Putative Genes Identified in Alkaloids Biosynthesis in Dendrobium officinale by Correlating the Contents of Main Bioactive Metabolites with Genes Expression between Protocorm-like Bodies and Leaves

Zhaojian Wang
Anhui University of Traditional Chinese Medicine

Shihai Xing (✉ xshshihai@163.com )
Anhui University of Traditional Chinese Medicine

Weimin Jiang
Hengyang Normal University

Yingying Liu
Anhui University of Traditional Chinese Medicine

Xiaoxi Meng
University of Minnesota

Xinglong Su
Anhui University of Traditional Chinese Medicine

Mengyang Cao
Anhui University of Traditional Chinese Medicine

Liping Wu
Anhui University of Traditional Chinese Medicine

Nianjun Yu
Anhui University of Traditional Chinese Medicine

Daiyin Peng
Anhui University of Traditional Chinese Medicine

Research article

Keywords: Dendrobium officinale, Protocorm-like body, Polysaccharide, Alkaloid, Flavonoid

DOI: https://doi.org/10.21203/rs.3.rs-88139/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** *Dendrobium officinale*, an endangered Chinese herb, has extensive therapeutic effects and contains bioactive ingredients including a large number of polysaccharides and alkaloids, and minimal flavonoids. Firstly, this study attempts to obtain the protocorm-like bodies of this plant through tissue culture to produce the main secondary metabolites whose distribution in each organelle and protocorm like bodies is analyzed. Then, analysis of the correlation between comparative transcriptome sequence and the metabolite content in different organs enables the discovery of putative genes encoding enzymes involved in the biosynthesis of polysaccharides and alkaloids, and flavonoids.

**Results:** The optimum condition for protocorm-like bodies (PLBs) induction and propagation of *D. officinale* is established. For protocorm induction, we use the seed as the explant, and the optimum medium formula for PLBs propagation is 1/2 MS + α-NAA 0.5 mg·L⁻¹ + 6-BA 1.0 mg·L⁻¹ + 2, 4-D 1.5-2.0 mg·L⁻¹ + potato juice 100 g·L⁻¹. The distribution of polysaccharides, alkaloids and flavonoids in *D. officinale* organs was clarified. Stems, PLBs and leaves have the highest content of polysaccharides, alkaloids and flavonoids, respectively. PLBs replace organs to produce alkaloids in *D. officinale*, and naringenin was only produced in stem. Hot water extraction (HWE) method was found outperforming the ultrasound-assisted extraction (UAE) method for polysaccharides from *D. officinale*. A comparative transcriptome analysis of the protocorm-like bodies and leaves of *D. officinale* showed genes encoding enzymes involved in polysaccharides, alkaloids and flavonoids biosynthetic pathway were differentially expressed. Putative genes encoding enzymes involved in polysaccharides, alkaloids and flavonoids synthetic pathway were identified. Notably, genes encoding enzymes of strictosidine β-glucosidase, geissoschizine synthase and vinorine synthase in alkaloids biosynthesis of *D. officinale* are first reported.

**Conclusions:** Our works, especially the identification of candidate genes encoding enzymes involved in metabolites biosynthesis will help to explore and protect the endangered genetic resources and will also facilitate further analysis of the molecular mechanism of secondary metabolites’ biosynthesis in *D. officinale*.

1. Background

*Dendrobium officinale* Kimura et Migo is one of the most famous species in the genus Dendrobium, which is a precious perennial epiphytic herb in China and other Asian countries [1]. Stem of *D. officinale* has a long history as a Traditional Chinese Medicine (TCM) since Tang dynasty about 1,300 years ago with effects on tonifying stomach and nourishing fluid, clearing heat and nourishing “yin”, and improving immunological function [2]. Furthermore, *D. officinale* plays an important role in treating atrophic gastritis, diabetes, cancer, cardiovascular, disease-cataract and delaying aging [3]. Its bioactive components include major *D. officinale* polysaccharides (DOPs) [4] and alkaloids [5], and minimal flavonoids [6]. DOPs have many pharmacological activities such as curving hypoglycemia [7], antioxidation [8], immunomodulation [9] and anti-tumor. Alkaloids of *D. officinale* exhibit a variety of pharmacological
activities such as antioxidant, anticancer, and neuroprotective activities [10], and flavonoids of *D. officinale* possess good functions on anti-cytotoxicity and antioxidant [11].

Polysaccharides are major medicinal components in *D. officinale* [12]. The biosynthesis pathway of polysaccharides in medicinal plants can be divided into three main stages [13, 14]. At the first stage, sucrose produced by through chloroplast photosynthesis in leaves is converted into two kind of monosaccharides such as glucose 1-phosphate (Glc-1P) and fructose-6-phosphate (Fru-6P) under a series of enzymes. In the second stage, these two monosaccharides are extended into various nucleotide-diphospho-sugars (NDP-sugars, and NDP includes UDP and GDP) by enzymatic reactions. Finally, all of the NDP-sugars are extended into macromolecular polysaccharides by successive catalysis of glycosyltransferases (GTs) [15]. Searches against KEGG database enabled prediction of the biosynthesis pathway of DOPs as shown in Additional file 1 [16, 17]. The conversion of NDP-sugars into many types of DOPs in Golgi Complex was catalyzed by different kinds of glycosyltransferases such as glucosyltransferases, fucosyltransferases, mannosytransferases and xylosyltransferases in *D. officinale* [12, 18]. Water-soluble polysaccharides, including major mannose and minimal glucose were confirmed as the dominant DOPs in *D. officinale* stems according to the previous study [19]. Mannan polysaccharide was under the action of mannan synthases which were encoded by *Cellulose synthase-like A* (*CslA*) family members [20, 21] (Additional file 1).

Alkaloid was another type of major medicinal components in *D. officinale*, most of which pertain to terpenoid indole alkaloids (TIAs) [22]. Strictosidine backbone was formed in a relatively conserved upstream biosynthetic pathway of TIAs in plants as the precursor for various types of specific downstream alkaloids [23]. Then, strictosidine is catalyzed by enzymes to form enormous TIAs of *D. officinale* [24].

In Dendrobium, there might be a series of P450 monoxygenases superfamily and aminotransferases following strictosidine in the downstream biosynthetic pathway of TIAs [25]. Cytochrome P450s are involved in the oxidation and hydroxylation reaction, and aminotransferases transfer amino acids to form alkaloids which is the nitrogenous amino acid derivatives [26]. Study had found that polyneuridine-aldehyde esterase (EC: 3.1.1.78) gene (PNAE) participated in 16-Epivellosimine biosynthesis in *D. officinale* [12]. Based on KEGG database and chemical structure of the compounds we propose a putative alkaloids biosynthesis pathway of *D. officinale* (Additional file 2) [18, 22, 27]. Among the putative downstream specific alkaloid biosynthetic pathway of *D. officinale*, strictosidine is converted to strictosidine aglycone catalyzed by a specific β-D-Glucosidase (EC3.2.1.21) (SG), and the conversion of strictosidine aglycone to 4,21-dehydro-geissoschizine, which can be reciprocally transformed to geissoschizine by geissoschizine synthase (GS), is catalyzed by successive unknown enzymes. Vinorine might be one of alkaloids predicted by us in *D. officinale*, which is synthesized under the catalysis of vinorine synthase (VS) from 16-Epivellosimine.

Flavonoids are all originated from two intermediates (L-tyrosine and L-phenylalanine) produced by the shikimate pathway in plants, and these intermediates can generate p-coumarinyl-CoA via p-cinnamic acid
and cinnamoyl-CoA by enzymes catalysis, respectively [28]. Then, different types of flavonoids, such as flavone, flavanone, flavonol and anthocyanin are produced in plants by various types of enzymatic reactions from p-coumaryl-CoA [29]. Among these flavonoids, naringenin is an important flavonoid which has much clearly effects on human body, such as inhibiting the inflammatory response of many cell types [30], treating hypotensive [31], anti-fibrotic and anti-cancer, as well as the effect on hepatoprotective [32]. Naringin is generated by chalcone isomerase (CHI) through a ring closure reaction from naringin chalcone, which is condensed from three molecules of malonyl-CoA and one molecule of p-coumaryl-CoA by chalcone synthase (CHS) [33]. The putative biosynthesis pathway of flavonoids in *D. officinale* was shown in Additional file 3 [34].

*D. officinale* has already been endangered because of overexploiting and habitat destruction since its high medicinal value. More importantly, the wild *D. officinale* is now on the verge of extinction and on the IUCN Red List of Threatened Species [34]. Therefore, it is imperative to conserve wild resources of *D. officinale* and seek alternative ways to propagate it. Tissue culture technology may be an available method for *D. officinale* breeding and the induction of protocorms from embryogenic tissues and protocorm-like bodies (PLBs) from non-embryogenic tissues culture in *D. officinale* has been studied [35]. It was found the content of the main bioactive compound content of *D. officinale* in its PLBs varied with the medium formulas and the effective elicitors as additives [36].

In this study, we aimed to establish a system for protocorms and PLBs induction; proliferation for *D. officinale* in solid medium; and compare the contents of main metabolites such as polysaccharides, alkaloids, flavonoids and naringenin among PLBs and organs of the original plant. Significant differences in contents of these metabolites were revealed between PLBs and organs of plant, suggesting that the PLB and leaf were good experimental materials for analyzing the expression levels of candidate enzyme-encoding genes involved in the biosynthetic pathways of these specific metabolites in *D. officinale*. Based on the differential expression of genes (DEGs) from transcriptome sequencing, the putative genes encoding enzymes involved in biosynthetic pathways of the main metabolites have been identified or verified. This study will help to conserve *D. officinale* resources and deeply analyze the biosynthesis and regulation of main metabolites of *D. officinale*.

### 2. Results

#### 2.1 Protocorms and PLBs Induction of *D. officinale*

The optimal explants for PLBs induction and the optimum formula for PLBs propagation were screened in order to construct a comprehensive system for PLBs proliferation and protocorms/PLBs induction of *D. officinale*. Five different explants were selected to induce protocorms or PLBs. Induction rate and the growth status of protocorms or PLBs determined the best explants. The results showed that seed is most suitable for inducing protocorms or PLBs, and capsules approaching maturity were collected and punctured to release seeds in sterile condition because of the tiny seeds (Table 1).
Table 1
Effects of different explants on protocorms and PLBs induction

| Explants          | Seeds  | Leaves | Stem fragments | Stem tips | Stem with nodes |
|-------------------|--------|--------|----------------|-----------|-----------------|
| Induction rate (%)| 100 ± 0.0 | 43 ± 0.9 | 25 ± 1.8     | 52 ± 1.6  | 31 ± 1.9        |
| Growth status     | Well   | Common | Not good      | Common    | Common          |

An $L_{16}(4^5)$ orthogonal experiment was performed for screening the best protocol for PLBs propagation. The results showed that the best formula for PLBs propagation of *D. officinale* was $A_3B_3C_2D_1(2)E_2$, and factor C (6-BA content) is the most important factor because of its highest $R$ value. A verification experiment with formula $A_3B_3C_2D_2E_2$ was conducted with 40 duplicates, and the propagation coefficient is up to 4.37 (Table 2, Fig. 1a). Based on what is stated above, the optimal formula for PLBs propagation of *D. officinale* is $1/2$ MS + α-NAA 0.5 mg·L$^{-1}$ + 6-BA 1.0 mg·L$^{-1}$ + 2, 4-D 1.5-2.0 mg·L$^{-1}$ + potato juice 100 g·L$^{-1}$. 
Table 2
The effects of different factors on PLBs proliferation

| Levels | Media (A) | α-NAA (B) mg·L⁻¹ | 6-BA (C) mg·L⁻¹ | 2,4-D (D) mg·L⁻¹ | Potato juice (E) g·L⁻¹ | Propagation coefficient |
|--------|-----------|------------------|-----------------|-----------------|----------------------|------------------------|
| 1      | 1/4MS     | 0.1              | 0.5             | 1.5             | 50.0                 | 2.14                   |
| 2      | 1/4MS     | 0.2              | 1.0             | 2.0             | 100.0                | 3.98                   |
| 3      | 1/4MS     | 0.5              | 1.5             | 2.5             | 150.0                | 3.23                   |
| 4      | 1/4MS     | 1.0              | 2.0             | 3.0             | 200.0                | 1.98                   |
| 5      | 1/2MS     | 0.1              | 1.0             | 2.5             | 200.0                | 4.54                   |
| 6      | 1/2MS     | 0.2              | 0.5             | 3.0             | 150.0                | 1.83                   |
| 7      | 1/2MS     | 0.5              | 2.0             | 1.5             | 100.0                | 4.57                   |
| 8      | 1/2MS     | 1.0              | 1.5             | 2.0             | 50.0                 | 3.11                   |
| 9      | MS        | 0.1              | 1.5             | 3.0             | 100.0                | 3.42                   |
| 10     | MS        | 0.2              | 2.0             | 2.5             | 50.0                 | 2.75                   |
| 11     | MS        | 0.5              | 0.5             | 2.0             | 200.0                | 4.63                   |
| 12     | MS        | 1.0              | 1.0             | 1.5             | 150.0                | 5.02                   |
| 13     | 2MS       | 0.1              | 2.0             | 2.0             | 150.0                | 2.65                   |
| 14     | 2MS       | 0.2              | 1.5             | 1.5             | 200.0                | 1.93                   |
| 15     | 2MS       | 0.5              | 1.0             | 3.0             | 50.0                 | 4.87                   |
| 16     | 2MS       | 1.0              | 0.5             | 2.5             | 100.0                | 2.53                   |
| K1     | 2.83      | 3.19             | 2.78            | 3.60            | 3.22                 |
| K2     | 3.51      | 2.62             | 4.60            | 3.59            | 3.62                 |
| K3     | 3.96      | 4.32             | 2.92            | 3.26            | 3.18                 |
| K4     | 3.00      | 3.16             | 2.99            | 3.02            | 3.27                 |
| R      | 1.13      | 1.16             | 1.82            | 0.52            | 0.44                 |

2.2 Total Polysaccharides, Alkaloids and Flavonoids Content in PLBs and Organs of *D. officinale*
Polysaccharides, flavonoids and alkaloids were determined in various organs of *D. officinale*, including the whole plant, root, stem, leaf, and PLB (Fig. 1b). Significant differences in contents of these three components among samples were detected (Fig. 2).

### 2.2.1 Differential Contents of Total Polysaccharides (DOPs)

The regression equations of the D-glucose standard curve obtained by UAE and HWE methods of polysaccharides are $Y = 0.0491X - 0.0035$ ($R^2 = 0.9997$) and $Y = 0.0684X - 0.0323$ ($R^2 = 0.9998$), respectively (Additional file 4), with determination wavelength at 488 nm. The contents of polysaccharides in PLBs and organs measured by two different extraction methods had the same content distribution of polysaccharides, but the extraction rate of HWE method was higher than that of the UAE method (Fig. 2a and b). It can be concluded that the HWE method is better for DOPs extraction.

Different distributions of DOPs in the organs of plant are as follows: stem > the whole plant > leaf > root. The stem of *D. officinale* has the highest content of polysaccharides in all organs at 402.93 mg·g$^{-1}$ (Dried weight). While, PLBs has the lowest content of polysaccharides than organs with average amount of 75.30 mg·g$^{-1}$ (Fig. 3).

### 2.2.2 Differential Contents of Total Alkaloids

The equation of standard curve with dendrobine as reference is $Y = 0.1154X - 0.0048$ ($R^2 = 0.9990$) with the absorbance detected at 620 nm (Additional file 4). Significant difference of total alkaloids content among different parts of *D. officinale* was not observed (all around 0.3 mg·g$^{-1}$ (dried weight)). While, the total alkaloids contents in the PLBs is up to 0.56 mg·g$^{-1}$ (dried weight), which is almost twice as high as that of *D. officinale* plant organs (Fig. 2c). PLBs can be used as alternative materials for dendrobine extraction.

### 2.2.3 Differential Contents of Total Flavonoids and Naringenin

The regression equation of the standard curve for total flavonoids measured with rutin as the reference is $Y = 0.0114X - 0.0009$ with $R^2 = 0.9992$, and the absorbance is determined at 510 nm. The results showed that the content of total flavonoid in the stem of *D. officinale* was the lowest with only 4.07 mg·g$^{-1}$, which was consistent with most findings reported previously [34]. While, leaf had highest content of total flavonoids among organs at 11.74 mg·g$^{-1}$ followed by root (Additional file 4). The content of total flavonoids in PLBs is higher than that in the whole plant and stem, while lower than that in the leaf and root (Fig. 2d). Considering the lower biomass of leaf and root in *D. officinale*, PLBs would be the most suitable materials for extracting total flavonoids of *D. officinale*.

Naringenin, an important flavonoid, had been determinate by HPLC-UV and the chromatogram showed the retention time (RT) of naringenin standard was at 44.8 min (Fig. 3a). The regression equation of the
standard curve for naringenin with its standard is \( Y = 2624X + 0.6714 \), \( R^2 = 0.9999 \) (Fig. 3b). The result showed that naringenin could only be quantified only in the stem of \( D. officinale \), which is consistent with the previous study [37], and the average content is about 0.029 mg·g\(^{-1}\). And no naringenin could be quantified in the PLB of \( D. officinale \), which indicates PLB can't be an alternative material used for naringenin extraction (Fig. 3c). Based on similarity in naringenin content between the whole plant about (0.021 mg·g\(^{-1}\)) and the stem (about 0.029 mg·g\(^{-1}\)), we could conclude that naringenin derives mainly from the stem. We also established that the biomass of \( D. officinale \) was mainly from stem (Additional file 4).

### 2.3 Sequencing, Assembly and Unigenes Functional Annotation

Totally, 39.11 Gb transcriptomic data was generated by the platform of BGISEQ-500. Clean reads were assembled by Trinity software, and then cluster the transcripts to remove redundancy and obtain unigenes by Tgicl. Finally, 157,901 unigenes were obtained (Additional file 5). The total length, average length, N50, and GC content were 202,415, 413 bp, 1,281 bp, 2,194 bp, and 40.39%, respectively.

All unigenes assembled were mapped to the major functional databases for annotation. There were 99,474 (NR: 63.00%), 96,634 (NT: 61.20%), 62,695 (SwissProt: 39.71%), 68,459 (KOG: 43.36%), 71,648 (KEGG: 45.38%), 68,281 (GO: 43.24%) and 65,117 (Pfam: 41.24%) unigenes with functional annotations (Additional file 6). Moreover, 73,300 CDS were detected by software of Transdecoder, 36,670 SSRs were found distributed among 28,910 unigenes, and 2,893 unigenes encoding transcription factors were predicted.

### 2.4 Identification and KEGG Pathway Enrichment of Differential Expression Genes (DEGs)

148,314 unigenes can be quantified from the 157,901 unigenes assembled based on FPKM ≥ 1.0 in this study. 71,307 differentially expressed genes (DEGs) between leaf and PLB were identified by Poisson distribution methods with the \( \log_2 \) FC values ≥ 1.0 or ≤ -1.0 and false discovery rate (FDR) ≤ 0.001. Among these DEGs, 39,597 unigenes were up-regulated and 31,710 unigenes were down-regulated in PLB (Fig. 4a). Based on a Poisson distribution, 27,242 unigenes were specifically expressed in leaves, 11,976 unigenes were expressed only in PLB, and 98,096 shared unigenes were identified in both of these two samples (Fig. 4b, Additional file 7). The large number of identified DEGs exhibited significant difference between leaves and PLBs in various aspects, which directly reflected from the critical difference on appearance.

KEGG enrichment analysis of 71,307 DEGs identified in PLBs versus leaves enabled the discovery of 135 mapped pathways (Additional file 8), which were mainly enriched in “phenylpropanoid biosynthesis” and “flavonoid and isoflavonoid”, “carbon fixation and photosynthesis” and “fructose and mannose metabolism”, “indole alkaloid biosynthesis”, etc. (Fig. 4c).
2.5 Differential Expression of Genes Involved in Polysaccharides Biosynthesis Pathway

From above, polysaccharide is more abundant in leaves, with much lower abundance in PLBs. Differences in expression of genes encoding enzymes involved in polysaccharides biosynthesis between PLBs and leaves were further revealed (Fig. 5). Some of genes were expressed much higher in leaves such as genes encoding INV, CsIA, FBPase, PFP, MPI, scrK, KK, USP, UGPA, PMM, GMPP and UXS1 etc. The higher level of gene expression in leaves is positively correlated with the higher content of polysaccharide content in the leaves. Compared to PLBs, INV gene was highly expressed and SuS was mildly increased in leaves, which indicated that more sucrose was necessarily required to synthesize more downstream polysaccharides. The underexpression of gene encoding CsIA was positively correlated with the lower polysaccharides (mainly composed by mannose monosaccharides) content of in PLB (Fig. 5a).

Meanwhile, there was no significant difference in the expression of some genes encoding enzymes such as PGM, HK, TK, UXE, PFK, TSTA3 and UER1 between leaf and PLB, suggested the catalysis reactions by these enzymes are not the restrictive steps for DOPs biosynthesis. Genes encoding enzymes such as UAE, UGD, XK, AR and GMDS are negatively correlated with total DOPs content, and the enzymes are all bifunctional and in closed loop pathway (Fig. 5a). This might indicate that these bifunctional enzymes have greater degradation ability than their biosynthesis effect in polysaccharides biosynthesis.

Glycosyltransferases (GTs) are a very widespread group of carbohydrate-active related enzymes participating in glycan and glycoside biosynthesis in higher plants [38]. Totally, 553 glycosyltransferase genes, including 305 glucosyltransferase genes, 29 fucosyltransferase genes, 110 mannosyltransferase genes and 106 xylosyltransferase genes, were identified using BLASTX in the transcriptomes. From which, 22 full-length glucosyltransferases genes, 5 fucosyltransferase genes, 9 mannosyltransferase genes and 8 xylosyltransferase genes were isolated. The expression levels of differential gene encoding enzymes between leaf and PL, such as CGT, UGT76C1_2, SGT1, ALG11, were in line with the expression pattern of DOPs, suggesting those could be candidate genes encoding enzymes involved in DOPs biosynthesis in D. officinale (Fig. 5b, Additional file 9).

2.6 Analysis of Differential Genes Expression in Alkaloids Biosynthesis

Previous study showed alkaloids in D. officinale are mainly terpenoid indole alkaloids (TIAs). The difference in expression of genes encoding enzymes in upstream of MVA, MEP and Shikimate pathway between leaves and PLBs was analyzed. The result showed significant differences between leaves and PLBs, but no consistent regulatory patterns were detected among these genes (Fig. 6a). Disorder of genes expression indicates that the intermediates in upstream of TIAs biosynthesis are not the restrictive products, which was further verified by the expression of two genes encoding LAMT and TDC involved in synthesizing loganin and tryptamine, respectively. Total alkaloids content is higher in PLBs than in leaves, but the expression level of genes encoding enzymes LAMT and TDC was the opposite of alkaloids.
content. There's no significant difference in expression of gene encoding SCS which catalyzed the conversion of loganin into secologanin between the PLB and leaf. From what is stated above, we can conclude that most intermediates, especially loganin and secologanin, as direct substrates for strictosidine, the precursor of TIAs are sufficient for downstream specific alkaloids biosynthesis. The expression quantity of gene encoding enzyme STR is much higher in PLBs than in leaves, which means higher content of precursor strictosidine is necessary for producing higher content of total alkaloids.

For the downstream biosynthesis pathway of specific alkaloids in *D. officinale*, the expression level of putative genes encoding enzymes SG (Strictosidine beta-glucosidase), GS (Geissoschizine synthase) and PNAE (Polyneuridine-aldehyde esterase) are much higher in PLBs than in leaves, and the expression level of gene encoding VS (Vinorine synthase) has a small increase in PLB (Fig. 6a). Genes encoding SG, GS and VS are first predicted in *D. officinale* alkaloids biosynthesis. Moreover, these predicted genes involved in TIAs biosynthesis in *D. officinale* indicates total alkaloids of *D. officinale* might include strictosidine aglycone, geissoschizine and vinorine.

Some of aminotransferases and P450s superfamily enzymes are necessary for alkaloids biosynthesis. By comparing the expression level between PLBs and leaves, we found some of aminotransferases and P450s were higher expressed in PLBs than in leaves (Fig. 6b, c, Additional file 9). We speculate that some highly expressed aminotransferases and P450s might participate in alkaloids biosynthesis in *D. officinale*.

### 2.7 Differential Expression of Genes in Flavones Biosynthesis Pathway

Expression quantity of predicted genes encoding enzymes in flavones biosynthesis between PLB and leaf was measured by the value of diffexp_log2fc_Cas-vs -Ctr (Cas is PLB, Ctr is leaf) of high amino acid sequence identity unigenes from RNA-seq database. There's significant difference in gene expression between PLBs and leaves (Fig. 7). TAL and PAL were higher expressed in leaves than in PLBs, which accorded with the higher content of flavonoids in leaves. ADT and PDT were expressed much more than PAL and TAL in PLBs, suggesting that PLBs need produce more L-phenylalanine as substrate for downstream flavonoids biosynthesis. The lower expression level of ANS and ANR in leaves suggested low content of Cyanidin and Epicatechin in leaves of *D. officinale*.

Naringenin can't be quantified in PLBs. Based on low expression of CHI, high expression of DFR, and relatively high expression of F3’H and F3’5’H in PLBs, we can conclude that naringenin is just an intermediate product converted to downstream metabolites and is not stored in PLBs. The naringenin content in leaves of *D. officinale* was too low to be detected, but higher expression of CHI, F3’H and F3’5’H, and lower expression of DFR indicated the naringenin might be transferred to stem and converted to downstream flavonoids biosynthesis. Naringenin can be detected only in the stem of *D. officinale*. There is possibility that a special organelle exists in the stem, functioning in accumulating naringenin and transferring naringenin from the leaf to stem in *D. officinale*. HCT and HST are higher expressed in PLBs.
than in leaves, suggesting most of Eriodictyol in PLBs comes from Caffeoyl-CoA, which consistent with relatively lower expression of F3’H and F3’5’H, and low naringenin content in PLBs.

2.8 Detection of Transcription Factor Families

Transcriptional factors (TFs) play an important role in the regulation of secondary metabolites by coordinating the expression of biosynthetic genes and controlling the production of secondary metabolites in plants [39]. In this study, by comparing to the NCBI database, a total of 1,113 unique putative transcripts encoding TFs were identified as differentially expressed genes, and these TFs belong to 51 known TF families (Additional file 10). The major types of these differentially expressed TFs are MYB, AP2/EREBP, bHLH and NAC etc. The number of upregulated and downregulated TFs from each type was shown in Fig. 8, and a striking number of differentially expressed TFs were identified in between PLBs and leaves.

3. Discussion

It had been reported that the extraction method has little effect on the types of monosaccharide, but influences the monosaccharide composition of heterogeneous polysaccharides [40]. By using the HWE method, the diffusion rate of polysaccharides could increase dramatically due to the raised temperature of water, hence the improvement of extraction efficiency [41]. Our result indicates that the HWE method is more suitable for DOPs extraction, which is consistent with the above conclusion.

Among all organs of D. officinale, the stem has the highest content of polysaccharides, which is consistent with the conclusion from previous researchers [22]. While, PLBs has the lowest content of polysaccharides, which indicates that the PLBs couldn’t substitute for the plant or organs of D. officinale in production of DOPs.

In spite of the relatively low content of total alkaloids in D. officinale, more and more attention has been paid to their broad pharmacological activities, such as antipyretic, analgesic, and anti-tumor activities [12]. PLBs were much more suitable for producing alkaloids in our study, which provided a new idea for the future large-scale production of alkaloids in D. officinale.

KEGG enrichment analysis indicated that comparative RNA-Seq between leaves and PLBs was consistent with the significant difference in content of DOPs, alkaloids and flavonoids between PLBs and leaves of D. officinale. It also demonstrates our reasonable prediction of the putative biosynthesis pathway of DOPs, alkaloids and flavonoids in D. officinale.

The characterization of the correlation between the content of DOPs, alkaloids and flavonoids and comparative transcriptome sequence results in discovering many putative genes encoding enzymes involved in biosynthesis of DOPs, alkaloids and flavonoids in D. officinale. Next, we are going to work on the clone of the genes encoding SG, GS, PNAE and VS from D. officinale, and the construction of expression vectors to make the proteins expressed. The experiment on enzyme activities in vitro will be
carried out to validate their function, and the substrate specificity of these putative enzymes and their subcellular localization also need to be studied.

Previous reports had indicated that TFs played important roles in regulating alkaloids biosynthesis, such as TFs of apetala2/ethylene response factors (AP2/ERFs) [42, 43], C2H2 zinc fingers [44], WRKYs [45] and basic helix-loop-helix (bHLH) [46] that are involved in regulating TIAs biosynthesis. ORCA3, an AP2/ERF TF, induces the expression of a number of TIAs biosynthetic genes such as G10H, CPR, STR, AS, TDC, and DXS in the vindoline or tryptamine branch pathway [47, 48]. CrMYC belongs to bHLH family, activates ORCA3, and then induces the expression of several TIAs biosynthetic genes, such as TDC and CPR [49]. CrBPF1 is a MYB transcription factor that regulates the expression of STR in *C. roseus* [50]. And in our study, some TFs such as MYB, AP2/EREBP, bHLH and NAC were found involved in alkaloids biosynthesis in *D. officinale*.

### 4. Conclusions

A best protocol for PLBs induction in *D. officinale* is using seed as explants, and 1/2 MS + α-NAA 0.5 mg·L⁻¹ + 6-BA 1.0 mg·L⁻¹ + 2, 4-D 1.5-2.0 mg·L⁻¹ + potato juice 100 g·L⁻¹ as the full medium.

The distribution of polysaccharides, alkaloids and flavonoids among the organs of *D. officinale* plant is clarified in this study. Content of Polysaccharides, alkaloids and flavonoids was the highest in stems, PLBs and leaves, respectively. Hot water extraction (HWE) is the better method for polysaccharides extraction in *D. officinale*. PLBs is an alternative for producing alkaloids of *D. officinale*. Naringenin is produced exclusively in stems, with no detectable amount from PLBs.

The comparative RNA-Seq analysis of *D. officinale* PLBs and leaves enabled the identification of some candidate genes encoding enzymes involved in polysaccharides, alkaloids and flavonoids biosynthesis. Especially, genes encoding enzymes such as SG, GS and VS in alkaloids biosynthesis of *D. officinale* are the first reported. Moreover, some TFs involved in regulating alkaloids biosynthesis were found differentially expressed in leaves and PLBs of *D. officinale*.

### 5. Methods

#### 5.1 Plant Materials and Reagents

Plants of *D. officinale* were grown at 25°C during day and 23°C at night with 60–70% relative humidity and a light/dark cycle of 14/10 h in the herbal garden of Anhui University of Chinese Medicine, Hefei, China.

Standard products of (+)-Glucose, Dendrobine, Rutin and Naringenin (> 98% purity) were bought from Chengdu Push Bio-technology CO., Ltd. Methanol (HPLC grade) were purchased from Oceanpak, and CTAB-PBIOZOL reagent used for total RNA extraction was purchased from Bioflux (Beijing, China). HPLC-grade water was prepared by water purification system from Pall Filter Co., Ltd. (Beijing, China). Plant
growth regulators (PGRs) include 6-BA (6-Benzylaminopurine), α-NAA (α-Naphthylacetic acid) and 2, 4-D (2, 4-dichlorophenoxyacetic acid) were purchased from Solarbio. Agents for PLBs induction and propagation were bought from Sinopharm group and all of the other agents were analytical grade.

5.2 Protocorms and PLBs of *D. officinale* Induction and Propagation

Capsules approaching mature, leaves, stem tips, stem fragments, and stems with nodes of *D. officinale* were collected, and sterilized to induce protocorms and PLBs (Capsules were punctured to release seeds as explants on medium under sterile condition), each explant was performed more than 30 duplicates. They were inoculated in 1/2 MS medium + 2, 4-D 0.5 mg·L\(^{-1}\) + 30 g·L\(^{-1}\) sucrose + 7 g·L\(^{-1}\) agar at pH 5.6 to 5.8, and placed in the tissue culture rooms at 70% relative humidity, 1600 lx illumination and a light/dark of 16/8 h for 60 d, to screen the optimal explants for protocorms and PLBs induction.

Protocorms and PLBs induced above were collected as explants to be cut into pieces of about 0.5 cm × 0.5 cm to proliferate much more PBLs, and the optimum culture protocol for the PLBs of *D. officinale* proliferation was screened. An L\(_{16}(4^5)\) orthogonal experiment without regarding the interactions among factors was performed, the factors and their levels were selected based on our previous experiments and the reports on literatures [51]. Four different gradients of MS medium such as 1/4MS, 1/2MS, MS and 2MS were used as test media (Factor A), the plant growth regulators (PGRs) combinations with different concentration of α-NAA (Factor B) with concentrations of 0.1 mg·L\(^{-1}\), 0.2 mg·L\(^{-1}\), 0.5 mg·L\(^{-1}\) and 1.0 mg·L\(^{-1}\), 6-BA (Factor C) with 0.5 mg·L\(^{-1}\), 1.0 mg·L\(^{-1}\), 1.5 mg·L\(^{-1}\) and 2.0 mg·L\(^{-1}\) and 2,4-D (Factor D) with 1.5 mg·L\(^{-1}\), 2.0 mg·L\(^{-1}\), 2.5 mg·L\(^{-1}\) and 3.0 mg·L\(^{-1}\) were arranged to screen the optimal PGRs combination. It was found that the materials in tissue culture with natural additives and fruit juice grow much better than those without natural additives and fruit juice [52, 53], in the PLBs proliferation processing potato juice was selected as additive to accelerate PLB growth and propagation. 200 g fresh potato was chopped into small pieces, added was mixed with purified water to 1 L and boiled to a mush. The mushy potato was passed through the gauze to collect the juice as additive. The concentration gradients of potato juice (Factor E) were 50 g·L\(^{-1}\), 100 g·L\(^{-1}\), 150 g·L\(^{-1}\) and 200 g·L\(^{-1}\). All of the full test media with pH 5.6 ~ 5.8 for PLBs proliferation were included basic medium, PRGs combination, potato juice, sucrose (30 g·L\(^{-1}\)) and agar (7 g·L\(^{-1}\)), and the design of orthogonal experiment table head was showed in Additional file 11. Each group has 10 bottles with 4 pieces, and the culture condition was light/dark for 16/8 h per day, the illuminance and the culture temperature were at 1600 lx and (25 ± 1) °C. Observations and statistics were carried out to analyze the propagation coefficient two months later since inoculation. The formula of propagation coefficient was calculated as follow: Propagation coefficient (%) = (FW\(_1\) − FW\(_0\))/ FW\(_0\) × 100%, FW\(_0\) and FW\(_1\) mean the fresh weight before inoculation and after culture, respectively.

5.3 Ultrasound-Assisted Extraction (UAE) or Hot Water Extraction (HWE) of Total Polysaccharides
Stems, leaves, roots and the whole plants from about 3 years old *D. officinale* and PLBs were harvested, cleaned, and dried in an oven at 50 °C to a constant weight. Samples were ground to a fine powder using the pulverizer and sieve the powder through a 40 meshes sieve. The water-soluble polysaccharides in *D. officinale* were extracted by UAE method [41] and HWE method [8] with a few modification and finally better one was selected for the extraction of polysaccharides in *D officinale* from these two methods.

The modified UAE method was as follows: 10 mL of double distilled water was added into each 0.2 g pulverous sample, and the samples were homogenized in the ultrasonic cleaner at 40 °C, 40 Hz for 0.5 h, then filtered through filter-paper and the filtrate was obtained. Repeat above steps one more time and collect all filtrate. The filtrate was concentrated at 55°C until its volume was down to 10 mL. 40 mL of ethanol was added into the concentrated filtrate, centrifuged at 4,000 rpm for 5 min, and discarded the supernatant. The samples were dissolved in 25 mL of pure water. 25 mL of savage reagent was added to remove impurities such as protein and nucleic acid, centrifuged at 1000 rpm for 5 min. 20 mL of the supernatant was transferred into a 50 mL volumetric flask and mixed with pure water to a constant volume of 50 mL for polysaccharides extraction use. The method HWE with a few modifications was as follows: firstly, 0.3 g samples were placed in a round-bottomed ask, and add 200 mL of water. Then the sample solution was heated and refluexed for 3 h. Let the reflux extract cool off to room temperature before filtering it. The filtrate was transferred into a 250 mL volumetric flask, and mixed with double distilled water to a constant volume of 250 mL. 2 mL of sample solution was taken into a 15 mL centrifugal tube and 10 mL of solute ethanol was added, shake it and refrigerate it for 1 h followed by centrifugation at 1000 rpm for 20 min. Discard the supernatant and wash the precipitate twice with 8 mL of 80% ethanol. The final precipitate was dissolved in heated water and transferred to a 25 mL volumetric flask. Let the solution cool off before mixing it with pure water to a constant volume of 25 mL for polysaccharides extraction use. Thus, the polysaccharide extracts by the two methods have been prepared, respectively.

### 5.4 Determination of Total Polysaccharides

1 mL of each polysaccharide extract was transferred into a 10 mL test tube, then add 1 mL of 5% phenol and vortex quickly. The solution was mixed thoroughly and 5 mL of concentrated sulfuric acid was added, shaken, and placed in a water bath at 100°C for 20 min. Then the solution was placed in ice bath to cool for 5 min. The absorbance of the sample solution was measured at wavelength of 488 nm using ultraviolet visible spectrophotometer with 1 mL of water as a blank and the parallel detection was conducted three times. The standard curve was prepared from the D-glucose reference. Each sample was assayed in three times and the content showed as the weight of polysaccharides to the dried weight of materials.

### 5.5 Extraction and Determination of Total Alkaloids

The sample powder (0.5 g) was put into a 125 mL ground-mouth flask and mixed with 30 mL of petroleum ether. The sample mixture was placed in a constant temperature water bath (35°C) to degrease for 30 min, then remove the supernatant and evaporate the petroleum ether. The pH value was
adjusted with a proper amount of ammonium hydroxide and add 10.0 mL of chloroform. Then the solution was refluxed for 2 h in a water bath at 80 °C. Let it cool for 20 min in room temperature before filtering. The filtrate was obtained as total alkaloids solution.

5 mL of total alkaloids solution was mixed with 5 mL of chloroform, then 5.0 mL of pH = 4.5 buffer and 2.0 mL of 0.04% bromocresol green solution were added sequentially. Shake the mixture vigorously for 3 min and let it stand for 30 min before filtering it. And 5.0 mL of filtrate was mixed with 1.0 mL of 0.01 M NaOH-ethanol solution and vortex it. The absorbance of the samples was measured at 620 nm using UV-visible spectrophotometer with chloroform as blank. The standard curve was prepared from the Dendrobine as reference (> 98% purity) bought from Chengdu Push Bio-technology CO., Ltd. (0.857, 1.714, 2.571, 3.429, 4.286, and 5.143 µg·mL⁻¹). The parallel detection was carried out three times and the content showed as the weight of total alkaloids to the dried weight of materials.

### 5.6 Extraction and Determination of Total Flavonoids

The samples powder (1.00 g) was dissolved in 50 mL of 70% ethanol and refluxed at 60°C for 2 h, then filtered. The filtrate was mixed with 70% ethanol again to a constant volume of 50 mL as total flavonoids extracts.

1 mL of total flavonoids extract was mixed with 5 mL of 70% ethanol. Then add 1 mL of 5% NaNO₂ and mix it well. Stand for 6 min and add 1 mL of 10% Al(NO₃)₃ and mix it well. Let it stand for 6 min before adding 10 mL of 1 M NaOH and mix it with 70% ethanol to 25 mL, and stand for 15 min. The absorbance of the sample solution was measured at 510 nm using ultraviolet visible spectrophotometer. The standard curve was prepared from Rutin as reference (2.208, 4.416, 6.624, 8.832, and 11.040 µg·mL⁻¹). Each sample was assayed in three times and the content showed as the weight of total flavonoids to the dried weight of materials.

### 5.7 Extraction and Quantification of Naringenin

Naringenin was extracted by the optimized methods: Solvent, methanol (20 mL) mixed with 20% hydrochloric acid (5 mL); Particle size of dry powder of *D. officinale* < 0.355 mm; temperature, 80°C; Condensation reflux extraction time, 90 min; 20 mL of supernatant was evaporated using the rotating evaporator until it was dry. 5 mL of methanol was added to dissolve naringenin and the solution was filtered through a 0.22 mm Nylon membrane filter before HPLC analysis.

HPLC analysis was performed on an Agilent HPLC system, including quaternary solvent management, sampler manager, separation system, detection systems. An Agilent C₁₈ column (4.6 mm × 250 mm, 5 µm) was used at a column temperature of 30°C. The standards and samples were separated using a gradient mobile phase consisting of 0.2% phosphate buffer (A) and methanol (B). The gradient elution program was as follows: 0.0–5.0 min, 0–25% B; 5.0–10.0 min, 25–30% B; 10.0–25.0 min, 30–40% B; 25.0–45.0 min, 40–55% B; 45.0–50.0 min, 55–60% B. The flow rate was 1 mL·min⁻¹, and the naringenin was detected at 280 nm. The injection volume was 10 µL.
Naringenin stock solution was prepared and diluted to an appropriate concentration for the preparation of calibration curve. The calibration curve was prepared according to the linear plots of the naringenin concentration versus the corresponding chromatographic peak area.

5.8 Total RNA Extraction, Construction of cDNA Libraries and RNA-Seq

Ethanol precipitation protocol and CTAB-PBIOZOL reagent was used to purify total RNA from PLBs and leaves according to the manual instructions, and each sample has three biological duplicates. 80 mg tissue samples were ground into powder in liquid nitrogen and the powdered samples were transferred into 1.5 mL of preheated 65°C CTAB-pBIOZOL reagents. The sample solution was incubated by thermostatic mixer at 65 °C for 15 min to completely dissociate the nucleoprotein complexes. The supernatant was obtained by centrifuging the solution at 12000 × g, at 4°C for 5 min. The supernatant was mixed with 400 uL of chloroform (per 1.5 mL of added CTAB-pBIOZOL reagent), and the sample mixture was centrifuged at 12000 × g, at 4°C for 10 min. The supernatant was transferred to a new 2 mL tube and added 700 µL of acidic phenol and 200 µL of chloroform into it followed by centrifugation at 12000 × g, at 4°C for 10 min. The aqueous supernatant was collected and an equal volume of chloroform was added into it followed by centrifugation at 12000 × g, at 4°C for 10 min. An equal volume of isopropyl alcohol was added into the supernatant and the mixture was placed at -20°C for 2 h to precipitate before centrifuging it at 12000 × g, at 4°C for 20 min, and the supernatant was removed. The RNA pellet was washed with 1 mL of 75% ethanol before being air-dried in a biosafety cabinet and dissolved in 50 uL of DEPC-treated water. Subsequently, qualification and quantification of total RNA were performed using a Nano Drop and Agilent 2100 bioanalyzer (Thermo Fisher Scientific, MA, USA).

Oligo(dT)-attached magnetic beads were used to purify mRNA. The purified mRNA was split up into small pieces with fragment buffer at an appropriate temperature. The random hexamer-primed reverse transcription was performed to synthesize the first-strand cDNA followed by the second-strand cDNA synthesis. Then, A-Tail Mix and RNA Index Adapter were added to end the repair. The cDNA fragments obtained from above mentioned steps were amplified by PCR. The products were purified by Ampure XP Beads and dissolved in EB solution. The product was verified for quality control on an Agilent 2100 bioanalyzer. The double-stranded PCR products were denatured and cycled by the splint oligo-nucleotide sequence to obtain the final library, and the single-stranded circular DNA (ssCir DNA) was formatted into the final library. The established library was amplified with phi29 to obtain DNA nanoball (DNB), each molecular of which had over 300 copies. The DNBs were added into the patterned nanoarray and 100 base pairs of reads generated on the BGSeq500 platform.

5.9 Functional Annotation of Unigenes by Reference Genome

The original sequencing data (original polymerase reads) produced from raw data by Pacific Bioscience Sequel were processed by the SMRT analysis package version 2.3.0 according to the IsoSeq protocol (Pacific Biosciences, https://www.pacb.com/products-and-services/analytical-software/smrt-analysis).
ROIs (reads of insert) were generated by original Polymerase reads, which had full passes > 0 and the predicted consensus accuracy > 0.75. According to whether the primers are 5’primers, 3’primers or poly-A tails, ROIs with a minimum length of 300 bp were divided into non-full-length and full-length transcribed sequences. The full-length sequences were processed to De novo consensus isoforms by ICE (Iterative Clustering for Error Correction) algorithm and then polished by Quvier quality-aware algorithm. The De novo consensus isoforms of high quality (the expected Quiver accuracy ≥ 0.95) from each library were combined and then rid redundancy by using CD-HIT [54] founded on the sequence similarity to get final unique full-length isoforms.

The final complete subtype was mapped to SwissProt (manually annotated and reviewed protein sequence database), NR (NCBI non-redundant protein sequence), KEGG (Kyoto Encyclopedia of Genes and Genomes), NT (NCBI non-redundant nucleotide sequence) and the KOG (Clusters of Eukaryotic Orthologous Groups) database through Blast software (version 2.2.23) [55] with default parameters (under a threshold E-value ≤ 10−5) to obtain the isoform annotations. GO (Gene Ontology) annotations and functional classifications were acquired using Blast2GO program (version 2.5.0, E-value ≤ 10−5) [56] according to NR annotations. InterProScan5 software (version 5.11-51.0) [57] was used to acquire annotations from InterPro database.

5.10 Identification of Differentially Expressed Genes (DEGs)

All clean reads were mapped to the reference full-length transcriptome using BLAST software. Gene expression levels were determined by the number of full-length transcriptions, which belong to a cluster after ice clustering and CD-HIT process. Then use the total isoforms’ counts to normalize the counts in each sample. Differentially expressed genes (DEGs) were acquired using DEseq2 with Q value (adjust P value) < 0.001 and fold change (FC) ≥ 2 or ≤ -2 [58]. These DEGs were then carried into GO and KEGG enrichment with Phyper in R package using Q value ≤ 0.05 as default.

5.11 Comparative Analysis of Putative Genes Expression in Metabolites Biosynthesis Pathway

The amino acid sequences of putative genes involved in polysaccharides, alkaloids and flavonoids biosynthetic pathway were searched in National Center for Biotechnology Information (NCBI) database (Additional file 12). All amino acid sequences of putative genes were mapped to the protein information from RNA-Seq using BLAST software to get the average FPKM value of each genes in each sample. Heatmaps of putative genes involving in polysaccharides, alkaloids and flavonoids biosynthesis were drawn based on FPKM value.

Statistical analysis

The experimental data are expressed as the mean ± standard deviation of three independent biological replicates. The statistical differences between samples were analyzed using two-way analysis of variance by SPSS (version 17.0). Values at P < 0.01 were considered statistically significant.

Abbreviations
TCM: Traditional Chinese Medicine; DOPs: *Dendrobium officinale* polysaccharides; Glc-1P: Glucose 1-phosphate; NDP-sugars: Nucleotide-diphospho-sugars; GTs: Glycosyltransferases; TIA:s: Terpenoid indole alkaloids; DEGs: Differential expression of genes; PGRs: Plant growth regulators; UAE: Ultrasound-assisted extraction; HWE: Hot water extraction; NCBI: National Center for Biotechnology Information; RT: Retention time; INV: β-fructofuranosidase; HK: Hexokinase; HK: Hexokinase; PGM: Phosphoglucomutase; XI: Xylose isomerase; AR: Aldose reductase; SDH: Sorbitol dehydrogenase; SDH: Sorbitol dehydrogenase; KK: Ketohexokinase; HK: Hexokinase; PFP: Diphosphate-dependent phosphofructokinase; PFK: 6-phosphofructokinase; FBPase: Fructose-1,6-bisphosphatase; X: Xylosidase; UGE: UDP-glucuronate 4-epimerase; UXS1: UDP-glucuronate decarboxylase; UDP-glucose-1-phosphate uridylyltransferase; UGE: UDP-arabinose 4-epimerase; UER1: 3,5-Epimerase/4-reductase; GT: UDP-galacturonate decarboxylase; RHM: UDP-glucose 4,6-dehydratase; scrK: Fructokinase; MPI: Mannose-6-phosphate isomerase; PMM: Phosphomannomutase; GMPP: Mannose-1-phosphate Guananyltransferase; GMDS: GDP-mannose 4,6-dehydratase; TSTA3: GDP-I-fucose synthase; UAE: UDP-glucuronate 4-epimerase; UGD: UDP-glucose 6-dehydrogenase; GGT: UDP-galacturonate decarboxylase; FKP: Fructose-2,6-bisphosphatase; SuS: Sucrose synthase; RMD: GDP-4-dehydro-D-ribose reductase; PGI: Glucose-6-phosphate isomerase; CslA: Cellulose synthase-like A; USP: UDP-sugar pyrophosphorylase; AACT: Acetyl-CoA acetyltransferase; HMGS: HMG-CoA synthase; HMG: HMG-CoA reductase; MVK: Mevalonate kinase; PMK: Phosphomevalonate kinase; MVD: Mevalonate diphosphate decarboxylase; IPI: IPP isomerase; DXS: 1-deoxyxylulose-5-phosphate synthetase; DXR: 1-deoxy-D-xylulose-5-phosphate reductoisomerase; CMS: 4-diphosphocytidyl-2C-methyl-D-erythritol synthase, CMK: 4-diphosphocytidyl-2C-methyl-D-erythritol kinase; HDR: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase; MCS: 2-C-methyl-D-erythritol 2,4-cyclophosphate synthase; HDS: 1-hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate synthase; GPPS: Geranyl diphosphate/farnesyl diphosphate synthase; GES: geranyl diphosphate diphosphatase; CPR: NADPH-hemoprotein reductase; G10H: Geraniol 8-hydroxylase; UGT8: 7-deoxyloganetic acid glucosyltransferase; LAMT: Loganate 0-methyltransferase; 7-DLH: 7-deoxloganate 7-hydroxylase; SCS: Secologanin synthase; DHS: 3-deoxy-D-xylulose-5-phosphate synthase; DHQS: 3-dehydroquinate synthase; DHQ: 3-dehydroquinate dehydratase; SKDH: Shikimate dehydrogenase; QSDH: Quinate/shikimate dehydrogenase; SK: Shikimate kinase; EPSP: 3-phosphoshikimate 1-carboxyvinyltransferase; CS: Chorismate synthase; AS: Anthranilate synthase; PRAI: Phosphoribosylanthranilate isomerase; TrpD: Anthranilate phosphoribosyltransferase; IGPS: Indole-3-glycerol phosphate synthase; TS: Tryptophan synthase; TDC: Tryptophan decarboxylase; STR:Strictosidine synthase; SG: Strictosidine beta-glucosidase; GS: Geissoschizine synthase; PNAE: Polyneuridine-aldehyde esterase; VS: Vinorine synthase; CM: chorismate mutase; ARAT: Aromatic-amino-acid transaminase; ADT: Arogenate dehydratase; PDT: Prephenate dehydratase; NADP+: Arogenate dehydrogenase; ADH: Cyclohexadienyl dehydrogenase; PAL: Phenylalanine ammonia lyase; 4CL: 4-coumarate CoA ligase; TAL: Tyrosine ammonia lyase; C4H: Cinnamate 4-hydroxylase; HCT: Quinate O-hydroxycinnamoyltransferase; HST: Shikimate O-hydroxy cinnamoyltransferase; CHS: Chalcone synthase; CHI: Chalcone isomerase; DFR: Dihydroflavonol-4-reductase; F3H: Flavonoid 3-hydroxylase; F3’H: Flavonoid 3’-monooxygenase; F3’5’H: Flavonoid 3’,5’-hydroxylase; DFR: Dihydroflavonol 4-reductase;
CCoAOMT: Caffeoyl-CoA O-methyltransferase; ANS: Anthocyanidin synthase; ANR: Anthocyanidin reductase; FLS: Flavonol synthase; C3'H: 5-O-(4-coumaroyl)-D-quinate 3'-monooxygenase.

**Declarations**

**Competing interests**

The authors declare the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Funding**

This work was supported by the National Natural Science Foundation of China (No. U19A2009), NSF of Anhui Province (No. 1908085MH268), Key Natural Science Research Projects in Anhui Universities (No. KJ2019A0453), Hunan Provincial Natural Science Foundation of China (No. 2018JJ3008), Returnee Program of Anhui People's Society Office (No. DT18100035) and Key Project of Scientific Research Fund of Anhui University of Traditional Chinese Medicine (No. 2018zzrzd05).

**Authors’ contributions**

Conceived, designed, and implemented the study: SX, ZW, XS, WJ; Statistics analysis: SX, ZW and YL; Reagents/materials/analysis tools: SX, DP, JY and YL; Drafted the manuscript: XS, ZW, XM, WJ; All authors edited the manuscript and approved the final version.

**Acknowledgements**

We would like to thank Dianyi Shi from FeiTai Bio-tech (Shanghai) Co., Ltd in China for bioinformatics analyzing, we also thank BGI-Shenzhen in China for RNA-Seq sequencing.

**References**

1. Li X, Ding X, Chu B, Zhou Q, Ding G, Gu S. Genetic diversity analysis and conservation of the endangered Chinese endemic herb Dendrobium officinale Kimura et Migo (Orchidaceae) based on AFLP. Genetica. 2008;133(2):159–66.

2. Xu M, Liu X, Wang JW, Teng SY, Shi JQ, Li YY, Huang MR. Transcriptome sequencing and development of novel genic SSR markers for Dendrobium officinale. Mol Breeding. 2017;37(2):18.

3. Xu XB, Ma XY, Lei HH, Song HM, Ying QC, Xu MJ, Liu SB, Wang HZ. Proteomic analysis reveals the mechanisms of Mycena dendrobii promoting transplantation survival and growth of tissue culture.
seedlings of Dendrobium officinale. J Appl Microbiol. 2015;118(6):1444–55.

4. Tang H, Zhao T, Sheng Y, Zheng T, Fu L, Zhang Y. Dendrobium officinale Kimura et Migo: A Review on Its Ethnopharmacology, Phytochemistry, Pharmacology, and Industrialization. Evid Based Complement Alternat Med. 2017;2017:7436259.

5. Li C, Qiu J, Ding L, Huang M, Huang S, Yang G, Yin J. Anthocyanin biosynthesis regulation of DhMYB2 and DhbHLH1 in Dendrobium hybrids petals. Plant Physiol Biochem. 2017;112:335–45.

6. Tao Y, Cai H, Li W, Cai B. Ultrafiltration coupled with high-performance liquid chromatography and quadrupole-time-of-flight mass spectrometry for screening lipase binders from different extracts of Dendrobium officinale. Anal Bioanal Chem. 2015;407(20):6081–93.

7. Liu Y, Yang L, Zhang Y, Liu X, Wu Z, Gilbert RG, Deng B, Wang K. Dendrobium officinale polysaccharide ameliorates diabetic hepatic glucose metabolism via glucagon-mediated signaling pathways and modifying liver-glycogen structure. J Ethnopharmacol 2020, 248.

8. Huang K, Li Y, Tao S, Wei G, Huang Y, Chen D, Wu C. Purification, Characterization and Biological Activity of Polysaccharides from Dendrobium officinale. Molecules 2016, 21(6).

9. Huang X, Nie S, Cai H, Zhang G, Cui SW, Xie M, Phillips GO. Study on Dendrobium officinale O-acetylgalactomannan (Dendronan): Part IV. Immunomodulatory activity in vivo. Journal of Functional Foods. 2015;15:525–32.

10. Ng TB, Liu J, Wong JH, Ye X, Wing Sze SC, Tong Y, Zhang KY. Review of research on Dendrobium, a prized folk medicine. Appl Microbiol Biotechnol. 2012;93(5):1795–803.

11. Wen L, Lin Y, Lv R, Yan H, Yu J, Zhao H, Wang X, Wang D. An Efficient Method for the Preparative Isolation and Purification of Flavonoids from Leaves of Crataegus pinnatifida by HSCCC and Pre-HPLC. Molecules 2017, 22(5).

12. Liang Y, Xiao W, Hui L, Yang T, Lian J, Yang R, Hao S: The Genome of Dendrobium officinale Illuminates the Biology of the Important Traditional Chinese Herb. 2015.

13. Xie Y, Zhou H, Liu C, Zhang J, Li N, Zhao Z, Sun G, Zhong Y. A molasses habitat-derived fungus Aspergillus tubingensis XG21 with high beta-fructofuranosidase activity and its potential use for fructooligosaccharides production. AMB Express. 2017;7(1):128.

14. Uematsu K, Suzuki N, Iwamae T, Inui M, Yukawa H. Expression of Arabidopsis plastidial phosphoglucomutase in tobacco stimulates photosynthetic carbon flow into starch synthesis. J Plant Physiol. 2012;169(15):1454–62.

15. Breton C, Šnajdrová L, Jeanneau C, Koča J, Imberty A. Structures and mechanisms of glycosyltransferases. Glycobiology. 2006;16(2):29R–37R.

16. Zhang J, He C, Wu K, Silva JATd, Zeng S, Zhang X, Yu Z, Xia H, Duan J. Transcriptome Analysis of Dendrobium officinale and its Application to the Identification of Genes Associated with Polysaccharide Synthesis. Front Plant Sci 2016, 7.

17. Yuan Y, Zhang J, Kallman J, Liu X, Meng M, Lin J. Polysaccharide biosynthetic pathway profiling and putative gene mining of Dendrobium moniliforme using RNA-Seq in different tissues. BMC Plant Biol. 2019;19(1):521.
18. Zhang G, Xu Q, Bian C, Tsai W, Yeh C, Liu K, Yoshida K, Zhang L, Chang S, Chen F, et al. The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Sci Rep. 2016;6(1):19029.

19. Xing X, Cui SW, Nie S, Phillips GO, Goff HD, Wang Q. Study on Dendrobium officinale O-acetyl-glucomannan (Dendronan®): Part II. Fine structures of O-acetylated residues. Carbohydr Polym. 2015;117:422–33.

20. He C, Wu K, Zhang J, Liu X, Zeng S, Yu Z, Zhang X, Teixeira da Silva JA, Deng R, Tan J, et al: Cytochemical Localization of Polysaccharides in Dendrobium officinale and the Involvement of DoCSLA6 in the Synthesis of Mannan Polysaccharides. In: Front Plant Sci. vol. 8; 2017: 173.

21. He C, Zhang J, Liu X, Zeng S, Wu K, Yu Z, Wang X, Silva JATd, Lin Z, Duan J. Identification of genes involved in biosynthesis of mannan polysaccharides in Dendrobium officinale by RNA-seq analysis. Plant Mol Biol. 2015;3(88):219.

22. Shen C, Guo H, Chen H, Shi Y, Meng Y, Lu J, Feng S, Wang H. Identification and analysis of genes associated with the synthesis of bioactive constituents in Dendrobium officinale using RNA-Seq. Sci Rep. 2017;7(1):187.

23. Rischer H, Oresic M, Seppanen-Laakso T, Katajamaa M, Lammertyn F, Ardiles-Diaz W, Van Montagu MCE, Inze D, Oksman-Caldentey K-M, Goossens A: Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in Catharanthus roseus cells. Proceedings of the National Academy of Sciences 2006, 103(14):5614–5619.

24. Zhu XX, Zeng XY, Sun C, Chen SL. Biosynthetic pathway of terpenoid indole alkaloids in Catharanthus roseus. Front Med. 2014;8(3):285–93.

25. Guo X, Li Y, Li CF, Luo H, Wang M, Qian LZ, Luo J, Xiang X, Song L, Sun JY. C et al: Analysis of the Dendrobium officinale transcriptome reveals putative alkaloid biosynthetic genes and genetic markers. Gene. 2013;527(1):131–8.

26. Li L, Cheng H, Gai J, Yu D. Genome-wide identification and characterization of putative cytochrome P450 genes in the model legume Medicago truncatula. Planta. 2007;226(1):109–23.

27. Chen Y, Wang Y, Lyu P, Chen L, Shen C, Sun C. Comparative transcriptomic analysis reveal the regulation mechanism underlying MeJA-induced accumulation of alkaloids in Dendrobium officinale. J Plant Res. 2019;132(3):419–29.

28. Zhang L, Sun X, Wilson IW, Shao F, Qiu D. Identification of the Genes Involved in Anthocyanin Biosynthesis and Accumulation in Taxus chinensis. Genes 2019, 10(12).

29. Yonekura-Sakakibara K, Higashi Y, Nakabayashi R. The Origin and Evolution of Plant Flavonoid Metabolism. Front Plant Sci 2019, 10.

30. Zeng W, Jin L, Zhang F, Zhang C, Liang W. Naringenin as a Potential Immunomodulator in Therapeutics. Pharmacol Res 2018:S1043661818311186-

31. Alam MA, Subhan N, Rahman MM, Uddin SJ, Reza HM, Sarker SD. Effect of citrus flavonoids, naringin and naringenin, on metabolic syndrome and their mechanisms of action. Adv Nutr. 2014;5(4):404–17.
32. Hernández-Aquino E, Muriel P. Beneficial effects of naringenin in liver diseases: Molecular mechanisms. World J Gastroenterol. 2018;24(16):4–32.

33. Zang Y, Zha J, Wu X, Zheng Z, Ouyang J, Koffas MAG. In Vitro Naringenin Biosynthesis from p-Coumaric Acid Using Recombinant Enzymes. J Agric Food Chem. 2019;67(49):13430–6.

34. Yuan Y, Zhang J, Liu X, Meng M, Wang J, Lin J. Tissue-specific transcriptome for Dendrobium officinale reveals genes involved in flavonoid biosynthesis. Genomics. 2019;112(2):1781–94.

35. Zha XQ, Luo JP, Jiang ST, Wang JH. Enhancement of polysaccharide production in suspension cultures of protocorm-like bodies from Dendrobium huoshanense by optimization of medium compositions and feeding of sucrose. Process Biochem. 2007;42(3):344–51.

36. Yang F, Wei N-N, Gao R, Piao X-C, Lian M-L. Effect of several medium factors on polysaccharide and alkaloid accumulation in protocorm-like bodies of Dendrobium candidum during bioreactor culture. Acta Physiologiae Plantarum 2015, 37(5).

37. Zhou G. Determination of naringenin in Dendrobium officinale by HPLC. China Journal of Chinese Materia Medica 2013.

38. Lao J, Oikawa A, Bromley JR, McInerney P, Suttangkakul A, Smith-Moritz AM, Plahar H, Chiu T-Y, González Fernández-Niño SM, Ebert B, et al: The plant glycosyltransferase clone collection for functional genomics. The Plant Journal 2014, 79(3):517–529.

39. Gantet P, Memelink J. Transcription factors: tools to engineer the production of pharmacologically active plant metabolites. Trends Pharmacol Sci 2002.

40. Yan Y, Li X, Wan M, Chen J, Li S, Cao M, Zhang D. Effect of extraction methods on property and bioactivity of water-soluble polysaccharides from Amomum villosum. Carbohydr Polym. 2015;117:632–5.

41. He L, Xiaoteng Y, Liang J, Li S, Hongrui X. Comparison of different extraction methods for polysaccharides from Dendrobium officinale stem. Carbohydrate Polymers Scientific & Technological Aspects of Industrially Important Polysaccharides 2018.

42. Menke FLH, Champion A, Kijne JW, Memelink J. A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. The EMBO Journal. 1999;18(16):4455–63.

43. Fits LVD, Zhang H, Menke FLH, Deneka M, Memelink J. A Catharanthus roseus BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene Str and is induced by elicitor via a JA-independent signal transduction pathway. Plant Mol Biol. 2000;44(5):675–85.

44. Rizvi NF, Weaver JD, Cram EJ, Lee-Parsons CWT, Sangeeta D. Silencing the Transcriptional Repressor, ZCT1, Illustrates the Tight Regulation of Terpenoid Indole Alkaloid Biosynthesis in Catharanthus roseus Hairy Roots. PLoS One. 2016;11(7):e0159712-.

45. Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L. The Transcription Factor CrWRKY1 Positively Regulates the Terpenoid Indole Alkaloid Biosynthesis in Catharanthus roseus. Plant Physiol. 2011;157(4):2081–93.
46. Van Moerkercke A, Steensma P, Schweizer F, Pollier J, Gariboldi I, Payne R, Vanden Bossche R, Miettinen K, Espoz J, Purnama PC et al. *The bHLH transcription factor BIS1 controls the iridoid branch of the monoterpenoid indole alkaloid pathway in* *Catharanthus roseus*. *Proceedings of the National Academy of Sciences* 2015, **112**(26):8130–8135.

47. Van Der Fits L, Memelink J. The jasmonate-inducible AP2/ERF-domain transcription factor ORCA3 activates gene expression via interaction with a jasmonate-responsive promoter element. *Plant J.* 2001;25(1):43–53.

48. Van Moerkercke A, Steensma P, Gariboldi I, Espoz J, Purnama PC, Schweizer F, Miettinen K, Vanden Bossche R, De Clercq R, Memelink J, et al. The basic helix-loop-helix transcription factor BIS2 is essential for monoterpenoid indole alkaloid production in the medicinal plant Catharanthus roseus. *Plant J.* 2016;88(1):3–12.

49. Zhang H, Hédhili S, Montiel G, Zhang Y, Chatel G, Pré M, Gantet P, Memelink J. The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in Catharanthus roseus. *Plant J.* 2011;67(1):61–71.

50. van der Fits L, Memelink J. ORCA3, a Jasmonate-Responsive Transcriptional Regulator of Plant Primary and Secondary Metabolism. *Science.* 2000;289(5477):295–7.

51. Wei M, Jiang S-T, Luo J-P. Enhancement of growth and polysaccharide production in suspension cultures of protocorm-like bodies from Dendrobium huoshanense by the addition of putrescine. *Biotechnol Lett.* 2007;29(3):495–9.

52. Straus J, Epp RR. Response of Cupressus funebris Tissue Cultures to Gibberellins. *Science.* 1960;131(3416):1806–7.

53. Steward FC, Caplin SM. A Tissue Culture from Potato Tuber: The Synergistic Action of 2,4-D and of Coconut Milk. *Science.* 1951;113(2940):518–20.

54. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012;28(23):3150–2.

55. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Blast (basic local alignment search tool). In: Encyclopedia of Genetics, Genomics, Proteomics and Informatics. Dordrecht: Springer Netherlands; 2008. pp. 221–1.

56. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics.* 2005;21(18):3674–6.

57. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics.* 2014;30(9):1236–40.

58. 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.