Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Metabolomic profiling reveals synchronized accumulation of biomass and bioactive flavonoids during in vitro development of Anoectochilus roxburghii

Shuangbin Fu
ZheJiang Academy of Agricultural Sciences
Yanping Yang
ZheJiang Academy of Agricultural Sciences
Zhiwei Kong
ZheJiang Academy of Agricultural Sciences
Zhen Ying
yingzhen0908@126.com

ZheJiang Academy of Agricultural Sciences

Research Article

Keywords:

Posted Date: March 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9040343/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

Background Anoectochilus roxburghii is a perennial orchid valued for its bioactive constituents. Traditionally, the harvest of in vitro cultured plants has relied primarily on plant height or approximate cultivation duration; however, identifying the optimal harvest window requires a precise understanding of the temporal dynamics of metabolite accumulation. Results In this study, we combined morphological phenotyping with widely targeted metabolomics to characterize the developmental trace of in vitro cultured A. roxburghii across five stages (120, 180, 240, 300, and 360 days after inoculation, DAI). Morphological characteristics generally showed a rapid growth from 120 DAI to 180 DAI and then to a stable plateau at 300–360 DAI. Time-series metabolomics demonstrated that lipids, terpenoids, and alkaloids were enriched in the early stages (120–180 DAI), whereas flavonoids and kinsenoside predominantly accumulated in the late stages (300–360 DAI). Weighted gene co-expression network analysis (WGCNA) identified distinct metabolite modules, revealing synchronized accumulation of biomass and flavonoids, alongside an inverse relationship between lipids and sugars and amino acids. These patterns indicate a metabolic transition from primary to secondary metabolism during development. Conclusion Our findings suggest that 300–360 DAI represents the optimal harvest window for balancing yield and quality, providing a scientific basis for precision in vitro cultivation.

1 Introduction

Anoectochilus roxburghii (Wall.) Lindl., commonly known as “Jinxianlian” in China, is a perennial orchid of significant economic and pharmaceutical value [1]. The whole plant serves as a source of traditional Chinese medicine and functional food ingredients. It exhibits diverse pharmacological activities, including hepatoprotection, anti-diabetes, and anti-inflammatory effects. These therapeutic properties are primarily attributed to its unique bioactive constituents, such as kinsenoside, flavonoids (e.g., quercetin, isorhamnetin), and polysaccharides [2, 3]. In particular, kinsenoside in *A. roxburghii* is a pharmacologically active compound that can repair damaged islet cells and restore normal insulin secretion [4]. Due to the scarcity of wild resources, strict environmental requirements, low propagation rate and the legal protection measures, artificial cultivation, particularly in vitro tissue culture has become the primary production method for meeting the commercial market demand [5, 6].

In medicinal plant cultivation, quality and yield are the two most critical factors, both of which are significantly influenced by growth duration [7, 8]. From the perspective of quality, this influence arises because the accumulation of bioactive ingredients, which forms the essential basis for clinical efficacy [5], is dynamic throughout plant development. Typically, most plants reach a specific growth stage during development at which the concentration of bioactive ingredients peaks, stabilizes, or begins to decline. Consequently, researchers routinely quantify these constituents at different developmental stages to establish the optimal harvest time. For instance, in *A. roxburghii*, analyses of flavonoids, polysaccharides, narcissoside, and kinsenoside have indicated that the optimal harvest time ranges from 4 to 8 months after planting, depending on cultivar, geographical region, and cultivation practices [9–13]. Based on fresh weight, drying rate, and *Bletilla striata* polysaccharide content, the recommended harvest

time of *Bletilla striata* is late September [14]. Metabolomic analysis of *Dendrobium officinale* stems revealed that flavonoids predominantly accumulated in biennial samples, whereas alkaloids are more abundant in triennial samples [15]. Similarly, in *Morinda officinalis* [16], various bioactive constituents reach their maximum concentrations after 3–4 years of growth. Therefore, analyzing active ingredients at different growth stages is crucial for identifying optimal harvest times and ensuring the quality of medicinal plants.

In contrast to numerous studies on field-cultivated medicinal plants that focus on time-dependent harvesting strategies, harvesting decisions for in vitro cultured *A. roxburghii* are primarily based on morphological indicators such as plant height or approximate cultivation duration (typically 6 months in our laboratory), and therefore lack scientific validation. Furthermore, in vitro cultured plants are maintained in isolated environments where indefinite growth is not feasible due to spatial constraints and nutrient depletion. Therefore, understanding bioactive compound accumulation patterns is beneficial for determine whether the plantlets are meant for direct sale, acclimatization, or subculture.

Therefore, in this study, we conducted widely targeted metabolomic analysis combined with morphological phenotyping on *A. roxburghii* at five time points (120, 180, 240, 300, 360 days after inoculation, DAI). We aimed to: (1) describe the morphological changes during the test periods; (2) map the temporal trend of metabolic accumulation; (3) identify key metabolites and hub metabolites using weighted gene co-expression network analysis (WGCNA); and (4) provide guidance on the optimal harvesting time.

2 Materials and methods

2.1 Materials

The initial plant materials of *A. roxburghii* (cultivar `Jianye') were asexually propagated plantlets obtained from the commercial nursery Jinxiu Agricultural Ecology in Fuzhou, Fujian, China. Subsequently, the plantlets were propagated via tissue culture and cultured in vitro for six months in our laboratory. Formal identification of the plant material was undertaken by Dr. Zhen Ying at the Zhejiang Institute of Subtropical Crops. A living voucher specimen (voucher number: J-2025) is currently maintained in the tissue culture room of the Zhejiang Institute of Subtropical Crops. The plant materials used in this study comply with the guidelines of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) [17] (https://cites.org/eng/app/appendices.php).

Internodes containing buds were excised from the same batch of plantlets and inoculated on Murashige and Skoog (MS) [18] medium (Kangbeisi, China) supplemented with 100 g/L banana pulp, 30 g/L sucrose (Morebetter, China), 1 g/L activated carbon (Morebetter, China), and 5 g/L agar (Morebetter, China). Approximately 150 mL of the medium was poured into each 650 mL glass bottle (caliber $\times$ chest diameter $\times$ height, 53 $\times$ 96 $\times$ 140 mm). Cultures were maintained at 25°C under a 16 h photoperiod provided by two LED light sources (300 lm, 14 W; OPPLE, China) for one year.

Samples were randomly collected at different time points, ranging from 120 DAI (when leaves became measurable) to 360 DAI (when plantlets reached the top of bottles). Fresh weight was measured using a precision balance (0.001 g; Sartorius BSA623S, Germany) after cleaning the attached media. Plant length was measured using a 20 cm ruler (Deli, China), with assistance from a tie when samples were bent. A vernier caliper (0–150 mm, Meinaite, Germany) was used to measure the lengths of leaves, roots, and internodes, as well as the widths of leaves and internodes. Leaves were photographed using a camera (Nikon Z 6II, Japan), and the area was determined using Fiji software [19]. Twenty biological replicates were measured for each morphological trait (1).

For metabolomic analysis, a separate set of plantlets was collected concurrently with the morphological measurements. Whole plants were snap-frozen in liquid nitrogen after cleaning the attached media and stored at − 80°C until further use. Three biological replicates were collected at each time point, with each replicate consisting of pooled tissues from three individual plantlets.

2.2 Sample extraction

The biological samples were placed in a lyophilizer (Scientz-100F), then grind (30 Hz, 1.5 min) the samples to powder form by using a grinder (MM 400, Retsch). Subsequently, 30 mg of sample powder was weighed using an electronic balance (MS105, Mettler Toledo, Switzerland) and add 1500 μL of − 20°C pre-cooled 70% methanolic aqueous internal standard extract (less than 30 mg added at the rate of 1500 μL extractant per 30 mg sample). Vortex once every 30 min for 30 s, for a total of 6 times. After centrifugation (rotation speed 12000 rpm, 3 min), the supernatant was aspirated, and the sample was filtered through a microporous membrane (0.22 μm pore size) and stored in the injection vial for UPLC-MS/MS analysis. Qualitative analysis of the metabolomic data was conducted by searching an internal database using a self-compiled metabolite index (Metware Biotechnology Inc., Wuhan, China; https://www.metwarebio.com/).

2.3 UPLC conditions

The sample extracts were analyzed using an UPLC-ESI-MS/MS system (UPLC, ExionLC AD) and Tandem mass spectrometry system (https://sciex.com.cn/). The analytical conditions were as follows, UPLC: column, Agilent SB-C18 (1.8 μm, 2.1 mm $\times$ 100 mm); The mobile phase consisted of solvent A, pure water with 0.1% formic acid, and solvent B, acetonitrile with 0.1% formic acid. Sample measurements were performed with a gradient program that employed the starting conditions of 95% A, 5% B. Within 9 min, a linear gradient to 5% A, 95% B was programmed, and a composition of 5% A, 95% B was kept for 1 min. Subsequently, a composition of 95% A, 5.0% B was adjusted within 1.1 min and kept for 2.9 min. The flow velocity was set as 0.35 mL per minute; The column oven was set to 40°C; The injection volume was 2 μL. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS.

2.4 ESI-Q TRAP-MS/MS

The ESI source operation parameters were as follows: source temperature 500°C; ion spray voltage (IS) 5500 V (positive ion mode)/ − 4500 V (negative ion mode); ion source gas I (GSI), gas II (GSII), curtain gas (CUR) were set at 50, 60, and 25 psi, respectively; the collision-activated dissociation (CAD) was high. triple quadrupole (QQQ) scans were acquired as multiple reaction monitoring (MRM) experiments with collision gas (nitrogen) set to medium. Declustering potential (DP) and the collision energy (CE) for individual MRM transitions were done with further DP and CE optimization. A specific set of MRM transitions were monitored for each period according to the metabolites eluted within this period.

2.5 Differential metabolites selected

For two-group analysis, differentially accumulated metabolites (DAMs) were determined by variable importance in projection $(VIP) > 1$, P-value < 0.05 (one-way analysis of variance, ANOVA), fold change ≥ 2 or fold change ≤ 0.5. VIP values were extracted from orthogonal partial least squares discrimination analysis (OPLS-DA) result, which also contained score plots and permutation plots, was generated using the R package MetaboAnalystR [20]. The data was log transform (Log_2) and mean centering before OPLS-DA. To avoid overfitting, a permutation test (200 permutations) was performed.

2.6 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).

2.7 Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed using the WGCNA R package [21], and all detected metabolites were included in the analysis. Since phenotype measurements ($n = 20$) and metabolite sampling ($n = 3$) were performed on different individuals within the same group, the mean values of phenotypic traits were assigned to each biological replicate of the metabolomic data for WGCNA analysis. The soft-thresholding power β was determined based on the scale-free topology criterion with the scale-free topology fit index (R^2) reaching 0.85.

Based on the selected power value, an adjacency matrix was constructed and subsequently transformed into a topological overlap matrix (TOM) to measure the network connectivity of metabolites. The corresponding dissimilarity (1-TOM) was used as the distance measure for metabolite hierarchical clustering.

Metabolite modules were identified using the dynamic tree cut algorithm with a minimum module size of 50. Modules with high similarity were merged by setting a cut height of 0.25. The module eigengene (ME), defined as the first principal component of each module, was calculated to summarize the overall metabolite expression within the module. Finally, Pearson correlation analysis was performed to assess

the relationships between MEs and phenotypic traits. Modules with significant correlation ($P < 0.05$) were selected as key modules for further analysis.

To screen key metabolites among different modules, correlations between metabolites and modules (module membership, MM) and between metabolites and phenotypic traits (gene significance, GS; retained here to maintain WGCNA terminology despite metabolite-based input data) were calculated. Metabolites meeting the criteria: $|MM| > 0.8$, $|GS| > 0.6$ and GS P-value < 0.05 were defined as key metabolites. These metabolites were subsequent calculated by using Euclidean distance approach. Specifically, the Euclidean distance (d) for each metabolite was calculated based on its MM and GS coordinates as $d = \sqrt{(MM - x)^2 + (GS - y)^2}$, where (x, y) represents the coordinates of the theoretical ideal key (typically (1, 1) or (-1, 1) depending on the correlation direction). Metabolites located at the extremes of the MM-GS distribution, the top five positively correlated and top five negatively correlated metabolites with the shortest Euclidean distances to the peak coordinates were labeled and focused.

To visualize the complex inter-relationships among metabolites within the modules, the weighted adjacency data were exported for co-expression network construction.

2.8 Statistical analysis

Data analysis was conducted using the R platform [22]. Data visualizations were primarily generated with the ggplot2 package [23]. Statistical significance of group differences was assessed using one-way ANOVA followed by Tukey's post hoc test implemented in the rstatix package [24]. For normally distributed variables exhibiting unequal variances, Welch's ANOVA with the Games-Howell post-hoc test was performed. Non-normally distributed variables were analyzed using the non-parametric Kruskal-Wallis test followed by Dunn's post-hoc test. The significance level was set at $P < 0.05$ for all statistical tests. Principal component analysis (PCA) and hierarchical clustering on principal components (HCPC) were performed using FactoMineR and visualized with factoextra [25]. UpSet plot was generated using the ggupset package [26]. Cluster analysis was carried out using the Mfuzz package [27], and the resulting metabolite clusters were visualized with ClusterGVis [28]. Heatmaps were generated using the ComplexHeatmap package [29].

3 Results

3.1 Dynamic changes in growth and morphological characteristics of *A. roxburghii* development

The fresh weight (2A) and plant length (2B) exhibited significant increases from 120 DAI to 300 DAI, followed by a slight stabilization or decrease at 360 DAI. The number of roots (2C) showed a rapid expansion phase between 180 and 300 DAI, followed by a stabilization at 360 DAI. Root length (2D) showed a stabilization after a growth from 120 DAI to 180 DAI. Number of leaves showed a gradually

increasing between 180 and 300 DAI, followed by a decrease at 360 DAI. Other leaf morphological traits including area, length, and width (2E–H) showed a rapid expansion phase between 120 and 180 DAI, followed by a stabilization after 300 DAI. Internodes characteristics including number, length, and width (2I–K) displayed consistent growth patterns, all reaching a plateau at 300–360 DAI. Overall, morphological characteristics generally showed a rapid growth from 120 DAI to 180 DAI, followed by a potential transition phase at 240 DAI. Subsequently, reaching a stable plateau at 300–360 DAI. This indicated that the period of 300–360 DAI, where morphological traits stabilized, represents the optimal harvest window for in vitro cultured *A. roxburghii*.

3.2 Metabolomic profiling and identification of DAMs

PCA revealed a clear separation among the groups (3A). The first principal component, accounting for 32.2% of total variance, distinguished the early developmental stages from the latest stage (360 DAI), suggesting a metabolic reprogramming occurs during this stage. This distinct clustering pattern was further corroborated by HCPC (3B) and correlation of samples (Figure S1), which grouped biological replicates tightly and separated samples according to their developmental chronology.

We then calculated the number and class of metabolites and found a total of 3361 metabolites were identified and classified into 13 classes. Among them, flavonoids (587), others (512), terpenoids (450), and alkaloids (434) were the predominant classes (3C). After applying OPLS-DA (Figure S2), we screened numerous DAMs across pairwise comparisons (3D). Notably, the number of DAMs increased with the temporal distance between stages. The comparison between early stages and 360 DAI yielded the highest number of DAMs, the top three comparisons with the highest number of DAMs all involved 360 DAI versus earlier stages. Conversely, adjacent developmental stages (e.g., 300 DAI vs. 240 DAI) showed fewer DAMs, indicating a gradual transition in the metabolic profile. Among these comparisons, flavonoids were generally the most numerous DAMs (Figure S3). The UpSet plot further revealed that comparisons involving 360 DAI shared the largest number of common DAMs with other pairwise comparisons, indicating that the metabolic profile at 360 DAI is most distinct from earlier developmental stages. Across all comparisons, two flavonoids, quercetin-3,4’-O-di-glucoside and homoeriodictyol-7,4’-di-O-β-D-glucopyranoside were consistently identified as common DAMs (Table S1).

3.3 KEGG pathway enrichment and dynamic accumulation of flavonoids

To investigate the metabolic shifts occurring throughout the time course, we performed a KEGG pathway enrichment analysis on the DAMs between different comparisons. As shown in 4, the DAMs were significantly enriched in pathways such as flavone and flavonol biosynthesis, and biosynthesis of quercetin aglycones (I and II). Particularly, biosynthesis of quercetin aglycones (I and II) pathways contained the most significantly enriched DAMs.

To further elucidate the accumulation patterns of these compounds, we then selected these DAMs enriched in biosynthesis of quercetin aglycones I and II pathways. The heatmap reveals a distinct

temporal trend in metabolite accumulation. From 120 DAI to 360 DAI, the abundance of most quercetin derivatives shows a general increasing trend (Figure S4). This indicates the metabolic flux is strongly directed toward the synthesis and accumulation of specific quercetin glycosides as the time post-inoculation increases.

3.4 Temporal accumulation patterns and functional characterization of metabolites

To further explore the dynamic accumulation patterns of metabolites, we performed Mfuzz clustering analysis, which classified the metabolites into eight distinct clusters (C1–C8) based on their temporal trends (Figure S5).

The clusters could be broadly categorized into two major trends: early accumulated and late accumulated patterns. Clusters C2, C3, and C6 exhibited an early accumulated pattern, where metabolite levels were highest at 120 or 180 DAI and subsequently declined. These DAMs mainly classify into lipids, alkaloids and terpenoids. Functional annotation revealed that these clusters were enriched in metabolic pathways, biosynthesis of stilbenoids II, biosynthesis of phenanthrenes and linoleic acid metabolism. In contrast, Clusters C1, C5, C7, and C8 displayed a later accumulated pattern, with metabolite abundance peaking at 300 or 360 DAI. These clusters were predominantly composed of flavonoids. KEGG enrichment analysis indicated that these late-accumulating clusters were significantly enriched in biosynthesis of stilbenoids I, biosynthesis of quercetin aglycones (I and II), tyrosine metabolism, citrate cycle (TCA cycle), glucosinolate biosynthesis and biosynthesis of anthocyanins II (5). This shift suggests a metabolic transition from primary growth to the synthesis of bioactive secondary metabolites as the plant grows.

3.5 Identification of key metabolite modules associated with growth traits via WGCNA

Sample dendrogram and trait heatmap analysis revealed a clear temporal trend in the phenotypic traits. Samples from 120 DAI showed lower values for traits. In contrast, samples from later stages showed significantly higher growth metrics. No significant outliers were detected, and all samples were retained for the weighted co-expression network (Figure S6).

The results of WGCNA identified 9 as the estimated soft threshold and classified all metabolites into 12 modules (Figure S6). Among these modules, the pink and magenta showed strong positive correlations with almost all growth traits. The part of turquoise module also exhibited a significant positive correlation (6A). Functional enrichment analysis indicated that these growth-correlated modules were significantly enriched in secondary metabolic pathways, such as biosynthesis of various plant secondary metabolites, biosynthesis of cinnamic acid derivatives and biosynthesis of quercetin aglycones (6B). Conversely, the black, blue, and purple modules showed significant negative correlations with growth traits (6A). These modules likely represent metabolites that are abundant in the early developmental

stages but decrease during later stages. Functional enrichment analysis indicated that these growth-correlated modules were significantly enriched in primary metabolic pathways, including pathways related to sugar, amino acids, inositol and lipids (6B).

To conduct a comprehensive analysis, we drew the scatter plot of above six modules based on GS versus MM. The plots demonstrated significant correlations (both positive and negative) between the traits and module metabolites (6C). Consequently, we then screened the top five metabolites from both the positive and negative quadrants using Euclidean distance (Table S1). These metabolites are candidate marker compounds for researching the correlations between growth and metabolites. Chemical classification further revealed that the positive modules were enriched with flavonoids, whereas the negative modules were predominantly composed of lipids (Figure S7).

3.6 Identification of hub metabolites via networks

To identify the core drivers within these modules, we selected top 100 weights and constructed co-expression networks for the six significant modules (7, S8). In the interaction networks (7A–D), we identified six hub metabolites based on their connectivity (degree). For instance, Safp010408 (Difenoconazole), WaYn011395 [] and Hmln005329 (Ephemeranthoquinone B) with the highest degrees were identified as central hubs in the pink, magenta and turquoise modules, respectively. Meanwhile, Zmxn110201 (2,6-Dimethoxy-4-hydroxyphenol-1-O-β-D-glucopyranoside), WapTKC008879 () and Sal001190 (Abealpha1-3Manalpha) were central to the black, blue and purple modules, respectively (annotations were listed on Table S2). In addition, we found in the pink and blue networks, lipids and terpenoids were the dominant metabolites, respectively.

3.7 Dynamic accumulation of major chemical components

To further and precisely trace the bioactive compounds, we selected kinsenoside and narcissoside. As shown in 8A, kinsenoside increased significantly from 120 DAI to 180 DAI, then remained relatively stable through the later stages. In contrast, narcissoside content showed a more dynamic and overall upward trend (8B). The levels increased significantly between 120 DAI and 240 DAI. After a significant decrease at 300 DAI, the narcissoside content reached its maximum at 360 DAI.

In addition, based on previous studies [30], we selected several identified components with similar structures. First, we compared the relative content of ursolic acid (terpenoids). The compounds exhibited a downward trend from 120 DAI to 240 DAI, followed by a sudden surge to the peak at 300 DAI. Subsequently, they experienced a decline, reaching nearly zero at 360 DAI (8C). Upon investigating identified compounds exhibiting partial matches to “quercetin”, “kaempferol”, “isorhamnetin”, and “rhamnazin”, all of which are flavonoids, we drew the heatmaps and observed that most of these compounds presented higher content at 240–360 DAI. In contrast, steroids, and organic acids and volatile compounds, these selected compounds presented scattered changes at all stages (Figure S9).

4 Discussion

In this study, we characterized morphological development and metabolic reprogramming in vitro cultured *A. roxburghii*, revealing a clear developmental transition from rapid growth to secondary metabolite accumulation. Most morphological traits increased before 300 DAI and subsequently plateaued. Whereas secondary metabolites, particularly flavonoids, were consistent with these morphological trends and continued to increase during later development. This developmental pattern is consistent with observations in other medicinal plants such as *Dendrobium officinale* [15], *Morinda officinalis* [16], which show enhanced investment in secondary metabolism at later stages. As review by Li et al., perennial herbs generally exhibit higher content and yield of secondary metabolites as they grow [31].

Furthermore, based on WGCNA, we found a synchronized metabolic program where biomass accumulation and secondary metabolite synthesis are synergistic. Accordingly, we propose that 300–360 DAI represents an ideal harvest window that achieves a “double peak” of yield and quality. In contrast, we observed a strong negative correlation between biomass and primary metabolites, including sugars, amino acids, inositol, and lipids, suggesting a decisive metabolic redirection during development. Similar developmental patterns have been reported in *Stellaria dichotoma* [32], *Polygala tenuifolia* [33] and *Prunella vulgaris* [34], where simultaneous biomass accumulation and secondary metabolite accumulation occur alongside a metabolic flux shift from primary metabolites to secondary metabolites. Mechanistically, the cause lies in the plant’s adaptation to different developmental stages and environmental demands. By utilizing transcription factors to regulate the expression of key enzymatic genes, the plant shifts the direction of metabolic flux, effectively reallocating limited resources between primary metabolism and secondary metabolism [7].

Through GS and MM analysis, we identified key metabolites within each modules, providing potential molecular markers for harvest time optimization. Regarding hub metabolites, we infer that they may function as metabolic intersections rather than terminal products from a topological perspective. Furthermore, their connectivity often extends across different chemical classes, indicating their role in mediating inter-pathway crosstalk. Based on profiling of these targeted compounds in previous studies, the hub metabolites could be classified as phytoalexins or compounds with pharmacological potential (Table S2). However, our analysis stops at the metabolite level; the underlying molecular machinery, specifically the transcriptional regulators and enzymatic drivers, remains to be fully elucidated. Future studies that integrate transcriptomics and proteomics data to construct a multi-layered regulatory network could provide a clearer understanding of these metabolites and the underlying networks.

In addition to flavonoids, we found some metabolites, such as lipids, alkaloids, terpenoids and steroids, mainly accumulated at early stages. Our decision not to select these metabolites as harvest indices does not imply that they lack value, even though flavonoids represent the primary functional constituents of *A. roxburghii*. As reviewed by Ye et al. [30], some compounds belonging to these classes are widely applied in disease treatment and health care. Therefore, we agree with previous conclusions emphasizing the importance of integrating multiple chemical markers, rather than relying on a single index compound, when determining harvest timing [35, 36]. Although we propose 300–360 DAI as the

optimal harvesting time, this recommendation has limitations, and harvesting strategies should be adjusted when specific compounds are targeted.

Compared with naturally cultivated *A. roxburghii*, in vitro cultured plantlets exhibit broadly similar patterns of metabolite accumulation, with increases in secondary metabolites (e.g., polysaccharides and flavonoids) and biomass as development progresses [11, 12, 37]. However, differences exist for specific compounds: tissue cultured plantlets are typically richer in kinsenoside and proteins, whereas field-grown plants accumulate higher levels of certain secondary metabolites [9, 38]. This difference may reflect the unique physiological environment of in vitro culture, where nutrient availability, light conditions, and stress exposure differ from field conditions.

An important difference from previous studies is that our study adopted whole-plant metabolomic strategy to evaluate overall quality. However, secondary metabolites often exhibit distinct tissue-specific distribution patterns [31, 39]. In *A. roxburghii*, previous studies have reported that flavonoids mainly accumulate in leaves, with relatively less accumulation in stem segments [40]. Therefore, our whole-plant sampling approach may partially mask the distinct accumulation peaks of specific organs. Future studies incorporating spatial metabolomics or organ-specific sampling could refine our understanding of metabolic allocation and enable more precise cultivation strategies, such as optimizing harvest timing based on leaf-to-stem ratios or developing targeted processing methods for different plant parts.

Indeed, our results showed that most flavonoids reached their maximum at 360 DAI with a significantly upward trend. This suggests the metabolic potential for flavonoid biosynthesis might not have been fully exhausted by the end of the experiment. However, under practical in vitro cultivation conditions, plants at 360 DAI had already reached the physical limits of the culture vessels. Such spatial constraints, together with nutrient depletion and potential accumulation of autotoxic metabolites, led to a decrease in biomass parameters, such as fresh weight and plant length. Future studies employing larger bioreactor systems could provide deeper insights into the ultimate capacity of *A. roxburghii* to synthesize bioactive constituents under controlled conditions. In addition, in Chinese agricultural practice, in vitro *A. roxburghii* plantlets are routinely transplanted to under forest environments for wild-simulated cultivation. Therefore, an integrated assessment combining performance data from both the prolonged in vitro phase and the subsequent under forest stage would provide comprehensive criteria for determining the optimal harvest time.

5 Conclusion

This study provides the first comprehensive metabolomic map of *A. roxburghii* development from 120 to 360 DAI, revealing distinct temporal patterns in primary and secondary metabolism. The strong positive correlation between growth traits and bioactive flavonoid accumulation suggests synergistic metabolic programming during later development stages. Based on morphological, metabolomic, and WGCNA analyses, we recommend 300–360 DAI as the optimal harvest window for balancing yield and quality in in vitro cultivated *A. roxburghii*.

Abbreviations

ANOVA, analysis of variance; DAI, days after inoculation; DAMs, differentially accumulated metabolites; ESI-Q TRAP-MS/MS, electrospray ionization-quadrupole trap-tandem mass spectrometry; GS, gene significance; KEGG, Kyoto Encyclopedia of Genes and Genomes; LC-MS/MS, liquid chromatography tandem-mass spectrometry; ME, module eigengene; MM, module membership; MS, Murashige and Skoog; OPLS-DA, orthogonal partial least squares discrimination analysis; PCA, principal components analysis; QC, quality control; SD, standard deviation; TOM, topological overlap matrix; UPLC, ultra performance liquid chromatography; VIP, variable importance in projection; WGCNA, weighted gene co-expression network analysis.

Declarations

Funding: This work was supported by the Zhejiang Provincial Key R&D Program: Vanguard and Leading Goose Initiative (2025C01133), Wenzhou Agricultural New Variety Breeding Cooperation Group (ZX2024005-5), Project of Zhejiang Academy of Agricultural Sciences (Chinese Medicinal Crops-Breeding and Cultivation of *Anoectochilus roxburghii* (Wall.) Lindl.) (2024LY27FX05).

Conflict of interest/Competing interest

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Author’s contributions

S.F and Z.Y conceived and designed the experiments; S.F and Y.Y performed the experiments; S.F, Z.K and Z.Y analyzed the data; S.F wrote the manuscript. All the authors have read and approved the final version of manuscript.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files. Detailed data and analysis process can contact the corresponding author.

References

1. Zhang Y, Li Y, Chen X, Meng Z, Guo S. Combined metabolome and transcriptome analyses reveal the effects of mycorrhizal fungus ceratobasidium sp. AR2 on the flavonoid accumulation in *anoectochilus roxburghii* during different growth stages. Int J Mol Sci. 2020; 21:564. https://doi.org/10.3390/ijms21020564.
2. Du XM, Sun NY, Tamura T, Mohri A, Sugiura M, Yoshizawa T, et al. Higher yielding isolation of kinsenoside in *Anoectochilus* and its antihyperliposis effect. Biol Pharm Bull. 2001; 24:65–9. https://doi.org/10.1248/bpb.24.65.
3. Cui SC, Yu J, Zhang XH, Cheng MZ, Yang LW, Xu JY. Antihyperglycemic and antioxidant activity of water extract from *anoectochilus roxburghii* in experimental diabetes. Exp Toxicol Pathol. 2013; 65:485–8. https://doi.org/10.1016/j.etp.2012.02.003.
4. Ito A, Kasai R, Yamasaki K, Sugimoto H. Aliphatic and aromatic glucosides from *Anoectochilus koshunensis*. PYTCAS. 1993; 33:1133–7. https://doi.org/10.1016/0031-9422(93)85037-R.
5. Li B, Tang MJ, Tang K, Zhao LF, Guo SX. Screening for differentially expressed genes in *anoectochilus roxburghii* (orchidaceae) during symbiosis with the mycorrhizal fungus *epulorhiza* sp. Sci China: Life Sci. 2012; 55:164–71. https://doi.org/10.1007/s11427-012-4284-0.
6. Zhang AL, Wang HZ, Shao QS, Xu MJ, Zhang WS, Li MY. Large scale *in vitro* propagation of *Anoectochilus roxburghii* for commercial application: Pharmaceutically important and ornamental plant. Ind Crops Prod. 2015; 70:158–62. https://doi.org/10.1016/j.indcrop.2015.03.032.
7. Yang H, Xin BH, Tong K, Tian ML. Medicinal material quality analysis and evaluation of paris plant. AMR. 2014;926–930:1152–8. https://doi.org/10.4028/www.scientific.net/AMR.926-930.1152.
8. Chu SS, Tan LL, Liu CC, Peng HS, Duan HY, Huang LQ. Growth rings in roots of medicinal perennial dicotyledonous herbs from temperate and subtropical zones in China. Microsc Res Tech. 2018; 81:365–75. https://doi.org/10.1002/jemt.22987.
9. Chen Y, Wang WY, Chen SN, Huang JF, Zhang X, Wu SS, et al. Changes in kinsenoside content of different strains of *anoectochilus roxburghii* at different growth periods. Chin Med. 2021; 23:1423–9. https://doi.org/10.13313/j.issn.1673-4890.20200911001.
10. Jiang YB, Li J, Ma YF, Huang YF. Dynamic change of polysaccharide in *Anoectochilus roxburghii* and *A.formosanus* in different collecting time. J Fujian Agric For Univ (Nat Sci Ed). 2014; 43:125–7. https://doi.org/10.13323/j.cnki.j.fafu(nat.sci.).2014.02.013.
11. Wei CH, Xie Y, Qin JB, Chen Q. Comparison of growth and dry matter content of plantlets in vitro of *anoectochilus roxburghii* under different cultivating periods. Subtrop Plant Sci. 2015; 44:107. https://doi.org/10.3969/j.issn.1009-7791.2015.02.004.
12. Zhang WJ, Zhao YQ, Liu BC, Huang YZ, Cai JR, Chen JY. Study on the optimal harvesting period of *Anoectochilus roxburghii* cultivated under forest conditions. J Chin Med Mater. 2024; 47:2957–65. https://doi.org/10.13863/j.issn1001-4454.2024.12.003.
13. Qiu DS, Luo YJ, Zheng WL, Li L. Comparative analysis of widely targeted metabolomics between greenhouse-cultivated and bottle-acclimated *anoectochilus roxburghii*. Subtrop Agric Res. 2024; 51:30–43. https://doi.org/10.13321/j.cnki.subtrop.agric.res.2024.02.001.

14. Wang J, Lv AM, Yao LN, Zhang M, Fang HL, Jiang JP, et al. Effects of harvest times on the production and quality of *bletilla striata*. Acta Physiol Plant. 2024; 46:107. https://doi.org/10.1007/s11738-024-03708-x.
15. Yuan YD, Zuo JJ, Zhang HY, Zu MT, Yu MY, Liu SA. Transcriptome and metabolome profiling unveil the accumulation of flavonoids in *dendrobium officinale*. Genomics. 2022; 114:110324. https://doi.org/10.1016/j.ygeno.2022.110324.
16. Yip KM, Xu J, Zhou SS, Lau YM, Chen QL, Tang YC, et al. Characterization of chemical component variations in different growth years and tissues of Morindae Officinalis Radix by integrating metabolomics and glycomics. J Agric Food Chem. 2019; 67:7304–14. https://doi.org/10.1021/acs.jafc.9b01910.
17. CITES Secretariat. Appendices I, II and III to the Convention on International Trade in Endangered Species of Wild Fauna and Flora. 2023; p. 62.
18. Murashige T, Skoog F. A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant. 1962; 15:473–97. https://doi.org/10.1111/j.1399-3054.1962.tb08052.x.
19. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: An open-source platform for biological-image analysis. Nat Methods. 2012; 9:676–82. https://doi.org/10.1038/nmeth.2019.
20. Chong J, Xia J. MetaboAnalystR: An R package for flexible and reproducible analysis of metabolomics data. Bioinformatics. 2018; 34:4313–4. https://doi.org/10.1093/bioinformatics/bty528.
21. Langfelder P, Horvath S. WGCNA: An R package for weighted correlation network analysis. BMC Bioinf. 2008; 9:559. https://doi.org/10.1186/1471-2105-9-559.
22. R Core Team. R: A language and environment for statistical computing. Manual. Vienna, Austria: R Foundation for Statistical Computing; 2024.
23. Wickham H. Ggplot2: elegant graphics for data analysis. Springer-Verlag New York; 2016.
24. Kassambara A. Rstatix: Pipe-friendly framework for basic statistical tests. 2023.
25. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J STAT SOFTW. 2008; 25:1–18. https://doi.org/10.18637/jss.v025.i01.
26. Ahlmann-Eltze C. Ggupset: Combination matrix axis for “ggplot2” to create “UpSet” plots. 2020.
27. Kumar L, Futschik ME. Mfuzz: a software package for soft clustering of microarray data. Bioinformation. 2007; 2:5–7. https://doi.org/10.6026/97320630002005.
28. Zhang J. ClusterGVis: one-step to cluster and visualize gene expression matrix. 2022.
29. Gu ZG, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016; 32:2847–9. https://doi.org/10.1093/bioinformatics/btw313.
30. Ye SY, Shao QS, Zhang AL. *Anoectochilus Roxburghii*: A review of its phytochemistry, pharmacology, and clinical applications. J Ethnopharmacol. 2017; 209:184–202. https://doi.org/10.1016/j.jep.2017.07.032.

31. Li YQ, Kong DX, Fu Y, Sussman MR, Wu H. The effect of developmental and environmental factors on secondary metabolites in medicinal plants. Plant Physiol Biochem. 2020; 148:80–9. https://doi.org/10.1016/j.plaphy.2020.01.006.
32. Li ZK, Feng L, Wang H, Zhang L, Li HS, Li YQ, et al. The impact of growth years on the medicinal material characteristics and metabolites of *stellaria dichotoma* l. Var. *Lanceolata* bge. reveals the optimal harvest age. Plants. 2023; 12:2286. https://doi.org/10.3390/plants12122286.
33. Xue Y, Li XW, Li ZY, Zeng ZP, Zhang FS, Li AP, et al. UPLC/Q-TOF MS and NMR plant metabolomics approach in studying the effect of growth year on the quality of *polygala tenuifolia*. Acta pharmaceutica Sinica. 2015; 50:340–7. https://doi.org/10.16438/j.0513-4870.2015.03.003.
34. Zhang ZM, Su Q, Xia BH, Li YM, Qin XY, Luo HS, et al. Integrative transcriptomic, proteomic and metabolomic analysis reveals the dynamic regulation of secondary metabolism upon development of *prunella vulgaris* L. Fitoterapia. 2022; 163:105334. https://doi.org/10.1016/j.fitote.2022.105334.
35. Ajdert P, Jan L, Burman R. Liquid chromatographic method for the quantification of salidroside and cinnamyl alcohol glycosides for quality control of golden root (*rhodiola rosea* l.). J Appl Res Med Aromat Plants. 2022; 26:100364. https://doi.org/10.1016/j.jarmap.2021.100364.
36. Fu L, Wang P, Sun YQ, Wang YY, Zhao J, Ye YT, et al. High performance liquid chromatography time of flight electrospray ionization mass spectrometry for quantification of sesquiterpenes in *chrysanthemi indici* flos active extract. Pharmacogn Mag. 2015; 11:740. https://doi.org/10.4103/0973-1296.165574.
37. Wei Q. Effects of wild-like planting density and harvest time under forest on the planting of *Anoectochilus roxburghii*. Journal of Green Science and Technology. 2023;79–82. https://doi.org/10.16663/j.cnki.lskj.2023.05.003.
38. He H, Liu HR, Qiao YJ, Zhang Y, Wang CF, Chen B jie, et al. Active ingredients and volatiles in *anoectochilus roxburghii* strains at various growth stages. Fujian J Agric Sci. 2023; 38:271–80. https://doi.org/10.19303/j.issn.1008-0384.2023.03.003.
39. Belkheir AK, Gaid M, Liu B, Hänsch R, Beerhues L. Benzophenone synthase and chalcone synthase accumulate in the mesophyll of *hypericum perforatum* leaves at different developmental stages. Front Plant Sci. 2016;921. https://doi.org/10.3389/fpls.2016.00921.
40. Lin HJ, Xu ZH, Lyu K, Dong LL, Pan PP, Wang ZH. In vitro antioxidant activity of *Anoectochilus roxburghii* and preliminary investigation of the synthesis pathway of flavonols. J Nucl Agric Sci. 2024; 38:1868–78. https://doi.org/10.11869/j.issn.1000-8551.2024.10.1868.

Figures

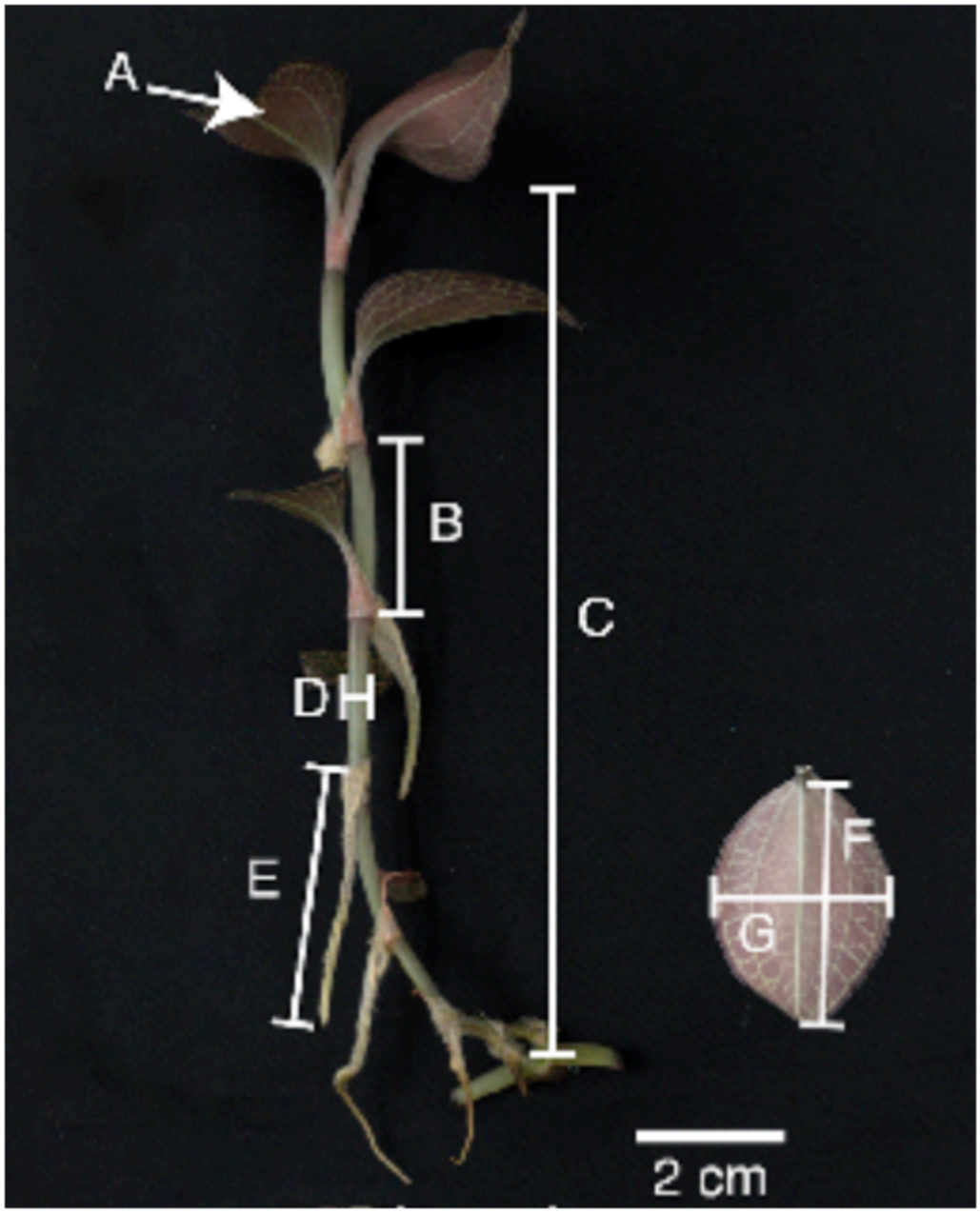


Figure 1

Schematic representation of morphological measurements taken for A. roxburghii. (A) Leaf; (B) Internode length; (C) Plant length; (D) Internode width; (E) Root length; (F) Leaf length; (G) Leaf width.

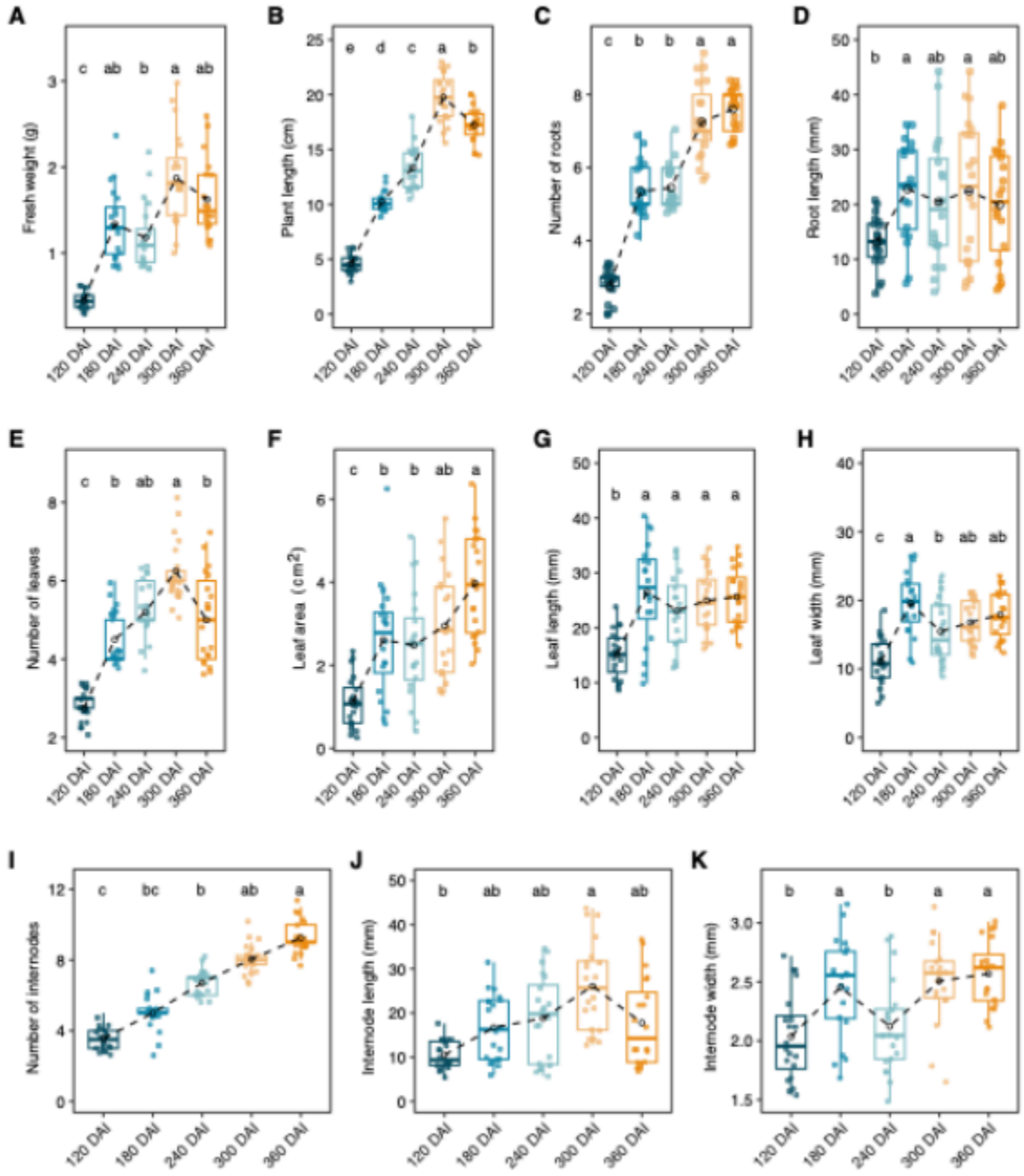


Figure 2

Temporal dynamic of growth and morphological characteristics in developmental stages. A. roxburghiiat different (A) Fresh weight. (B) Plant length. (C) Number of roots. (D) Root length. (E) Number of leaves. (F) Leaf area. (G) Leaf length. (H) Leaf width. (I) Number of internodes. (J) Internode length. (K) Internode width. The colored points represent individual biological replicates (n=20), and the black open circles connected by dashed lines denote the mean values for each time point. Different lowercase letters above the boxes indicate statistically significant differences between sampling times determined by Methods: (1) Kruskal-Wallis with Dunn's test (applied to A, C, E, I, and J); (2) Welch's ANOVA with

Games-Howell test (applied to B, D, F, and G); and (3) one-way ANOVA with Tukey's post-hoc test. P <0.05 (applied to F and K).



Figure 3

(A) PCA score plot of metabolite profiles derived from standardized data. QC: quality control samples. (B) HCPC dendrogram. (C) Classification of the identified metabolites categorized by chemical class. (D) The number of DAMs in pairwise comparisons between developmental stages. (E) UpSet plot. The

vertical stacked bars indicate the number of shared DAMs in the specific intersections, with colors representing the chemical classification of these metabolites.

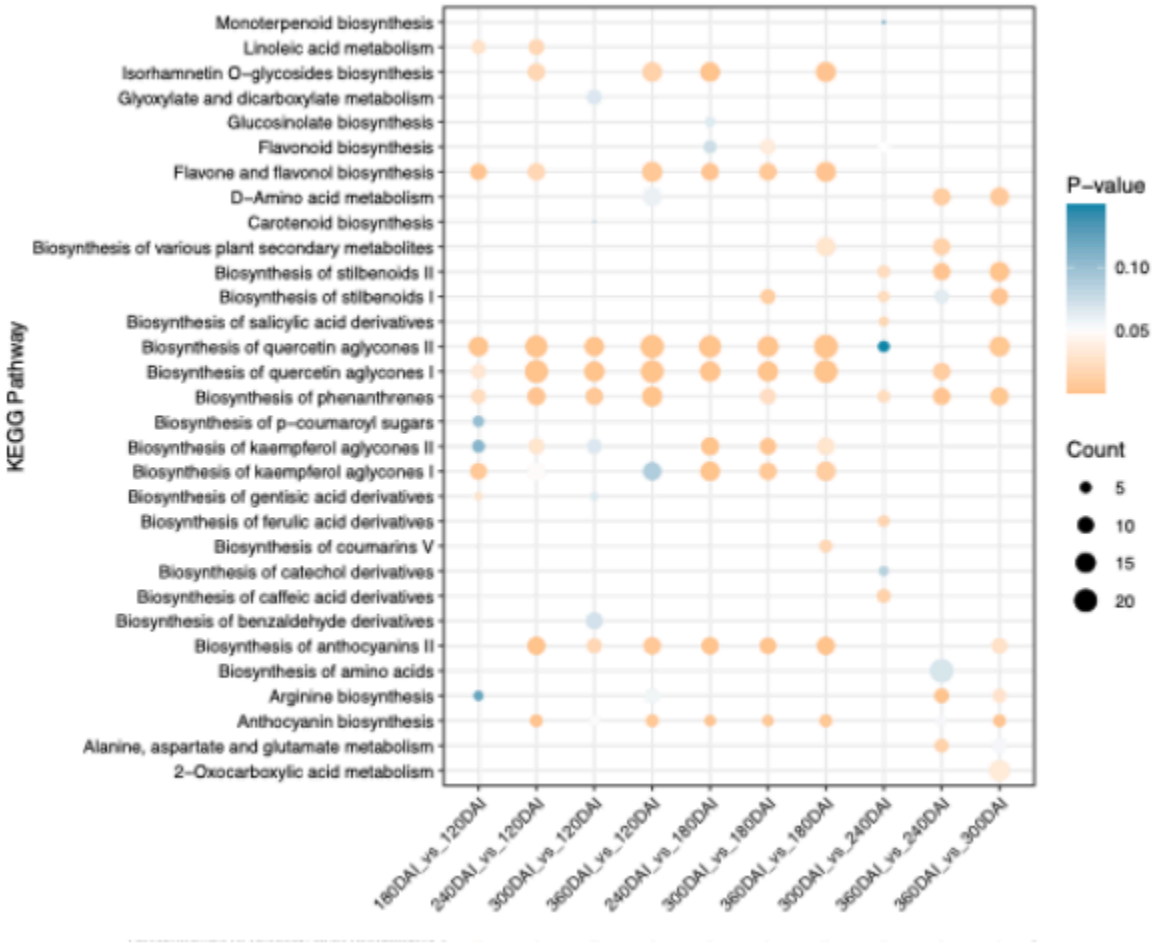


Figure 4

Bubble plot of KEGG pathway enrichment analysis for differential metabolites across different comparison groups.



Figure 5

Time-series changes across five development stages. The central heatmap displays the standardized relative abundance (Z-score) of metabolites across 5 time points. In the boxes on the left, the red line graphs represent the average expression trend (mean Z-score) of all metabolites within each cluster. The class column displays the chemical categories of the metabolites, with the font size proportional to the number of metabolites in that category; the rightmost column lists the significantly enriched KEGG pathways associated with the metabolites in each cluster. The font size is proportional to the significance of enrichment.

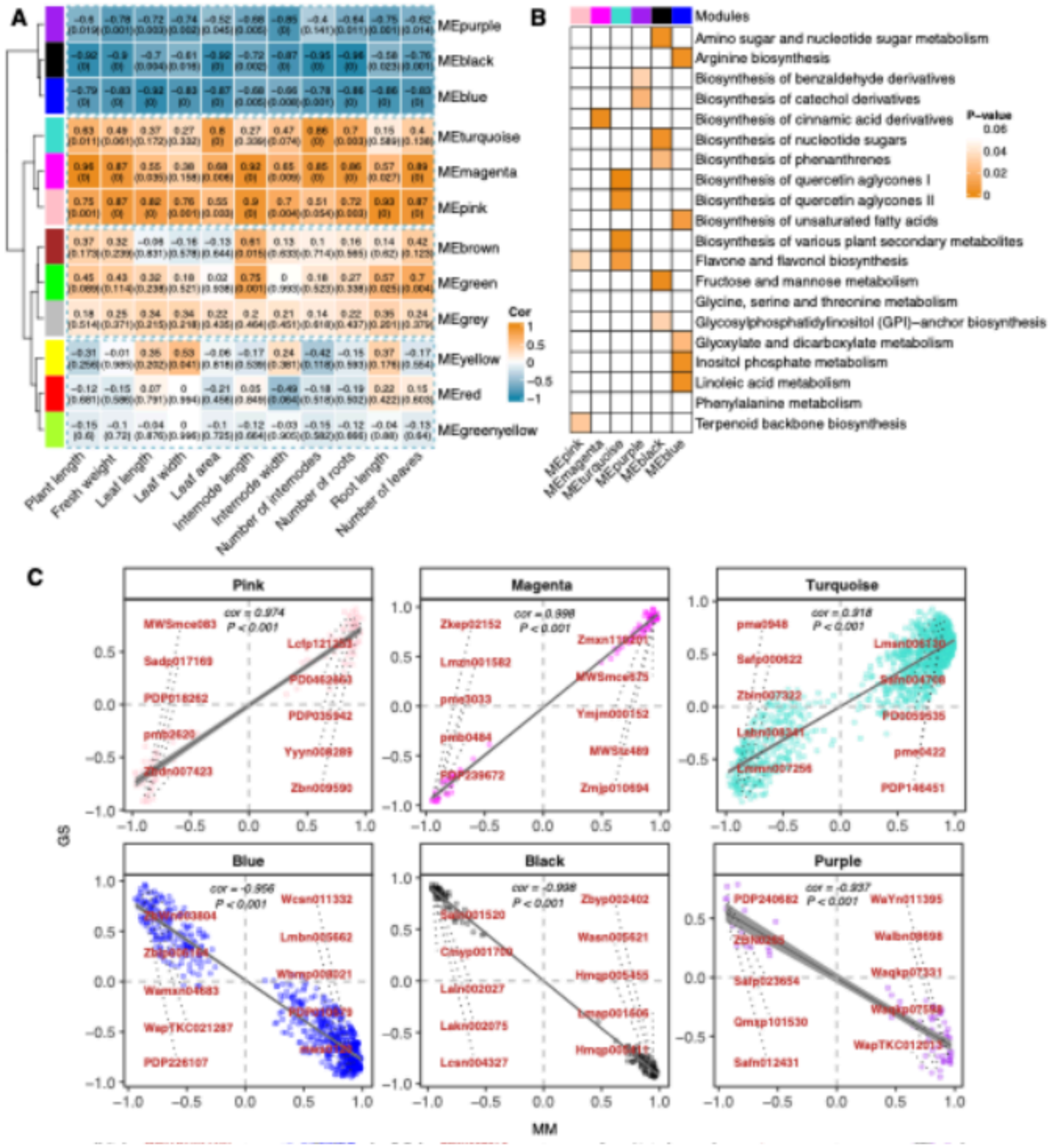


Figure 6

(A) Module-trait relationship heatmap. Each cell displays the Pearson correlation coefficient (top) and the corresponding P-value (bottom, in parentheses). (B) KEGG pathway enrichment analysis of the key modules (top five significant enrichment). (C) Scatter plots of GS versus MM for the six representative modules. Labeled points (by index) indicate the top-ranking metabolites identified based on Euclidean distance. Five metabolites were selected from each of the upper-right (positive GS and MM) and lower-

left (negative GS and MM) quadrants. Pearson correlation coefficients (cor) and statistical significance (P-values) are shown for each plot.

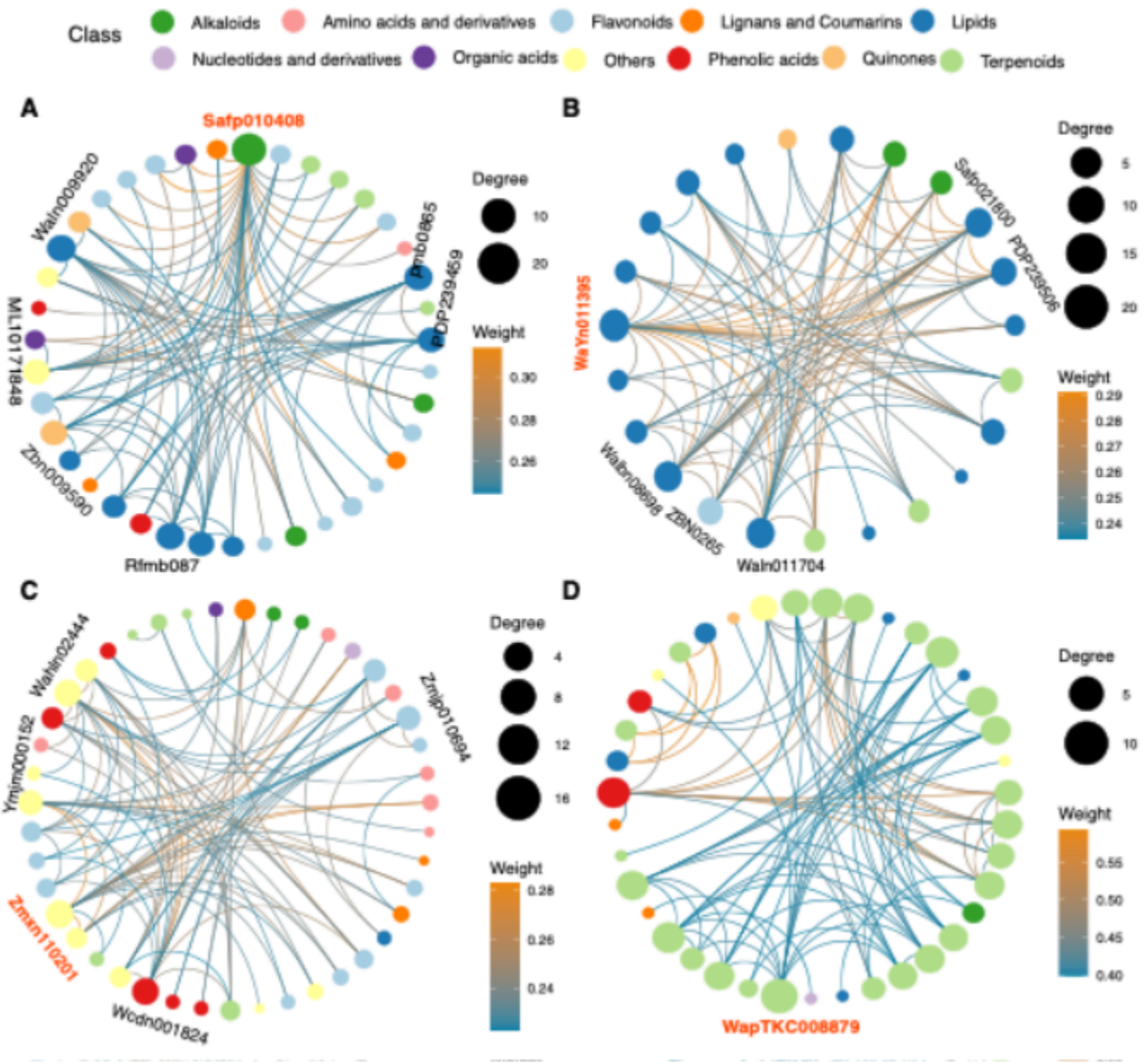


Figure 7

(A) Pink, (B) magenta, (C) black, and (D) blue modules. Nodes represent individual metabolites, with the size proportional to their degree. Edges represent the interaction weights between metabolites, with the color presenting the weight number. The red labels highlight the hub metabolites in each network (hub metabolites were labeled by index, nodes with a degree greater than 10 were selected and labeled. Other metabolites were not labeled to keep off overlap).

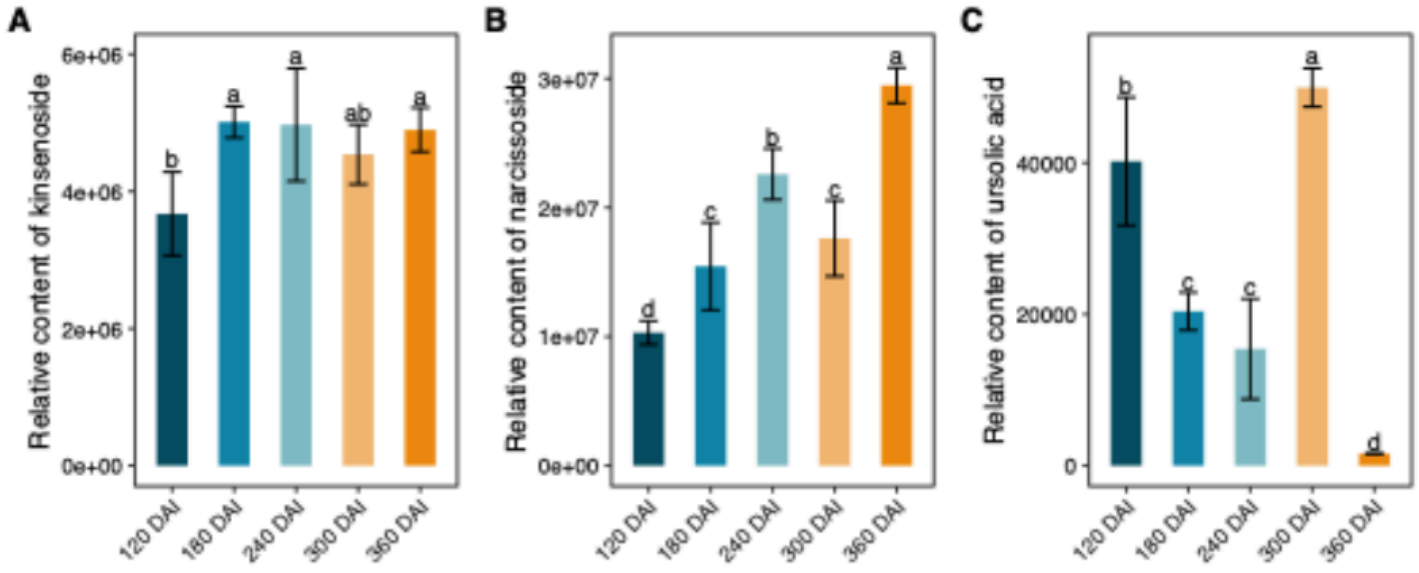


Figure 8

Relative contents of (A) kinsenoside, (B) narcissoside and (C) ursolic acid. Data are presented as mean ± SD (standard deviation). Different lowercase letters above the bars indicate significant differences at P <0.05 according to Tukey's test.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Sup.pdf
- TableS1.xlsx
- TableS2.xlsx