Combined Multi-omics and GLK Gene Editing Offer Insights into Coloration in Golden-leaf Tree

Ruqian Wu[1,†], Ting Guo[1,†], Xiong Yang[1], Sai Huang[1], Deyu Miao[1], Tingting Chen[1], Yinxuan Xue[1], Juan Li[1], Kai Gao[1], Xiaoyu Yang[1], Bin Guo[2], Xinmin An[1,*]

(1 State Key Laboratory of Tree Genetics and Breeding, National Engineering Research Center of Tree Breeding and Ecological Restoration, Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China. 2 Shanxi Academy of Forestry and Grassland Sciences, Taiyuan 030012, China)

ABSTRACT:*Koelreuteria paniculata* is an important tree species with ornamental, medicinal, and ecological values, it is widely distributed in East Asia and introduced to Europe and North America. *K. paniculata* 'Jinye' is a mutant variety that has a golden-color leaf phenotype. Although similar leaf color variants occur in plants, little is known of the underlying mechanism. We performed physiological, anatomical, microRNA sequencing, transcriptomic and metabolomic analyses of the golden leaf variation in the mutant, and used gene editing technology verify the function of candidate genes. Compared with wild type, the golden leaf mutant exhibited 76.05% and 44.32% decreased chlorophyll *a* (Chl *a*) and chlorophyll *b* (Chl *b*) contents, respectively, and significantly increased carotenoid content. Analysis of leaf ultrastructure revealed an abnormal chloroplast morphology and fewer lamellae in the mutant. Fifty-nine differentially expressed genes (DEGs), forty transcription factors (TFs) and forty-nine differentially expressed miRNAs (DEmiRs) involved in pigment metabolism, chloroplast development and photosynthesis, were identified. *GLK* and *petC* genes were downregulated and are involved in chloroplast development and chlorophyll synthesis, respectively. Further, the function of candidate gene *GLK* was validated or identified through gene editing technology in *Populus tomentosa*. The upregulation of *PSY* and *PDS* genes, and the downregulation of *NCED* gene lead to accumulation of carotenoid in mutant leaf. Varieties of chalcones and flavonols were upregulated in the mutant. Consequently, the carotenoid to chlorophyll ratio increased by more than 75%, and the accumulation of chalcones and flavonols all together lead to the golden leaf phenotype in the mutant of *K. paniculata*.

KEYWORDS:*Koelreuteria paniculata*, leaf color variation, chloroplast, combined omics, gene editing

INTRODUCTION

Leaf color mutants are widely used for landscaping. They have bright colors, long viewing cycle, can be used to form large color blocks, and replace flowers with leaves in the light flower season. Therefore, leaf color mutant plants—such as *Ulmus pumila* 'Jinye', *Acer rubum*, *Cotinus coggygria*, *Populus deltoids* 'Quanhong', and *Photinia fraseri*—are preferred for landscaping and road greening (Zhang *et al*., 2017). Although the genetic patterns of colored-leaf plants are varied, at the metabolic level, the leaf color of higher plants depends on the content and ratio of pigments (mainly including chlorophyll, carotene, and flavonoids) in leaves. Chlorophyll is an important photosynthetic pigment, and mutations in genes linked to its biosynthesis, chloroplast protein transport, or the plant pigment regulation pathway (Wen *et al*., 2016; Zhang *et al*., 2017) can result in leaf coloration mutants. Plant carotenoids range in color from yellow to red, are involved in light harvesting, and are indispensable for

photoprotection against excess light. An increase in the carotenoid-to-chlorophyll ratio may result in yellow leaf traits, such as the yellow-striped leaves of a *Ginkgo biloba* L mutant (Li *et al*., 2018). Anthocyanins are naturally occurring pigments responsible for the red, purple, and blue color of plant organs, and are essential in multiple biological processes. The formation of red leaves is the result of anthocyanin accumulation (Li *et al*., 2015). Therefore, studies of the molecular mechanism of leaf color mutants have focused on pigment metabolism.

Koelreuteria paniculata is an arbor species widely distributed in Asia and introduced to Europe and North America. This species is strongly adaptable to the environment and suitable for phytoremediation in heavy metal-contaminated areas (Tian *et al*., 2009). Its crude extract has medicinal and antimicrobial properties, and its leaves are used as antifungal and antibacterial agents by local people (Yang *et al*., 2018). Because of its rich flavonoids, *K. paniculata* is used in both medicine and landscaping (Kim *et al*., 2017; Lyu *et al*., 2017). *K. paniculata* 'Jinye' is a new variety of *K. paniculata* bred by seedling mutation. Its physiological characteristics are similar to those of *K. paniculata* except that the golden leaves remain yellow during the growing period, enhancing its ornamental value. It also serves as a source of materials for study of the molecular mechanism and secondary metabolites of the *K. paniculata* leaf-color mutation.

The integration of metabolomics and transcriptomics can reveal the biosynthetic mechanisms of key differential functional pathways in plants (Li *et al*., 2020). For example, metabolomics and transcriptomics revealed the metabolic and transcriptional differences between the 'Rougui' protogreen leaf variety and its yellow leaf mutant, deepening our understanding of the mechanism of tea leaf coloration (Wang *et al*., 2020). Investigation of changes in the transcriptome and metabolome of jujube peel at various maturity stages and the mechanism of jujube peel coloring revealed the metabolic pathways and genes linked to jujube peel color change (Zhang *et al*., 2020). Indeed, combined -omics methods are now widely used in plant research.

To gain insight into the biological basis of leaf color and metabolite changes, we performed metabolomic, RNA-seq, microRNA sequencing and iso-seq analyses of the physiological and transcriptomic characteristics of golden leaf coloration in *K. paniculata*. The key candidate genes for leaf color variation were functionally verified by gene editing technique. Our findings provide reference information for studies of leaf coloration in other plant species.

RESULTS

Cytological changes in mutant leaves

We compared chloroplast ultrastructure in normal and mutant leaves. In healthy plants, the chloroplast was oblong or fusiform, close to the cell membrane, the stromal lamella and the basal lamella were closely arranged, and the lamellae were stacked neatly. There were a few osmophilic granules and starch granules on the surface, and the chloroplast membrane was intact (Fig. 1e, f). In mutant leaves, chloroplast morphology was abnormal, the thylakoid structure was disrupted, and the grana lamella was loose and broken or missing (Fig. 1h). Correspondingly, some chloroplasts contained irregularly arranged vesicles. Furthermore, there were fewer starch granules in chloroplasts of GL, but these were filled with numerous plastoglobuli (Fig. 1g), indicating a low photosynthetic capacity and a disrupted chloroplast membrane.

Physiological changes in GL leaves

We analyzed changes in the pigment contents of WT and GL (Fig. 2e, f, g). Compared with the WT, the contents of Chl *a* and Chl *b* in leaves of GL in May decreased by 87.38% and 44.19%, respectively; the total chlorophyll content decreased by 72.81%, and the anthocyanin content decreased by 10%. The content of carotenoids increased by 26.43%; the carotenoid-chlorophyll ratio in the WT was 0.35, which was significantly lower than in the mutant leaves (1.63). In July, with the increase of daylight time and illumination intensity，the GL leaves turns yellow-green. The Chl *a* and Chl *b* contents in *K. paniculata* 'Jinye' leaves decreased by 76.87% and 44.32%, respectively, compared with the WT whereas the contents of carotenoids increased by 30.22%, and the contents of anthocyanins remained unchanged (Fig. 2c). The carotenoid-chlorophyll ratio in GL increased by 78.49%. Compared with the WT, the Chl *a*, Chl *b*, and anthocyanin contents in GL decreased significantly in September, whereas the carotenoid content and carotenoid-chlorophyll ratio increased significantly. In May and September, there was no significant difference in the contents of chlorophyll intermediates between WT and mutant leaves (Fig. 2b, d), but in July, the Urogen Ⅲ and Pchlide contents increased significantly, whereas that of coprogen Ⅲ decreased significantly.

Using LC-MS with PCA and PLS-DA analyses (Fig. S1), we detected a total of 1471 metabolites. Using a VIP ≥ 1 and fold change ≥ 2 or fold change ≤ 0.5 as criteria, 128 significantly changed metabolites (SCMs) were detected and grouped into 11 classes (Table S1). Among them, the abundances of 95 and 33 SCMs were increased and decreased, respectively. Organic acids, lipids, and oxygenous organic compounds were major contributors to the separation between the two samples. These are likely related to the basic growth metabolism of plants caused by leaf color changes. However, phenylpropane and polyketides were upregulated in GL (Fig. S2). To assess their contribution to leaf color variation, we generated a clustering heat map of 128 SCMs using the TBtools method (Fig. 2a). The GL contained higher levels of flavonoids. Notably, the top enriched KEGG terms among the SCMs detected were phenylpropanoid biosynthesis, glutathione metabolism, and flavone and flavonol biosynthesis (Fig. S3). Also, chalcone, an important intermediate of flavonoid metabolites, and several flavonols, but not anthocyanins, accumulated significantly.

DEmiRs related to pigment metabolism and predicted target genes in WT and GL

Most DEmiRs and their predicted targets were primarily identified between WT and GL in pigment metabolism pathway. Further, the correlation between DEmiRs and DEGs was analyzed. A total of 77 relevant miRNA-mRNA interaction pairs, including 45 in the chlorophyll metabolic pathway (Fig. 3a), 14 in the carotenoid metabolic pathway (Fig. 3b), and 18 in the flavonoid synthesis pathway (Fig. 3c), were predicted. Further, analyses were carried out to identify whether these interactions were either coherent (the expression level of target mRNA is more when that of miRNA is less; the "DU" and "UD" patterns) or non-coherent (miRNA and its target mRNA have similar expression profiles) (Li *et al*., 2022).

Analyses revealed that 77 miRNA-mRNA pairs, involving 49 miRNAs and 27 targets, were associated expressed with transcriptome data between WT and GL (Table S2, S3, S4). In most combinations, the higher the target gene expression, the lower the microRNA expression, for example miR166b-GLK, miR3711a-petC pairs etc.; a few combinations indicated similar microRNAs and target genes expression, for example miR157d-psaO etc..

Overview of SMRT and Illumina sequencing

Using SMRT sequencing technology, 20 Gb of subreads were detected in the roots, stems, leaves, flowers, and seeds of *K. paniculata*. By filtering based on a length > 50 bp, full passes ≥ 0, and quality > 0.80, we screened 355,606 RoIs (Fig. S4), including 93.81% (333,598) FLNC reads and 3.80% (13,512) NFL reads (Table S5). The FLNC sequences were clustered to obtain nonredundant isoforms for use as a reference for sequence alignment. The clustering sequence statistics are shown in Table S7.

From WT and GL leaves, we obtained 38.34–41.38 million raw reads, which yielded 35.57–38.32 million clean reads after quality control. The Q30 of the raw reads ranged from 89.1% to 91.69%, indicating the high quality of the transcriptome data (Table S6). Furthermore, 79,393 unigenes were annotated in six databases (Fig. 4b, c, d; Fig. S5). In the NR database, 75,344 annotations were obtained, accounting for 94.9%; the minimum number of annotations in the GO database was 26,954, accounting for 33.95%; 13.73% of the unigenes were annotated in all databases (Table S8).

Identification and verification of DEGs

A differential expression analysis yielded 3793 DEGs in the leaves of WT compared to GL, including 2190 upregulated and 1630 downregulated genes (Fig. 4a).

GO analysis indicated that among biological processes, “carbohydrate metabolic process” and “inositol biosynthetic process” were the top enriched terms, whereas among cellular components, most DEGs were enriched in “cell wall” and “external encapsulating structure.” Among molecular functions, most DEGs were enriched in “catalytic activity” and “inositol-3-phosphate synthase activity” (Fig. S6). We speculate that a change in leaf color affects photosynthetic efficiency and alters physiological metabolism. “Organic acid metabolic process”, “L-phenylalanine metabolic process”, and “tetrapyrrole binding” were the major leaf color-related GO terms across all GO categories. Among these, tetrapyrrole binding is related to chlorophyll synthesis. “Organic acid metabolic process” and “L-phenylalanine metabolic process” are upstream pathways of flavonoid synthesis. In short, in response to leaf color mutation, there may be a variety of changes in biological response in *K. paniculata*.

The KEGG pathway analysis showed a significant separation between the WT and GL, indicating that changes in metabolite accumulation during development are tightly governed by differential gene expression. The WT *vs.* GL DEGs mapped to 121 KEGG pathways. Among them, primary metabolic processes were significantly enriched, demonstrating that leaf-color changes affect plant growth and development. Six pathways related to leaf color were identified: “phenylpropane biosynthesis”, “flavonoid biosynthesis”, “porphyrin and chlorophyll metabolism”, “carotenoid biosynthesis”, “photosynthesis”, and “photosynthetic antenna protein”. Of the top KEGG enrichment pathways, two were related to flavonoid metabolism (“phenylpropanoid biosynthesis” and “flavonoid biosynthesis”) (Fig. 4e). In mutant leaves, many downregulated DEGs were enriched in “energy metabolism” and “carbohydrate metabolism” whereas several upregulated DEGs were enriched in “porphyrin and chlorophyll metabolism” and “carotenoid biosynthesis.” A few DEGs were enriched in the photosynthesis pathway.

DEGs and SCMs related to pigment metabolism in WT and GL

Normal green leaves depend on the balance of chlorophyll, carotenoids, and flavonoids. Although the content

of Coprogen III in yellow leaves was about 0.77-fold that of green leaves, the contents of other intermediates such as UrogenⅢ, Proto Ⅸ, Mg-ProtoIX, and Pchlide were not decreased in yellow leaves. This is consistent with the finding that 17 DEGs annotated in the porphyrin and chlorophyll metabolism pathway were upregulated in mutant leaves (Fig. 5a). However, the Chl *a* and Chl *b* contents were lower in yellow leaves compared to green leaves. RNA-seq showed that the expression of *CHLG*, which catalyzes chlorophyllide production, was not significantly affected in yellow leaves. *GLK* (golden2-like) genes are key regulators of chloroplast development. In this study, the expression of two DEGs annotated as *GLKs* was lower in yellow leaves. According to the KEGG pathway results, 15 unigenes in the photosynthesis pathway were annotated as 10 genes. Compared with normal green leaves, the expression levels of *psaA*, *psaF*, and *psaO* (related to the reaction center subunit of photosystem I [PSI]), as well as *ATPB* (related to F-ATPase) were upregulated. However, the mRNA levels of *ATPD*, *ATPG*, and *petC* (associated with the cytochrome b6-f complex) were significantly downregulated, and *petC* was a highly significantly downregulated.

The total carotenoid content differed significantly between WT and GL, and carotenoid composition may influence leaf coloration. We identified 15 DEGs regulating carotenoid biosynthesis and degradation based on KEGG pathway annotations (Fig. 5b). Two genes involved in carotenoid metabolism were upregulated in mutant leaves; one was annotated as PHYTOENE SYNTHASE (*PSY*) and the other as 15-CIS-PHYTOENE DESATURASE (*PDS*). Nine carotenoid cleavage dioxygenase (*CCD*) genes encode synthetic proteases, which mediate oxidative degradation of carotenoids. Three *CCD1/CCD4/NCED* genes, which are closely related to carotenoid degradation, were significantly downregulated in the mutant. *ABA2*, *CYP707A1*, and *CYP707A4* were related to the biosynthesis of abscisic acid (ABA) downstream of the carotenoid synthesis pathway, and all were downregulated. The upregulation of carotenoid biosynthesis genes and downregulation of carotenoid degradation genes likely lead to carotenoid accumulation in mutant leaves.

Based on the KEGG enrichment results, 2 of the top 10 metabolic pathways were related to flavonoid synthesis. We constructed a network to assess the relationships between gene expression and metabolite accumulation (Fig. 5c). Compared with the WT, the abundance of cinnamic acid and *p*-coumaric acid in GL was increased. The intermediates naringenin and kaempferol, as well as other flavonoids and flavonols, accumulated significantly in mutant leaves. We identified 17 DEGs associated with flavonoid biosynthesis. The expression of nine DEGs was upregulated in GL leaves, including flavonoid synthesis precursor synthase genes (*e.g.*, *PAL*, *4CL*, and *CYP73A*), early biosynthesis genes (*e.g.*, *CHS, CHI*), and a late biosynthesis gene anthocyanin biosynthesis gene (*DFR*). This is consistent with the higher flavonoid and flavonol contents in the leaves of GL.

Transcription factors involved in leaf coloration

In this study, 381 DEGs were identified as transcription factors (TFs) belonging to 39 TF families, some of which were significantly upregulated and downregulated. Therefore, leaf-color variation affects the regulation of gene expression via TFs. The most abundant TF family was the MYB ultrasound family (53, 13.91%), followed by the C3H (36, 9.45%), NAC (30, 7.87%), ERF (27, 7.09%), bHLH (19, 4.99%), and WRKY (17, 4.46%) families (Fig. S7). MYB and bHLH transcription factors regulate flavonoid biosynthesis. In this study, the MYB family was represented by 43 DEGs upregulated in GL. Also 19 DEGs related to bHLH transcription factors were identified,

most of which were upregulated in GL. In addition, the expression of DEGs encoding WRKY, TCP, and C2H2 family TFs was significantly higher in GL leaves than in WT leaves.

Verification of DEGs by RT-qPCR

To verify the transcriptome data, we performed RT-qPCR analysis of selected DEGs at three developmental stages (May, July, and September). In total, 15 DEGs were detected, comprising 5 genes involved in chlorophyll biosynthesis, 2 in chlorophyll development, 2 in photosynthesis, 3 in carotenoid metabolism, and 3 genes involved in flavonoid biosynthesis. The expression levels of 12 genes detected by RT-qPCR showed patterns similar to the transcriptome data (Fig. 5d). In May, July, and September, the expression levels of several genes in the mutant were significantly decreased, such as *GLK* and *FtsH* (related to chloroplast development) and *NCED* (related to carotenoid degradation). *PDS* and *CYP73A* were upregulated in the mutant. Compared with the WT, *HEMA* (involved in chlorophyll biosynthesis) was significantly upregulated in GL in May and September but downregulated in July. *HEMC* was significantly upregulated in May and downregulated in July and September. In May and July, photosynthesis-related genes were significantly downregulated, but significantly upregulated in September. By contrast, *psaL* was significantly upregulated in May and July and downregulated in September. *PSY* and *DFR* were consistent with the trend of *psaL*. In general, the RT-qPCR results were consistent with RNA-seq data, indicating the transcriptome data to be valid and reliable.

Knocking out one allele of GLK generating the yellow-green phenotype in Populus tomentosa

We successfully knocked out one allele of *GLK* in *Populus tomentosa* using the CRISPR/Cas9 genome editing system and obtained two mutants *Populus tomentosa* (KO5, KO6). Through continuous identification of regenerated lines, 5 transgenic positive plants were identified from 9 regenerated plants (Fig. 6c). Among them, the mutant line KO5 and KO6 was found by Sanger sequencing. As shown in Fig 6a. , we designed two CRISPR/Cas9 sgRNAs for the secord exon of *GLK*, but only target 1 was mutated in KO5 and KO6. In addition, the KO5 and KO6 lines were found to contain monoallelic mutations, with a base insertion in KO6 and a base mismatch in KO5. A base insertion of KO6 leads to early termination of amino acid sequence and loss of conserved motifs of MYB domain (Fig. 6d). We observed that the leaves of KO6 turned yellow-green than WT (Fig. 6e).

DISCUSSIONS

Similar leaf color variations (from Green to Yellow) are universal phenomena, and often observed in nature. For examples, *Sophora japonica*, *Ulmus pumila*, and so on (Fig. S8). We boldly speculate that such phenomena follow similar regulatory mechanisms. We investigated a yellow leaf mutant of *K. paniculata*. The mutant phenotypic features (golden color leaf) of *K. paniculata* is stable by continuous several years observation, this mutant has been propagated and planted in field for over five years (Fig. S9). To explore mechanisms of the leaf color mutation in *K. paniculata* ‘Jinye’, numerous color-related genes were identified and their expression patterns were investigated. We focused on the genes related to pigment synthesis, chloroplast development, TFs, and photosynthesis based on enrichment results of transcriptome and metabolome analyses.

Chlorophyll metabolism and chloroplast development are responsible for leaf color change

Chlorophyll is an important pigment in the thylakoid that captures and transfers light energy in photosynthesis (Nakajima *et al*., 2012). Leaf color formation is closely related to chlorophyll metabolism and chloroplast development. Similar to other yellow-leaf mutants, *K. paniculata* 'Jinye' is a chlorophyll-deficient chlorina mutant. Compared with the WT, the Chl *a* and Chl *b* contents of GL were significantly lower by 76.05% and 44.32%, respectively. Therefore, the decrease in total chlorophyll content was responsible for the yellow leaf phenotype.

In higher plants, the synthesis of chlorophyll starts from glutamyl tRNA, and can be divided into three parts: the synthesis of 5-aminolevulinic acid (ALA) to protoporphyrin IX is completed in the chloroplast matrix, and the synthesis of magnesium protoporphyrin to chlorophyllide is completed in the chloroplast membrane. Finally, Chl *a* and Chl *b* are synthesized on the thylakoid membrane (Matringe *et al*., 1992). Most leaf color mutants have an altered thylakoid membrane structure. For example, in yellow leaf mutant *G. biloba*, chloroplast ultrastructure was markedly altered, chloroplast-like membranes were broken, vesicles were dense, and there were no inner membrane structures (Li *et al*., 2018). Similarly, an abnormal internal capsule membrane was found in *P. deltoides*, bamboo, and *Anthurium andraeanum* leaf mutants (Yang *et al*., 2015). In this study, the chloroplast structure of GL differed significantly from that of the WT. The chloroplasts of mutant leaves were irregularly shapes, in the basic structure, the thylakoid membranes were ruptured and chloroplast grana were loosely arranged and showed fractures and deletions. Also, there were more plastoglobuli and fewer starch particles; the latter indicates decreased photosynthetic capacity, and osmium-containing granules indicate disruption of the chloroplast membrane. These structural abnormalities in chloroplasts affect photosynthesis and chlorophyll accumulation.

In this study, we found that the precursors of chlorophyll did not decrease in the mutant, and no down-regulation of any structural gene at transcript level in the chlorophyll biosynthesis pathway. However, as an important TF for chloroplast formation and chlorophyll accumulation, *GLK* was significantly down-regulated in GL, similarly studies in other species. For example, low expression of *dmGLK2* and lacking expression of *dmGLK1* hampered chloroplast development and accumulation in the chlorophyll content, resulting in the yellow-green leaf phenotype of *Chrysanthemum* (Chang, 2013). A 40-bp deletion in *BpGLK1* of birch hampered chloroplast development, caused decrease of the chlorophyll content, and eventually led to a yellow-green leaf mutant (Gang *et al*., 2019). Therefore, the change of *GLK* gene expression led to abnormality of plastid development and malfunction, further influenced chlorophyll content.

Photosynthesis is responsible for the leaf-color change

Light is the most important environmental factor affecting plant leaf color (Zhang *et al*., 2019). The leaf color of *K. paniculata* 'Jinye' was yellow-green in the shade and after shading treatment, the leaf color changed from golden yellow to yellow-green. This suggests *K. paniculata* 'Jinye' to be a photosensitive mutant. The leaf color of the GL mutant is affected by light intensity (Zhang *et al*., 2019). However, the mechanism of the light response of the GL mutant was unclear. Light is necessary for chloroplast morphogenesis. Light controls gene transcription, chlorophyll synthesis, and protein degradation, thus regulating chloroplast development. The phenotype of chlorophyll-deficient mutants was affected significantly by environmental factors such as light (Song *et al*., 2017). During photosynthesis, light energy is captured by the pigment in the LHC protein and transferred to the reaction center complex of photosystem I (PSI) and photosystem II (PSII). The cytochrome b6-f (Cyt b6-f) complex responds

to the electron flow balance from PSII to PSI via the plastid quinone pool and regulates the activity of PSII kinase light-harvesting complex. In a chlorophyll-deficient mutant, LHC protein was significantly reduced or lost, hampering granule accumulation in chloroplasts (Kim *et al*., 2009). In this study, the expression of three DEGs in the LHC gene family was decreased in GL, indicating that the light-harvesting chlorophyll protein was reduced compared with the reaction center complex. PSI catalyzes light-driven electron transfer from luminosome cyanin to matrix ferredoxin, which consists of more than 10 subunits including PsaC and PsaL. The PsaL subunit is responsible for most of the interactions, and its mutants show a slightly smaller functional size of the photosynthetic antenna and a lower excitation level (Klodawska *et al*., 2013). Activity of the Cyt b6-f complex, which is encoded by chloroplast and nuclear genes, affects the electron transfer rate. The deletion mutation of *NTA1* has been shown to severely impair the accumulation of the Cyt b6-f complex in the thylakoid membrane and co-precipitate with components of the Cyt b6-f complex, resulting in a short phenotype, systemic yellow and inability to support photoautotrophic growth in *Arabidopsis thaliana* (Li *et al*., 2023). The Cyt b6-f complex Rieske Fe/S precursor protein of all photosynthetic eukaryotes is encoded by the nuclear *petC* gene (Veit *et al*., 2012). Site-directed knockout of the *petC* gene in *Arabidopsis thaliana* blocks the electron transport chain and produces a yellow leaf trait (Maiwald *et al*., 2003). In a *Forsythia* GL mutant, *petC* expression was downregulated and not affected by light intensity (Shen, 2019). In this study, *psaL* and *petC* expression in the Cyt b6-f complex decreased highly significantly (log2foldchange –9.1). Therefore, downregulation of gene expression in photosynthesis may explain the decreased chlorophyll content. *petC* (c69066_g1) is key for the leaf color change in *K. paniculata* ‘Jinye’.

Carotenoid biosynthesis is responsible for the leaf color change

Carotenoids have important effects on photomorphogenesis and photosynthesis (Cazzonelli and Pogson, 2010; Li *et al*., 2009). Carotenoids have multiple conjugated double bonds and can absorb light in the range of 400–500 nm. Therefore, the accumulation of carotenoids turns plants yellow, orange, and red (Zhou *et al*., 2020). Compared with the WT, the carotenoid content in GL leaves increased significantly. Yellow leaf mutants typically have a significantly increased carotenoid content. Carotenoid accumulation and pigmentation are dependent on the expression of carotenoid biosynthesis genes (Hao *et al*., 2020). As the first committed enzyme in carotenoid synthesis, phytoene synthase (PSY) converts two molecules of geranylgeranyl pyrophosphate (GGPP) to phytoene (Kang *et al*., 2014). Overexpression of *PSY1* increased the carotenoid content of tomato fruit (Fraser *et al*., 2007). PDS catalyzes the biosynthesis of lycopene from phytoene (Fu *et al*., 2016). In this study, five DEGs (*PSY*, *PSY1*, and *PDS*) related to carotenoid biosynthesis were upregulated in GL and likely contribute to carotenoid accumulation.

Carotenoid accumulation is also related to carotenoid degradation. Plants have two major carotenoid degradation pathways mediated by carotenoid dioxygenase (CCD) and 9-*cis*-epoxidized carotenoid dioxygenase (NCED). The degradation of carotenoid is suppressed in *CCD1* mutant Arabidopsis, the seed of which showed an increased carotenoid content (Schwartz, 2001). Moreover, the β-carotene content of the flesh of a *CCD4*-deficient mutant changed at the late-ripening stage, accompanied by yellowing (Ma *et al*., 2014). *NCED* is the rate limiting step from 9'-*cis* neoxanthin to ABA. The transcriptional regulation of *NCED* is a focus of research on ABA metabolism. RNAi-mediated inhibition of *NCED* expression in tomato increased the lycopene and β-carotene

contents of mature fruit and reduced the ABA content (Ji *et al*., 2014). The mRNA levels of other ABA synthesis-related genes, such as *ABA2* and *CYP707A*, were significantly decreased in GL. As a downstream product of the carotenoid pathway, ABA may be a feedback regulator of carotenoid metabolism (Mohd *et al*., 2013). Exogenous ABA induces important genes involved in carotenoid metabolism, such as *PSY3* in maize and *DSM2*, a *β-Carotene Hydroxylase* gene in rice (Du *et al*., 2010; Li *et al*., 2008). Moreover, some TFs in citrus, such as *CrMYB68*, indirectly affect carotenoid metabolism by directly inhibiting ABA biosynthesis. The effect of ABA on carotenoid metabolism needs further study.

Flavonoid biosynthesis is responsible for leaf color changes

Flavonoids are an important part of leaf color formation and protect leaves from damage caused by sunlight-derived UV irradiation (Agati *et al*., 2012). In this study, GO and KEGG analyses identified 17 DEGs in flavonoid metabolic pathways and 10 flavonoid-related DEGs in phenylpropanoid biosynthesis. Most showed significant changes in expression level. The transcript abundances of flavonoid synthesis precursor synthases (*e.g.*, PAL, 4CL, and CYP73A), EBGs (*e.g.*, CHS and FLS), and LBGs (*e.g.*, DFR) were higher in GL compared to WT. Phenylalanine ammonia lyase (PAL) catalyzes the deamination of L-phenylalanine (L-Phe) to yield cinnamic acid as the first committed step in phenylpropanoid synthesis. *4CL* encodes 4-coumaric acid coenzyme A ligase, which catalyzes the conversion of phenylalanine to 4-coumaric acid coenzyme A as a precursor for flavonoid synthesis. As key enzymes in flavonoid biosynthesis, PAL and 4CL affect the total flavonoid pathway flux through a variety of phenylacetone pathways (Winkel-Shirley, 2001). The upregulation of upstream gene expression may lead to accumulation of reaction substrates (Liu *et al*., 2019). High expression of EBGs leads to the accumulation of chalcones and flavonols. Although DFR was upregulated, no anthocyanins were found among SCMs. This may be a result of redirection of carbon flux to flavonoid branches, which is consistent with the significant increase in kaempferol content in the GL.

As secondary metabolites of phenylpropanes, flavonoids have a basic structure of C6-C3-C6 and are classified into several groups (*e.g.*, chalcones, flavonols, flavones, and anthocyanins). Flavonoids have the widest color range, from light yellow to blue. Anthocyanins are responsible for the plant coloration range from orange to blue and are used as natural food pigments. Chalcones and aurones are the main sources of yellow color in plants. For example, naringenin is widely found in yellow plants, whereas aurones are endemic to *Caryophylla* (Tanaka *et al*., 2008). Some flavonols and flavones are light yellow, affecting anthocyanin color development as co-pigments. In this study, most differential metabolites were chalcones, flavonols, and flavones, but not anthocyanins, which explains the yellow color of the leaves. Our results shed light on the regulatory network of flavonoid biosynthesis. The combination of metabolomics and transcriptomics enables investigation of the relationships between key genes and metabolites in biosynthetic pathways. We identified candidate genes and metabolites involved in the flavonoid biosynthesis pathway, providing insight into the leaf color variation of *K. paniculata*.

CONCLUSIONS

We investigated differences in coloration between normal green leaves of *K. paniculata* and golden leaves of *K. paniculata* ‘Jinye’ at the physiological and molecular levels. On the one hand, compared with the WT, the

expression of *GLK* controlling chloroplast development was significantly down-regulated in the mutant, leading to the chloroplast ultrastructure was abnormal and lamellar structure was reduced. Meanwhile, the content of Chl *a* and Chl *b* decreased by 76.05% and 44.32%, respectively, which affected the normal photosynthesis of the mutant, and the expression of *petC* related to photosynthesis was also significantly decreased. On the other hand, the expression of *PSY* and *PDS*, genes related to carotenoid synthesis, was up-regulated, while *CCD1*, *CCD4* and *NCED* for related to carotenoid degradation, were down-regulated, resulting in the accumulation of carotenoid content. Therefore, the ratio of carotenoid to total chlorophyll content in the mutant was increased by more than 75%. Metabolite content determination displayed that flavonol and chalcone expression were up-regulated in the mutant. This leads to accumulation of carotenoids, flavonoids, and flavonols, and a decrease in chlorophyll content, resulting in the yellow leaf color (Fig. 7). Our findings provide a reference for studying the mechanism of plant leaf color variation.

MATERIALS AND METHODS

Plant materials

The green-leaf *K. paniculata* cultivar (wild type, WT) and the golden-leaf *K. paniculata* 'Jinye' cultivar (golden leaf mutant, GL) were used with three independent biological replicates. They were grown under natural conditions for over five years and the mutant phenotypic features (golden color leaf) is stable by continuous several years observation in nursery in Xiangyuan County, Shanxi Province, China. From May to September, the sunshine duration and light intensity increased first and then decreased, and the GL leaves changed from yellow to yellow-green and then to yellow, therefore leaf tissues were collected in May, July and September in 2018. For cytological analysis, RNA-seq, and metabolomics experiments, green leaves and mutant leaves were sampled separately on July (Fig. 1a-d). For physiological and RT-qPCR experiments, mutant leaves and green leaves collected at three stages (May to September) were used. At least 10 tender leaves at the third position from the top of a branch were sampled from three plants of each type, immediately frozen in liquid nitrogen, and stored at −80°C for further analysis.

Measurement of chlorophyll and carotenoid contents

Approximately 0.1 g of wild-type and mutant leaves were cut into pieces and extracted with 15 mL of 80% acetone at 4°C for 24 h in the dark. The extract was measured spectrophotometrically at 470, 646, and 663 nm. The chlorophyll *a* (Chl *a*), chlorophyll *b* (Chl *b*), carotenoid, and cyanidin contents were determined as described elsewhere (Gang *et al*., 2019). To measure the contents of chlorophyll intermediates, leaves were dissolved in nine volumes of 0.01 M phosphate-buffered saline in an ice bath and centrifuged (30 min at 2500 × *g*). The supernatant was assayed separately using an ELISA Kit (HengYuan Biological Technology Co., Ltd, Shanghai, China). Three biological replicates were analyzed per sample. Data were analyzed by *t*-test using SPSS software (ver. 17.0; SPSS Inc., Chicago, IL), and means were compared at significance levels of 0.01 and 0.05.

Transmission electron microscopy (TEM)

Mature leaves (third pair of leaves at the top of the plant) of the WT and GL were collected for TEM analysis.

Avoiding the main vein, we cut the fresh tissue into 1–2-mm^3 pieces and transferred them to 2.5% (v/v) glutaraldehyde for vacuum infiltration. Next, the pieces were pre-fixed in 2.5% glutaraldehyde for 24 h at 4°C, followed by 1% OsO_4 for 2 h. After dehydration, infiltration, and embedding, the pieces were sectioned using an EM UC6 ultramicrotome (Leica Microsystems GmbH, Wetzlar, Germany), and observed using a JEOL 1200 transmission electron microscope (JEOL Ltd., Tokyo, Japan).

Metabolite profiling

Metabolites were extracted from WT, mutant leaves and the equivalent mixture of two materials taken as samples. Samples (0.1 g) were analyzed by liquid chromatography electrospray ionization tandem mass spectroscopy (LC-ESI-MS/MS) with SIMCA-P data-analysis software. Next, a principal component analysis (PCA), partial least-squares discrimination analysis (PLS-DA), and an orthogonal partial least-squares discrimination analysis (OPLS-DA) were performed to verify the reliability of the data. Fold-change (FC) (FC > 2 or FC < 0.5 and P < 0.05) and variable importance for projection (VIP) (VIP > 1 and P < 0.1) criteria were applied to identify differential metabolites.

Small RNA sequencing

Two clones of the same cultivars in both the WT and the GL libraries were used for the small RNA sequencing. Small RNA libraries were constructed using the KAITAI-BIO Small RNA Library Prep Kit (KAITAI-BIO, China) according to the manufacturer's instructions. The library preparations were sequenced on an Illumina HiSeq platform.

The raw reads obtained were processed. Fastp software is used to filter Raw Data, mainly including removing primer and connector sequences; low quality sequences at the end were removed (>20% bases <Q30, and N base >10%); and those shorter than 18 nt and longer than 30 nt were also discarded.

Identification and annotation of the miRNAs

The derived reads were further screened against rRNA, tRNA, snRNA, snoRNA and other ncRNA as well as repeat sequences by mapped to RNA database, which is Rfam (http://rfam.xfam.org/) and the remaining unannotated reads were then mapped onto the *Sapindus mukorossi* reference genome.

Importantly and notably, there was not miRNA annotation in the *K. paniculata* full-length transcriptome. Therefore, to identify known miRNAs, the mapped reads with the *K. paniculata* reference genome were mapped onto the miRNAs of *Sapindus mukorossi*, which is the most evolutionarily close to *K. paniculata*.

The unaligned unique reads were further used for novel miRNA prediction using miRDeep-P2 (Kuang *et al*., 2018) and newly updated criteria for plant miRNA criteria (Axtell and Meyers, 2018), referring to the method described by Guo et al. (2020). Furthermore, the identified miRNAs were clustered into families based on sequence similarity.

Differential expression analysis of miRNAs

To identify differentially expressed miRNAs (DEmiRs), edgeR was used. To filter out miRNAs with low expressions, those with $|\log2(FC)| \geq 1.00$ and a Benjamini-Hochberg FDR corrected P - value < 0.05 were assigned as differentially expressed. MiRNAs related to chloroplast development, photosynthesis and pigment metabolism

pathway were screened from differentially expressed miRNAs, and analyzed in heat maps.

Target genes prediction

Potential miRNA target genes were predicted using the psRNA Target, RNAhybrid Target and miranda Target. Target genes of miRNAs related to chloroplast development, photosynthesis and pigment metabolism pathway were predicted and counted.

PacBio library preparation and sequencing

To obtain integrated full-length transcriptome sequences of *Koelreuteria*, we pooled total RNA of the root, stem, leaf, seed, and flower in equal quantities to construct sequencing libraries. Next, we used the SMARTer PCR cDNA Synthesis Kit to synthesize full-length cDNA. After end repair, adaptor ligation, and index code addition, PCR amplification was conducted. The polymerase-bound template was bound to MagBeads and sequencing was performed on a PacBio Sequel platform to obtain polymerase reads.

By filtering out the adapter sequences and subreads < 1000 bp and the raw polymerase read fragment sequences < 50 bp or sequence accuracy of < 0.80, we extracted the suitable subreads from the polymerase reads. The reads of insert (RoIs) obtained by screening were classified into full-length non-chimeric (FLNC), non-full-length (NFL), full-length-chimeric, and short (< 300 bp) reads depending on whether a 5′-primer, 3′-primer, or poly-A tail was detected. Finally, similar sequences among the FLNC sequences were clustered to nonredundant isoforms after removing redundant high-quality consensus FLNC reads using CD-HIT program with a threshold of 0.99 identity.

RNA-Seq library preparation, sequencing, and expression estimation

Total RNA was isolated from the wild type and mutant with a Quick RNA Isolation Kit (Omega) and used for RNA-seq. We performed 1% agarose gel electrophoresis and used the Agilent 2100 Bioanalyzer to assess RNA integrity. RNA purity and concentration were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher). Library construction and RNA-seq were performed by HTHealth Biotechnology Co., Ltd. (Beijing, China). The libraries were subjected to next-generation sequencing on the Illumina HiSeq platform. Finally, we used in-house Perl scripts to process the low-quality data to obtain clean data.

We combined the unigenes with the PacBio Iso-Seq data resulting in a final reference transcriptome. Isoform and unigene expression levels were quantified using RSEM software (https://github.com/deweylab/rsem), and transcript lengths were normalized to FPKM values. The unigenes were mapped to five public databases: NCBI nonredundant protein sequences (NR), Swiss-Prot, Gene Ontology (GO), evolutionary genealogy of genes: Non-supervised Orthologous Groups (eggNOG), and Kyoto Encyclopedia of Genes and Genomes (KEGG), using BLAST software to obtain annotation information. We used DESeq software (Anders *et al*., 2010) to analyze gene expression with |log2foldchange| > 1 and $p < 0.05$ as the screening criteria. Finally, the DEGs were subjected to GO and KEGG analyses.

Quantitative RT-qPCR

RNA-seq samples were subjected to RT-qPCR to assess the reliability of the transcriptome data. RNA extraction and cDNA synthesis were conducted as above. The 2× SYBR Green qPCR Mix Kit (TaKaRa, Japan) and the cDNA concentration and primers described above were used; the reaction volume was 25 μL and PCR was

conducted using the 7500 Real-Time PCR System (Applied Biosystems). Three independent biological replicates per sample and three technical replicates per biological replicate were analyzed. The primer sequences are listed in Supplementary Table S1. DEG expression was normalized by the $2^{-\Delta\Delta Ct}$ method (Livak *et al*., 2001).

Vector construction of gene editing and sanger sequencing of on-target mutations

According to the sequence information of *Populus trichocarpa GLK* gene (Genebank: XM_024592840) in NCBI website, primers (F: 5'- ATG CTA GCT GTA TCA CCC TTGA -3'; R: 5'- TCA AGC ACG AGT TGG TGG TATC -3') were designed to clone *GLK* of *Populus tomentosa* (Fig. 6a). Used Vector pDIRECT_22C express the Csy4-multi-gRNA system19 to modify several targets synchronously (Li *et al*., 2018). gRNAs were designed to target regions of the coding sequences: gRNA-1 (5'- ATT GGA CAC CAG AGC TGC AC -3') and gRNA-2 (5'- TGG ATA AGG CAG TGC CTT CG -3') targeting the secord exon in *GLK*, respectively. The gRNA array in construct pDIRECT_22C_6gR_A was arranged in the following order: gRNA-1–gRNA-2 (Fig. 6b). Specific primers (Cas9-F: 5'- TGT GGA ATT GTG AGC GGA TAAC -3'; Cas9-R: 5'- AGT CTT TAC GGC GAG TTC TGTT -3') were used to identify stable transgenic lines. Mutant type was identified using *PtoGLK* gene-specific primers (*PtoGLK* arget-F: 5'- AAT CAG CAG CTC ATG CCA AGAA -3'; *PtoGLK* arget-R: 5'- CCT GGG GTT AGT CCG TTC GGA -3'). The 2 × Phanta® Max Master Mix(Vazyme, China) was used to amplify potential mutation regions in the secord exon of *PtoDLK*, and the PCR products were inserted into the pGEM®-T Easy Vector (Promega, America) to select single clones for sequencing.

AUTHOR CONTRIBUTIONS

Xinmin An and Ting Guo designed the experiments. Ruqian Wu, Ting Guo drafted the manuscript. Ting Guo, Ruqian Wu, Xiong Yang, Sai Huang and Deyu Miao performed experiments. Bin Guo and Yinxuan Xue collected experimental materials. Juan Li, Kai Gao, Tingting Chen and Xiaoyu Yang analyzed the transcriptomic data. All authors read and approved the final manuscript.

DATA AVAILABILITY

The transcriptome raw sequence data have been submitted to the NCBI Sequence Read Archive database under accession bioproject number PRJNA851565.

ACKNOWLEDGEMENTS

This work was supported by the Natural Science Foundation of China (31870652), the Opening Foundation of Key Laboratory of Urban Agriculture (North China), Ministry of Agriculture, P. R. China (kf2018012). We thank Mr. Wang Guodong for providing experimental materials. The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see: http://www.textcheck.com/certificate/te8dE6.

CONFLICT OF INTEREST STATEMENT

The authors have declared that no competing interests exist.

REFERENCES

Agati G, Azzarello E, Pollastri S, Tattini M. (2012). Flavonoids as Antioxidants in Plants: Location and Functional Significance. Plant Science. 196: 67-76. DOI: 10.1016/j.plantsci.2012.07.014

Anders S, Huber W. (2010). Differential Expression Analysis for Sequence Count Data. Genome Biology. 11(10): R106. DOI: 10.1186/gb-2010-11-10-r106

Axtell M J, Meyers B C. (2018). Revisiting Criteria for Plant MicroRNA Annotation in the Era of Big Data. Plant Cell. 30: 272–284. DOI: 10.1105/tpc.17.00851

Cazzonelli C I, Pogson B J. (2010). Source to Sink: Regulation of Carotenoid Biosynthesis in Plants. Trends in Plant Science. 15: 266-274. DOI: 10.1016/j.tplants.2010.02.003

Chang Q S, Zhang L X, Chen L, Chen S M, Liu Z L, Fang W M, Chen F D. (2013). Characteristics of Photosynthesis and Spectra Properties of Thylakoid Membranes in the Yellow-Green Leaf Mutant of Chrysanthemum. Scientia Silvae Sinicae. 49(02): 72-78. DOI: 10.11707/j.1001-7488.20130211

Du H, Wang N L, Cui F, Li X H, Xiao J H, Xiong L Z. (2010). Characterization of the Beta-carotene Hydroxylase Gene *DSM2* Conferring Drought and Oxidative Stress Resistance by Increasing Xanthophylls and Abscisic Acid Synthesis in Rice. Plant Physiology. 154: 1304-1318. DOI: 10.1104/pp.110.163741

Fraser P D, Enfissi E M, Halket J M, Truesdale M R, Yu D, Gerrish C. (2007). Manipulation of Phytoene Levels in Tomato Truit: Effects on Isoprenoids, Plastids, and Intermediary Metabolism. Plant Cell. 19 (10): 3194-3211. DOI: 10.1105/tpc.106.049817

Fu C C, Han Y C, Fan Z Q, Chen J Y, Chen W X, Lu W J, Kuang J F. (2016). The Papaya Transcription Factor *CpNAC1* Modulates Carotenoid Biosynthesis through Activating Phytoene Desaturase Genes *CpPDS2/4* during Fruit Ripening. Journal of Agricultural & Food Chemistry. 64(27): 5454-63. DOI: 10.1021/acs.jafc.6b01020

Gang H, Li R, Zhao Y, Liu G, Chen S, Jiang J. (2019). Loss of *GLK1* Transcription Factor Function Reveals New Insights in Chlorophyll Biosynthesis and Chloroplast Development. Journal of Experimental Botany. 70: 3125-3138. DOI: 10.1093/jxb/erz128

Guo Z, Kuang Z, Wang Y, Zhao Y, Tao Y, Cheng C, Yang J, Lu X, Hao C, Wang T, et al. (2020). PmiREN: a Comprehensive Encyclopedia of Plant MiRNAs. Nucleic Acids Research. 48: D1114-D1121. DOI: 10.1093/nar/gkz894

Hao Z, Liu S, Hu L, Shi J. (2020). Transcriptome Analysis and Metabolic Profiling Reveal the Key Role of Carotenoids in the Petal Coloration of *Liriodendron tulipifera*. Horticulture Research. 7: 70. DOI: 10.1038/s41438-020-0287-3

Ji K, Kai W, Zhao B, Wang H, Guo Y, Leng P, et al. (2014). *SlNCED1* and *SlCYP707A2*: Key Genes Involved in ABA Metabolism during Tomato Fruit Ripening. Journal of Experimental Botany. 65: 5243-5255. DOI: 10.1093/jxb/eru288

Kang B, Gu Q, Tian P, Xiao L, Cao H, Yang W. (2014). A Chimeric Transcript Containing *Psy1* and a Potential mRNA is Associated with Yellow Flesh Color in Tomato Accession PI 114490. Planta. 240: 1011-1021. DOI: 10.1007/s00425-014-2052-z

Kim S C, Baek S H, Hong K N, Lee J W. (2017). Characterization of the Complete Chloroplast Genome of *Koelreuteria paniculata* (Sapindaceae). Conservation Genetics Resources. 10: 69-72. DOI: 10.1007/s12686-017-0767-4

Klodawska K, Kis M, Malec P, Kovacs L, Gombos Z, Strzalka K. (2013). Morphological and Physiological Characterization of the *Delta-Psal* Mutant of Cyanobacterium Synechocystis Pcc6803. Polyhedron. 18: 289-297. DOI: 10.1016/S0277-5387(98)00298-8

Kuang Z, Wang Y, Li L, and Yang X. (2018). Mirdeep-P2: Accurate and Fast Analysis of the Microrna Transcriptome in Plants. Bioinformatics. 35: 2521–2522. DOI: 10.1093/bioinformatics/bty972

Li B J, Zheng B Q, Wang J Y, Qiao H J, Liu Z J, Wang Y. (2020). New Insight into the Molecular Mechanism of Colour Differentiation among Floral Segments in Orchids. Communication Biology. 3(1): 89. DOI: 10.1038/s42003-020-0821-8

Li F, Vallabhaneni R, Wurtzel E T. (2008). *PSY3*, A New Member of the Phytoene Synthase Gene Family Conserved in the Poaceae and Regulator of Abiotic Stress-induced Root Carotenogenesis. Plant Physiology. 146: 1333-1345. DOI: 10.1104/pp.107.111120

Li N, Wong W S, Feng L, Wang C M, Wong K S, Zhang N H, Yang W, Jiang Y M, Jiang L W, He J X. (2023). The thylakoid membrane protein NTA1 is an assembly factor of the cytochrome b6f complex essential for chloroplast development in *Arabidopsis*. Plant Communications. 4(1): 100509. DOI: 10.1016/j.xplc.2022.100509

Li T D, Yang X P, Yu Y, Si X M, Zhai X W, Zhang H W, Dong W X, Gao C X, Xu C. (2018). Domestication of Wild Tomato is Accelerated by Genome Editing. Nature Biotechnology. 36: 1160–1163. DOI: 10.1038/nbt.4273

Li W, Yang S, Lu Z, He Z, Ye Y, Zhao B, Wang L, Jin B. (2018). Cytological, Physiological, and Transcriptomic Analyses of Golden Leaf Coloration in *Ginkgo Biloba* L. Horticulture Research. 5: 12. DOI 10.1038/s41438-018-0015-4

Li Y, Zhang D Q, Zhang S Q, Lou Y F, An X M, Jiang Z H, Gao Z M. (2022). Transcriptome and Mirnaome Analysis Reveals Components Regulating Tissue Differentiation of Bamboo Shoots. Plant Physiology. 00: 1–17. DOI: 10.1093/plphys/kiac018

Li Y, Zhang Z, Wang P, Wang S, Ma L, Li L, Yang R, Ma Y, Wang Q. (2015). Comprehensive Transcriptome Analysis Discovers Novel Candidate Genes Related to Leaf Color in a Lagerstroemia Indica Yellow Leaf Mutant. Genes & Genomics. 37: 851-863. DOI: 10.1007/s13258-015-0317-y

Liu Y Y, Chen X R, Wang J P, Eliphaz N, Sun M Q, Li Y H. (2019). Transcriptomic Analysis Reveals Flavonoid Biosynthesis of Syringa Oblata Lindl. in Response to Different Light Intensity. BMC Plant Biology. 19(1): 487. DOI: 10.1186/s12870-019-2100-8

Livak K J, Schmittgen T D. (2001). Analysis of Relative Gene Expression Data Using Real-time Quantitative Pcr and the $2^{-\Delta\delta c(T)}$ Method. Methods. 25: 402-408. DOI: 10.1006/meth.2001.1262

Lyu Y Z, Dong X Y, Huang L B, Huang L S. (2017). De Novo Assembly of *Koelreuteria* Transcriptome and Analysis of Major Gene Related to Leaf Etiolation. South African Journal of Botany. 113: 355-361. DOI: 10.1016/j.sajb.2017.09.004

Ma J, Li J, Zhao J, Hui Z, Han Y. (2014). Inactivation of a Gene Encoding Carotenoid Cleavage Dioxygenase (*CCD4*) Leads to Carotenoid-Based Yellow Coloration of Fruit Flesh and Leaf Midvein in Peach. Plant Molecular Biology Reporter. 32: 246-257. DOI: 10.1007/s11105-013-0650-8

Maiwald D, Dietzmann A, Jahns P, Pesaresi P, Joliot P, Joliot A, Levin J Z, Salamini F, Leister D. (2003). Knock-out of the Genes Coding for the Rieske Protein and the ATP-synthase Delta-subunit of *Arabidopsis*. Effects on Photosynthesis, Thylakoid Protein Composition, and Nuclear Chloroplast Gene Expression. Plant Physiology. 133(1): 191-202. DOI: 10.1104/pp.103.024190

Matringe M, Camadro J M, Block M A, Joyard J, Scalla R, Labbe P, Douce R. (1992). Localization within Chloroplasts of Protoporphyrinogen Oxidase, the Target Enzyme for Diphenylether-like Herbicides. Journal of Biological Chemistry Mar. 267(7): 4646-51. DOI: 10.1016/S0021-9258(18)42882-7

Mohd I, Hawa J. (2013). Abscisic Acid Induced Changes in Production of Primary and Secondary Metabolites, Photosynthetic Capacity, Antioxidant Capability, Antioxidant Enzymes and Lipoxygenase Inhibitory Activity of Orthosiphon Stamineus Benth. Molecules. 18(7): 7957-7976. DOI: 10.3390/molecules18077957

Nakajima S, Ito H, Tanaka R, Tanaka A. (2012). Chlorophyll b Reductase Plays an Essential Role in Maturation and Storability of *Arabidopsis* Seeds. Plant Physiology. 160: 261. DOI: 10.2307/23274693

Schwartz S H. (2001). Characterization of a Novel Carotenoid Cleavage Dioxygenase from Plants. Journal of Biological Chemistry. 276: 25208-25211. DOI: 10.1074/jbc.M102146200

Shen J S. (2019). Molecular Mechanism of Yellow Leaf Formation in Yellow-leafed Forsythia. Beijing Forestry University. DOI: 10.26949/d.cnki.gblyu.2019.001198

Song L, Ma Q, Zou Z, Sun K, Yao Y, Tao J, Kaleri N A, Li X. (2017). Molecular Link between Leaf Coloration and Gene Expression of Flavonoid and Carotenoid Biosynthesis in Camellia Sinensis Cultivar 'Huangjinya'. Frontiers in Plant Science. 8: 803. DOI: 10.3389/fpls.2017.00803

Tanaka Y, Sasaki N, Ohmiya A. (2008). Biosynthesis of Plant Pigments: Anthocyanins, Betalains and Carotenoids. Plant Journal. 54: 733-749. DOI: 10.1111/j.1365-313X.2008.03447.x

Tian D, Zhu F, Yan W, Xi F, Peng C, et al. (2009). Heavy Metal Accumulation by Panicled Goldenrain Tree (*Koelreuteria Paniculata*) and Common Elaeocarpus (*Elaeocarpus Decipens*) in Abandoned Mine Soils in Southern China. Journal of Environmental Sciences. 21: 340-345.

Veit S, Takeda K, Tsunoyama Y, et al. (2012). Structure of a Thermophilic Cyanobacterial b6f-Type Rieske Protein. Acta Crystallographica. 68(10): 1400-1408. DOI: 10.1107/S0907444912034129

Wang F, Huang Y, Wu W, Zhu C, Zhang R, Chen J, Zeng J. (2020). Metabolomics Analysis of the Peels of Different Colored Citrus Fruits (*Citrus Reticulata* Cv. 'Shatangju') during the Maturation Period Based on UHPLC-QQQ-MS. Molecules. 25(2): 396. DOI: 10.3390/molecules25020396

Wen L, Sha T, Zhang S, Shan J, Diao X. (2016). Gene Mapping and Functional Analysis of the Novel Leaf Color Gene *Siygl1* in Foxtail Millet [*Setaria italica* (L.) P. Beauv]. Physiologia Plantarum. 157: 24-37. DOI: 10.1111/ppl.12405

Winkel S B. (2001). Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. Plant Physiology. **126**: 485-493. DOI: 10.1104/pp.126.2.485

Yang H Y, Xia X W, Fang W, Fu Y, An M M, Zhou M B. (2015). Identification of Genes Involved in Spontaneous Leaf Color Variation in *Pseudosasa Japonica*. Genetics and Molecular Research. **14**: 11827-11840. DOI: 10.4238/2015.October.2.16

Yang L P, Zhu J, Wang P, Zeng J, Tan R, Yang Y Z, Liu Z M. (2018). Effect of Cd on Growth, Physiological Response, Cd Subcellular Distribution and Chemical Forms of *Koelreuteria Paniculata*. Ecotoxicology and Environmental Safety. **160**: 10-18. DOI: 10.1016/j.ecoenv.2018.05.026

Zhang J, Guo S, Ren Y, He H, Liu F, Xu Y, et al. (2017). High-level Expression of a Novel Chromoplast Phosphate Transporter *ClPHT4;2* is Required for Flesh Color Development in Watermelon. New Phytologist. **213(3)**: 1208-1221. DOI: 10.1111/nph.14257

Zhang Q, Wang L, Liu Z, Zhao Z, Zhao J, Wang Z, Zhou G, Liu P, Liu M. (2020). Transcriptome and Metabolome Profiling Unveil the Mechanisms of *Ziziphus jujuba* Mill. Peel Coloration. Food Chemistry. 312: 125903. DOI: 10.1016/j.foodchem.2019.125903

Zhang S, Wu X, Cui J, Zhang F. (2019). Physiological and Transcriptomic Analysis of Yellow Leaf Coloration in *Populus deltoides* Marsh. Plos One. 14(5): e0216879. DOI: 10.1371/journal.pone.0216879

Zhang S, Zuo L, Zhang J, et al. (2017). Transcriptome Analysis of *Ulmus Pumila* 'Jinye' Responses to Different Shading Involved in Chlorophyll Metabolism. Tree Genetics & Genomes. **13(3)**: 64. DOI: 10.1007/s11295-017-1139-7

Zhou W, Niu Y, Ding X, Zhao S, Li Y, Fan G, Zhang S, Liao K. (2020). Analysis of Carotenoid Content and Diversity in Apricots (*Prunus Armeniaca* L.) Grown in China. Food Chemistry. 330: 127223. DOI: 10.1016/j.foodchem.2020.127223

FIGURE LEGENDS

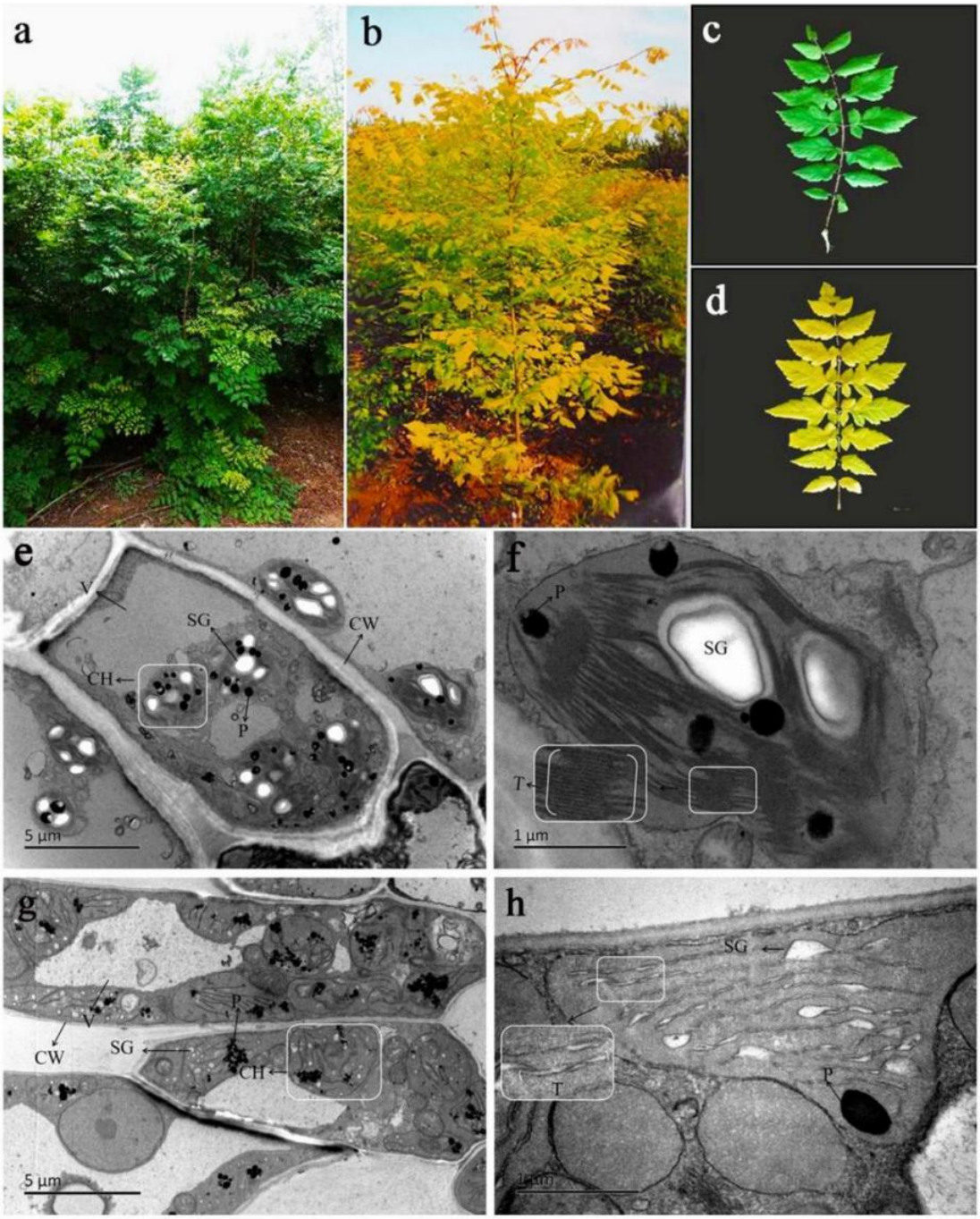

Fig. 1 Phenotypes of leaves and chloroplast ultrastructure on *K. paniculata* and *K. paniculata* 'Jinye' in July. (a, c) Phenotype of the normal green leaves. (b, d) Phenotype of the mutant leaves. (e, f) The chloroplast ultrastructure of normal green leaves showed typical structure and distinct thylakoid membrane. (g, h) Abnormal chloroplast ultrastructures contained irregularly arranged vesicles in the mutant. Bars=5 μm (e, g), 1 μm (f, h). CH chloroplast, CW cell wall, V vacuole, P plastoglobuli, T thylakoid grana, SG starch granule.

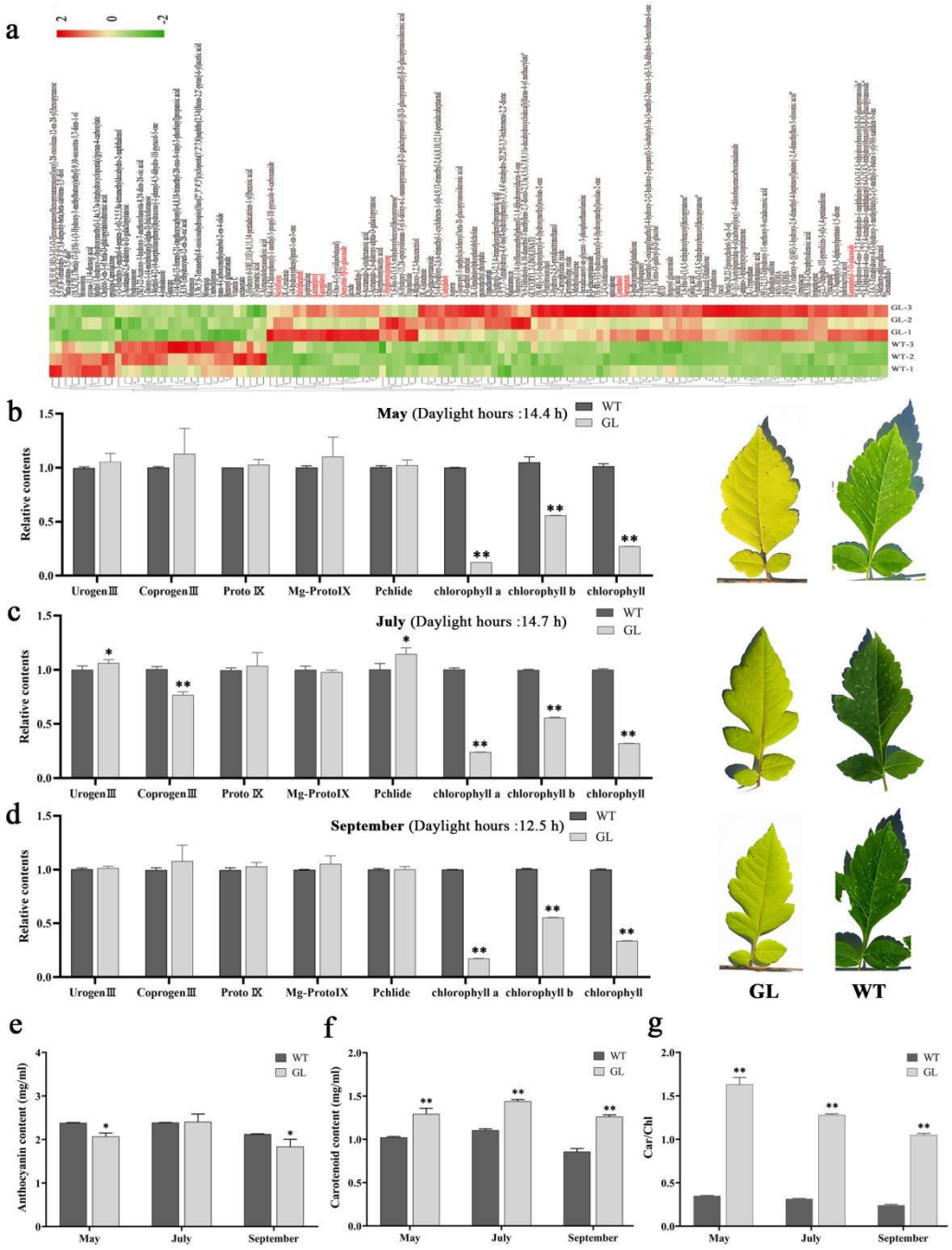

Fig. 2 Determination of pigment contents and hierarchical clustering of metabolites in *K. paniculata* and *K. paniculata* 'Jinye'.

(a) Intensity values were adjusted by log transformation and then normalized. Flavonoid with significant differences in abundance among differently colored leaves are indicated in red text. (b) Relative content of chlorophyll and chlorophyll intermediaries between normal green and mutant leaves in May. (c) Relative content of chlorophyll and chlorophyll intermediaries between normal green and mutant leaves in July. (d) Relative content of chlorophyll and chlorophyll intermediaries between normal green and mutant leaves in September. Asterisks indicate: (*) P⩽0.05, (**) P⩽0.01. (e) Anthocyanin contents of normal green and mutant leaves. (f) Carotenoid contents of normal green and mutant leaves. (g) The ratio of carotenoid to chlorophyll content of normal green and mutant leaves.

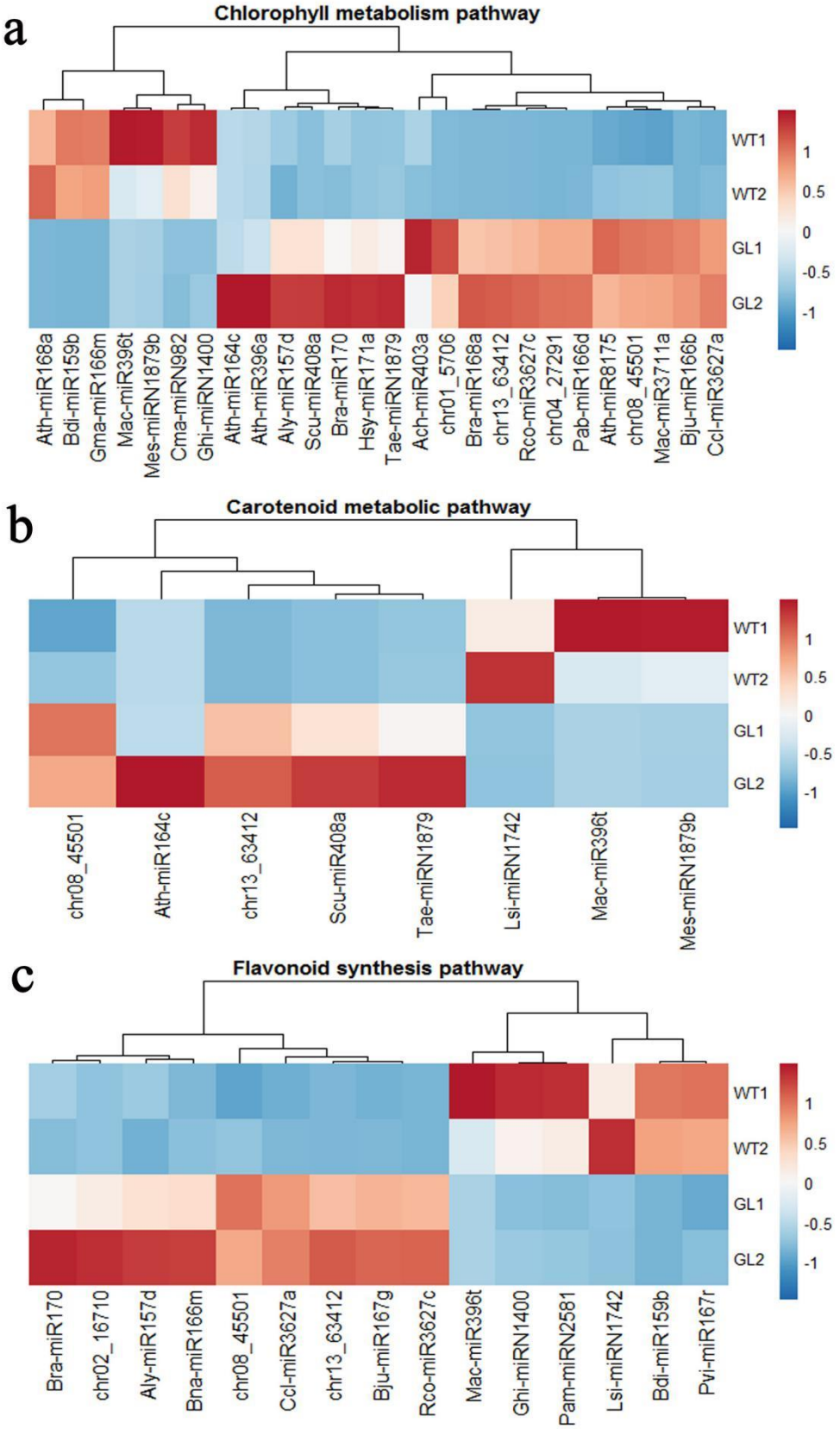

Fig. 3 Differentially expressed miRNAs in the pigment metabolic pathway.

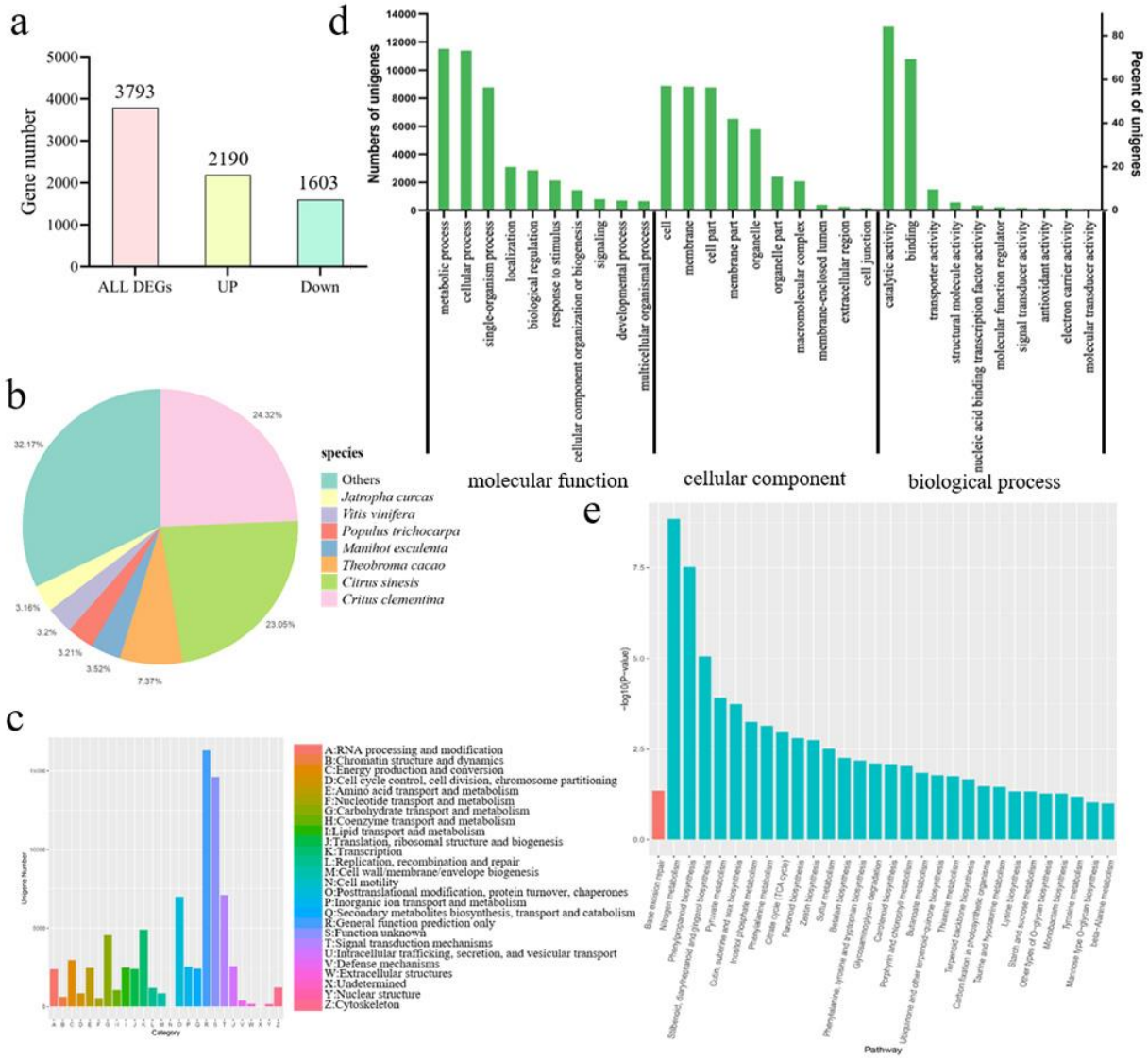

Fig. 4 Functional annotation of unigenes in leaf transcriptomes of *K. paniculata* among different samples. (a) Summary of the transcriptome DEGs. (b) Top 7 species distribution of *K. paniculata* unigenes. (c) KOG classification of *K. paniculata* unigenes. (d) GO classification of *K. paniculata* unigenes. (e) KEGG pathway enrichment of DEGs.

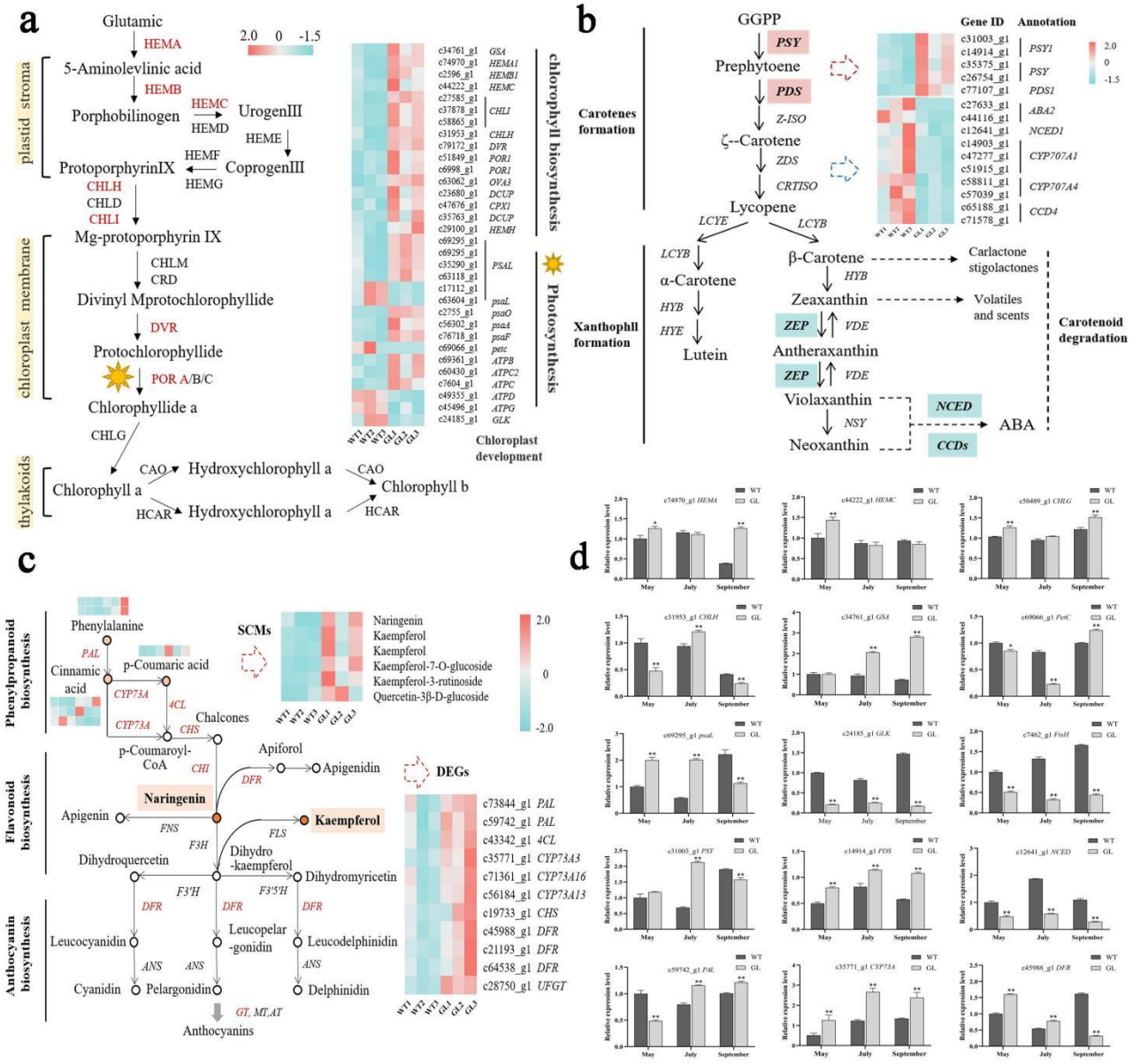

Fig. 5 Differentially expressed genes and RT-qPCR analysis. (a) Chlorophyll metabolic pathway. (b) Carotenoid metabolic pathway. (c) Flavonoid synthesis pathway. (d) RT-qPCR analysis of the expression of fifteen DEGs at different developmental stages between *K. paniculata* and *K. paniculata* 'Jinye'. Asterisks indicate: (*) P⩽0.05, (**) P⩽0.01

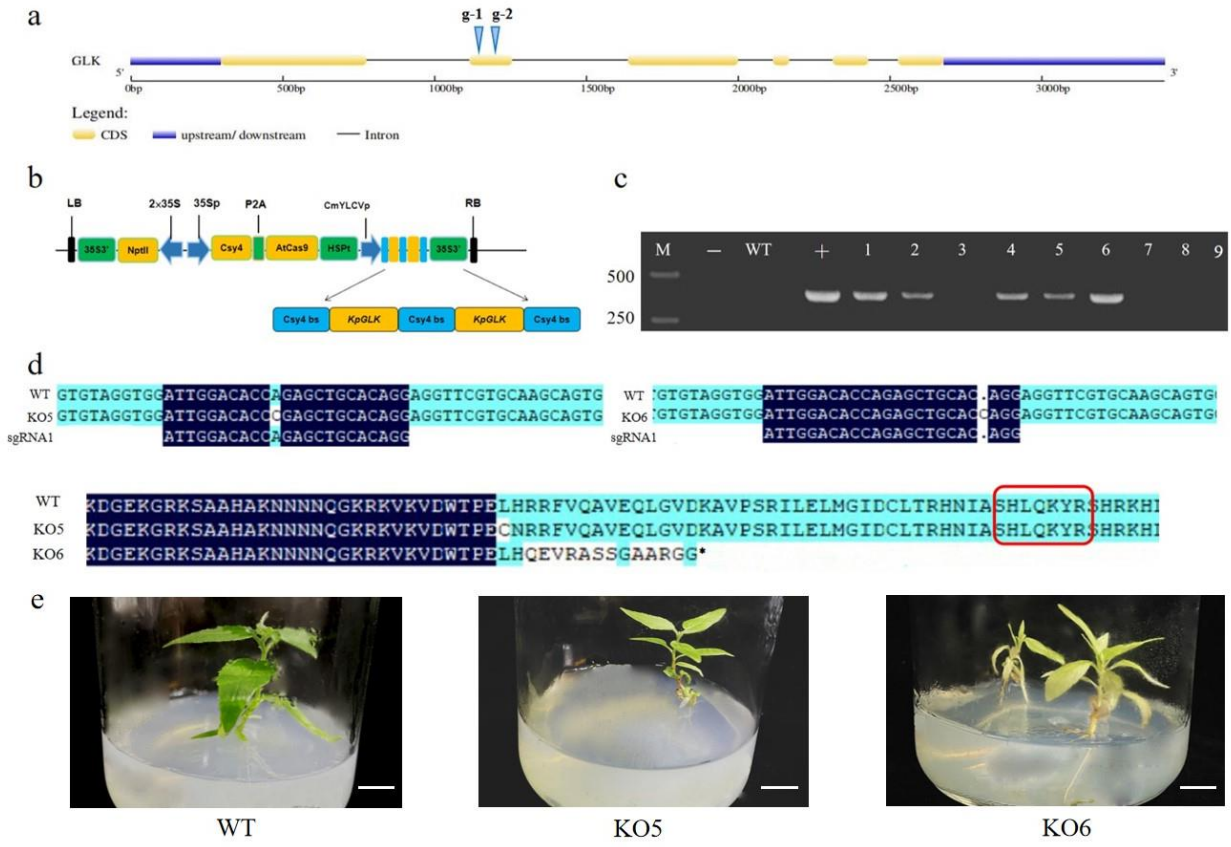

Fig. 6 Knocking out one allele of *PtoGLK* influences chloroplast development in *populus tomentosa*. (a) The target mutagenesis site in gene structure. (b) Construction sequence map of gene editing vector. (c) Identification of the T-DNA insertions in 9 regenerated plants. (d) The differences of base and amino acid sequences between the mutant and wild type strains. The red box represents the MYB domain of the *GLK* gene. (e) Yellow-green leaf phenotype of one-month-old transgenic (KO5 and KO6) and WT plants. Scale bar = 1 cm.

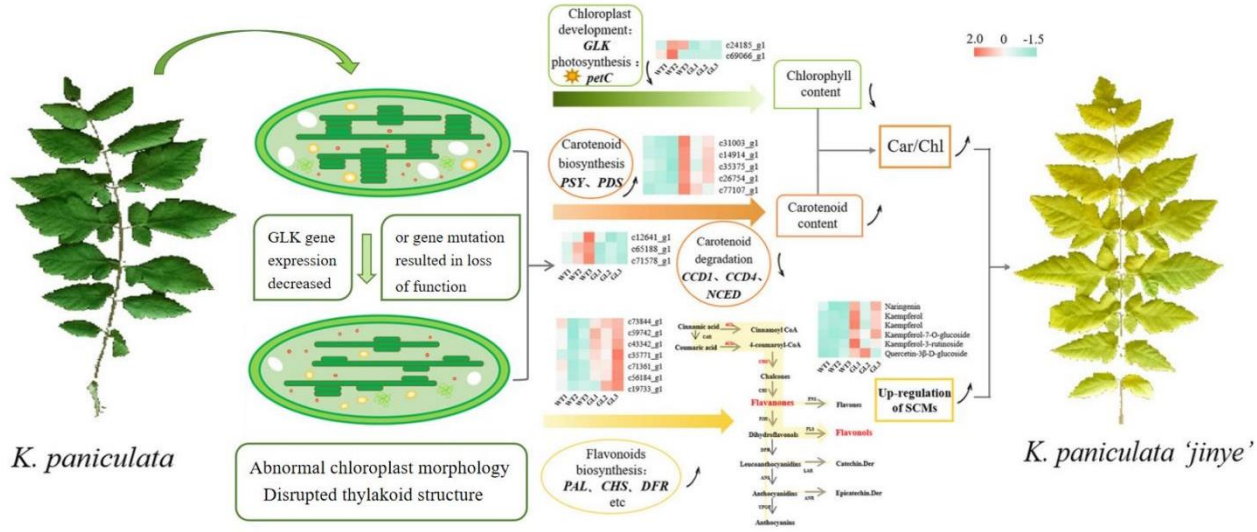

Fig. 7 The proposed pathway of mutant leaf coloration in *K. paniculata*

SUPPLEMENTAL INFORMATION