Medicinal Plant Biology

ARTICLE

Open Access

https://doi.org/10.48130/mpb-0025-0022

Medicinal Plant Biology **2025**, 4: e024

Selection and functional identification of berberine biosynthetic genes in *Berberis amurensis*

Syed Basit Ali Shah[1#], Lingzhe Kong[1#], Yuzhuo Dong[1], Yaning Fu[2], Ya Tian[1], Yupeng Du[1], Zhichao Xu[1], Ge Bai[2,3*] and Zhoujie An[1*]

[1] *Key Laboratory of Saline-alkali Vegetation Ecology Restoration (Northeast Forestry University), Ministry of Education, College of Life Science, Northeast Forestry University, Harbin 150040, China*

[2] *Beijing Life Science Academy, Beijing 102209, China*

[3] *Tobacco Breeding and Biotechnology Research Center, Yunnan Academy of Tobacco Agricultural Sciences, Kunming 650021, China*

[#] Authors contributed equally: Syed Basit Ali Shah, Lingzhe Kong

[*] Corresponding authors, E-mail: 30431@163.com; zhoujie_an@163.com

Abstract

Berberine, a pharmacologically important benzylisoquinoline alkaloid (BIA) exhibiting antimicrobial and anti-inflammatory properties, accumulates abundantly in *Berberis amurensis*. However, there is a lack of reports concerning the berberine biosynthesis pathway in *B. amurensis*. In this study, berberine biosynthetic genes were systematically identified in *B. amurensis*, and key functional enzymes (BaNCS4, Ba6OMT4, Ba4'OMT2, and BaSOMT1) were characterized through transcriptomic and biochemical analyses. Notably, Ba4'OMT2 was found to exhibit superior catalytic efficiency compared to its ortholog Cc4'OMT, highlighting its potential as a valuable molecular tool for heterologous biosynthesis platforms. These findings not only elucidate the key enzymes of BIA skeleton formation and berberine *O*-methylation modification, as well as their evolutionary conservation but also offer potential molecular elements for achieving the biosynthesis of berberine through synthetic biology approaches.

Citation: Shah SBA, Kong L, Dong Y, Fu Y, Tian Y, et al. 2025. Selection and functional identification of berberine biosynthetic genes in *Berberis amurensis*. *Medicinal Plant Biology* 4: e024 https://doi.org/10.48130/mpb-0025-0022

Introduction

Huanglumu (derived from *Berberis amurensis*), a traditional Chinese herb with a millennium-long medicinal history, is used to treat gastrointestinal inflammation (enteritis, dysentery), cardiovascular disorders (hypertension), and hepatic conditions (jaundice)[1,2]. The medicinal value of *B. amurensis* primarily stems from its production of benzylisoquinoline alkaloids (BIAs)[2–4]. BIAs constitute a structurally diverse class of nitrogen-containing plant metabolites, comprising over 2,500 identified variants. These compounds are phylogenetically widespread in Ranunculales, especially in the Berberidaceae, Ranunculaceae, Papaveraceae, and Menispermaceae families[5–7]. Nine bioactive BIAs have been isolated from *B. amurensis*, of which berberine is common in *Berberis* species[8,9]. Recent pharmacological research has indicated that berberine, a main active ingredient in *B. amurensis* (with a fresh weight content of 2.6%)[3,10,11], has a series of important functions such as antibacterial, anti-inflammatory, antioxidant, hypoglycemic, lipid-lowering, neuroprotective, and antitumor activities[12–25]. These multifaceted bioactivities underscore the significance of berberine in traditional and modern pharmacopeias.

Extensive research has elucidated the biosynthetic pathways of structurally diverse BIAs in various plant species, including berberine in *Coptis chinensis* (Ranunculaceae) as well as sanguinarine, papaverine, and morphine in *Papaver somniferum* (Papaveraceae), with the key catalytic enzymes responsible for their structural diversification being well-characterized[26–31]. Despite their significant structural diversity, the biosynthesis of the benzylisoquinoline alkaloid (BIA) core skeleton initiates from L-tyrosine[6]. L-tyrosine undergoes decarboxylation by tyrosine decarboxylase (TYDC) and hydroxylation by tyramine 3-hydroxylase (3OHase) to form dopamine, a crucial precursor for the tetrahydroisoquinoline moiety[32]. Concurrently, tyrosine aminotransferase (TyrAT) and 4-hydroxyphenylpyruvate decarboxylase (HPPDC) convert L-tyrosine to 4-hydroxyphenylacetaldehyde (4-HPAA), which provides the benzyl moiety[32]. Through Pictet–Spengler condensation, dopamine, and 4-HPAA form the tetrahydroisoquinoline and benzyl moieties of (*S*)-norcoclaurine, respectively[33,34]. These two key intermediates then undergo Pictet-Spengler condensation, catalyzed by pathogenesis-related (PR)10/Bet v1-type (*S*)-norcoclaurine synthase (NCS), to form (*S*)-norcoclaurine - the universal precursor of all BIAs[35]. Notably, this NCS-mediated condensation represents the first committed and rate-limiting step in BIA biosynthesis, exerting strict control over the metabolic flux toward downstream derivatives. Based on the basic skeleton formed by this enzymatic reaction, structural diversification of BIAs is achieved through subsequent enzymatic modifications, including methylation and oxidation mediated by methyltransferases and oxidases, respectively[36,37]. (*S*)-norcoclaurine is converted to the pivotal intermediate (*S*)-reticuline through four enzymatic steps: 6-*O*-methylation, *N*-methylation, 3'-hydoxylation, and 4'-*O*-methylation. These steps are catalyzed by (*S*)-norcoclaurine 6-*O*-methyltransferase (6OMT), (*S*)-coclaurine *N*-methyltransferase (CNMT), (*S*)-*N*-methylcoclaurine 3'-hydroxylase (NMCH, CYP80B), and (*S*)-3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase (4'OMT), respectively[38–42]. Notably, *O*-methyltransferases (OMTs) are crucial in producing structural diversity and enhancing the biological properties of BIAs. For instance, OMTs in *Coptis* species catalyze position-specific methylations that are critical for protoberberine alkaloid formation[37]. Genome-guided mining in *Corydalis tomentella* identified novel methyltransferases involved in cavidine biosynthesis[43]. The catalytic versatility and substrate

promiscuity exhibited by *O*-methyltransferases from diverse medicinal plants—including CtOMTs (*Coptis teeta*), CyOMTs (*Corydalis yanhusuo*), PsOMTs (*Papaver somniferum*), and StOMT2 (*Stephania tetrandra*)—significantly contribute to the structural diversification of benzylisoquinoline alkaloids (BIAs). The enzymatic flexibility enables the rational exploitation of their known modification capabilities, facilitating the directed biosynthesis of novel BIA derivatives with potential pharmacological activities[27,44–48].

Recent advances in omics have accelerated the discovery of active BIA pathway genes in non-model species[49–51]. Despite the notable accumulation of berberine in *B. amurensis*, the molecular mechanisms underlying its biosynthesis remain poorly characterized due to limited genetic information, hindering the identification of key pathway enzymes. In this study, transcriptome sequencing, assembly, and functional annotation were performed on tissues derived from the roots, stems, and leaves of *B. amurensis*. Through systematic analysis, five key biosynthetic genes were cloned and characterized: One NCS, one 6OMT, two 4'OMT, and one SOMT gene. Functional characterization revealed that: (1) BaNCS4 facilitates the condensation reaction between dopamine and 4-HPAA, resulting in the formation of (S)-norcoclaurine; (2) Ba6OMT4 subsequently converts (*S*)-norcoclaurine to (*S*)-coclaurine; (3) Ba4'OMT2 converts (*S*)-3'-hydroxy-*N*-methylcoclaurine to (*S*)-reticuline; and (4) BaSOMT1 successively methylates (*S*)-scoulerine at the C9 and C2 position, generating (*S*)-tetrahydrocolumbamine and (*S*)-tetrahydropalmatine as products. These findings provide molecular foundation for future metabolic engineering efforts aimed at optimizing benzylisoquinoline alkaloid production, with particular relevance to berberine biosynthesis.

Materials and methods

Plant materials and chemicals

Plant materials of *B. amurensis* were collected in July 2022 from their natural habitat in Yichun City, Jiangxi Province, China. Species identification was confirmed through DNA barcoding analysis. Fresh tissues, including leaves (L), stems (S), and roots (R), were immediately frozen in liquid nitrogen and stored at -80°C until RNA extraction. Authentic standards—(*S*)-norcoclaurine, *N*-methylcoclaurine, (*S*)-reticuline, and (*S*)-scoulerine—were obtained from BioBioPha Co., Ltd. (Kunming, Yunnan, China). While (*S*)-tetrahydrocolumbamine, (*S*)-tetrahydropalmatine and (*S*)-coclaurine were obtained from Shanghai Yuanye Bio-Technology Co., Ltd. (Shanghai, China). The suppliers of 4-HPAA and 3'-hydroxy-*N*-methylcoclaurine were Toronto Research Chemicals Inc. (Toronto, Canada), and S-adenosyl-L-methionine (SAM) was procured from Sigma-Aldrich (St. Louis, MO, USA). These commercial compounds had a purity level of at least 98%. Chromatographically pure reagents were employed in the compound detection tests.

Transcriptome sequencing and de novo assembly

In *Berberis amurensis*, transcriptome sequencing was performed on the roots, leaves, and stems using Illumina HiSeq 2000 sequencing technology. Frozen leaf, root, and stem samples were used to extract total cellular RNAs (Supplementary Method S1). The Illumina TruSeq RNA Sample Prep Kit v3-cBot-HS was utilized to create a cDNA library, and the Illumina HiSeq platform form Novogene (Beijing, China) was utilized for paired-end sequencing (Supplementary Method S2). Following RNA-seq, adaptors, low-quality nucleotides, and unknown nucleotides were eliminated from the raw reads to create clean reads. Raw sequencing reads were quality-filtered using Trimmomatic (v0.39)[52] with the following parameters: removal of bases with Phred quality scores < 20, exclusion of reads shorter than 40 bp after trimming, elimination of reads containing terminal 'N' nucleotides, and adapter sequence clipping using the ILLUMINACLIP algorithm with a 2 bp seed mismatch allowance. FastQC software[53] was used for quality control both before and after cutting. The transcriptome sequences were constructed using the de novo RNA-Seq Assembly Pipeline (https://github.com/abims-sbr/drap) with Trinity (v2.13.2)[54], employing the default settings. The CD-HIT-EST program[55] eliminated duplicate contigs. TransDecoder (v3.0.1) was finally used to predict the coding region and the read counts. The completeness of the *B.amurensis* transcriptome assembly was evaluated using the Embryophyta_odb10 dataset, which contains 1,614 evolutionarily conserved embryophyte genes, through Benchmarking Universal Single-Copy Orthologs (BUSCO) tool (v5.7.1)[56].

Functional annotation, CDS prediction, and phylogenetic analysis

Transcriptomic data from *B. amurensis* leaves, stems, and roots were compared against publicly available databases, including the nonredundant protein (NR), SwissProt, Pfam, the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and Gene Ontology (GO)[57,58]. After GO annotations were obtained, GO functional classification was performed using WEGO software[59]. The transcriptome sequences were blasted with an E-value cutoff of 1×10^{-5} to identify the sequences encoding the berberine biosynthetic enzymes. ClustalW[60] was employed for aligning the protein sequences, followed by phylogenetic analysis. Using MEGA X (v10.2.4)[61], a maximum likelihood phylogenetic tree was constructed to identify NCSs and OMTs from *B. amurensis*.

Identification of differentially expressed genes

For transcriptomic analysis, gene expression levels were quantified as FPKM (fragments per kilobase of exon per million fragments mapped) values using R packages[62,63]. The resulting expression matrix was visualized as a heatmap using OmicStudio tools (www.omicstudio.cn). To ensure data stability and comparability, normalization was performed by adding a pseudo-count of 1 to each FPKM value, followed by a $\log_{10}$ transformation. Differential expression analysis was conducted using edgeR (v4.6.1) with p value ≤ 0.05 and $|\log_2 FC| > 1$. Functional annotation of DEGs was performed through GO and KEGG pathway enrichment analysis, with statistical significance set at $p \leq 0.05$.

Recombinant protein expression and purification

Utilizing an RNAprep pure plant kit from TIANGEN (Beijing, China), total RNA was extracted from the leaf, stem, and root tissues of *B. amurensis*. The TransScripte One-Step gDNA Removal and cDNA Synthesis SuperMix from TransGen Biotech (Beijing, China) was used to synthesize first-strand cDNA.

Total RNA was isolated from *B. amurensis* leaf, stem, and root tissues using the RNAprep Pure Plant Kit (TIANGEN, Beijing, China). First-strand cDNA was synthesized from 1 μg total RNA using TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, Beijing, China) according to the manufacturer's protocol. Full-length open reading frames (ORFs) of putative *BaNCS* and *BaOMT* genes were amplified from cDNA templates using 2× Phanta Flash Master Mix (Vazyme, Nanjing, China) with gene-specific primers containing overlapping sequences for subsequent cloning (Supplementary Table S1). PCR products were purified and cloned into the pET-28a(+) prokaryotic expression vector using E.Z.N.A.® Gel Extraction Kit (Omega) and the Basic Seamless Cloning and Assembly Kit (TransGen Biotech). Recombinant plasmids were transformed into *E. coli* DH5α competent cells for propagation and subsequently purified using the TIANprep Mini Plasmid Kit

(TIANGEN). The verified recombinant plasmids were transformed into *E. coli* BL21(DE3) for heterologous protein expression. Cultures were grown to $OD_{600} \approx 0.6$ at 37 °C, then induced with 0.5 mM IPTG at 15 °C for 20 h with shaking (140 rpm). The cells were harvested by centrifugation (8,000 × g, 10 min, 4 °C) and resuspended in lysis buffer (100 mM Tris-HCl, pH 7.5) supplemented with 1 mM PMSF. The cell disruption was achieved by sonication (10 min, 5 s pulse/5 s rest) on ice, followed by centrifugation (12,000 × g, 20 min, 4 °C) to remove cell debris. The supernatant containing soluble proteins was incubated with Ni-NTA resin (CWBIO, Beijing, China) for 1 h at 4 °C (Supplementary Method S3). After washing with 20 mM imidazole in the binding buffer, target proteins were eluted with 250 mM imidazole, as previously optimized[48]. Target protein purity was assessed by 12% SDS-PAGE with Coomassie Brilliant Blue staining, and concentrations were determined by A_{280} measurement using a Spark multimode microplate reader (Tecan, Switzerland).

In vitro enzyme assays and functional characterization

The enzyme assay was conducted on the activity of the genes encoding NCS, 6OMT, 4'OMT, and SOMT using a reaction mixture containing 100 μl of 100 mM Tris-HCl (pH 7.4), 25 mM sodium ascorbate, 10% (v/v) glycerol, 1 mM β-mercaptoethanol, 2 mM SAM, 0.5 mM potential substrate, and crude enzyme (or 10 μg of purified recombinant enzyme). The CcNCS, Cc6OMT, Cc4'OMT, and CcSMT were used as the positive control, whereas the empty vector was used as the negative control. Enzymatic assays were performed at 37 °C and 200 rpm for 2 h. The reaction was terminated by adding 50% (v/v) methanol, followed by centrifugation at 12,000 rpm for 20 min. The resulting supernatant was collected, spin-dried, and reconstituted in 150 μL of methanol for Liquid chromatography with high performance (HPLC) and Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) analysis.

Identification of enzymatic products using HPLC and LC-MS/MS

Chromatographic separation was performed on a Shimadzu LC-2050C 3D system (Shenyang, China) using an Ultimate® UHPLC XB-C18 column (2.1 mm × 100 mm, 1.8 μm) maintained at 40 °C. The mobile phase consisted of 0.1% (v/v) formic acid in water (A) and acetonitrile (B), with the following gradient program at 0.5 mL/min: 5% B (0–1 min), linear increase to 30% B (1–3 min), 60% B (3–15 min), 100% B (15–20 min), returning to 5% B (20–25 min), and held for 5 min (25–30 min). Detection was performed at 282 nm with a 10 μL injection volume. For LC-MS/MS analysis, the Shimadzu LCMS-9030 Q-TOF system employed identical mobile phases with an optimized gradient: 15% B (0–1 min), linear increase to 45% B (1–8 min), 80% B (8–10 min), returning to 15% B (10–11 min) and held for 4 min (11–15 min) at 0.2 mL/min (column temperature, 35 °C). Mass spectrometric analysis was performed in data-dependent acquisition (DDA) mode with the following parameters: full MS scans (*m/z* 100–800), followed by top-10 MS/MS scans. MS detection used positive electrospray ionization (ESI+) with nitrogen nebulizing gas (3.0 L/min) and drying gas (10 L/min at 300 °C), scanning *m/z* 100–1,000. Each sample was injected at a volume of 1 μL.

Results

De novo assembly and functional annotation

To systematically characterize the *B. amurensis* transcriptome, three strand-specific cDNA libraries (BaL, BaS, and BaR) were constructed from leaf, stem, and root tissues, respectively. Following quality control (removal of adapter sequences and low-quality reads), 21,665,463 (3.2 Gb), 25,527,459 (3.8 Gb), and 18,494,454 (2.8 Gb) clean reads were obtained for BaL, BaS, and BaR libraries, respectively (Supplementary Table S2). The GC content of clean reads of BaL, BaS, and BaR were 45%, 44%, and 44%, respectively. *De novo* assembly using a standardized short-read pipeline generated 55,016 coding genes with an N50 of 1,170 bp. Benchmarking against the Embryophyta_odb10 dataset revealed exceptional assembly completeness, with 90.8% of the 1,614 conserved single-copy orthologs being fully represented (Table 1, Supplementary Table S3). This high-quality transcriptome provides a robust foundation for downstream gene discovery and expression analyses.

The coding genes were annotated by searching the sequences against the following five public protein databases: NR, Swiss-Prot, GO, KEGG, and Pfam. In *B. amurensis*, these databases contained 53,926 genes (98.02%) that were annotated overall. In addition, 48,464 (88.09%), 35,005 (63.63%), 14,735 (27.32%), 9,376 (17.04%), and 25,133 (57.76%) genes were annotated that were annotated in the NR, Swiss-Prot, GO, KEGG, and Pfam databases, respectively (Table 2, Supplementary Data S1). A total of 14,735 genes from *B. amurensis* were categorized into three classes in the GO analyses (Supplementary Fig. S1). Using the KEGG database for annotation, 265 biological pathways were ascribed to 10,811 gene from *B. amurensis*, with 3,443 genes representing the largest annotated set for metabolic pathways. This included 1,523 genes involved in the biosynthesis of secondary metabolites (ko01110), including BIA biosynthesis (Supplementary Fig. S2).

Identification of differentially expressed genes

The differential expression analysis (p value ≥ 0.05 and $|\log_2 FC| \geq 1$) identified 23,794 DEGs across three tissue comparisons (BaL, BaS, BaR). The BaR vs BaL comparison showed the highest number of DEGs (18,095), with 10,564 up-regulated and 7,531 down-regulated genes. Similarly, BaS vs BaL revealed 14,837 DEGs (9,714 up/5,123 down), while BaR vs BaS had 12,315 DEGs (4,761 up/7,554 down) (Supplementary Table S4). Further analysis of the up-regulated DEGs exhibited significant differences in the number of tissue-specific DEGs across comparisons, with BaS vs BaL showing the highest number of tissue-specific DEGs (4,057), followed by BaR vs BaL (2,039) and BaR vs BaS (836). Intersection analysis identified substantial overlap between comparisons, with 4,600 DEGs co-upregulated

Table 1. Overview of de novo assembly statistics.

Item	Transcriptome assembly	Unigenes (the longest isoform)	Coding genes
Total sequences	325,523	262,191	55,016
Complete BUSCOs	92.40%	92.00%	90.80%
Total bases	229,930,964	111,153,438	48,237,144
Min sequence length (bp)	178	201	255
Max sequence length (bp)	18,516	15,698	12,783
Average sequence length (bp)	706.34	423.94	876.78
Median sequence length (bp)	385	294	618
N50 length (bp)	1,186	435	1,170
GC content	41.10%	40.02%	43.83%

Table 2. Statistics on the number of gene annotated for function by KEGG, NR, GO, PFAM, and Swissprot databases.

Database	Number of genes	Percentage
KEGG	9,376	17%
NR	48,464	88%
GO	14,735	34%
Pfam	25,133	58%
Swiss-Prot	35,005	64%
Total	53,926	98%

in both BaR vs BaL and BaS vs BaL contrasts, and 2,868 genes shared between BaL-BaR and BaS-BaR (Supplementary Fig. S3a). These results facilitate the identification of candidate genes potentially involved in berberine biosynthesis.

To further characterize the DEGs, KEGG and GO enrichment analyses were performed. KEGG analysis identified 7,336 (BaR vs BaL), 5,856 (BaS vs BaL), and 5,929 (BaR vs BaS) DEGs, with 'biosynthesis of secondary metabolites' being the most significantly enriched pathway across all comparisons. Notably, the isoquinoline alkaloid biosynthesis (ko00950) pathway was consistently enriched and ranked among the top 20 pathways in all three comparisons. Specifically, BaR vs BaL showed 91 DEGs (75 up-regulated/16 down-regulated) in ko00950 pathway, along with 61 DEGs (32 up-regulated/29 down-regulated) in the tyrosine metabolism (ko00350) pathway and 44 DEGs (19 up-regulated/33 down-regulated) in nitrogen metabolism (ko00910) pathway (Fig. 1a). Similar patterns were observed in BaS vs BaL (79 DEGs: 71 up-regulated/8 down-regulated in ko00950 pathway; 51 DEGs: 34 up-regulated/17 down-regulated in ko00350 pathway; 38 DEGs: 10 up-regulated/28 down-regulated in ko00910 pathway) and BaR vs BaS (69 DEGs: 39 up-regulated/30 down-regulated in ko00950 pathway; 40 DEGs: 18 up-regulated/32 down-regulated in ko00350 pathway; 33 DEGs: 14 up-regulated/19 down-regulated in ko00910 pathway), all potentially related to berberine biosynthesis (Fig. 1b–c). Functional characterization through GO enrichment identified oxidoreductase activity (GO:0016491) as the most significantly represented molecular function across BaR vs BaL, BaS vs BaL, and BaR vs BaS comparisons (Supplementary Fig. S3b–S3d), highlighting the importance of redox reactions in the observed differential gene expression patterns.

Identification and expression pattern analysis of berberine biosynthetic genes

The berberine biosynthetic pathway initiates with the stereospecific condensation of dopamine and 4-HPAA, catalyzed by NCS to form the pivotal intermediate (*S*)-norcoclaurine. This tyrosine-derived scaffold then undergoes sequential modifications through a series of reactions involving multiple methyltransferases (OMTs and NMTs) and oxidases (CYP450, BBE, and STOX), ultimately yielding berberine (Fig. 2). Comprehensive transcriptome analysis revealed distinct expression profiles of 37 genes encoding nine key enzymatic components of the berberine pathway in *B. amurensis*, including four *NCSs,* six *6OMTs*, twelve *CNMTs*, two *CYP80Bs*, two *4'OMTs*, two *BBEs*, three *SOMTs*, three *CYP719A21s*, and three *STOXs* (Fig. 2). Tissue-specific expression patterns demonstrated clear partitioning of the biosynthetic active enzymes. The early pathway genes (NCS, 4'OMT, and CYP80B) showed predominant expression in stems, suggesting this organ as a major site for intermediate production (Fig. 2, Supplementary Table S5). In contrast, downstream enzymes (BBE and STOX) exhibited root-enriched expression, indicating roots as the primary site for late-stage modifications and berberine accumulation. Notably, *BaCYP719A2*, as the only leaf-specific

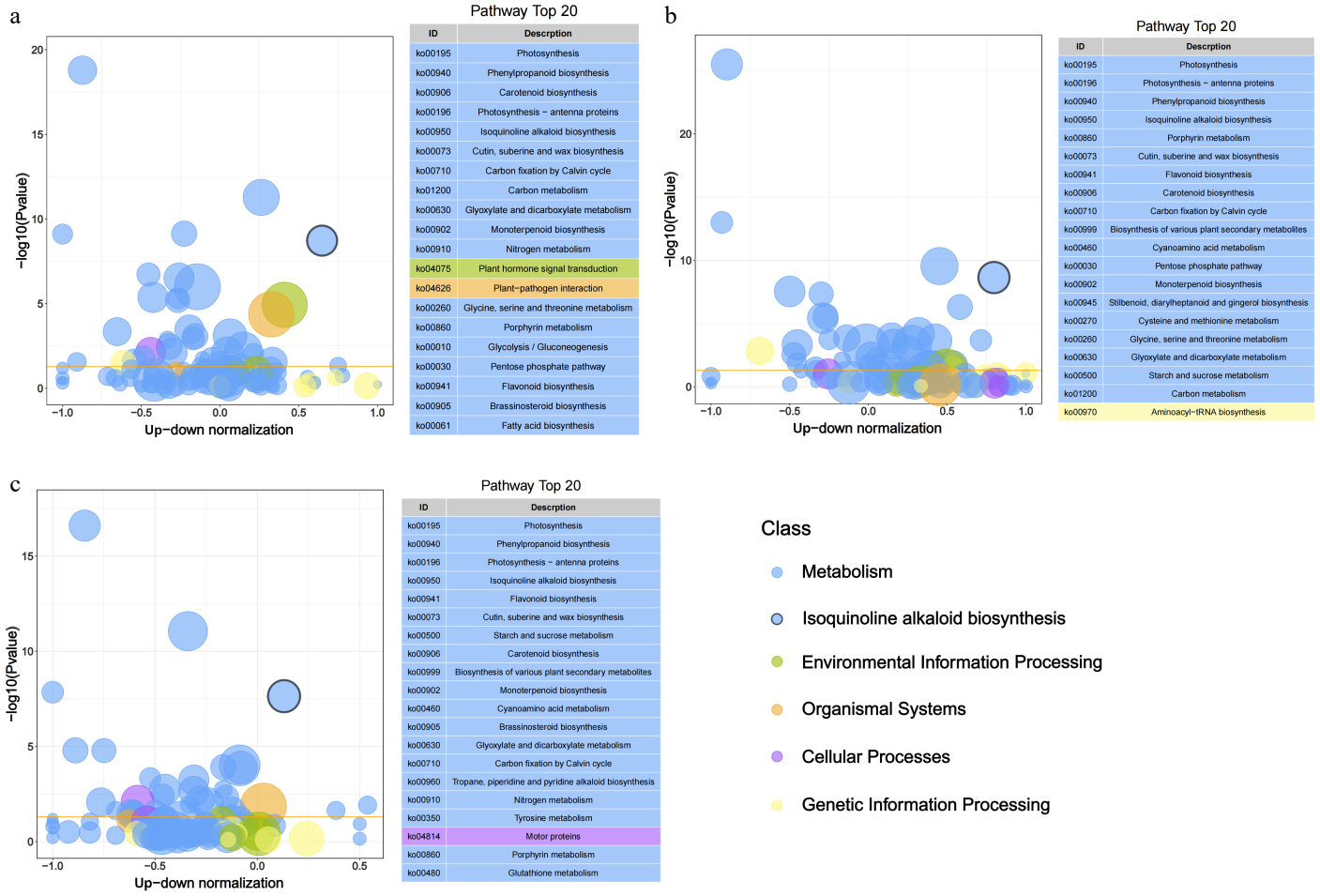

Fig. 1 Bubble plot of KEGG pathway enrichment analysis based on DEGs across tissues. The x-axis represents the up-down normalization value calculated as (Number of up-regulated genes − Number of down-regulated genes)/(Number of up-regulated genes + Number of down-regulated genes), reflecting the directional bias of pathway regulation. Bubble size indicates the number of DEGs per pathway, while color denotes different pathway classes. The yellow line represents the threshold value of *p* value = 0.05. On the right is the path list of the top 20 with *p* values.

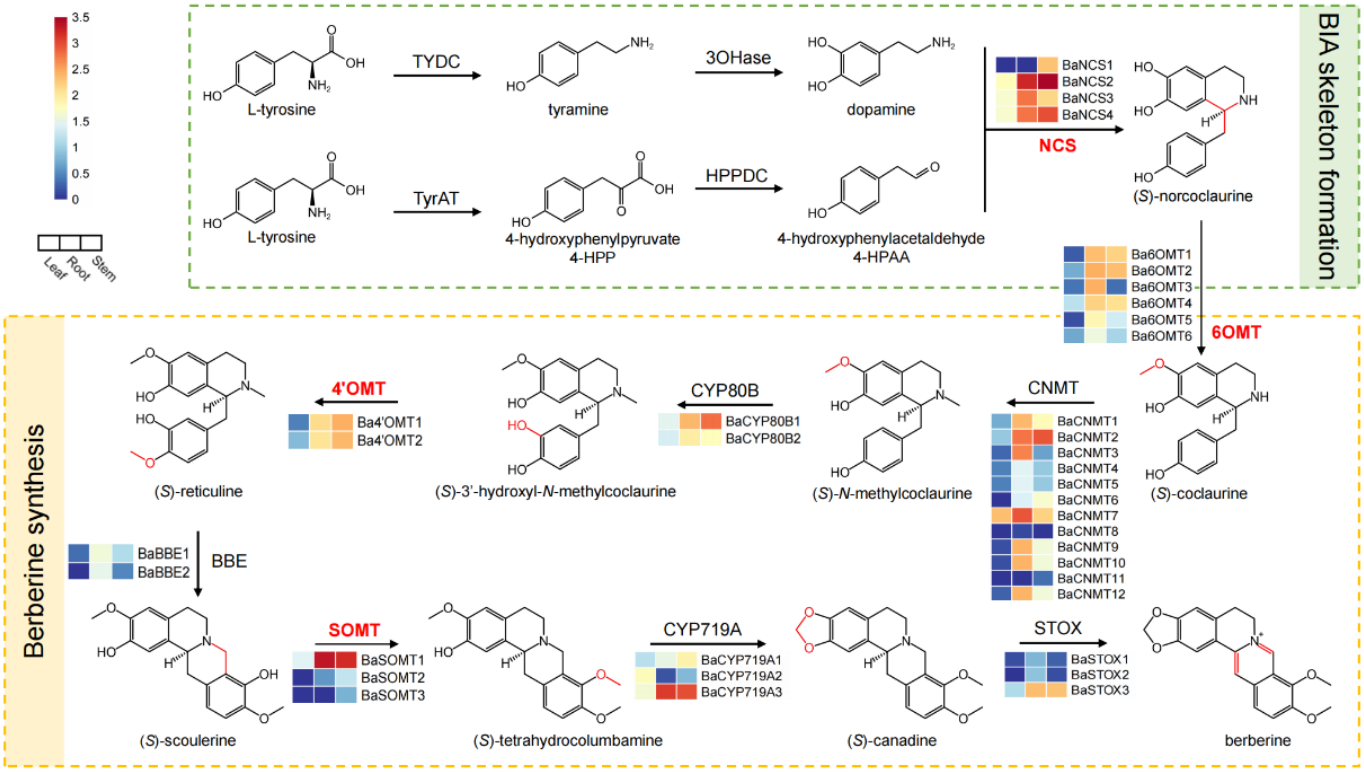

Fig. 2 Gene expression in the biosynthetic pathway of berberine in *B. amurensis*. The columns from left to right in each heat map display the expression levels of the leaf, root, and stem of *B. amurensis*. The $\log_{10}$(FPKM+1) values were used for normalization.

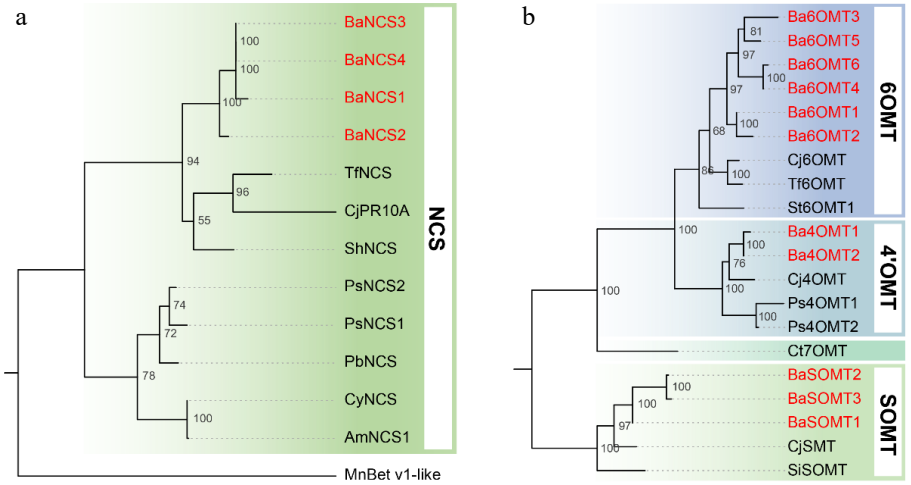

Fig. 3 Phylogenetic analysis of (a) NCS, and (b) OMT protein sequences. The maximum-likelihood tree was constructed using aligned amino acid sequences from *B. amurensis* (highlighted in red) and homologous enzymes from other plant species. Bootstrap values (1000 replicates) are shown at branch nodes. A higher value indicates that the branch is more reliable. Abbreviations and GenBank accession numbers for the sequences are listed as follows: TfNCS (ACO90248.1), CjPR10A (BAF45338.2), ShNCS (AIT42265.1), PsNCS2 (AAX56304.1), PsNCS1 (AAX56303.1), PbNCS (ACO90258.1), CyNCS (WDS97777.1), AmNCS1 (ACJ76785.1), MnBet_v1-like (EXB29146.1), Cj6OMT (BAB08004.1), Tf6OMT (AAU20765.1), St6OMT1 (6OMT from *S. tetrandra*)[48], Cj4'OMT (BAB08005.1), Ps4'OMT1 (AAP45313.1), Ps4'OMT2 (AAP45314.1), CjSMT (BAA06192.1), and SiSOMT (MK749415).

isoform, may contribute to organ-specific alkaloid profiles, given the established role of CYP719A in methylenedioxy bridge formation (Fig. 2). These transcriptional patterns strongly correlated with the tissue distribution of BIAs, where roots and stems showed the highest berberine accumulation.

Phylogenetic analysis of NCSs and OMTs involved in berberine biosynthesis

In this study, four BaNCSs and 11 BaOMTs (6 Ba6OMTs, 2 Ba4'OMTs, 3 BaSOMTs) from the transcriptome of *B. amurensis* (Fig. 3). The NCS-derived phylogenetic tree revealed a close relationship between the NCSs of Berberidaceae (four BaNCSs from *B. amurensis* and ShNCS from *Sinopodophyllum hexandrum*) and Ranunculaceae (CjPR10A from *Coptis japonica* and TfNCS form *Thalictrum flavum*), while the NCS of Papaveraceae (PsNCS1 and PsNCS2 from *P. somniferum*, PbNCS from *Papaver bracteatum*, CyNCS from *Corydalis yanhusuo*, and AmNCS from *Argemone mexicana*) is a separate branch of divergence, basically consistent with the phylogenetic relationships between species (Fig. 3a)[64]. Therefore, these

BaNCSs in this clade might be specifically related to the NCS activity in *B. amurensis*. Notably, the sequence identity among the four genes was more than 90%. The OMT-derived phylogenetic tree indicated distinct functional clades among BaOMTs, including 6OMT, 4'OMT, and SOMT. It showed significant gene expansion, particularly in the 6OMT branch (Fig. 3b). Ct7OMT (norcoclaurine 7-*O*-methyltransferase) is a separate branch involved in jatrorrhizine biosynthesis that contains an unusual 7-*O*-methylation pattern[27]; however, 7OMT was not annotated in our transcriptome dataset. These findings indicate gene expansion and evolutionary conservation of NCS and OMT within the lineage.

To elucidate the molecular basis of berberine biosynthesis in *B. amurensis*, five candidate genes (*BaNCS4*, *Ba6OMT4*, *Ba4'OMT1*, *Ba4'OMT2*, and *BaSOMT1*) were cloned and heterogenously expressed in *E. coli*. Both crude extracts and purified recombinant proteins were subsequently used for *in vitro* enzymatic assays to characterize their catalytic activities (Supplementary Fig. S4).

Functional characterization of NCS *in vitro*

In vitro enzymatic assays were conducted to characterize the catalytic activity of BaNCS4 with dopamine and 4-HPAA as substrates. Using CcNCS as the positive control, the results showed the same product peak under the catalysis of CcNCS and BaNCS4 (Supplementary Fig. S5). LC-MS analysis revealed a reaction product with mass-to-charge ratio (*m/z*) 272 $[M+H]^+$, whose retention time (8.42 min) and MS/MS fragmentation pattern matched authentic (*S*)-norcoclaurine standards (Fig. 4). These results demonstrate that BaNCS4 catalyzes the stereospecific condensation of dopamine and

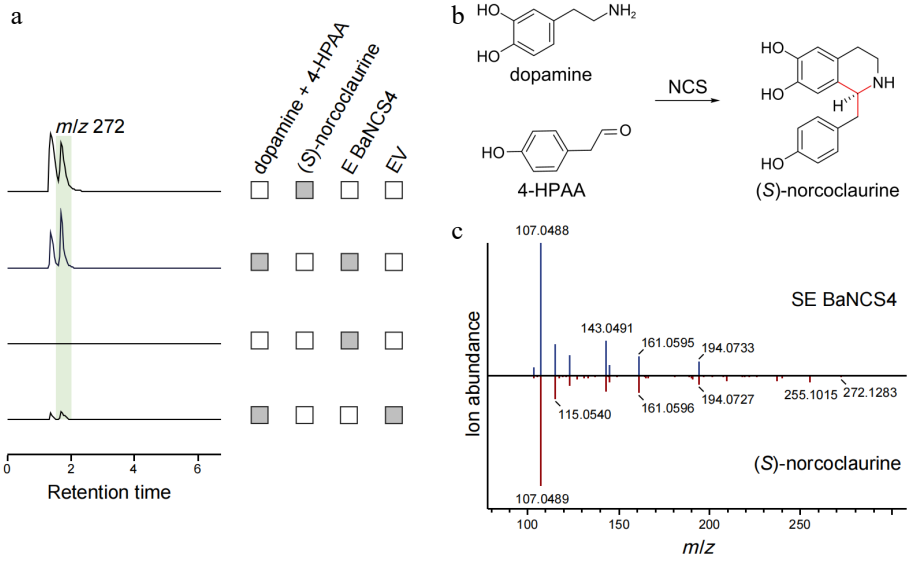

Fig. 4 Functional characterization of BaNCS related to berberine biosynthesis in *B. amurensis*. (a) The LC-MS extracted ion chromatogram at *m/z* 272 represents the catalytic product generated by BaNCS. The reaction system without substrate and pET-28a (+) empty vector (EV) was used as the negative control. (b) NCS mediates the Pictet-Spengler condensation of dopamine and 4-HPAA, yielding (*S*)-norcoclaurine. (c) Comparison of MS/MS fragments between the product from BaNCS (above) and authentic (*S*)-norcoclaurine (below).

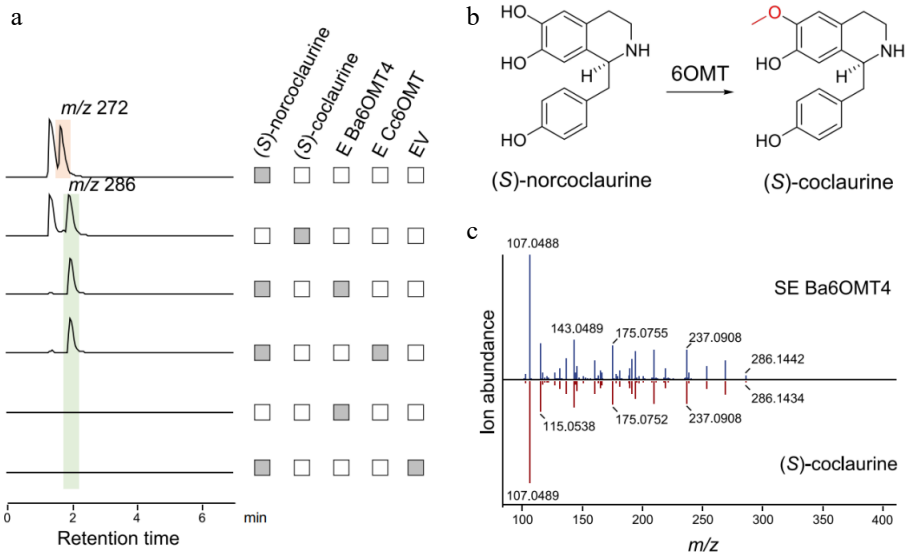

Fig. 5 Functional characterization of Ba6OMT related to berberine synthesis in *B. amurensis*. (a) The LC-MS extracted ion chromatogram at *m/z* 272 and 286 represent the substrate (*S*)-norcoclaurine and the catalytic product generated by Ba6OMT, respectively. The pET-28a (+) empty vector (EV) was used as the negative control. (b) The synthesis of (*S*)-coclaurine catalyzed by 6OMT. (c) Comparison of MS/MS fragments between the product from Ba6OMT (above) and authentic (*S*)-coclaurine standard (below).

4-HPAA to form (*S*)-norcoclaurine, the universal precursor of benzylisoquinoline alkaloids.

Functional characterization of 6OMT *in vitro*

The *in vitro O*-methyltransferase activity of the candidate 6OMT protein, the first modifying enzyme after the formation of the BIA basic skeleton, was assessed using (*S*)-norcoclaurine as substrate, with Cc6OMT from *C. chinensis* as the positive control (Supplementary Fig. S5). LC-MS analysis showed that a product peak with a *m/z* of 286 appeared at the same position under the catalysis of Ba6OMT4 and Cc6OMT purified protein (Fig. 5a). After comparing the retention time and mass spectrum information of the product with the standard, the product was identified as (*S*)-coclaurine (Fig. 5b and c). The result indicated that Ba6OMT has (*S*)-norcoclaurine 6OMT activity and could catalyze the *O*-methylation of the C6 hydroxyl group to convert (*S*)-norcoclaurine to (*S*)-coclaurine. This modification represents the first committed step in the diversification of benzylisoquinoline alkaloid scaffolds.

Functional characterization of 4'OMT *in vitro*

In vitro catalytic activity of Ba4'OMT1 and Ba4'OMT2 was assessed using (*S*)-3'-hydroxy-*N*-methylcoclaurine as substrate, with the Cc4'OMT from *C. chinensis* as the positive control, and it was found that only Ba4'OMT2 exhibited activity (Supplementary Fig. S5). The LC-MS results showed that under the catalysis of Ba4'OMT2 and Cc4'OMT, a product peak with the *m/z* of 330 appeared at the same position (Fig. 6a). After comparing the retention time and mass spectrum information of the product with the standard, the product was identified as (*S*)-reticuline (Fig. 6b & c). The result demonstrated that Ba4'OMT2 could catalyze the methylation of the C4' hydroxyl group in (*S*)-3'-hydroxy-*N*-methylcoclaurine to form (*S*)-reticuline and has higher catalytic activity than Cc4'OMT.

Functional characterization of SOMT *in vitro*

Furthermore, HPLC analysis showed that, using (*S*)-scoulerine as the substrate, BaSOMT had the same catalytic function as CcSOMT (Supplementary Fig. S5). The LC-MS results showed that both enzymes could catalyze (*S*)-scoulerine to produce product peaks with *m/z* of 342 and 356 (Fig. 7a). The two products were identified as (*S*)-tetrahydrocolumbamine and (*S*)-tetrahydropalmatine, respectively, based on comparison with the mass spectrometry information of the standard sample (Fig. 7b & c). This shows that BaSOMT can catalyze the successive methylation of the C9 and C2 hydroxyl groups to convert (*S*)-scoulerine to (*S*)-tetrahydrocolumbamine and (*S*)-tetrahydropalmatine. The identical retention times and mass spectra of the methylated derivatives confirmed conserved functional characteristics between these orthologous enzymes.

Discussion

The roots and stems of *B. amurensis* accumulates high levels of berberine—a prominent BIA with broad pharmacological activities and a key bioactive component in traditional medicine[10]. The transcriptomic analysis revealed significant differential expression of all identified genes involved in berberine biosynthesis in various tissues. The enrichment of 'oxidoreductase activity' (GO:0016491) and the prominence of the isoquinoline alkaloid biosynthesis pathway (ko00950) in DEGs support the tissue-specific pattern of BIA metabolism in *B. amurensis*. Notably, the spatial separation of early and late biosynthetic genes suggests potential inter-organ transport of pathway intermediates in *B. amurensis*, a phenomenon well-documented in other berberine-producing plants[65].

Given that the NCS enzyme catalyzes the first rate-limiting step of the berberine pathway, and methyltransferases catalyze the biochemical reactions involved in the biosynthesis and modification of bioactive BIAs in plants, the NCS and OMT functional enzymes were identified in *B. amurensis*. Based on the assembled transcriptome, the gene expression profile of 37 candidate genes was observed. In addition to *BaNCS4,* which showed high expression levels in the stem, *Ba6OMT4*, *Ba4'OMT2*, and *BaSOMT1* showed high expression in the roots, which may play a direct role in the substantial accumulation of BIAs observed in both the roots and

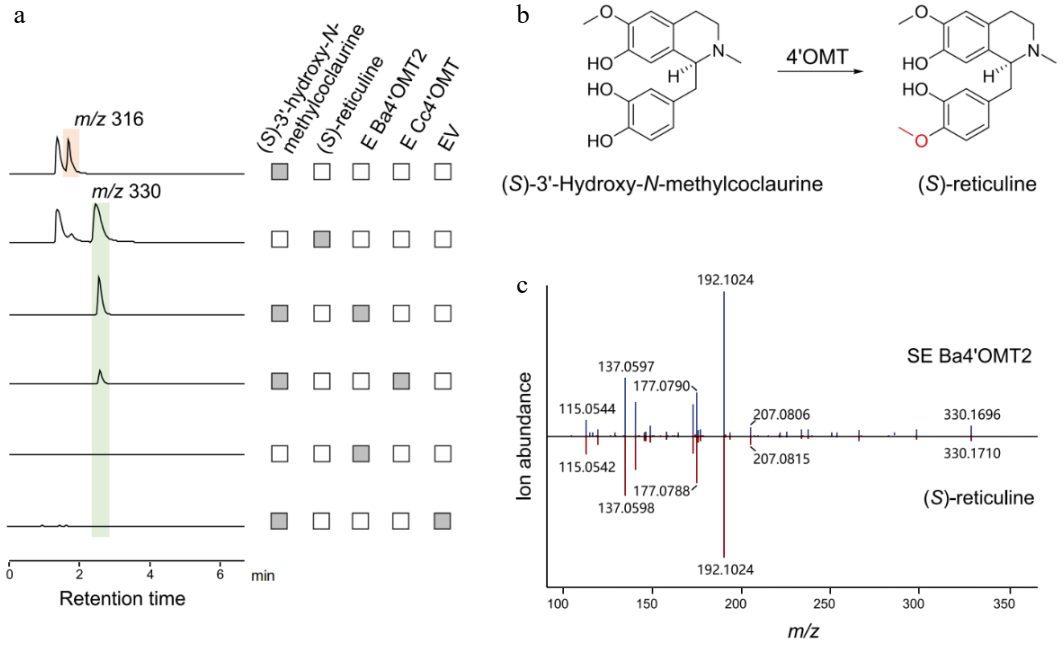

Fig. 6 Functional characterization of Ba4'OMT2 related to berberine biosynthesis in *B. amurensis*. (a) The LC-MS extracted ion chromatogram at *m/z* 316 and *m/z* 330 represent the substrate (*S*)-3'-hydroxy-*N*-methylcoclaurine and the catalytic product generated by Ba4'OMT2 enzymes, respectively. (b) The synthesis of (*S*)-reticuline catalyzed by 4'OMT. (c) Comparison of MS/MS fragments between the product from Ba4'OMT2 (above) and authentic (*S*)-reticuline (below).

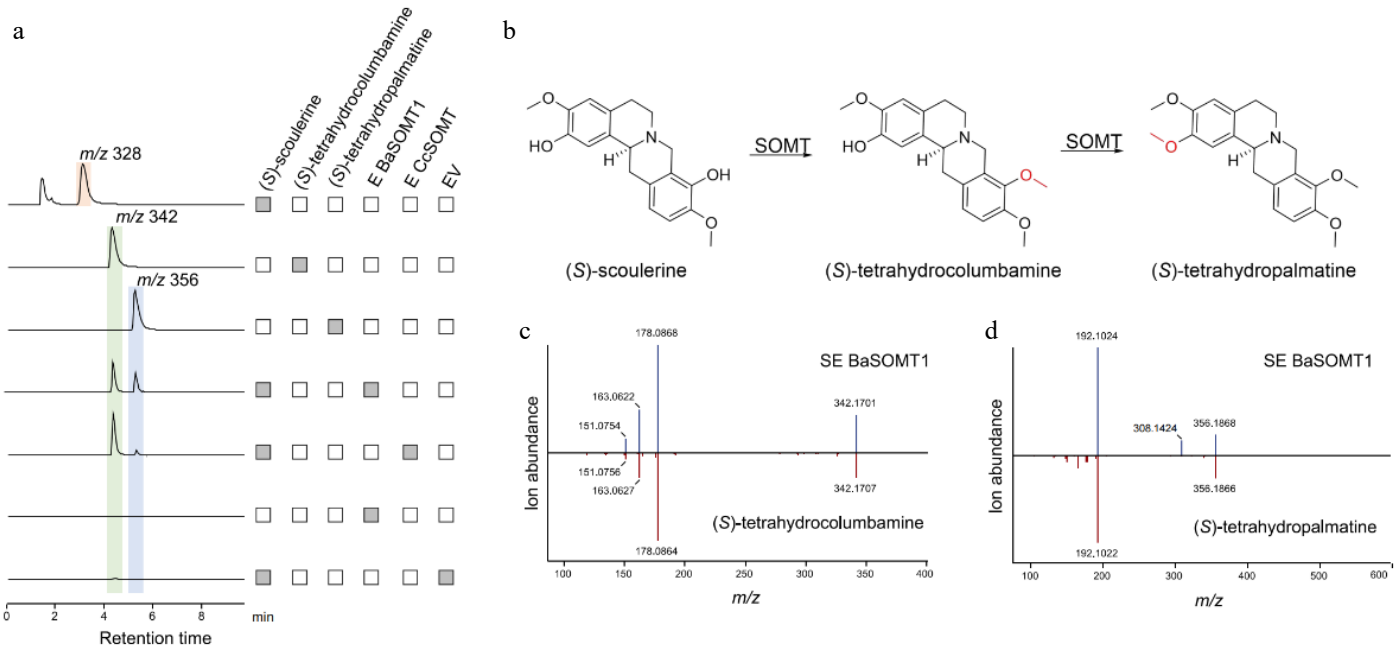

Fig. 7 Functional characterization of BaSOMT related to berberine synthesis in *B. amurensis*. (a) The LC-MS extracted ion chromatogram at *m/z* 328, 342, and 356 represent the substrate (*S*)-scoulerine and two continuous catalytic products generated by BaSOMT, respectively. (b) The synthesis of (*S*)-tetrahydrocolumbamine and (*S*)-tetrahydropamatine catalyzed by SOMT. (c), (d) Comparison of MS/MS fragments between the product from BaSOMT (above) and their corresponding standards (below).

stems of *B. amurensis*. The biosynthetic pathways for diverse arrays of BIAs with significant pharmacological activity have been established in several plants, such as noscapine and morphine in opium poppy[28,66], berberine in *C. chinensis*[26], nelumboferine in sacred lotus[67], berberine in *Phellodendron amurense*[49], etc. Among them, the upstream pathway of BIA is highly conserved in Ranunculales. In the same way, these findings confirm the evolutionary conservation of NCS and OMT in *Berberis*.

The growing demand for bioactive natural products has outpaced traditional plant extraction methods, driving increasing interest in heterologous biosynthesis approaches. To establish sustainable microbial production platforms, a critical focus must be placed on engineering high-performance enzymes capable of supporting efficient biomanufacturing processes. For instance, due to the low enzymatic conversion efficiency of (*S*)-*N*-methylcoclaurine to (*S*)-reticuline, Rao et al.[68] streamlined biosynthesis by eliminating NMCH and 4'OMT-mediated methylation steps in traditional protoberberine pathways, achieving high yields of rotundine at the gram level. In this study, *in vitro* functional verification revealed Ba4'OMT2 with significantly higher activity than Cc4'OMT, providing a potentially efficient biosynthetic element. This discovery holds significant promise for applications in synthetic biology. Integration of Ba4'OMT2 into microbial hosts (e.g., *Saccharomyces cerevisiae* or *E. coli*) may enhance titers of protoberberine alkaloids, as evidenced by recent successes with plant-derived OMTs in heterologous systems[68–70]. The superior activity of Ba4'OMT2 could increase the flow to the target alkaloid without changing the natural substrate and eliminating the methylation step.

In summary, based on transcriptome analysis, this study systematically revealed high conservation with homologs from other berberine-producing plants. Four functional enzymes were identified *in vitro*: BaNCS4 for fundamental skeleton formation, along with three methyltransferases (Ba6OMT4, Ba4'OMT2, and BaSOMT1) for subsequent modifications. In the first committed step of berberine biosynthesis in *B. amurensis*, dopamine and 4-HPAA condense to produce (*S*)-norcoclaurine, catalyzed by BaNCS4. Likewise, these OMTs are vital for the diversification of BIA. Notably, a highly catalytically active Ba4'OMT2 enzyme was discovered. These findings clarify the enzymatic framework for berberine biosynthesis in *B. amurensis* and provide functional genetic elements to address current yield limitations in microbial production of these medicinally important alkaloids. These results advance the understanding of specialized metabolism in *Berberis* species and offer practical resources for metabolic engineering applications.

Author contributions

The authors confirm their contributions to the paper as follows: study design and coordination: Xu Z, Kong L, An Z; sample preparation of plant materials : Kong L; bioinformatic analysis: Shah SBA, Kong L, Fu Y; experiments conduction and data analysis: Shah SBA, Dong Y, Tian Y; draft manuscript preparation: Bai G, An Z, Shah SBA, Dong Y, Xu Z, Du Y. All authors reviewed the results and approved the final version of the manuscript.

Data availability

Raw sequencing reads have been deposited in the Genome Sequence Archive in the Beijing Institute of Genomics Data Center, Chinese Academy of Sciences, under accession number CRA024471, which are publicly accessible at https://ngdc.cncb.ac.cn/gsa/s/vYxYc7xZ. The primer sequences and the data on gene expression levels are available in the supplementary information.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos 82274037, 32460117), and the Beijing Life Science Academy (BLSA) (Grant No. 2024500CD0120), and (Grant Nos YNDG202402ZY02, YNDG202302ZY02, CNTC-110202201016 (JY-16), and YNTC-2023530000241009). The author sincerely thanks the China Association for Science and Technology (CAST) for the

academic support through the Youth Talent Lift Project (PhD Student Special Support Program).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/mpb-0025-0022)

Dates

Received 27 April 2025; Revised 17 May 2025; Accepted 20 May 2025; Published online 23 July 2025

References

1. Han Y. 2010. Tai Ping Sheng Hui Fang yu Song Dai She Hui (《太平圣惠方》与宋代社会) [Taiping Holy Prescriptions for Universal Relief and the compilation during the early Song Dynasty]. *Zhonghua yi shi za zhi [Chinese Journal of Medical History]* 40:198–205
2. Liao CP, Liu XC, Dong SQ, An M, Zhao L, et al. 2021. Investigation of the metabolites of five major constituents from *Berberis amurensis* in normal and pseudo germ-free rats. *Chinese Journal of Natural Medicines* 19(10):758–71
3. Roy NS, Park NI, Kim NS, Park Y, Kim BY, et al. 2022. Comparative transcriptomics for genes related to berberine and berbamine biosynthesis in Berberidaceae. *Plants* 11:2676
4. Liu X, Jin X, Ou H, Qian C, Wu H, et al. 2022. The direct STAT3 inhibitor 6-ethoxydihydrosanguinarine exhibits anticancer activity in gastric cancer. *Acta Materia Medica* 1(3):365–80
5. Farrow SC, Hagel JM, Facchini PJ. 2012. Transcript and metabolite profiling in cell cultures of 18 plant species that produce benzylisoquinoline alkaloids. *Phytochemistry* 77:79–88
6. Tian Y, Kong L, Li Q, Wang Y, Wang Y, et al. 2024. Structural diversity, evolutionary origin, and metabolic engineering of plant specialized benzylisoquinoline alkaloids. *Natural Product Reports* 41:1787–810
7. An Z, Gao R, Chen S, Tian Y, Li Q, et al. 2024. Lineage-specific CYP80 expansion and benzylisoquinoline alkaloid diversity in early-diverging eudicots. *Advanced Science* 11:e2309990
8. Yusupov MM, Karimov A, Shakirov R, Gorovoi PG, Faskhutdinov MF, et al. 1993. Berberis alkaloids. XXVI. An investigation of the alkaloids of *Berberis amurensis*. *Chemistry of Natural Compounds* 29:338–40
9. Roy NS, Choi IY, Um T, Jeon MJ, Kim BY, et al. 2021. Gene expression and isoform identification of PacBio full-length cDNA sequences for berberine biosynthesis in *Berberis koreana*. *Plants* 10:1314
10. Fan RF, Li HT, Kuang HY, Li W. 2019. Huanglumu de Sheng Yao Xue Yan Jiu ji Yan Suan Xiao Bo Jiang de Han Liang Ce Ding (黄芦木的生药学研究及盐酸小檗碱的含量测定) [Pharmacognosical studies and content determination of berberine hydrochloride of *Berberis amurensis* Rupr]. *Shizhen Guo Yi Guo Yao (时珍国医国药) [Lishizhen Medicine and Materia Medica Research]* 30:2074–75
11. Srivastava S, Srivastava M, Misra A, Pandey G, Rawat A. 2015. A review on biological and chemical diversity in *Berberis* (Berberidaceae). *EXCLI Journal* 14:247–67
12. Li T, Wang P, Guo W, Huang X, Tian X, et al. 2019. Natural berberine-based Chinese herb medicine assembled nanostructures with modified antibacterial application. *ACS Nano* 13:6770–81
13. Wang K, Yin J, Chen J, Ma J, Si H, et al. 2024. Inhibition of inflammation by berberine: molecular mechanism and network pharmacology analysis. *Phytomedicine* 128:155258
14. Jain S, Tripathi S, Tripathi PK. 2023. Antioxidant and antiarthritic potential of berberine: *In vitro* and *in vivo* studies. *Chinese Herbal Medicines* 15:549–55
15. Zhao MM, Lu J, Li S, Wang H, Cao X, et al. 2021. Berberine is an insulin secretagogue targeting the KCNH6 potassium channel. *Nature Communications* 12:5616
16. Hu Y, Ehli EA, Kittelsrud J, Ronan PJ, Munger K, et al. 2012. Lipid-lowering effect of berberine in human subjects and rats. *Phytomedicine* 19:861–67
17. Wang Y, Tong Q, Ma SR, Zhao ZX, Pan LB, et al. 2021. Oral berberine improves brain dopa/dopamine levels to ameliorate Parkinson's disease by regulating gut microbiota. *Signal Transduction and Targeted Therapy* 6:77
18. Chen CY, Zhang Y. 2025. Berberine: An isoquinoline alkaloid targeting the oxidative stress and gut-brain axis in the models of depression. *European Journal of Medicinal Chemistry* 290:117475
19. Phogat A, Singh J, Malik V. 2024. Pharmacological effects of berberine – a Chinese medicine, against xenobiotics toxicity. *Pharmacological Research - Modern Chinese Medicine* 13:100507
20. Lagoa R, Silva J, Rodrigues JR, Bishayee A. 2020. Advances in phytochemical delivery systems for improved anticancer activity. *Biotechnology Advances* 38:107382
21. Yi LT, Zhu JX, Dong SQ, Chen M, Li CF. 2021. Berberine exerts antidepressant-like effects via regulating miR-34a-synaptotagmin1/Bcl-2 axis. *Chinese Herbal Medicines* 13:116–23
22. Bao S, Wang X, Ma Q, Wei C, Nan J, et al. 2022. Mongolian medicine in treating type 2 diabetes mellitus combined with nonalcoholic fatty liver disease via FXR/LXR-mediated P2X7R/NLRP3/NF-κB pathway activation. *Chinese Herbal Medicines* 14:367–75
23. Zhang C, Song Y, Sun X, Liu Q, Li Z, et al. 2023. Photoredox-catalyzed reaction as a powerful tool for rapid natural product Gem-dimethylation modification: discovery of potent anti-cancer agents with improved druggability. *Acta Materia Medica* 2:400–8
24. Sun C, Sha S, Shan Y, Gao X, Li L, et al. 2025. Intranasal delivery of BACE1 siRNA and berberine via engineered stem cell exosomes for the treatment of Alzheimer's disease. *International Journal of Nanomedicine* 20:5873–91
25. Lei Y, Xie J, Xie Z, Zhao X, Huang J, et al. 2025. Comparative drug-drug interactions of berberine and astragaloside IV in normal and type 2 diabetes mellitus rats based on UPLC-QqQ-MS/MS. *Acta Materia Medica* 4:51–69
26. Liu Y, Wang B, Shu S, Li Z, Song C, et al. 2021. Analysis of the *Coptis chinensis* genome reveals the diversification of protoberberine-type alkaloids. *Nature Communications* 12:3276
27. He SM, Liang YL, Cong K, Chen G, Zhao X, et al. 2018. Identification and characterization of genes involved in benzylisoquinoline alkaloid biosynthesis in *Coptis* species. *Frontiers in Plant Science* 9:731
28. Guo L, Winzer T, Yang X, Li Y, Ning Z, et al. 2018. The opium poppy genome and morphinan production. *Science* 362:343–47
29. Hagel JM, Beaudoin GAW, Fossati E, Ekins A, Martin VJJ, et al. 2012. Characterization of a flavoprotein oxidase from opium poppy catalyzing the final steps in sanguinarine and papaverine biosynthesis. *Journal of Biological Chemistry* 287:42972–83
30. Yang H, Liu Z, Yu C, Song C, Wang C. 2023. Expression relationship between microRNA and transcription factors in *Stephania japonica*. *Medicinal Plant Biology* 2:7
31. Liu T, Zhang W, Wang S, Tian Y, Wang Y, et al. 2024. Metabolome and transcriptome association study reveals biosynthesis of specialized benzylisoquinoline alkaloids in *Phellodendron amurense*. *Chinese Herbal Medicines* 17:178–88
32. Hagel JM, Facchini PJ. 2013. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant & Cell Physiology* 54:647–72
33. Samanani N, Facchini PJ. 2002. Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzylisoquinoline alkaloid biosynthesis in plants. *The Journal of Biological Chemistry* 277:33878–83
34. Samanani N, Liscombe DK, Facchini PJ. 2004. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. *The Plant Journal* 40:302–13
35. Lee EJ, Facchini P. 2010. Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. *The Plant Cell* 22:3489–503

36. Lee S, Park NI, Park Y, Park KC, Kim ES, et al. 2024. *O*- and *N*-Methyltransferases in benzylisoquinoline alkaloid producing plants. *Genes & Genomics* 46:367−78
37. Wu L, Zhao B, Deng Z, Wang B, Yu Y. 2024. A biosynthetic network for protoberberine production in Coptis chinensis. *Horticulture Research* 11:uhad259
38. Sato F, Tsujita T, Katagiri Y, Yoshida S, Yamada Y. 1994. Purification and characterization of *S*-adenosyl-L-methionine: norcoclaurine 6-*O*-methyltransferase from cultured *Coptis japonica* cells. *European Journal of Biochemistry* 225:125−31
39. Choi KB, Morishige T, Shitan N, Yazaki K, Sato F. 2002. Molecular cloning and characterization of coclaurine *N*-methyltransferase from cultured cells of *Coptis japonica*. *Journal of Biological Chemistry* 277:830−35
40. Pauli HH, Kutchan TM. 1998. Molecular cloning and functional heterologous expression of two alleles encoding (*S*)-*N*-methylcoclaurine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. *The Plant Journal* 13:793−801
41. Morishige T, Tsujita T, Yamada Y, Sato F. 2000. Molecular characterization of the *S*-adenosyl-L-methionine: 3'-hydroxy-*N*-methylcoclaurine 4'-*O*-methyltransferase involved in isoquinoline alkaloid biosynthesis in *Coptis japonica*. *The Journal of Biological Chemistry* 275:23398−405
42. Huang P, Cheng P, Sun M, Liu X, Qing Z, et al. 2024. Systemic review of *Macleaya cordata*: genetics, biosynthesis of active ingredients and functions. *Medicinal Plant Biology* 3:e020
43. Xu Z, Li Z, Ren F, Gao R, Wang Z, et al. 2022. The genome of *Corydalis* reveals the evolution of benzylisoquinoline alkaloid biosynthesis in Ranunculales. *The Plant Journal* 111:217−30
44. Hafner J, Payne J, MohammadiPeyhani H, Hatzimanikatis V, Smolke C. 2021. A computational workflow for the expansion of heterologous biosynthetic pathways to natural product derivatives. *Nature Communications* 12:1760
45. Bu J, Zhang X, Li Q, Ma Y, Hu Z, et al. 2022. Catalytic promiscuity of *O*-methyltransferases from *Corydalis yanhusuo* leading to the structural diversity of benzylisoquinoline alkaloids. *Horticulture Research* 9:uhac152
46. Ounaroon A, Decker G, Schmidt J, Lottspeich F, Kutchan TM. 2003. (*R,S*)-Reticuline 7-*O*-methyltransferase and (*R,S*)-norcoclaurine 6-*O*-methyltransferase of *Papaver somniferum* – cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy. *The Plant Journal* 36:808−19
47. Li H, Wei W, Xu H. 2022. Drug discovery is an eternal challenge for the biomedical sciences. *Acta Materia Medica* 1:1−3
48. Li Q, Bu J, Ma Y, Yang J, Hu Z, et al. 2020. Characterization of *O*-methyltransferases involved in the biosynthesis of tetrandrine in Stephania tetrandra. *Journal of Plant Physiolog* 250:153181
49. Xu Z, Tian Y, Wang J, Ma Y, Li Q, et al. 2024. Convergent evolution of berberine biosynthesis. *Science Advances* 10:eads3596
50. Leng L, Xu Z, Hong B, Zhao B, Tian Y, et al. 2024. Cepharanthine analogs mining and genomes of *Stephania* accelerate anti-coronavirus drug discovery. *Nature Communications* 15:1537
51. Bi Z, Li H, Liang Y, Sun D, Liu S, et al. 2025. Emerging paradigms for target discovery of traditional medicines: a genome-wide pan-GPCR perspective. *The Innovation* 6:100774
52. Raghavan V, Kraft L, Mesny F, Rigerte L. 2022. A simple guide to *de novo* transcriptome assembly and annotation. *Briefings in Bioinformatics* 23:bbab563
53. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. *Bioinformatics* 26(15):1968−71
54. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644−52
55. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658−59
56. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO Update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution* 38:4647−54
57. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution* 38:5825−29
58. Kanehisa M. 2016. KEGG bioinformatics resource for plant genomics and metabolomics. *Methods in Molecular Biology* 1374:55−70
59. Ye J, Fang L, Zheng H, Zhang Y, Chen J, et al. 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research* 34:293−97
60. Hung JH, Weng Z. 2016. Sequence alignment and homology search with BLAST and ClustalW. *Cold Spring Harbor Protocols* 2016:pdb.prot093088
61. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35:1547−49
62. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, et al. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology* 28:511−15
63. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, et al. 2020. TBtools: an integrative Toolkit developed for interactive analyses of big biological data. *Molecular Plant* 13:1194−202
64. The Angiosperm Phylogeny Group, Chase MW, Christenhusz MJM, Fay MF, Byng JW, et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181(1):1−20
65. Zhong F, Huang L, Qi L, Ma Y, Yan Z. 2020. Full-length transcriptome analysis of *Coptis deltoidea* and identification of putative genes involved in benzylisoquinoline alkaloids biosynthesis based on combined sequencing platforms. *Plant Molecular Biology* 102:477−99
66. Winzer T, Gazda V, He Z, Kaminski F, Kern M, et al. 2012. A Papaver somniferum 10-gene cluster for synthesis of the anticancer alkaloid noscapine. *Science* 336:1704−8
67. Menéndez-Perdomo IM, Facchini PJ. 2023. Elucidation of the (*R*)-enantiospecific benzylisoquinoline alkaloid biosynthetic pathways in sacred lotus (*Nelumbo nucifera*). *Scientific Reports* 13:2955
68. Li F, Yuan Z, Gao Y, Deng Z, Zhang Y, et al. 2025. A concise enzyme cascade enables the manufacture of natural and halogenated protoberberine alkaloids. *Nature Communications* 16:1904
69. Han J, Li S. 2023. *De novo* biosynthesis of berberine and halogenated benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Communications Chemistry* 6:27
70. Jiao X, Fu X, Li Q, Bu J, Liu X, et al. 2024. *De novo* production of protoberberine and benzophenanthridine alkaloids through metabolic engineering of yeast. *Nature Communications* 15:8759