

Full Length Article

Metabolomic Analysis of Rhizosphere Soil Fertility in Maize (Zea mays) at Milking Stage

Yunpeng Luan^{1,2,3†}, Yunqi Luan¹, Lifang He⁶, Li Zheng³, Dechang Mao⁴, Xiao-Guang Yue⁵, Hongmin Wang⁴, Peng Guo¹, Nadezda Vladimirovna Verkhovtseva^{1*} and Vladimir Mikhailovich Goncharov^{1*}

¹Department of Soil Lomonosov, Moscow State University, Moscow, Russia

²Key Laboratory for Forest Resources Conservation and Use in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming 650224, P. R. China

³Southwest China Eco-development Academy, Southwest Forestry University, Kunming 650224, Yunnan, P. R. China

⁴School of Life Science, Southwest Forestry University, Kunming 650224, Yunnan, P. R. China

⁵European University Cyprus, Nicosia 1516, Cyprus

⁶Yunnan Yunhe Pharmaceutical Co., Ltd., Yunnan, P. R. China

^{*}For correspondence: Team_PI_Luan@163.com; kmzypaper@163.com

[†]Contributed equally to this work and are co-first authors

Received 31 August 2020; Accepted 09 November 2020; Published 25 January 2021

Abstract

In this study rhizosphere soil of corn at a milk stage was collected to investigate characteristic metabolites and their potential functions. Total nitrogen, organic matters, ammonium nitrogen, pH values, available phosphorus and potassium were determined by semimicro-Kjeldahl method, potassium dichromate (external heating) method, indophenol blue colorimetric method, potentiometry, NaHCO₃ leaching-molybdenum-antimony colorimetric method and NH₄OAc leaching-flame spectrometry, respectively. In addition, UPLC-Q/TOF-MS was adopted for non-targeted metabolomic analysis. As revealed by results, the total nitrogen contents in soils collected from Dongchuang (*i.e.*, DCMRS for short) 0.67 \pm 0.14 mg/kg lower than from Fumin (*i.e.*, FMMRS for short); moreover, both DCMRS and FMMRS were acid soils. DCMRS contains higher levels of AN (Ammonium nitrogen), SOC (Soil organic carbon), and AP (Available phosphorus) than FMMRS. The amount of TN (Total nitrogen) contained in FMMRS soil was 2.410 \pm 0.422 mg/kg, which is higher than DCMRS. All data derived from UPLC-Q/TOF-MS met the corresponding requirements for further analysis. Metabolites such as 2-methyl-1-propylamine, gamma-butyrolactone and 3-methyl-1-butylamine were detected in DCMRS and FMMRS samples. Several pathways were included, such as lipid metabolism, xenobiotics biodegradation and metabolism, terpenoids and polyketides, and amino acid metabolism. Through comparison of FMMRS and DCMRS, metabolic pathways associated with nitrogen, carbon, and antibiotic metabolism including iron transport were significantly different between them. Taken together, FMMRS is more fertile, less acidic, and higher in nitrogen than DCMRS. © 2021 Friends Science Publishers

Keywords: Corn; Red soil; Rhizosphere soil; Untargeted metabolomics

Introduction

Corn (*Zea mays* L.) is a crop extensively planted all over the world, serving as raw materials of fodder, grains and energy (Byrt *et al.* 2011). Compared with rice, wheat and other food crops, corn is better in drought, cold, barren tolerance and excellent environmental adaptability. Corn is an excellent food crop with a high nutritional value. It is an important source of feed in animal husbandry, aquaculture, etc., and also one of the indispensable raw materials for food, medical and health, light industry, chemical industry, etc. (Byrt *et al.* 2011; Ma 2019). It plays an important role in ensuring national food security. 500 g of corn kernels

contain 365 g of carbohydrates, which is slightly lower than rice; 21.5 g fat, more than any other cereal crops; 42.5 g protein, second to millet only; higher vitamin B than other crops (Chen *et al.* 2004; Ou 2016). Corn also contains more cellulose. As an important industrial raw material, more than 500 kinds of industrial products have been made directly or indirectly from corn kernels and by-products (Ge *et al.* 2017; Miao 2018; Wang 2018). The main product of modern corn industry is corn starch, which is widely used in food, medicine, textile and other industrial fields (Liu 2014; Lu *et al.* 2018; Tong 2019). Glucose, liquor, beer, acetone and so on can be made from corn seeds (Wang 2016; Wang and Wang 2016). Corn stalks can be used to make fiberboard,

To cite this paper: Luan Y, Y Luan, L He, L Zheng, D Mao, XG Yue, H Wang, P Guo, NV Verkhovtseva, VM Goncharov (2021). Metabolomic Analysis of rhizosphere soil fertility in maize (*Zea mays*) at milking stage. *Intl J Agric Biol* 25:555–566

paper, rayon, electrical insulation and chemical rubber plate. Besides, corn has a wide range of applications in medicine. For example, corn starch is an important raw material to produce penicillin, streptomycin, aureomycin and other antibiotics. Corn has high yield and wide use, and has high economic value as food, feed, industrial and pharmaceutical raw materials. In China, cultivated area of corns ranks only second to wheat and paddy rice. Located in Southwest China, Yunnan is one of the provinces where corns are introduced at the earliest (Huang 2012). According to statistics, the cultivated area and yield of corn in Yunnan rank first and second, respectively. For instance, the cultivated area of corn reached 1.7177 million hectares in 2016; and the yield of the same year was up to 9.196 million tons (Li et al. 2018). As corns are extremely rich, cheap, easy to obtain, strongly environment adaptable, and also have many biological activities, such as anti-oxidation, antitumor, immunomodulation and bacteriostasis, etc., it has a broad development and application prospect (Yang et al. 2019).

As a major medium in which crops are planted, soil contains not only nutrients required by growth of crops, but also substances (e.g., plant hormones, amino acid and antibiotics) secreted by soil microbials and secretions (e.g., polysaccharose and amino acids) (Kuang et al. 2003). All these substances have certain influences on metabolism of soil microbials and plants. Soil resources are precious for human survival, and soil fertility is its essential attribute (Wang and Fu 2007; Pang 2009; Huang et al. 2017). Soil fertility is an indicator of the ability of the soil to provide various nutrients required for crop growth. It is a comprehensive performance of various basic soil properties and differentiates the soil from soil parent materials and other natural bodies. The material basis of soil proves a natural resource and means of agricultural production (Xiong 2001; Huang and Sun 2006; Marschner and Rengel 2007; Lazcano 2011). Soil fertility is the basic property and essential characteristic indicating the ability of soil to supply and coordinate nutrients, water, air and heat for plant growth, and the comprehensive response of soil physical, chemical and biological properties (Liu 2010; Buckland and Grime 2010; Ling-An et al. 2011; Zhao and Bai 2013). The evaluation of soil fertility contains qualitative and quantitative methods. Qualitative descriptions of soil fertility quality are relatively simple, such as intuitive description of how the soil looks, feels and smells. While the quantitative method refers to the calculation of the "score" of soil quality according to the quantitative soil properties, and the best soil usually gets the highest score (Song 2011). Soil fertility index includes soil chemical, physical, biological and environmental condition index, and all the factors are expressed in numerical value. In this way, a series of scores will be involved in soil fertility evaluation. It is difficult to find the internal relationship between each index from these data, which is difficult to achieve by processing. Therefore, the comprehensive manual

evaluation of soil fertility must be carried out from the perspective of multiple factors by means of mathematical analysis (Luo et al. 2002; Yang et al. 2016). Different soil types correspond to different vegetational forms and microbial compositions. High throughput testing and data processing as its approaches, mass spectrometry (MS) and nuclear magnetic resonance (NMR) as its analytical platforms, and information modeling and system integration as its objectives. At present, metabolomics has become one of the powerful research tools (Ram et al. 2005; Baker 2011; Geyer 2013; Suthar et al. 2013). Ultra-high Performance Liquid Chromatography-Quadrupole-Time-of-Flight/Mass Spectrometry (UPLC/Q-TOF-MS) is a chromatographic separation system that uses Ultra-high Performance Liquid Chromatography (UPLC) as a chromatographic separation system. TOF-MS is a mass spectrometry technology formed by series analyzers (Zhang et al. 2017). UPLC/Q-TOF-MS is one of the most effective methods for multi-component analysis and identification of complex matrices in recent years (Lacina et al. 2010). Compared with traditional highperformance liquid chromatography (High Performance Liquid Chromatography, HPLC), the speed, sensitivity and resolution of UPLC are 9 times, 3 times and 1.7 times higher respectively (Wren 2005; Lucie et al. 2006). Q-TOF-MS tandem technology is an important breakthrough in mass spectrometry technology. It has the advantages of high sensitivity, high selectivity, multi-stage mass spectrometry, and high information acquisition speed. Therefore, the UPLC/Q-TOF-MS combination technology can effectively solve the problems of complex composition and quantitative difficulties in soil composition analysis, and can analyze the metabolites in the soil quickly, accurately, comprehensively, and reliably. For the past few years, metabolomics techniques have been widely applied in the research of soil. It is also increasingly combined with other system biology technology, such as proteomics, transcriptomics and genomics (Chen et al. 2013; Mandal and Singh 2013). MS is deemed as an analytical method dependent on mass-tocharge ratio determination for ions. Thanks to its sensitivity and specificity, MS gradually surmounts NMR and becomes the most powerful tool for qualitative and quantitative analysis on metabolites in metabolomics-related investigations (Dunn and Ellis 2005; Griffiths and Wang 2009) Since red soils are extensively distributed in Yunnan Province (Zhou 1983) corns in Yunnan are mostly planted in red soils. Moreover, red soils with high viscidity contain metallic oxides such as iron and aluminum. Being highly acidic, red soils also contain a few amounts of organic matters (Huang 2012; Zhao et al. 2019).

In this paper high performance liquid chromatography/ tandem high-resolution mass spectrometry was performed to test metabolome of rhizosphere soils of corn at milk stage in Dongchuan and Fumin regions. In combination with bioinformation analysis, MS data were interpreted for presenting metabolites in such soils and providing references for improvement of soils where corns are planted in both regions.

Materials and Methods

Sample collection

Dongchuang is located at the northeastern part of Yunnan, Fumin, as a county, is situated in the northeast of central Yunnan. Being 200 km apart, both regions have a subtropical monsoon plateau climate. Rhizosphere soils of corns at a milking stage were sampled from red soils in Dongchuang District and Fumin County of Yunnan Province. While those collected in Dongchuan were labeled as DCMRS, those collected in Fumin were labeled as FMMRS. Corn plants that grow well were collected, of which the roots about 5 cm below the surface layer were cut off, with soils attached. After which, they were put in a preservation box at 4°C and brought back to the laboratory. Totally, 6 soil samples were gathered in Dongchuang and Fumin, respectively. On an ultra-clean bench, excess soils were shaken off (Riley and Barber 1970; Hui et al. 2016) and the soils still attached on the roots were collected. Subsequently, impurities (including plant residues and fibrous root systems) in the soils were removed. Then, each soil sample was further equally divided into two portions, viz. one was stored at 4°C temporarily and used for general analysis on soil compositions, the other portion was stored at -80°C and used for non-targeted metabolomic analysis.

Analysis soils physico-chemical properties

Soil pre-treatment was carried out according to requirements in determination of total nitrogen, organic matters, ammonium nitrogen, pH, available phosphorus and available potassium. To be specific, the method proposed by (Li et al. 2018) in Tropical Forestry was adopted to implement soil determination. However, a semi-micro Kieldahl method was selected for total nitrogen determination. The organic matter content in soils was measured by the potassium dichromate (external heating) method. The contents of available phosphorus and potassium were determined by the NaHCO₃ (0.5 mol/L) leaching-molybdenum-antimony colorimetric method and NH₄OAc (1.0 mol/L) leaching-flame spectrometry. By following Li (2011), the potentiometric method was utilized to measure the soil pH, in which water-soil ratio was set to 2.5:1. At last, ammonium nitrogen in soils was determined by the indophenol blue colorimetric method. It should be noted that each test was repeated for three times.

Metabolite leaching

The collected samples were thawed on ice, and metabolites were extracted with 50% methanol buffer. Briefly, 20 μ L of sample was extracted with 120 μ L of precooled 50% methanol, vortexed for 1 min, and incubated at room temperature for 10 min. Then, the extraction mixture was stored overnight at -20°C, followed by centrifugation at 4,000 g for 20 min. After which, the supernatants were

transferred into new 96-well plates, stored at -80°C prior to the LC-MS analysis. In addition, pooled QC samples were also prepared by mixing $10 \,\mu$ L of each extraction mixture.

LC/MS analysis

All samples were acquired by the LC-MS system according to machine orders. Firstly, all chromatographic separations were performed using an ultra-performance liquid chromatography (UPLC) system (SCIEX, UK). An ACQUITY UPLC T3 column (100 mm*2.1 mm, 1.8 μ m, Waters, U.K.) was used for the reversed phase separation. The column oven was maintained at 35°C. The flow rate was 0.4 mL/min. The mobile phase consisted of solvent A (water, 0.1% formic acid) and solvent B (Acetonitrile, 0.1% formic acid). Gradient elution conditions were set as follows: 0~0.5 min, 5% B; 0.5~7 min, 5% to 100% B; 7~8 min, 100% B; 8~8.1 min, 100 to 5% B; 8.1~10 min, 5% B. The injection volume for each sample was 4 μ L.

A high-resolution tandem mass spectrometer TripleTOF5600plus (SCIEX, UK) was used to detect metabolites eluted form the column. The Q-TOF was operated in both positive and negative ion modes. The curtain gas was set to 30 psi, Ion source gas1 was set to 60 PSI, Ion source gas 2 was set to 60 PSI, and an interface heater temperature was to 650°C. For positive ion mode, the Ionspray voltage floating was set to 5000 V. For negative ion mode, the Ionspray voltage floating was set to -4500V. The mass spectrometry data were acquired in IDA mode. The TOF mass range was from 60 to 1200 Da. The survey scans were acquired in 150 ms and a total of 12 product ion scans were collected if exceeding a threshold of 100 counts per second (counts/s) with a 1+ charge-state. Total cycle time was fixed to 0.56 s. Four time bins were summed for each scan at a pulser frequency value of 11 kHz through monitoring of the 40 GHz multichannel TDC detector with four-anode/channel detection. Dynamic exclusion was operated for 4 s. During the acquisition, the mass accuracy was calibrated once every 20 samples. Furthermore, in order to evaluate the stability of the LC-MS during the whole acquisition process, a quality control sample (Pool of all samples) was acquired for every 10 samples.

Information analysis

The pretreatments of acquired MS data including peak picking, grouping, retention time correction, second peak grouping, and annotation of isotopes and adducts were performed using XCMS software (Want *et al.* 2006). LC–MS raw data files were converted into mzXML format and then processed by the XCMS, CAMERA and metaX (Wen *et al.* 2017) toolbox of R software. Each ion was identified by combining retention time (RT) and m/z data. Intensity of each peak was recorded and a three-dimensional matrix consisting of arbitrarily assigned peak indices (retention time-m/z pairs), sample names (observations) and

ion intensity information (variables) was generated.

The online KEGG database was used to annotate the metabolites by matching the exact molecular mass data (m/z) of samples with those from database. If a mass difference between observed value and the database value was less than 10 ppm, the metabolite would be annotated and the molecular formula of metabolites would further be identified and validated by the isotopic distribution measurements. An in-house fragment spectrum library of metabolites was also used to validate the metabolite identification.

The intensity of peak data was further preprocessed by metaX. Those features that were detected in less than 50% of QC samples or 80% of biological samples were removed; the remaining peaks with missing values were imputed with the k-nearest neighbor algorithm to further improve the data quality. PCA was performed for outlier detection and batch effects evaluation based on the pre-processed dataset. Quality control-based robust LOESS signal correction was fitted to the QC data according to the order of injection to minimize the drift of signal intensity over time. In addition, the relative standard deviations of the metabolic features among all QC samples were calculated, and those > 30%were then removed. Before quantitative analysis on metabolites, quality control over their intensity was carried out to obtain quantitative information of high-quality ion peaks. Based on CV and hierarchical clustering of QC samples, data quality control can be achieved.

Statistical analysis

SPSS 20.0 (SPSS, Inc., Chicago, IL, USA) was used for data statistics and analysis, while GraphPad 8.0.2 was utilized to plot the relevant histograms. SIMCA_P 13.0 was used to implement principal component analysis (PCA) and orthogonal to partial least squares discriminant analysis (OPLS-DA). The R software was adopted to plot heat maps of metabolites and corresponding calculations were carried out by means of hierarchical clustering (HCL). As for distance computing, the Euclidean distance was adopted. It is worth noting that ward. D2 was used as the clustering method.

Student t-tests were conducted to detect difference in metabolite concentration between 2 phenotypes. The *P* value was adjusted for multiple tests using an FDR (Benjamini–Hochberg). Supervised PLS-DA was conducted using metaX to discriminate the variables between groups. The VIP value was calculated. A VIP cut-off value of 1.0 was used as a criterion to select important features.

Results

Soil physico-chemical properties

Nitrogen, phosphorus and potassium are three major soil nutrients essential to plant growth. These three nutrients are not only applied to the most during agricultural production, but also consumed the most by plant growth. In this study, the contents of total nitrogen, ammonium nitrogen, soil organic carbon, available phosphorus and available potassium in DCMRS and FMMRS were measured (Table 1; Fig. 2A–D). **Soil metabolic substances:** The extracted metabolic substances were detected in both positive and negative ion modes using a high-resolution mass spectrometer. The total ion count and the number of substances annotated by primary and secondary mass spectrometry data can be found in positive and negative ion modes (Table 2).

Detection control of metabolic substances: The only mass-to-charge ratio and chromatographic retention time were revealed in mass spectra of each metabolite. With the goal of controlling data quality, quality control over substances obtained by XCMS was provided, including total ion chromotogram, m/z-rt distribution of metabolites, m/z difference ranges generated by peak alignment of each substance, and rt difference ranges thus produced. As shown in the total ion chromotogram, the detected metabolites can be well isolated in the chromatographic condition set for this study. A great number of features were found when the retention time reached 1-1.5 min, 2-3.5 min or 8 min approximately. Among them, the number of features generated at a retention time of 2-3.5 min reached its peak. In terms of m/z, the maximum number of features was produced when it ranged between 200 and 400. Triple TOF 5600, a high-resolution mass spectrometer (resolution: 30, 000 and above), was utilized to detect the metabolites. Subsequently, alignment could be implemented with the help of XCMS. It turned out that the aligned m/z had a range of -0.015-0.015 and the aligned retention time all varied between -0.5 min and 0.5 min (Fig. 2A-D).

Identification and quantification of metabolic substances

Identification of metabolic substances: According to the detected substances, the primary m/z was matched in the KEGG database to acquire primary identification results. During primary identification, one m/z correspond to multiple metabolites. In order to produce more accurate identification results, secondary mass spectrogram of inhouse metabolites was compared with their secondary mass spectrometry data (Table 2).

When several metabolites share identical chemical composition, a substance may still have different isomerides due to different types of element sorting. Considering that these isomerides have the same molecular weight, it is less likely to distinguish them based on mass spectra. For this reason, one substance may correspond to multiple possible metabolites during primary identification of substances. Hence, one-to-many statistics were made for metabolite identification, (Fig. 3A–B).

After metabolite matching in the KEGG database, the metabolites were categorized (Fig. 3C–D). In line with global and overview maps of KEGG pathway level 2, the number of metabolites associated with lipid metabolism, xenobiotics biodegradation and metabolism, metabolism of

Table 1: Soil pH, and contents of total nitrogen, organic matters, ammonium nitrogen, available phosphorus and available potassium in soils (Mean \pm SD)

	TN (Total nitroge	n) AN	(Ammonium	SOC (Soil	organic carbon)	AP	(Available A	K (Available	potassium) pH
	(mg/kg)	nitrogen) (m	g/kg)	(mg/kg)		phosphorus) (m	g/kg) (1	mg/kg)	
DCMRS	0.67 ± 0.14	39.22 ± 3.32		18.20 ± 0.89)	51.31 ± 5.76	4	18.83 ± 17.92	5.79±0.43
FMMRS	2.10 ± 0.42	19.493 ± 1.4	7	19.94 ± 2.27	,	23.93 ± 3.51	4	78.97 ± 12.93	6.77±0.76
Notes: DCYRS stands for rhizosphere soil for corns in a seedling stage in Dongchuan District, DCMRS for rhizosphere soil for corns at a milk stage in Dongchuan District, and									
FMMRS for rhizosphere soil for corns at a milk stage in Fumin County									

Table 2: The total ion count and identification statistics of metabolites

Mode	All feature	All annotated	MS2	KEGG	
POS	4189	2477	85	1899	
NEG	2739	1230	34	957	

Notes: Mode: An ion mode in which the substance is detected under a mass spectrometer; POS: Positive ion mode; NEG: Negative ion mode; All feature: The number of substances extracted by the XCMS software; All annotated: The number of substances obtained by primary and secondary mass spectrometry data; and, MS2; The number of secondary ions identified, that is the number of substances that can be matched with both primary m/s and secondary fragment ions m/s of a substance in the database



Fig. 1: A: Total nitrogen; B: Available phosphorus; C: Organic matters; D: Ammonium nitrogen; E: Available potassium; and, F: Soil pH. DCMRS refers to rhizosphere soil for corns at a milk stage in Dongchuan District, and FMMRS refers to rhizosphere soil for corns at a milk stage in Fumin County

terpenoids and polyketides, and amino acid metabolism was above 160. More particularly, global and overview maps corresponded to probably the most metabolites. Regarding lipid metabolism, xenobiotics biodegradation and metabolism, metabolism of terpenoids and polyketides, and amino acid metabolism, the number of potential metabolites may range from 165 to 200. Among the first 20 KEGG pathways, there were metabolic pathways, biosynthesis of secondary metabolites, microbial metabolism in diverse environments, biosynthesis of antibiotics, degradation of aromatic compounds, arachidonic acid metabolism, and sesquiterpenoid and triterpenoid biosynthesis, etc.

According to peak areas of DCMRS and FMMRS, the top first 20 metabolites are respectively listed in Table 3 and 4, including their MZ, RT, peak areas, names and formulas. It could be found that such metabolites had similar rankings in both DCMRS and FMMRS. For example, 2-methyl-1propylamine had the largest peak area, followed by gammabutyrolactone and 3-methyl-1-butylamine successively. Except proline and valine, peaks areas of other metabolites in DCMRS were larger than those in FMMRS.

Ouantitative analysis on metabolic substances: Quantitative information of metabolites was derived from primary chromatographic peak areas of the substance. Moreover, no quantitative analysis was made on differences of the ions with CV > 30% for a reason they may vary substantially during the experiment. Quantitative statistical information of metabolites is listed in Table 5. The number of features extracted by XCMS was 4,189 in a positive ion mode or 3,622 in a negative ion mode. After the number of ions of a substance with a missing value < 80% in the sample or < 50% in the OC sample was acquired, the number of high quality features can be figured out, which was 3,622 in the positive ion mode or 2,389 in the negative ion mode.

As shown in Fig. 4, quality control analysis was carried out prior to quantitative analysis on metabolites. In this figure, A stands for CV distribution of the sample. After normalization, such data were proven to be consistent with post-processing requirements. In addition, C represents distribution intensity evaluation of differential metabolites A Boxlot diagram was plotted after log 2 normalization of relevant intensity values, where log 2 (i.e., intensity) ranged from 10 to 12. B indicates that hierarchical clustering analysis on metabolite intensity of FMMRS and DCMRS was implemented by means of ward D2. At last, principal component analysis (PCA) was made on FMMRS, DCMRS and QC samples. More particularly, FMMRS, DCMRS and QC samples can be preferably differentiated in the PCA diagram. This indicates there was certain difference in differential metabolites between FMMRS and DCMRS.

As shown in Fig. 5, mean intensity of metabolites undergoing secondary mass spectrum identification (A) is given. Here, bar length of the second ring represents diverse mean intensity of the metabolites in the mass spectrum. Moreover, mean intensity of which should be multiplied

Table 3: First 20 substances with maximum peak areas in DCMRS

Number	MZ	PT	Average neak area	Name	Formula
1	220 1050022	2.04505			Toffildia C.H. N
1	228.1959023	3.04505	136183.8769	2-Methyl-1-propylamine	$C_4H_{11}N$
2	256.2631784	5.566516667	131652.8086	Gamma-Butyrolactone	$C_4H_6O_2$
3	284.2949678	6.489666667	125826.1855	3-Methyl-1-butylamine	$C_5H_{13}N$
4	132.1009235	1.766241667	112990.2737	2,5-Dihydro-2,4-dimethyloxazole	C ₅ H ₉ NO
5	229.1405197	2.7169	87674.70663	(2R,3S)-2-methyl-3-propyloxirane	$C_6H_{12}O$
6	135.0796726	5.277983333	76720.33604	Triethylamine	$C_6H_{15}N$
7	228.1956866	6.213366667	57819.67183	Triethylamine	$C_6H_{15}N$
8	279.1587938	4.5003	57116.42804	Choline	C ₅ H ₁₄ NO
9	172.1328279	2.528366667	50148.55729	1-Piperidinecarboxaldehyde	C ₆ H ₁₁ NO
10	87.04354563	3.111133333	46848.60368	Proline	C ₅ H ₉ NO ₂
11	310.3105369	6.595866667	44256.53949	Valine	$C_5H_{11}NO_2$
12	250.1773982	6.213366667	43742.21084	Indane	C_9H_{10}
13	136.0204317	6.441	36122.42364	Nicotinic acid	$C_6H_5NO_2$
14	118.0852842	0.889216667	32412.89428	Malonic acid	$C_3H_4O_4$
15	226.1799047	6.220416667	32183.90691	Malonic acid	$C_3H_4O_4$
16	413.266458	7.251366667	31642.44589	Hypotaurine	$C_2H_7NO_2S$
17	353.2299268	2.996966667	28868.39866	Dihydrothymine	$C_5H_8N_2O_2$
18	337.2349649	5.83	28080.3092	Pipecolate	$C_6H_{11}NO_2$
19	540.5357826	6.34005	27069.31943	Isoleucine	$C_6H_{13}NO_2$
20	149.0224496	4.50395	24963.68949	Leucine	$C_6H_{13}NO_2$



Fig. 2: Metabolite detection control. **A**: Total ion chromotogram (TIC) of metabolites, where the overall mass spectrum signal strength of the sample is under control. In this map, x-axis represents time and the y-axis stands for sums of all ion intensity at each time point in the mass spectrogram; **B**: A mz-rt distribution diagram, where x-axis refers to the substance retention time, y-axis refers to m/z of the substance, each point in the diagram represents a substance; **C**: The m/z alignment ranges of metabolites; **D**: The <u>rt</u> alignment ranges of metabolites

with a log10 function. A bar stands for a type of metabolites color represents the type of metabolites. Fig. 5B shows the correlation of metabolites subjected to secondary mass spectrum identification. In this study, their correlation is embodied in different colors and various shades. For example, the darker the red is, the stronger their correlation will be; the darker the blue is, the weaker the correlation will be. In addition, two correlated areas are marked by a green triangle in dashed lines. Fig. 5C shows the intensity of metabolites for secondary mass spectrum identification,

where x-axis refers to DCMRS, FMMRS and QC samples, while y-axis represents the type of metabolites. Each square in the figure represents a metabolite. Likewise, darker red corresponds to a stronger correlation, but darker blue represents a weaker correlation.

Differential metabolite screening & differential metabolic pathway analysis

Differential metabolite screening: In this paper, q-value

Number	MZ	RT	FMMRS Average peak area	Metabolite	Formula
1	228.1959023	3.04505	134457.8221	2-Methyl-1-propylamine	$C_4H_{11}N$
2	256.2631784	5.566516667	127645.9067	Gamma-Butyrolactone	$C_4H_6O_2$
3	284.2949678	6.489666667	123988.4382	3-Methyl-1-butylamine	$C_5H_{13}N$
4	132.1009235	1.766241667	86018.61235	2,5-Dihydro-2,4-dimethyloxazole	C ₅ H ₉ NO
5	229.1405197	2.7169	55237.28711	(2R,3S)-2-methyl-3-propyloxirane	$C_6H_{12}O$
6	135.0796726	5.277983333	54484.58411	Triethylamine	C ₆ H ₁₅ N
7	228.1956866	6.213366667	53744.10721	Triethylamine	C ₆ H ₁₅ N
8	279.1587938	4.5003	50242.79421	Choline	C ₅ H ₁₄ NO
9	172.1328279	2.528366667	49425.78368	1-Piperidinecarboxaldehyde	C ₆ H ₁₁ NO
10	87.04354563	3.111133333	48750.38422	Proline	C ₅ H ₉ NO ₂
11	310.3105369	6.595866667	44981.94301	Valine	$C_5H_{11}NO_2$
12	250.1773982	6.213366667	40121.6433	Indane	C_9H_{10}
13	136.0204317	6.441	32067.15209	Nicotine acid	C ₆ H ₅ NO ₂
14	118.0852842	0.889216667	30676.72433	Malonic acid	$C_3H_4O_4$
15	226.1799047	6.220416667	29604.93344	Malonic acid	$C_3H_4O_4$
16	413.266458	7.251366667	27758.16764	Hypotaurine	$C_2H_7NO_2S$
17	353.2299268	2.996966667	26494.2727	Dihydrothymine	$C_5H_8N_2O_2$
18	337.2349649	5.83	25551.76821	Pipecolate	$C_6H_{11}NO_2$
19	540.5357826	6.34005	24993.91453	Isoleucine	C ₆ H ₁₃ NO ₂
20	149.0224496	4.50395	24497.43798	leucine	$C_6H_{13}NO_2$

Table 4: Top 20 substances with maximum peak areas in FMMRS



Fig. 3: Primary m/z identification for metabolite. **A**: One-to-many statistical chart of metabolite identification, where x-axis represents the number of identified metabolites corresponding to one feature, and y-axis is the number of features; **B**: Approximate categories of metabolites identified, where x-axis stands for the metabolite category, y-axis stands for a proportion taken by each category of metabolites in all metabolites; **C**: Classification of KEGG pathways, where x-axis refers to items of KEGG pathway level 2, y-axis refers to the number of possible metabolites that involve such classification; **D**: First 20 pathways in which all possible metabolites may participate

was acquired after BH correction is conducted based on differential folds of univariate analysis and *t*-test. Moreover, the value of Variable Important for the Projection (VIP) was acquired through PLS-Discriminant Analysis (PLS-DA). The number of differential metabolites selected is shown in Fig. 6A. It can be seen that the number of up-regulated and down-regulated metabolites of FMMRS was 222 and 261, respectively (Fig. 6B).

Fig. 6C shows the volcano plot of univariate statistical test evaluation. The green and red points stand for down-

Table 5: Quantitative statistics of metabolites

Mode	All feature	High quality feature
POS	4189	3622
NEG	2739	2389

Notes: Mode: An ion mode in which the substance is detected under a mass spectrometer; POS: Positive ion mode; NEG: Negative ion mode; All features: The number of substances extracted by the XCMS software; High quality feature: The number of ions in a substance with missing value < 80% in the sample or < 50% in the QC sample



Fig. 4: Quantitative quality control over metabolites. **A:** CV distribution diagram, where CV=SD/Mean; and, the x-axis represents value of CV and the y-axis represents a proportion taken by the number of ions. Generally, it is considered that if $CV \le 30\%$, the sample shows good repeatability; **B**: Intensity of each metabolite in each sample is presented in the heat map, where, each scale is obtained by log transformation of intensity values; the Euclidean distance is introduced in calculation; and, ward.D2 serves as the clustering method; **C**: Boxplot of intensity distribution evaluation for metabolites in each sample, where, y-axis represents log2 transformation of intensity in the mass spectrum; **D**: A chart of PCA scores, where, each point represents a sample; if the points are rather close to each other, it indicates that the samples are very similar; otherwise, it reveals that certain differences can be found among these samples

regulated and up-regulated substances, respectively (Fig. 6D). As a supervised discriminant analytical statistical method, PLS-DA has the potential to reveal differences among different groups to the greatest extent. By using such method, a relation model for metabolite expression quantity and sample category was constructed by means of PLS regression (PLSR) so that modeling and prediction of sample categories can be realized. Additionally, VIP was worked out



Fig. 5: Quantitative information of metabolites in DCMRS and FMMRS. **A**: An annular chart of the mean intensity of metabolites undergoing secondary mass spectrum identification, where the color represents the type of metabolites; **B**: Correlation of metabolites subjected to secondary mass spectrum identification, where correlation between two metabolites is shown; the darker the red is, the stronger their correlation will be; but if the blue is darker, the correlation will be weaker; **C**: Metabolite intensity heat map of secondary mass spectrum identification, where represent different groups of DCMRS, FMMRS and QC samples, but color blocks on the left refer to different types of metabolites; in the middle, each color block represents a metabolite; similarly, darker red corresponds to stronger correlation, but darker blue represents weaker correlation

for the purpose of evaluating influential intensity and explanatory ability of various metabolite expression patterns for sample classification and discrimination. This may contribute to screening of metabolic markers (screening condition: VIP ≥ 1.0). R² and Q², as two parameters of the PLS-DA model, were calculated to be 0.99 and 0.94. The closer their values are to 1, the more reliable the PLS-DA model will be. On this basis, it is deemed that the relation model constructed is rather reliable.

Differential metabolism analysis: Differential metabolic pathway analysis of differential substances and KEGG was carried out. Through comparison between FMRMRS and DCMRS, 151 differential metabolic pathways were obtained, including 76 up-regulated ones and 75 down-regulated ones. It was found that some of these metabolic pathways were associated with nitrogen cycling (Fig. 7), such as carbon

metabolic pathway, amino acid metabolic pathway, antibiotics anabolism, and biosynthesis of siderophore group non-ribosomal peptides. Some differential metabolic pathways were in a positive ion mode (Table 6).

Discussion

Red soil containing metallic oxides such as iron and aluminum is widely distributed in Yunnan Province, China (Huang and Fu 2002; Zhao *et al.* 2019). Corn is an important fodder and food crop of Yunnan. In this study, soil conditions are preliminarily evaluated by measuring soil pH and the contents of total nitrogen, organic matters, ammonium nitrogen, available phosphorus and potassium in DCMRS and FMMRS for corns at a milk stage in Dongchuan District of Yunnan. Overall, FMMRS is more fertile than DCMRS,



Fig. 6: Differential metabolite screening. **A**: A statistical chart of differential metabolites from FMMRS/DCMRS; **B**: A heat map of differential metabolite expression; **C**: A volcano chart of univariate statistical test evaluation, where x-axis represents intensity values and y-axis is log10 (i.e., q-value corrected through statistical testing); **D**: A diagram of PCA scores from multivariate statistics, where x-axis and y-axis respectively refer to the first and the second principal components, that is PC1 and PC2 for short Note: In these figures, each point in this figure stands for a sample; the dispersion degrees of two colors represent distribution tendencies of samples in two groups along PC1 and PC2 axes

but less acidic (Table 1 and Fig. 1). These components in the red soil may be produced by fertilizers applied prior to the milk stage of corns. For example, chemical fertilizers can increase the contents of available phosphorus and available potassium in the soil. Moreover, once organic fertilizers (*e.g.*, decaying manure from the livestock) are applied, the contents of total nitrogen and organic matters in the soil may go up (Zhang 2011).

The source of metabolic substances in rhizosphere soil is in exudates of rhizosphere microbes and plant root systems. Matters detected by soil metabolomic analysis consist of these metabolic substances. Plant root exudates have the capability to enhance stress tolerance of plants and promote growth of soil microbes (Bashir et al. 2016). According to relevant findings, corn root exudates contain asparaginic acid, tyrosine, starch and flavonoids (Kuang et al. 2003; Carvalhais et al. 2011; Zhu et al. 2016; Luo et al. 2017). In this study, amino acid, saccharides and flavonoids were detected in root soil DCMRS and FMMRS. The influence of soil microbes on rhizosphere soil cannot be ignored either. The reason is that microbes are able to produce secondary metabolites by using plant root exudates, including antibiotics and acids. Through comparison between FMMRS and DCMRS, multiple antibiotic synthesis associated metabolic pathways have been annotated. In this course, the metabolic pathways associated with iron transport were

observed. In addition, another investigation on microbiological compositions in corn rhizosphere soil was carried out as well. Likewise, not only were bacteria associated with antibiotic synthesis detected, but also functional genes and metabolic pathways associated with iron element metabolism and transport in soil were also found. It is expected that some significant information can be provided for improving the soil where corns are planted. The pH of DCMRS was lower than of FMMRS. However, if the pH value of rhizosphere soil were always low, it would affect the accumulation of crop biomass. Before the mature stage of corn, alkaline chemical fertilizer can be properly added to improve soil pH value, which is more conducive to the accumulation of dry matter in corn. Rhizosphere soil metabolites affect many aspects of crop growth. Proper application of organic fertilizer or microbial soil inoculants can improve the content of soil metabolites by adjusting the composition of soil microorganisms, thus improving the stress resistance and yield of corn.

Conclusion

Concerning the non-targeted metabolome, metabolites in samples can be comprehensively analyzed to confirm an analytical method for differential metabolites. Under the circumstance that massive data are generated during non-

#	Pathway	Difft metabolites (241)	All metabolites (1176)	P value	Pathway ID
1	ABC transporters	23	34	1.95E-09	map02010
2	Microbial metabolism in diverse environments	79	238	1.45E-07	map01120
3	Starch and sucrose metabolism	8	8	2.83E-06	map00500
4	Cyanoamino acid metabolism	9	15	0.000871116	map00460
5	Aminoacyl-tRNA biosynthesis	7	10	0.000954765	map00970
6	Ascorbate and aldarate metabolism	7	10	0.000954765	map00053
7	Carbapenem biosynthesis	4	4	0.001728987	map00332
8	Phenylalanine, tyrosine and tryptophan biosynthesis	5	6	0.001745478	map00400
9	Carbon metabolism	7	13	0.00771802	map01200
10	Inositol phosphate metabolism	3	3	0.008521438	map00562
11	Phenylalanine metabolism	10	23	0.009946065	map00360
12	Caprolactam degradation	6	12	0.02120928	map00930
13	Degradation of aromatic compounds	22	73	0.02876643	map01220
14	Nitrogen metabolism	3	4	0.02889879	map00910
15	Toluene degradation	6	13	0.03277322	map00623
16	Tryptophan metabolism	6	13	0.03277322	map00380
17	Caffeine metabolism	4	7	0.0357948	map00232
18	Pentose and glucuronate interconversions	4	7	0.0357948	map00040
19	Bacterial chemotaxis	2	2	0.04185845	map02030
20	Terpenoid backbone biosynthesis	2	2	0.04185845	map00900

Table 6: Top 20 metabolic pathways ranked according to difference significance through comparison of FMRMRS and DCMRS



Fig. 7: Metabolic pathway Map 01120 annotated by the KEGG pathway, a pathway of microorganism in different environments; those marked by a red rectangle represent annotated nitrogen metabolic pathway

targeted metabolomic detection, the selected separation and detection technique may affect the data. As an analytical approach, UPLC-Q/TOF-MS is featured with high sensitivity and high through-put. Therefore, it is adopted in this study to detect and analyze metabolites in DCMRS and FMMRS. As an extraction solvent exerting a great influence on leaching of metabolites from the samples, methanolwater was adopted in this study to extract most watersoluble and fat-soluble substances from the samples. Although, these substances can largely reveal components of metabolites in the soil, errors caused by the detection technique or calculations may lead to data loss to a certain extent. To solve such a defect, new detection techniques and calculation methods are needed in the future research with the goal of reducing errors.

Author Contributions

Yunpeng Luan proposed the experimental idea of the article.

Yunpeng Luan, Nadezda Vladimirovna Verkhovtseva and Vladimir Mikhailovich Goncharov designed and supervised the experiment process. Yunqi Luan, Lifang He, Li Zheng, Dechang Mao, Xiao-Guang Yue, Hongmin Wang and Peng Guo participated Experiments and sorting out and analyzing the data. Yunpeng Luan mainly writes the article. All authors participate in writing and revising the article.

References

- Baker M (2011). Metabolomics: From small molecules to big ideas. Nat Meth 8:117–121
- Bashir O, K Khan, KR Hakeem, NA Amir, GH Rather, R Mohiuddin (2016). Soil microbe diversity and root exudates as important aspects of rhizosphere ecosystem. *In: Plant, Soil and Microbes*, pp:337–357. Springer, Cham, Switzerland
- Buckland SM, JP Grime (2010). The effects of trophic structure and soil fertility on the assembly of plant communities: A microcosm experiment. *Oikos* 91:336–352
- Byrt CS, CP Grof, RT Furbank (2011). C4 Plants as biofuel feedstocks: Optimizing biomass production and feedstock quality from a lignocellulosic perspective free access. J Integr Plant Biol 53:120–135

- Carvalhais LC, PG Dennis, D Fedoseyenko (2011). Root exudation of sugars, amino acids, and organic acids by corn as affected by nitrogen, phosphorus, potassium, and iron deficiency. J Plant Nutr Soil Sci 174:3–11
- Chen CX, J Wang, H Guo, WY Hou, N Yang, BR Biao, M Liu, HQ Dai, XT Liu, FH Song, LX Zhang (2013). Three antimycobacterial metabolites identified from a marine-derived *Streptomyces* spp. MS100061. *Appl Microbiol Biotechnol* 97:3885–3892
- Chen YX, DW Zhou, YF Zhang (2004). Temporal and spatial variations of chemical constituents in maize. J Appl Ecol 15:1589–1593
- Dunn WB, DI Ellis (2005). Metabolomics: Current analytical platforms and methodologies. *TrAc Trend Anal Chem* 24:285–294
- Ge WX, XW Niu, WJ Zhang (2017). The nutritional value, processing and utilization of corn cob and its application in animal husbandry production. Cer Feed Industr 11:51–54
- Geyer T (2013). Modeling metabolic processes between molecular and systems biology. *Curr Opin Struct Biol* 23:218–223
- Griffiths WJ, Y Wang (2009). Mass spectrometry: From proteomics to metabolomics and lipidomics. *Chem Soc Rev* 38:1882–1896
- Huang J, XJ Jiang, YH Zeng (2017). Research progress on evaluation methods and indexes of paddy soil fertility. *Soils Fert Sci Chin* 6:1–8
- Huang JM (2012). Research on present situation and development strategies of corn industry in Yunnan province. *Hortic Seed* 8:57–62
- Huang Y, BC Fu (2002). The change properties of the chemical compositions for Laterite. J Kunning Univ Sci Technol 27:63–66
- Huang Y, WJ Sun (2006). Soil science-change trend of organic carbon content of farmland topsoil in mainland China in the past 20 years. *Abstr Chin Acad J* 51:1
- Hui J, XQ Wu, DX Cai (2016). Soil nitrogen content and enzyme activities in rhizosphere and non-rhizosphere of summer corn under different nitrogen application rates. *J Appl Ecol* 27:1917–1924
- Kuang YW, DZ Wen, CW Zhong, GY Zhou (2003). Root Exudates and Their Roles in Phytoremediation. Acta Phytoecol Sin 27:709–717
- Lacina O, J Urbanova, J Poustka (2010). Identification/quantification of multiple pesticide residues in food plants by ultra-high-performance liquid chromatography-time-of-flight mass spectrometry. J Chromatogr A 1217:648–659
- Lazcano C (2011). The use of vermicompost in sustainable agriculture: Impact on plant growth and soil fertility. Soil Nutr 10:1–23
- Li JF, HY Wan, XY Wang (2018). Analysis of soil chemical properties of five different forest types in Wanning City. *Trop For* 46:22–25
- Li P (2011). Analyses on rice soil nutrients in Ningde Zhangwan town. J Xiamen Univ-Nat Sci 50:49–52
- Ling-An N, H Jin-Min, Z Bao-Zhong (2011). Influences of long-term fertilizer and tillage management on soil fertility of the North China plain. *Pedosphere* 21:813–820
- Liu H (2014). Corn industry: A Multi-pronged Approach to Get Out of Trouble. Agric Knowledge, China
- Liu J (2010). Study on Soil Quality Characteristics and Fertility Control Technology of Degraded Red Soil in the Red Soil Region of Central and South Human. Hum Agric Univ, China
- Lu QM, ZC Chen, AL He (2018). Application and development of corn gluten meal, a by-product of corn starch processing. J Food Saf Qual 9:15–22
- Lucie N, M Ludmila, P Solich (2006). Advantages of application of UPLC in pharmaceutical analysis. *Talanta* 68:908–918
- Luo DQ, J Bai, DT Xie (2002). Discussion on soil fertility evaluation index and method. *Ecol Environ Sci* 11:202–205
- Luo Q, SY Wang, LN Sun, H Wang (2017). Metabolic profiling of root exudates from two ecotypes of *Sedum alfredii* treated with Pb based on GC-MS. *Sci Rep* 7; Article 39878
- Ma XH (2019). Research on quality characteristics and comprehensive utilization of maize. *Cer Oils* 273:1–3
- Mandal K, B Singh (2013). Persistence of fipronil and its metabolites in sandy loam and clay loam soils under laboratory conditions. *Chemosphere* 91:1596–1603

- Marschner P, Z Rengel (2007). Contributions of the rhizosphere interactions to soil biological fertility. *In: Soil Biological Fertility*, pp:81–98. Springer, The Netherlands
- Miao CR (2018). Research on the Nutritional Value of Corn and Its Postprocessed Products. J Chin Food Saf 215:175
- Ou JL (2016). An overview of corn nutrition and key points of quality control. *Guangdong Feed* 25:38–40
- Pang YM (2009). Research Progress of Soil Fertility Evaluation. J Shanxi Agric Sci 2:87–89
- Ram RJ, NC Verberkmoes, MP Thelen, GW Tyson, BJ Baker, RC Blake, M Shah, RH Hettich, JF Ban (2005). Community proteomics of a natural microbial biofilm. *Science* 308:1915–1920
- Riley D, SA Barber (1970). Salt accumulation at the soybean (*Glycine max* (L.) Merr) root-soil interface. *Soil Sci Soc Amer J* 34:154–155
- Song SS (2011). Research on Soil Fertility Evaluation Method. Northwest A&F University, Xianyang, China
- Suthar MS, MM Brassil, G Blahnik, M Aimee, HJ Ramos, SC Proll, SE Belisle, MG Katze, MJ Gale (2013). A systems biology approach reveals that tissue tropism to West Nile virus is regulated by antiviral genes and innate immune cellular processes. *PLoS Pathog* 9; Article e1003168
- Tong Y (2019). The development course and prospect of China's corn starch and starch sugar industry technology. *Food Ferment Indust* 45:294– 298
- Wang LM (2018). Thoughts on corn production and seed industry development and breeding goals under the new situation. In: Proceedings of the 2018 Shandong Crop Society Academic Annual Meeting, Shandong, China
- Wang SD (2016). Using corn stover to produce bio-butanol and its separation and purification. DaLian Univ Technol China
- Wang Y, YJ Wang (2016). Application of Corn Pellets as Auxiliary Material in Beer Brewing. Chinese Foreign Wine Industry and Beer Technology, China
- Wang ZL, Q Fu (2007). Research progress on comprehensive evaluation of soil fertility. *Soil Crop* 23:15–18
- Want EJ, G Maille, R Abagyan (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 78:779–787
- Wen B, Z Mei, C Zeng (2017). Meta X: A flexible and comprehensive software for processing metabolomics data. BMC Bioinform 18; Article 183
- Wren SAC (2005). Peak capacity in gradient ultra-performance liquid chromatography (UPLC). J Pharmaceut Biomed 38:337–343
- Xiong SG (2001). Basic Soil Science. Chin Agric Univ Press, China
- Yang XQ, H Zhi, H Zhang (2019). Research progress on chemical constituents, pharmacological activity and utilization status of different parts of corn. *Jilin J Tradit Chin Med* 6:837–840
- Yang YQ, GF Chen, HL Guo (2016). Research on big data processing technology in soil fertility evaluation. J Chin Agric Mech 4:233–236
- Zhang Y, BM Feng, X Lu (2017). Application progress of UPLC/Q-TOF-MS combined technology in drug analysis. Nat Prod Res Dev 29:1992–1996
- Zhang ZD (2011). Research on Soil Fertilization and Changes of Components During Decomposition of Cow Manure Mixed with Green Manures. Jilin Agric University, Jilin, Chin
- Zhao H, HG Bai (2013). Erratum to: The critical soil P levels for crop yield, soil fertility and environmental safety in different soil types. *Plant Soil* 327:39–39
- Zhao KL, BR Wang, MG Xu (2019). Changes in pH with depths of soils derived from different parent materials and analysis of acidification in Southern China. J Plant Nutr Fert 25:1308–1315
- Zhou LF (1983). Characteristics and zonality of soil distribution in Yunnan Province. *J Mount Res* 4:33–40
- Zhu SS, JM Vivanco, DK Manter (2016). Nitrogen fertilizer rate affects root exudation, the rhizosphere microbiome and nitrogen-use-efficiency of corn. Appl Soil Ecol 107:324–333