

RESEARCH

Open Access



# Integrated transcriptome and endogenous hormone analyses reveal the factors affecting the yield of *Camellia oleifera*

Yayan Zhu<sup>1†</sup>, Da Huo<sup>1†</sup>, Minggang Zhang<sup>1</sup>, Gang Wang<sup>1</sup>, Feng Xiao<sup>2</sup>, Jiajuan Xu<sup>1</sup>, Fang Li<sup>1</sup>, Qinmeng Zeng<sup>1</sup>, Yingying Wei<sup>1</sup> and Jie Xu<sup>1\*</sup>

## Abstract

*Camellia oleifera* is an important woody oil tree in China, in which the flowers and fruits appear during the same period. The endogenous hormone changes and transcription expression levels in different parts of the flower tissue (sepals, petals, stamens, and pistils), flower buds, leaves, and seeds of Changlin 23 high-yield (H), Changlin low-yield (L), and control (CK) *C. oleifera* groups were studied. The abscisic acid (ABA) content in the petals and stamens in the L group was significantly higher than that in the H and CK groups, which may be related to flower and fruit drops. The high N6-isopentenyladenine (iP) and indole acetic acid (IAA) contents in the flower buds may be associated with a high yield. Comparative transcriptome analysis showed that the *protein phosphatase 2C* (PP2C), *jasmonate-zim-domain protein* (JAZ), and *WRKY-related* differentially expressed genes (DEGs) may play an important role in determining leaf color. Gene set enrichment analysis (GSEA) comparison showed that jasmonic acid (JA) and cytokinin play an important role in determining the pistil of the H group. In this study, endogenous hormone and transcriptome analyses were carried out to identify the factors influencing the large yield difference in *C. oleifera* in the same year, which provides a theoretical basis for *C. oleifera* in the future.

**Keywords** *Camellia oleifera*, Transcriptome, Comparative analysis, Endogenous hormone, Yield

## Introduction

*Camellia oleifera* (Fam.: *Theaceae*; Gen.: *Camellia*) is one of the most important woody oil species in China. Unsaturated fatty acids and oleic acid are the main compounds in *C. oleifera* seeds, and the contents of squalene, tea polyphenols, tocopherols, and phytosterols are high [1], conferring this plant important economic and medicinal

values. Currently, research on *C. oleifera* mainly focuses on the primary economic traits of its fruit, oil quality, and the expression of genes related to oil transformation and accumulation during different periods. *C. oleifera* has the characteristic of “holding the child and bearing the child,” that is, when the fruit is ripe in autumn, it flowers. The development of fruits and flowers at the same time causes a high nutritional demand, resulting in a large difference in yield between different individual plants of the same variety. Significant yield variations between different years and individual trees have been observed. The yield of *C. oleifera* is a critical factor constraining the development of the industry. The factors affecting yield include climatic conditions, varietal characteristics, the number of floral buds, tree nutrition, and management practices. Among these factors, having fewer floral

<sup>†</sup>Yayan Zhu and Da Huo have contributed equally to this work and should be considered as co-first authors.

\*Correspondence:

Jie Xu

404715698@qq.com

<sup>1</sup> Guizhou Academy of Forestry, Guiyang, Guizhou 550005, China

<sup>2</sup> Institute for Forest Resources and Environment of Guizhou, Guizhou University, Guiyang 550025, China



buds directly causes the alternate bearing phenomenon in *C. oleifera*. Research on the correlations between the growth of *C. oleifera* flowers, fruits, and spring shoots has revealed that the numbers of floral buds and spring shoots in the same year are positively correlated, but the former is negatively correlated with the fruit yield and fruit-setting rate of that year; trees with a high fruit yield often have significantly fewer floral buds, while those with more floral buds tend to have a relatively lower fruit-setting rate [2]. An increase in fruit production inhibits floral bud differentiation, leading to the alternate bearing phenomenon [3]. Research on the optimal leaf-to-fruit ratio for five-year-old 'Huaxin' *C. oleifera* has shown that different leaf-to-fruit ratios significantly affect the growth, fruit quality, and photosynthetic physiology of the plants, while leaf color varies noticeably under different leaf-to-fruit ratios [4]. Changlin 23 is an improved variety approved by the State Forestry Administration. During this preliminary test, the yield and phenotype of different individual strains of the same variety were inconsistent. In the high-yield group (H group, Changlin 23), the average canopy area yield was 2.28 kg/m<sup>2</sup>, which is higher than that (1.2 kg/m<sup>2</sup>) stipulated in the national standard for code of practice on breeding techniques of *C. oleifera*, the flowers in the canopy area covered less than 20/m<sup>2</sup>, and the leaves were obviously yellow-green. The average yield of 7 a Changlin 23 plant exceeded 3 kg, which was 30.34% ~ 194.87% greater than that of the control [5]. Intraspecific hybridization, i.e., Changlin No. 4 × Changlin 23, resulted in a higher single-fruit weight, an improved seed rate of fresh fruit, a better kernel rate of dry seeds, and higher contents of oil in the kernels and oleic acid in the oil [6]. Comprehensive assessment indicated that Changlin 55, 153, and 23 were the most excellent cultivated varieties [7]. The average yield of the low-yield group (group L) was only 0.1567 kg/m<sup>2</sup>. However, the number of flowers in the crown area covered more than 2500/m<sup>2</sup>, and the leaves were emerald green. Changlin 40 is a stable variety and had the same yield as the control group (CK group). The leaves of this variety are bright green, and the average crown area yield is 1.827 kg/m<sup>2</sup>. The flowers in the crown area are reported to cover more than 2000/m<sup>2</sup>, making it a high-yielding and good pollinator variety [8].

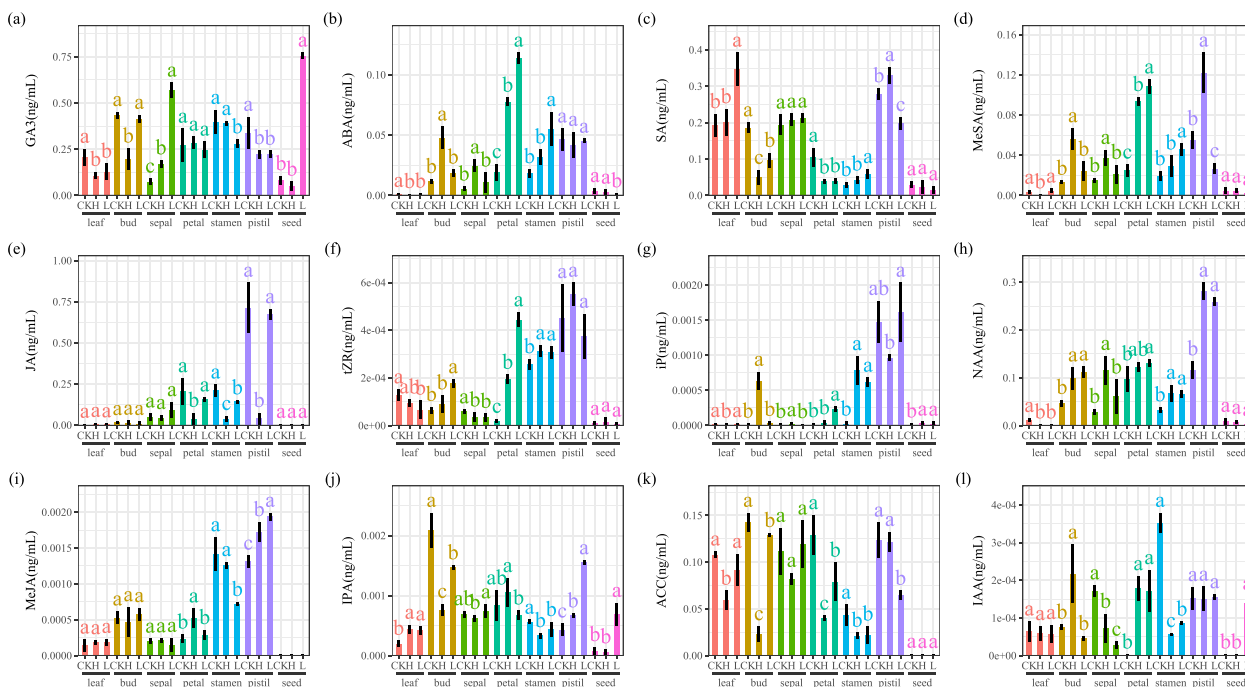
Plant growth depends on synergistic interactions between internal and external signals, and the yield potential of crops is a manifestation of how these complex factors interact, particularly at critical stages of development [9]. RNA-seq has been used to analyze the differentially expressed genes (DEGs) between high- and low-oleic-acid varieties, and the coordinated regulation of multiple genes between the upstream synthesis and downstream utilization of oleic acid ensures

the accumulation of oleic acid in *C. oleifera* with a high content of oleic acid [10]. More gibberellins and steroids accumulated in the field-cultivated Panax ginseng root, which might be responsible for its quick vegetative growth [11]. Auxin, brassinosteroid, jasmonic acid hormone signaling and cellulose, sucrose, and starch synthase metabolic DEGs were mapped on known seed cotton yield quantitative trait loci (QTLs) [12]. A comparative transcriptome analysis of the flower buds and fruits of high-yield and low-yield varieties of *C. oleifera* has shown that some DEGs involved in fatty acid biosynthesis, *FabB*, *FabF*, *FabZ*, and *AccD*, are highly expressed in flower buds and are considered to be related to the high oil yield of fruits [13]. However, it is difficult to determine whether the genes are consistently overexpressed by evaluating only the differences in buds and fruits. To explain the genetic mechanism of yield differences between different varieties, three varieties (high-yield Changlin 23; low yield variety, and Changlin 40 with an average consistent yield, labeled CK; and different parts, including the flower bud, leaf, petal, sepal, stamen, pistil, and seed) were subjected to liquid chromatography–mass spectrometry (LC–MS) to determine the endogenous hormone content. Transcriptome sequencing was performed using Illumina. The purpose of this study was to explore the reasons for the young age of *C. oleifera* and to provide a molecular basis for the clonal breeding and high and stable yield of *C. oleifera*.

## Results

### Analysis of the endogenous hormone content in different parts of different *C. oleifera* varieties

In *C. oleifera*, the gibberellic acid (GA<sub>3</sub>) contents in the leaves and petals of the high- and low-yielding groups H and L were similar, while the GA<sub>3</sub> contents in the flower buds, sepals, and seeds of the high-yielding group were significantly lower than those of the low-yielding group. The GA<sub>3</sub> content in the stamens of the high-yielding group was significantly higher than that of the low-yielding group (Fig. 1a). The abscisic acid (ABA) contents in the pistils of the high-yield, low-yield, and control groups were relatively consistent, and the ABA content in the petals and stamens of group L was significantly higher than those of groups H and CK (Fig. 1b). The salicylic acid (SA) contents in the sepals and seeds in the three plants were relatively consistent (Fig. 1c), and the SA and MeSA contents in the pistils of group H were higher than those of groups L and CK (Fig. 1c,d). In the pistils, the jasmonic acid (JA) content in group H was significantly lower than those in groups L and CK (Fig. 1e). The trans-Zeatin-riboside (tZR) contents in the pistils of the three plants were relatively



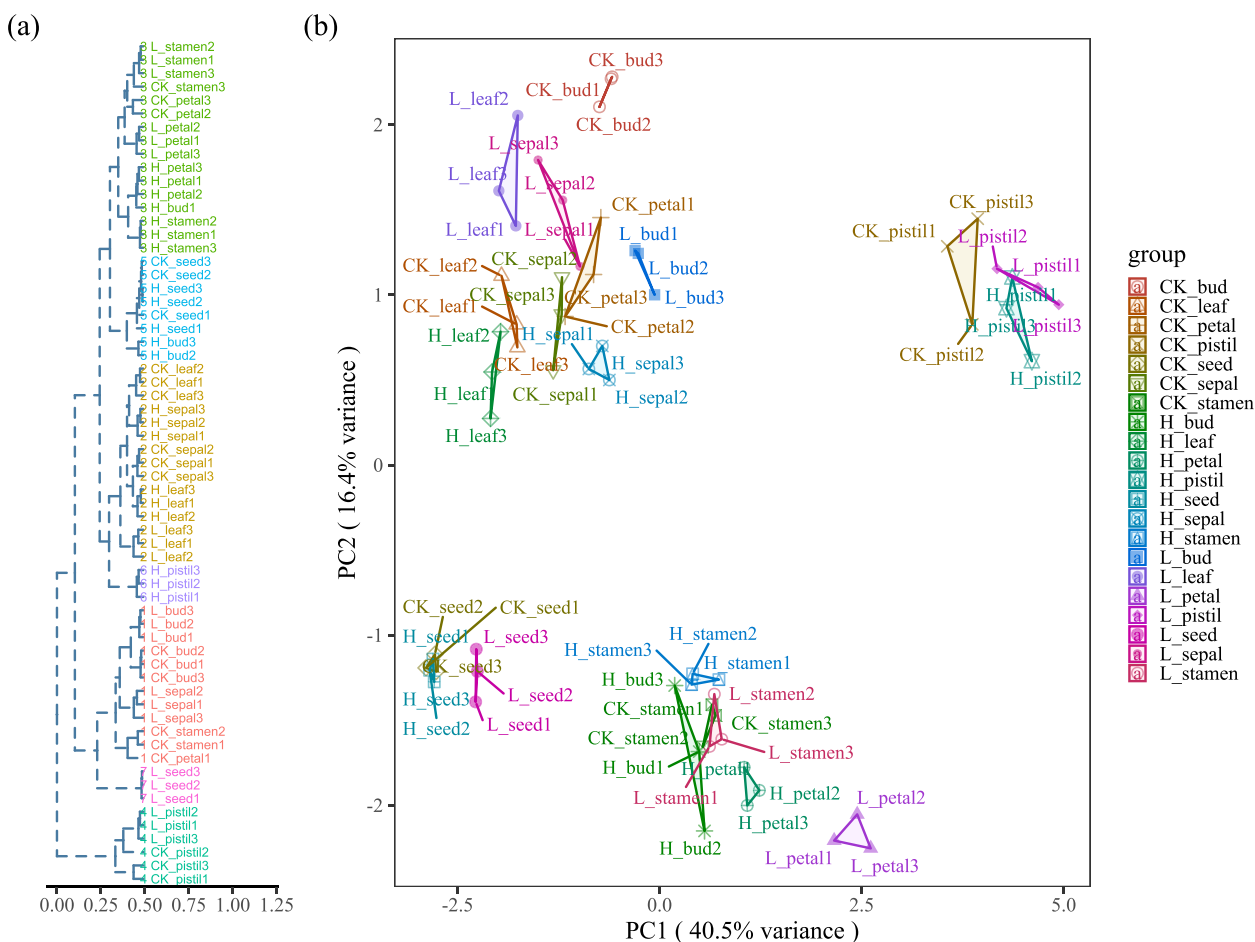
**Fig. 1** Endogenous hormone levels in different tissues of *C. oleifera*. **a–k** represented different hormones, CK represented Changlin 40, H represented Changlin 23, L represented a variety with consistent low yield. A one-way analysis of variance (ANOVA) was performed on different parts, and multiple comparisons were made using the least significant difference (LSD) method. Different letters represented significant differences ( $p < 0.05$ )

consistent (Fig. 1f). In the flower buds, the N6-isopenentenyladenine (iP) content in group H was significantly higher than those in groups L and CK (Fig. 1g). The 1-Naphthaleneacetic acid (NAA) contents in the pistils and flower buds in groups H and L were significantly higher than that in group CK (Fig. 1h). The methyl jasmonate (MeJA) contents in the pistils and flower buds of groups H and L were relatively consistent (Fig. 1i). The 3-Indolepropionic acid (IPA) content in the flower buds in group CK was higher than those in groups H and L (Fig. 1j). The 1-Aminocyclopropanecarboxylic acid (ACC) content of group L was significantly lower than those of groups H and CK (Fig. 1k). In the pistils, the indole acetic acid (IAA) contents of the three plants were relatively consistent, and the IAA content of the flower buds of group H was significantly higher than those of groups L and CK (Fig. 1l). Among the samples (Fig. 2a), the flower buds of groups L and CK were clustered together, while the flower buds of group H were clustered together with the CK and H seeds, indicating that the flower bud parts of the high- and low-yield groups had inconsistent hormone levels. Principal component analysis (PCA) showed that the distance between the samples of the same plant was small, and repeatability was good (Fig. 2b).

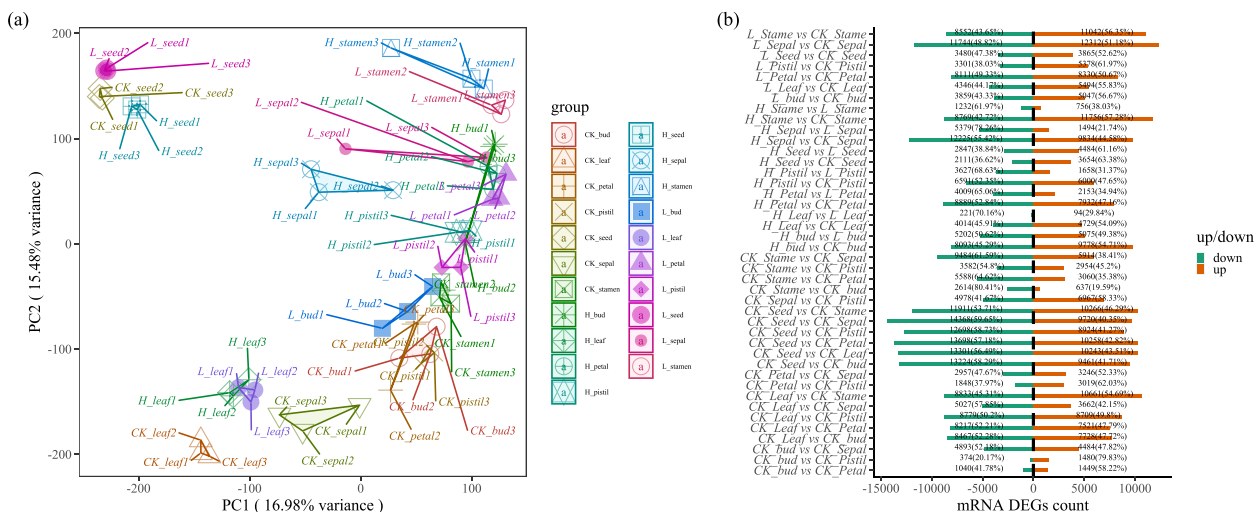
**Transcriptome assembly, annotation, and DEG identification**

A total of 146,594 unigenes with an average length of 875 bp were obtained after Trinity assembly, and redundancy was removed. After comparison with various databases, 92,410 unigenes (63.04%) matched with the non-redundant (NR) protein database, and 69,695 unigenes (47.54%) matched with the Gene Ontology (GO) database. A total of 30,471 genes (20.79%) were matched to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. After clean data comparison and quantification, the PCA of the gene expression of all the samples was performed (Fig. 3a). The distance between the samples in the same group was small, and the samples of the same parts of different plants were close, indicating the good repeatability of the samples.

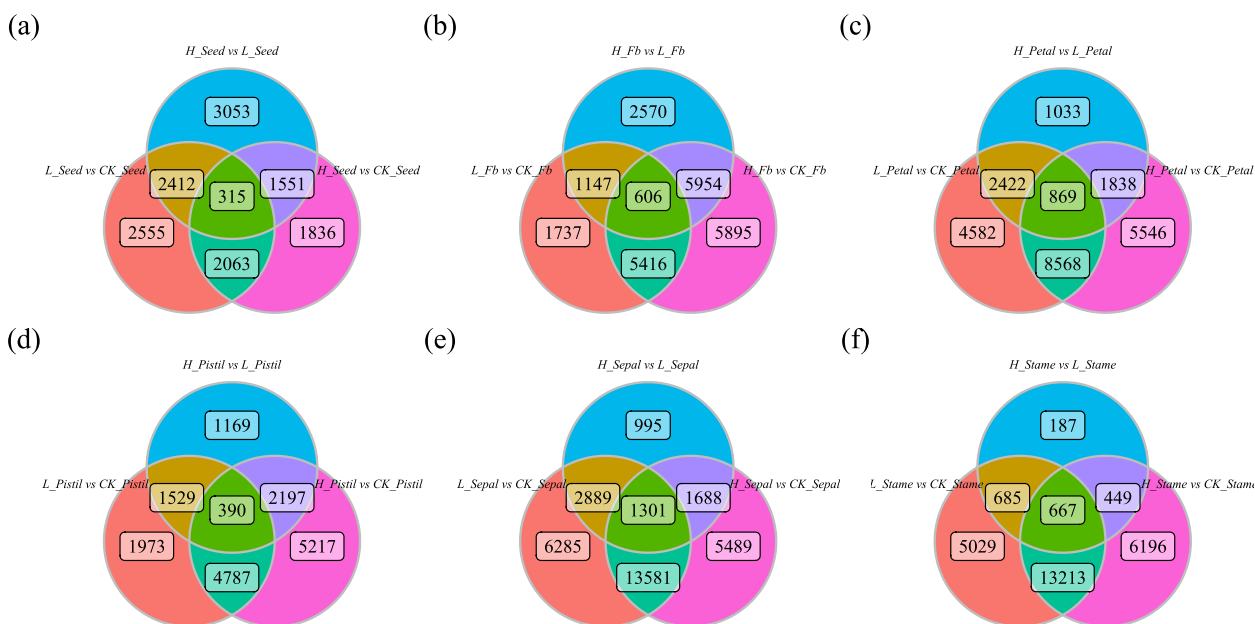
The DEGs were identified for the different plants based on the differential gene screening criteria (Fig. 3b). In the comparison of H\_stame vs L\_stame, 1,232 genes (61.97%) were downregulated, and 756 (38.03%) were upregulated. In the comparison of H\_bud vs L\_bud, 5,202 (50.62%) genes were downregulated, and 5,075 (49.38%) were upregulated. The Venn diagram analysis of the DEGs between different plants showed that there were 64, 315, 606, 869, 1301, 667, and 390 DEGs in the leaves, seeds, flower buds, petals, sepals, stamens, and pistil parts of the different samples, respectively (Fig. 4).



**Fig. 2** Hormone indicator clustering and principal component analysis (PCA) in *C. oleifera* **a** The cluster analysis. **b** The principal component analysis. Note, in a, the clustering method used was "hclust", and the number of clusters was set to 7. In b, different colors and shapes represented different combinations



**Fig. 3** PCA analysis of gene expression in *Camellia oleifera* transcriptome samples and statistics of the number of DEGs **a** PCA analysis of sample gene expression; **b** Statistics of the number of DEGs; Note In **a**, different colors and shapes represented different groups. In **b**, the left side of the graph represented downward adjustment, and the right side of the graph represented upward adjustment

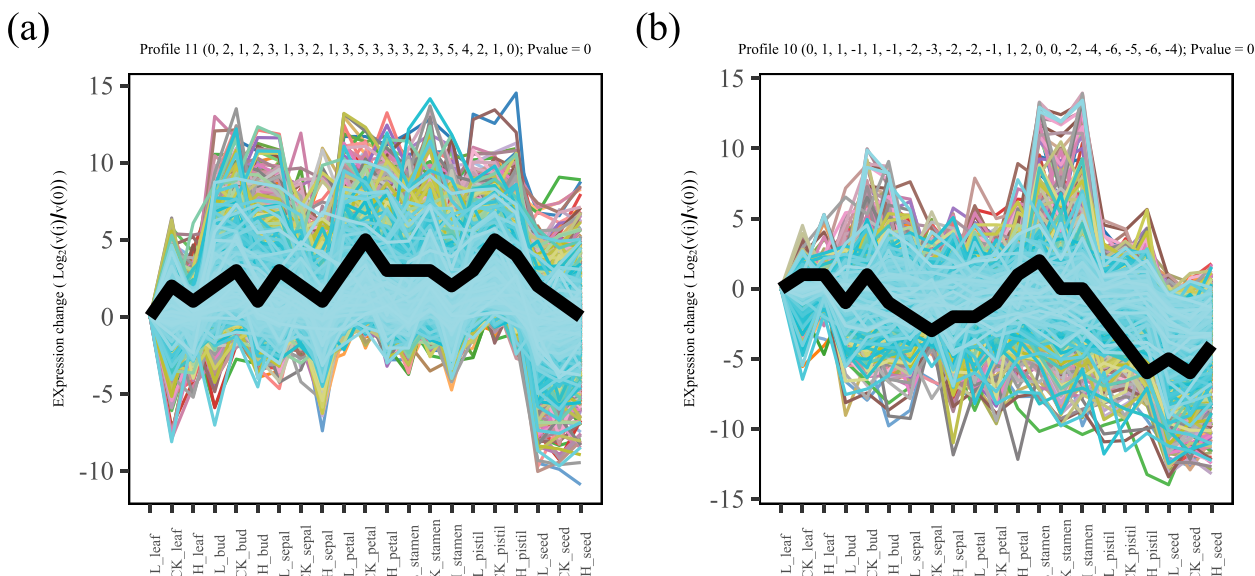


**Fig. 4** Identification of shared and unique DEGs among different combinations of *Camellia oleifera* transcriptome samples. Note: In **a-f**, each subfigure was a venn subfigure

**Short time-series Expression Miner (STEM) and Gene set enrichment analysis (GSEA)**

STEM analysis was performed based on the DEGs between all the plants (Fig. 5), and 14 profiles of significance ( $p < 0.05$ ) were identified. The expression level of Profile 11 (0, 2, 1, 2, 3, 2, 1, 3, 5, 3, 3, 2, 3, 5, 4, 2, 1, 0),

Profile 10 (0, 1, 1, 1, 1, 1, 2, 3, 2, 2, 1, 1, 2, 0) was relatively high in CK\_pistil and CK\_petal. KEGG analysis revealed enrichment in phenylalanine, tyrosine, and tryptophan biosynthesis (ko00400), flavonoid biosynthesis (ko00941), plant hormone signal transduction (ko04075), and other metabolic pathways. The expression level of Profile 10 (0, 1, 1, 1, 1, 1, 2, 3, 2, 2, 1, 1, 2,



**Fig. 5** DEGs STEM analysis of *Camellia oleifera* transcriptome samples. **a** A significant Profile 11 in the STEM analysis of the gene expression pattern; **b** A significant Profile 10 in the STEM analysis of the gene expression pattern. Note: The upper part of the subfigure showed the number and corresponding trend of the modules. Black line represented the trendline. Each line represented the relative expression trend of a gene

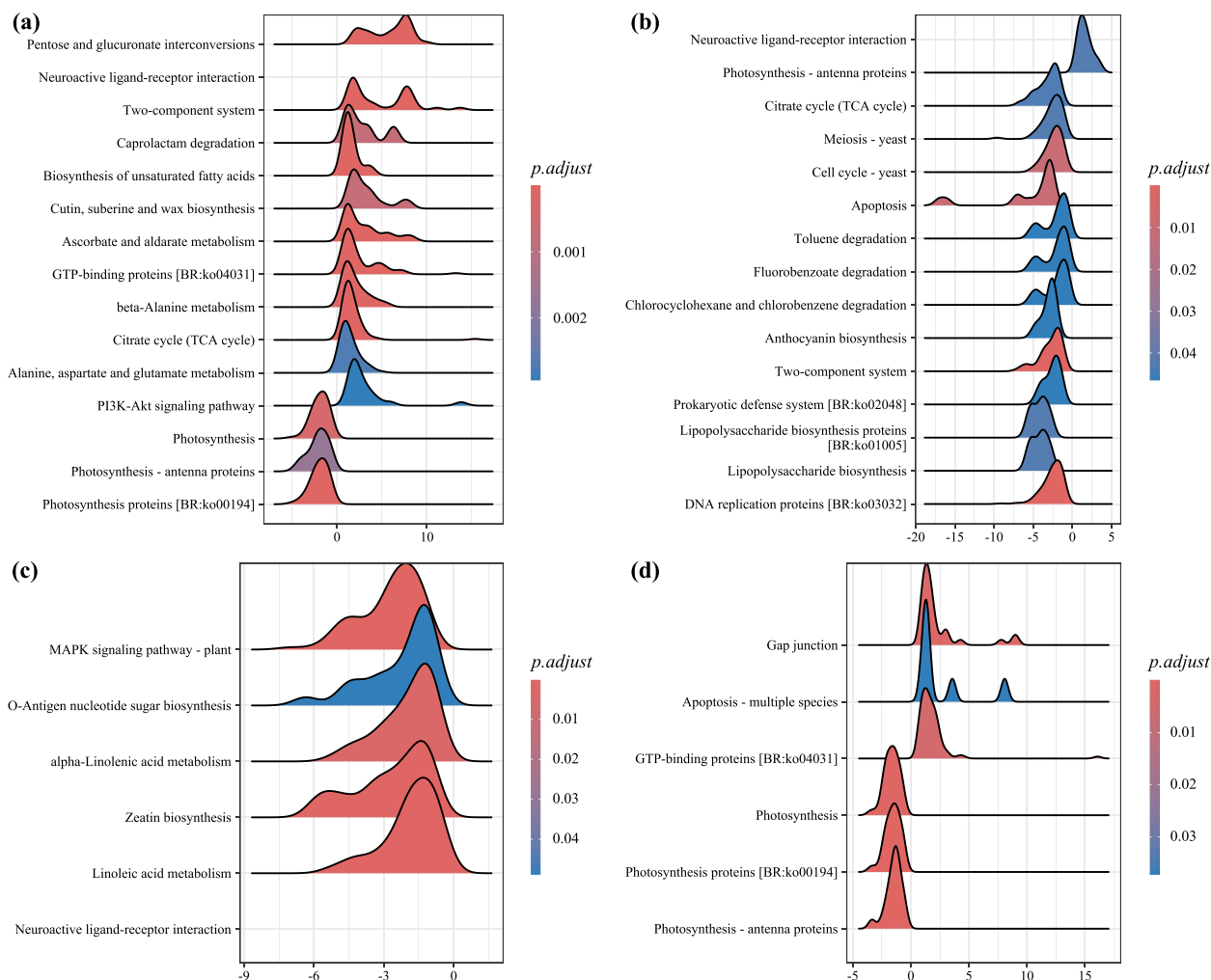
0, 0, 2, 4, 6, 5, 6, 4) was relatively high in the leaves. The expression level in the pistils and seeds was relatively low (Fig. 4c). KEGG enrichment analysis was performed on the gene sets of Profile 10, and the GTP-binding proteins (ko04031), phagosomes (ko04145), lysosomes (ko04142), and carotenoid biosynthesis (ko00906) were enriched. GO enrichment analysis showed that the plastid inner membrane (GO:0009528), chloroplast inner membrane (GO:0009706), and other cellular component (CC) terms were enriched.

The GSEA of the GO terms and KEGG pathways was performed for the different plant parts. In CK\_Stame vs CK\_Pistil (Fig. 6a), pentose and glucuronate interconversions and cutin, suberine, and wax biosynthesis were activated in CK\_Stame. In the H\_Pistil vs CK\_Pistil comparison (Fig. 6b), the TCA cycle and lipopolysaccharide biosynthesis were activated in H\_Pistil. In the

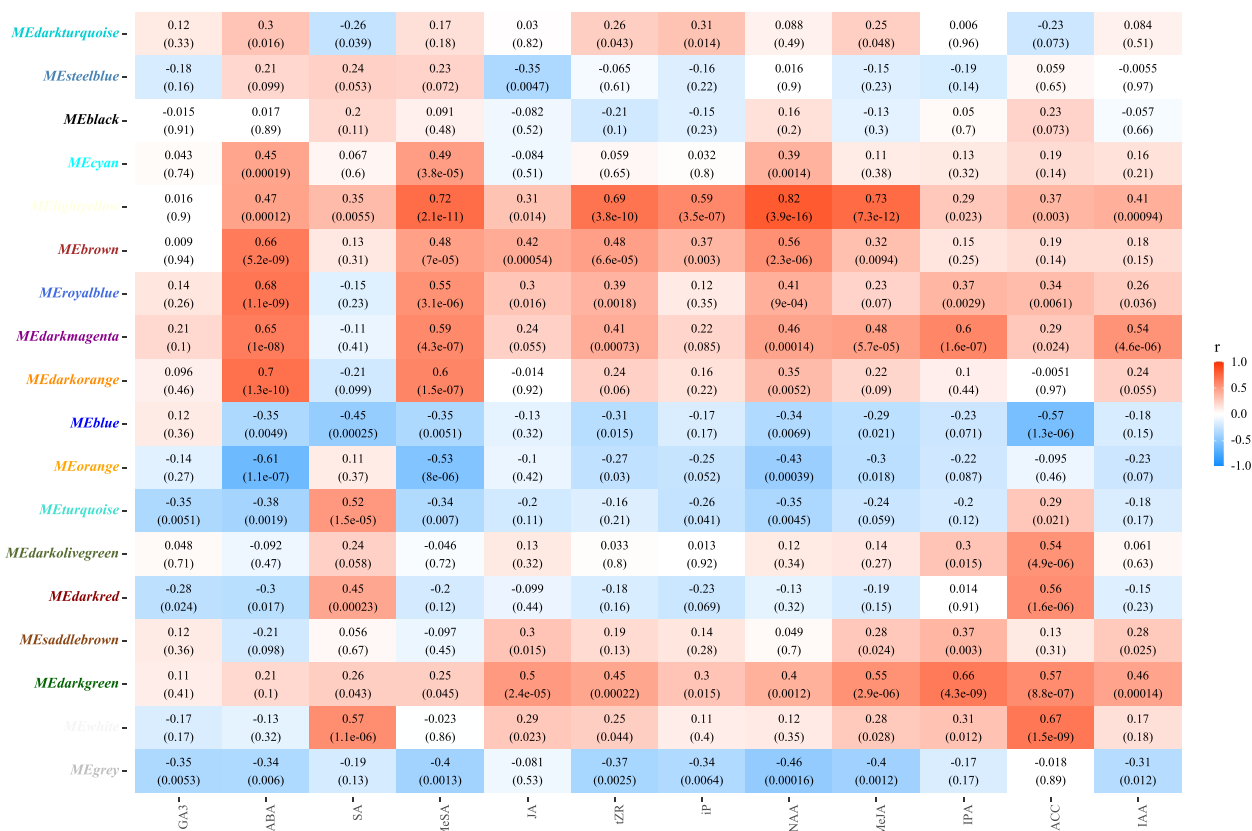
comparison H\_Pistil vs L\_Pistil (Fig. 6c), the plant MAPK signaling pathway, alpha-linolenic acid metabolism, and zeatin biosynthesis were activated in H\_Pistil. In the H\_Stame vs L\_Stame comparison (Fig. 6d), photosynthesis-related pathways (photosynthesis and photosynthesis-antenna proteins) were inhibited in H\_Stame. We performed multi-group GSEA for pollen development and the response to the auxin categories. Compared with CK and L, the pollen development categories were inhibited in H pistil (NES < 1, *p*-value < 0.05).

**Weighted correlation network analysis (WGCNA)**

WGCNA showed that 17 co-expressed gene modules related to these traits could be classified (Fig. 7). The number of genes in the light yellow module was 117, which was similar among the groups. MeSA (*r* = 0.72, *p* = 2.1e - 11), tZR (*r* = 0.69, *p* = 3.8e - 10), NAA (*r* = 0.82,



**Fig. 6** GSEA analysis of different comparative transcriptional combinations in *Camellia oleifera*. **a** GSEA of KEGG pathway between CK\_Stame vs CK\_Pistil; **b** GSEA of KEGG pathway between H\_Pistil vs CK\_Pistil comparison; **c** GSEA of KEGG pathway between H\_Pistil vs L\_Pistil; **d** GSEA of KEGG pathway between H\_Stame vs L\_Stame



**Fig. 7** Module-trait relationships from the WGCNA analysis. Note: The correlation coefficients and corresponding correlation coefficients and p-values between modules and different traits are displayed in each cell; the left panel showed the 17 modules with different colors (The genes in the gray modules were gene sets that cannot be divided into any modules), the colour scale on right showed module-trait correlation from -1 (blue) to 1 (red)

$p=3.9e-16$ ), and MeJA ( $r=0.73$ ,  $p=7.3e-12$ ) were significantly positively correlated. The genes in the light yellow module were highly expressed in the three pistil combinations (H, L, and CK) and H\_bud (Figure S1). GO enrichment analysis showed that lipid transport (GO:0006869) and localization (GO:0010876) and P-type ion (GO:0015662) and P-type transmembrane transporter activities (GO:0140358) were enriched. KEGG analysis indicated that glutathione metabolism (ko00480) was enriched. The number of genes in the dark turquoise module was 79, which was highly expressed in H\_stamen and L\_stamen and was related to stamen development. GO enrichment analysis showed that microtubule cytoskeleton organization (GO:0000226), the structural constituent of the cytoskeleton (GO:0005200), and cytidyltransferase activity (GO:0070567) were enriched.

### Discussion

Many factors affect the yield of *C. oleifera*, including self-incompatibility [14]. When the pollen tube of *C. oleifera* grows to the base of the flower column, distortion

and clustering occur, and the pollen tube cannot reach the embryo sac to complete double fertilization. Self-incompatibility results in a relatively low ratio of *C. oleifera* fruits/seeds to flowers [15]. The declining age of *C. oleifera* and other factors, such as the management mode, also lead to a low yield [16]. Clonal propagation is an important means of shortening the breeding cycle, which is conducive to *C. oleifera* entering the fruiting stage in advance and accelerating the development of the industry. High and stable yield traits are the core of *C. oleifera* breeding [17]. The differences in metabolites between the high- and low-yield groups may be a metabolic marker for the selection of excellent varieties of *C. oleifera*. Using GC-MS, the relative squalene content in the high-yield group was found to be 4.07%, while that in the low-yield group was 35.87% [18]. The flower amount can be used to determine the yield difference in *C. oleifera* at an early stage. The preliminary test of Changlin 23 varieties showed that the high-yield group had a higher yield, fewer flowers, and yellow leaves. The low-yield group showed a lower yield and more flowers. Studies

have shown that the artificial thinning of a large number of flowers can significantly reduce the fruit yield of *C. oleifera* and the ratios of palmitic, palmitoleic, linoleic, and linolenic acids to fatty acids [19]. The endogenous hormones of plants are closely related to the processes of flowering, fruiting, and seed development. In the bud differentiation stage, the levels of IAA and ZR in flowering plants are significantly higher than those in the non-flowering plants of *Glycyrrhiza uralensis* [20]. In *Lycium ruthenicum*, the IAA, ABA, and GA<sub>3</sub> contents in the buds and leaves increase significantly from the open stage to the senescent stage, while the ZR content decreases significantly. The increase in soluble sugar and sucrose in the buds and leaves is believed to enhance the energy base, and the higher the content, the more conducive it is to flower bud differentiation [21]. Different flower tissues exhibit different hormonal changes during flower development. Before flowering, the cytokinin and auxin levels increase in most cases in petals, while they decrease in the middle and late stages of aging [22]. When pollen tubes enter the embryo sac, high levels of GA<sub>3</sub> can inhibit the growth of pollen tubes and release gametophytes, but high levels of ZR can promote the normal growth of pollen tubes in the stigma. High levels of ABA can also promote the growth and fertilization of pollen tubes [23]. The leaves, flower buds, sepals, petals, stamens, pistils, and seeds of Changlin 23 were collected from different tissues during the same period to determine their endogenous hormone content. The GA<sub>3</sub> contents in the leaves and petals of the high- and low-yielding groups H and L were similar, while the GA<sub>3</sub> content in the flower buds, sepals, and seeds of the high-yielding group was significantly lower than that of the low-yielding group. The GA<sub>3</sub> content in the stamens of the high-yielding group was significantly higher than that of the low-yielding group (Fig. 1a). Priming the induction of flowers requires a high level of energy metabolism [24]. The comparison of endogenous hormone contents between two early-spring-flowering plants (*Hylomecon japonica* and *Smilacina japonica*) and two early-spring-flowerless plants (*Cimicifuga foetida* and *Veratrum nigrum*) showed that the IAA content in the flowering plants in early spring was significantly lower than that in the non-flowering plants, while the GA<sub>3</sub> and ABA contents were higher. The contents of D-ribonium, L-valine, L-isoleucine, alanine, and other compounds involved in the key pathways of flowering plants in early spring were significantly reduced, indicating that flowering plants in early spring consume a lot of energy to regulate flowering [25]. The high level of GA<sub>3</sub> in the *C. oleifera* flower buds may be related to the number of blooms. A high ABA content is associated with falling flowers and fruits [26, 27]. The ABA contents in the pistils of the high-yield,

low-yield, and control groups were relatively consistent, and the ABA content in the petals and stamens of group L was significantly higher than those of groups H and CK (Fig. 1b), which may be related to flower and fruit drops. The flower bud IP content of group H was significantly higher than those of groups L and CK (Fig. 1g), while the IAA contents of the three combined pistils were relatively consistent. The IAA content of the flower buds of group H was significantly higher than those of groups L and CK (Fig. 1f). This suggests that the high IP and IAA contents in the flower buds may be associated with a high yield. In conclusion, the high GA<sub>3</sub> content during *C. oleifera* flower bud differentiation may be correlated with the number of blooms, and the high IP and IAA contents during flower bud differentiation may be correlated with a high yield. The ABA content in petals and stamens contributes to flower shedding and is related to flower and fruit drops.

A transcriptome is used to identify the DEGs associated with the target trait among different varieties or traits. Based on the transcriptome assembly, 146,594 genes with an average length of 875 bp were obtained after redundancy was removed. After the expression was quantified, all the samples were analyzed using PCA (Fig. 3a). The distance between the samples in the same group and the samples of the same parts of different plants was small, indicating the good repeatability of the samples. Based on DEG screening (Fig. 3b), in the H\_stame vs L\_stame comparison, 1232 genes (61.97%) were downregulated, and 756 (38.03%) were upregulated. In the H\_bud vs L\_bud comparison, 5,202 (50.62%) genes were downregulated and 5,075 (49.38%) were upregulated. The Venn diagram analysis of DEGs between different comparisons showed that there were 64, 315, 606, 869, 1301, 667, and 390 DEGs in the leaves, seeds, buds, petals, sepals, stamens, and pistils of the different plants, respectively (Fig. 4). CYPs might play critical roles in the determination of leaf color in *Gleditsia sinensis* via iron ion and tetrapyrrole binding [28]. The sets and number of DEGs in the leaves of the different plants were the lowest. ABCG39 (ABC transporter G family member 39-like), WRKY31, WRKY24, WARY70, JAZ (jasmonate-zim-domain protein), PP2C (protein phosphatase 2C), and other gene sets were annotated. The transcriptome has been used to analyze the expression of genes related to leaf aging in autumn *Ginkgo biloba*, and the expression levels of most abscisic acid-, jasmonic acid-, and autophagy-related genes and the WRKY and NAC genes were increased [29]. Transcriptome and metabolome data on the autumn leaf transformation of *Quercus dentata* have shown that argonin-3-O-glucoside, anthocyanin-3-o-arabinoside, and anthocyanin-3-o-glucoside are the main pigments involved in leaf color transformation,



and the MYB-bHLH-WD40 (MBW) transcriptional activation complex is the core of anthocyanin biosynthesis regulation [30]. Abscisic acid-related genes *PP2C*, *jasmonate-zim-domain protein*, and *WRKY-related* DEGs may play an important role in the leaf color of *C. oleifera*.

The genes *TFL1*, *SUP*, *AP1*, *CRY2*, *CUC2*, *CKX1*, *TAA1* and *PIN1* are associated with reproductive phase transition in *Jatropha curcas* [31]. After the initiation of flower bud differentiation, nitrogenous compounds (such as alkaline amino acids and heterocyclic imino acids, especially arginine, Pro, Ala, and l-cysteine) and a higher content of soluble sugars favored morphological differentiation [32]. The MADS-box gene family is involved in flower development. The ABCDE model of flower development indicates [33] that class A genes [*AETALA1* (AP1) and *APETALA2* (AP2)] determine sepal characteristics in the outer domain of the flower meristem. The B genes [*APETALA3* (AP3) and *PISTILLATA* (PI)] combine with the A genes to regulate petals; the classes B and C genes [*AGAMOUS* (AG)] regulate the stamen, while the class C genes regulate the fourth pistil alone [34]. Time-Ordered Gene Co-expression Network analysis revealed that *MADS2* and *MADS26* may play important roles in the development of female and male *Trachycarpus fortune* plants, respectively [35]. With the use of transcriptome transcription factor annotation and screening, 50 genes were annotated as members of the MADS-box gene family. *Agamous-like MADS-box protein TM6* (*MADS6*, TRINITY\_DN4862\_c1\_g2) (Figure S2), belonging to the AP3 class B genes, was highly expressed in the bud, petal, and stamen tissues [36]. These results suggest that *MADS6* may play a significant role in stamen development.

WGCNA can screen out the key regulatory genes by association between gene sets and phenotypic traits and has been used in several species. The key nodes of the transcriptomes of different colors of *Chrysanthemum morifolium* were screened using WGCNA, and *CmMYB305*, *CmMYB29*, *CmRAD3*, *CmbZIP61*, *CmAGL24*, and *CmNAC1* may be related to the carotene metabolic pathways and plasmid development [37]. Conducting the WGCNA of the transcriptome of *C. oleifera* for different light durations, eight key node hub genes, *GI*, *AP2*, *WRKY65*, *SCR*, *SHR*, *PHR1*, *ERF106*, and *SCL3*, were identified and considered to be related to photoperiod-sensitive genes [38]. The high expression level of *GA20ox oxidase* (gibberellin 20 oxidase), a key component of gibberellin biosynthesis, the low expression level of gibberellic acid insensitive (gibberellic acid insensitive), and the increased expression of *9-cis-epoxycarotenoid dioxygenase* (NCED) may contribute to the accumulation of oil in *C. oleifera* fruit [39]. The *CoGA20ox1* transgene of *C. oleifera* results in important

phenotypes, such as seed enlargement and early flowering [40]. Transcriptome-based analysis revealed that the DEGs between pistillate and staminate flowers (SFs) in phase I (the differentiation of male and female primordia) (involved in ABA, auxin (AUX), CK, ethylene (ET), and GA biosynthesis and signaling) showed higher expression levels in males than in females in general, whereas the DEGs involved in the JA and SA pathways displayed opposite expression patterns [41]. In vitro experiments further verified that *MYC2-1* alone promotes the expression of FT, and *JAZ1-3* and *MYC2-1* play key roles in the differentiation of female flower buds of chestnut [42]. Floral bud differentiation is related to sex determination; the genes *DELLA*, *MYC2* and *CYCD3* might be related to the stamen abortion and female flower sex determination during the inflorescence bud stage of *Jatropha nigrovien-srugosus* [43]. WGCNA showed that the number of genes in the light yellow module was 117, which was similar among the groups. MeSA ( $r=0.72$ ,  $p=2.1e-11$ ), tZR ( $r=0.69$ ,  $p=3.8e-10$ ), NAA ( $r=0.82$ ,  $p=3.9e-16$ ), and MeJA ( $r=0.73$ ,  $p=7.3e-12$ ) were significantly positively correlated. The genes in the light yellow module were highly expressed in the pistils (H, L, and CK) and H\_bud. In the light yellow module, *ARF6* (*Auxin response factor 6*, TRINITY\_DN2955\_c2\_g2, and TRINITY\_DN2955\_c2\_g3), *Calcium-transporting ATPase 5* (TRINITY\_DN34809\_c0\_g1), *Calcium-transporting ATPase 10* (TRINITY\_DN9394\_c0\_g1), and *Calcium-transporting ATPase 1* (TRINITY\_DN152457\_c0\_g1) were annotated. GO enrichment analysis showed that lipid transport (GO:0006869) and localization (GO:0010876) and the P-type ion (GO:0015662) and transmembrane transporter activities (GO:0140358) were enriched. In H\_Pistil vs L\_Pistil (Fig. 6c), the plant MAPK signaling pathway, alpha-Linolenic acid metabolism, and zeatin biosynthesis were activated in H\_Pistil. These results indicate that JA and cytokinin play important roles in the pistils of high-yielding *C. oleifera*.

## Conclusions

A low yield limits the development of the *C. oleifera* industry, and the yields of the different varieties are significantly different. High-yield, low-yield, and control CK groups were selected as experimental objects to study the endogenous hormone changes and transcriptional expression levels of seven different tissues, such as the leaves, petals, and stamens. The ABA content in the petals and stamens in the L group was significantly higher than those in the H and CK groups, which was related to flower and fruit drops. The higher IP and IAA contents in the buds may be associated with a high yield. Comparative transcriptome analysis showed that the ABA-related genes *PP2C* and *JAZ* and the *WRKY-related* DEGs may

play important roles in leaf color. GSEA comparison showed that JA and cytokinin play important roles in the pistils of high-yielding *C. oleifera*.

## Materials and methods

### Materials

The materials were collected from Yuping County, Tongren city, Guizhou Province, China, which is one of the main production areas of *C. oleifera*. Ten-year-old *C. oleifera* varieties with relatively stable and healthy growth were selected, and collection was completed between 9 and 10 a.m. Different parts, including the flower buds, leaves, petals, sepals, stamens, pistils, and seeds, were collected (Figure S3), washed in distilled water, and quickly frozen in liquid nitrogen for at least half an hour. The samples were then stored in an ultra-low-temperature freezer at  $-80^{\circ}\text{C}$  for future use.

### Determination of endogenous hormones using LC-MS/MS

Samples weighing 50–100 mg were added to 1 mL 50% acetonitrile solution, ground at  $4^{\circ}\text{C}$  for 6 min, ultrasonicated at a low temperature for 30 min ( $5^{\circ}\text{C}$ , 40 kHz), left at  $4^{\circ}\text{C}$  for 30 min, and centrifuged at 14,000 rcf for 15 min at  $4^{\circ}\text{C}$ . The supernatant was removed and purified using an HLB column. The perforated liquid and eluent were collected and blown dry in a 2 mL centrifuge tube with nitrogen. Then, 50  $\mu\text{L}$  of 30% acetonitrile aqueous solution was added and swirled to mix it in. The samples were ultrasonicated and centrifuged at 14,000 rcf for 15 min at  $4^{\circ}\text{C}$ , and 40  $\mu\text{L}$  of the supernatant was collected in a sample vial. LC-ESI-MS/MS (UHPLC-Qtrap) was used to conduct the qualitative and quantitative detection of target substances in the samples. Chromatography was performed on a Waters C18 (2.1 mm  $\times$  150 mm, 1.7  $\mu\text{m}$ ) liquid chromatography column with an ExionLC AD system. The column temperature was  $50^{\circ}\text{C}$ . The sample size was 10  $\mu\text{L}$ . Mobile phase A consisted of 0.01% formic water, and mobile phase B comprised 0.01% formic acetonitrile. The mass spectrum conditions were as follows: AB SCIEX QTRAP 6500+ with positive/negative mode detection; curtain gas (CUR), 35; collision gas (CAD), medium; ion spray voltage (IS), +5500/−4500; temperature (TEM), 550; ion source gas 1 (GS1), 50; and ion source gas 2 (GS2), 50. With the use of AB Sciex quantitative software OS, the ion fragments were automatically identified and integrated using the default parameters, and manual inspection was also performed. With the mass spectrum peak area of the analyte as the vertical coordinate and the concentration of the analyte as the horizontal coordinate, a linear regression standard curve was drawn, and the concentration was calculated according to the standard curve.

### RNA extraction and library construction

The total RNA was extracted from all the samples using Trizol (Invitrogen, Waltham, MA, USA). After purification, the integrity and concentration of the total RNA were determined. Nanodrop2000 was used to detect the concentration and purity of the RNA, agarose gel electrophoresis was used to detect the RNA integrity, and Agilent5300 was used to determine the RIN value. An Illumina Truseq<sup>TM</sup> RNA sample prep kit was used to construct the library. The 3' end of eukaryotic mRNA had a poly tail structure. Magnetic beads with Oligo (dT) were used to pair the A–T bases with poly(A), and a fragmentation buffer was added to the mRNA at random, generating small fragments of about 300 bp. Single-stranded cDNA was synthesized using random primers, and then two strands were synthesized to form a stable double-stranded structure. End Repair Mix was added to complete the flat end, and an A base was added to the 3' end to connect the Y-shaped joint. After adapter connection, the product was purified, and the fragments were sorted. The sorted product was amplified using PCR, and the final library was purified. The cDNA libraries were sequenced using the Illumina HiSeq high-throughput sequencing platform.

### Transcriptome assembly, annotation, and analysis of differentially expressed genes

Fastp v0.20.1 software was used to control the quality of the original data [44]. The transcriptome assemblies were generated using Trinity software [45]. The assembled sequences were merged and clustered using CD-HIT to remove redundancy [46]. The unigenes of the redundant data were used as the dataset for subsequent analysis. The unigenes and NR (<http://www.ncbi.nlm.nih.gov/>), Gene Ontology (GO), SwissProt, KEGG (<http://www.genome.jp/kegg/>), Clusters of Orthologous Groups of Proteins (COGs), and eggNOG databases were used for comparison and annotation. Bowtie2 v2.4.3 was used for comparison, and RSEM v1.2.12 [47] was used for quantification based on TPM expressions. Principal component analysis (PCA) was performed based on the expression levels of all the samples. The DEGs were identified using the R package DEseq2 v1.40.1 [48], and the screening criteria were set as  $|\log_2\text{FC}| > 1$  and  $p\text{-value} \leq 0.05$ . The R package clusterProfiler v4.8.1 was used to conduct the GO and KEGG enrichment analyses of the DEGs [49]. A short time-series Expression Miner (STEM) was used for gene expression trend analysis [50], and the number of modules was set to 50.

### WGCNA analysis

The R package WGCNA v1.72.1 [51] was used for weighted gene co-expression network analysis (WGCNA). The first 5000 genes were screened according to the median absolute deviation, and the parameters were set as follows: power = 10, minModuleSize = 30, and MEDissThres = 0.25.

### Quantitative real-time PCR (qRT-PCR) analysis

qRT-PCR analysis was performed to verify the expression of RNA-seq data. The total RNA (isolated from RNA-Seq samples) was used to synthesize the first-strand cDNA. The real-time CFX96 Touch Real-Time PCR System (Bio-Rad, USA) was used to perform qRT-PCR analysis. Each 50  $\mu$ L reaction included 25  $\mu$ L of 2X SYBR<sup>®</sup> Green PCR Master Mix, 1  $\mu$ L of forward/reverse primer (10  $\mu$ M), 2.0  $\mu$ L of cDNA template, and 21  $\mu$ L of RNase-free water. The PCR conditions were as follows: preheating at 95  $^{\circ}$ C for 30 s, 40 cycles of heat denaturation at 95  $^{\circ}$ C for 5 s, and annealing at 60  $^{\circ}$ C for 34 s. The gene relative expression levels were calculated using the  $2^{-\Delta\Delta CT}$  method [52]. Three independent biological replicates were created. The *GAPDH* gene was used as the reference (Table S1).

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-024-10795-0>.

Supplementary Material 1: Table S1. The primer list of qRT-PCR.

Supplementary Material 2: Figure S1. Histogram of gene expression trends within each module of WGCNA.

Supplementary Material 3: Figure S2. Circular heatmap of MADS box gene expression in *C. oleifera*.

Supplementary Material 4: Figure S3. Sampling schematic diagram. a: high-yield Changlin 23; b: low-yielding; c: Changlin 40 with an average consistent yield; d-f: different part of tissues.

Supplementary Material 5: Figure S4. Comparison of qRT-PCR results with RNA seq.

### Acknowledgements

Not applicable.

### Authors' contributions

Conceptualization, YYZ and DH; methodology, YYZ, DH and JJX; software, FX, FL, GW, JX; validation, YYZ, QMZ and MGZ; formal analysis, YYZ and DH; investigation, YYZ, MGZ; resources, MGZ, FX, YYW, QMZ and JJX; data curation, YYZ, FX; writing—original draft preparation, YYZ, FX; writing—review and editing, JX; visualization, YYZ and DH; supervision, YYZ; project administration, YYZ; funding acquisition, JX. All authors have read and agreed to the published version of the manuscript.

### Funding

This research was funded by the Science and Technology Department of Guizhou Province (grant number: Qiankehejichu-ZK[2024]zhongdian091, Qiankehingtairencai[2019]5643, Qiankehezhicheng[2022]yiban166 and Qiankefuqi [2018]4003) and Forestry Bureau of Guizhou Project (grant number: Gui[2023]TG32 and Telinyan 2020–15).

### Availability of data and materials

The raw reads generated via Illumina sequencing were deposited in the NCBI SRA database (BioProject accession number: PRJNA973966).

### Declarations

#### Ethics approval and consent to participate

All plants samples were collected on improved variety base, we had obtained the permissions from that unit. Experimental research and field studies on plants including the collection of plant material are comply with relevant guidelines and regulation. All samples were identified by Yayan Zhu and Jie Xu, and all voucher specimens were deposited in the Guizhou Academy of Forestry (voucher ID numbers: CoH23202212 for Changlin 23 high-yield, CoL202212 for Changlin low-yield, CoCK40202212 for Changlin stable variety).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 8 April 2024 Accepted: 11 September 2024

Published online: 20 September 2024

### References

- Zhang F, Zhu F, Chen B, Su E, Chen Y, Cao F. Composition, bioactive substances, extraction technologies and the influences on characteristics of *Camellia oleifera* oil: a review. *Food Res Int.* 2022;156:111159.
- Chen L, Chen Y, Xu Y, Peng Y, Peng S. Research on correlation of flower bud, fruit and spring shoot growth of *Camellia oleifera* Abel. *J Cent South Univ Forest Technol.* 2018;38:1–5.
- Jia T, Su S, Ma L, Su Q. Response of flower bud differentiation to different sink-source relationships in *Camellia oleifera*. *J Northeast For Univ.* 2018;46(9):50–3.
- Zhang X. Study on the optimum leaf-fruit ratio of 5-year-old *Camellia oleifera* "Huaxin." Master: Central South University of Forestry and Technology; 2023.
- Xiangzhen F, Zilong Z, Chenyu Y, Zhiheng X, Zhifeng Z. Effects of Changlin *Camellia oleifera* varieties on yield and quality in the mountainous areas of Suichang county. *J Sichuan For Sci Technol.* 2020;41(2):69–73.
- Jing-yu C, Kai-liang W, Xiao-hua Y, Jian-hua T, Ping L. Genetic analysis of the fruit and oil related traits on hybrid offspring of nested mating of *Camellia oleifera*. *For Res.* 2023;36(1):1–10.
- Xue-yong H, Chao W, Jian-xun L, Yun-jie G, Fu-rong L. Comprehensive evaluation on the characteristics of different *Camellia oleifera* varieties cultivated in Sichuan Basin. *J Sichuan For Sci Technol.* 2018;39(5):13–6.
- Chen LJ, Zhang Y, Liao GW, Cheng YM, Xiao C, Zhang H, Wang SQ, Hui W, Du KB, Shu CQ. Pollination varieties selection of 3 main cultivated *Camellia oleifera* in Hubei Province. *Chinese J Oil Crop Sci.* 2022;44(5):1030–6.
- Sato Y, Antonio B, Namiki N, Motoyama R, Sugimoto K, Takehisa H, Minami H, Kamatsuki K, Kusaba M, Hirochika H. Field transcriptome revealed critical developmental and physiological transitions involved in the expression of growth potential in japonica rice. *BMC Plant Biol.* 2011;11:1–15.
- Wu B, Ruan C, Han P, Ruan D, Xiong C, Ding J, Liu S. Comparative transcriptomic analysis of high-and low-oil *Camellia oleifera* reveals a coordinated mechanism for the regulation of upstream and downstream multigenes for high oleic acid accumulation. *3 Biotech.* 2019;9:1–19.
- Fan H, Li K, Yao F, Sun L, Liu Y. Comparative transcriptome analyses on terpenoids metabolism in field-and mountain-cultivated ginseng roots. *BMC Plant Biol.* 2019;19:1–15.
- Shahzad K, Zhang X, Guo L, Qi T, Bao L, Zhang M, Zhang B, Wang H, Tang H, Qiao X. Comparative transcriptome analysis between inbred and hybrids reveals molecular insights into yield heterosis of upland cotton. *BMC Plant Biol.* 2020;20:1–18.

13. Xie Y, Wang X. Comparative transcriptomic analysis identifies genes responsible for fruit count and oil yield in the oil tea plant *Camellia chekiangoleosa*. *Sci Rep*. 2018;8(1):6637.
14. Li C, Long Y, Lu M, Zhou J, Wang S, Xu Y, Tan X. Gene coexpression analysis reveals key pathways and hub genes related to late-acting self-incompatibility in *Camellia oleifera*. *Front Plant Sci*. 2023;13:1065872.
15. He Y, Song Q, Chen S, Wu Y, Zheng G, Feng J, Yang Z, Lin W, Li Y, Chen H. Transcriptome analysis of self-and cross-pollinated pistils revealing candidate unigenes of self-incompatibility in *Camellia oleifera*. *J Hortic Sci Biotechnol*. 2020;95(1):19–31.
16. Liu J, Long W, Chen S, Jin X, Nan H. Study on the relationship between the present situation of *Camellia oleifera* forest resources and the proportion of scale management in Jiangxi Province. *J Fujian For Sci Technol*. 2022;49(4):113–9.
17. Kailiang W, Xiaohua Y, Huadong R, Pingan Z. Evaluation of yield and stability of *Camellia oleifera* Abel clones. *J Nanjing For Univ (Natural Science Edition)*. 2016;40(1):174–8.
18. Cao H, He Z, Li Z, Lu S, Zheng D, Chen Y, Lai B. Identification of Potential Metabolic Markers for the High and Low Yield Clones of *Camellia oleifera*. *Journal of Fujian Forestry Science and Technology*. 2021;48(4):8–12.
19. Ye T, Liu X, Liang X, Zhu X, Bai Q, Su S. Flower Thinning Improves Fruit Quality and Oil Composition in *Camellia oleifera* Abel. *Horticulturae*. 2022;8(11):1077.
20. Yan B, Hou J, Cui J, He C, Li W, Chen X, Li M, Wang W. The effects of endogenous hormones on the flowering and fruiting of *Glycyrrhiza uralensis*. *Plants*. 2019;8(11):519.
21. Guo Y, An L, Yu H, Yang M. Endogenous hormones and biochemical changes during flower development and florescence in the buds and leaves of *Lycium ruthenicum* Murr. *Forests*. 2022;13(5):763.
22. Arrom L, Munné-Bosch S. Hormonal changes during flower development in floral tissues of *Lilium*. *Planta*. 2012;236:343–54.
23. Ming D, Xudong Y, Fanhua W. Dynamic changes of endogenous hormones in self-pollinated and cross-pollinated pistils of two *Camellia* species in Hainan. *Journal of Tropical Biology*. 2023;14(2):173–7.
24. Guo H, Zhong Q, Tian F, Zhou X, Tan X, Luo Z. Transcriptome analysis reveals putative induction of floral initiation by old leaves in tea-oil tree (*Camellia oleifera* changlin53). *Int J Mol Sci*. 2022;23(21):13021.
25. Yan X, Liu J, Wu KX, Yang N, Pan L-B, Song Y, Liu Y, Tang ZH. Comparative analysis of endogenous hormones and metabolite profiles in early-spring flowering plants and unflowered plants revealing the strategy of blossom. *J Plant Growth Regul*. 2022(41):2421–34
26. Nyoman Rai I, Poerwanto R, Darusman L, Purwoko B. Flower and fruit ABA, IAA and carbohydrate contents in relation to flower and fruit drop on mangosteen trees. In: IV International symposium on tropical and subtropical fruits 975: 2008. 2008. p. 323–8.
27. Shao-ze L, Zhi-gang Z, Ping C, Lu Y, Hong L, Ya-qin R, Dong-liang W, Zi-long W. The change of endogenous hormone of apricot tree in flowering and setting period. *Xinjiang Agric Sci*. 2019;56(2):258.
28. Wu C, Yang X, Feng L, Wang F, Tang H, Yin Y. Identification of key leaf color-associated genes in *Gleditsia sinensis* using bioinformatics. *Hortic Environ Biotechnol*. 2019;60:711–20.
29. Li W, Wang L, He Z, Lu Z, Cui J, Xu N, Jin B, Wang L. Physiological and transcriptomic changes during autumn coloration and senescence in *Ginkgo biloba* leaves. *Hortic Plant J*. 2020;6(6):396–408.
30. Wang WB, He XF, Yan XM, Ma B, Lu CF, Wu J, Zheng Y, Wang WH, Xue WB, Tian XC. Chromosome-scale genome assembly and insights into the metabolome and gene regulation of leaf color transition in an important oak species. *Quercus dentata* *New Phytol*. 2023;238:2016–32.
31. Gangwar M, Sood H, Chauhan RS. Genomics and relative expression analysis identifies key genes associated with high female to male flower ratio in *Jatropha curcas* L. *Mol Biol Rep*. 2016;43:305–22.
32. Zhang W, Li J, Zhang W, Njie A, Pan X. The changes in C/N, carbohydrate, and amino acid content in leaves during female flower bud differentiation of *Juglans sigillata*. *Acta Physiol Plant*. 2022;44:1–12.
33. Theißen G. Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol*. 2001;4(1):75–85.
34. Sun W, Huang W, Li Z, Song C, Liu D, Liu Y, Hayward A, Liu Y, Huang H, Wang Y. Functional and evolutionary analysis of the AP1/SEP/AGL6 superclade of MADS-box genes in the basal eudicot *Epimedium sagittatum*. *Ann Bot*. 2014;113(4):653–68.
35. Xiao F, Zhao Y, Wang X, Mao Y, Jian X. Comparative transcriptome analysis of dioecious floral development in *Trachycarpus fortunei* using Illumina and PacBio SMRT sequencing. *BMC Plant Biol*. 2023;23(1):536.
36. Gioppato HA, Dornelas MC. When Bs are better than As: The relationship between B-Class MADS-Box gene duplications and the diversification of perianth morphology. *Tropic Plant Biol*. 2019;12:1–11.
37. Lu C, Pu Y, Liu Y, Li Y, Qu J, Huang H, Dai S. Comparative transcriptomics and weighted gene co-expression correlation network analysis (WGCNA) reveal potential regulation mechanism of carotenoid accumulation in *Chrysanthemum morifolium*. *Plant Physiol Biochem*. 2019;142:415–28.
38. Yan J, He J, Li Ja, Ren S, Wang Y, Zhou J, Tan X: Analysis of *Camellia oleifera* transcriptome reveals key pathways and hub genes involved during different photoperiods. *BMC Plant Biol*. 2022;22(1):1–13.
39. Song Q, Ji K, Mo W, Wang L, Chen L, Gao L, Gong W, Yuan D. Dynamics of sugars, endogenous hormones, and oil content during the development of *Camellia oleifera* fruit. *Botany*. 2021;99(8):515–29.
40. Wang Y, Li Ja, Guo P, Liu Q, Ren S, Juan L, He J, Tan X, Yan J. Ectopic expression of *Camellia oleifera* Abel. gibberellin 20-oxidase gene increased plant height and promoted secondary cell walls deposition in *Arabidopsis*. *Planta*. 2023;258(3):65.
41. Li W, Chen J, Dong X, Liu M, Wang G, Zhang L. Flower development and sexual dimorphism in *Vernicia montana*. *Hortic Plant J*. 2024;10(2):586–600.
42. Cheng H, Zha S, Luo Y, Li L, Wang S, Wu S, Cheng S, Li L. JAZ1-3 and MYC2-1 synergistically regulate the transformation from completely mixed flower buds to female flower buds in *Castanea mollissima*. *Int J Mol Sci*. 2022;23(12):6452.
43. Xiao F, Wang XR, Zhao Y, He H. Flowering related comparative transcriptomics between *Jatropha curcas* and *Jatropha nigroviensrugosus*. *Int J Agric Biol*. 2018;20(7):1523–32.
44. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–90.
45. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat Biotechnol*. 2011;29(7):644.
46. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22(13):1658–9.
47. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323.
48. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
49. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)*. 2021;2(3):100141.
50. Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics*. 2006;7(1):1–11.
51. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
52. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>ΔΔCT method. *Methods*. 2001;25(4):402–8.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.