Transcriptomic and metabolomic profiling reveals the effect of LED light quality on morphological traits, and phenylpropanoid-derived compounds accumulation in *Sarcandra glabra* seedlings

Dejin Xie

Fujian Agriculture and Forestry University <https://orcid.org/0000-0003-4879-4981>

Lingyan Chen

Fujian Agriculture and Forestry University

Chengcheng Zhou

Fujian Agriculture and Forestry University

Muhammad Waqqas Khan Tarin

Fujian Agriculture and Forestry University

Deming Yang

Fujian Agriculture and Forestry University

Ke Ren

Fujian Agriculture and Forestry University

Tianyou He

Fujian Agriculture and Forestry University

Jundong Rong

Fujian Agriculture and Forestry University

Yushan Zheng (✉ zys1960@163.com)

Research article

Keywords: *Sarcandra glabra*, LED lights, Transcriptomic profiling, Metabolomic profiling, Phenylpropanoid biosynthesis

Posted Date: September 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-26873/v2>

License: (cc) (i) This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Plant Biology on October 15th, 2020. See the published version at <https://doi.org/10.1186/s12870-020-02685-w>.

Abstract

Background: *Sarcandra glabra* is an evergreen and traditional Chinese herb with anti-oxidant, anti-bacterial, anti-inflammatory, and anti-tumor effects. Recently, artificial light-emitting diodes (LEDs) have been widely used as energy for plants in greenhouse. However, little is known regarding the effects of the different LED lights on plant growth and the regulatory mechanism of phenylpropanoid-derived compounds accumulation patterns in *S. glabra*.

Results: Compared to white LED light (WL), the red LED light (RL) significantly increased the plant height and decreased the stem diameter and leaf area, whereas blue LED light (BL) significantly suppressed the height and leaf area of *S. glabra*. According to the results of transcriptomic profiling, 861, 378, 47, 10,033, 7917, and 6379 differentially expressed genes (DEGs) were identified among the groups of leaf tissue under BL (BY) vs leaf tissue under RL (RY), BY vs leaf tissue under WL (WY), RY vs WY, root tissue under WL (WG) vs WY, stem tissue under WL (WJ) vs WG, and WJ vs WY, respectively. We identified 46 unigenes encoding for almost all known enzymes involved in phenylpropanoid biosynthesis, e.g., phenylalanine ammonia lyase (PAL), 4-coumaroyl CoA ligase (4CL), chalcone synthase (CHS), flavanone 3-hydroxylase (F3H), and flavonol synthase (FLS). We found 53 unigenes encoding R2R3-MYB proteins and bHLH proteins, respectively, where several were related to flavonoids biosynthesis. Based on metabolomic profiling, a total of 454 metabolites were identified, of which 44, 87, and 296 compounds were differentially produced in WY vs RY, WY vs BY, and WY vs WG, respectively. In BY, the production of esculetin, caffeic acid, isofraxidin, and fraxidin were significantly reduced, while the yields of quercitrin and kaempferol were significantly up-regulated. In RY, the contents of cryptochlorogenic acid, cinnamic acid, and kaempferol were significantly decreased. In addition, the production of metabolites (e.g., chlorogenic acid, cryptochlorogenic acid, and scopolin) were significantly reduced, whereas the yields of metabolites (e.g., esculetin, fraxetin, isofraxidin, and fraxidin) were significantly promoted in WG.

Conclusion: These results provide further insight into the regulatory mechanism of phenylpropanoid-derived compounds accumulation patterns in *S. glabra* under different light conditions, enabling the development of optimum breeding conditions for this plant.

Background

Sarcandra glabra (Thunb.) Nakai (Caoshanhu or Zhongjiefeng in Chinese medicine) belongs to the Chloranthaceae family, a widely distributed taxonomic group in southern China and throughout Southeast Asia. *S. glabra* is an evergreen and traditional Chinese herb that is propagated by seeds or division [1, 2]. The Chinese pharmacopoeia records that the whole dry plant has been used as medicine [3], which contains abundant bioactive phytochemicals that not only contribute to protecting the plant against abiotic or biotic stress [4], but also confer anti-oxidant [5], anti-bacterial [6], anti-inflammatory [7], and anti-tumor [8] properties. Previous studies on this species mainly focused on isolating and extracting its phytochemical components, which can be divided into six groups: organic acids [9, 10], flavonoids [9], coumarins [2], terpenoids [11], phenolic acids [5, 12], and polysaccharides [1]. Among these components,

flavonoids (e.g., quercitrin and kaempferol), coumarins (e.g., isofraxidin and scopoletin), and phenolic acids (e.g., rosmarinic acid) are closely associated with the phenylpropanoid biosynthesis.

To address the limited natural resources and meet consumer demand for safe and high-quality herbal products, China has initiated the project of Good Agricultural Practice (GAP) based on Chinese materia medica to improve herb's resources and quality of cultivation. Thus, the production of *S. glabra* is gradually being shifted from the old cultivation model to more modern and sustainable practices, such as stereoscopic cultivation [13] or greenhouse cultivation [14]. Greenhouse cultivation is considered as an effective agricultural method to produce a large amount of herbs ensuring product quality [15]. Light is a pivotal element in the greenhouse environment, especially the new-generation of LEDs that provide favorable conditions for plant growth and development [16, 17].

Recently, a large number of studies have shown that specific LED lights significantly affect the morphogenesis in plants [18, 19, 21]. For instance, both the stem length and the leaf area of the tomato seedlings were significantly restricted at the higher BL/RL ratio [20, 21]. Similarly, the hypocotyl length and leaf area were suppressed when cucumber seedlings were exposed to a high percentage of BL [22]. In addition, the different LED lights have shown a significant impact on the yields of the primary and species-specific secondary metabolites in plants. For instance, the BL significantly increased the yields of phenolic compounds in pea sprouts and basil [23, 24], flavonoids production in *Cyclocarya paliurus* [25], and the amount of salidroside in callus culture of *Rhodiola imbricata* [26]. By contrast, the red and far-red parts of the light spectrum induce the production of salicylic acid, which in turn increased flavonoids accumulation in *Ginkgo biloba* [27]. Furthermore, it was found that yellow LED light had beneficial effects on flavonoids concentration in *Epimedium pseudowushanense* [28]. Collectively, these findings have shown that flavonoids and other phenolic compounds derived from the phenylpropanoid pathway accumulated in response to the monochromatic LED light irradiation. The phenylpropanoid pathway is also responsible for the production of coumarins and lignins [29].

So far, many efforts have been made in elucidating the phenylpropanoid biosynthesis, and the results showed that this complex regulatory network relied on various factors, such as enzymatic interactions, transcription factors (TFs), and the environmental factors [29, 30]. For TFs, MYB, bHLH, WDR, and the MBW complex (comprising of MYB, bHLH, and WD) are well conserved in plants and play a vital role in the regulation of the phenylpropanoid pathway [31-35].

This study was aimed to experimentally compare the effects of WL, RL, and BL on *S. glabra* seedlings, and analyze the transcriptome and metabolomic profiling within these treatments. We also identify the functional genes and TFs (R2R3-MYB and bHLH) involved in the biosynthesis of phenylpropanoid-derived compounds, as well as their expression patterns. The findings of this study will enable a better understanding of the effects of the different monochromatic LED lights on plant growth and secondary metabolites accumulation, and promote the quantity and quality of *S. glabra* cultivation by fine-tuning the optimum light condition.

Results

Effects of different light qualities on plant growth

At the final harvest (day 60), the leaf area was significantly lower in the RL and BL treatments than in the WL treatment (Fig. 1A). The plant height was significantly higher and lower in the RL and BL treatments, respectively, compared to the WL (Fig. 1B). Finally, on day 60, the stem diameter of samples in the WL and BL treatments was significantly larger than those in RL treatment (Fig. 1C). All statistical data were shown in Additional file 1: Table S1.

Illumina sequencing, *de novo* assembly, and annotation of the reference transcriptome

In total, 1,025 million raw reads were generated from the 15 cDNA libraries with a total of 1,009 million high-quality clean reads. To be specific, each library produced about 55.9-92.7 million raw reads (CRA002486). The relevant parameters of *de novo* assembly were shown in Additional file 2: Table S2. The results of the functional annotation from the seven public databases showed that 46.9 % of unigenes were annotated in at least one database, while only 3.0 % of unigenes were annotated in all the databases (Additional file 3: Table S3). A total of 33,626 unigenes were categorized into 56 functional groups among three categories: biological process, cellular component, and molecular function (Additional file 4: Fig. S1).

DEGs analysis

To identify DEGs in the 15 samples, a pairwise comparison was performed and filed with $\text{padj} < 0.05$ and $|\log_2\text{FoldChange}| > 1$. The results showed that 861 (BY vs RY), 378 (BY vs WY), 47 (RY vs WY), 10,033 (WG vs WY), 7917 (WJ vs WG), and 6379 (WJ vs WY) unigenes were significantly expressed.

The enrichment analysis was performed using the GO ontology analysis (Additional file 5: Fig. S2a-e). A large number of DEGs were predominant in five terms, namely: metabolic process, single-organism metabolic process, catalytic activity, ion binding, and transferase activity. Meanwhile, all DEGs were mapped in the KEGG database (Additional file 6: Fig. S3a-f). Interestingly, in almost every group (apart from RY vs WY) the DEGs were enriched in the biosynthesis of phenylpropanoids (including flavonoids). The phenylpropanoid-derived compounds are just originated from this biosynthetic pathway. Therefore, the follow-up experiment and analysis were performed to describe this regulatory mechanism based on transcriptomic and metabolomic profiling.

Analysis of metabolites

Based on the results obtained from the UPLC-MS/MS platform and self-compiled database, a total of 454 metabolites were detected, including various classes of lignans and coumarins, flavonoids, phenolic acids and other. Moreover, the majority of them are phenylpropanoid-derived compounds, including flavonoids (90), phenolic acids (66), lignans and coumarins (11), and tannins (13) (Table 1 and Additional file 7: Table S4). Among them, the major ingredients of fraxetin, scopolin, isofraxidin, fraxidin

and esculin from the lignans and coumarins class and rosmarinic acid, rosmarinyl glucoside, chlorogenic acid, and esculetin from the phenolic acids class were isolated from *S. glabra* samples.

Table 1 Classes of the detected metabolites in *S. glabra*

Metabolite class	Number of metabolites	Metabolic class	Number of metabolites
Flavonoids	90	Organic acids	32
Alkaloids	39	Lipids	60
Phenolic acids	66	Tannins	13
Lignans and Coumarins	11	Terpenoids	10
Amino acids and derivatives	52	Others	53
Nucleotides and derivatives	28		

The PCA (Fig. 2A) results showed that the samples tested were significantly different and that replication was generally consistent with that of others in the same group. The HCA (Fig. 2B) presented metabolic profiling and distributed them into 10 classes in all of the samples. Compared to WY and RY, most flavonoids contents (e.g. quercitrin and kaempferol) were higher in BY. Moreover, the chemicals distribution was significantly different between the root (WG) and leaves (WY, RY, and BY) as a whole. Leaf tissues contained a higher abundance of compounds such as the classes of flavonoids, phenolic acids, and tannins, while the classes of lipids, amino acids and derivative alkaloids, organic acids, nucleotides and derivatives, lignans and coumarins, terpenoids were much more abundant in the root tissue.

The results of the comparative analysis are shown in Table 2 and Fig. 3. There were 44, 87, and 296 compounds that were differentially produced in WY vs RY, WY vs BY, and WY vs WG, respectively (Fig. 3A, B, and C). Besides, compared with WY, 11 flavonoids (e.g., kaempferol) and 6 phenolic acids (e.g., cryptochlorogenic acid) were significantly down-regulated in RY (Table 2 and Fig. 3D). Among them, only 3 flavonols compounds (kaempferol, kaempferin, and cynaroside) were mapped in KEGG pathway related to flavonoids biosynthesis. In WY vs BY group, 40 flavonoids, 11 phenolic acids, and 2 lignans and coumarins had significant differences. Among them, 40 of these compounds were up-regulated and 13 compounds were down-regulated (Table 2 and Fig. 3D). In addition, only 8 out of 40 flavonoids were mapped in the flavonoid biosynthesis, and 6 of which increased by 3.2- to 6.2-fold in BY, such as quercitrin and kaempferol. However, the content of cynaroside and protocatechuic aldehyde were approximately reduced by 4.5-fold in BY. For phenolic acids, seven metabolites contents were down-regulated, including caffeic acid, esculetin, and sinapinaldehyde, whereas others increased by 2.5- to 3.3-fold. For lignans and coumarins compounds, in BY, daphnetin production decreased by 9.5-fold and oxypeucedanin increased by 2.1-fold. In terms of WY vs WG group, 129 phenylpropanoid-derived metabolites were significantly up- or down-regulated (Table 2 and Fig. 3D). Within the group of WY vs WG, compounds derived from flavonoids biosynthesis mainly accumulated in WY rather than WG (Fig. 2B). In contrast, some coumarins (esculin, isofraxidin, and fraxidin) were more abundant in WG, but the production of other coumarins (scopolin and oxypeucedanin) and most phenolic acids compounds were much higher in WY.

Furthermore, we selected the main active constituents and compared their accumulation patterns in WY, RY, BY, and WG tissues (Fig. 4). Relative to WY, the production of esculetin, caffeic acid, isofraxidin, fraxidin, fumaric acid, and maleic acid were significantly reduced (fold change ≤ 0.5), while the yields of quercitrin and kaempferol were significantly up-regulated in BY (fold change ≥ 2). Besides, the production of cryptochlorogenic acid, cinnamic acid, esculin, and rutin were increased by 1.5-fold in BY. Meanwhile, compared to WY, the production of sinapic acid, esculetin, esculin, astilbin, fumaric acid, and maleic acid were increased by 1.3- to 1.7-fold, whereas the contents of cryptochlorogenic acid, cinnamic acid, and kaempferol were significantly decreased in RY (fold change ≤ 0.5). Finally, compared to WY, the production of chlorogenic acid, cryptochlorogenic acid, cinnamic acid, coumaric acid, rosmarinyl glucoside, scopolin, quercitrin, kaempferol, astilbin, phloretin 2'-O-glucoside, quercetin, and rutin were significantly reduced (fold change ≤ 0.5), whereas the yields of sinapic acid, esculetin, fraxetin, esculin, isofraxidin, fraxidin, fumaric acid, and maleic acid were significantly promoted in WG (fold change ≥ 2).

Table 2 The quantitative statistics of identified Phenylpropanoid-derived metabolites in different groups

Group name	Identified Metabolites	Phenylpropanoid-derived Metabolites	Phenylpropanoid-derived Metabolites (significant)	Up-regulated	Down-regulated
WY vs RY	450	180	17	0	17
WY vs BY	450	180	53	40	13
WY vs WG	454	180	129	30	99

Candidate genes involved in the phenylpropanoid biosynthesis

We have identified 46 candidate unigenes encoding most of the enzymes involved in the flavonoids, coumarins, and phenolic acid biosynthesis (Table 3 and Additional file 8: Table S5). Moreover, five unigenes relating to rosmarinic acid biosynthesis were also detected. Whatever flavonoids, phenolic acids, and most of the coumarins biosynthesis, they all shared the core phenylpropanoid pathway, comprising the three committed steps catalyzed by PAL, C4H, and 4CL enzymes. In this study, four *PAL* unigenes, four *4CL* unigenes, and one *C4H* unigene were identified, of which the *PAL* unigene (Cluster-22883.50216), *4CL* unigene (Cluster-22883.47381), *C4H* unigene (Cluster-22883.50271) exhibited the highest expression levels compared to other members and showed higher expression levels in BY and RY. In the chlorogenic acid pathway, five unigenes encoding hydroxycinnamoyltransferases were found, and the expression level of unigene (Cluster-22883.48487) was much higher than other members. Besides, the unigene (Cluster-22883.48487) was highly expressed in BY. Furthermore, the unigene (Cluster-22883.21708) was only highly expressed in WG compared to other samples. The scopoletin biosynthesis is derived from a key precursor, namely ferulic CoA, produced by the enzymatic reactions of

3-O-methyltransferase (COMT) and 4CL on caffeic acid or by Caffeoyl CoA 3-O-methyltransferase (CCoAOMT) reacting with caffeoyl CoA. In this pathway, only one *COMT* unigene (Cluster-22883.56077) and one *CCoAOMT* unigene (Cluster-22883.54879) were detected, which both of them have highest expression levels in WG and down-regulated in RY and BY compared to WY. Furthermore, five unigenes encoding 2-oxoglutarate and Fe(II)-dependent dioxygenases family proteins were characterized. Particularly, the F6'H and S8H proteins belong to 2-oxoglutarate and Fe(II)-dependent dioxygenases family. The unrooted phylogenetic tree showed that the proteins encoded by these five unigenes were highly similar to F6'H or S8H proteins (Additional file 9: Fig. S4). Among them, the protein encoded by unigene (Cluster-22883.51443) from *S. glabra* showed maximum similarity with S8H protein (XP_030962790.1) from *Quercus lobata*. In addition, among these unigenes, the unigene (Cluster-22883.51443) was highly expressed in WG, while another unigene (Cluster-22883.50772) had a significantly higher expression level in BY compared to other samples. The pathway to the product of rosmarinic acid consists of four consecutive enzyme reactions, starting from the precursor (tyrosine) and four enzymes tyrosine aminotransferase (TAT), hydroxyphenylpyruvate reductase (HPPR), rosmarinate synthase (RAS), and cytochrome P450 reductase (CYP), which encoded by unigenes of Cluster-22883.49301, Cluster-22883.50853, and Cluster-22883.50532, respectively. As for the rosmarinate synthase (RAS) protein, the unrooted phylogenetic tree showed that RAS protein encoded by the unigene (Cluster-22883.29589) from *S. glabra* had the maximum similarity with RAS proteins from *Juglans regia* (XP_018826314.1) and *Arachis ipaensis* (XP_016205435.1) Additional file 10: Fig. S5). The unigene of Cluster-22883.49301 showed no significantly differential expression level in WY, RY, and BY, but Cluster-22883.50853 was highly expressed in BY. The unigene of Cluster-22883.50532 had a significantly higher expression level in WG, whereas its expression level was less in BY, compared to WY and RY.

Table 3 Summary of identified unigenes in the transcriptome of *S. glabra*

Gene	Number of unigenes	Gene	Number of unigenes	Gene	Number of unigenes
<i>PAL</i>	4	<i>CHS</i>	2	<i>ANR</i>	1
<i>4CL</i>	4	<i>CHI</i>	2	<i>ANS</i>	1
<i>C4H</i>	1	<i>F3H</i>	2	<i>TAT</i>	1
<i>C3H</i>	2	<i>F3'H</i>	1	<i>HPPR</i>	1
<i>COMT</i>	1	<i>FLS</i>	1	<i>HPPD</i>	1
<i>CCoAOMT</i>	1	<i>UFGT</i>	1	<i>RAS</i>	1
<i>F6'H</i>	4	<i>RT</i>	2	<i>CPR</i>	1
<i>S8H</i>	1	<i>DFR</i>	1	<i>Hydroxycinnamoyl-transferase</i>	5
<i>SGTF</i>	1	<i>LAR</i>	1	<i>O-MT</i>	2

The flavonoids biosynthesis shared the core pathway incorporating CHS, CHI, and F3H enzymes, and the expression levels of these unigenes encoding above-mentioned enzymes were higher in RY and BY. When the intermediate compound (dihydrokaempferol) was produced by the core enzymatic reactions, it would be flowed into several branches to generate anthocyanins, proanthocyanidins, flavonols, etc. The

biosynthesis of kaempferol is catalyzed by flavonol synthase (FLS) enzyme on dihydrokaempferol, and FLS enzyme was encoded by unigene of Cluster-22883.47725, the expression level of Cluster-22883.47725 was significantly up-regulated in BY. Rutin also originates from the precursor of dihydrokaempferol through the enzymatic reactions by flavonoid 3'-hydroxylase (F3'H, encoded by Cluster-22883.55303), FLS, UDPG flavonoid O-glucosyltransferase (UFGT, encoded by Cluster-22883.49005) and rhamnosyl transferase (RT, encoded by Cluster-22883.85233 or Cluster-22883.49258). These unigenes (except Cluster-22883.85233) had similar expression patterns that BL had the potential to induce the higher expression levels. Interestingly, the unigene of Cluster-22883.85233 was mainly expressed in WG to synthesize rutin, but Cluster-22883.49258 showed a higher expression level in WY and WJ. A series of enzymes, including dihydroflavonol 4-reductase (DFR), leucoanthocyanidin reductase (LAR), anthocyanin synthase (ANS), and anthocyanidin reductase (ANR), are responsible for the production of catechin and epicatechin, which eventually synthesize the product of proanthocyanidins (PAs). The unigenes of *DFR* (Cluster-22883.42375) and *LAR* (Cluster-22883.41029) had higher expression levels in RY, while unigene of Cluster-22883.47730 encoding ANR had a higher expression level in BY. Additionally, the unigene of *ANS* (Cluster-22883.50399) was highly expressed in RY and BY. Overall, the BL and RL cultivation increased the expression levels in a majority of functional genes, such as some isoforms in the *PAL* gene family, *4CL* gene family, *CHS*, *CHI*, *F3H*, *C4H*, and *ANS*. The phenylalanine metabolic pathways were shown in Fig. 5A and the heat map diagram of the aforementioned unigenes with the values of Log₂(FPKM) were presented in Fig. 5B.

Candidate transcription factors of MYB and bHLH

A total of 142 raw MYB (MYB and MYB-related) unigenes and 61 raw bHLH unigenes were collected. For MYB proteins, after screening, 92 MYB unigenes had different numbers of highly conserved DNA-binding domains and were divided into 4 classes: 1R-MYB, R2R3-MYB, 3R-MYB, and 4R-MYB proteins including 35, 53, 3, and 1 unigenes, respectively. Since the *Arabidopsis* genome was published, it has provided the first insight into the description and classification of plant MYB TFs. In the plants, R2R3-MYB proteins are the largest class in MYB family. They are involved in plant-specific processes and have been divided into subgroups in *Arabidopsis* based on the conservation of the DNA-binding domain and of amino acid motifs in the C terminal domain [32]. Here, we have identified all the primary structures of 53 R2R3-MYB domains in *S. glabra* (Additional file 11: Fig. S6). We constructed the phylogenetic tree with these 53 R2R3-MYB amino acid sequences from *S. glabra* and 125 full-length R2R3-MYB amino acid sequences from *Arabidopsis thaliana* (Fig. 6A). Numerous studies have shown that some subgroups of R2R3-MYB proteins have been involved in the regulation of flavonoids biosynthesis in *Arabidopsis thaliana*, including subgroups of S4, S5, S6, and S7. In subgroup 7, AtMYB11, AtMYB12, and AtMYB111 were closely related to flavonols biosynthesis. SgMYB38 was clustered with AtMYB111. SgMYB53 showed a high degree of functional similarity within the subgroup of S6 (AtMYB75/90/113/114), regulating the anthocyanin pigments pathway. A subset of six R2R3-MYB proteins of SgMYB8/14/28/43/27/39 were in the same branch of the PAs biosynthetic pathway with AtMYB123 in subgroup 5. Moreover, we identified two R2R3-MYB proteins: SgMYB6 with a high amino acid sequence similarity to AtMYB38 (RAX2, S14),

and *SgMYB32* clustered with members of subgroup 16. The heat map showed the expression patterns of selected 10 unigenes mentioned earlier according to the $\text{Log}_2(\text{FPKM})$ values (Fig. 6B). Based on the transcriptome results, the expression level of *SgMYB38* (Cluster-22883.68401) was up-regulated in BY. In subgroup 5, the six unigenes had different expression patterns among different samples. The unigenes of *SgMYB28* (Cluster-22883.46964), *SgMYB8* (Cluster-22883.824), and *SgMYB39* (Cluster-22883.68770) exhibited the highest expression levels in WY, compared to RY and BY. Meanwhile, the expression level of *SgMYB43* (Cluster-22883.34211) displayed a significant increase in BY. Remarkably, the unigene of *SgMYB27* (Cluster-22883.3887) presented a highest expression level in WJ, while the unigenes of *SgMYB6* (Cluster-22883.63255) and *SgMYB32* (Cluster-22883.11906) were highly expressed in WG.

The bHLH proteins are another ancient and functionally diverse superfamily of TFs, which have been widely investigated in the past. In this study, we have identified 53 unigenes encoding bHLH proteins that have the distinguishable signature domain. We constructed the phylogenetic tree with these 53 bHLH amino acid sequences from *S. glabra* and 182 full-length *AtbHLH* protein sequences from *Arabidopsis thaliana* (Fig. 7A). The results of multiple sequences alignment of SgbHLH domains were displayed in Additional file 12: Fig. S7. As shown in Fig. 7A, all the bHLH proteins were divided into 21 subclasses indicating that the node values between different subclasses were low, but values within each subclass were high, suggesting that the latter have strong evolutionary relationships. Specially, a previous study showed that the bHLH proteins were classified into 32 subclasses [43], with subclasses of S2, S5, and S24 (corresponding to subclasses of S8, S7, and S15 of *AtbHLH* proteins) involved in the regulation of flavonoid biosynthesis. In the subclass of S8, we have identified a group of SgbHLH proteins, namely SgbHLH22/24/25/43/45, which showed high sequence similarity with other *AtbHLH* protein members. Compared to the similar gene-expression change between the WY and RY, the relative expression levels of these five unigenes encoding SgbHLH proteins from S8 were significantly lower in BY. Particularly, the unigene of *SgbHLH45* (Cluster-22883.52750) was significantly up-regulated in WJ and WG. Within the subclass of S7, SgbHLH13 and SgbHLH12 were clustered with the subset of At1g63650 (EGL), At5g41315 (GL3), At4g00480 (MYC1), and At4g09820 (TT8) that associated with some R2R3-MYB members to participate in the regulation of flavonoids metabolism, trichome formation, epidermal cell fate specification, and in the formation of hair and non-hair cells in root epidermis cells [43]. The unigene of *SgbHLH12* (Cluster-22883.46083) showed no significantly differential expression pattern in WY, RY, BY, and WG, but it had a lower expression level in WJ. However, the unigene of *SgbHLH13* (Cluster-22883.47507) had the highest expression level in WJ. Furthermore, three bHLH proteins of SgbHLH4/27/37 belonged to the subclass of S15. The expression level of *SgbHLH4* (Cluster-22883.57401) was slightly reduced in BY and RY, compared to WY, but *SgbHLH27* (Cluster-22883.76231) showed a lower level in BY, WJ, and WG. The unigene of *SgbHLH37* (Cluster-22883.51340) had a much higher expression level in BY (Fig. 7B).

Validation of the transcript expression using qRT-PCR

To confirm the accuracy of the RNA-seq results, a total of 24 unigenes (five R2R3-MYB unigenes, five bHLH unigenes, and another 14 functional genes related to flavonoids biosynthesis) were selected

(Additional file 8: Table S5). The qRT-PCR results revealed that the expression levels of 24 selected unigenes were mostly consistent with those obtained by RNA-seq (Fig. 8). Furthermore, for each selected unigene, the high correlation coefficient ($r > 0.600$) was found between qRT-PCR and RNA-seq. Therefore, the qRT-PCR results showed that RNA-seq data were accurate and could be used in future experiments.

Discussion

Effects of different LED lights on plant morphological traits

In this study, we showed that monochromatic LED lights had a significant influence on the morphological characteristics of *S. glabra*. Specifically, compared to white light, the red light significantly increased the plant height and decreased the stem diameter and leaf area, whereas blue light significantly suppressed the height and leaf area of *S. glabra*. Similar findings were also obtained in previous studies, for instance, radish stems were found to be thin and elongated under red light [38]. In experiments on the effects of blue and red light on stem elongation of tomato seedlings and other species, results showed that stem length decreased with increasing blue to red light ratio [39, 40]. However, another report on cucumber showed that seedlings showed increased height, hypocotyl, and epicotyl under blue light alone compared to other treatments [41]. Red and blue light are both significant components of the spectrum of photosynthetically active radiation (PAR) and hence play a fundamental role in plant photomorphogenesis. Plants sense and transmit light signals through photoreceptors such as phytochromes (Phy) [42, 43], cryptochromes (Cry) [44] and phototropins (Phot) [45], which in turn trigger plant growth and development by signal transduction. Phytochromes are R/FR light photoreceptors that can constantly perceive and respond to the surrounding environment [46]. Under irradiation by red light, phytochromes suppress the activity of CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) and PHYTOCHROME INTERACTING FACTORS (PIFs) proteins. The expression level of unigene (Cluster-22883.48222) encoding the COP1 protein was significantly reduced in RY compared to WY and BY, implying that red light accelerated Phys to break down COP1 and thus induced several photomorphogenesis-promoting TFs (e.g., HY5 and LAF1), as evidenced by their higher expression level [42]. In this study, we found that the unigenes of Cluster-22883.57986 encoding HY5 and *SgMYB32* (Cluster-22883.11906) encoding LAF1 showed higher expression levels to various degrees.

The Cry family members, Cry1, Cry2, and Cry3 (Cry-DASH), which primarily mediate blue light and involve in the inhibition of hypocotyl elongation, de-etiolation of seedlings, floral initiation, root development, and other photoresponses in *Arabidopsis thaliana* [44], whereas the Cry family members are diverse among different plants [40]. Meanwhile, the Phot family members (Phot1 and Phot2) are recognized another Ultra-violet (UV) and blue-light photoreceptors. Phot1 and Phot2 are involved in the regulation of phototropism in different regulatory patterns: Phot1 activation needs a low threshold of blue light intensity, while Phot1 and Phot2 are mediated coherently by higher blue light intensities [40, 45]. In terms of hypocotyl elongation, previous research demonstrated that its inhibition was controlled by the combined action of Cry1, Cry2, and Phot1 during the de-etiolation stage of young seedlings [40, 44]. In this study, we found that four unigenes of Cluster-22883.51204, Cluster-22883.48125, Cluster-

22883.53309, and Cluster-22883.50402 putatively encoding Phot1, Phot2, Cry1, and Cry2 proteins, respectively, and they were involved in the perception and response to blue light irradiation (Additional file 14: Fig. S8a-b). Among these unigenes, the relative expression levels of *Phot1* (Cluster-22883.51204), and *Cry1* (Cluster-22883.53309) genes were up-regulated in BY, however, the *Cry2* (Cluster-22883.50402) gene was down-regulated. As mentioned earlier, the higher expression level of *Phot1* and *Cry1* genes inhibited hypocotyl elongation in *S. glabra* seedlings, which were subsequently integrated with other signaling molecules to mediate the regulatory network of those blue light-regulated genes by transcriptional and post-translational mechanisms [44, 47]. In a previous study using a kinetic analysis method, it was reported that the rate of stem growth was determined by multiple environmental signals, such as hormones and endogenous rhythms [48]. The reason why different phenotypic and aerial architecture development responses to red, blue, and blue: red light was different among different plants is complex and requires more investigation. Furthermore, the detailed mechanism of phytochrome-mediated regulation (e.g., stem elongation and bud outgrowth, also warrants further research [46].

The conjoint analyses of the production of flavonoids and coumarins biosynthesis

All of the flavonoid compounds identified in this study are thought to be derived directly from the reaction between 4-coumaroyl CoA and three units of malonyl CoA catalyzed by CHS, CHI, F3'H, F3H, and FLS enzymes (encoded by early biosynthetic genes, EBGs). The intermediate products of this reaction are dihydroflavonols (e.g., dihydrokaempferol or dihydroquercetin) which are subsequently used as substrates in other pathways involved in flavonol, anthocyanin, and proanthocyanidin biosynthesis regulated by downstream genes [designated late biosynthetic genes (LBGs), Fig. 5A] [49-52]. Our analysis of metabolism profiling (Fig. 4) revealed that the yield of some flavonols (e.g., quercitrin and kaempferol) were significantly higher in BY. They play an important role in pollen fertility, flower color, and UV-B protection [53]. In flavonol metabolism, FLS, belonging to 2-oxoglutarate-dependant dioxygenase, is a branch-point enzyme with multiple isoforms that catalyzes the conversion from dihydrokaempferol to products in a tissue-specific and inducible manner [54, 55]. In this study, the unigene (Cluster-22883.47725) encoding FLS enzyme was identified. It was also shown that flavonol production varied according to its relative expression in WY, RY, BY, and WG. Given the higher relative expression of *FLS* gene in the blue light treatment, the production of quercitrin, kaempferol, and rutin increased drastically in the leaves of *S. glabra*. Cinnamic acid and 4-coumaric acid are both upstream products of the first two steps of the core phenylpropanoid pathway, their production thus regulates the flow of precursors to all side branches in connection with the biosynthesis of lignans, flavonoids, and phenolic acids. Among these, lignins are synthesized by the conversion from 4-coumaric acid to caffeic acid catalyzed by C3H enzyme, and caffeic acid subsequently enters the lignin pathway. Meanwhile, the 4CL enzyme catalyzes the conversion of 4-coumaric acid to 4-coumaroyl CoA, a substrate assigned into the chlorogenic pathway in combination with shikimate or quinic acid, and it also proceeds to all classes of flavonoids production together with malonyl CoA. According to metabolite analysis, compared to WY, the initial products of cinnamic acid and 4-coumaric acid contents were increased by 1.47-fold and 1.08-fold in BY, respectively, but the bypass of caffeic acid production was inhibited in BY. Therefore, the increasing production of 4-coumaric acid led to more synthesis of 4-coumaroyl CoA, which in turn generated a

higher production of flavonoids and chlorogenic (e.g., chalcone and naringenin). Generally, the accumulation of initial precursors is closely related to the expression level of biosynthetic genes, including *PAL*, *C4H*, and *4CL* genes that are the three consecutively committed enzymes in core phenylpropanoid pathway leading to varieties of flavonoids. In *A. thaliana*, the *PAL* gene family (*AtPAL 1-AtPAL 4*), *4CL* gene family (*At4CL 1-At4CL 4*) and only one member of *AtC4H* have been identified and elaborated, some of which-like *AtPAL 1* and *AtPAL 2* -are thought to have a major role in the phenylpropanoids pathway [29, 56], while *At4CL 3* is thought to has preference for the flavonoids biosynthesis [57]. The unrooted phylogenetic tree showed that the unigenes of Cluster-22883.50216, Cluster-22883.50271, and Cluster-22883.47381 were highly homologous with *PAL 1/2*, *C4H*, and *4CL 3* genes, respectively, which exhibited the up-regulated expression levels under blue and red light condition, suggesting that these consecutive genes served to the flavonoids biosynthesis (Additional file 15: Fig. S9a-b).

Coumarins are widely distributed in numerous plant species and encompass multiple bioactivities in defense against phytopathogens, abiotic stress, oxidative stress, clinical diseases, and are the mainly specialized metabolites of *S. glabra* [4, 5]. Due to the crucial enzymatic step related with coumarins biosynthesis are still largely unknown in plants including *S. glabra*, this study reviewed the current probable pathways and analyzed the unigenes encoding F6'H, S8H, and scopoletin glucosyltransferase (SGTF) involved in the biosynthesis of scopoletin, fraxetin, and scopolin. In Fig. 4, except scopolin and scopoletin, the productions of fraxetin, esculin, isofraxidin, and fraxidin were most abundant in WG. Regrettably, the isofraxidin and fraxidin biosynthesis are not The substantial accumulation of fraxetin was closely related to the higher expression level of the unigene (Cluster-22883.51443) encoding S8H in WG. Furthermore, the yield of scopolin was decreased in RY, BY, and WG because of the lower expression level of *SGTF* gene (Cluster-22883.45883), which catalyzes scopolin formation. As for scopoletin, apart from the points on the formation of scopoletin derived from ferulic acid by enzymatic reaction of 4CL and F6'H1 in turn [58]. Other studies revealed that scopoletin could also be produced from the conversion of esculetin via the methylation of O-methyltransferase (O-MT) [4]. Under the blue light condition, the reduced production of scopoletin might be caused by the less yields of its both precursors of ferulic acid and esculetin. Similar to scopoletin, the biosynthetic pathway of esculetin originated from umbelliferone or caffeic acid on earth that is still unclear [4].

Some R2R3-MYB and bHLH TFs regulate the expression levels of EBGs and LBGs

To date, there have been various studies on TFs of R2R3-MYB, bHLH, WDR, and MBW complexes as regulators of the phenylpropanoid biosynthesis in plants [36, 51, 59-62]. These studies highlighted the importance of these TFs and MBW complexes in the regulation of flavonoids biosynthesis in plant tissues and the seed coat. The whole flavonoids pathway relies on the coordinated expression of EBGs and LBGs. Moreover, in *A. thaliana*, the subclasses (S4, S5, S6, and S7) of R2R3-MYB proteins [32], the subgroups (S8, S7, and S15) of bHLH proteins [34], and several WD40 proteins [35, 63] (e.g., TTG1) are proved to be involved in the regulation of the expression of EBGs and LBGs in the flavonoids biosynthesis in a certain way. In flavonol biosynthesis, members of AtMYB11/12/111 proteins from subgroup 7

regulate the EBG genes, in turn controlling flavonol production [50, 64]. SgMYB38 was assigned to subgroup 7 and was likely to control the expression level of *CHS*, *CHI*, *F3H*, *F3'H*, and *FLS* genes. Compared to WY and RY, *SgMYB38* gene was significantly up-regulated in BY, which directly increased the expression levels of EBGs and promoted the accumulation of some flavonol compounds. In addition, the results of qRT-PCR and RNA-seq data of these EBGs have demonstrated it. The accumulation of PAs is mainly distributed in the innermost cell layer of the seed coat regulated by more than four MBW complexes, i.e., AtMYB123 (TT2)-TT8/GL3/EGL3-TTG1, and AtMYB5-TT8-TTG1, which target the different LBGs, such as *ANS* (*LDOX*), *DFR*, and *UFGT* [31]. The MBW complexes of AtMYB75/90/113/114-TT8/GL3/EGL3-TTG1 also determine anthocyanin biosynthesis via the regulation of the spatial and temporal expression of LBGs in vegetative tissues [65]. These results suggest that it is reasonable to assume that MBW complexes target gene specificity. Therefore, in this paper, we screened and identified some unigenes encoding R2R3-MYB, bHLH and WDR TFs. Among these TFs, SgMYB8/14/27/28/43/39 might encode the homologous protein TT2 belonging to subgroup 5 of R2R3-MYB proteins, while the protein encoded by *SgMYB53* showed a high degree of functional similarity within members in subgroup 6 of R2R3-MYB proteins. SgbHLH12/13 were predicted to encode TT8/GL3/EGL3/MYC1 belonging to subgroup 7 of bHLH proteins in *Arabidopsis*, while the unigene of Cluster-22883.52183 was largely thought to encode TTG1. It is reasonable to assume that the assembled MBW complexes of SgMYB8/14/27/28/43/39-SgbHLH12/13-TTG1 might be involved in the synthesis of PAs, while SgMYB53-SgbHLH12/13-TTG1 were likely to be associated with anthocyanins biosynthesis. In this study, we just identified these genes encoding TFs and proposed potential MBW complexes in *S. glabra* according to the studied model plants. However, some aspects of the flavonoid biosynthesis still remain unknown: for instance, due to the unavailable genomic data for *S. glabra*, the copy numbers of important genes involved in the phenylpropanoid pathway have not yet been identified. An in-depth analysis of the MBW complexes structure, function, and regulation in *S. glabra* is still required.

Conclusions

In conclusion, *S. glabra* seedlings cultured under different monochromatic LED lights (WL, RL, and BL) showed significant differences in plant growth. Additionally, a combined analysis of transcriptomic and metabolomic profiling demonstrated that LED lights had a significant influence on metabolites accumulation and the expression patterns of functional genes and TFs (R2R3-MYB and bHLH) involving in phenylpropanoid biosynthesis. Compared to WL, BL showed higher expression levels in EBGs and some LBGs under the regulation of specific TFs, leading to a significant increase in flavonoid production in BY. *FLS* gene acted as a limiting factor, reducing the yield of flavonol compounds in the RL condition. Furthermore, the production of desired compounds (e.g., esculetin, isofraxidin, fraxidin, fumaric acid, and maleic acid) was significantly reduced in BY, while RL stimulated the accumulation of some vital compounds (e.g., sinapic acid, esculetin, esculin, scopoletin, astilbin, fumaric acid, and maleic acid) by 1.24- to 1.77-fold in RY. This study gave valuable insights into the regulatory mechanism of the phenylpropanoid biosynthesis in *Sarcandra glabra*.

Methods

Treatments and samples collection

Seeds of *S. glabra* were collected from the Louyuan state-owned forest farm in Sanming City, Fujian Province, China. The seeds were then cultivated at constant temperature and humidity incubator (25 °C, and 60 % humidity) in the Bamboo Institute of the Fujian Agriculture and Forestry University, Fuzhou, Fujian Province, China (26°5' N, 119°13' E). When the first pair of true leaves were fully expanded, the 90 seedlings were transplanted into black Poly Vinyl Chloride (PVC) rectangular pots (10 × 10 × 8 cm) using Pindstrup substrate (pH 5.5) for cultivation. Seedlings were arranged randomly in the culture room under three different light treatments (WL, 380~760 nm; RL, 656 nm; BL, 450 nm). Each treatment (WL, RL, and BL) had 30 seedlings, making 90 seedlings in total, and in each treatment (30 seedlings), we further divided into 10 seedlings in the same treatments. The LED lights, quantum sensor, and light meter (HR-450) were purchased from Hipoint Corporation. The photosynthetic photon flux density (PPFD) of the lights was set as 80 $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$, and the photoperiod was 16/8h (day/night). The respective intensities of the PPFD and light spectra were monitored using a HR-450 machine. After 60 days, we selected nine healthy seedlings from each treatment that were pulled out from substrates carefully, washed, and dried on filter paper. In each treatment, among those nine seedlings, three seedlings were selected to make a composite sample (one replicate). Thereafter, leaves, stems, and roots from each replicate were separated and the same tissues in each treatment were mixed. The detail of treatments and replicates have been presented in Additional file 16: Figure S10. In this study, we primarily focused on the effects of WL, RL, and BL on leaves, as well as the influence of WL on different tissues *S. glabra*. Therefore samples of WY, RY, BY, WJ, and WG were selected with three replication and immediately frozen in liquid nitrogen and kept at -80°C until further analysis.

Determination of morphological traits

Plant height and stem diameter were measured by using a measuring tape and a vernier caliper, respectively throughout the cultivation period. The upper fully expanded leaves were measured using a portable leaf area meter (LI-3000C, Ecotek) 60 days after seeds were planted.

RNA extraction and transcriptome profiling

Total RNA was extracted from the samples using RNeasy Pure Plant Kit (DP441, TIANGEN BIOTECH). These RNA samples were then treated with DNase I (Takara) to remove genomic DNA contamination and sent to the Novogene Bioinformatics Technology Company (Beijing, China) for cDNA libraries construction and Illumina sequencing. The cDNA libraries were generated using NEBNext Ultra™ RNA Library Prep Kit for Illumina (NEB, USA) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, mRNA was purified from total RNA using poly-Toligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X) . The first-strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H-). Secondstrand

cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 250-300 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 µl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq platform and paired-end reads were generated.

***De novo* assembly and annotation**

For the quality of assembly, clean reads were obtained by removing reads containing adapter, reads containing ploy-N and low quality reads from raw data. The Trinity method was employed to *de novo* assemble the clean reads [66]. Seven public databases or programs were used to annotate the unigenes, including the NCBI non-redundant protein (Nr), the NCBI nucleotide sequences (Nt), the protein family (Pfam), the euKaryotic Ortholog Groups (KOG), the Swiss-Prot database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Gene Ontology (GO).

Gene expression analysis

The fragments per kilobase of transcript per million mapped transcript (FPKM) values were used to analyze gene expression. The differential expression analysis of any two sets of treatments was measured using the R package *DEGseq2* [67]. An adjusted p-value of 0.05 was set as the threshold to determine significant differences in DEGs. The enrichment analysis was performed with GO ontology using the R package *GOseq* [68]. The KEGG pathway was annotated using the KEGG database. The corrected P-value < 0.05 and $|\log_2(\text{FoldChange})| \geq 1$ were set as the threshold to determine the significant differences under the GO and KEGG enrichment analyzes. Lastly, the heat map diagrams of the selected unigenes were constructed using TBtools software with $\log_2$ conversion of FPKM values [69].

Candidate transcription factors of MYB and bHLH

To identify MYB and bHLH TFs, we used the PlnTFDB database (<http://planttfdb.cbi.pku.edu.cn/>) according to the transcriptome data of *S. glabra* [70]. Based on previous studies by Dubos and Stracke who have identified the characteristics and classification of the *R2R3-MYB* gene family in *Arabidopsis thaliana* [32, 37], we identified and classified the unigenes encoding R2R3-type MYB proteins in *S. glabra*. In addition, we learned from the previous researches [34, 71] on the characteristics of the bHLH proteins in *Arabidopsis thaliana* to identify the distribution and predicted DNA binding features of the *bHLH* proteins in *S. glabra*.

Validation of RNA-seq data by qRT-PCR

First-strand cDNA was generated from 1 µg of total RNA using the PrimeScriptTM RT reagent Kit with gDNA Eraser (RR047A, Takara, Japan). The gene-specific primers and a reference gene (clathrin adaptor complexes, CAC) were designed by Primer Premier 5.0 software (Primer, Canada), and they were synthesized by the SunYa Company (Fuzhou, China) (Additional file 17: Table S7). All analyses were repeated twice for each biological replicate. The qRT-PCR was performed on the ABI 7500 Real-Time PCR system (Applied Biosystems, USA) using SYBR Green premix Ex Taq Kit (RR820A, Takara, Japan). The expression levels of the tested reference genes were determined from CT values and calculated using the $2^{-\Delta\Delta CT}$ method.

Phylogenetic analysis

A multiple sequence alignment analysis was performed using the DNAMAN software. Phylogenetic trees were constructed by using MEGA7.0 with a neighbor-joining method and 1000 bootstrap replicates after alignment by ClustalW [72].

Metabolomic profiling Analysis

Sample preparation and extraction

Meanwhile, samples of WY, RY, BY, and WG with three replication were sent to Metware Technology Company (Wuhan, China) for metabolomic profiling. The freeze-dried leaf and root were crushed using a mixer mill with a zirconia bead for 90 s at 30 Hz. 100mg powder was weighted and extracted overnight at 4 °C using 0.6 ml 70 % aqueous methanol. Following centrifugation at 10, 000 g for 10 min, the extracts were absorbed and filtrated before UPLC-MS/MS analysis.

UPLC Conditions

The sample extracts were analyzed using a UPLC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 4500Q TRAP). The analysis conditions were as follows, UPLC: column, Waters ACQUITY UPLC HSS T3 C18 (1.8 µm, 2.1 mm*100 mm); The mobile phase consisted of solvent A, pure water with 0.04 % acetic acid, and solvent B, acetonitrile with 0.04 % acetic acid. Sample measurements were performed under a gradient program that employed the starting conditions of 95 % A and 5 % B. Within 10 min, a linear gradient to 5 % A and 95 % B was programmed, and a composition of 5 % A, 95% B was maintained for 1min. Subsequently, a composition of 95 % A and 5.0 % B was adjusted within 0.10 min and maintained for 2.9 min. The column oven was set to 40 °C, and the injection volume was 4 µl. The effluent was alternatively connected to an ESI-triple quadrupole-linear ion trap (QTRAP)-MS. A quality-control sample was made of an equal mixture of all samples. During the process, the quality-control sample was run every ten injections to monitor the stability of the analytical conditions.

ESI-Q TRAP-MS/MS

LIT and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap mass spectrometer (Q TRAP), API 4500 Q TRAP UPLC/MS/MS System, equipped with an ESI Turbo Ion-Spray interface, operating in positive and negative ion mode and controlled by Analyst 1.6.3 software (AB Sciex). The ESI source operation parameters were as follows: an ion source, turbo spray; source temperature 550 °C; ion spray voltage (IS) 5500 V (positive ion mode)/-4500 V (negative ion mode); ion source gas I (GSI), gas II (GSII), curtain gas (CUR) were set at 50, 60, and 30.0 psi, respectively; the collision gas (CAD) was high. Instrument tuning and mass calibration were performed with 10 and 100 µmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as MRM (multiple reaction monitoring) experiments with collision gas (nitrogen) set at 5 psi. DP and CE for individual MRM transitions were done with further optimization of the two parameters. Specific sets of MRM transitions were monitored for each period according to the metabolites eluted within the period.

Metabolite identification and analysis

Metabolite identification was conducted by the self-compiled MWDB database (MetWare biological science and Technology Co., Ltd., Wuhan, China) and publicly available metabolite databases. Quantitative analysis of metabolites was done based on the MRM mode, and the characteristic ions of each metabolite were screened through the QQQ mass spectrometer to obtain signal strengths. Integration and correction of chromatographic peaks were performed using MultiQuant version 3.0.2 (AB SCIEX, Canada). The corresponding relative metabolite contents were shown as chromatographic peak area integrals. To investigate the metabolic differences and the degree of variation between samples and quality control sample (a mixer sample pooled together with equal mass from the aforementioned samples) with three replications, the unsupervised principal component analysis (PCA) was performed using the R package *prcomp*. Data was unit variance scaled before performing PCA. To investigate metabolic changes in different samples, a hierarchical cluster analysis (HCA) was carried out using the R package *heatmap*, and the raw data was disposed of unit variance (UV) scaling before HCA analysis. The significantly differential metabolites were set with the thresholds of VIP (variable importance projection) ≥ 1 and $\text{Log}_2(\text{Fold change}) \geq 1$ or ≤ -1 .

Statistical analysis

Statistical analyses were performed by using the SPSS 20.0 software (IBM, USA). Differences in comparison of mean between different treatments were performed by analysis of variance (ANOVA) at α 0.05, 0.01, 0.001, and 0.0001 probability level. *Represent $P < 0.05$, **Represent $P < 0.01$, ***Represent $P < 0.001$, ****Represent $P < 0.0001$. Graphpad Prism 8.0 and MS excel were used for graphs and tables, respectively.

Abbreviations

LED: Light-emitting diodes; WL: White LED light; RL: Red LED light; BL: Blue LED light; DEGs: Differentially expressed genes; WG: Root tissue under WL; WJ: Stem tissue under WL; BY: Leaf tissue under BL; WY:

Leaf tissue under WL; RY: Leaf tissue under RL; GAP: Good Agricultural Practice; TFs: Transcription factors; MBW: MYB-bHLH-WD40; Nr: Non-redundant protein; Nt: NCBI nucleotide sequences; Pfam: The protein family; KOG: The euKaryotic Ortholog Groups; KEGG: The Kyoto Encyclopedia of Genes and Genomes; GO: the Gene Ontology; FPKM: Fragments per kilobase of transcript per million mapped transcript; PVC: Poly Vinyl Chloride; PPFD: Photosynthetic photon flux density; qRT-PCR: quantitative Real-time PCR; PCA: Principal component analysis; HCA: Hierarchical cluster analysis; UV: unit variance; PAL: Phenylalanine ammonia lyase; C4H: Cinnamate 4-hydroxylase; 4CL: 4-coumaroyl CoA ligase; C3H: Coumarate 3-hydroxylase; HCT: Hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase; HQT: Hydroxyl-cinnamoyl CoA quinate hydroxycinnamoyl transferase; COMT: Caffeic acid 3-O-methyltransferase; CCoAOMT (CCOMT): Caffeoyl CoA 3-O-methyltransferase; F6'H: Feruloyl-CoA 6'-Hydroxylase; S8H: Scopoletin 8-hydroxylase; SGTF: Scopoletin glucosyltransferase; O-MT: O-methyltransferase; CHS: Chalcone synthase; CHI: Chalcone isomerase; F3H: Flavanone 3-hydroxylase; F3'H: Flavonoid 3'-hydroxylase; FLS: Flavonol synthase; UFGT: UDPG flavonoid O-glucosyltransferase; RT: Rhamnosyl transferase; DFR: Dihydroflavonol 4-reductase; LAR: Leucoanthocyanidin reductase; ANR: Anthocyanidin reductase; ANS or LDOX: Anthocyanin synthase or Leucoanthocyanidin dioxygenase; TAT: Tyrosine aminotransferase; HPPR: Hydroxyphenylpyruvate reductase; HPPD: 4-hydroxyphenylpyruvate dioxygenase; RAS: Rosmarinate synthase (rosmarinic acid synthase); CYP (CPR): Cytochrome P450 reductase. CA2H: Caffeic acid 2-hydroxylase, PAs: Proanthocyanidins; r: Correlation coefficient; PAR: Photosynthetically active radiation; Phy: Phytochromes; Cry: Cryptochromes; Phot: Phototropins; COP1: CONSTITUTIVE PHOTOMORPHOGENIC 1; PIFs: PHYTOCHROME INTERACTING FACTORS; EBGs: Early biosynthetic genes; LBGs: Late biosynthetic genes; VIP: variable importance projection.

Declarations

Acknowledgements

We are very grateful to Ph.D Shaoyan Shen, Ph.D Junming Zheng, and Professor Lianfeng Gu for providing useful suggestions, discussion to this study. We are very grateful to the editor and reviewers for evaluating our manuscript and providing comments for its improvement.

Author's contributions

YZ made the research plan. TH collected the materials. DX, CZ, M.W.K, DY, and KR performed the experiments. DX, LC, CZ, and M.W.K analyzed the data and wrote the manuscript. All authors discussed the results and comment on the manuscript. All authors gave final approval for publication.

Funding

This work was supported by the Science and Technology Innovation Platform of Fujian province of China (2008Y2001). GAP Engineering Research Center of Fujian province of China.

Availability of data and materials

15 transcriptome raw data produced by Illumina NovaSeq 6000 have been deposited in the Genome Sequence Archive (SRA) database (<https://bigd.big.ac.cn/gsa/>) under the accession number CRA002486.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing Interest

The authors declare no conflict of interest.

References

1. Li H, Gong X, Wang Z, Pan C, Zhao Y, Gao X, Liu W. Multiple fingerprint profiles and chemometrics analysis of polysaccharides from *Sarcandra glabra*. *Int J Biol Macromol*. 2019; 123:957-967.
2. Zhou H, Liang J, Lv D, Hu Y, Zhu Y, Si J, Wu S. Characterization of phenolics of *Sarcandra glabra* by non-targeted high-performance liquid chromatography fingerprinting and following targeted electrospray ionisation tandem mass spectrometry/time-of-flight mass spectrometry analyses. *Food Chem*. 2013; 138(4):2390-2398.
3. The state pharmacopoeia commission of People's Republic of China. *Pharmacopoeia of the People's Republic of China*, vol. 1: Beijing: Chemical Industry Press; 2015.
4. Bourgaud F, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, Matern U. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. *Phytochem Rev*. 2006; 5(2):293-308.
5. Liu J, Li X, Lin J, Li Y, Wang T, Jiang Q, Chen D. *Sarcandra glabra* (Caoshanhu) protects mesenchymal stem cells from oxidative stress. A bioevaluation and mechanistic chemistry. *BMC Complem Altern M*. 2016; 16:423.
6. Cao H, Tan R, He R, Tang L, Wang X, Yao N, Duan W, Hu Y, Yao X, Kurihara H. *Sarcandra glabra* Extract Reduces the Susceptibility and Severity of Influenza in Restraint-Stressed Mice. *Evid Based Complement Alternat Med*. 2012; 2012:236539.
7. Wei S, Chi J, Zhou M, Li R, Li Y, Luo J, Kong L. Anti-inflammatory lindenane sesquiterpenoids and dimers from *Sarcandra glabra* and its upregulating AKT/Nrf2/HO-1 signaling mechanism. *Ind Crop Prod*. 2019; 137:367-376.
8. Guo X, Shen L, Tong Y, Zhang J, Wu G, He Q, Yu S, Ye X, Zou L, Zhang Z *et al*. Antitumor activity of caffeic acid 3,4-dihydroxyphenethyl ester and its pharmacokinetic and metabolic properties. *Phytomedicine*. 2013; 20(10):904-912.

9. Li X, Zhang Y, Zeng X, Yang L, Deng Y. Chemical profiling of bioactive constituents in *Sarcandra glabra* and its preparations using ultra-high-pressure liquid chromatography coupled with LTQ Orbitrap mass spectrometry. *Rapid Commun Mass Spectrom*. 2011; 25(17):2439-2447.
10. Zheng W, Wang S, Chen X, Hu Z. Analysis of *Sarcandra glabra* and its medicinal preparations by capillary electrophoresis. *Talanta*. 2003; 60(5):955-960.
11. Yaermaimaiti S, Wang P, Luo J, Li R, Kong L. Sesquiterpenoids from the seeds of *Sarcandra glabra* and the potential anti-inflammatory effects. *Fitoterapia*. 2016; 111:7-11.
12. Zhou H, Liang J, Lv D, Hu Y, Zhu Y, Si J, Wu S. Characterization of phenolics of *Sarcandra glabra* by non-targeted high-performance liquid chromatography fingerprinting and following targeted electrospray ionisation tandem mass spectrometry/time-of-flight mass spectrometry analyses. *Food Chem*. 2013; 138(4):2390-2398.
13. Liao P, Liu P, Wang Y, Huang C, Lan L, Yang Y, Cui X. Stereoscopic cultivation of *Panax notoginseng*: A new approach to overcome the continuous cropping obstacle. *Ind Crop Prod*. 2018; 126:38-47.
14. Ding J, Tu H, Zang Z, Huang M, Zhou S. Precise control and prediction of the greenhouse growth environment of *Dendrobium candidum*. *Comput Electron Agr*. 2018; 151:453-459.
15. Yano A, Cossu M. Energy sustainable greenhouse crop cultivation using photovoltaic technologies. *Renew Sust Energ Rev*. 2019; 109:116-137.
16. Tarakanov I, Yakovleva O, Konovalova I, Paliutina G, Anisimov A. Light-emitting diodes: On the way to combinatorial lighting technologies for basic research and crop production. *Acta Hortic*. 2012; 956:171-178.
17. Yeh N, Chung J. High-brightness LEDs—Energy efficient lighting sources and their potential in indoor plant cultivation. *Renew Sust Energ Rev*. 2009; 13(8):2175-2180.
18. De Keyser E, Dhooghe E, Christiaens A, Van Labeke M, Van Huylenbroeck J. LED light quality intensifies leaf pigmentation in ornamental pot plants. *Sci Hortic*. 2019; 253:270-275.
19. Berkovich YA, Konovalova IO, Smolyanina SO, Erokhin AN, Avercheva OV, Bassarskaya EM, Kochetova GV, Zhigalova TV, Yakovleva OS, Tarakanov IG. LED crop illumination inside space greenhouses. *Reach*. 2017; 6:11-24.
20. Thwe AA, Kasemsap P, Vercambre G, Gay F, Phattaralerphong J, Gautier H. Impact of red and blue nets on physiological and morphological traits, fruit yield and quality of tomato (*Solanum lycopersicum* Mill.). *Sci Hortic*. 2020; 264:109185.
21. Nanya K, Ishigami Y, Hikosaka S, Goto E. Effects of blue and red light on stem elongation and flowering of tomato seedlings. *Acta Hortic*. 2012; 956:261-266.
22. Hernández R, Kubota C. Physiological responses of cucumber seedlings under different blue and red photon flux ratios using LEDs. *Environ Exp Bot*. 2016; 121:66-74.
23. Liu H, Chen Y, Hu T, Zhang S, Zhang Y, Zhao T, Yu H, Kang Y. The influence of light-emitting diodes on the phenolic compounds and antioxidant activities in pea sprouts. *J Funct Foods*. 2016; 25:459-465.

24. Taulavuori K, Pyysalo A, Taulavuori E, Julkunen-Tiitto R. Responses of phenolic acid and flavonoid synthesis to blue and blue-violet light depends on plant species. *Environ Exp Bot.* 2018; 150:183-187.
25. Liu Y, Fang S, Yang W, Shang X, Fu X. Light quality affects flavonoid production and related gene expression in *Cyclocarya paliurus*. *J Photoch Photobio B.* 2018; 179:66-73.
26. Kapoor S, Raghuvanshi R, Bhardwaj P, Sood H, Saxena S, Chaurasia OP. Influence of light quality on growth, secondary metabolites production and antioxidant activity in callus culture of *Rhodiola imbricata* Edgew. *J Photoch Photobio B.* 2018; 183:258-265.
27. Ni J, Dong L, Jiang Z, Yang X, Sun Z, Li J, Wu Y, Xu M. Salicylic acid-induced flavonoid accumulation in *Ginkgo biloba* leaves is dependent on red and far-red light. *Ind Crop Prod.* 2018; 118:102-110.
28. Yang Q, Pan J, Shen G, Guo B. Yellow light promotes the growth and accumulation of bioactive flavonoids in *Epimedium pseudowushanense*. *J Photoch Photobio B.* 2019; 197:111550.
29. M.Fraser C, Chapple C. The Phenylpropanoid Pathway in Arabidopsis. *The Arabidopsis Book.* 2011; 9:e152.
30. Vogt T. Phenylpropanoid Biosynthesis. *Mol Plant.* 2010; 3(1):2-20.
31. Xu W, Grain D, Bobet S, Gourrierc J, Thévenin J, Kelemen Z, Lepiniec L, Dubos C. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR complexes and their targets in Arabidopsis seed. *New Phytol.* 2014; 202(1):132-144.
32. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* 2010; 15(10):573-581.
33. Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* 2004; 40(1):22-34.
34. Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis Basic/Helix-Loop-Helix Transcription Factor Family. *Plant Cell.* 2003; 15(8):1749-1770.
35. Smith TF, Gaitatzes C, Saxena K, Neer EJ. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci.* 1999; 24(5):181-185.
36. Carretero-Paulet L, Galstyan A, Roig-Villanova I, Martínez-García JF, Bilbao-Castro JR, Robertson DL. Genome-Wide Classification and Evolutionary Analysis of the bHLH Family of Transcription Factors in Arabidopsis, Poplar, Rice, Moss, and Algae. *Plant Physiol.* 2010; 153(3):1398-1412.
37. Stracke R, Werber M, Weisshaar B. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol.* 2001; 4(5):447-456.
38. Samuolienė G, Sirtautas R, Brazaitytė A, Sakalauskaitė J, Sakalauskienė S, Duchovskis P. The impact of red and blue light-emitting diode illumination on radish physiological indices. *Cent Eur J of Biol.* 2011; 6(5):821-828.
39. Nanya K, Ishigami Y, Hikosaka S, Goto E. Effects of blue and red light on stem elongation and flowering of tomato seedlings. *Acta Hortic.* 2012; 956:261-266.

40. Huché-Thélier L, Crespel L, Gourrierc JL, Morel P, Sakr S, Leduc N. Light signaling and plant responses to blue and UV radiations—Perspectives for applications in horticulture. *Environ Exp Bot.* 2016; 121:22-38.
41. Hernández R, Kubota C. Physiological responses of cucumber seedlings under different blue and red photon flux ratios using LEDs. *Environ Exp Bot.* 2015; 29(1):66-74.
42. Li J, Li G, Wang H, Wang Deng X. Phytochrome signaling mechanisms. *The Arabidopsis Book.* 2011; 9:e148.
43. Franklin KA, Quail PH. Phytochrome functions in Arabidopsis development. *J Exp Bot.* 2010; 61(1):11-24.
44. Yu X, Liu H, Klejnot J, Lin C. The Cryptochrome Blue Light Receptors. *The Arabidopsis Book.* 2010; 8:e135.
45. Christie JM. Phototropin Blue-Light Receptors. *Annu Rev Plant Biol.* 2007; 58(1):21-45.
46. Demotes-Mainard S, Péron T, Corot A, Bertheloot J, Le Gourrierc J, Pelleschi-Travier S, Crespel L, Morel P, Huché-Thélier L, Boumaza R *et al.* Plant responses to red and far-red lights, applications in horticulture. *Environ Exp Bot.* 2016; 121:4-21.
47. Sellaro R, Hoecker U, Yanovsky M, Chory J, Casal JJ. Synergism of Red and Blue Light in the Control of Arabidopsis Gene Expression and Development. *Curr Biol.* 2009; 19(14):1216-1220.
48. Parks BM, Folta KM, Spalding EP. Photocontrol of stem growth. *Curr Opin Plant Biol.* 2001; 4(5):436-440.
49. Koes R, Quattrocchio F, Mol J. The flavonoid biosynthetic pathway in plants: Function and evolution. *Bioessays.* 1994; 16:123-132.
50. Falcone Ferreyra ML, Rius SP, Casati P. Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Front Plant Sci.* 2012; 3:222.
51. Xu W, Dubos C, Lepiniec L. Transcriptional control of flavonoid biosynthesis by MYB–bHLH–WDR complexes. *Trends Plant Sci.* 2015; 20(3):176-185.
52. Grotewold E. Plant metabolic diversity: a regulatory perspective. *Trends Plant Sci.* 2005; 10(2):57-62.
53. Wang Z, Wang S, Xiao Y, Li Z, Wu M, Xie X, Li H, Mu W, Li F, Liu P *et al.* Functional characterization of a HD-ZIP IV transcription factor NtHDG2 in regulating flavonols biosynthesis in *Nicotiana tabacum*. *Plant Physiol Bioch.* 2020; 146:259-268.
54. Owens D, Alerding A, Crosby K, Bandara A, Westwood J, Winkel B. Functional Analysis of a Predicted Flavonol Synthase Gene Family in *Arabidopsis*. *Plant Physiol.* 2008; 147(3):1046-1061.
55. Lukacin R, Wellmann F, Britsch L, Martens S, Matern U. Flavonol synthase from *Citrus unshiu* is a bifunctional dioxygenase. *Phytochemistry.* 2003; 62:287-292.
56. Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y, Yu J, Chen Z. Functional analysis of the *Arabidopsis* PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiol.* 2010; 153(4):1526-1538.

57. Ehrling J, Büttner D, Wang Q, Douglas CJ, Kombrink E. Three 4-coumarate: coenzyme A ligases in *Arabidopsis thaliana* represent two evolutionarily divergent classes in angiosperms. *Plant J.* 1999; 19(1):9-20.
58. Siwinska J, Siatkowska K, Olry A, Grosjean J, Ihnatowicz A. Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in *Arabidopsis*. *J Exp Bot.* 2018; 69(7):1735-1748.
59. 59. Ambawat S, Sharma P, Yadav NR, Yadav RC. MYB transcription factor genes as regulators for plant responses: an overview. *Physiol Mol Biol Pla.* 2013; 19(3):307-321.
60. 60. Liu J, Osbourn A, Ma P. MYB transcription factors as regulators of phenylpropanoid metabolism in p *Mol Plant.* 2015; 8(5):689-708.
61. 61. Bipei, Zhang, Martin, Hülkamp. Evolutionary Analysis of MBW Function by phenotypic rescue in *Arabidopsis thaliana*. *Front Plant Sci.* 2019; 10:375.
62. 62. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* 2011; 66(1):94-116.
63. 63. Ramsay N, Glover B. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci.* 2005; 10(2):63-70.
64. 64. Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, Weisshaar B. Analysis of production of flavonol glycosides-dependent flavonol glycoside accumulation in *Arabidopsis thaliana* plants reveals MYB11-, MYB12- and MYB111-independent flavonol glycoside accumulation. *New Phytol.* 2010; 188(4):985-1000.
65. 65. Gonzalez A, Zhao M, Leavitt JM, Lloyd AM. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.* 2008; 53(5):814-827.
66. 66. Grabherr M, Haas B, Yassour M, Levin J, Thompson D, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q *et al.* Full-Length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* 2011; 29:644-652.
67. 67. 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.
68. 68. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 2010; 11(2):R14.
69. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant.* 2020; 13(8):1194-1202.
70. 70. Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B. PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 2010; 38:D822-D827.
71. 71. Atchley WR, Terhalle W, Dress A. Positional dependence, cliques, and predictive motifs in the bHLH protein d *J Mol Evol.* 1999; 48(5):501-516.

Figures

Figure 1

Box plots of the effects on leaf area, plant height, and stem diameter under different LED lights. (A) the leaf area of the seedlings on the 60th day; (B) the plant height; (C) the stem diameter; (D) the *S. glabra* seedlings grown under different LED lights; (E) the whole plant morphology under WL, RL, BL; Maximum and minimum values are represented at the upper and lower ends of the whisker, respectively. The 75th and 25th percentiles are represented at the upper and lower ends of the box, respectively. *Represent $P <$

0.05, **Represent $P < 0.01$, ***Represent $P < 0.001$, ****Represent $P < 0.0001$. WL: White LED light; RL: Red LED light; BL: Blue LED light;

Figure 2

Results of PCA (A) and HCA (B) analyses of 12 samples from *S. glabra*. PC1 (61.25% variance) and PC2 (12.8% variance) were used to generate the scores plot. Samples of WY, RY, BY, WG, and mix are marked in green, purple, pink, brown, and blue, respectively (Fig. 2A). In HCA plot, all the identified metabolites were divided into 11 classes, and each coloured rectangle represented the specific compound content (Fig. 2B).

Figure 3

Metabolite profiling of identified compounds in different groups. Volcano plot (Fig.3A, B, and C) showing all the differential compounds in groups of WY vs RY, WY vs BY, and WY vs WG. Violin plot (Fig.3D) showing significantly differential phenylpropanoid-derived compounds in groups of WY vs RY, WY vs BY, and WY vs WG. If $VIP \geq 1$ and $\text{Log}_2(\text{Fold change}) \geq 1$ or ≤ -1 , they were regarded as the differential metabolites.

Figure 4

Heat map diagram of the relative concentration of the 24 principle active metabolites. In the upper legend, different colors represent different metabolite classes. The metabolite names are listed at the bottom. A: the fold change of relative content between RY and WY; B: the fold change of relative content between BY and WY; C: the fold change of relative content between WG and WY. Fold change ≥ 2 or ≤ 0.5 , they were regarded as the differential metabolites.

Figure 5

Phenylalanine metabolic pathways (A) and the heat map diagram of the selected unigenes (B). The name of unigenes are listed on the vertical line, and the samples are listed horizontally. Dotted lines indicate that some steps of enzymatic reactions are not completely confirmed. The scale represents the logarithms of the FPKM values of these unigenes based on the transcriptome data. The grey rectangles without values mean that the FPKM value is zero.

Figure 6

Phylogenetic trees and heat map diagrams with R2R3-MYB amino acid sequences. Bootstrap values are displayed as percentages (1000 replicates) when >50 %; (A) The subgroups (S1-S25) of R2R3-MYB highlighted with different colors were designated as previously reported [32, 37], and 53 unigenes encoding R2R3-MYB proteins in *S. glabra* were numbered randomly from 1 to 53. All the labeled R2R3-MYB amino acid sequences are given in Additional file 13: Table S6; (B) All the selected unigenes are listed on the vertical line. The scale represents the logarithms of the FPKM values of these unigenes based on the transcriptome data. The grey rectangles without values mean that the FPKM value is zero

Figure 7

Phylogenetic trees and heat map diagrams with bHLH amino acid sequences. Bootstrap values are displayed as percentages (1000 replicates) when >50 %; (A) The Subgroups S1-S21 for bHLH highlighted with different colors were designated as previously reported [34], and 53 unigenes encoding bHLH proteins of *S. glabra* were numbered randomly from 1 to 53. All the labeled bHLH amino acid sequences are given in Additional file 13: Table S6; (B) All the selected unigenes are listed on the vertical line. The scale represents the logarithms of the FPKM values of these unigenes based on the transcriptome data. The grey rectangles without values mean that the FPKM value is zero.

Figure 8

Real-time quantitative PCR confirmation of the 24 selected unigenes. The left y-axis and white legend represent the relative expression levels determined by RT-qPCR, while the right-axis and slash legend represent the FPKM values obtained by RNA-seq. Expression values were adjusted by setting the expression of WY to 1 for each unigene. All RT-qPCRs for each unigene used triplicates with two repeats per experiment. Error bars indicate SD, and different lower case letters (a-e) represent a significant difference among the five samples at $p < 0.05$. All the individual points were plotted in the bar plots. The

correlation coefficient (r) between RT-qPCR and RNA-seq for each unigene is shown with its corresponding significance level (**p < 0.01, *p < 0.05).

Supplementary Files

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

- [Additionalfile1TableS1.xls](#)
- [Additionalfile2TableS2.doc](#)
- [Additionalfile8TableS5.xls](#)
- [Additionalfile7TableS4.xlsx](#)
- [Additionalfile4FigureS1.doc](#)
- [Additionalfile3TableS3.doc](#)
- [Additionalfile13TableS6.txt](#)
- [Additionalfile9FigureS4.doc](#)
- [Additionalfile10FigureS5.doc](#)
- [Additionalfile5FigureS2.doc](#)
- [Additionalfile11FigureS6.doc](#)
- [Additionalfile16TableS7.doc](#)
- [Additionalfile12FigureS7.doc](#)
- [Additionalfile14FigureS8.doc](#)
- [Additionalfile15FigureS9.doc](#)
- [Additionalfile6FigureS3.doc](#)
- [Additionalfile16FigureS10.doc](#)