

Article

Not peer-reviewed version

---

# PlantMetSuite: A User-Friendly Web-Based Tool for Metabolomics Analysis and Visualisation

---

[Yu Liu](#) , [Hao-Zhuo Liu](#) , Ding-Kang Chen , Hong-Yun Zeng , [Yi-Li Chen](#) , [Nan Yao](#) \*

Posted Date: 10 July 2023

doi: 10.20944/preprints202307.0466.v1

Keywords: plant metabolomics; metabolite identification; data visualisation; omics data; bioinformatics tools



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Article*

# PlantMetSuite: A User-Friendly Web-Based Tool for Metabolomics Analysis and Visualisation

Yu Liu, Hao-Zhuo Liu, Ding-Kang Chen, Hong-Yun Zeng, Yi-Li Chen and Nan Yao \*

State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resource, School of Life Sciences, Sun Yat-sen University, Guangzhou 510275, China; liuy233@mail2.sysu.edu.cn (Y.L.); 1214800593@qq.com (H-Z.L.); cdkon@outlook.com (D-K.C.); zenghy7@mail2.sysu.edu.cn (H-Y.Z.); chenyl25@mail2.sysu.edu.cn (Y-L.C.)

\* Correspondence: yaonan@mail.sysu.edu.cn.

**Abstract:** The advancement of mass spectrometry technologies has revolutionised plant metabolomics research by enabling the acquisition of raw metabolomics data. However, the identification, analysis, and visualisation of these data require specialised tools. Existing solutions lack a dedicated plant-specific metabolite database and pose usability challenges. To address these limitations, we developed PlantMetSuite, a web-based tool for comprehensive metabolomics analysis and visualisation. PlantMetSuite encompasses interactive bioinformatics tools and databases specifically tailored for plant metabolomics data, facilitating upstream-to-downstream analysis in metabolomics and supporting integrative multi-omics investigations. PlantMetSuite can be accessed directly through a user's browser without the need for installation or programming skills. The tool is freely available at <https://plantmetsuite.verysgenome.com/> and will undergo regular updates and expansions to incorporate additional libraries and newly published metabolomics analysis methods. The tool's significance lies in empowering researchers with an accessible and customisable platform for unlocking plant metabolomics insights.

**Keywords:** plant metabolomics; metabolite identification; data visualisation; omics data; bioinformatics tools

## 1. Introduction

Metabonomics refers to the comprehensive analysis of endogenous metabolic small molecules in organisms, specifically studying the changes in metabolites within cells, tissues, or organs [1]. Plant metabolomics, a subset of metabonomics, focuses on identifying and quantifying metabolites involved in various cell regulatory processes, which comprise over 200,000 known plant metabolites, with each plant harbouring more than 5,000 metabolites [2]. This field finds applications in studying plant responses to biotic and abiotic stress stresses, ensuring food safety, nutrition, pharmaceuticals, and cosmetics, as well as supporting metabolic engineering and molecular breeding [3]. However, the study of plant metabolomics presents significant challenges due to their diverse chemical compositions and wide dynamic range. Several separation and detection technologies have been developed and applied to plant metabolomics research. Among them, liquid chromatography-mass spectrometry (LC-MS) has emerged as the dominant technique for metabolomics research due to its inherent advantages, including the absence of derivatisation requirements and broader coverage of metabolite classes compared to gas chromatography [4]. LC-MS combines the separation power of LC with the high resolution and sensitivity of MS, enabling the detection and identification of a wide range of metabolites, including polar and non-polar compounds, small molecules, and larger biomolecules. This versatility has solidified LC-MS as the preferred technique in plant metabolomics research, allowing for comprehensive metabolomic profiling and the study of complex biological systems.

While several methods and software tools have been developed for plant metabolomics research, no single metabolomic tool is applicable to the analysis of the entire metabolome due to the diverse nature of metabolites, significant variation in content across tissues and cells, and the complexity of physical and chemical properties of the metabolites prevent the existence of a single tool applicable to the entire metabolome [5]. Thus, data obtained from metabolomics experiments require pre-processing using software packages such as XCMS [6], MS-DIAL [7] and MZmine2 [8]. Further cleaning of the metabolomics data, including missing value imputation, data standardisation, and quality assessment, can be performed using MetaboAnalystR [9]. Subsequently, metabolite identification, a critical step in plant metabolomics, is accomplished by comprehensive analysis combining retention time (RT), mass-to-charge ratio and tandem mass spectrometry (MS/MS) matching [10]. XCMS is a web-based metabolomics analysis tool that allows users to upload raw data for data processing and statistical analysis, while also providing an R package for local implementation, enabling MS/MS-based metabolite peak identification and data statistics [6]. MS-DIAL, a user-friendly graphical interface-based metabolomics analysis software, supports targeted and untargeted LC-MS and GC-MS analysis. It not only comes with a wide collection of metabolomics and lipidomics spectral databases but also allows users to import their own libraries in MSP format for comprehensive searching and analysis of metabolites [7]. Following metabolite identification, univariate and multivariate analyses are conducted. Univariate analysis focuses on examining individual variables over a specific period. It involves the application of various statistical methods, such as the *t*-test, analysis of variance (ANOVA), and the Mann-Whitney U test. In contrast, multivariate analysis is more effective in exploring relationships between multiple variables and commonly employs principal component analysis (PCA), partial least squares discriminant analysis (PLSDA), and canonical correlation analysis (CCA). Notable open-source software packages, such as MetaboAnalystR [9] and MetaboDiff [11], facilitate both univariate and multivariate analyses in metabolomics studies.

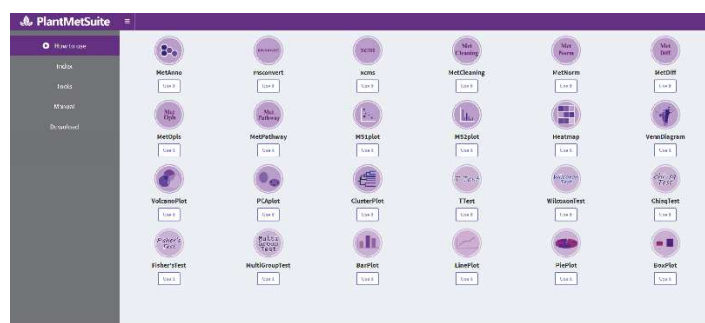
Despite the success of combining various metabolomics analysis tools for plant metabolome analysis, the field still lacks a comprehensive and streamlined analytical platform that can effectively decode plant metabolites. Currently, such a platform is still in its developmental stage, necessitating the use of disparate software and databases [12]. To address this gap, we have developed PlantMetSuite, an online analytical platform that provides comprehensive compound annotation by integrating information from major databases. PlantMetSuite supports the analysis of plant metabolomics data, facilitates the accurate identification of plant metabolites and visualisation of data results, and simplifies routine analysis of plant metabolomics without the need for coding expertise. This platform is now accessible at <https://plantmetsuite.verygenome.com/>.

## 2. Results

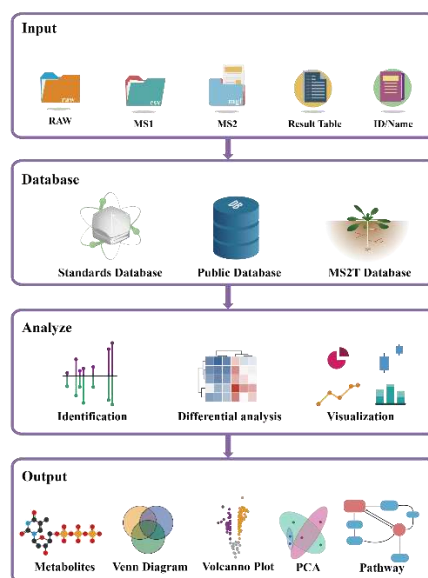
### 2.1. Workflow for the metabolomics analysis in PlantMetSuite

PlantMetSuite offers a comprehensive suite of visually-oriented applications specifically designed for metabolomics analysis (Figure 1a), providing a wide array of functionalities. Each application in PlantMetSuite is accompanied by test data, sample results, and video tutorials to ensure user-friendliness and accessibility. These resources are designed to facilitate ease of use for all users. A detailed overview of the capabilities of each analysis application is presented in Supplementary Table S1. PlantMetSuite adopts a streamlined data processing workflow comprising four primary steps (Figure 1b). To begin using PlantMetSuite, users are initially prompted to upload data in several formats: (1) Major MS vendor formats, including those from AB Sciex (.Wiff), Thermo Fisher Scientific (.RAW), Bruker Daltonics (.D), Agilent Technologies (.D), and Waters (.RAW). These data formats necessitate conversion to mzXML and mgf formats using the msconvert application within PlantMetSuite or the ProteoWizard (version 3.0.22167) [13] installed locally by the user. (2) Pre-processed data files, including MS1 peak table files (in .CSV format or .XLS format) and MS2 spectrogram files (in .MSP format or .MGF format). MS1 files can be obtained from the xcms application in PlantMetSuite, XCMS Online [14], the locally installed XCMS by the user [6], MS-DIAL

[7], or MZmine2 [8], while MS2 files can be derived from raw data conversion or MS-DIAL analysis. (3) Metabolite content tables suitable for analysis with various analytical software, including the MetAnno application within PlantMetSuite, XCMS Online [7], and MS-DIAL [7]. (4) Metabolite ID in the databases (KEGG, Metlin, Pubchem, etc.) or the metabolite name. Furthermore, users are required to select the preferred database for metabolite identification. PlantMetSuite encompasses three database types: (1) The internally constructed standards database, incorporating  $m/z$ , MS/MS, and RT information derived from standards. (2) Public databases such as MoNA (<http://mona.fiehnlab.ucdavis.edu>), MassBank [15], HMDB [16], KEGG [17], Respect [18], and CASMI [19]. (3) The LC-MS-based MS/MS spectral tag database, encompassing identified and unidentified metabolic profiles of *Arabidopsis thaliana* cauline leaves, rosette leaves, stems, and inflorescence parts. This database proves valuable for investigating tissue specificity and novel functions of secondary metabolites in plants [20]. Subsequently, users selectively employ diverse applications as per their requirements for data analysis, including data pre-processing, data cleaning, metabolite identification, differential metabolite analysis, pathway analysis, and data visualisation. Finally, upon completion of the analysis, tabular results and vector graphics can be exported. Additionally, if provided, the results are conveyed to the user's email address. Raw data utilised for the analysis are securely retained on the server for 24 h only, after which they are entirely purged to ensure data security.



(a)



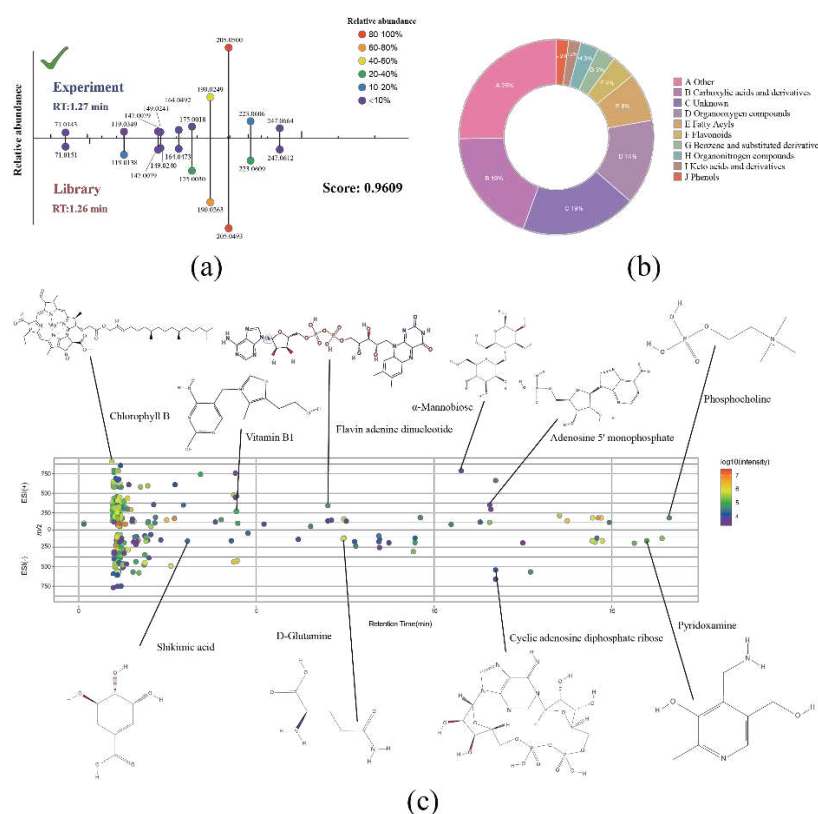
(b)

**Figure 1.** PlantMetSuite user interface design and functionalities. (a) Screenshot of the PlantMetSuite interface. Users can assess an instructional video by clicking on the "How to use" tab. The Index tab leads to the PlantMetSuite home page, the Tools tab provides direct access to PlantMetSuite applications, the Manual tab allows users to consult the help manual, and the Download tab enables downloading the data already analysed. (b) PlantMetSuite data processing workflow. The analysis process consists of four steps: data entry, database selection, analysis tool selection, and result export.

In the data entry step, users can input MS1 files, MS2 files, quantitative tables, and metabolite names or IDs. The database selection step allows users to choose from standards, public, and MS2T databases. The analysis tool selection step enables the selection of tools for metabolite identification analysis, difference analysis, and visualisation analysis. In the result export step, users can export various outputs such as metabolite identification results, Venn diagrams, volcano diagrams, PCA results, and pathway enrichment results.

## 2.2. Plant metabolite annotation based on a high-quality plant-specific library

Metabolite annotation in metabolomics heavily relies on robust scoring of spectral matches between MS/MS spectra and library spectra. The fragmentation patterns observed during MS/MS are greatly influenced by the instrument platform and impact energy parameters utilized. Therefore, the construction of a reliable in-house standards library becomes indispensable to enhance the precision of metabolite annotation. A detailed methodology outlining the curation of our high-quality plant-specific MS/MS spectral library is provided in the Methods section. Our in-house MS/MS spectral library includes a total of 1,122 plant-specific metabolites, thus offering comprehensive coverage of plant-derived metabolites beyond what is available in public spectral libraries. To exemplify the identification process, we presented a detailed explanation using the identification of 1-O-beta-D-Glucopyranosyl sinapate as an example (Figure 2a). In the negative ionisation mode, peak M386T77 exhibited a mass-to-charge ratio ( $m/z$ ) of 385.1118 and RT of 1.27 min. Employing a 25 ppm cutoff threshold, the MS1 identification confidently assigned this peak to 1-O-beta-D-Glucopyranosyl sinapate. Fragment alignment between the experimental and standard spectra was performed using a 10 ppm threshold, where the aligned fragments were visually represented as dots, with their relative abundance distinguished by various colours. Subsequent scoring based on Eq.1 yielded an MS/MS score of 0.9606. RT was assessed using the RT evaluation formula (Eq. 2), resulting in an RT score of 1. Integrating the scores using the score integration formula (Eq.3), along with user-defined weights, yielded a final score of 0.9688. Consequently, the peak was conclusively identified as 1-O-beta-D-Glucopyranosyl sinapate. The standards library incorporated into PlantMetSuite covers a wide range of plant metabolites, including carboxylic acids and derivatives (19%), organooxygen compounds (14%), fatty acyls (8%), flavonoids (4%), benzene and substituted derivatives (3%), organonitrogen compounds (3%), keto acids and derivatives (2%), and phenols (2%) (Figure 2b). Leveraging both positive and negative modes, our platform allows for the simultaneous quantification of numerous metabolites, reaching hundreds in a single analysis. Furthermore, representative metabolites from various classes are visually highlighted on the graph, with the colour-coded dots indicating their relative abundance (Figure 2c).



**Figure 2.** Metabolite identification process and capabilities of PlantMetSuite. (a) Illustration of the metabolite identification process using the example of 1-O-β-D-glucopyranosyl mustardate. The top part shows the spectrum from the experiment, and the bottom part displays the spectrum from the database. Colours represent the relative abundance of fragments, and RT and the final scoring are marked on the image. (b) Pie charts depicting mass proportions of different metabolite classes in the standard library. (c) PlantMetSuite enables the identification of hundreds of different plant metabolites in a single experiment.

### 2.3. Metabolomics analysis of wild-type *Arabidopsis thaliana* and ceramide kinase mutant *acd5*

Metabonomic profiling was conducted on wild-type *Arabidopsis thaliana* and the ceramide kinase mutant *acd5* using the aforementioned workflow. A total of 4,831 peaks were detected in the positive mode, while 5,459 peaks were detected in the negative mode. Among these peaks, 518 metabolites were identified, with 313 metabolites in the positive mode and 205 in the negative mode (Figure S1). PCA was employed, and the resulting scatter plot demonstrated the well-grouped biological replicates, indicating high experimental stability. Notably, a clear separation between the *acd5* samples and wild-type (Col-0) samples was observed in the plots, suggesting pronounced metabolomic differences between the mutant and wild-type plants (Figure S2a). These findings were supported by the results of the partial least squares-discriminant analysis (PLS-DA) (Figure S2b, S2c). Univariate analysis revealed 35 differentially abundant metabolites out of the 518 identified metabolites, with statistical thresholds set at a P-value < 0.05 and an absolute log2foldchange > 1 (Figure S2d). The volcano plot illustrated the distribution of these significantly different metabolites (Figure 3b). To identify and characterise the dysregulated pathways, pathway enrichment analysis was performed. For each pathway, the hypergeometric test was used to calculate the corresponding P-value, with pathways exhibiting P-values below 0.05 (default threshold) considered enriched. A total of ten dysregulated pathways were detected (Figure 3c), primarily associated with amino acid and sugar metabolism. Furthermore, several key differential metabolites underwent further re-examination (Figure 3d).



plant metabolomics data. In this investigation, metabolite identification adhered to the rigorous standards outlined by the Metabolomics Standards Initiative (MSI), employing either a standards database for level 1 identification or a public database or MS2T database for level 2 identification [22, 23]. PlantMetSuite surpasses the prescribed orthogonality parameters established by the MSI through its integration of precise mass, RT, and MS/MS fragment comparison. Our approach demonstrated comparable accuracy in identification and quantification to conventional methodologies while presenting distinctive attributes. These attributes include a high-quality standards database, supplemented by an expansive public database that will be expanded in forthcoming updates. Moreover, PlantMetSuite offers a comprehensive suite of analytical procedures, enhancing versatility in analysing diverse plant metabolism data, along with supplementary visualization tools. Consequently, PlantMetSuite emerges as an indispensable tool for the analysis of metabolomics data in plants, catering to researchers with varying levels of bioinformatics expertise. PlantMetSuite is available at <https://plantmetsuite.verygenome.com/>.

## 4. Materials and Methods

### 4.1. Construction of plant standards spectral library

The construction of the standards spectral library in PlantMetSuite was based on a modified method derived from the NIST [24]. Employing Sciex TripleTOF 5600 instruments, MS2 spectra were acquired using commercially available metabolite standards. To acquire the MS2 spectra for each metabolite, a flow injection method was employed, utilizing target product ion scans. To ensure a robust representation, at least 16 MS2 spectra were selected for each metabolite. Clusters of MS2 spectra exhibiting high similarity ( $DP > 0.8$ ) were carefully aggregated to generate a consensus MS2 spectrum. This acquisition process involved employing varying collision energies, including 10, 20, 30, 40, 50, 60, 70, and  $35 \pm 15$  eV. MS1 and RT data were obtained using the same LC-MS method described below. The current iteration of the plant standards spectral library in PlantMetSuite comprises a collection of 1,122 metabolites, consisting of 1,122 for the positive mode and 1,119 for the negative mode.

### 4.2. Annotation algorithms for metabolites

Metabolite annotation in PlantMetSuite follows a systematic, algorithmic approach. Initially, the MS2 spectra were matched to the corresponding MS1 peaks based on their RT ( $\pm 8$  s) and  $m/z$  ( $\pm 20$  ppm) using the imported MS2 data and MS1 peak table. In cases where multiple MS2 spectra corresponded to a single MS1 peak, the algorithm selected the most abundant MS2 spectrum. For quantifying the abundance of the MS2 spectrum, PlantMetSuite calculated the cumulative intensity of the top ten fragment ions, ensuring an accurate representation. When using MS-DIAL for peak detection and MS2 data output in the .MSP format, PlantMetSuite bypassed the MS2-MS1 matching step, thereby expediting the process. The resulting MS1/MS2 pairs are then compared to our in-house standard spectral library. This facilitates metabolite annotation. By default, a tolerance of  $\pm 25$  ppm was employed for the MS1  $m/z$  value. A tailored DP function was applied to assess the similarity between the experimental and standards spectra, yielding a similarity score ranging from 0 to 1 (Eq. 1). This score provides a quantitative measure of the degree of alignment, ranging from none to perfect alignment. For consistency, fragment ion intensities in MS2 spectra were rescaled, normalising the highest fragment ion to 1.

$$\text{score.MSMS} = \frac{\sum ([\text{intensity.library}][\text{mz.library}])([\text{intensity.experiment}][\text{mz.experiment}])}{\sqrt{\sum ([\text{intensity.library}][\text{mz.library}])^2([\text{intensity.experiment}][\text{mz.experiment}])^2}} \quad (1)$$

Where score.MSMS represents the score of MS2, and intensity.library and mz.library correspond to the abundance and the  $m/z$  of the product ions in the data-base, respectively. Similarly, intensity.experiment and mz.experiment represent the abundance and  $m/z$  of the product ions in the experimental data. score.MSMS is de-rived from the combination of forward and reverse matches.

A trapezoidal function was used to evaluate the RT match score in PlantMetSuite (Eq. 2). This particular function has gained considerable acceptance in assessing RT in gas chromatography analysis [25].

$$\text{score.RT} = \begin{cases} 1, & |RT.\text{experiment} - RT.\text{library}| < RT.\text{Tol.min} \\ 1 - \frac{|RT.\text{experiment} - RT.\text{library}| - RT.\text{Tol.min}}{RT.\text{Tol.max} - RT.\text{Tol.min}}, & RT.\text{Tol.min} \leq |RT.\text{experiment} - RT.\text{library}| \leq RT.\text{Tol.max} \\ 0, & |RT.\text{experiment} - RT.\text{library}| > RT.\text{Tol.max} \end{cases} \quad (2)$$

Where score.RT represents the RT score, RT.library represents the RT in the database, RT.experiment represents RT in the experiment, RT.Tol.min represents the minimum RT tolerance, RT.Tol.max represents the maximum RT tolerance.

An integrated score, calculated by combining the scores from different dimensions, was determined using a linear weighting function as described by Eq. 3:

$$\text{score.integrate} = \text{Weight.MSMS} \times \text{score.MSMS} + \text{Weight.RT} \times \text{score.RT} \quad (3)$$

Where score.RT and score.MSMS are the scores for RT and MS/MS matches calculated in the previous steps. Weight.RT and Weight.MSMS are the weights assigned by the user for RT and MS/MS match scores. Only metabolite candidates with integrated scores higher than the user-defined threshold were considered valid match results.

#### 4.3. Software implementation of PlantMetSuite

The interface of PlantMetSuite was implemented using Shiny Server (Version 1.5.20.1002). The software tools in PlantMetSuite were developed using a combination of programming languages, including R, Perl, and Python. The web server hosting the PlantMetSuite platform is currently deployed on a high-performance Linux server boasting 32 cores (3.2GHz CPU) and 256 GB RAM, ensuring optimal performance. Rigorous testing was conducted using the latest stable versions of web browsers, namely Chrome, Firefox, and Microsoft Edge.

#### 4.4. Reagents, plant culture, and sample preparation

All chemical standards were purchased from MedChemExpress (New Jersey, USA), and the list of these standards can be found in Supplementary Table S2. Acetonitrile (ACN) and methanol (MeOH) of LC-MS grade were procured from Honeywell (Muskegon, MI, USA). Ammonium hydroxide (NH<sub>4</sub>OH) and ammonium acetate (NH<sub>4</sub>OAc) were obtained from Sigma-Aldrich (St. Louis, USA).

Wild-type *Arabidopsis thaliana* (Col-0) and *acd5* mutants [26] were stratified at 4 °C for 3 d and subsequently cultivated in a greenhouse under controlled conditions of 22 °C, 50% relative humidity, and a photoperiod of 16 hours light and 8 hours dark. The plants were illuminated with PAK lamps (PAK090311), providing a light intensity of 4800-6000 lux. Mature leaves of *Arabidopsis thaliana* were harvested and promptly placed in microcentrifuge tubes, which were immediately snap-frozen in liquid nitrogen and stored at -80 °C until further metabolite extraction. Prior to extraction, plant samples were thawed on ice and homogenized using a TissueLyser 192 (Jingxin Technologies) with beads (3mm in diameter) and 200µL of water (H<sub>2</sub>O) for homogenization. Subsequently, a mixture of ACN:MeOH (1:1, v/v; 800µL) was added to the samples and vortexed for 30 seconds. The samples were then rapidly frozen for 1 minute in liquid nitrogen, followed by thawing on ice. This vortexing, freezing, and thawing cycle was repeated three times. To precipitate proteins, the samples were incubated at -20 °C for 1 h, followed by centrifugation at 15,800g for 15 min at 4 °C. The collected supernatant was then subjected to solvent evaporation using a vacuum concentrator (Labconco, USA) until complete dryness was achieved. The resulting dried extracts were reconstituted by adding a mixture of ACN:H<sub>2</sub>O (1:1, v/v; 100 µL) and sonication (50 Hz, 4 °C) for 10 min. The solutions were

centrifuged at 15,800 g for 5 min at 4 °C to remove any remaining insoluble debris. Prior to LC-MS/MS analysis, the clarified supernatants were transferred to glass HPLC vials and stored at - 80°C.

#### 4.5. LC-MS/MS analysis of wild-type *Arabidopsis thaliana* and ceramide kinase mutant *acd5*

Metabolomics data for *Arabidopsis thaliana* samples were collected using a quadrupole time-of-flight mass spectrometer (TripleTOF 5600, AB SCIEX, USA) coupled with a UHPLC system (Shimadzu 30A; SHIMADZU Corporation, Japan). LC separation was achieved using a Waters ACQUITYUPLC BEHAmide column (particle size, 1.7 µm; 100 mm (length) x 2.1 mm (inner diameter)), with the column temperature maintained at 25 °C. Mobile phase A consisted of 25 mM ammonium hydroxide (NH<sub>4</sub>OH) + 25 mM ammonium acetate (NH<sub>4</sub>OAc) in water. Mobile phase B consisted of ACN, used for both positive (ESI+) and negative (ESI-) modes. The gradient was configured as follows: with a flow rate of 0.3 mL/min, the elution conditions were set as follows: 0-1 minute with 95% solvent B, 1-14 minutes with a linear decrease from 95% solvent B to 65% solvent B, 14-16 minutes with a linear decrease from 65% solvent B to 40% solvent B, 16-18 minutes with a constant 40% solvent B, 18-18.1 minutes with a linear increase from 40% solvent B to 95% solvent B, and finally, 18.1-23 minutes with a constant 95% solvent B. The injection volume was fixed at 2 µL, and all samples were collected randomly for data acquisition. Information-Dependent Acquisition (IDA) was employed for data collection. The source parameters were set as follows: gas 1 (GAS1) at 60 psi, gas 2 (GAS2) at 60 psi, curtain gas (CUR) at 30 psi, temperature (TEM) at 600°C, and declustering potential (DP) at either 60 V in positive mode or -60 V in negative mode.; and ion spray voltage float (ISVF), 5,500 or -4,000 V in positive or negative mode, respectively. The TOF MS scan parameters covered a mass range of 60-1200 Da, with an accumulation time of 200 ms and dynamic background subtraction enabled. For product ion scans the mass range was set to 25-1200 Da with an accumulation time of 50 ms and a collision energy of 30 or -30 V in positive or negative mode. Additional parameters were collision energy spread (0), resolution (units), charge state (1 to 1), intensity (100 cps), isotopic exclusion within 4 Da, mass tolerance (10 ppm), maximum number of candidate ions to monitor per cycle (6) and exclusion of previous target ions for 4 s after two occurrences.

#### 4.6. Data analysis

An MS1 peak table (.csv format) and MS2 data files (.mgf or .msp format) are required for import into PlantMetSuite. The MS1 peak table lists the metabolic peaks together with the annotated *m/z* and RT. This table was generated by extracting data from the raw MS files using popular peak-picking software like XCMS and MS-DIAL. The original MS files were converted into MS2 data files (.mgf format) using ProteoWizard (version 3.0.6150). Alternatively, users have the option to execute all aforementioned steps within the PlantMetSuite framework. For data analysis using PlantMetSuite, users upload MS1 peak tables (.xls) and MS2 data (.mgf) were uploaded to PlantMet Suite for data analysis. Metabolite identification is performed based on user-defined parameters, encompassing the following specifications: Polarity: positive/negative; instrument platform: Sciex; collision energy: 30; *m/z* tolerance: 20 ppm; RT match: checked; MS/MS match: checked; RT minimum tolerance: 30 s; RT maximum tolerance: 60 s; MS/MS mass range: 50-1250; MS/MS score cut-off: 0.6; RT score weight: 0.2; MS/MS score weight: 0.8; score cut-off: 0.6. Default parameters were used for all other analyses, including data pre-processing, data cleaning, metabolite identification, differential metabolite analysis, pathway analysis, and data visualisation.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Figure S1: Total ion chromatograms of ceramide kinase mutant *acd5* and wild type Col-0 in positive and negative ion modes; Figure S2: Multivariate statistical analysis of metabolome in *acd5* mutants and wild-type plants; Table S1: List of applications and description of PlantMetSuite; Table S2: Standards used to construct the standards library in PlantMetSuite.

**Author Contributions:** Conceptualization, Y.L. and N.Y.; methodology, Y.L.; software, Y.L. and H.L.; validation, Y.L., H.L., D.C. and Y.C.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and

editing, H.Z. and N.Y.; visualization, Y.L.; supervision, N.Y.; project administration, N.Y.; funding acquisition, N.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (32070196), the Natural Science Foundation of Guangdong Province (2019B1515120088, 2017A030311005), and Sun Yat-sen University (Project 33000-31143406).

**Data Availability Statement:** All the pertinent data can be located in the article or supplementary materials and are also accessible through the PlantMetSuite website (<https://plantmetsuite.verygenome.com/>). PlantMetSuite source code is available on Github at <https://github.com/YaonanLab/PlantMetSuite> for scientific research purposes. All other data in support of the results of this study are available at reasonable request from the corresponding author. .

**Acknowledgments:** We would like to thank Guangzhou VeryGenome Technology Co., Ltd for the use of computational resources and hosting PlantMetSuite.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Jordan, K. W.; Nordenstam, J.; Lauwers, G. Y.; Rothenberger, D. A.; Alavi, K.; Garwood, M.; Cheng, L. L., Metabolomic characterization of human rectal adenocarcinoma with intact tissue magnetic resonance spectroscopy. *Dis Colon Rectum* **2009**, *52* (3), 520-525.
- Dixon, R. A.; Strack, D., Phytochemistry meets genome analysis, and beyond. *Phytochemistry* **2003**, *62* (6), 815-6.
- Maia, M.; Figueiredo, A.; Cordeiro, C.; Sousa Silva, M., FT-ICR-MS-based metabolomics: A deep dive into plant metabolism. *Mass spectrometry reviews* **2021**.
- Lu, X.; Zhao, X.; Bai, C.; Zhao, C.; Lu, G.; Xu, G., LC-MS-based metabonomics analysis. *Journal of Chromatography B* **2008**, *866* (1-2), 64-76.
- Razzaq, A.; Sadia, B.; Raza, A.; Khalid Hameed, M.; Saleem, F., Metabolomics: a way forward for crop improvement. *Metabolites* **2019**, *9* (12), 303.
- Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G., XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Analytical Chemistry* **2006**, *78* (3), 779-787.
- Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M., MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature methods* **2015**, *12* (6), 523-526.
- Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC bioinformatics* **2010**, *11* (1), 1-11.
- Chong, J.; Xia, J., MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics* **2018**, *34* (24), 4313-4314.
- Kind, T.; Tsugawa, H.; Cajka, T.; Ma, Y.; Lai, Z.; Mehta, S. S.; Wohlgemuth, G.; Barupal, D. K.; Showalter, M. R.; Arita, M., Identification of small molecules using accurate mass MS/MS search. *Mass spectrometry reviews* **2018**, *37* (4), 513-532.
- Mock, A.; Warta, R.; Dettling, S.; Brors, B.; Jäger, D.; Herold-Mende, C., MetaboDiff: an R package for differential metabolomic analysis. *Bioinformatics* **2018**, *34* (19), 3417-3418.
- Piasecka, A.; Kachlicki, P.; Stobiecki, M., Analytical methods for detection of plant metabolomes changes in response to biotic and abiotic stresses. *International journal of molecular sciences* **2019**, *20* (2), 379.
- Adusumilli, R.; Mallick, P., Data conversion with proteowizard msconvert. *Methods in Molecular Biology* **2017**, *1550*, 339-368.
- Tautenhahn, R.; Patti, G. J.; Rinehart, D.; Siuzdak, G., XCMS Online: a web-based platform to process untargeted metabolomic data. *Analytical chemistry* **2012**, *84* (11), 5035-5039.
- Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K., MassBank: a public repository for sharing mass spectral data for life sciences. *Journal of mass spectrometry* **2010**, *45* (7), 703-714.
- Wishart, D. S.; Guo, A.; Oler, E.; Wang, F.; Anjum, A.; Peters, H.; Dizon, R.; Sayeeda, Z.; Tian, S.; Lee, B. L., HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Research* **2022**, *50* (D1), D622-D631.
- Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M., The KEGG resource for deciphering the genome. *Nucleic acids research* **2004**, *32* (suppl\_1), D277-D280.

18. Sawada, Y.; Nakabayashi, R.; Yamada, Y.; Suzuki, M.; Sato, M.; Sakata, A.; Akiyama, K.; Sakurai, T.; Matsuda, F.; Aoki, T., RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry* **2012**, *82*, 38-45.
19. Oberacher, H., Applying tandem mass spectral libraries for solving the critical assessment of small molecule identification (CASMI) LC/MS challenge 2012. *Metabolites* **2013**, *3* (2), 312-324.
20. Matsuda, F.; Yonekura-Sakakibara, K.; Niida, R.; Kuromori, T.; Shinozaki, K.; Saito, K., MS/MS spectral tag-based annotation of non-targeted profile of plant secondary metabolites. *The Plant Journal* **2009**, *57* (3), 555-577.
21. Kranner, I.; Birtić, S.; Anderson, K. M.; Pritchard, H. W., Glutathione half-cell reduction potential: A universal stress marker and modulator of programmed cell death? *Free Radical Biology and Medicine* **2006**, *40* (12), 2155-2165.
22. Salek, R. M.; Steinbeck, C.; Viant, M. R.; Goodacre, R.; Dunn, W. B. J. G., The role of reporting standards for metabolite annotation and identification in metabolomic studies. **2013**, *2* (1), 2047-217X-2-13.
23. Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W.-M.; Fiehn, O.; Goodacre, R.; Griffin, J. L. J. M., Proposed minimum reporting standards for chemical analysis: chemical analysis working group (CAWG) metabolomics standards initiative (MSI). **2007**, *3*, 211-221.
24. Yang, X.; Neta, P.; Stein, S. E. J. A. c., Quality control for building libraries from electrospray ionization tandem mass spectra. *Analytical chemistry* **2014**, *86* (13), 6393-6400.
25. Smith, D. H.; Achenbach, M.; Yeager, W. J.; Anderson, P. J.; Fitch, W. L.; Rindfleisch, T. C., Quantitative comparison of combined gas chromatographic/mass spectrometric profiles of complex mixtures. *Analytical Chemistry* **1977**, *49* (11), 1623-1632.
26. Greenberg, J. T.; Silverman, F. P.; Liang, H., Uncoupling salicylic acid-dependent cell death and defense-related responses from disease resistance in the Arabidopsis mutant *acd5*. *Genetics* **2000**, *156* (1), 341-350.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.