*Acta Biochim Biophys Sin* xxxx, xx(x): xx-xx
https://cstr.cn/32197.14.abbs.2026113
https://doi.org/10.3724/abbs.2026113
Original Article

Original Article

# Discovery of high-efficiency (*S*)-*N*-methylcoclaurine 3′-hydroxylase involved in the biosynthesis of benzylisoquinoline alkaloids in Berberidaceae

**Lingzhe Kong[1,2,†], Xiao He[2,3,†], Chenghua Gong[1], Maolun Gao[1], Ziyan Xie[1], Ruibing Chen[2,*], Zhichao Xu[1,*], and Lei Zhang[2,4,*]**

[1]Key Laboratory of National Forestry and Grassland Administration on Chinese Herbal Medicine, College of Life Science, Northeast Forestry University, Harbin 150040, China, [2]Department of Pharmaceutical Botany, School of Pharmacy, Naval Medical University, Shanghai 200433, China, [3]Jiangxi University of Chinese Medicine, Nanchang 330006, China, and [4]College of Life Sciences and Medicine, Key Laboratory of Plant Secondary Metabolism and Regulation of Zhejiang Province, Zhejiang Sci-Tech University, Hangzhou 310018, China

[†]These authors contributed equally to this work.

*Correspondence address. E-mail: rbchenstar@163.com (R.C.); zcxu@nefu.edu.cn (Z.X.); Leizhang100@163.com (L.Z.)

Received 27 April 2026 Accepted 11 May 2026 Published 13 July 2026

**Abstract**

Benzylisoquinoline alkaloids (BIAs) represent a class of plant-derived compounds with significant pharmacological importance. The biosynthesis of most BIAs originates from a common precursor (*S*)-reticuline. (*S*)-*N*-methylcoclaurine 3′-hydroxylase (NMCH), encoded by CYP80B members, has always been regarded as a rate-limiting step that restricts the efficient synthesis of (*S*)-reticuline, thereby leading to the low accumulation of BIAs. The highly-efficiency NMCH still needs to be further explored. Species within the Berberidaceae family are known to accumulate diverse BIAs at relatively high levels, making them an ideal plant resource for exploring highly active enzymes involved in BIA synthesis. Here, by integrating transcriptome and metabolite analysis across 11 Berberidaceae plants, we mine and characterize candidate genes involved in (*S*)-reticuline biosynthesis. We further establish an engineered yeast platform for functional screening of NMCH genes and evaluate the catalytic activity of all CYP80B candidates. Among these, MbNMCH, isolated from *Mahonia bealei*, exhibits significantly higher catalytic activity for (*S*)-reticuline production in yeast compared to previously reported NMCH enzymes. Our findings provide abundant genetic and metabolic information on Berberidaceae plants and identify a highly efficient enzymatic tool for BIA production in microbial cell factories, facilitating the sustainable manufacturing of diverse valuable alkaloids.

**Key words** Berberidaceae, benzylisoquinoline alkaloids, NMCH, microbial cell factories

## Introduction

Benzylisoquinoline alkaloids (BIAs) represent a major class of plant-derived compounds that are primarily distributed in magnoliids and early diverging eudicots, including the orders Piperales, Laurales, Magnoliales, Ranunculales, and Proteales [1,2]. The structural diversity of BIAs is remarkable, with over 2500 identified compounds classified into nine principal categories, including 1-benzylisoquinolines (1-BIAs), aporphines, protoberberines (BBRs), benzophenanthridines (BZPs), phthalideisoquinolines (PQs), morphinans, protopines, bisbenzylisoquinolines (bisBIAs), and pavines [1]. Numerous BIAs possess significant pharmacological properties. For instance, morphine and codeine serve as widely utilized narcotic analgesics [3], while papaverine is employed as a vasodilator and antispasmodic agent [4,5]. Additionally, antimicrobial alkaloids, including sanguinarine, coptisine, tetrandrine, and berberine, have been well characterized [6–8]. Other BIAs, such as noscapine, demonstrate antitussive and anticancer activities [9,10]. Furthermore, boldine and rotundine have demonstrated antischistosomal properties [11]. With the exception of bisBIAs, all the bioactive compounds mentioned above are derived from a common

© The Author(s) 2026. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/).

precursor, (*S*)-reticuline. Through a series of reactions, including isomerization, coupling, rearrangement, and methylation, (*S*)-reticuline can give rise to almost all structural types of BIAs. Due to its central role, (*S*)-reticuline has become a major focus in synthetic biology research on BIAs, aiming to synthesize (*S*)-reticuline with high yields in microbial cell factories. Current advances in synthetic biology research have achieved microbial production of (*S*)-reticuline with the highest titer of 4.6 g/L in bioreactors [12]. However, since the catalytic efficiency of the rate-limiting enzyme NMCH is relatively low, increasing its copy number remains the primary strategy for enhancing reaction efficiency [12–15]. Continuously increasing the number of copies of an enzyme in microbial cell factories will significantly increase the difficulty of genetic manipulation, cytotoxicity, and robustness. Consequently, the discovery of high-activity NMCH variants is essential for further progress.

The biosynthetic pathway of (*S*)-reticuline originates from the aromatic amino acid *L*-tyrosine [1] (Figure 1). Through a series of enzymatic transformations, *L*-tyrosine undergoes decarboxylation and hydroxylation via tyrosine decarboxylase (TYDC) and polyphenol oxidases (PPOs) to generate dopamine, which serves as one precursor for norcoclaurine formation [16]. Concurrently, a parallel metabolic branch involving tyrosine aminotransferase (TyrAT) and 4-hydroxyphenylpyruvate decarboxylase (HPPDC) converts *L*-tyrosine to 4-hydroxyphenylacetaldehyde (4-HPAA), the second precursor. Norcoclaurine synthase (NCS) then catalyzes the Pictet-Spengler condensation of dopamine and 4-HPAA to form (*S*)-norcoclaurine, the universal precursor for all BIAs [17–19]. Subsequently, (*S*)-norcoclaurine undergoes a four-step enzymatic modification involving three methyltransferases and one cytochrome P450 monooxygenase (CYP), leading to the formation of the central intermediate (*S*)-reticuline. These sequential reactions include 6-*O*-methylation catalyzed by (*S*)-norcoclaurine 6-*O*-methyltransferase (6OMT) [20], *N*-methylation by (*S*)-coclaurine *N*-methyltransferase (CNMT) [21], 3′-hydroxylation by (*S*)-*N*-methylcoclaurine 3′-hydroxylase (NMCH, CYP80B) [22], and finally 4′-*O*-methylation by (*S*)-3′-hydroxy-*N*-methylcoclaurine 4′-*O*-methyltransferase (4′OMT) [23]. The structural diversification of BIAs proceeds through various cyclizing reactions catalyzed by the CYP and berberine bridge enzyme (BBE)-mediated C-C and C-O coupling, yielding various structural skeletons. CYPs constitute a superfamily of hemoproteins that play oxidative roles in BIA biosynthesis, participating in hydroxylation, isomerization, and coupling reactions [16,24]. Among the CYP families involved in BIA modification, CYP80, CYP82, and CYP719 have been functionally characterized [1]. The CYP80 family plays a crucial role in the diverse modifications of BIA and specifically catalyzes hydroxylation, intermolecular C-O, and intramolecular C-C phenol coupling reactions, predominantly accepting 1-benzylisoquinoline substrates. CYP80B, encoding NMCH, specifically catalyzes the 3′-hydroxylation of (*S*)-*N*-methylcoclaurine to yield (*S*)-3′-hydroxy-*N*-methylcoclaurine [22,25–27].

Berberidaceae contains a diverse group of angiosperms comprising 19 genera and approximately 700 species worldwide [28]. Various organs of these plants have been utilized in the food, dyeing, and pharmaceutical industries, with a long history of application in traditional medicine [29,30]. These species possess numerous medicinal properties, including wound healing, hepatoprotective, antioxidant, antibacterial, and antihyperlipidemic activities, due to the presence of abundant secondary metabolites [31,32]. Among the specialized metabolites in Berberidaceae, BIAs are the most representative and specific type, especially protoberberines, aporphines, 1-BIAs, and bis-benzylisoquinoline alkaloids. However, the identification and application of enzyme resources within Berberidaceae for BIA biosynthesis remain limited. To address this knowledge gap, we generated comprehensive transcriptome datasets from eleven Berberidaceae and identified

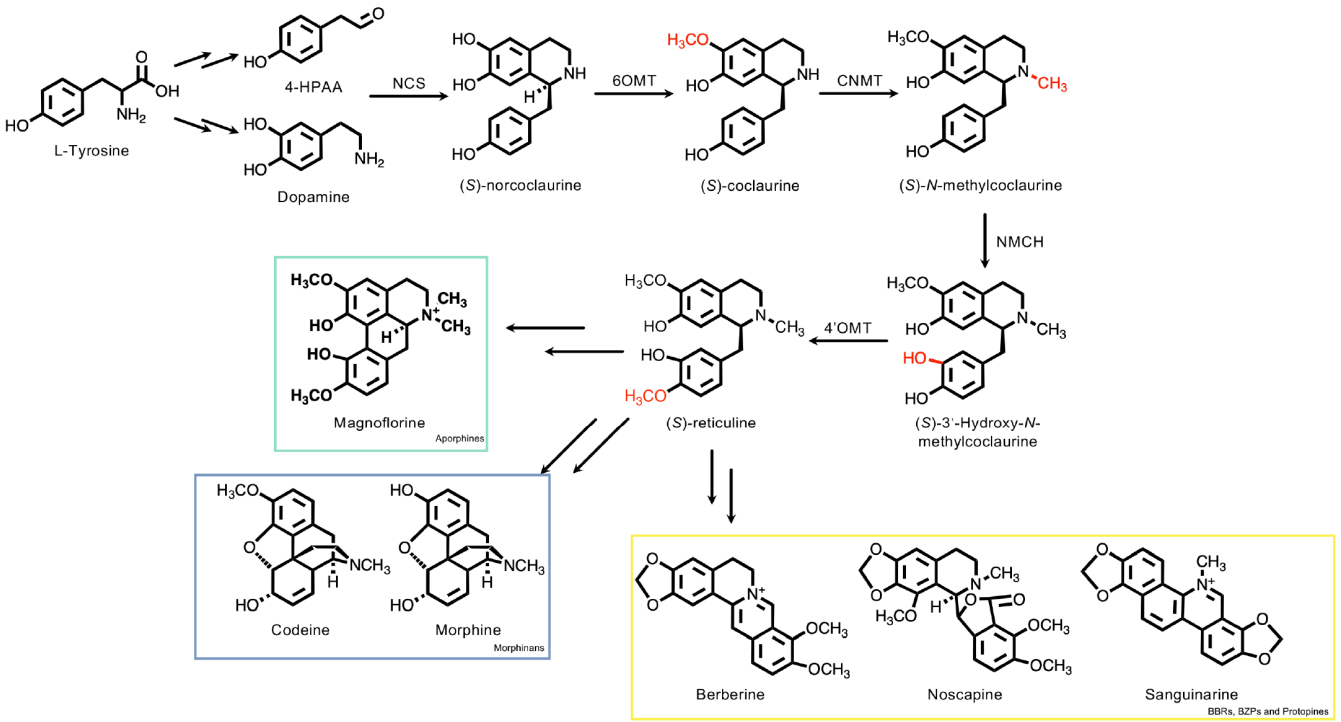


**Figure 1. Biosynthesis of multiple kinds of BIAs derived from (*S*)-reticuline**

candidates potentially involved in BIA biosynthesis. We subsequently performed *in vivo* functional characterization of CYP80B from different Berberidaceae species, and thus, the one with the highest activity was identified. Finally, we tested its application in a yeast cell factory, which significantly improved the synthesis efficiency of (*S*)-reticuline. Collectively, these resources provide biochemical insights into BIA metabolism in these underexplored systems, facilitating pathway elucidation and the discovery of novel catalytic elements. Moreover, the identification of functional enzymes from diverse plant species expands the genetic toolbox essential for synthetic biology approaches aimed at BIA production.

## Materials and Methods

### Plant materials, chemicals, reagents, and strains

Plant materials of Berberidaceae were collected in July 2023 from the natural habitat in Baoshan city, Yunnan province, China and Heilongjiang Provincial Botany Garden. 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 of (*S*)-*N*-methylcoclaurine and (*S*)-3′-hydroxy-*N*-methylcoclaurine were obtained from Shanghai Yuanye Bio-Technology Co., Ltd. (Shanghai, China). Chromatographically pure reagents were employed in the compound detection tests.

The LB medium was composed of 10 g/L peptone, 5 g/L yeast extract, and 10 g/L sodium chloride, serving as a standard medium for cultivating *Escherichia coli*. The YPD medium was composed of 20 g/L peptone, 10 g/L yeast extract, and 20 g/L glucose, serving as a standard medium for cultivating yeast and preparing competent cells. *E. coli* strains containing URA3-based plasmids and cassettes were selected on SD medium (20 g/L glucose and 6.7 g/L yeast nitrogen base without amino acids) with histidine. For plate preparation, agar at a concentration of 20 g/L was added as necessary. Shake-flask fermentation was performed in 20 mL of minimal medium (Delft-D) containing 2.5 g/L $(NH_4)_2SO_4$, 14.4 g/L $KH_2PO_4$, 0.5 g/L $MgSO_4 \cdot 7H_2O$, 20 g/L glucose, 2 mL of trace metal and 1 mL of vitamin solution supplemented with 40 mg/L histidine and 60 mg/L uracil if needed. All of these reagents and chemicals were purchased from Sangon Biotech Co., Ltd. (Shanghai, China) unless stated otherwise.

*Saccharomyces cerevisiae* strain CEN.PK113-11C* (*MATa*, *SUC2*, *MAL2-8c*, *his3Δ*, *ura3Δ*, *gal80Δ*, *XI5*::($P_{TEF}$-*CAS9*-$T_{CYC1}$) derived from CEN.PK113-11C (*MATa*, *SUC2*, *MAL2-8c*, *his3Δ1*, *ura3Δ-52*) was used as the background strain for all genetic manipulations and strain construction [33]. The *E. coli* Trans1 T1 strain from TransGen Biotech (Beijing, China) was used for routine plasmid assembly. *S. cerevisiae* strains and plasmids in this study were all from our laboratory.

### RNA library construction, sequencing and analysis of BIA candidate genes

RNA from different tissues was extracted using TRIzol Reagent (Thermo Fisher Scientific, Carlsbad, USA). RNA quality was based on UV absorption ratios (Thermo Fisher Scientific), where only samples with ratios above 2.0 (260/280 nm) and 2.2 (260/230 nm) were used. Then, cDNA was generated using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (AT311-02; TransGen Biotech), according to the manufacturer's protocol. The Illumina HiSeq platform from Novogene (Beijing, China) was utilized for paired-end sequencing. Following RNA-seq, adaptors, low-quality nucleotides, and unknown nucleotides were eliminated from the raw reads to create clean reads. Raw sequencing reads were subjected to quality control using Trimmomatic (v0.39) [34]. Adapter sequences, reads containing ambiguous bases, and low-quality reads were removed to obtain clean reads. The quality-filtered reads were then used for subsequent *de novo* transcriptome assembly and functional annotation analyses. FastQC software [35] 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) [36], employing the default settings. The CD-HIT-EST program [37] was used to eliminate duplicate contigs. TransDecoder (v3.0.1) was finally used to predict the coding region and the read counts. The completeness of these transcriptome assemblies was evaluated using the Embryophyta_odb10 dataset, which contains 1614 evolutionarily conserved embryophyte genes, through the Benchmarking Universal Single-Copy Orthologs (BUSCO) tool (v5.7.1) [38].

The transcriptome sequences were blasted with an E-value cutoff of $1 \times 10^{-5}$ to identify the sequences encoding the berberine biosynthetic enzymes. ClustalW [39] was employed for aligning the protein sequences, followed by phylogenetic analysis. Based on the reference genes of BIA-producing plants and the obtained candidate genes, phylogenetic analysis was conducted using MEGA11 [40] through maximum likelihood estimation. Gene expression levels were quantified as FPKM (fragments per kilobase of exon per million fragments mapped) values using R packages [41]. The resulting expression matrix was visualized as a heatmap using TBtools [42]. 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.

### Cloning of candidate genes and expression cassette construction

Total RNA was isolated from candidate plant leaf, stem, and root tissues using TRIzol Reagent. First-strand cDNA was synthesized from 1 μg of total RNA using the TransScript One-Step gDNA Removal and cDNA Synthesis SuperMix kit (TransGen Biotech) according to the manufacturer's protocol. All candidate genes were amplified from total cDNA using the primers listed in Supplementary Table S1 and 2× Phanta Flash Master Mix (Vazyme, Nanjing, China) according to the manufacturer's instructions. Genes were transformed into competent *E. coli* Trans1 T1 and identified using colony PCR (TransGen Biotech) and Sanger sequencing. The identified positive recombinant plasmids were extracted from Trans 1 T1 using the TIANprep Mini Plasmid Kit (TIANGEN, Beijing, China).

All gene overexpression cassettes, consisting of a promoter, a target gene, a terminator and a block of the next cassette for homologous recombination, were constructed by overlapping extension PCR [43]. The promoters and terminators were cloned from genomic DNA *S. cerevisiae* CEN.PK113-11C. All primers and expression cassette details used for engineering yeast construction are listed in Supplementary Tables S1 and S2.

### Heterologous expression and functional validation in *S. cerevisiae*

The overexpression cassettes and a guide RNA (gRNA) plasmid

were cotransformed into the (*S*)-*N*-methylcoclaurine expression strain LZ38 for cassette knock-in by the CRISPR-Cas9 method as previously reported [33]. High-efficiency yeast transformation was performed using the LiAc/SS carrier DNA/PEG method [44], and transformants were selected on plates containing solid synthetic medium (SD) supplemented with 0.02 g/L histidine (SD + histidine). Clones were verified by colony PCR and sequencing. Subsequently, verified clones were cultivated overnight in SD + histidine liquid medium and then plated on solid medium containing SD + histidine + uracil + 5-fluoroorotic acid (5-FOA, 1 g/L) to loop out the gRNA vectors and recycle the URA3 marker.

### Alkaloid extraction, strain cultivation and metabolite quantitation

All plant samples were freeze dried overnight and then ground into powder and filtered with a 60-mesh sieve. Fifty milligrams of root, stem, and leaf powders were mixed with 2 mL of methanol, and then the extracts were sonicated overnight. After sonication, the samples were centrifuged at 15,000 *g* for 15 min. The supernatant was collected in a clean tube and then filtered through a 0.22-μM nylon syringe filter for analysis.

Yeast cells were normally cultivated in YPD media. Strains containing URA3- and/or HIS3-based plasmids/cassettes were selected on SD media without uracil or histidine, respectively. The URA3 marker plasmid was removed on SD + uracil + 5-FOA plates. Shake flask fermentations for the production of BIAs were performed in 20 mL of Delft-D medium. Cultures were inoculated at an initial $OD_{600}$ of 0.1 in 100-mL flasks and cultivated at 200 rpm and 30°C for 72 h. After fermentation, metabolites were extracted from the culture. One milliliter of culture was filtered through a 0.22-μM organic membrane. Samples were then stored at –80°C until analysis.

Chromatographic separation was performed on an LC-2030C 3D system (Shimadzu, Shenyang, China) equipped with an Ultimate® UHPLC XB-C18 column (2.1 mm × 100 mm, 1.8 μM) from Welch Materials Co., Ltd. (Shanghai, China) maintained at 30°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 (3–10 min), 100% B (10–15 min), returning to 5% B (15–20 min), and holding for 10 min (20–30 min). Detection was performed at 280 nm with an injection volume of 5 μL. For LC-MS/MS analysis, the LCMS-9030 Q-TOF system (Shimadzu) employed identical mobile phases with an optimized gradient: 15% B (0–1 min), a linear increase to 45% B (1–8 min), 80% B (8–10 min), a return to 15% B (10–11 min) and a hold 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. Positive electrospray ionization (ESI+) was used for MS detection with nitrogen nebulizing gas (3.0 L/min) and drying gas (10 L/min at 300°C), scanning *m*/*z* 100–1000. Each sample was injected at a volume of 1 μL. The target product peak was identified by comparing its retention time with the corresponding authentic standard. Peak areas were integrated from the HPLC chromatograms under identical analytical conditions and used to evaluate relative product accumulation among strains. All samples were analyzed using three biologically independent replicates.

### Statistical analysis

All experiments were performed with three independent biological replicates. Data are presented as the mean ± standard deviation (SD). Statistical analyses were conducted using GraphPad Prism 9.0 (GraphPad Software, San Diego, USA). Differences between two groups were evaluated using a two-tailed Student's *t*-test. A $P$ value $< 0.05$ was considered statistically significant.

## Results

### Metabolite profiling of BIAs in Berberidaceae

Berberidaceae species are recognized as a key plant family producing a diverse array of BIA compounds, including berberine, magnoflorine, palmatine, and protopine [45–47]. While previous investigations on BIA biosynthesis in Berberidaceae have primarily focused on the genus *Berberis*, studies on other genera within this family remain scarce. To our knowledge, no systematic investigation has been conducted to profile the content of BIAs or the genes associated with their biosynthesis across diverse genera of Berberidaceae. To address this research gap, we selected 11 representative species from Berberidaceae, including *Berberis amurensis*, *Berberis sieboldii*, *Caulophyllum robustum*, *Dysosma delavayi*, *Dysosma versipellis*, *Epimedium brevicornu*, *Mahonia bealei*, *Mahonia fortunei*, *Nandina domestica*, *Plagiorhegma dubia*, and *Sinopodophyllum hexandrum*. Roots, stems, and leaves from all candidate species were collected for metabolite profiling and transcriptome assembly. Berberidaceae, as a member of the magnoliid plants, is capable of synthesizing the type I aporphine alkaloid magnoflorine. Concurrently, studies have reported the presence of bisBIAs such as berbamine and protoberberine-type BIA in *berberis* species. Therefore, to characterize the distribution of different types of BIAs in Berberidaceae, we selected five representative molecules of these three categories of BIAs as characteristic markers for detecting alkaloids in berberidaceae: magnoflorine, berbamine, jatrorrhizine, coptisine, and berberine (Figure 2A). All plant extracts and standards were analyzed using liquid chromatography, and a representative chromatogram is shown in Figure 2B and Supplementary Figures S1–S11. The peak areas of different compounds were quantified to reflect their accumulation levels within the plants (Figure 2C). Magnoflorine was detected in all samples, and its distribution pattern was consistent across all these species. Protoberberine alkaloids were detected in *Berberis amurensis*, *Berberis sieboldii*, *Mahonia bealei*, *Mahonia fortunei*, *Nandina domestica*, and *Plagiorhegma dubia*. Berberine and coptisine were predominantly distributed in the roots and stems, whereas less jatrorrhizine was found in roots, stems, and leaves compared to the other two compounds. Notably, berbamine was not detected in *Berberis* plants, which is inconsistent with previous reports on bisBIAs in Berberidaceae. We speculate that this is due to the low content of bisBIAs in the samples we analyzed, which did not reach the detection limit of liquid chromatography. These results suggest that aporphine and protoberberine alkaloids constitute the principal BIA classes in Berberidaceae, along with a low content of bisBIAs.

### Transcriptome sequencing, assembly, and analysis of Berberidaceae

Considering that the distribution of BIAs in plants shows significant tissue specificity (Figure 2C), we collected roots, stems, and leaves for transcriptome analysis. Following quality control (removal of

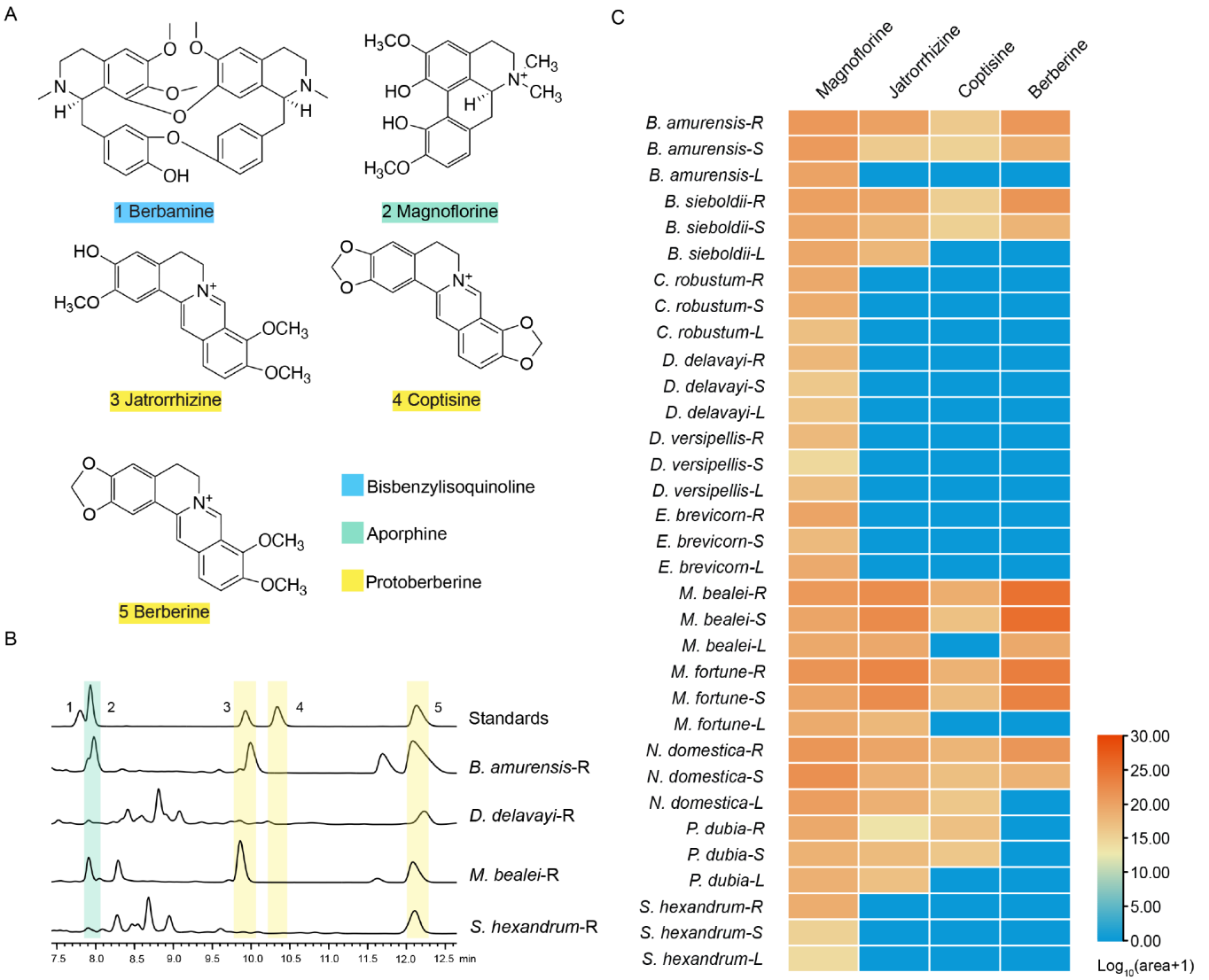


**Figure 2. Determination of BIAs in berberidacea** (A) The number and structure of the five BIAs. (B) HPLC chromatogram of representative plant tissues. (C) Heatmap analysis of the peak area of BIAs in Berberidaceae.

adapter sequences and low-quality reads), clean reads were obtained for three tissue libraries (root-R, stem-S, and leaf-L). Information on the *de novo* assembly, N50 length and GC content is shown in Supplementary Table S3. Benchmarking Universal Single-Copy Ortholog (BUSCO) assessment [38] demonstrated that these transcriptome assemblies contained highly complete BUSCOs, with completeness scores ranging from 88.3% to 93.7% (Supplementary Table S4). These high-quality transcriptomic resources provide a robust foundation for subsequent gene discovery and expression profiling analyses.

With the exception of the previously reported *BaNCS4* and *BsCYP80A* [48,49], the biosynthetic pathways for BIAs in Berberidaceae remain largely uncharacterized. Given that (*S*)-reticuline is the precursor for the synthesis of most BIAs, the four steps catalyzing (*S*)-norcoclaurine to generate (*S*)-reticuline received the most attention in this study (Figure 3A). We identified candidate genes exhibiting high sequence similarity to previously characterized BIA biosynthetic genes, particularly *CYP80B*. As a result, a total of 82 genes related to the enzymes in the (*S*)-reticuline biosynthetic pathway were identified, including 29 6OMTs, 27 CNMTs, 9 NMCHs and 17 4′OMTs (Figure 3B).

To find more active NMCHs for application in cell factories, a phylogenetic tree of NMCH was constructed by using *AtCYP51G1* as an outgroup. Given the sequence similarity between the CYP76G and CYP80 families, we included *CYP76G* genes to enhance the resolution of candidate gene identification. Detailed information regarding these protein sequences were provided in Supplementary Table S5. Genes with sequence homology levels exceeding 65% were selected as strong candidates, and 9 genes were identified across most species, with the exception of *D. delavayi* and *D. versipellis* (Figure 3B,C). The absence of candidate genes in these two species may be caused by the incompleteness of transcriptome sequencing or genuine genetic differences. To analyze the expression patterns of these candidate transcripts across different tissues of Berberidaceae, we generated a heatmap based on the FPKM values of the candidate genes (Figure 3B). In plants with high BIA content, such as *B. amurensis*, *B. sieboldii*, *M. bealei*, *M. fortunei*, and *N. domestica*, NMCH exhibited predominant expression in the

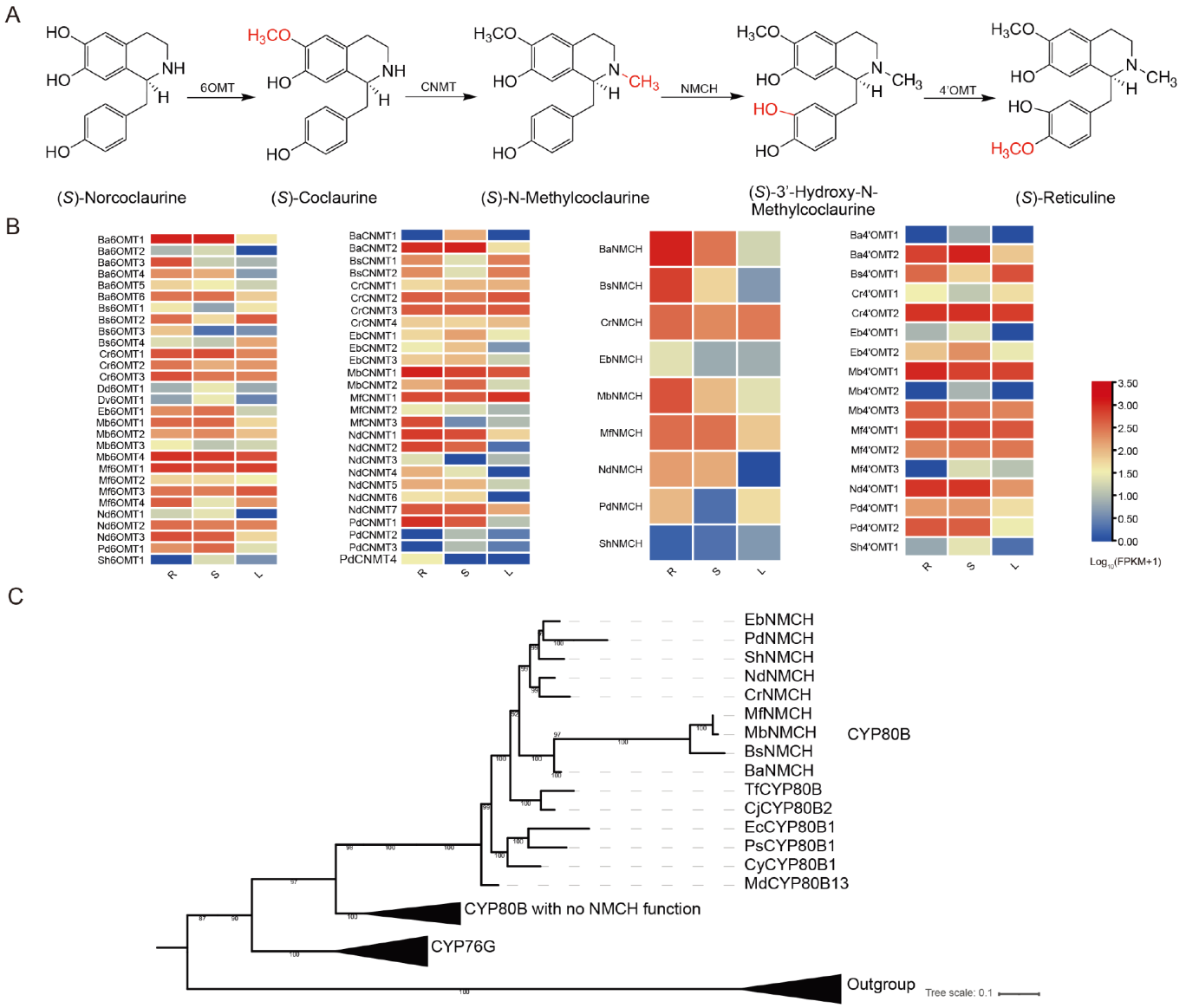


**Figure 3. Expression pattern analysis of genes involved in (*S*)-reticuline biosynthesis** (A) Biosynthesis pathway of (*S*)-reticuline. (B) Heatmap analysis of the expression pattern of genes. (C) Phylogenetic tree of NMCH in Berberidaceae.

roots and stems. In contrast, NMCH expression was barely detectable in plants with low BIA accumulation, including *E. brevicornum*, *P. dubia*, and *S. hexandrum*. After verifying these sequences and excluding identical gene sequences, we successfully cloned all these candidate genes from cDNA. Following sequence confirmation, we expressed them in yeast strains for functional characterization.

## Functional verification of CYP80B in Berberidaceae

To elucidate the function of CYP80B in Berberidaceae, we performed *in vivo* functional validation. CYP80B is the enzyme that specifically catalyzes the 3′-hydroxylation of (*S*)-*N*-methylcoclaurine to produce (*S*)-3′-hydroxy-*N*-methylcoclaurine, a precursor for numerous structurally complex BIAs (Figure 4A). For *in vivo* functional characterization, we constructed each candidate gene into a recombination cassette and introduced it into an engineered *S. cerevisiae* strain capable of synthesizing (*S*)-*N*-methylcoclaurine, the substrate for the CYP80B-mediated reaction. This substrate-producing strain named LZ37 was constructed to determine NMCH. The genotype of LZ37 is listed in Supplementary Table S2. Based on the yeast strain CEN. PK113-11C*, we expressed the enzymes for synthesizing (*S*)-norcoclaurine, along with Ps6OMT and PsCNMT, together with AtCPR1 from *A. thaliana* to ensure proper expression of CYP. The production of LZ37 was detected by LC–MS, and analysis showed that a product peak with a m/*z* of 300 appeared at the same retention time as the (*S*)-*N*-methylcoclaurine standard (Figure 4B,C). Each cassette contained a neutral locus sequence for genomic integration, along with the necessary promoter and terminator for gene expression. Following the assembly of all expression cassettes and their subsequent transformation into the yeast chassis, we selected six individual clones from each recombinant strain for transformant verification. Functional analyses were conducted using three biological replicates for each transgenic strain. Using EcNMCH and CyNMCH as positive

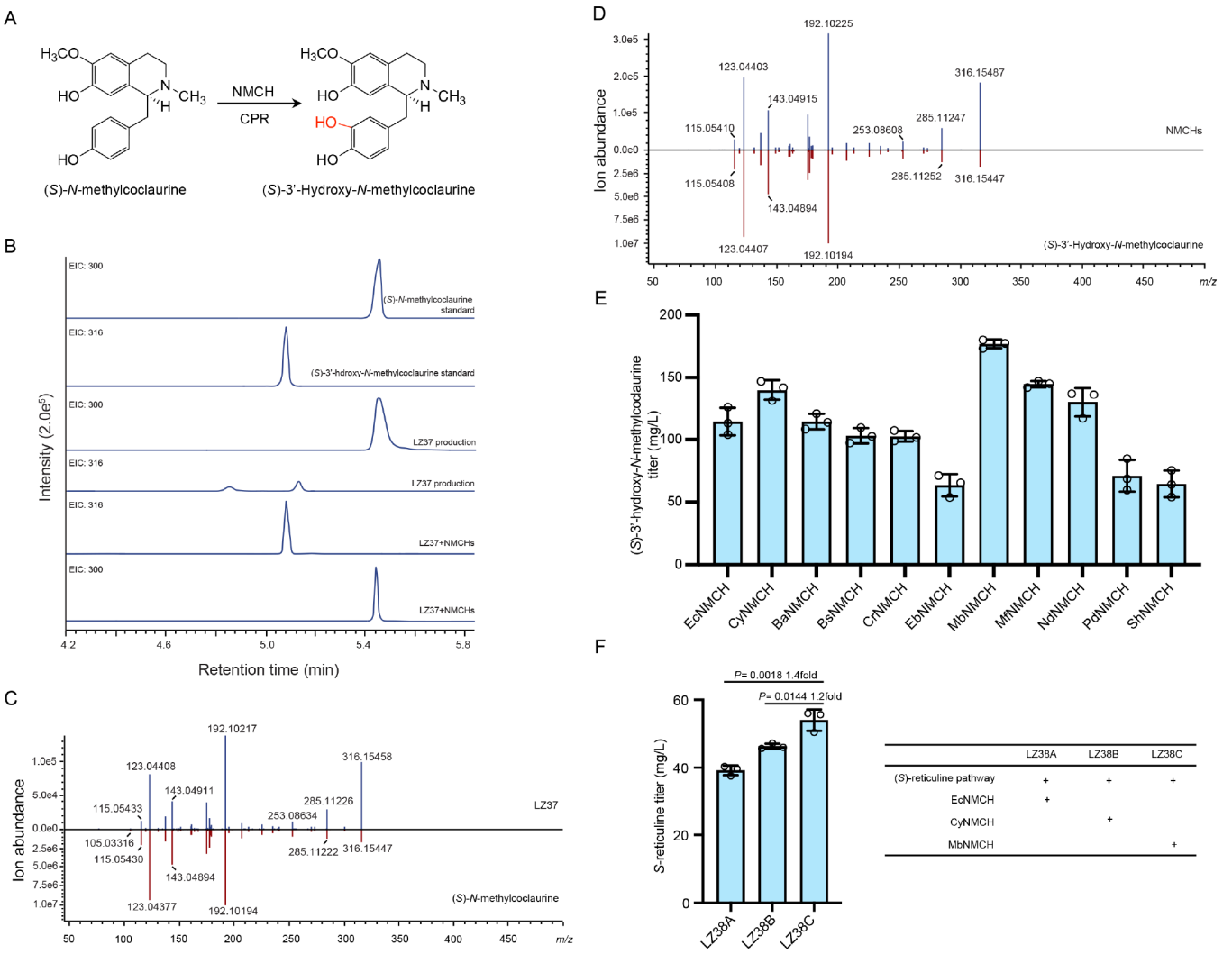


**Figure 4. Functional analysis of NMCH candidates from Berberidaceae** (A) The synthesis of (*S*)-3′-hydroxy-*N*-methylcoclaurine catalyzed by NMCH. (B) The LC-MS extracted ion chromatograms at *m*/*z* 300 and *m*/*z* 316 represent the substrate (*S*)-*N*-methylcoclaurine and the catalytic product generated by NMCH, respectively. (C) Comparison of MS/MS fragments between the product generated by the substrate-producing strain (above) and the (*S*)-*N*-methylcoclaurine standard (below). (D) Comparison of MS/MS fragments between the product from the NMCH-expressing strain (above) and the (*S*)-3′-hydroxy-*N*-methylcoclaurine standard (below). (E) Peak area of each NMCH-expressing strain. (F) (*S*)-reticuline titers of strains expressing different NMCH enzymes. For (E) and (F), the data are shown as the mean ± SD from $n = 3$ biologically independent samples with experiments conducted in shake flasks. Statistical analysis was conducted using Student's *t* test.

controls, we demonstrated that all candidate genes exhibited enzymatic activity. LC-MS analyses revealed the emergence of a product peak with *m*/*z* 315 at identical retention times across all active transformations. By comparing the retention time and mass spectral data of this product with a standard, the product was unambiguously identified as (*S*)-3′-hydroxy-*N*-methylcoclaurine (Figure 4B,D). These results confirm that the identified NMCH enzymes catalyze the 3′-hydroxylation of (*S*)-*N*-methylcoclaurine to form (*S*)-3′-hydroxy-*N*-methylcoclaurine. To quantitatively compare the catalytic activities of the NMCH variants, we analyzed the peak areas of the product formed by each candidate gene alongside those of the previously characterized EcNMCH and CyNMCH, the latter being utilized in multiple BIA production systems. The titer of (*S*)-3′-hydroxy-*N*-methylcoclaurine indicated that MbNMCH yielded the highest product peak area, suggesting that MbNMCH possesses superior catalytic activity compared to the two reference enzymes (Figure 4E). In previous BIA synthetic biology research, EcNMCH and CyNMCH have been widely employed as highly efficient modules for the synthesis of various products. Compared to these two enzymes, MbNMCH demonstrates significantly higher catalytic activity, highlighting its potential for applications in synthetic biology studies. To further validate the catalytic activity of MbNMCH, we constructed (*S*)-reticuline-producing strains with three different NMCHs. Under identical conditions with coexpression of 6OMT, CNMT, and 4′OMT combined with the three NMCH enzymes, the strain expressing MbNMCH produced significantly higher levels of (*S*)-reticuline than the other two NMCHs (Figure 4F). In the (*S*)-reticuline-producing strain expressing MbNMCH, the yield increased by 1.4-fold and 1.2-fold compared to strains expressing EcNMCH and CyNMCH, respectively. This result demonstrates that MbNMCH can significantly enhance the catalytic efficiency of microbial cell factories for the synthesis of (*S*)-

reticuline and its downstream derivations. MbNMCH possesses superior catalytic efficiency and can be advantageously employed in synthetic biology applications.

## Discussion

The majority of BIAs have been progressively elucidated, yet the complex enzymatic transformations from 1-benzylisoquinolines to diverse BIA scaffolds, encompassing C-C/C-O coupling, hydroxylation, methylation, and acylation, continue to represent a dynamic frontier in plant specialized metabolism [1]. Our study provides the first comprehensive transcriptomic landscape of eleven species within the Berberidaceae family. By integrating metabolite profiling with transcriptome analysis, we identified and functionally characterized a suite of candidate genes implicated in BIA biosynthesis. *In vivo* functional validation confirmed that the identified CYP80B homologs specifically catalyze the 3′-hydroxylation of (*S*)-*N*-methylcoclaurine, a vital reaction of central metabolites toward a multitude of structurally complex BIAs. Our metabolite analysis revealed the distinct accumulation patterns of five representative BIAs across the studied species. The ubiquitous presence of magnoflorine, an aporphine-type alkaloid, suggests a highly conserved and active biosynthetic branch for aporphines across Berberidaceae. In contrast, the restricted occurrence of protoberberines, including berberine and jatrorrhizine, in specific genera (*e.g.*, *Berberis* and *Mahonia*) may reflect lineage-specific evolution of the regulatory networks and biosynthetic enzymes associated with protoberberine production. Notably, the absence of detectable berbamine levels in any species, contrary to some previous reports [49–51], could be attributed to several factors: concentrations falling below our analytical detection threshold, the specific developmental stages or environmental conditions of the sampled tissues, or a genuine lack of the requisite CYP80A-like enzymes for the C-O phenolic coupling necessary for bisbenzylisoquinoline assembly in these particular accessions.

Then, we sequenced the transcriptome of 11 species of Berberidaceae using next-generation sequencing, followed by *de novo* assembly and functional annotation. Previous studies have shown that CYPs are involved in all biosynthetic pathways of BIAs, specifically CYP80, which participates in hydroxylation and C-O and C-C coupling reactions [52–55]. Through gene annotation and phylogenetic tree construction, we identified candidate genes in the same branch from NMCH. The heatmap shows that the majority of genes were expressed in roots, which is the same as metabolite accumulation.

The enormous demand for bioactive BIAs far exceeds the capacity and sustainability of traditional plant extraction, thereby intensifying the focus on developing heterologous microbial production platforms. A critical bottleneck in many engineered pathways is the catalytic efficiency of plant-derived enzymes. In this context, our functional screening yielded a significant discovery: MbNMCH from *Mahonia bealei* demonstrated superior catalytic activity compared to the frequently referenced CyNMCH and EcNMCH. The identification of a high-performance enzyme is a valuable asset for synthetic biology. To investigate the differences in activity among MbNMCH, CyNMCH, and EcNMCH, we performed multiple sequence alignment of these three enzymes to explore whether variations in key amino acids contribute to the differences in activity. The results are shown in Supplementary Figure S12. Among the three NMCHs, the reported conserved Thr residue is identical and conserved, and the remaining conserved motifs are also present. Therefore, structural elucidation of MbNMCH is necessary in future studies. Integrating MbNMCH into optimized microbial chassis could enhance metabolic flux through the critical reticuline node, thereby boosting the titers of a wide spectrum of valuable BIA derivatives, as evidenced by successful precedents employing engineered plant CYPs in yeast. The superior activity of MbNMCH offers a route to intensify precursor flux without resorting to non-native bypasses, preserving the authenticity of the pathway and potentially simplifying the overall engineering strategy.

In conclusion, this study systematically unravels the transcriptional and metabolic underpinnings of BIA biosynthesis in the under-explored Berberidaceae family. We identified and validated a set of functional enzymes, most notably the highly active MbNMCH, responsible for a key hydroxylation step in the BIA backbone. Our findings not only clarify the enzymatic framework for BIA biosynthesis in these species but also provide a portfolio of validated genetic elements to address the yield limitations that currently constrain the microbial production of these medicinally indispensable alkaloids. The resources and insights presented herein significantly advance our understanding of specialized metabolism evolution in Berberidaceae and pave the way for more efficient biotechnological production of BIA-based pharmaceuticals.

## Supplementary Data

Supplementary data are available at *Acta Biochimica et Biophysica Sinica* online.

## Data Availability

The raw RNA-seq data are available via the CNGB Genome Sequence Archive under BioProject accession number PRJCA060102. All other data supporting the findings of this study, including raw data and original images, are available from the corresponding author upon reasonable request.

## Author Contributions

L.Z., R.C., and Z.X. conceptualized the original research. L.K. designed the experiments. L.K., X.H., C.G., and M.G. performed the experiments and analyzed the data. L.K. and X.H. wrote the paper.

## Funding

This work was supported by the grants from the National Key Research and Development Program of China (No. 2023YFC3504800 to R.C.), the National Natural Science Foundation of China (Nos. 82274037, 82225047, and 82522090 to L.Z.), the Open Fund Project of State Key Laboratory of Synthetic Biology (No. HSK202505 to Z.X.), Beijing Life Science Academy (BLSA, No. 2024500CD0120 to Z.X.) and the Special Project on Achievement Transformation of the Qinghai Provincial Science and Technology Department (No. 2025-SF-137 to R.C.).

## Conflict of Interest

The authors declare that they have no conflict of interest.

## References

1. Tian Y, Kong L, Li Q, Wang Y, Wang Y, An Z, Ma Y, *et al.* Structural diversity, evolutionary origin, and metabolic engineering of plant

specialized benzylisoquinoline alkaloids. *Nat Prod Rep* 2024, 41: 1787–1810
2. Leibetseder L, Bindics J, Buyel JF. Benzylisoquinoline alkaloid production: moving from crop farming to chemical and biosynthesis. *Biotechnol Adv* 2025, 85: 108700
3. Beaudoin GAW, Facchini PJ. Benzylisoquinoline alkaloid biosynthesis in opium poppy. *Planta* 2014, 240: 19–32
4. Baltsavias G, Yella S, Al Shameri RA, Luft A, Valavanis A. Intra-arterial administration of papaverine during mechanical thrombectomy for acute ischemic stroke. *J Stroke Cerebrovasc Dis* 2015, 24: 41–47
5. Girard DS, Sutton Iii JP, Williams TH, Crumbley Iii AJ, Zellner JL, Kratz JM, Crawford FA. Papaverine delivery to the internal mammary artery pedicle effectively treats spasm. *Ann Thoracic Surg* 2004, 78: 1295–1298
6. Qu Q, Huang X, Zhu Z, Wang J, Zhao M, Cui W, Zheng Y, ***et al.*** Targeting membrane integrity and imidazoleglycerol-phosphate dehydratase: sanguinarine multifaceted approach against *Staphylococcus aureus* biofilms. *Phytomedicine* 2025, 138: 156428
7. Inoue N, Terabayashi T, Takiguchi-Kawashima Y, Fujinami D, Matsuoka S, Kawano M, Tanaka K, ***et al.*** The benzylisoquinoline alkaloids, berberine and coptisine, act against camptothecin-resistant topoisomerase I mutants. *Sci Rep* 2021, 11: 7718
8. Park JJ, Seo SM, Kim EJ, Lee YJ, Ko YG, Ha J, Lee M. Berberine inhibits human colon cancer cell migration via AMP-activated protein kinase-mediated downregulation of integrin β1 signaling. *Biochem Biophys Res Commun* 2012, 426: 461–467
9. Chen X, Dang TTT, Facchini PJ. Noscapine comes of age. *Phytochemistry* 2015, 111: 7–13
10. Kumar S, Sagar B, Gaur A, Shukla S, Pandey E, Gulati S. Insight into the tubulin-targeted anticancer potential of noscapine and its structural analogs. *Anticancer Agents Med Chem* 2023, 23: 624–641
11. Marchant JS, Harding WW, Chan JD. Structure-activity profiling of alkaloid natural product pharmacophores against a *Schistosoma serotonin* receptor. *Int J Parasitology-Drugs Drug Resist* 2018, 8: 550–558
12. Pyne ME, Kevvai K, Grewal PS, Narcross L, Choi B, Bourgeois L, Dueber JE, ***et al.*** A yeast platform for high-level synthesis of tetrahydroisoquinoline alkaloids. *Nat Commun* 2020, 11: 3337
13. Jiao X, Fu X, Li Q, Bu J, Liu X, Savolainen O, Huang L, ***et al.*** *De novo* production of protoberberine and benzophenanthridine alkaloids through metabolic engineering of yeast. *Nat Commun* 2024, 15: 8759
14. Liu T, Gou Y, Zhang B, Gao R, Dong C, Qi M, Jiang L, ***et al.*** Construction of ajmalicine and sanguinarine *de novo* biosynthetic pathways using stable integration sites in yeast. *Biotech & Bioeng* 2022, 119: 1314–1326
15. Gou Y, Li D, Zhao M, Li M, Zhang J, Zhou Y, Xiao F, ***et al.*** Intein-mediated temperature control for complete biosynthesis of sanguinarine and its halogenated derivatives in yeast. *Nat Commun* 2024, 15: 5238
16. Hagel JM, Facchini PJ. Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world. *Plant Cell Physiol* 2013, 54: 647–672
17. Samanani N, Facchini PJ. Purification and characterization of norcoclaurine synthase. *J Biol Chem* 2002, 277: 33878–33883
18. Samanani N, Liscombe DK, Facchini PJ. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. *Plant J* 2004, 40: 302–313
19. Lee EJ, Facchini P. Norcoclaurine synthase is a member of the pathogenesis-related 10/Bet v1 protein family. *Plant Cell* 2010, 22: 3489–3503
20. Sato F, Tsujita T, Katagiri Y, Yoshida S, Yamada Y. Purification and characterization of *S*-adenosyl-l-methionine:norcoclaurine 6-*O*-methyltransferase from cultured *Coptis japonica* cells. *Eur J Biochem* 1994, 225: 125–131
21. Choi KB, Morishige T, Shitan N, Yazaki K, Sato F. Molecular cloning and characterization of coclaurinen-methyltransferase from cultured cells of *Coptis japonica*. *J Biol Chem* 2002, 277: 830–835
22. Pauli HH, Kutchan TM. 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. *Plant J* 1998, 13: 793–801
23. Morishige T, Tsujita T, Yamada Y, Sato F. Molecular characterization of the S-adenosyl-L-methionine: 3′-hydroxy-*N*-methylcoclaurine 4′-O-methyltransferase involved in isoquinoline alkaloid biosynthesis in *Coptis japonica*. *J Biol Chem* 2000: 23398–23405
24. Nguyen TD, Dang TTT. Cytochrome P450 enzymes as key drivers of alkaloid chemical diversification in plants. *Front Plant Sci* 2021, 12: 682181
25. Huang FC, Kutchan TM. Distribution of morphinan and benzo[c] phenanthridine alkaloid gene transcript accumulation in *Papaver somniferum*. *Phytochemistry* 2000, 53: 555–564
26. Liu X, Bu J, Ma Y, Chen Y, Li Q, Jiao X, Hu Z, ***et al.*** Functional characterization of (S)–N-methylcoclaurine 3′-hydroxylase (NMCH) involved in the biosynthesis of benzylisoquinoline alkaloids in *Corydalis yanhusuo*. *Plant Physiol Biochem* 2021, 168: 507–515
27. An Z, Gao R, Chen S, Tian Y, Li Q, Tian L, Zhang W, ***et al.*** Lineage-specific CYP80 expansion and benzylisoquinoline alkaloid diversity in early-diverging eudicots. *Adv Sci* 2024, 11: 2309990
28. Hsieh CL, Yu CC, Huang YL, Chung KF. *Mahonia* vs. *Berberis* unloaded: generic delimitation and infrafamilial classification of Berberidaceae based on plastid phylogenomics. *Front Plant Sci* 2021, 12: 720171
29. Srivastava S, Srivastava M, Misra A, Pandey G, Rawat A. A review on biological and chemical diversity in *Berberis* (Berberidaceae). *EXCLI J* 2015: 247–267
30. Rahimi-Madiseh M, Lorigoini Z, Zamani-Gharaghoshi H, Rafieian-Kopaei M. *Berberis vulgaris*: specifications and traditional uses. *Iran J Basic Med Sci* 2017: 569–587
31. Alamgeer, Uttra AM, Hasan UH. Anti-arthritic activity of aqueous-methanolic extract and various fractions of *Berberis orthobotrys Bien ex Aitch*. *BMC Complement Altern Med* 2017, 17: 371
32. Jo A, Yoo HJ, Lee M. Robustaflavone isolated from *Nandina domestica* using bioactivity-guided fractionation downregulates inflammatory mediators. *Molecules* 2019, 24: 1789
33. Yang S, Cao X, Yu W, Li S, Zhou YJ. Efficient targeted mutation of genomic essential genes in yeast *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2020, 104: 3037–3047
34. Raghavan V, Kraft L, Mesny F, Rigerte L. A simple guide to *de novo* transcriptome assembly and annotation. *Brief BioInf* 2022, 23: bbab563
35. Brown J, Pirrung M, McCue LA, Wren J. FQC Dashboard: Integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics* 2017, 33: 3137–3139
36. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, ***et al.*** Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 2011, 29: 644–652
37. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006, 22: 1658–1659
38. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM, Kelley J. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and

viral genomes. *Mol Biol Evol* 2021, 38: 4647–4654
39. Hung JH, Weng Z. Sequence alignment and homology search with BLAST and clustalW. *Cold Spring Harb Protoc* 2016, 2016: pdb.prot093088
40. Tamura K, Stecher G, Kumar S, Battistuzzi FU. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol* 2021, 38: 3022–3027
41. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, ***et al.*** Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 2010, 28: 511–515
42. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol Plant* 2020, 13: 1194–1202
43. Zhou YJ, Gao W, Rong Q, Jin G, Chu H, Liu W, Yang W, ***et al.*** Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J Am Chem Soc* 2012, 134: 3234–3241
44. Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* 2007, 2: 31–34
45. Ali S, Badshah G, Da Ros Montes D'Oca C, Ramos Campos F, Nagata N, Khan A, de Fátima Costa Santos M, ***et al.*** High-resolution magic angle spinning (HR-MAS) NMR-based fingerprints determination in the medicinal plant *Berberis laurina*. *Molecules* 2020, 25: 3647
46. Qin Z, Liao D, Chen Y, Zhang C, An R, Zeng Q, Li X. A widely metabolomic analysis revealed metabolic alterations of *Epimedium Pubescens* leaves at different growth stages. *Molecules* 2020, 25: 137
47. Park NI, Roy NS, Park Y, Choi BS, Jeon MJ, Oh JY, Kim BY, ***et al.*** Isolation and characterization of the genes involved in the berberine synthesis pathway in asian blue cohosh, *Caulophyllum robustum*. *Plants* 2023, 12: 1483
48. Shah SBA, Kong L, Dong Y, Fu Y, Tian Y, Du Y, Xu Z, ***et al.*** Selection and functional identification of berberine biosynthetic genes in *Berberis amurensis*. Medicinal Plant Biology 2025, 4: e024
49. Kraus PF, Kutchan TM. Molecular cloning and heterologous expression of a cDNA encoding berbamunine synthase, a C-O phenol-coupling cytochrome P450 from the higher plant *Berberis stolonifera*. *Proc Natl Acad Sci USA* 1995, 92: 2071–2075
50. Roy NS, Park NI, Kim NS, Park Y, Kim BY, Kim YD, Yu JK, ***et al.*** Comparative transcriptomics for genes related to berberine and berbamine biosynthesis in berberidaceae. *Plants* 2022, 11: 2676
51. Roy NS, Choi IY, Um T, Jeon MJ, Kim BY, Kim YD, Yu JK, ***et al.*** Gene expression and isoform identification of Pacbio full-length cDNA sequences for berberine biosynthesis in *Berberis koreana*. *Plants* 2021, 10: 1314
52. Li Q, Jiao X, Li X, Shi W, Ma Y, Tan X, Gan J, ***et al.*** Identification of the cytochrome P450s responsible for the biosynthesis of two types of aporphine alkaloids and their *de novo* biosynthesis in yeast. *J Integr Plant Biol* 2024, 66: 1703–1717
53. Hao C, Yu Y, Liu Y, Liu A, Chen S. The CYP80A and CYP80G are involved in the biosynthesis of benzylisoquinoline alkaloids in the sacred lotus (*Nelumbo nucifera*). *Int J Mol Sci* 2024, 25: 702
54. Meng F, Zhang S, Su J, Zhu B, Pan X, Qiu X, Cui X, ***et al.*** Characterization of two CYP80 enzymes provides insights into aporphine alkaloid skeleton formation in *Aristolochia contorta*. *Plant J* 2024, 118: 1439–1454
55. Shi W, Li Q, Li X, Luo L, Gan J, Ma Y, Wang J, ***et al.*** Transcriptome analysis of *Stephania yunnanensis* and functional validation of CYP80s involved in benzylisoquinoline alkaloid biosynthesis. *Molecules* 2025, 30: 259