Innovative Metabolomics Insights for Better Health

Carotenoids Metabolomics Assay Final Report

Metware Biotechnology Inc.

www.metwarebio.com

Contents

MWXS-23-xxx Carotenoids Targeted Metabolomics Assay Final Report

1 Abstract

Carotenoids are a class of yellow, orange-red, or reddish polyenes, generally consisting of eight isoprenoid units. Carotenoids have antioxidant, anti-cancer, and night blindness prevention properties. In plants, carotenoids act as auxiliary pigments for photosynthesis in chloroplasts, protect chlorophyll from bright light damage, and are also precursors for the synthesis of ABA. The biosynthetic pathway of carotenoids has been well characterized - their synthesis is based on isopentenyl pyrophosphate as the precursor, which is catalyzed by enzymes including IPI, GGPS, PSY, PDS, ZDS, LycB, LycE, and others to produce various types of carotenoids.

There are more than 700 varieties of carotenoids discovered so far, which can be categorized into two groups based on their chemical structures: carotenes (containing only two elements, carbon, and hydrogen, without oxygen) and luteins (containing hydroxyl, ketone, carboxyl, methoxyl, and other oxygen-containing functional groups, such as lutein and astaxanthin). In plants, carotenes exist in their free state. In contrast, lutein analogs combine with different fatty acids to form carotenoid esters in plants and thus exist in free and esterified forms. METWARE has established two carotenoid assays tailored to customer needs: carotenoid assay and saponified carotenoid assay. In contrast to the conventional carotenoid assay, carotenoid esters are hydrolyzed in the saponified carotenoid assay to detect carotenoids resulting from the hydrolysis reaction.

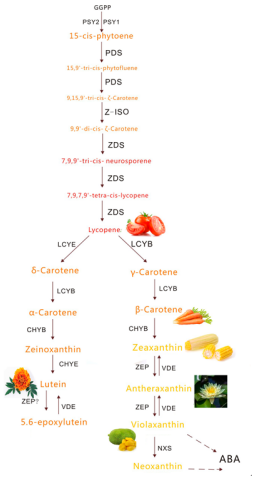

Fig 1: Diagram of the carotenoid synthesis pathway

2 The experimental process

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) can detect and quantify compounds with high polarity and poor thermal stability, and accurately quantify them. The overall process is as follows:

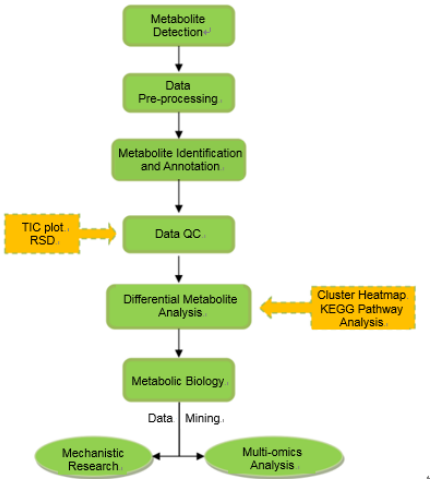

Fig 2: Flow chart of metabolomics analysis

Compounds to be detected:

Table 1: List of compounds in the panel

Number	Compounds	Index
1	α-carotene	Carotenoid_01
2	lycopene	Carotenoid_02
3	γ-carotene	Carotenoid_03
4	β-carotene	Carotenoid_04
5	phytofluene	Carotenoid_05
6	(E/Z)-phytoene	Carotenoid_06
7	ε-carotene	Carotenoid_07
8	antheraxanthin dipalmitate	Carotenoid_08
9	lutein caprate	Carotenoid_09
10	lutein laurate	Carotenoid_10
11	lutein myristate	Carotenoid_11
12	lutein palmitate	Carotenoid_12
13	lutein stearate	Carotenoid_13
14	5,6epoxy-luttein dilaurate	Carotenoid_14
15	lutein dilaurate	Carotenoid_15
16	5,6epoxy-lutein-caprate-palmitate	Carotenoid_16
17	lutein dimyristate	Carotenoid_17
18	lutein dipalmitate	Carotenoid_18
19	lutein distearate	Carotenoid_19
20	lutein dioleate	Carotenoid_20
21	lutein oleate	Carotenoid_21
22	neochrome palmitate	Carotenoid_22
23	rubixanthin caprate	Carotenoid_23
24	rubixanthin laurate	Carotenoid_24
25	rubixanthin myristate	Carotenoid_25
26	rubixanthin palmitate	Carotenoid_26
27	violaxanthin dibutyrate	Carotenoid_27
28	violaxanthin laurate	Carotenoid_28

Table 1: List of compounds in the panel

Number	Compounds	Index
29	violaxanthin myristate	Carotenoid_29
30	violaxanthin palmitate	Carotenoid_30
31	violaxanthin palmitoleate	Carotenoid_31
32	violaxanthin dilaurate	Carotenoid_32
33	violaxanthin-myristate-caprate	Carotenoid_33
34	violaxanthin-myristate-laurate	Carotenoid_34
35	violaxanthin dimyristate	Carotenoid_35
36	violaxanthin-myristate-palmitate	Carotenoid_36
37	violaxanthin dipalmitate	Carotenoid_37
38	violaxanthin-myristate-oleate	Carotenoid_38
39	violaxanthin dioleate	Carotenoid_39
40	zeaxanthin myristoleate	Carotenoid_40
41	zeaxanthin palmitate	Carotenoid_41
42	zeaxanthin-caprate-laurate	Carotenoid_42
43	zeaxanthin dilaurate	Carotenoid_43
44	zeaxanthin-laurate-myristate	Carotenoid_44
45	zeaxanthin dimyristate	Carotenoid_45
46	zeaxanthin-laurate-palmitate	Carotenoid_46
47	zeaxanthin-myristate-palmitate	Carotenoid_47
48	zeaxanthin-palmitate-stearate	Carotenoid_49
49	zeaxanthin-oleate-palmitate	Carotenoid_50
50	β-cryptoxanthin laurate	Carotenoid_51
51	β-cryptoxanthin myristate	Carotenoid_52
52	β-cryptoxanthin palmitate	Carotenoid_53
53	β-cryptoxanthin oleate	Carotenoid_54
54	antheraxanthin	Carotenoid_55
55	zeaxanthin	Carotenoid_56
56	violaxanthin	Carotenoid_57
57	neoxanthin	Carotenoid_58
58	lutein	Carotenoid_59
59	β-cryptoxanthin	Carotenoid_60
60	astaxanthin	Carotenoid_61
61	8'-apo-beta-carotenal	Carotenoid_62
62	capsanthin	Carotenoid_63
63	α-cryptoxanthin	Carotenoid_64
64	capsorubin	Carotenoid_65
65	canthaxanthin	Carotenoid_66
66	echinenone	Carotenoid_67
67	β-citraurin	Carotenoid_68

Original file path: Final report/data/metware_component.xlsx

2.1 Sample information

This project has 24 samples divided into 4 groups. Sample information is shown in the following table:

Table 2: Sample information table

Species	Tissues	MW_ID	Sample_ID
_	_	A1	A1
_	_	A2	A2
_	_	A3	A3
_	_	A4	A4
_	_	A5	A5
_	_	A6	A6
_	_	B1	B1
_	_	B2	B2
_	_	B3	B3
_	_	B4	B4
_	_	B5	B5
_	_	B6	B6
_	_	C1	C1
_	_	C2	C2
_	_	C3	C3
_	_	C4	C4
_	_	C5	C5
_	_	C6	C6
_	_	D1	D1
_	_	D2	D2
_	_	D3	D3
_	_	D4	D4
_	_	D5	D5
_	_	D6	D6

Original file path: Final report/0.data/sample_info.xlsx

2.2 Reagents and instruments

Table 3: Instrument information

Device	Model	Brand
LC-MS/MS	Tripe Quad 6500+	SCIEX
Centrifuge	5424R	Eppendorf
Electronic balance	AS 60/220.R2	RADWAG
Ball mill	MM400	Retsch
Centrifugal concentrator	CentriVap	LABCONCO
Multi-tube vortex oscillator	Vortex Mixer，MI0101002	Four E's
Ultrasonic cleaner	Ultrasonic Cleaner，CD-F15	Olenyer

Table 4: Information of standards and reagents

Reagent	Grade	Brand
Ethanol	Chromatographic pure	Merck
Methanol	Chromatographic pure	Merck
Acetonitrile	Chromatographic pure	Merck
N-hexane	Chromatographic pure	CNW
2,6-Di-tert-butyl-4-methylphenol	Chromatographic pure	aladdin
Sodium Chloride	Analytically pure	RHAWN
Potassium Hydroxide	Analytically pure	Hushi Reagent
Acetone	Chromatographic pure	Sinopharm Chemical Reagent
Formic Acid	Chromatographic pure	Sigma-Aldrich
Methyl Tertiary Butyl Ether	Chromatographic pure	CNW

2.3 Sample Pre-processing

2.3.1 Solid Sample

(1) TCryopreserved samples were collected (if there was no special requirement, plant samples defaulted to freeze-dried samples, and animal samples defaulted to fresh samples) and milled to powder form with a ball mill (30 Hz, 1 min);

(2) A 50 mg of the pulverized sample was weighed and extracted with 0.5 mL of hexane/acetone/ethanol mixture (1:1:1, v/v/v) containing 0.01% BHT (g/mL);

(3) The extract was vortexed for 20 min at room temperature and then centrifuged at 12000 r/min for 5 min at 4°C. The supernatant was removed, and the extraction was repeated once before centrifugation to combine the supernatants;

(4) The obtained extract was concentrated, re-dissolved with 100 μL dichloromethane, passed through a 0.22 μm filter membrane, and stored in a brown injection vial for LC-MS/MS analysis.

Note: Refer to references 9 for pre-processing methods.

2.4 Chromatography - Mass Spectrometry Acquisition Conditions

The data acquisition system mainly consisted of Ultra Performance Liquid Chromatography (UPLC) (ExionLC™ AD, https://sciex.com.cn/) and Tandem Mass Spectrometry (MS/MS) (QTRAP® 6500+, https://sciex.com.cn/).

The primary liquid-phase conditions include:

1) Chromatographic column: YMC C30 (3 μm, 100 mm×2.0 mm i.d.);
2) Mobile phase: Phase A, methanol/acetonitrile (1:3, v/v) with 0.01% BHT and 0.1% formic acid; Phase B, methyl tert-butyl ether with 0.01% BHT;
3) Gradient elution program: 0 min A/B at 100:0 (v/v), 3 min at 100:0 (v/v), 5 min at 30:70 (v/v), 9 min at 5:95 (v/v), 10 min at 100:0 (v/v), 11 min at 100:0 (v/v);
4) Flow rate 0.8 mL/min; column temperature 28°C; injection volume 2 μL.

The primary mass spectrometry conditions include:

Atmospheric pressure chemical ionization source (APCI) temperature 350°C, curtain gas (CUR) 25 psi. In the Q-Trap 6500+ LC-MS/MS system, each ion pair was scanned for detection based on optimized declustering potential (DP) and collision energy (CE).

Note: Refer to references 5, 7 and 8 for chromatography and mass spectrometry methods.

2.5 Qualitative and quantitative principles of metabolites

Metabolites were quantified by multiple reaction monitoring (MRM) using triple quadrupole mass spectrometry. In MRM mode, the first quadrupole screened the precursor ions for the target substance and excluded ions of other molecular weights. After ionization induced by the impact chamber, the precursor ions were fragmented, and a characteristic fragment ion was selected through the third quadrupole to exclude the interference of non-target ions. After obtaining the metabolite spectrum data from different samples, the peak area was calculated on the mass spectrum peaks of all substances and analyzed by standard curves.

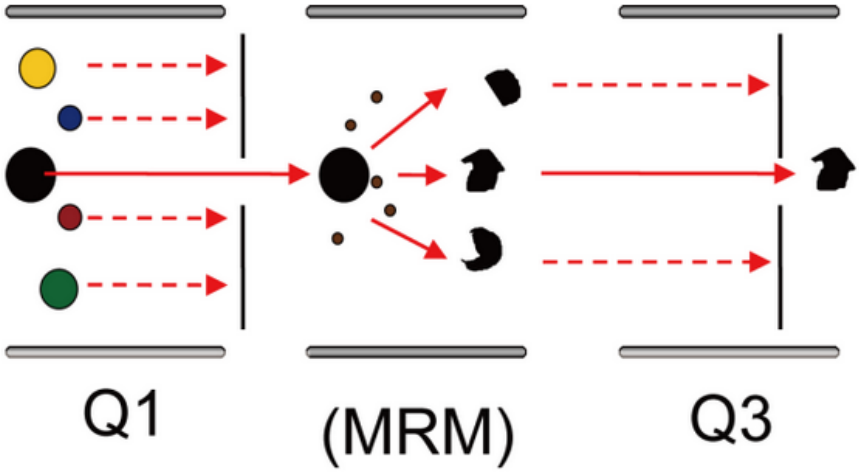

Fig 3:
Schematic diagram of multiple reaction monitoring mode by mass spectrometry

3 Data evaluation

3.1 Data pre-processing

Analyst 1.6.3 was used to process mass spectrum data. The following figure shows the total ions current (TIC) and MRM metabolite detection multi-peak diagram (XIC) of the mixed QC samples. The X-axis shows the Retention time (RT) from metabolite detection, and the Y-axis shows the ion flow intensity from ion detection (intensity unit: CPS, count per second).

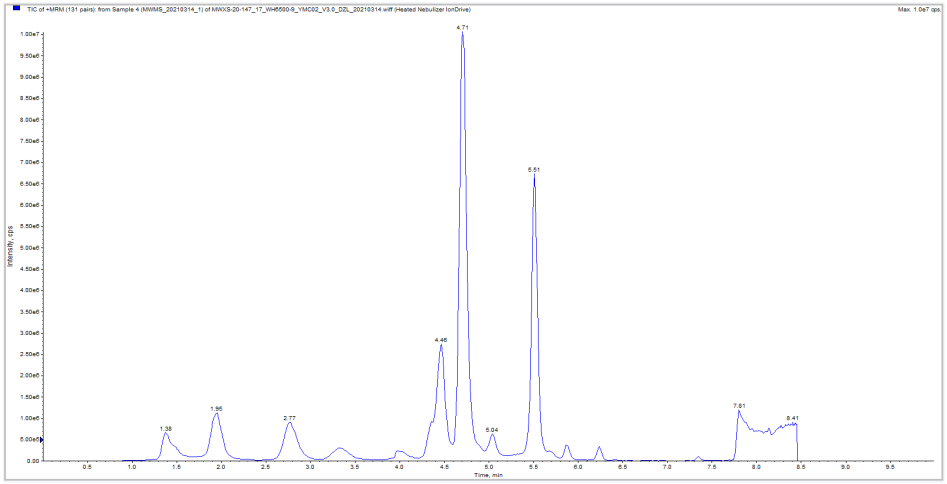

Fig 4: Total ion current diagram of mixed phase mass spectrum analysis

Original file path: Final report/0.data/QC/*QC_MS_TIC.png

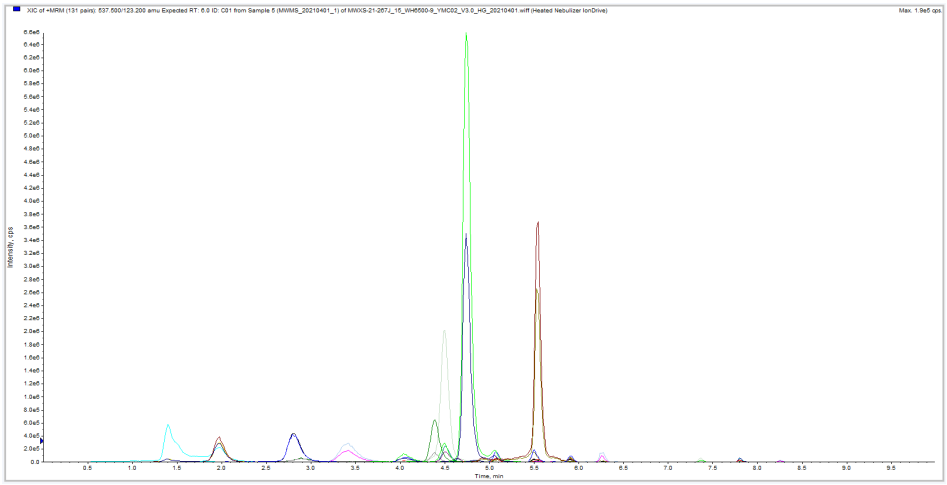

Fig 5: Extraction ion flow chromatogram

Original file path: Final report/0.data/QC/*MRM_detection_of_multimodal_maps*

The mass spectrometry data was analyzed using MultiQuant 3.0.3 software. The mass spectrum peaks detected in different samples were scored and corrected based on retention time and peak shape of the standard. The figure below shows the correction results of quantitative analysis of a substance randomly selected from different samples.

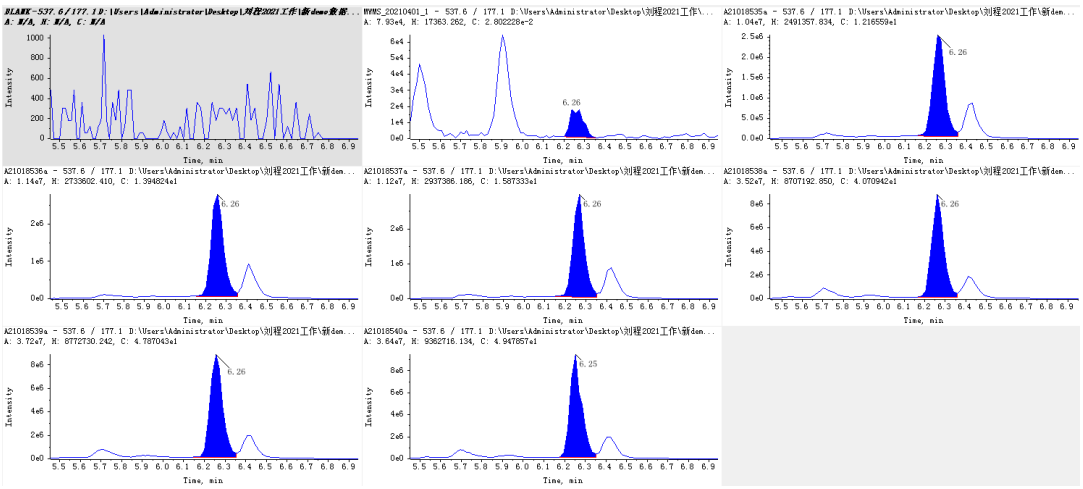

Fig 6: Scoring correction diagram for quantitative analysis of metabolites

Note: The figure shows the quantitative analysis integral correction results of randomly selected metabolites in different samples. The x-axis is the retention time (min) of metabolite detection, the y-axis is the ion flow intensity (CPS) of a certain metabolite ion detection, and the peak area represents the relative content of the substance in the sample.

Original file path: Final report/0.data/QC/*Integral_correction.png

3.2 Standard Curve

The carotenoid reference standards were prepared at concentrations of 0.001 μg/mL, 0.005 μg/mL, 0.01 μg/mL, 0.05 μg/mL, 0.1 μg/mL, 0.5 μg/mL, 1 μg/mL, 5 μg/mL, 10 μg/mL, 50 μg/mL, 100 μg/mL, 250 μg/mL, 400 μg/mL. The mass spectral peak intensity data for each concentration of these reference standards were obtained. The standard curves of different substances were plotted by taking the concentrations (Concentration) of the standards as the horizontal coordinates and the peak areas (Area) of the standards as the vertical coordinates. The linear equations and correlation coefficients of the standard curves of the substances detected in this project are shown in the table below.

Table 5: Equation of calibration curve

Index	Compounds	Class	RT
Carotenoid_05	phytofluene	carotenes	1.89
Carotenoid_06	(E/Z)-phytoene	carotenes	4.93
Carotenoid_07	ε-carotene	carotenes	5.51
Carotenoid_01	α-carotene	carotenes	5.92
Carotenoid_04	β-carotene	carotenes	6.28
Carotenoid_03	γ-carotene	carotenes	7.39
Carotenoid_02	lycopene	carotenes	8.33
Carotenoid_57	violaxanthin	xanthophylls	1.56
Carotenoid_58	neoxanthin	xanthophylls	1.91
Carotenoid_68	β-citraurin	xanthophylls	2.76

Final report/0.data/equation.xlsx

3.3 Quantification Results

Compound quantification was obtained by substitute the area under the peak in the detected samples into the equation of calibration curve for calculation.

3.3.1 Solid Sample

Carotenoid content in the sample (μg/g) = c*V/1000/m

Meaning of each letter in the formula:

c: the concentration value (μg/mL) obtained by substituting the integrated peak area obtained from this sample into the standard curve;

V: volume of solution used in redissolution (μL);

m: mass of the sample weighed (g).

The metabolite ID, concentration and corresponding metabolite names of some metabolites detected in this experiment are shown in the following table:

Table 6: Statistical Table of metabolite quantity

Index	A1	A2
Carotenoid_01	10.7679	12.0129
Carotenoid_02	28.1869	31.0509
Carotenoid_03	65.0646	38.4195
Carotenoid_04	85.812	60.9604
Carotenoid_05	111.002	91.8609
Carotenoid_06	44.4245	29.7026
Carotenoid_07	96.9489	80.6602
Carotenoid_50	21.9222	32.1499
Carotenoid_49	15.262	18.2121
Carotenoid_47	15.8869	14.3142

Original file path: Final report/0.data/*level.xlsx

3.4 Sample Quality Control Analysis

3.4.1 Total Ion Chromatogram Analysis

Using the mixed solution as the QC sample, one QC sample was inserted every 10 detection samples for analysis during the detection by the system. The stability of the device during the detection of the project can be assessed by analyzing the overlapped total ion flow chromatograms (TICs) obtained from the mass spectrometry detection and analysis of the same QC samples. The high stability of the testing device is a vital safeguard for the reproducibility and reliability of the data.

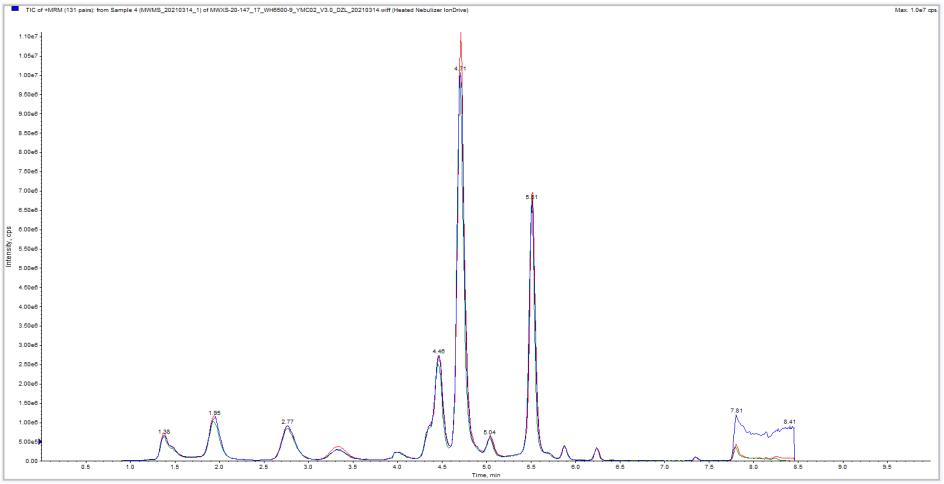

Fig 7: TIC overlap diagram detected by QC sample essence spectrum
Note: Superimposed spectrum from different QC samples. The results showed that the spectrum of total ion flow were highly consistent indicating that the signal stability was good when the same sample was detected at different times by mass spectrometry.

Original file path: Final report/0.data/picture/*QC_MS_tic_overlap*

3.4.2 CV value distribution of all samples

The Coefficient of Variation (CV) value is the ratio between the standard deviation of the original data and the mean of the original data, which can reflect the degree of data dispersion. The Empirical Cumulative Distribution Function (ECDF) can be used to analyze the frequency of CV of substances that is smaller than the reference value. The higher the proportion of substances with low CV value in QC samples is, the more stable the experimental data is. The proportion of substances with CV value less than 0.3 in QC samples was higher than 80% , indicating that the experimental data were relatively stable. The proportion of substances with CV value less than 0.2 in QC samples was higher than 80%, indicating that the experimental data were very stable.

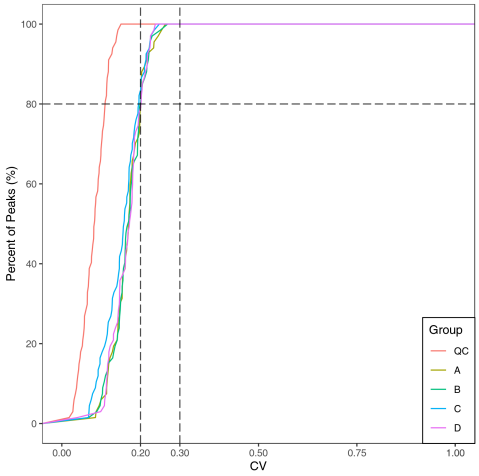

Fig 8: CV distribution of each group

Note: The X-axis represents the CV value, the Y-axis represents the proportion of metabolites with CV value less than a corresponding reference value. Different colors represent different sample groups. QC indicates quality control samples. The two dash lines on X-axis correspond to 0.2 and 0.3; the two dash lines on Y-axis correspond to 80% .

Original file path: Final report/1.Data_Assess/CV/*ECDF*

3.5 Sample quantification histogram

The results of sample content are grouped by statistics, and the statistical results are shown in the following table.

Table 7: Statistical results table

Index	Group	N	Mean
Carotenoid_01	D	6	105.12
Carotenoid_01	C	6	60.589
Carotenoid_01	B	6	39.359
Carotenoid_01	A	6	11.335
Carotenoid_02	D	6	43.253
Carotenoid_02	C	6	24.404
Carotenoid_02	B	6	62.356
Carotenoid_02	A	6	28.471
Carotenoid_03	D	6	74.852
Carotenoid_03	C	6	47.919

Original file path: Final report/1.Data_Assess/histogram/groups*.xlsx

The bar chart below shows the content difference of each substance in different groups.

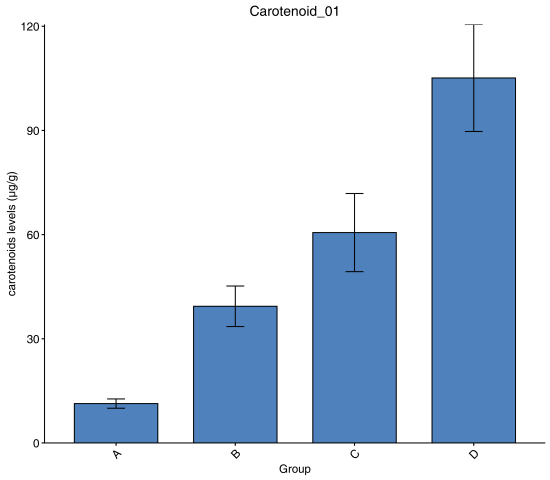

Fig 9: Sample content histogram

Note:The x-axis is the groups, the y-axis is the content, error bars are standard deviations.

Original file path: Final report/1.Data_Assess/histogram/histogram_compounds/*.png

3.6 Hierarchical Cluster Analysis

3.6.1 Principles of cluster analysis

Hierarchical Cluster Analysis (HCA) is a type of multivariate statistical analysis method. The samples are classified according to their features such that highest homogeneity is achieved between sample from the same group and highest heterogeneity is achieved between samples from different groups. In this report, the compound quantification data was normalized (Unit Variance Scaling, UV Scaling) and heatmaps were drawn by R software Pheatmap package. Hierarchical Cluster Analysis (HCA) was used to cluster the samples.

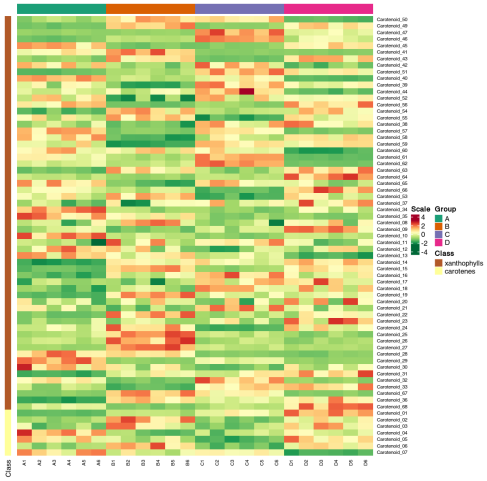

Fig 10: Sample clustering diagram

Note: X-axis indicates the sample name and the Y-axis are the metabolites. Group indicates sample groups. Z-Score indicates the relative quantification of each metabolite with red representing higher content and green representing lower content. Cluster analysis was performed on both metabolites (vertical cluster tree) and samples (horizontal cluster tree). “all_heatmap_class”: Heat map based on metabolite classification; “all_heatmap_no_cluster”: Showing only heatmap.

Original file path: Final report /1.Data_Assess/*all_heatmap*

4 Analysis results

4.1 Differential metabolite screening

Metabolomics data is characterized by “high-dimensional and massive”, requiring a large number of replicate samples to accurately uncover differential metabolites. Variance screening was performed based on the fold change value. Screening criteria:

For two sets of comparisons:

1.Metabolites with Fold Change ≥ 2 and Fold Change ≤ 0.5 were considered as significant and selected.

A partial result from the screening criteria is seen below:

Table 8: Screening results of differential metabolites

Index	Compounds	Type
Carotenoid_01	α-carotene	down
Carotenoid_02	lycopene	down
Carotenoid_04	β-carotene	up
Carotenoid_50	zeaxanthin-oleate-palmitate	down
Carotenoid_49	zeaxanthin-palmitate-stearate	down
Carotenoid_45	zeaxanthin dimyristate	up
Carotenoid_43	zeaxanthin dilaurate	down
Carotenoid_51	β-cryptoxanthin laurate	down
Carotenoid_40	zeaxanthin myristoleate	up
Carotenoid_39	violaxanthin dioleate	up

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/group-ID*_vs_group-ID*filter.xlsx.

4.1.1 Differential metabolite statistics

The number of different metabolites in each group is shown in the table below:

Table 9: Statistical table of differential metabolites

group name	All sig diff	down regulated	up regulated
A_vs_B	33	19	14

Statistical table of differential metabolites:Final report/2.Basic_analysis/Difference_analysis/sigMetabolitesCount.xl

4.2 Functional annotation and enrichment analysis of differential metabolites in KEGG database

KEGG (Kyoto Encyclopedia of Genes and Genomes) is a database that integrates compounds and genes into metabolic pathways. The KEGG database enabled researchers to study genes with their expression information and compounds with their abundances as a complete network.

4.2.1 Functional annotation of differential metabolites

Metabolites are annotated using the KEGG database, and only metabolic pathways containing differential metabolites are shown. Detailed results are found in the attached results. A portion of the results is shown below:

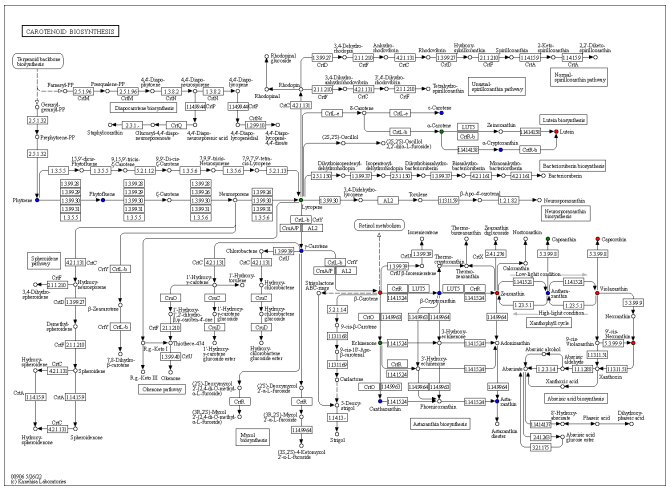

Fig 11: KEGG pathway of metabolites

Note: Red circles indicate that the metabolite content was significantly up-regulated in the experimental group; the blue circles indicate that the metabolite content was detected but did not change significantly; Green circles indicate that the metabolite content was significantly down-regulated in the experimental group. The orange circles indicate a mixture of both up-regulated and down-regulated metabolites.This allows searching for metabolites that may contribute to the phenotypic differences.

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/enrichment/Graph/ko*.

Statistical analysis of KEGG database annotation of screened metabolites with significant differences. Some of the results are as follows:

Table 10: KEGG annotations for differential metabolites

Index	Compounds	Type	cpd_ID
Carotenoid_01	α-carotene	down	C05433
Carotenoid_02	lycopene	down	C05432
Carotenoid_04	β-carotene	up	C02094
Carotenoid_50	zeaxanthin-oleate-palmitate	down	–
Carotenoid_49	zeaxanthin-palmitate-stearate	down	–
Carotenoid_45	zeaxanthin dimyristate	up	–
Carotenoid_43	zeaxanthin dilaurate	down	–
Carotenoid_51	β-cryptoxanthin laurate	down	–
Carotenoid_40	zeaxanthin myristoleate	up	–
Carotenoid_39	violaxanthin dioleate	up	–

Table 11: Enrichment Statistics of KEGG annotations for differential metabolites

ko_ID	Sig_compound	compound	Sig_compound_all	compound_all
ko00906	10	19	11	20
ko01100	8	13	11	20
ko01110	11	19	11	20
ko01240	2	2	11	20
ko00999	2	2	11	20

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/enrichment/group-ID*_vs_group-ID*_filter_kegg.xlsx.

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/enrichment/group-ID*_vs_group-ID*_KEGG.xlsx.

4.2.2 KEGG classification of differential metabolites

The significant differential metabolites were classified based on pathway annotation . The results are as follows:

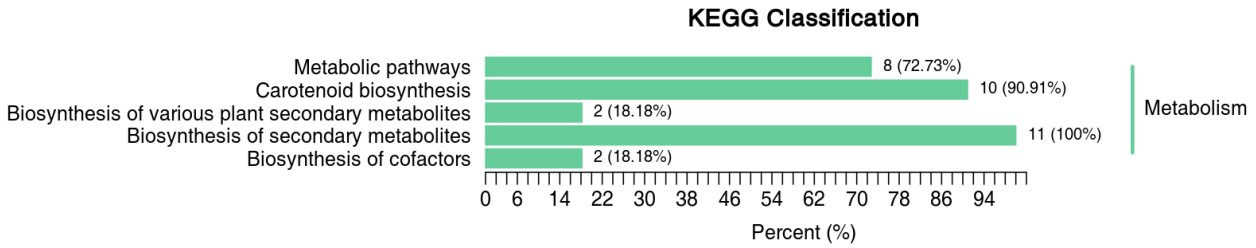

Fig 12: KEGG classification of differential metabolites
Note: the Y-axis shows the name of the KEGG pathway. The number of metabolites and the proportion of the total metabolites are shown next to the bar plot.

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/enrichment/group-ID*_vs_group-ID_KEGG_barplot.*.

4.2.3 KEGG enrichment analysis of differential metabolites

KEGG pathway enrichment analysis was conducted based on the annotation results. We calculated the Rich Factor for each pathway, which is the ratio of the number of differenetial metabolites in the corresponding pathway to the total number of metabolites annotated in the same pathway. The greater the Rich Factor, the greater the degree of enrichment. P-value is the calculated using hypergeometric test as shown below:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}$$

N represents the total number metabolites with KEGG annotation, n represents the number of differential metabolites in N, M represents the number of metabolites in a KEGG pathway in N, and m represents the number of differential metabolites in a KEGG pathway in M. The closer the p-value to 0, the more significant the enrichment. The size of the dots in the figure represents the number of significantly different metabolites enriched in the corresponding pathway. The results are shown below:

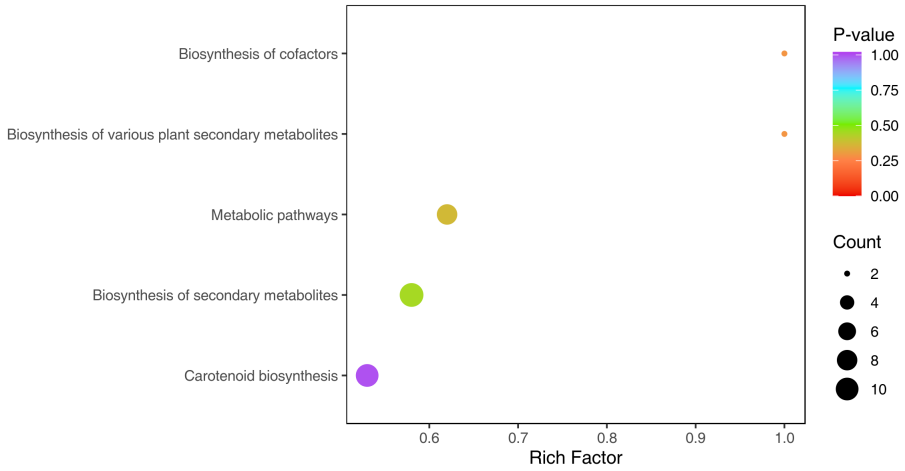

Fig 13: KEGG enrichment diagram of differential metabolites

Note: The X-axis represents the Rich Factor and the Y-axis represents the pathway. The color of points reflects the p-value. The darker the red, the more significant the enrichment. The size of the dot represents the number of enriched differential metabolites.

Final report/2.Basic_analysis/Difference_analysis/group-ID*_vs_group-ID*/enrichment/group-ID*_vs_group-ID*_KEGG_Enrichment.*.

5 References

1. Edge R, McGarvey DJ, Truscott TG. The carotenoids as anti-oxidants-a review. Journal of Photochemistry and Photobiology B Biology.1998, 41(3):189-200.
2. Krinsky NI, Mayne ST, Sies H. Carotenoids in Health and Disease. CRC Press Inc. 2004.
3. Bartleyg E, Scolnik PA. Plant carotenoids: Pigments for photoprotection, visual attraction and human health. Plant Cell. 1995, 7:1027-1038.
4. Geyer R, Peacock AD, White DC, Lytle C, Van GJ. Atmospheric pressure chemicalionization and atmospheric pressure photoionization forsimultaneous mass spectrometric analysis of microbial respiratory ubiquinones and menaquinones. Journal of Mass Spectrom. 2004, 39:922–929.
5. Inbaraj BS, Lu H, Hing CF, Wu CF, Lin CL, Chen BH. Determination of carotenoids and their esters in fruits of Lycium barbarum Linnaeus by HPLC-DAD-APCI-MS.Journal of Pharmaceutical and Biomedical Analysis. 2008, 47:812–818.
6. Petry FC and Mercadante AZ. New method for carotenoid extraction and analysis by HPLC-DAD-MS/MS in freeze-dried Citrus and Mango pulps. Journal of the Brazilian Chemical Society. 2018, 1:205-215.
7. Rodrigues DB, MariuttiL RB and Mercadante AZ. Two-step cleanup procedure for the identification of carotenoid esters by liquid chromatography-atmospheric pressure chemical ionization-tandem mass

spectrometry. J Chromatogr A. 2016, 1457:116-124.

8. Rivera SM, Christou P and Canela-Garayoa R. Identification of carotenoids using mass spectrometry. Mass Spectrom Rev. 2014, 33(5):353-372.
9. P Regal, KT Amorim-Carrilho, A Cepeda, C Fente. Review of methods for analysis of carotenoids. Trends in Analytical Chemistry. 2014, 56:49-73.
10. Murillo E, et al. Native carotenoids composition of some tropical fruits. Food Chem. 2013. 140(4):825-836.

6 Appendix

6.1 Analytical methods

1.Hierarchical Cluster Analysis

The HCA (hierarchical cluster analysis) results of samples and metabolites were presented as heatmaps with dendrograms. HCA was carried out by R package pheatmap. For HCA, normalized signal intensities of metabolites (unit variance scaling) are visualized as a color spectrum.

2.Differential metabolites selected

Significantly regulated metabolites between groups were determined by absolute Log_2FC (fold change).

3.KEGG annotation and enrichment analysis

Identified metabolites were annotated using KEGG compound database (http://www.kegg.jp/kegg/compound/), annotated metabolites were then mapped to KEGG Pathway database (http://www.kegg.jp/kegg/pathway.html). Pathways with significantly regulated metabolites mapped to were then fed into MSEA (metabolite sets enrichment analysis), their significance was determined by hypergeometric test's p-values.

6.2 List of software and versions

Data processing methods were mainly adopted in the analysis process in two ways:

(1) unit variance scaling (UV)

Unit variance Scaling (UV) is also called Z-Score standardization, i.e., auto scaling. This method standardizes data according to mean and standard deviation of original data. The processed data conform to the standard normal distribution, that is, the mean value is 0 and the standard deviation is 1.

Calculation method: Divide the original data center by standard deviation.

The formula is as follows:

$$x' = \frac{x - \mu}{\sigma}$$

Where μ is the mean value and σ is the standard deviation.

(2) Centralization/zero-mean-centered (Ctr)

Calculation method: subtract the mean of the variables from the original data.

The formula is as follows:

$$x' = x - \mu$$