

Comparative chemical profiling of callus and seeds of *Theobroma cacao* by precolumn derivatization GC-MS and LC-MS

Xiufang Wang

Hebei University of Chinese Medicine

Jiabao Chen

Hebei University of Chinese Medicine

Guangtian Yang

Hebei University of Chinese Medicine

Qian Zheng

Shijiazhuang Traditional Chinese Medicine Hospital

Yuguang Zheng

Hebei Chemical and Pharmaceutical College

Lei Wang

lei.wang@hebcm.edu.cn

Hebei University of Chinese Medicine

Research Article

Keywords: *Theobroma cacao*, tissue culture, chemical analysis, mass spectrometry, multivariate statistical analysis

Posted Date: April 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-8390136/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Theobroma cacao L. is a vital tropical economic crop. The seeds of cacao serve as core raw material for chocolate industry. However, conventional breeding for cacao production was challenged by long-term constraints from pests, diseases, climate changes, and soil pollution. Plant tissue culture offers a viable approach for cacao variety improvement, resource conservation and metabolism regulation research. In the present study, cacao callus was successfully induced from hypocotyls of aseptically germinated cacao seeds. The obtained cacao callus was maintained on Murashige & Skoog medium supplemented with 0.1 mg/L 1-naphthaleneacetic acid and 2 mg/L 6-benzylaminopurine for rapid and stable cacao callus proliferation. Subsequently, precolumn derivatization GC-MS and LC-MS coupled with multivariate statistical analysis were employed to conduct metabolomic profiling of cacao callus and cacao seeds. A total of 182 metabolites were annotated, including amino acids, organic acids, sugars, flavonoids, alkaloids and others. All the annotated metabolites were detected in both cacao callus and cacao beans but with their concentrations varied. The established cacao tissue culture system could not only provide stable materials for investigating cacao metabolic regulation mechanisms but also facilitate the preservation, rejuvenation, and genetic transformation of elite germplasm which would be helpful for the healthy development of cacao industry.

1 Introduction

Theobroma cacao L. is a perennial plant in the family of Malvaceae (Meza-Sepúlveda et al. 2025). It is endemic to tropical regions, thriving in warm and humid climates (Rocha et al. 2025). The seeds of cacao, normally called cacao beans, are the materials for production of chocolates and cacao drinks. Modern chemical and pharmacological research unveiled bioactive gradients of cacao beans, which further extended its applications in pharmaceuticals and cosmetics (Nieves-Orduña et al. 2023). Cacao products exhibit complex chemical compositions, comprising lipids, proteins, carbohydrates, polyphenols, xanthines, carboxylic acids, amines, vitamins, volatile compounds, and minerals (Alotaibi et al. 2024, Azizah et al. 2007, Edo et al. 2023). These bioactive constituents demonstrated broad pharmacological effects, including significant efficacy in preventing cardiovascular diseases and cancer, as well as notable anti-inflammatory, antimicrobial, and antioxidant properties (Hurst et al. 2002, Oracz et al. 2015). As an economic crop and the cornerstone of the chocolate industry, cacao also serves as a vital agricultural commodity sustaining the livelihoods of millions worldwide (Kongor et al. 2024). Annual global production of cacao reached approximately 4.5 million metric tons, generating over 10 billion USD income for farming households (Paisic-Ramirez et al. 2024). However, cacao production faces multiple constraints, such as pest infestations, diseases, climate change, soil contamination, and socioeconomic challenges (Amponsah-Doku et al. 2022, Darmawan et al. 2024, Kongor et al. 2024). Biotechnological breeding provides solutions for variety breeding and improvement.

Plant tissue culture refers to an aseptic cultivation of isolated plant cells, tissues, organs, or their components under controlled physical and chemical conditions. This technology holds significant values in fundamental scientific research and commercial applications across agriculture, horticulture, and pharmaceuticals. For most of land plants, wounded tissues regenerate by forming undifferentiated callus tissue. Such natural phenomenon provides the biological basis for plant *in vitro* culture. Under sterile conditions, various explants, such as roots, stems, leaves, anthers, hypocotyls, and embryos, can be induced to form callus. Through regular subculturing, callus can maintain its proliferative capacity *in vitro* for extended periods. The differentiation potential of callus depends on culture conditions. With optimized hormone ratios and growing conditions, callus could undergo redifferentiation, form adventitious shoots or somatic embryos, and ultimately develop into a complete plant. Such characteristic endowed tissue culture with applications in micropropagation, genetic transformation, and germplasm conservation. Plant tissue culture technology also served as an essential tool for studying plant development and laid a solid foundation for modern agricultural research (Efferth et al. 2019). Especially for woody plants that characterized by a long growth cycle, slow reproduction and complex genetic background, tissue culture can shorten the breeding period from several years or decades down to just a few days. Callus that induced from tissue culture could serve as a standardized *in vitro* model which is more manipulable and axenic than plant or mature seeds which are vulnerable to complex environments. However, different from plants and mature seeds, callus are undifferentiated cells. It is still unclear whether callus retains a core metabolite repertoire that overlaps with that of seeds and to which extent callus can recapitulate the metabolic traits of seeds or plants. Thus, it is essential to systematically comparing the chemical profiles of cacao callus and cacao seeds before using the callus as materials for metabolism regulation study.

Characteristic chemical composition of cacao endowed its products with unique flavor and extended its applications in food, pharmaceutical and cosmetic industries. Metabolomics aims to capture a comprehensive snapshot of chemical composition within a biological matrix (Pu et al. 2025). With technological advancements, high throughput annotation of bioactive compounds greatly promoted research in natural products, chemotaxonomy, and related disciplines (Gonda et al. 2020, Mhlongo et al. 2020, Szűcs et al. 2019). Nowadays, advanced metabolomic techniques across various instrumental platforms allow in-depth analysis of chemical diversity. In this study, a high-throughput metabolomic platforms based on mass spectrometry (MS) were applied to systematically compare the metabolic profiles of cacao callus and cacao seeds. Multivariate statistical analysis was further conducted to elucidated their chemical variations. The established rapid propagation system of cacao could provide sufficient material for molecular breeding. Chemical comparison of cacao seeds and cacao callus would be helpful in clarifying their chemical distinctions and essential for subsequent research on metabolic regulation of cacao.

2 Materials and methods

2.1 Chemicals and reagents

HPLC grade methanol, acetonitrile and formic acid were purchased from Fisher Scientific (Pittsburgh, PA, USA). Ultrapure water was prepared by synergy water purification system (Millipore, Billerica, United States). The reserpine standards (HPLC grade) and GC grade derivatizing reagent N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA), methoxyamine hydrochloride were purchased from Sigma-Aldrich (St. Louis, MO, USA). GC grade *n*-hexane and anhydrous pyridine were obtained from Aladdin Biochemical Technology Co., Ltd. (Shanghai, China). Salicylic acid, luteolin, apigenin and rosmarinic acid standards were provided by Yuanye Bio-Technology Co., Ltd. (Shanghai, China). Reference standards of luteolin-7-O- glucoside, scutellarin, luteolin-7-O-

glucuronide, apigenin-7-O-glucoside and apigenin-7-O-glucuronide were purchased from Shanghai Standard Technology Co., Ltd. (Shanghai, China). The purities of all standards were determined to be higher than 98%. Murashige & Skoog Medium were purchased from Qingdao Hope Bio-Technology Co., Ltd. (Shandong, China). Other chemicals and reagents were analytical grade.

2.2 Plant material

Cacao pods (Forastero) were purchased from Wanning, Hainan Province, China, and identified by Prof. Yuguang Zheng. The fresh cacao seeds were taken from the pods with their pulp removed. Some of the cacao seeds were stored at the Innovation Centre for Traditional Chinese Medicine Processing Technology of Hebei Province, Hebei University of Chinese Medicine as voucher specimens with specimen number of TC202002.

2.3 Induction and maintenance of cacao callus

Fresh cacao seeds were washed twice with sterilized water. The cleaned cacao seeds were firstly surface-sterilized with 75% ethanol for 1 min and rinsed three times with sterilized water to remove the residue disinfectants. Subsequently, the cacao seeds were disinfected with sodium hypochlorite solution (available chlorine 5%) for 4 min and then rinsed 6 times with sterilized water. Finally, the surface sterilized cacao seeds were placed on Murashige & Skoog medium (MS medium) and cultured at $25^{\circ}\text{C} \pm 2^{\circ}\text{C}$ in dark to promote the growth of hypocotyl. The developing hypocotyls were excised aseptically into 0.5 ~ 1 cm pieces and cultured on MS medium supplemented with 0.1 mg/L 1-naphthaleneacetic acid (NAA) and 2 mg/L 6-benzylaminopurine (6-BA) for callus induction. The generated cacao callus was sub-cultured every two weeks and incubated at $25^{\circ}\text{C} \pm 2^{\circ}\text{C}$ under a photoperiod of 16h light/8h dark cycle.

2.4 Metabolite extraction and profiling

Cacao callus and cacao seeds were ground to fine powder in liquid nitrogen with mortar and pestle. 100 mg of each sample powder was mixed with 1 mL extraction solvent that composed of methanol, water, and formic acid in a ratio of 70:28:2. After vortex vigorously, the mixture was sonicated (300W, 40 kHz) 15 min at room temperature. The extracts were separated from the residue by centrifugation at 13000 rpm for 10 min.

Pre-column derivatization and GC-MS analysis was performed with an Agilent 7890B- 5977B GC-MS (Agilent, Santa Clara, CA, USA) coupled with a HP-5MS capillary column (30 m $\times$ 0.25 mm, 0.25 mm film thickness, Agilent, Santa Clara, CA, USA) basing on a previously validated chromatographic method (Chen et al. 2022) with a scan range of m/z 30–400. UHPLC-Q/TOF-MS analysis was conducted using an Agilent 1290 UHPLC system coupled with an Agilent 6545 quadrupole time-of-flight mass spectrometer (Agilent Technologies, Santa Clara, CA, USA). Chromatographic separation was achieved on an Agilent ZORBAX SB-C18 column (4.6 $\times$ 50 mm, 1.8 μm particle size) according to a reported chromatographic method (Chen et al. 2022). Mass spectrometric detection was performed in positive ionization mode with a capillary voltage of 4000 V and collision energies of 35 eV, scanning from m/z 50 to 1000.

2.5 Data processing and multivariate statistical analysis

For qualitative analysis, the metabolites detected by GC-MS with a similarity over 80% to the NIST17 standard library were annotated using the Agilent MassHunter analysis program (Agilent, Santa Clara, CA, USA). The raw LC-MS data files were converted into mzML format by MSConvert. The MS/MS spectrum of each feature was searched by MS-DIAL5 software based on accurate precursor masses and fragment masses. Intensity of each peak was recorded. For quantitative analysis, the integrated peak area was considered to be a variable for analysis and normalized to internal standard (0.2 mg/mL salicylic acid in GC-MS and 0.5 $\mu\text{g/L}$ reserpine in LC-MS analysis). The combined GC-MS and LC-MS dataset was transformed by Min-Max Scaling. Principal component analysis (PCA) and clustering analysis (CA) were performed with Origin Pro 2020 (OriginLab Corporation, USA) software. ANOVA was performed with IBM SPSS Statistics 23.0 (IBM, USA) software.

3 Results and discussion

3.1 Induction and maintenance of cacao callus

Surface disinfected cacao seeds were placed on MS medium and cultured at 25°C in dark to promote the growth of hypocotyl (Fig. 1A). The developing hypocotyls were excised aseptically into 0.5 ~ 1 cm pieces then cultured on MS medium supplemented with 0.1 mg/L NAA and 2 mg/L 6-BA (Fig. 1B). The cutting surfaces of hypocotyl pieces turned brown instantly (Fig. 1B). After around two weeks cultivation, the hypocotyl pieces began to expand and pale callus clusters started to form at the cutting surface (Fig. 1C). The obtained amorphous, faint yellow and friable callus aggregates were subcultured every two weeks. Two weeks old callus (Fig. 1D) were collected for further chemical analysis.

3.2 Chemical analysis of cacao callus and cacao beans

Stably growing callus are useful materials for production of bioactive metabolites and investigation of their biosynthetic regulations. To clarify the metabolomic overviews of cacao callus and cacao seed, GC-MS and LC-MS were applied for high throughput chemical analysis. The typical total ion chromatograms (TICs) obtained from GC-MS and LC-MS analyses revealed distinct metabolomic profiles of cacao seeds and callus (Fig. 1).

Pre-column derivatization coupled with GC-MS analysis allowed annotation of 104 metabolites by comparing the mass spectra of detected peaks with those in the NIST 17 database. The annotated metabolites included sugars (32), amino acids (14), organic acids (32), terpenoids (5) and others with their detail information shown in Table 1.

LC-MS analysis enabled characterization of another 77 metabolites, predominantly flavonoids and alkaloids, with detailed information such as retention times, molecular formulas, and fragment ions provided in Table 2. All the annotated metabolites were detected in both cacao seeds and cacao callus. However, the corresponding peak intensities of the characterized metabolites varied between cacao seeds and callus, indicating differences in their relative contents .

Table 1
Annotation of compounds analyzed by GC-MS

No.	Retention time (min)	Compound name	Molecular weight	Molecular formula	Classification
1	3.503	Chrysanthemol (TMS)	154.25	C ₁₀ H ₁₈ O	Terpenoids
2	3.633	1,3-Propanediol (2TMS)	76.09	C ₃ H ₈ O ₂	Others
3	3.689	Lactic Acid (2TMS)	90.08	C ₃ H ₆ O ₃	Organic acids
4	3.788	Hexanoic acid (TMS)	116.16	C ₆ H ₁₂ O ₂	Organic acids
5	3.832	Glycolic acid (2TMS)	76.05	C ₂ H ₄ O ₃	Organic acids
6	3.928	1,3-Butanediol (2TMS)	90.12	C ₄ H ₁₀ O ₂	Others
7	4.157	L-Alanine (2TMS)	89.09	C ₃ H ₇ NO ₂	Amino acids
8	4.59	Oxalic acid (2TMS)	90.03	C ₂ H ₂ O ₄	Organic acids
9	4.662	Ethylene glycol (2TMS)	62.07	C ₂ H ₆ O ₂	Others
10	4.817	1,4-Butanediol (2TMS)	90.12	C ₄ H ₁₀ O ₂	Others
11	5.341	Propanedioic acid (2TMS)	104.06	C ₃ H ₄ O ₄	Organic acids
12	5.506	L-Valine (2TMS)	117.15	C ₅ H ₁₁ NO ₂	Amino acids
13	5.74	Formamide (2TMS)	45.041	CH ₃ NO	Others
14	5.879	Benzoic acid (TMS)	122.12	C ₇ H ₆ O ₂	Organic acids
15	6.039	Octanoic acid (TMS)	144.21	C ₈ H ₁₆ O ₂	Organic acids
16	6.147	Ethanolamine (3TMS)	61.08	C ₂ H ₇ NO	Others
17	6.212	L-Leucine (2TMS)	131.17	C ₆ H ₁₃ NO ₂	Amino acids
18	6.255	Glycerol (3TMS)	92.09	C ₃ H ₈ O ₃	Others
19	6.494	L-Isoleucine (2TMS)	131.17	C ₆ H ₁₃ NO ₂	Amino acids
20	6.537	L-Proline (2TMS)	115.13	C ₅ H ₉ NO ₂	Amino acids
21	6.609	Maleic acid (2TMS)	116.07	C ₄ H ₄ O ₄	Organic acids
22	6.659	Glycine (3TMS)	75.07	C ₂ H ₅ NO ₂	Amino acids
23	6.698	Succinic acid (2TMS)	118.09	C ₄ H ₆ O ₄	Organic acids
24	6.991	Glyceric acid (3TMS)	106.08	C ₃ H ₆ O ₄	Organic acids
25	7.105	Fumaric acid (2TMS)	116.07	C ₄ H ₄ O ₄	Organic acids
26	7.25	Nonanoic acid (TMS)	158.24	C ₉ H ₁₈ O ₂	Organic acids
27	7.335	Amphetamine (2TMS)	135.21	C ₉ H ₁₃ N	Others
28	7.365	L-Serine (3TMS)	105.09	C ₃ H ₇ NO ₃	Amino acids
29	7.445	Lavandulol (TMS)	154.25	C ₁₀ H ₁₈ O	Terpenoids
30	7.606	4-Hydroxybutanoic acid (2TMS)	104.1	C ₄ H ₈ O ₃	Organic acids
31	7.647	4-Hydroxyisovaleric acid (2TMS)	118.13	C ₅ H ₁₀ O ₃	Organic acids
32	7.708	L-Threonine (3TMS)	119.12	C ₄ H ₉ NO ₃	Amino acids
33	7.829	Methylsuccinic acid (2TMS)	132.11	C ₅ H ₈ O ₄	Organic acids
34	8.116	Butyric acid (TMS)	88.11	C ₄ H ₈ O ₂	Organic acids
35	8.163	L-Alanine (3TMS)	89.09	C ₃ H ₇ NO ₂	Amino acid s
36	8.259	Butanoic acid (3TMS)	88.11	C ₄ H ₈ O ₂	Organic acids

No.	Retention time (min)	Compound name	Molecular weight	Molecular formula	Classification
37	8.406	2,3-Butanediol (2TMS)	90.12	C ₄ H ₁₀ O ₂	Others
38	8.451	Decanoic acid (TMS)	172.26	C ₁₀ H ₂₀ O ₂	Organic acids
39	8.48	Homoserine (3TMS)	119.12	C ₄ H ₉ NO ₃	Amino acids
40	8.787	2,2-Dimethylbutanoic acid (2TMS)	116.16	C ₆ H ₁₂ O ₂	Organic acids
41	8.96	Malic acid (3TMS)	134.09	C ₄ H ₆ O ₅	Organic acids
42	9.069	3-Amino-2-piperidone (2TMS)	114.15	C ₅ H ₁₀ N ₂ O	Others
43	9.206	Salicylic acid (2TMS)	138.12	C ₇ H ₆ O ₃	Internal standard
44	9.261	Erythritol (4TMS)	122.12	C ₄ H ₁₀ O ₄	Sugars
45	9.345	L-Pyrogutamic acid (2TMS)	129.11	C ₅ H ₇ NO ₃	Amino acids
46	9.428	4-Aminobutanoic acid (3TMS)	103.12	C ₄ H ₉ NO ₂	Organic acids
47	9.533	Dihydroactinidiolide (2TMS)	180.24	C ₁₁ H ₁₆ O ₂	Others
48	9.656	Elemicin	208.25	C ₁₂ H ₁₆ O ₃	Terpenoids
49	9.702	2,3,4-Trihydroxybutyric acid (4TMS)	136.1	C ₄ H ₈ O ₅	Organic acids
50	9.778	Dodecanol (TMS)	186.33	C ₁₂ H ₂₆ O	Others
51	9.924	Tartronic acid (3TMS)	120.06	C ₃ H ₄ O ₅	Organic acids
52	10.045	2-Aminooctanoic acid (2TMS)	159.23	C ₈ H ₁₇ NO ₂	Organic acids
53	10.157	2,2,4-Trimethyl-1,3-pentanediol diisobutyrate	286.41	C ₁₆ H ₃₀ O ₄	Others
54	10.491	L-Glutamic acid (3TMS)	147.13	C ₅ H ₉ NO ₄	Amino acids
55	10.526	D-Ribofuranose (4TMS)	150.13	C ₅ H ₁₀ O ₅	Sugars
56	10.582	L-Phenylalanine (2TMS)	165.19	C ₉ H ₁₁ NO ₂	Amino acids
57	10.66	D-Xylonolactone (3TMS)	148.11	C ₅ H ₈ O ₅	Sugars
58	10.717	D-Arabinolactone (3TMS)	148.11	C ₅ H ₈ O ₅	Sugars
59	10.74	Dodecanoic acid (TMS)	200.32	C ₁₂ H ₂₄ O ₂	Organic acids
60	10.812	Tartaric acid (4TMS)	150.09	C ₄ H ₆ O ₆	Organic acids
61	10.864	3,4-Dihydroxyphenylglycol (4TMS)	170.16	C ₈ H ₁₀ O ₄	Others
62	11.055	D-Arabinose (4TMS)	150.13	C ₅ H ₁₀ O ₅	Sugars
63	11.098	L-Asparagine (3TMS)	132.12	C ₄ H ₈ N ₂ O ₃	Amino acids
64	11.124	D-Xylose (4TMS)	150.13	C ₅ H ₁₀ O ₅	Sugars
65	11.176	9,12-Octadecadiynoic acid (TMS)	276.4	C ₁₈ H ₂₈ O ₂	Organic acids
66	11.267	11-Hexadecyn-1-ol	238.41	C ₁₆ H ₃₀ O	Others
67	11.441	3-Deoxypentitol (4TMS)	136.15	C ₅ H ₁₂ O ₄	Sugars
68	11.493	D-Arabinitol (5TMS)	152.15	C ₅ H ₁₂ O ₅	Sugars
69	11.762	Xylitol (5TMS)	152.15	C ₅ H ₁₂ O ₅	Sugars
70	11.818	D-Fucitol (5TMS)	166.17	C ₆ H ₁₄ O ₅	Sugars
71	11.965	1-Tetradecanol (TMS)	214.39	C ₁₄ H ₃₀ O	Others
72	12.091	3-Deoxyhexitol (5TMS)	166.17	C ₆ H ₁₄ O ₅	Sugars
73	12.234	5-(1,2-Dihydroxyethyl)-3-hydroxyoxolan-2-one (3TMS)	162.14	C ₆ H ₁₀ O ₅	Sugars

No.	Retention time (min)	Compound name	Molecular weight	Molecular formula	Classification
74	12.268	L-Fructosonic acid (5TMS)	194.14	C ₆ H ₁₀ O ₇	Sugars
75	12.435	Ribonic acid (5TMS)	166.13	C ₅ H ₁₀ O ₆	Sugars
76	12.529	Erythrofuranose (3TMS)	120.1	C ₄ H ₈ O ₄	Sugars
77	12.629	Methyl galactoside (4TMS)	194.18	C ₇ H ₁₄ O ₆	Sugars
78	12.702	Protocatechoic acid (3TMS)	154.12	C ₇ H ₆ O ₄	Organic acids
79	12.791	Citric acid (4TMS)	192.12	C ₆ H ₈ O ₇	Organic acids
80	12.871	Myristic acid (TMS)	228.37	C ₁₄ H ₂₈ O ₂	Organic acids
81	12.915	2-Deoxyribonic acid (4TMS)	150.13	C ₅ H ₁₀ O ₅	Sugars
82	12.971	Cubenol (TMS)	222.37	C ₁₅ H ₂₆ O	Terpenoids
83	13.057	Occidentalol (TMS)	220.35	C ₁₅ H ₂₄ O	Terpenoids
84	13.159	Isobutyl phthalate	278.34	C ₁₆ H ₂₂ O ₄	Others
85	13.318	D-galactono-1,4-lactone (4TMS)	178.14	C ₆ H ₁₀ O ₆	Sugars
86	13.327	Quininic acid (5TMS)	203.19	C ₁₁ H ₉ NO ₃	Organic acids
87	13.491	D-Fructose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
88	13.565	D-Gluconolactone (4TMS)	178.14	C ₆ H ₁₀ O ₆	Sugars
89	13.613	D-Tagatose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
90	13.725	D-Galactose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
91	13.795	D-Glucose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
92	13.869	D-Talofuranose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
93	13.969	D-Mannose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
94	14.012	D-Allose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
95	14.107	D-Talose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
96	14.211	Ethyl glucoside (4TMS)	208.21	C ₈ H ₁₆ O ₆	Sugars
97	14.228	Dibutyl phthalate	278.34	C ₁₆ H ₂₂ O ₄	Others
98	15.048	D-Gluconic acid (6TMS)	196.16	C ₆ H ₁₂ O ₇	Sugars
99	15.146	Palmitic acid (TMS)	256.42	C ₁₆ H ₃₂ O ₂	Organic acids
100	15.217	Galactaric acid (6TMS)	210.14	C ₆ H ₁₀ O ₈	Sugars
101	15.335	Scyllo-Inositol (6TMS)	180.16	C ₆ H ₁₂ O ₆	Others
102	16.086	Myo-Inositol (6TMS)	180.16	C ₆ H ₁₂ O ₆	Others
103	16.708	D-Allose (6TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
104	16.778	D-Altrose (5TMS)	180.16	C ₆ H ₁₂ O ₆	Sugars
105	17.563	Stearic acid (TMS)	284.5	C ₁₈ H ₃₆ O ₂	Organic acids

Table 2
Annotation of metabolites analyzed by UPLC-ESI-Q-TOF-MS/MS

No.	RT (min)	Adduct ions (m/z)	Molecular ions(m/z)	Fragment ions in MS/MS (m/z)	Molecular formula
1	1.098	[M+NH ₄] ⁺	360.1476	127.0375,97.0274,85.0275,69.0330,55.0174	C ₁₂ H ₂₂ O ₁₁
2	1.134	[M + H] ⁺	138.0538	105.0207,92.0485,78.0331,65.0384,53.0384	C ₇ H ₇ NO ₂
3	1.150	[M + H] ⁺	130.086	84.0800,70.0651,56.0494	C ₆ H ₁₁ NO ₂
4	1.181	[M + H] ⁺	136.0608	119.0337,109.050292.0234,65.0382,55.0289	C ₅ H ₅ N ₅
5	1.196	[M + H] ⁺	268.1024	136.0605,119.0339,71.0124	C ₁₀ H ₁₃ N ₅ O ₄
6	1.243	[M + H] ⁺	284.0971	194.0382,152.0547,135.0294,110.0334,85.0273	C ₁₀ H ₁₃ N ₅ O ₅
7	1.273	[M + H] ⁺	182.0799	136.0730,119.0478,107.0484,91.0535,77.0381,65.0379	C ₉ H ₁₁ NO ₃
8	1.609	[M + H] ⁺	177.1007	159.0666,143.0702,132.0794,115.0529,105.0681,91.0531	C ₁₀ H ₁₂ N ₂ O
9	1.627	[M + H] ⁺	189.1008	174.0765,160.0738,146.0586,127.0525,117.0559,91.0536,77.0379	C ₁₁ H ₁₂ N ₂ O
10	1.767	[M + Na] ⁺	447.1449	315.1025,301.1155	C ₁₇ H ₂₈ O ₁₂
11	1.859	[M + H] ⁺	166.0851	120.0796,103.0534,91.0534,77.0382	C ₉ H ₁₁ NO ₂
12	1.877	[M + H] ⁺	103.0532	77.0381,51.0226	C ₃ H ₆ N ₂ O ₂
13	1.968	[M + Na] ⁺	267.0961	164.9987,130.0485,121.0271,93.0326,84.0437,65.0382	C ₁₅ H ₁₆ O ₃
14	2.046	[M + H] ⁺	199.0589	171.1027,140.0448,125.0218,97.0276,77.0381,65.0382	C ₉ H ₁₀ O ₅
15	2.048	[M + H] ⁺	181.0731	163.0595,138.0648,122.0574,108.0544,83.0595,67.0287	C ₇ H ₈ N ₄ O ₂
16	2.369	[M + H] ⁺	298.095	152.0409,136.0605,119.0339,75.0256,61.0103	C ₁₁ H ₁₅ N ₅ O ₃ S
17	2.478	[M + H] ⁺	201.1005	185.0698,160.0740,132.0795,82.0280,68.0486	C ₁₂ H ₁₂ N ₂ O
18	2.618	[M + H] ⁺	122.0957	103.0535,77.0378,51.0225	C ₈ H ₁₁ N
19	2.725	[M + H] ⁺	413.1386	162.0392,136.0604,120.0647	C ₁₅ H ₂₀ N ₆ O ₈
20	2.819	[M+NH ₄] ⁺	448.1791	151.0365,137.0584,85.0274	C ₁₉ H ₂₆ O ₁₁
21	3.065	[M + H] ⁺	205.0957	170.0576,143.0714,118.0640,91.0533,77.0376	C ₁₁ H ₁₂ N ₂ O ₂
22	3.234	[M + H] ⁺	287.0526	269.0407,241.0467,185.0567,137.0218,123.0427,109.0265	C ₁₅ H ₁₀ O ₆
23	3.250	[M + H] ⁺	453.1354	273.0715,165.0528,147.0414,139.0372,123.0427,85.0273	C ₂₁ H ₂₄ O ₁₁
24	3.255	[M + H] ⁺	449.1046	287.0527	C ₂₁ H ₂₀ O ₁₁
25	3.917	[M + H] ⁺	225.0741	192.0387,149.0214,119.0480,91.0533,77.0381,65.0381	C ₁₁ H ₁₂ O ₅
26	4.010	[M+NH ₄] ⁺	496.1989	263.0889,185.0792,170.0552,153.0532,125.0581,73.0278	C ₂₀ H ₃₀ O ₁₃
27	4.196	[M + H- H ₂ O] ⁺	177.0534	134.0347,117.0318,89.0377,78.0457,63.0227	C ₁₀ H ₁₀ O ₄
28	4.257	[M + H] ⁺	222.111	1335.0422,118.0406,107.0476,91.0534,77.0380	C ₁₂ H ₁₅ NO ₃
29	4.273	[M + H] ⁺	144.0796	127.0534,115.0531,91.0533,77.0381,65.0380	C ₁₀ H ₉ N
30	4.288	[M+NH ₄] ⁺	344.1321	147.0424,119.0478,85.0276,69.0329	C ₁₅ H ₁₈ O ₈
31	4.353	[M + H] ⁺	161.1058	143.0716,117.0556,104.0621,91.0529,77.0381,65.0380,53.0385	C ₁₀ H ₁₂ N ₂
32	4.678	[M + H] ⁺	287.053	269.0424213.0527,185.0577,137.0225,109.0265,81.0327	C ₁₅ H ₁₀ O ₆
33	4.680	[M + H] ⁺	419.094	287.0526,137.0225	C ₂₀ H ₁₈ O ₁₀

No.	RT (min)	Adduct ions (m/z)	Molecular ions(m/z)	Fragment ions in MS/MS (m/z)	Molecular formula
34	4.813	[M + H] ⁺	307.0784	139.0374,111.0428,93.0319,65.0370	C ₁₅ H ₁₄ O ₇
35	4.942	[M + H] ⁺	195.0858	180.0556,138.0645,123.0417,110.0699,83.0596,69.0433,56.0491	C ₈ H ₁₀ N ₄ O ₂
36	5.854	[M + H] ⁺	579.1457	289.0683,259.0572,163.0370,139.0376,123.0427	C ₃₀ H ₂₆ O ₁₂
37	6.995	[M + H] ⁺	165.0528	119.0473,103.0536,91.0529,77.0383,65.0380,51.0232	C ₉ H ₈ O ₃
38	7.000	[M + H] ⁺	291.0847	189.0517,161.0578,147.0425,139.0374,123.0428,119.0479,111.0423,77.0375,55.0174	C ₁₅ H ₁₄ O ₆
39	7.387	[M + H] ⁺	481.093	319.0424,209.0433,123.0431,74.0292	C ₂₁ H ₂₀ O ₁₃
40	7.523	[M + H] ⁺	289.0686	163.0348,135.0425,123.0427,107.0476,93.0328,77.0380,65.0384	C ₁₅ H ₁₂ O ₆
41	8.127	[M + H] ⁺	212.1169	195.0908,119.0594,94.0638,77.0381,65.0383	C ₁₃ H ₁₃ N ₃
42	8.436	[M+NH ₄] ⁺	490.1678	147.0427,119.0480	C ₂₄ H ₂₄ O ₁₀
43	9.504	[M + H] ⁺	198.0745	135.0447,117.0339,107.0478,93.0327,77.0378,65.0383	C ₉ H ₁₁ NO ₄
44	9.513	[M + H] ⁺	360.1066	163.0377,145.0274,135.0430,117.0326,107.0480,89.0379	C ₁₈ H ₁₇ NO ₇
45	9.691	[M+NH ₄] ⁺	760.298	383.1447,265.1046,235.0945,217.0843,205.0843,167.0689	C ₃₄ H ₄₆ O ₁₈
46	10.324	[M + H] ⁺	146.0586	128.0487,101.0375,91.0532,77.0383,65.0383	C ₉ H ₇ NO
47	10.436	[M + H] ⁺	275.0886	139.0364,107.0488,55.0535	C ₁₅ H ₁₄ O ₅
48	11.193	[M + H] ⁺	305.0636	231.0635,213.0530,153.0171,123.0426,93.0327	C ₁₅ H ₁₂ O ₇
49	12.521	[M + H] ⁺	449.1052	287.0531,153.0162,85.85.0269	C ₂₁ H ₂₀ O ₁₁
50	12.957	[M+NH ₄] ⁺	600.262	249.1100,187.0731,85.0272	C ₂₈ H ₃₈ O ₁₃
51	13.034	[M + H] ⁺	273.074	153.0169,119.0480,107.0476,91.0539,69.0325	C ₁₅ H ₁₂ O ₅
52	13.700	[M + H] ⁺	451.1209	289.0680,225.0504,163.0375	C ₂₁ H ₂₂ O ₁₁
53	14.428	[M + H] ⁺	303.0483	257.0414,229.0474,201.0526,153.0167,137.0216,121.0261,68.9965	C ₁₅ H ₁₀ O ₇
54	14.437	[M + H] ⁺	465.1006	303.0474,85.0277	C ₂₁ H ₂₀ O ₁₂
55	14.613	[M + H] ⁺	449.1048	287.0528	C ₂₁ H ₂₀ O ₁₁
56	16.870	[M + H] ⁺	435.0887	303.0468,107.0475,73.0273	C ₂₀ H ₁₈ O ₁₁
57	16.873	[M + H] ⁺	419.1676	315.0825,217.0839,190.0607,173.0573,145.0629,117.0687,81.0321	C ₂₂ H ₂₆ O ₈
58	16.888	[M+NH ₄] ⁺	598.2463	235.0943,217.0829,205.0842,190.0601,173.0584	C ₂₈ H ₃₆ O ₁₃
59	17.585	[M + H] ⁺	271.0586	215.0672,197.0585,153.0169,119.0469,91.0532,68.9962	C ₁₅ H ₁₀ O ₅
60	17.605	[M + H] ⁺	433.1107	215.0672,197.0585,153.0169,119.0469,91.0532,68.9962	C ₂₁ H ₂₀ O ₁₀
61	17.925	[M + H] ⁺	328.117	147.0425,119.0478,91.0531	C ₁₈ H ₁₇ NO ₅
62	17.956	[M + H] ⁺	655.2234	428.1416,147.0426	C ₃₀ H ₃₈ O ₁₆
63	18.791	[M + H] ⁺	317.0628	302.0408,274.0417,229.0470,217.0463,153.0161	C ₁₆ H ₁₂ O ₇
64	18.801	[M + H] ⁺	479.115	317.0622,85.0282	C ₂₂ H ₂₂ O ₁₂
65	19.071	[M + H] ⁺	449.1053	287.0513,123.0410	C ₂₁ H ₂₀ O ₁₁
66	19.938	[M + H] ⁺	463.1214	301.0676,286.0444,163.0744	C ₂₂ H ₂₂ O ₁₁
67	19.848	[M + H] ⁺	523.1404	361.0875,301.0706,186.8257	C ₂₄ H ₂₆ O ₁₃

No.	RT (min)	Adduct ions (m/z)	Molecular ions(m/z)	Fragment ions in MS/MS (m/z)	Molecular formula
68	20.247	[M + H] ⁺	289.0681	163.0371,153.0169,135.0419,117.0322,89.0371,65.0384	C ₁₅ H ₁₂ O ₆
69	30.278	[M + H] ⁺	235.1677	202.8846,179.0639,141.0677,134.0343,122.1019,109.1002,93.0682,77.0384,67.0537,58.0650	C ₁₅ H ₂₂ O ₂
70	32.100	[M + H] ⁺	301.0683	286.0455,212.0448,184.0498,168.0012,118.0404,57.0698	C ₁₆ H ₁₂ O ₆
71	32.987	[M + H] ⁺	361.0888	331.0421,303.0483,282.0490	C ₁₈ H ₁₆ O ₈
72	33.049	[M+NH ₄] ⁺	812.4734	439.3541,393.3473,249.1827,203.1775,191.1774	C ₄₂ H ₆₆ O ₁₄
73	34.033	[M + H] ⁺	609.2762	448.1933,397.2085,195.0635,174.0898	C ₃₃ H ₄₀ N ₂ O ₉
74	37.568	[M-H ₂ O + H] ⁺	415.2089	119.0846,107.0845,91.0536	C ₂₄ H ₃₂ O ₇
75	38.110	[M-H ₂ O + H] ⁺	471.3444	201.1615,189.1619,173.1309,133.0999,119.0840,107.0844,95.0846,81.0690,69.0692	C ₃₀ H ₄₈ O ₅
76	38.404	[M + H] ⁺	433.2193	247.1304,232.1071,216.1126,195.0994,181.0843,169.0841	C ₂₄ H ₃₂ O ₇
77	38.868	[M + H] ⁺	455.3494	398.3244,189.1619,158.0794,116.0695,95.0844,70.0647	C ₃₀ H ₄₆ O ₃
78	41.253	[M-H ₂ O + H] ⁺	439.354	203.1778,189.1618,147.1150,133.0996,107.0843,95.0846	C ₃₀ H ₄₈ O ₃

3.3 Multivariate statistical analysis for screening of differential metabolites between cacao callus and cacao beans

3.3.1 Principal component analysis

Principal component analysis (PCA), an unsupervised pattern recognition method, was firstly employed for further comparison the relative contents of annotated metabolites in cacao callus and cacao beans. PCA plot was performed with the integrated GC-MS and LC-MS dataset containing 181 independent variables which were normalized to internal standards and transformed by Min-Max scaling. As shown in the PCA plot (Fig. 2), cacao callus and cacao seed samples were separated into two groups indicating chemical composition differences of them. Principal component 1 (PC1) accounted for 79.3% of the variance, while PC2 explained 8.1%, yielding a cumulative variance contribution rate of 87.4% (Fig. 2).

3.3.2 Clustering analysis

Then hierarchical clustering analysis was performed with the results presented as heatmaps (Fig. 3) for an intuitive comparison of the contents of annotated metabolites in cacao callus and cacao seeds. Samples of cacao seeds and cacao callus were clustered separately into two distinct branches which is consistent with the PCA results. For further data interpretation, amino acids, organic acids, sugars and derivatives, flavonoids, alkaloids, terpenoids and other metabolites that with their contents varied significantly between cacao callus and cacao seeds were displayed respectively in Fig. 3B-3G.

3.3.2.1 Sugars and sugar derivatives

Sugars represent central components of plant primary metabolism, serving as the fundamental carbon carriers for all photosynthetically fixed carbon. Sugar metabolism provide essential carbon skeletons and energy required for plant growth, differentiation, and cellular maintenance. Virtually, all organic compounds in plants can be traced back to sugar-derived carbon precursors, underscoring the tightly regulated nature of sugar metabolism and its integration with developmental processes (Afoufa-Bastien et al. 2010). As demonstrated in Fig. 3B, the contents of most sugars and sugar derivatives were significantly higher in cacao callus than in cacao seeds. This observation aligned with the physiological expectation that actively growing callus, generate substantial quantities of sugars and related derivatives through their photosynthetic activity.

3.3.2.2 Amino acids

Amino acids are normally generated from glycolysis, TCA cycle or pentose phosphate pathway. They serve as fundamental building blocks for generation of secondary metabolites and proteins which are essential biomolecules responsible for cellular structure and function in all living organisms (Giri et al. 2013). As shown in Fig. 3C, relative contents of 6 amino acids, including cycloserine, phenylalanine, tryptophan, tyrosine, methylproline, and leucine were significantly higher in cacao seeds than in cacao callus. Aromatic amino acids such as phenylalanine plays critical roles in seed germination and synthesis of secondary metabolites in the phenylpropanoid pathway to support organelle development and plant defense (Perkowski et al. 2019). Conversely, the relative abundance of other amino acids was markedly higher in cacao callus. Among these amino acids, glycine, asparagin, and serine were reported to play pivotal roles in callus proliferation, energy metabolism, or intercellular signaling (Wang et al. 2023).

3.3.2.3 Organic acids

Organic acids are widely utilized in the food, pharmaceutical, and chemical industries (Sun et al. 2020). Cacao contains substantial amounts of organic acids. The contents of most of the detected organic acids were higher in cacao callus than in cacao seeds except decanoic acid, coumaric acid, syringic acid, quinic acid, tartaric acid, and oxalic acid (Fig. 3D). From the perspective of biosynthetic pathways, organic acids such as citric acid and succinic acid that participate in the tricarboxylic acid (TCA) cycle contribute to the generation of ATP and reducing power which are crucial for cellular biosynthesis (Niehaus. 2021). As shown in Fig. 3D, the concentrations of citric acid and succinic acid were significantly higher in cacao callus than in seeds. This metabolic profile aligns well with the physiological state of actively proliferating callus during *in vitro* culture.

Phenolic acids are another category of organic acids that originate from phenylpropanoid pathway with diverse structures and bioactivities (Herrmann et al. 1989). These compounds serve as precursors for various secondary metabolites, including benzoic acids, flavonoids, lignins, stilbenes, and coumarins (Deng. 2017). Phenolic acids exhibited a broad spectrum of biological activities, including antioxidant, anti-inflammatory, and antimutagenic properties (Lee et al. 2005). As shown in Fig. 3D, the levels of ferulic acid, coumaric acid-4-O-glucoside, sinapic acid, and protocatechuic acid were significantly higher in cacao callus than in seeds. Accumulated phenolic acids could be used for biosynthesis of various of secondary metabolites which would be helpful in bolstering stress resistance under hostile environment.

Fatty acids such as stearic acid, palmitic acid, and myristic acid are essential components of cellular lipids and play critical roles in plant cell membrane functionality, plant growth and development (Ariizumi et al. 2011). Higher contents of these fatty acids were observed in cacao callus than in seeds which consisted with their growth status. Callus were underwent a more vigorous cell proliferation thus requires more fatty acids to facilitate rapid membrane assembly and sustained callus growth.

3.3.2.4 Flavonoids and lignans

Flavonoids, originated from the phenylpropanoid pathway, represent one of the largest classes of plant secondary metabolites (Liu et al. 2021). Research has demonstrated that flavonoids played multifaceted roles in plant physiology and participated in defense responses against biotic and abiotic stresses (Mittler et al. 2022). Cacao is particularly rich in diverse flavonoids, including flavan-3-ols, flavonols, and their derivatives. These compounds have attracted significant research interest due to their remarkable physiological activities, such as potent antioxidant effects, cardiovascular protective properties, and potential cholesterol-regulating functions (Harrington et al. 2011). As shown in Fig. 3E, the relative contents of core flavonoids such as dihydrokaempferol, quercetin, kaempferol, and luteolin, as well as flavanols such as epigallocatechin, afzelechin, epicatechin, and procyanidin B1, were significantly higher in cacao seeds than in callus tissues. In contrast, the relative levels of major flavonoid glycosides such as isorhamnetin-3-O-glucoside, apigenin-7-O-glucoside, cyanidin-3-O-glucoside, luteolin-7-O-glucoside, and lyoniresinol-9-O-glucoside were markedly higher in cacao callus compared to seeds.

Lignans are phenolic oligomers in plants acting as plants' chemical arsenal or phytoalexins (Saleem et al. 2005). Pharmacological studies of lignans displayed promising metabolic regulatory effects (Wang et al. 2024, Huang et al. 2024). Comparative analysis revealed significant higher levels of lignans such as grandisin, syringaresinol, acanthoside B, lyoniresinol-9-O-glucoside, and liriodendrin in cacao callus than in seeds (Fig. 3E). Higher contents of lignans and other bioactive metabolites in cacao callus making it potential materials for bioactive compounds production.

3.3.2.5 Alkaloids and terpenoids

Alkaloids represent a major class of nitrogen-containing secondary metabolites that are widely distributed in plants. These compounds serve dual roles as defensive toxins and pharmacological active agents (Muro-Villanueva et al. 2023), exhibiting notable antioxidant, anticancer, and neuroprotective properties (Ng et al. 2012). In cacao, the predominant alkaloids were theobromine and caffeine, both of which showed higher accumulation in cacao seeds than in cacao callus (Fig. 3G).

Terpenoids with structure diversity underpins multiple applications in pharmaceutical, agricultural, and material science (Brahmkshatriya et al. 2013, Jaeger et al. 2016, Martin-Smith et al. 1963, Yazaki et al. 2017). Detected terpenoids such as chrysanthemol, oleanolic acid, spinasaponin, and curcumenol were more abundant in callus tissues than in seeds (Fig. 3G).

4 Conclusion

In the present study, cacao (Forastero) callus was induced from hypocotyl successfully. Mass spectrometry-based profiling allowed systematic chemical comparison of cacao callus and beans. Chemical comparison revealed that all the 182 annotated metabolites were detected in both cacao callus and cacao beans with most of their contents varied significantly. The obtained stable cacao callus and the revealing of its comprehensive metabolomic profile would provide continuous materials and theoretical basis for cacao breeding research and biosynthesis of bioactive metabolites which demonstrating significant implications for promoting sustainable development of the cacao industry.

Statements and Declarations

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research was funded by Natural Science Foundation of Hebei Province (C2020423047, H2025423024); Research Foundation for The Excellent Youth Scholars of Hebei University of Chinese Medicine (YQ2019008); Science Research Project of Hebei Education Department (QN2024291); Special Project of Basic Scientific Research of Provincial Universities (TDZR2025004); The earmarked fund for Hebei Agriculture Research System (HARS) (HBCT2024110201 and HBCT2024110205).

Authorship contribution statement

X.W. acquired the data and wrote the main manuscript. J.C. analyzed the data and prepared figures and tables. G.Y. and Q.Z. revised the manuscript critically. Y.Z. and L.W. supervised the project, reviewed and edited the manuscript, acquired funding to support the project. All authors reviewed the manuscript.

References

1. Afoufa-Bastien, D., Medici, A., Jeauffre, J., Coutos-Thevenot, P., Lemoine, R., Atanassova, R. and Laloi, M. (2010) The *Vitis vinifera* sugar transporter gene family: phylogenetic overview and macroarray expression profiling. *BMC Plant Biol.* 10, 245. <https://doi.org/10.1186/1471-2229-10-245>
2. Alotaibi, R. F., AlTilasi, H. H., Al-Mutairi, A. M., & Alharbi, H. S. (2024). Chromatographic and spectroscopic methods for the detection of cocoa butter in cocoa and its derivatives: A review. *Heliyon*, 10(11), e31467. <https://doi.org/10.1016/j.heliyon.2024.e31467>
3. Amponsah-Doku, B., Daymond, A., Robinson, S., Atuah, L., & Sizmur, T. (2022). Improving soil health and closing the yield gap of cocoa production in Ghana – A review. *Scientific African*, 15, e01075. <https://doi.org/10.1016/j.sciaf.2021.e01075v>
4. Azizah, O., Ismail, A., Abdul Ghani, N., & Adenan, I. (2007). Antioxidant capacity and phenolic content of cocoa beans. *Food Chemistry*, 100(4), 1523-1530. <https://doi.org/10.1016/j.foodchem.2005.12.021>
5. Brahmikshatriya, P. P., & Brahmikshatriya, P. S. (2013). Terpenes: Chemistry, biological role, and therapeutic applications. In K. Ramawat & J. M. Mérillon (Eds.), *Natural products* (pp. 2665-2691). Springer. https://doi.org/10.1007/978-3-642-22144-6_120
6. Chen, J., Guo, L., Yang, G., Yang, A., Zheng, Y., & Wang, L. (2022). Metabolomic profiling of developing perilla leaves reveals the best harvest time. *Frontiers in Plant Science*, 13, 989755. <https://doi.org/10.3389/fpls.2022.989755>
7. Darmawan, D., & Mutalib, A. (2024). Evaluation of environmental impact on cocoa production and processing under life cycle assessment method: From beans to liquor. *Environmental Advances*, 15, 100481. <https://doi.org/10.1016/j.envadv.2023.100481>
8. Deng, Y., and Lu, S. (2017). Biosynthesis and regulation of phenylpropanoids in plants. *Crit. Rev. Plant Sci.* 36, 257–290. <https://doi.org/10.1080/07352689.2017.1402852>
9. Edo, G. I., Samuel, P. O., Oloni, G. O., Ezekiel, G. O., Onoharigho, F. O., Oghenegueke, O. O., Nwachukwu, S. C., Rapheal, O. A., Ajokpaoghene, M. O., Okolie, M. C., Ajakaye, R. S., Ndudi, W., & Igbo, P. C. (2023). Review on the biological and bioactive components of cocoa (*Theobroma cacao*). Insight on food, health and nutrition. *Natural Resources for Human Health*, 3(4), 426–448. <https://doi.org/10.53365/nrfhh/174302>
10. Efferth, T. (2019). Biotechnology applications of plant callus cultures. *Engineering*, 5(1), 50–59. <https://doi.org/10.1016/j.eng.2018.11.006>
11. Gonda, S. (2020). Special Issue: Plant Metabolomics. *Metabolites*, 10(11), 467. <https://doi.org/10.3390/metabo10110467>
12. Gu, F., Tan, L., Wu, H., Fang, Y., Xu, F., Chu, Z., & Wang, Q. (2013). Comparison of cocoa beans from China, Indonesia and Papua New Guinea. *Foods*, 2 (2), 183–197. <https://doi.org/10.3390/foods9020183>
13. Harrington, W. L. (2011). *The effects of roasting time and temperature on the antioxidant capacity of cocoa beans from Dominican Republic, Ecuador, Haiti, Indonesia, and Ivory Coast* (Publication No. 1494378) [Master's thesis, University of Tennessee]. ProQuest Dissertations and Theses Global. https://trace.tennessee.edu/utk_gradthes/976
14. Herrmann, K., & Nagel, C. W. (1989). Occurrence and content of hydroxycinnamic and hydroxybenzoic acid compounds in foods. *Critical Reviews in Food Science and Nutrition*, 28(4), 315-347. <https://doi.org/10.1080/10408398909527504>
15. Huang W., Huang G.P., Zhang L.X., Yu E.D., Yang W.K., Ye M., Zou S.Q., Ni L., He H.Q. Lignan-rich extract from *Cinnamomum camphora* leaf attenuates metabolic syndrome by modulating glycolipid metabolism and gut microbiota in T2DM mice. *Phytomedicine*. 135:156118. <https://doi.org/10.1016/j.phymed.2024.156118>.
16. Hurst, W. J., Tarka, S. M., Powis, T. G., Valdez, F., & Hester, T. R. (2002). Cacao usage by the earliest Maya civilization. *Nature*, 418 (6895), 289. <https://doi.org/10.1038/418289a>
17. Jaeger, R., & Cuny, E. (2016). Terpenoids with special pharmacological significance: A review. *Natural Products Communications*, 11(9). <https://doi.org/10.1177/1934578X1601100946>
18. Kongor, J. E., Owusu, M., & Oduro-Yeboah, C. (2024). Cocoa production in the 2020s: Challenges and solutions. *CABI Agriculture and Bioscience*, 5, 102. <https://doi.org/10.1186/s43170-024-00310-6>
19. Lee, Y. T., Don, M. J., Liao, C. H., Chiou, H. W., Chen, C. F., & Ho, L. K. (2005). Effects of phenolic acid esters and amides on stimulus-induced reactive oxygen species production in human neutrophils. *Clinica Chimica Acta*, 352(1–2), 135–141. <https://doi.org/10.1016/j.cccn.2004.09.010>
20. Liu, W., Feng, Y., Yu, S., Fan, Z., Li, X., Li, J., & Yin, H. (2021). The flavonoid biosynthesis network in plants. *International Journal of Molecular Sciences*, 22(23), 12824. <https://doi.org/10.3390/ijms222312824>
21. Lopes, U. V., Monteiro, W. R., Pires, J. L., Clement, D., Yamada, M. M., & Gramacho, K. P. (2011). Cacao breeding in Bahia, Brazil: Strategies and results. *Crop Breeding and Applied Biotechnology*, 11(spe), 73-81. <https://doi.org/10.1590/S1984-70332011000500011>
22. Martin-Smith, M., & Khatoon, T. (1963). Biological activity of the terpenoids and their derivatives. In E. Jucker (Ed.), *Fortschritte der Arzneimittelforschung / Progress in Drug Research* (Vol. 5, pp. 279–346). Birkhäuser. https://doi.org/10.1007/978-3-0348-7050-4_4
23. Meza-Sepúlveda, D. C., Hernandez-Urrea, C., & Valencia-Sanchez, H. (2025). Thermochemical characterization of cellulose extracted from the cocoa pod husk of *Theobroma cacao* L. *Industrial Crops and Products*, 226, 120579. <https://doi.org/10.1016/j.indcrop.2025.120579>
24. Mhlongo, M. I., Piater, L. A., Steenkamp, P. A., Labuschagne, N., & Dubery, I. A. (2020). Metabolic profiling of PGPR-treated tomato plants reveal priming-related adaptations of secondary metabolites and aromatic amino acids. *Metabolites*, 10(5), 210. <https://doi.org/10.3390/metabo10050210>

25. Mittler, R., Zandalinas, S. I., Fichman, Y., & Van Breusegem, F. (2022). Reactive oxygen species signalling in plant stress responses. *Nature Reviews Molecular Cell Biology*, 23(10), 663–679. <https://doi.org/10.1038/s41580-022-00499-2>
26. Muro-Villanueva, F., & Nett, R. S. (2023). Emerging functions within the enzyme families of plant alkaloid biosynthesis. *Phytochemistry Reviews*. Advance online publication. <https://doi.org/10.1007/s11101-023-09901-z>
27. Ng, T. B., Liu, J., Wong, J. H., Ye, X., Sze, S. C. W., Tong, Y., & Zhang, K. Y. (2012). Review of research on *Dendrobium*, a prized folk medicine. *Applied Microbiology and Biotechnology*, 93(5), 1795-1803. <https://doi.org/10.1007/s00253-011-3829-7>
28. Niehaus, T. D. (2021). Phosphatases are involved in modulating the TCA cycle in plants. *Molecular Plant*, 14 (7), 1036–1037. <https://doi.org/10.1016/j.molp.2021.05.019>
29. Nieves-Orduña, H. E., Krutovsky, K. V., & Gailing, O. (2023). Geographic distribution, conservation, and genomic resources of cacao *Theobroma cacao* L. *Crop Science*, 63, 1750–1778. <https://doi.org/10.1002/csc2.20959>
30. Orazc, J., Nebesny, E., & Żyżelewicz, D. (2015). Changes in the flavan-3-ols, anthocyanins, and flavanols composition of cocoa beans of different *Theobroma cacao* L. groups affected by roasting conditions. *European Food Research and Technology*, 241(5), 663-681. <https://doi.org/10.1007/s00217-015-2494-y>
31. Paisic-Ramirez, R., Hernández-Amasifuen, A. D., Sánchez-Aguilar, W. D., Corazon-Guivin, M. A., Bobadilla, L. G., Mansilla-Córdova, P. J., Caetano, A. C., Zuta, M. Z. S., & Guerrero-Abad, J. C. (2024). Effect of osmoregulatory on the secondary somatic embryogenesis of cocoa (*Theobroma cacao* L.). *Journal of Applied Biology and Biotechnology*, 12 (3), 177-183. <https://doi.org/10.7324/JABB.2024.180092>
32. Perkowski, M. C., & Warpeha, K. M. (2019). Phenylalanine roles in the seed-to-seedling stage: Not just an amino acid. *Plant Science*, 289, 110223. <https://doi.org/10.1016/j.plantsci.2019.110223>
33. Prasath, K. G., Alexpandi, R., Parasuraman, R., Pavithra, M., Ravi, A. V., & Pandian, S. K. (2021). Anti-inflammatory potential of myristic acid and palmitic acid synergism against systemic candidiasis in *Danio rerio* (Zebrafish). *Biomedicine & Pharmacotherapy*, 133, 111043. <https://doi.org/10.1016/j.biopha.2020.111043>
34. Pu, J., Liu, Y., Wu, H., Gui, S., Tian, L., Yu, Y., Song, X., Zhong, X., Chen, X., Chen, W., Zheng, P., Zhang, H., Gong, X., Liu, L., Wu, J., Wang, H., & Xie, P. (2025). Characterizing metabolomic and proteomic changes in depression: A systematic analysis. *Molecular Psychiatry*. Advance online publication. <https://doi.org/10.1038/s41380-025-02919-z>
35. Rocha, G. H. A. M., Almeida, M. C., Silva, L. L. P., Flores, I. S., Castiglioni, G. L., Oliveira, T. F., & Pereira, J. (2025). Food-related properties and composition of cocoa honey (*Theobroma cacao* L.): An integrated investigation. *Food Research International*, 202 , 115694. <https://doi.org/10.1016/j.foodres.2025.115694>
36. Saleem M., Kim H.J., Ali M.S., Lee Y.S. (2005) An update on bioactive plant lignans. *Nat Prod Rep*. 22(6):696-716. <https://doi.org/10.1039/b514045p>.
37. Sun, L., Gong, M., Lv, X., Huang, Y., Wang, Z., & Li, Y. (2020). Current advance in biological production of short-chain organic acid. *Applied Microbiology and Biotechnology*, 104(21), 9109–9124. <https://doi.org/10.1007/s00253-020-10917-0>
38. Szűcs, Z., Cziáky, Z., Kiss-Szikszai, A., Sinka, L. T., Vasas, G., & Gonda, S. (2019). Comparative metabolomics of *Tilia platyphyllos* Scop. bracts during phenological development. *Phytochemistry*, 167 , 112084. <https://doi.org/10.1016/j.phytochem.2019.112084>
39. Wang, G., Liu, Y., Gao, Z., Li, H., & Wang, J. (2023). Effects of amino acids on callus proliferation and somatic embryogenesis in *Litchi chinensis* cv. 'Feizixiao'. *Horticulturae*, 9 (12), 1311. <https://doi.org/10.3390/horticulturae9121311>
40. Wang S., Hu Y., Liu B., Li Y., Wang M., Sun Q. (2024) Lignan intake and type 2 diabetes incidence among US Men and Women. *JAMA Netw Open*. 7(8):e2426367. <https://doi.org/10.1001/jamanetworkopen.2024.26367>.
41. Yazaki, K., Arimura, G.-I., & Ohnishi, T. (2017). 'Hidden' terpenoids in plants: Their biosynthesis, localization and ecological roles. *Plant and Cell Physiology*, 58(9), 1615-1621. <https://doi.org/10.1093/pcp/pcx123>

Figures

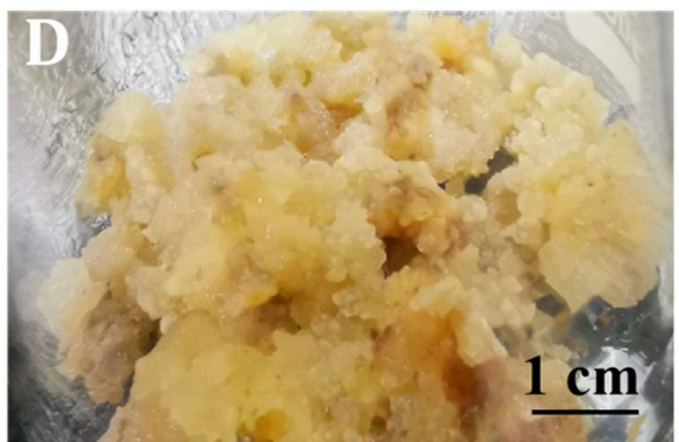
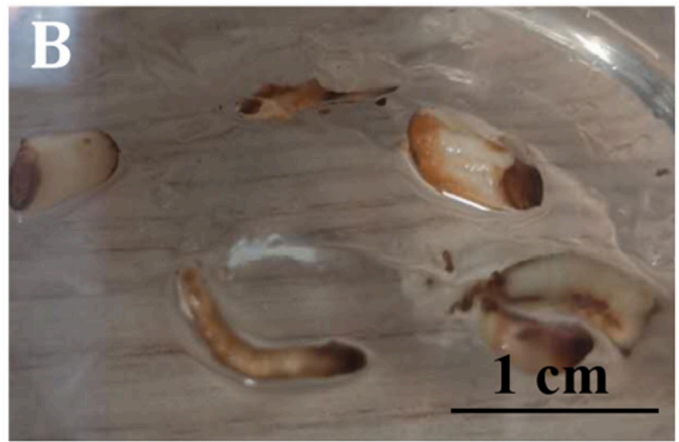


Figure 1

Induction and maintenance of cacao callus. (A) Germinated cacao seeds; (B) The cutted hypocotyl pieces; (C) Initiation of cacao callus; (D) Two weeks old callus

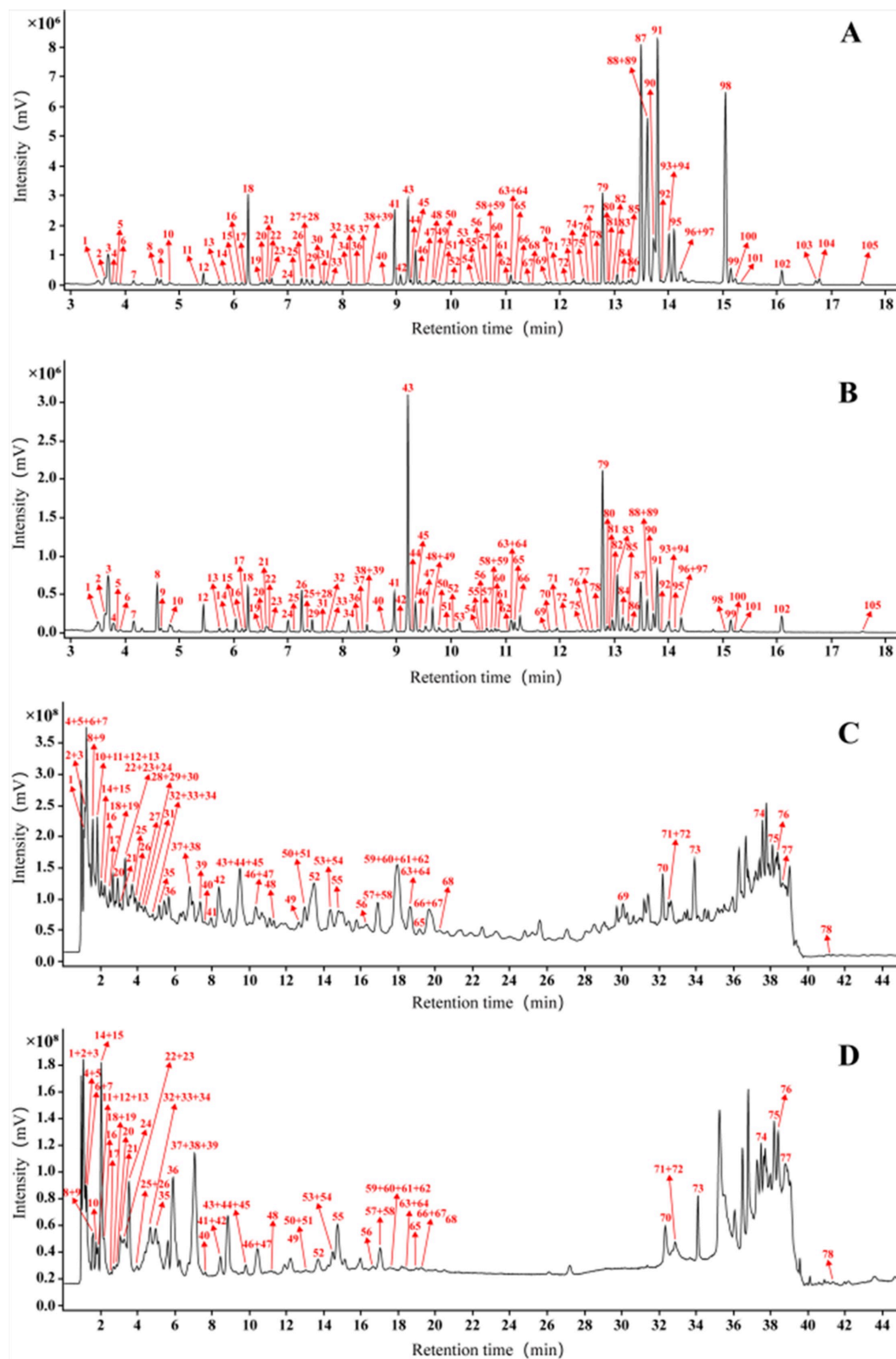


Figure 2

Figure 1. Typical total ion chromatograms (TICs) of cacao callus and cacao seeds obtained by GC-MS and LC-MS analyses. (A) TIC of cacao callus extract analyzed by GC-MS; (B) TIC profile of cacao seed extract analyzed by GC-MS; (C) TIC profile of cacao callus extract analyzed by LC-MS; (D) TIC profile of cacao seed extract analyzed by LC-MS.

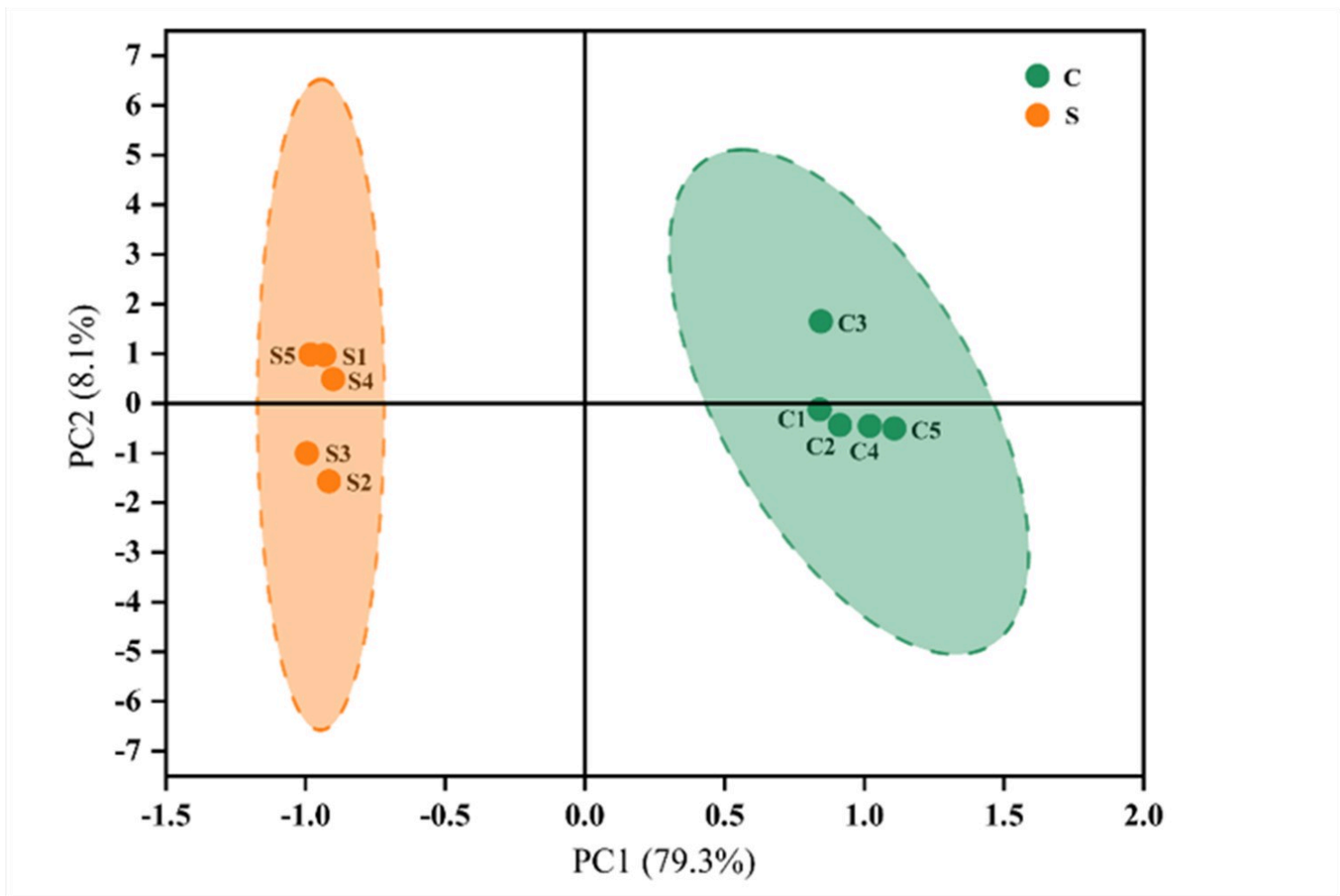


Figure 3

Figure 2. PCA plot of cacao callus and cacao seeds (C: cacao callus; S: cacao seed).

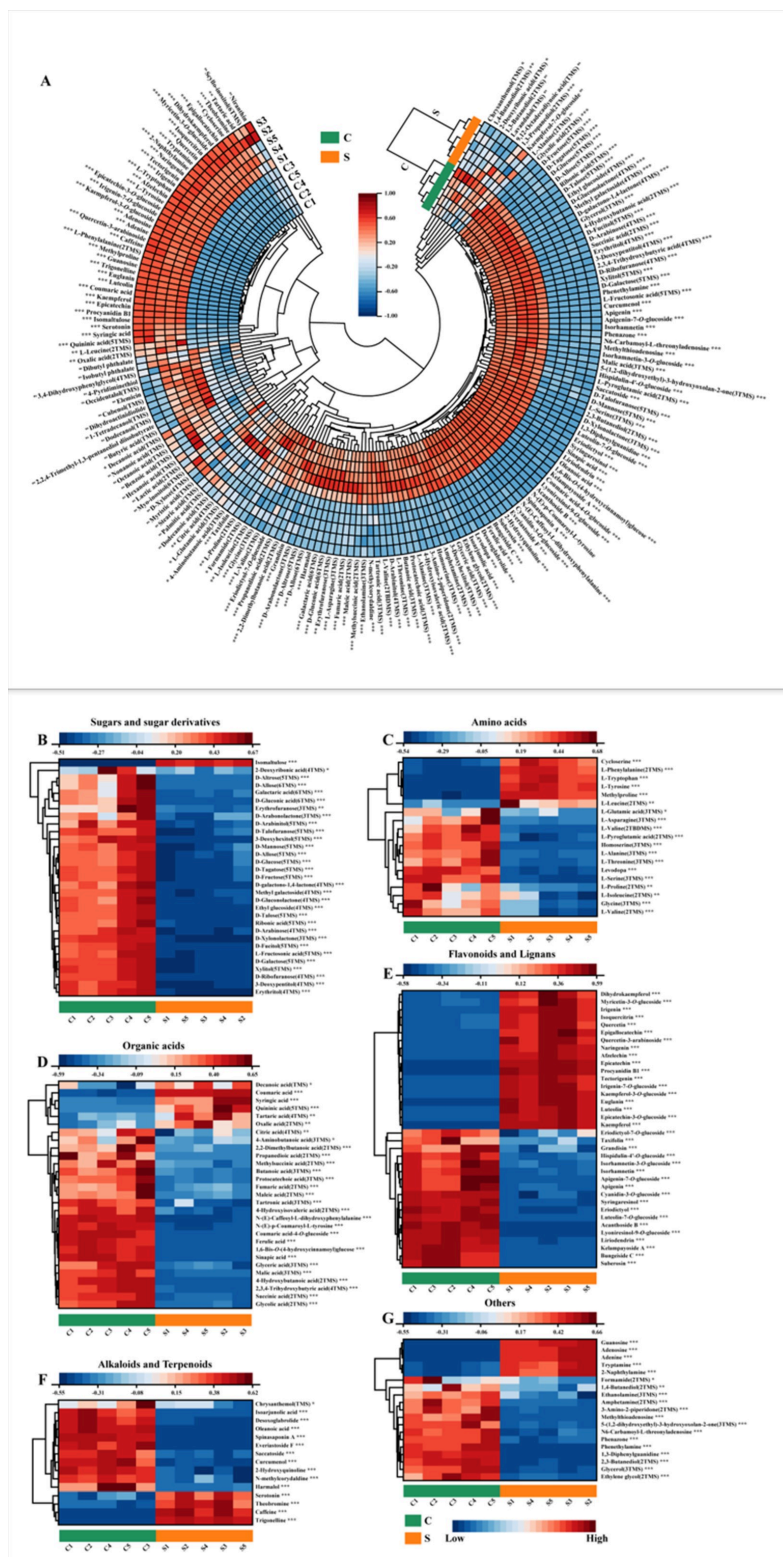


Figure 4

Figure 3. Clustering analysis of annotated metabolites in cacao callus and cacao seeds displaying with bi-hierarchical-clustering heat maps. Overview of all the annotated metabolites (A), sugars and sugar derivatives (B), amino acids (C), organic acids (D), flavonoids and lignans (E), alkaloids and terpenoids (F), and others (G) (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$; C: cacao callus; S: cacao seeds).