X = 0.743, R2Y = 1, Q2 = 0.975



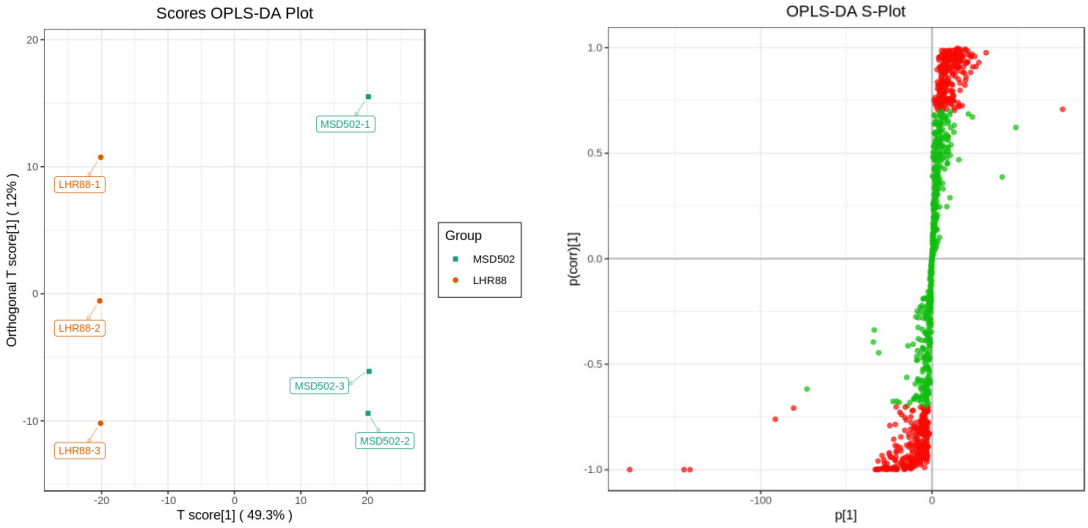
LHR88 vs LHHG

B R2X = 0.682, R2Y = 1, Q2 = 0.967



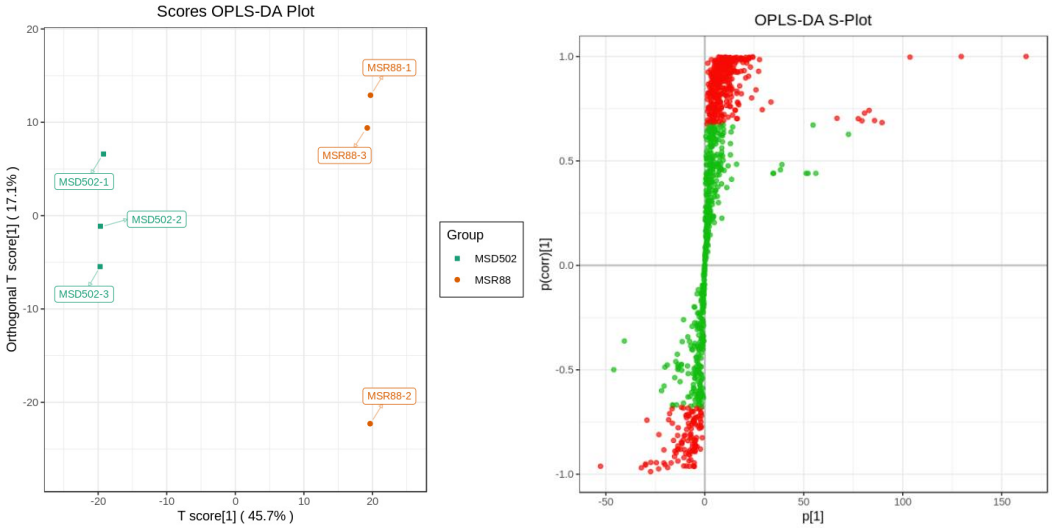
MSD502 vs LHHG

C R2X = 0.613, R2Y = 1, Q2 = 0.944



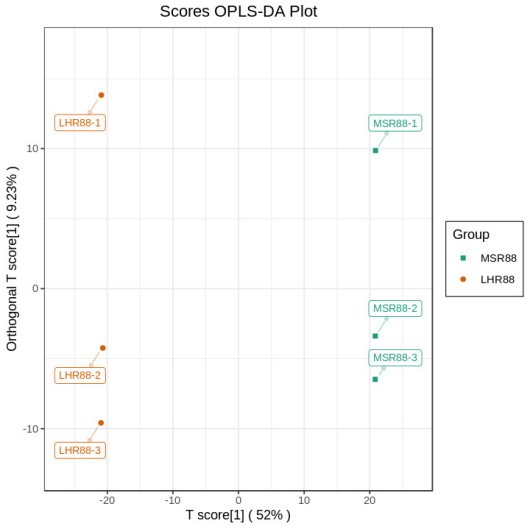
MSD502 vs LHR88

D R2X = 0.628, R2Y = 1, Q2 = 0.918



MSD502 vs MSR88

E R2X = 0.613, R2Y = 1, Q2 = 0.954


Scores OPLS-DA Plot
LHR88-1
MSR88-1
MSR88-2
MSR88-3
LHR88-2
LHR88-3
Orthogonal T score[1] (9.23%)
T score[1] (52%)
Group
MSR88
LHR88

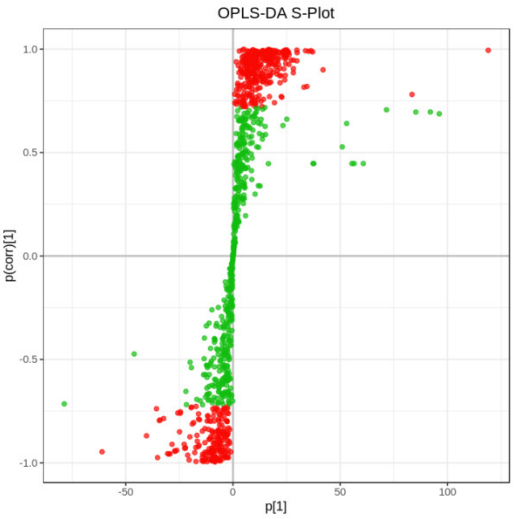

OPLS-DA S-Plot
p(corr)[1]
p[1]

Figure 3(on next page)

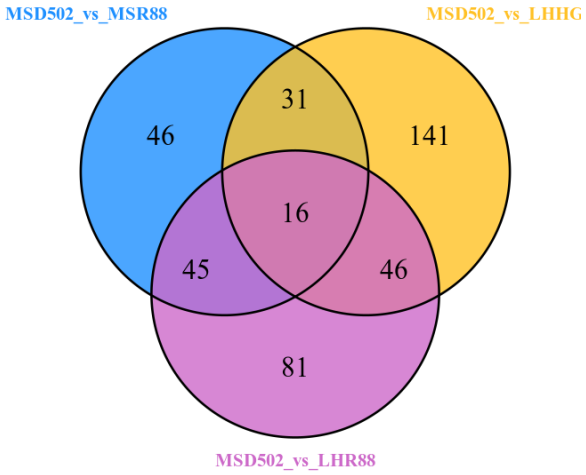
Visual representations including Venn diagrams and bubble diagrams were utilized to illustrate comparisons among three distinct sets of differentially expressed metabolites (MSD502_vs_MSR88, MSD502_vs_LHHG, MSD502_vs_LHR88)

In Figure A), a Venn diagram was employed to showcase both shared and unique metabolites across these comparison groups. Additionally, Figures B-D present KEGG enrichment analyses for differentially expressed metabolites within each group (MSD502 vs. MSR88/LHHG/LHR88). Each individual bubble within these figures corresponds to a specific metabolic pathway; its position on an axis as well as its size collectively indicate its level of influence within that particular pathway. Larger bubbles signify greater influence factors. Furthermore, variations in bubble color reflect p values obtained from enrichment analysis with darker colors indicating higher degrees of enrichment.

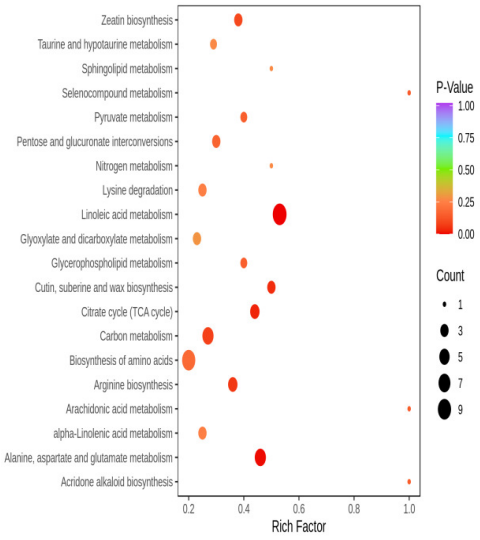
Fig. 3. Visual representations including Venn diagrams and bubble diagrams were utilized to illustrate comparisons among three distinct sets of differentially expressed metabolites (MSD502_vs_MSR88, MSD502_vs_LHHG, MSD502_vs_LHR88). In Figure A), a Venn diagram was employed to showcase both shared and unique metabolites across these comparison groups. Additionally, Figures B-D present KEGG enrichment analyses for differentially expressed metabolites within each group (MSD502 vs. MSR88/LHHG/LHR88). Each individual bubble within these figures corresponds to a specific metabolic pathway; its position on an axis as well as its size collectively indicate its level of influence within that particular pathway. Larger bubbles signify greater influence factors. Furthermore, variations in bubble color reflect p values obtained from enrichment analysis with darker colors indicating higher degrees of enrichment.

A

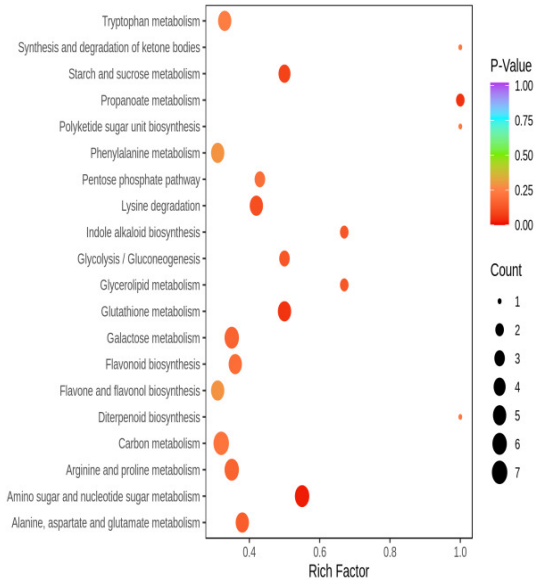
MSD502_vs_MSR88 KEGG Enrichment



B



C MSD502_vs_LHHG KEGG Enrichment Enrichment



D MSD502_vs_LHR88 KEGG

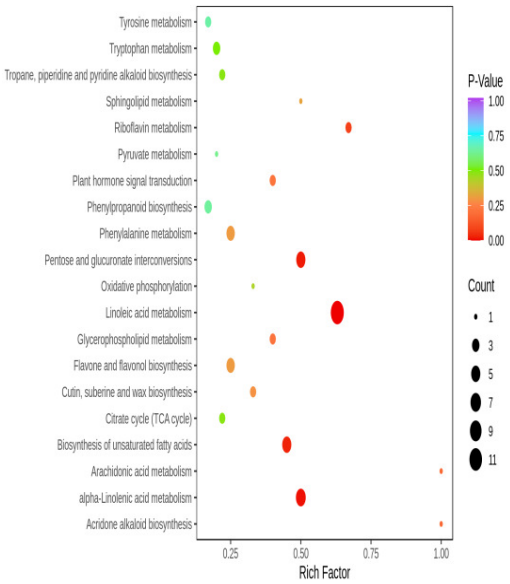
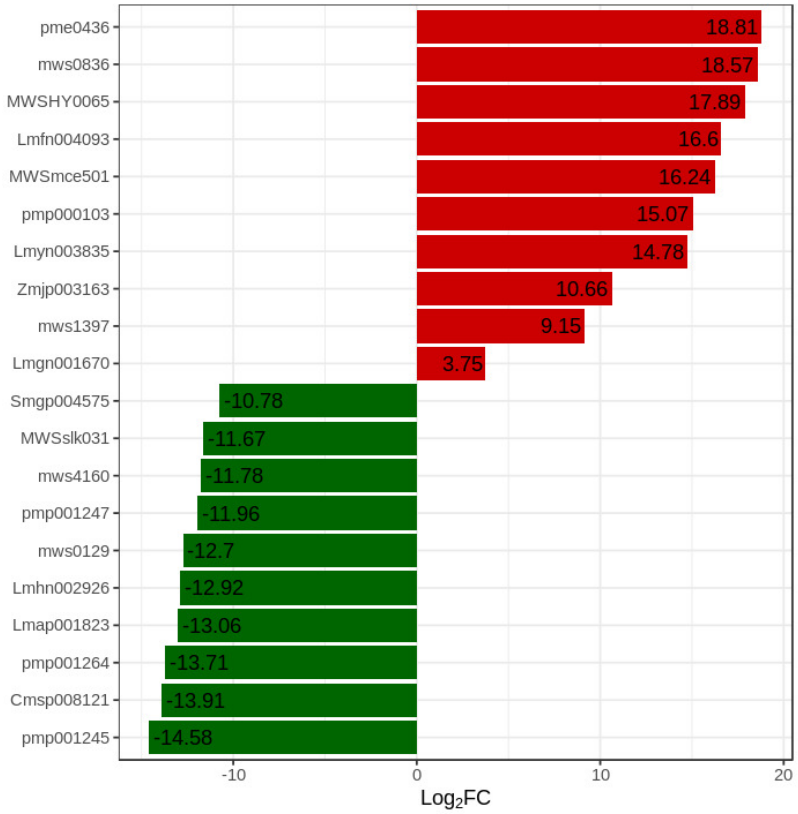


Figure 4(on next page)

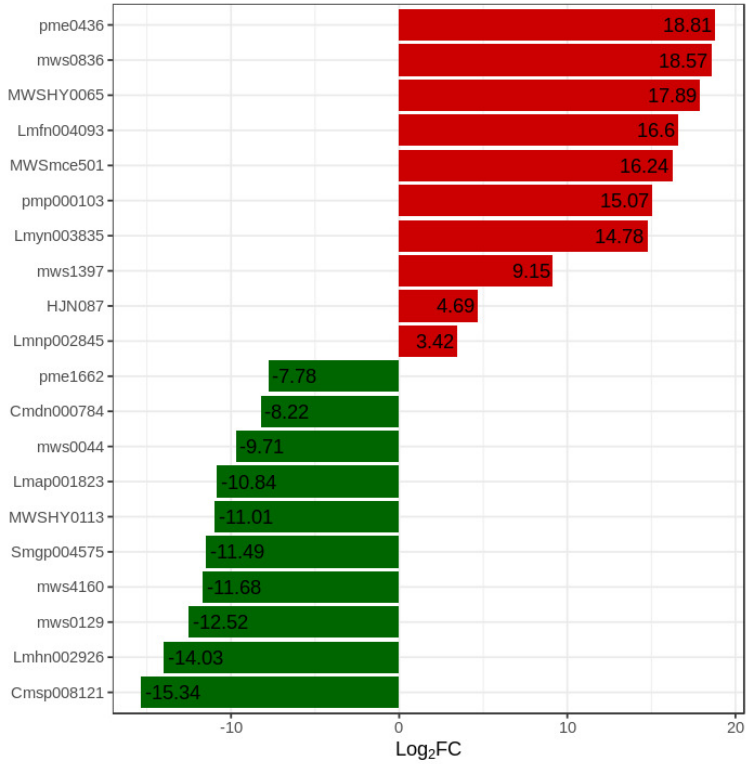
The highest fold change was observed in the top 10 metabolites that were up-regulated and down-regulated in each comparison group

Group A compared LHR88 to LHHG, Group B compared MSD502 to LHHG, Group C compared MSD502 to LHR88, Group D compared MSD502 to MSR88, and Group E compared MSR88 to LHR88. The up-regulated metabolites are represented by red bar charts while the down-regulated metabolites are represented by green bar charts.

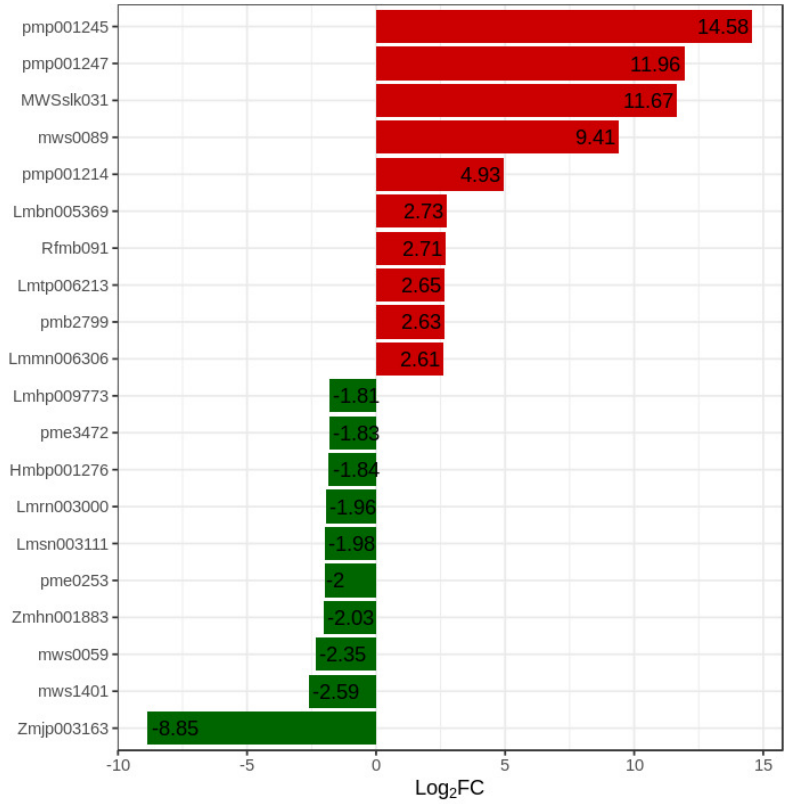
Fig. 4. The highest fold change was observed in the top 10 metabolites that were up-regulated and down-regulated in each comparison group. Group A compared LHR88 to LHHG, Group B compared MSD502 to LHHG, Group C compared MSD502 to LHR88, Group D compared MSD502 to MSR88, and Group E compared MSR88 to LHR88. The up-regulated metabolites are represented by red bar charts while the down-regulated metabolites are represented by green bar charts.



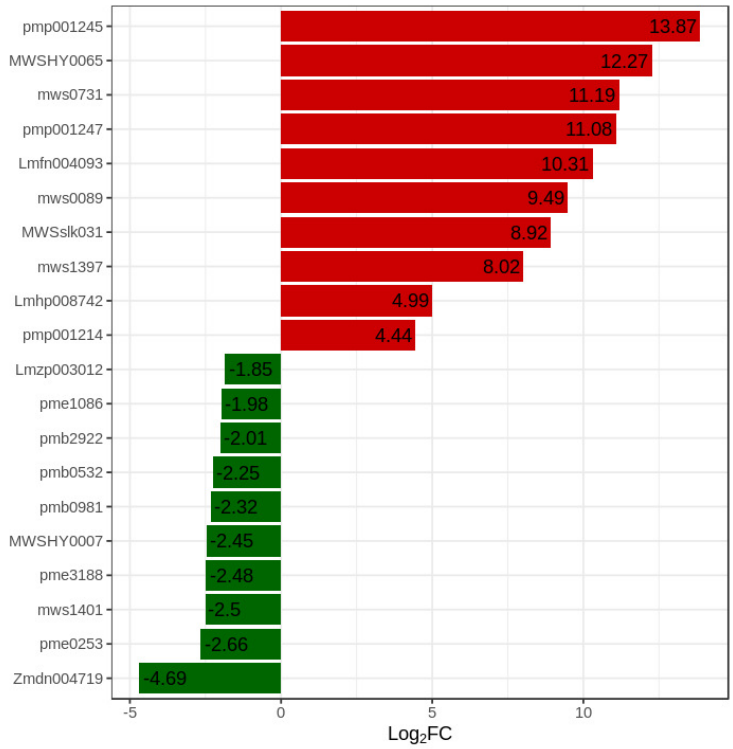
（A）：LHR88 vs LHHG



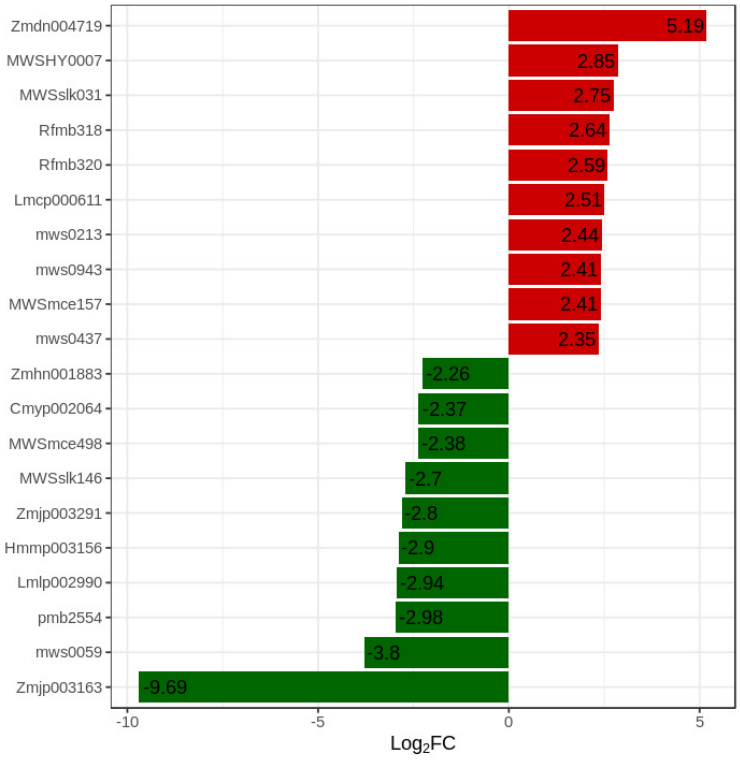
（B）: MSD502 vs LHHG



（C）: MSD502 vs LHR88



（D）: MSD502 vs MSR88



（E）: MSR88 vs LHR88