Synthetic and Systems Biotechnology 8 (2023) 437–444

Contents lists available at ScienceDirect

Synthetic and Systems Biotechnology

journal homepage: www.keaipublishing.com/en/journals/synthetic-and-systems-biotechnology

Original Research Article

Functional characterization of triterpene synthases in *Cibotium barometz*

Zhongju Ji [1], Baolian Fan [1], Yidu Chen, Jingyang Yue, Jiabo Chen, Rongrong Zhang, Yi Tong, Zhongqiu Liu, Jincai Liang [**], Lixin Duan [*]

Guangdong Provincial Key Laboratory of Translational Cancer Research of Chinese Medicines, Joint International Research Laboratory of Translational Cancer Research of Chinese Medicines, International Institute for Translational Chinese Medicine, School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China

ARTICLE INFO

Keywords:
2,3-oxidosqualene cyclase
Squalene cyclase
Cibotium barometz
Triterpene biosynthesis

ABSTRACT

Cibotium barometz (Linn.) J. Sm., a tree fern in the Dicksoniaceae family, is an economically important industrial exported plant in China and widely used in Traditional Chinese Medicine. *C. barometz* produces a range of bioactive triterpenes and their metabolites. However, the biosynthetic pathway of triterpenes in *C. barometz* remains unknown. To clarify the origin of diverse triterpenes in *C. barometz*, we conducted *de novo* transcriptome sequencing and analysis of *C. barometz* rhizomes and leaves to identify the candidate genes involved in *C. barometz* triterpene biosynthesis. Three *C. barometz* triterpene synthases (CbTSs) candidate genes were obtained. All of them were highly expressed in *C. barometz* rhizomes, consisting of the accumulation pattern of triterpenes in *C. barometz*. To characterize the function of these CbTSs, we constructed a squalene- and oxidosqualene-overproducing yeast chassis by overexpressing all the enzymes in the MVA pathway under the control of GAL-regulated promoter and disrupted the GAL80 gene in *Saccharomyces cerevisiae* simultaneously. Heterologous expressing CbTS1, CbTS2, and CbTS3 in engineering yeast strain produced cycloartenol, dammaradiene, and diploptene, respectively. Phylogenetic analysis revealed that CbTS1 belongs to oxidosqualene cyclase, while CbTS2 and CbTS3 belong to squalene cyclase. These results decipher enzymatic mechanisms underlying the origin of diverse triterpene in *C. barometz*.

1. Introduction

Cibotium barometz (Linn.) J. Sm. (a species of fern in the Dicksoniaceae family) is a precious medical plant named “gou-ji” in Chinese, commonly known as the “golden thread”. As a tropical tree fern, *C. barometz* has a high ornamental quality, making it an economically significant industrial exported plant in China [1]. In addition, *C. barometz* has also been utilized as a medicinal herb by various ethnic groups in China and Vietnam for treating lumbao, sciatica, rheumatism, polyuria, leukorrhea, and limb pain [2]. The rhizome extracts of *C. barometz* exhibit various pharmacological activities, including anti-osteoporosis, antioxidant activity, anti-inflammatory, and stimulating chondrocyte proliferation [3]. Another medicinal part of *C. barometz* is the hairs covered on the prostrate rhizome. It has been used to stop bleeding, and its extracts also possess gastro protective effects on ethanol-induced gastric ulcers in Sprague-Dawley rats [4]. Due to the deterioration of the ecological environment and great demand, the wild resources of *C. barometz* have significantly been reduced. *C. barometz* has been listed in the category II of national essential protected wild plants in China.

Plants can produce an abundance of phytochemicals, many of these unique secondary metabolites with potent bioactivities. The main active components of *C. barometz* are anthraquinones, flavonoids, phenolics, tannins, phytosterols, and triterpenoids [5]. Triterpenes are isoprenoidal phytochemicals that occur overall in the plant kingdom, ranging from ferns to higher plants (gymnosperms and angiosperms). Triterpenes and their metabolites, such as triterpenoids and triterpenoid saponins, play essential roles in plant physiology, especially in plant defense against invading phytopathogens and pests. Notably, plant triterpene and its derivatives also play crucial roles in the pharmaceutical

Peer review under responsibility of KeAi Communications Co., Ltd.
* Corresponding author.
** Corresponding author.
E-mail addresses: 530547204@qq.com (J. Liang), nlizn@gzucm.edu.cn (L. Duan).
[1] These authors contributed equally to this article.

https://doi.org/10.1016/j.synbio.2023.06.005
Received 7 March 2023; Received in revised form 12 June 2023; Accepted 16 June 2023
Available online 25 June 2023

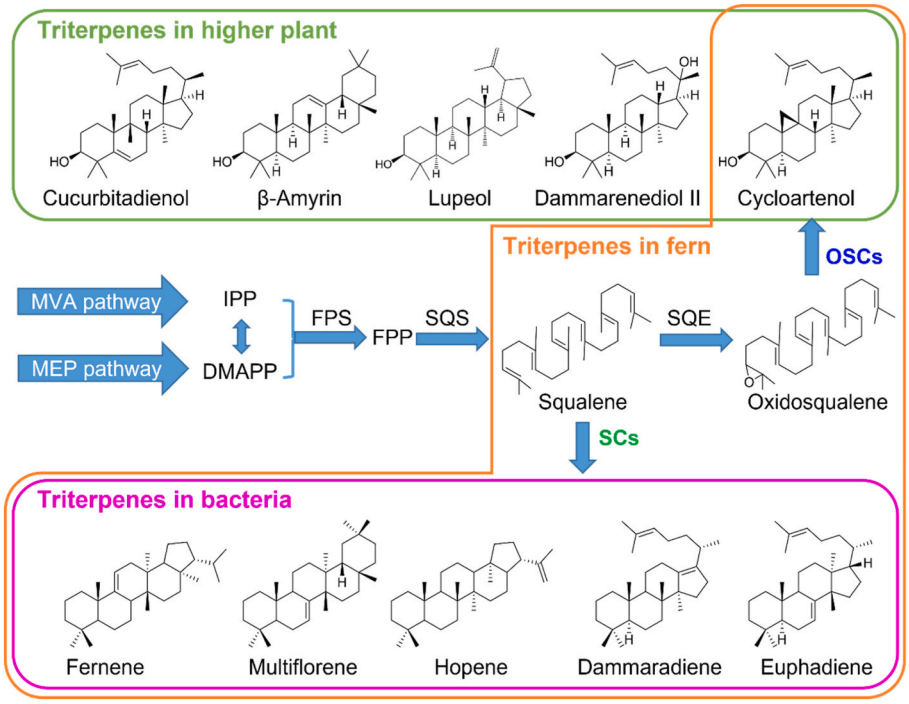

Fig. 1. Biosynthetic pathway of triterpenes in ferns, higher plants, and bacteria.

industry because of their valuable pharmacological properties. Intriguingly, ferns include both oxygen-containing sterol triterpenoids and oxygen-free hopanoids, which is rarely found in higher plants [6]. Hopanoids are essential triterpenoid lipids to regulate the membrane fluidity of prokaryotic cells like bacteria [7,8]. Sterol triterpenoids were the components in plant cells particularly. These mixture types of triterpenes play an essential role in the chemotaxonomy of ferns.

In plants, triterpene biosynthesis arises from the cytosolic mevalonate pathway (MVA pathway) and the plastidial methylerythritol 4-phosphate pathway (MEP pathway). The MVA pathway converts acetyl-CoA into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), while the MEP pathway converts pyruvate to IPP and DMAPP. IPP and DMAPP were then catalyzed into geranyl diphosphate, farnesyl diphosphate, squalene, and 2,3-oxidosqualene by various enzymes. In higher plants, the cyclic triterpene skeletons are commonly generated from 2,3-oxidosqualene catalyzed by oxidosqualene cyclases (OSCs) [9]. This process consists of substrate binding and folding, the protonation and opening of epoxide to initiate the polyene cyclization and rearrangement, and deprotonation or water capture to terminate the reaction. The common cyclization products of oxidosqualene cyclases, such as cycloartenol, lupeol, and beta-amyrin, are characteristic of containing a beta-hydroxyl group at C-3 [10]. In contrast, with a few exceptions, ferns generally produce triterpene from squalene instead of 2,3-oxidosqualene catalyzed by squalene cyclase (SCs), and the resulting cyclization products, like diploptene, fernane, and dammaradiene, lack the hydroxyl group at C-3 [11]. Both squalene cyclase and oxidosqualene cyclase contribute to the structural diversity of plant triterpenes (Fig. 1). In Ferns, bacterial-type SCs and higher plant-type OSCs are coexistence [12,13]. Triterpene synthases catalyze the most important step, directly deciding triterpenoids and hopanoids skeleton. To decipher the unusual hopanoids skeleton biosynthesis pathway, we use *C. barometz* as the object to discover the intermediary of hopanoids in *C. barometz*.

In these years, the production of valuable triterpene compounds applying synthetic biology has emerged as a potential way. *Saccharomyces cerevisiae* is a model eukaryotic cell that not only has the characteristics of fast reproduction, easy culture, and convenient genetic engineering operation but also has the functions of protein processing and post-translational modification. Thus, *S. cerevisiae* has unique advantages as a heterologous synthetic host for natural products. Hu et al. refactored the whole biosynthetic pathways of 3*β*-*O*-Glc-dammarenediol and 20*S*–*O*-Glc-dammarenediol by multistep metabolic engineering strategies in *S. cerevisiae*, the resulting strain could produce 2.4 g/L 3*β*-*O*-Glc-dammarenediol and 5.6 g/L 20*S*–*O*-Glc-dammarenediol [14]. Yu et al. obtained over 1 g/L of α-amyrin in engineered *S. cerevisiae* by remodeling αAS MdOSC1 and expanding the storage pool [15]. Wang et al. constructed a protopanaxadiol-producing chassis and established a series of cell factories resulting high yield of 2.25 g/L ginsenoside Rh2 in 10 L fed-batch fermentation [16]. In the present study, to decipher the triterpene cyclases underlying *C. barometz* triterpene biosynthesis, we conducted *de novo* transcriptome sequencing and analysis of rhizomes and leaves of *C. barometz* to screen the candidate genes involved in triterpene biosynthesis in this fern. Three *C. barometz* triterpene synthase (CbTSs) candidate genes were obtained by combining bioinformatic analysis and targeted metabolite analysis. Their biochemical function was evaluated by heterologous expression in an engineering yeast chassis. Additionally, the phylogenetic evolution of CbTSs was discussed as well. These results fill gaps in our knowledge of *C. barometz* triterpene biosynthesis.

2. Materials and methods

2.1. Chemicals, microbial strains, and plant material

Squalene and cycloartenol were purchased from Aladdin (Shanghai, China); N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA) was purchased from Sigma-Aldrich Co., Ltd. (Shanghai, China); Hexane, methanol, ethanol, and KOH, were purchased from Xilong Scientific Co., Ltd. (Guangzhou, China).

Escherichia coli strain Stbl3 was purchased from ThermoFisher (Shanghai, China); *S. cerevisiae* strain CEN.PK2-1D was purchased from EUROSCARF (Oberursel, Germany). *Cibotium barometz* (Linn.) J. Sm. (Dicksoniaceae) at the two-year-old stage was collected from Yaowang Hill of Guangzhou University of Chinese Medicine (Guangzhou, China).

The specie identity was confirmed by Dr. Yi Tong, co-author of this paper.

2.2. RNA isolation, RNA library construction, and sequencing

Rhizomes and leaves were sampled from two-year-old *C. barometz* and frozen in liquid nitrogen immediately. RNA isolation was carried out using the RNAprep Pure Plant Plus Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's instructions. The RNA quality was checked with both NanoDrop 2000C spectrophotometer and electrophoresis. The values of RNA Integrity Number were ensured to be greater than 7.5 assessed by the BioAnalyzer 2100 (Agilent Technologies). VAHTS ® Universal V8 RNA-seq Library Prep Kit for Illumina was to use to prepare the RNA library. After mRNA enrichment using VAHTS mRNA capture beads, the first and second strand cDNA was synthesized and end-repaired with dA-tailing. Finally, the resulting cDNA library sequencing was performed on the Illumina HiSeq 2000 platform.

2.3. Transcriptome data analysis and assembly

Raw reads were cleaned up by removing the adaptor sequences, low-quality sequences that Q values less than 20 in more than 20% of the bases and ambiguous sequences with more than 10% unknown sequences. The resulting clean read data were *de novo* assembly using Trinity software, and the unigenes were obtained after removing redundancy in the TGI Clustering tools. All the unigenes were blasted to the public databases for annotation, including Swiss-Prot protein database, NCBI non-redundant protein database (Nr), Kyoto Encyclopedia of Genes and Genome (KEGG), Clusters of Orthologous Groups database (COG), and Gene Ontology database (GO). CDS prediction and translation of unigenes were based on Blastx and ESTScan search tools.

2.4. KEGG pathway enrichment and GO analysis

KEGG pathway enrichment was established by mapping the unigenes to the KEGG classifications according to KO number, including "organismal systems", "environmental information processing", genetic information processing", cellular processes", and "metabolism". For GO analysis, the unigenes were annotated in the Blast2GO program and then assigned with the Gene Ontology annotations in the WEGO software.

2.5. Identification and sequence alignment of CbTSs

For identification of CbTSs candidate genes from *C. barometz* transcriptome assembly, ten representatives of squalene cyclase and oxidosqualene cyclase that were previously reported were employed as query sequences to blast against the *C. barometz* transcriptome using the tblastn search tool of the BioEidt program. The resulting candidate genes were subsequently aligned with squalene cyclase and oxidosqualene cyclase reported in a previous study to check the integrality of all the CbTSs conserved motifs. Multiple protein sequence alignment was conducted on MEGA7 with the ClustalW algorithm using default settings.

2.6. Analysis of CbTSs expression levels and triterpene contents of C. barometz

Total RNA of rhizome and leave samples were extracted using RNAprep Pure Plant Plus Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's instructions. After RNA isolation, first-stand cDNA synthesis was conducted using the PrimeScriptTM ^{II} 1st Strand cDNA Synthesis kit (Takara, Beijing, China). Gene-specific primers of all the *CbTSs* and internal reference genes (the housekeeping gene ubiquitin) were designed in Vector NIT Advance 11.5.1 software. For qRT-PCR experiments, a 20 μL qPCR reaction mix was prepared, which contained 10 ng of template cDNA, 200 nM of each gene-specific primer, and 10 μL of 2 × SYBR Green Master Mix Reagent (Takara). qRT-PCR analysis was performed on Roche Light cycler 480 real-time PCR system in triplicate following the reaction conditions: 95 °C for 30 s, 42 cycles of 95 °C for 10 s, and 60 °C for 30 s. The data were processed by the $2^{-\Delta\Delta CT}$ method, and the average data from triplicate repetitive.

For analysis the triterpene contents of *C. barometz*, roots and leaves were sampled and ground into a fine powder with liquid nitrogen. About 0.2 g of the powder was weighed and extracted with 1.5 mL of boiled 50% (vol/vol) ethanol and 20% (wt/vol) potassium hydroxide, and the supernatant was extracted three times with an equal volume hexane. The resulting hexane extracts were then combined and evaporated to obtain the crude extracts. After derivation with 50 μL of N-methyl-N-trimethylsilyl-trifluoroacetamide (MSTFA) at 90 °C for 15 min, the crude extracts were analyzed by gas chromatography mass spectrometry (GC-MS).

GC-MS analysis was conducted on an Agilent 7890B GC instrument equipped with an HP-5MS column and 5977B single quadrupole mass spectrometer. The injector temperature was set to 280 °C, and the oven temperature program was as follows: 80 °C with 2 min hold, first gradient from 80 °C to 280 °C at 25 °C/min, second gradient from 280 °C to 320 °C at 10 °C/min with 3 min hold. The carrier gas was helium, and the flow rate was 1 mL/min. The ionization energy was set at 70 eV, and the mass spectral data range (m/z) of 50–600 was acquired.

2.7. Construction of squalene synthetic pathway in S. cerevisiae, CbTSs cloning, and plasmid construction

To construct the high-yield substrate engineering strain in *S. cerevisiae*, the genomic DNA (gDNA) of *S. cerevisiae* was extracted using the DNAios Kit (Takara Biomedical Technology Co., Ltd., Beijing, China). All genes of *S. cerevisiae* MVA pathway, including *ERG10*, *ERG13*, *tHMGR1* (a truncated version of yeast 3-hydroxy-3-methylglutaryl coenzyme A reductase, consisting of just the catalytic domain), *ERG12*, *ERG8*, *IDI1*, *ERG20*, *ERG19*, and *ERG9*, were amplified from the gDNA of *S. cerevisiae*, and then cloned into the pESC-URA vector (Agilent) between BamHI and NheI restriction site to create the gene expression cassettes under control of the *GAL1* promoter. To integrate all the expression cassettes into *GAL80* gene locus simultaneously, we modularized the integration units by tethering the adjacent expression cassettes to each other via overlap extension PCR. These integration units contain 500 bp homologous region of each adjacent gene in MVA pathway. The gRNA block of *GAL80* was designed using E-CRISP-Version 5.4 [17,18] and synthesized by GenScript (Nanjing, China), and then cloned into pCRCT (Addgene plasmid # 60621; http://n2t.net/addgene:60621; RRID: Addgene_60621) using the Golden Gate assembly method to yield pCRCT-GAL80. DNA fragment and plasmid transformation were carried out using LiAc/SS carrier DNA/PEG method. Briefly, 2 μg plasmid or 500 ng of each fragment was mixed and transformed into 100 μL of *S. cerevisiae* competent cells. After transformation, cells were spread on synthetic complete media plates and incubated at 30 °C for 3–4 days. The resulting single clones of each transformation were picked up to check with colony PCR and DNA sequencing. These integrated genes improve the production of 2,3-oxidosqualene, the common precursor for triterpene biosynthesis. This high-yielding oxidosqualene-producing engineering yeast strain was named p24.

The full-length open reading frames of all *CbTSs* candidate genes were amplified from a *C. barometz* rhizome cDNA library. The resulting amplicons were cloned into the pESC-URA vector between BamHI and SmaI restriction site to yield pESC-URA-P_{Gal1}-CBTSs-T_{CYC1}. To construct the plasmids for multiple-copy integration into *S. cerevisiae* chromosomes, the P_{Gal1}-CBTSs-T_{CYC1} expression cassettes of pESC-URA- P_{Gal1}-CBTSs-T_{CYC1} were subcloned into pCfb2988, which was a gift from Irina Borodina & Jerome Maury (Addgene plasmid # 63638; http://n2t.net/addgene:63638; RRID: Addgene_63638) between the NheI and HindIII restriction site to afford pCfb2988-P_{Gal1}-CBTSs-T_{CYC1}. All DNA

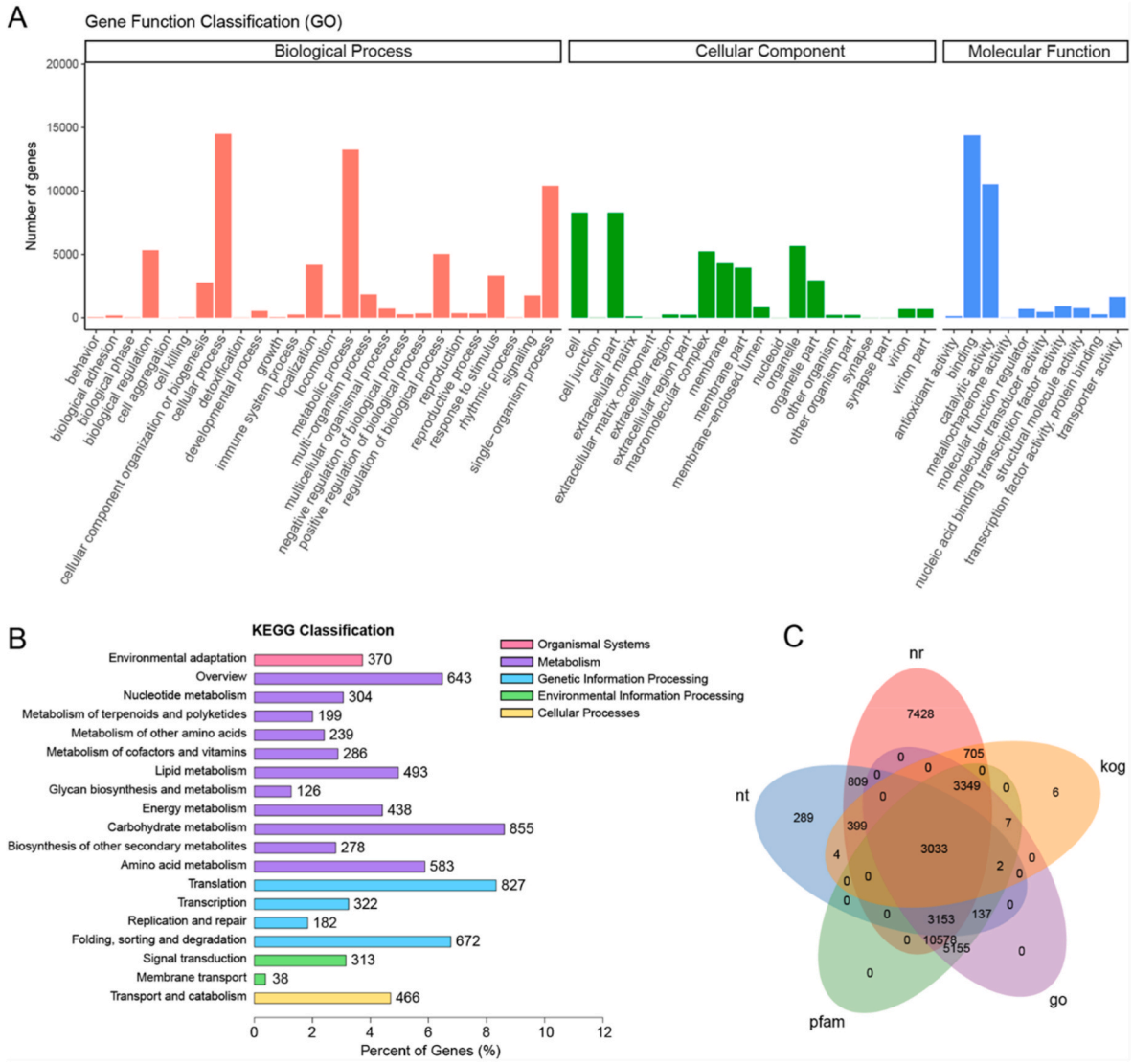

Fig. 2. Gene function classification of *C. barometz* (A) GO analysis of unigenes. (B) KEGG analysis of unigenes. (C) Venn diagram of the unigenes annotation.

fragments were amplified using Phanta MAX Super-Fidelity DNA Polymerase (Vazyme Biotech, Inc.). All plasmids were constructed using ClonExpress II one-step cloning kit (Vazyme Biotech) and verified by complete sequencing. Primers and plasmids used in cloning and construction are listed in Table S1. The sequences of *CbTS1*, *CbTS2*, *CbTS3*, and *CbTS4* have been deposited in GenBank under accessions numbers OP326758, OP326759, OP326760, and OP326761, respectively.

2.8. Media and growth conditions of S. cerevisiae strains

For shake-flask culture, the single clones of recombinant strains were firstly inoculated in 3 mL of synthetic dropout (SD) media without corresponding amino acids (6.7 g/L of yeast nitrogen base without amino acids, 20 g/L of dextrose, and 1.3 g/L of amino acid dropout powder) at 220 rpm 30 °C for overnight. These seed cultures were then transferred into 250 mL flasks containing 50 mL SD media for additional growth until 72 h.

For bioreactor fed-batch culture, the strains were inoculated in 5 mL of SD media at 220 rpm 30 °C for 24 h and then transferred into 100 mL of seed YPD media (5% glucose, 1% yeast extract, 3% peptone, 0.8% KH_2PO_4, and 0.6% $MgSO_4{\cdot}7H_2O$) in a 500 mL flask growing for an addition 24 h. The resulting cultures were inoculated into 2 L of batch YPD media (4% glucose, 1% galactose, 1% yeast extract, 3% peptone, 0.8% KH_2PO_4, and 0.6% $MgSO_4{\cdot}7H_2O$.) in a 7 L bioreactor (Baoxing, Shanghai, China). The galactose fed-batch media contains galactose (400 g/L), KH_2PO_4 (9 g/L), $MgSO_4{\cdot}7H_2O$ (5.12 g/L), K_2SO_4 (3.5 g/L), Na_2SO_4 (0.28 g/L), 12 mL/L vitamin solution (1 g/L calcium pantothenate, 0.05 g/L biotin, 25 g/L inositol, 1 g/L thiamine HCl, 1 g/L nicotinic acid, 1 g/L pyridoxal HCl and 0.2 g/L p-aminobenzoic acid), and 10 mL/L trace metals solution (0.47 g/L $CoCl_2{\cdot}6H_2O$, 5.75 g/L $ZnSO_4{\cdot}7H_2O$, 0.48 g/L $Na_2MoO_4{\cdot}2H_2O$, 0.32 g/L $MnCl_2{\cdot}4H_2O$, 2.8 g/L $FeSO_4{\cdot}7H_2O$, 2.9 g/L $CaCl_2{\cdot}2H_2O$, and 80 mL/L 0.5 M EDTA, pH 8.0). The galactose fed fermentation processes were automatically triggered when the carbon sources were completely depleted. The dissolved oxygen was maintained at 40%, and the pH was maintained at 5.5 by adding 10 M ammonium hydroxide automatically.

2.9. Metabolite extraction and content analysis

For metabolite analysis of recombinant yeast strains, 1 mL of flask cultures or fed-batch cultures were collected by centrifugation. The pellet was extracted with 1 mL of hot 50% (vol/vol) ethanol and 20% (wt/vol) potassium hydroxide for 5 min. The supernatant was extracted with 1 mL of hexane. Pipetted 100 μL of hexane extracts into a glass tube to evaporate and derivate with MSTFA at 90 °C for 15 min. GC-MS analyzed the resulting samples, and the method was equal to *2.6*.

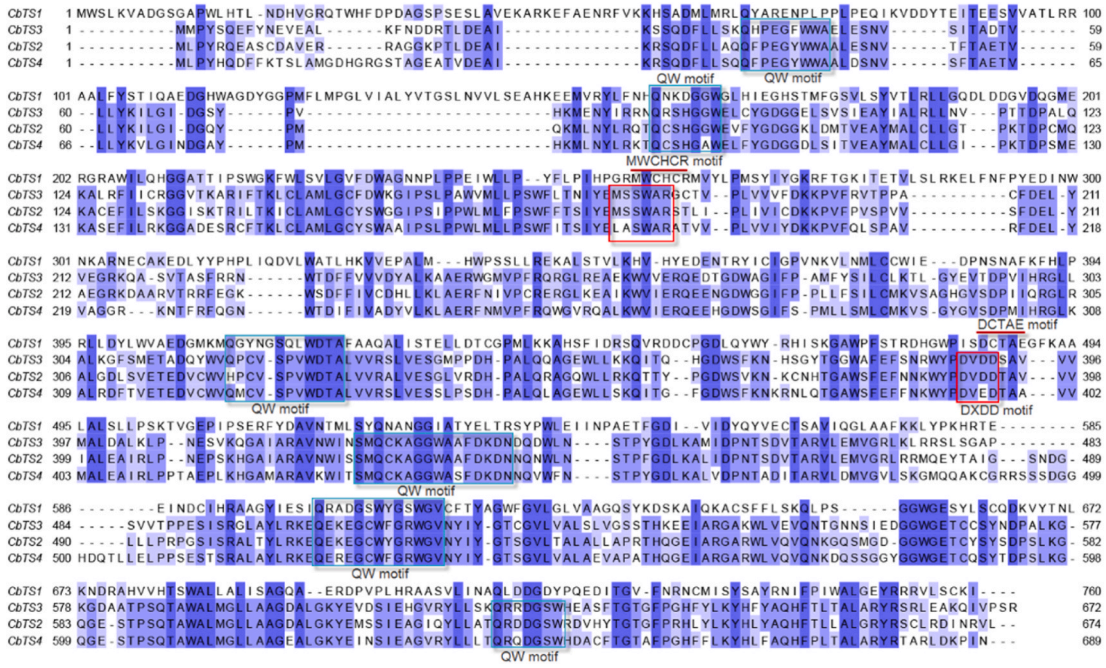

Fig. 3. Amino acid alignment and motif features for CbTS1~4.

2.10. Compound isolation and structural identification

After fed-batch fermentation (5L), yeast cells of the relevant recombinant strains were collected and extracted as described above. The hexane extracts were combined and dried under vacuum. The crude extracts were loaded into silica gel column (50 cm × 3 cm, 300–400 mesh particle size) and eluted with a hexane/ethyl acetate gradient solvent system (50:1 to 8:1). The resulting fractions were analyzed by GC-MS and further purified by preparative thin layer chromatography and/or reversed-phase high performance liquid chromatography (RP-HPLC). Purified products were dissolved in deuterated chloroform, and NMR analysis was conducted on a Bruker Avance III HD 400 MHz instrument using solvent signals as the standards to determine chemical shifts.

2.11. Phylogenetic analysis

All CbTSs and other charactered plant triterpene synthases were used for phylogenetic analysis. Multiple protein sequence alignment was conducted on MEGA7 with the ClustalW algorithm using the default settings. A neighbor-joining (NJ) tree was then constructed using the JTT matrix-based model, and the tree was validated by the bootstrap method with 1000 replicates.

3. Results and discussion

3.1. Transcriptome sequencing and screening of the CbTSs sequences

To screen the candidate unigenes and excavate the biosynthetic pathway of triterpenes in *C. barometz*, the roots and leaves of this plant were conducted for the transcriptome analysis(Table S2). In total, 124,648,860 and 130,382,284 raw sequencing reads were generated from cDNA libraries of rhizome and leaf with Q20 > 98% (Table S3). Totally, 60,634 unigene sequences were aligned to the Nr, Nt, Swissprot, KEGG, COG, and GO databases for annotation. GO analysis was performed (Table S4), and the unigenes were functionally categorized into three categories: “Biological Process”, “Cellular Component” and “Molecular Function”. The KEGG enrichment analysis was scanned based on the KO number using the KEGG database. The metabolic pathways were further categorized into different groups, including “Metabolism of terpenoids and polyketides” (Fig. 2).

Since the protein sequences of plant OSCs and SCs are highly conserved, some charactered plant OSCs and SCs amino acid sequences were used as a query to blast search against the transcriptome data (Fig. S1, Table S5). Six matching unigenes were screened, but only 4 of them were annotated as oxidosqualene cyclases or squalene cyclases. Using primer pairs for the four selected *CbTSs*, four new ORF sequences (*CbTS*1, *CbTS*2, *CbTS*3, and *CbTS*4) were obtained by sequencing the vector pESC-URA-*CbTSs* with the *CbTS* genes. The sequence of CbTS1 (2280 bp) was slightly larger than others (2016–2067 bp). The feature QW motifs [QXXXGXW] responsible for strengthening the enzyme structure and stabilizing its carbocation intermediates differed between CbTS1 and CbTS2-4 (Fig. 3). CbTS1 showed a conserved region DCTAE for the OSCs family, and CbTS2-4 displayed a DXDDTA motif correspondingly. These two motifs were verified to participate into the substrate binding and initiate the polycyclization of 2,3-oxidosqualene and squalene, respectively [19]. In addition, another motif related to the product specificities also was different, the conserved MWCHCR in CbTS1 indicated it could be cycloartenol cyclases [20,21].

3.2. Expression levels of CbTSs and triterpene contents in leaves and roots

To better understand the functional divergence of *CbTSs*, the transcript levels of their encoding genes in leaves and roots of *C. barometz* were assessed (Fig. S2). All the *CbTS* genes were expressed in leaf and root, but the expression levels differed. Among them, the transcript level of *CbTS1* was higher than other *CbTSs* in both leaves and roots. The expression of *CbTS2* and *CbTS3* were equally low in leaves and roots, but the expression in roots was slightly higher than in the leaves.

To test the accumulation patterns of triterpenes in *C. barometz*, the saponified extracts were detected in *C. barometz* by GC-MS (Fig. S3 and Table S6). Various sterols and their derivatives appear in leaves and roots, but triterpenols are rarely found. Two common sterols stigmastanol and lanosterol showed obvious different accumulation patterns.

3.3. Functional characterizations of candidate CbTSs

To functionally characterize these four *CbTS* genes, each ORF

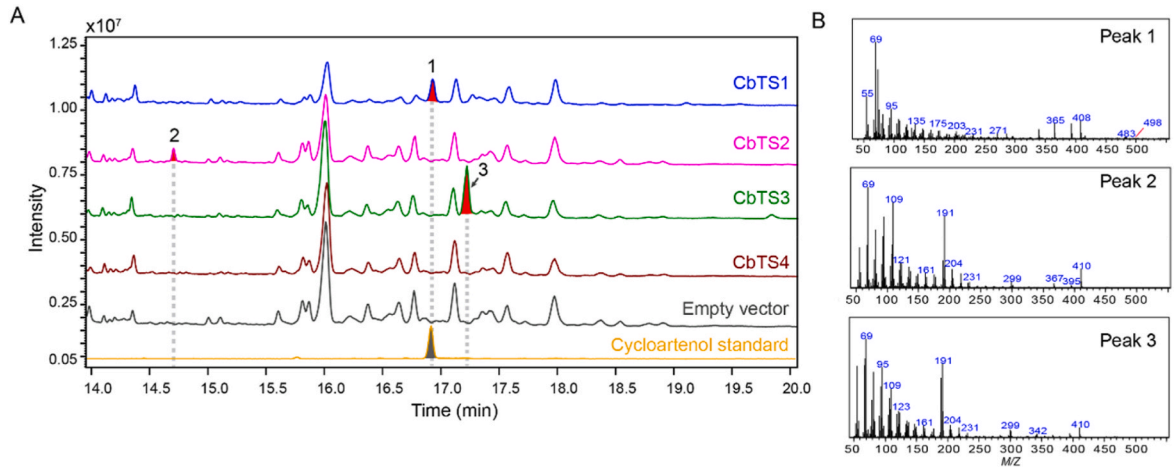

Fig. 4. Functional characterization of CbTSs in engineering yeast chassis. (A) GC analysis of the in vivo reaction products catalyzed by CbTS1∼4. (B) MS spectrum of component 1, 2, and 3.

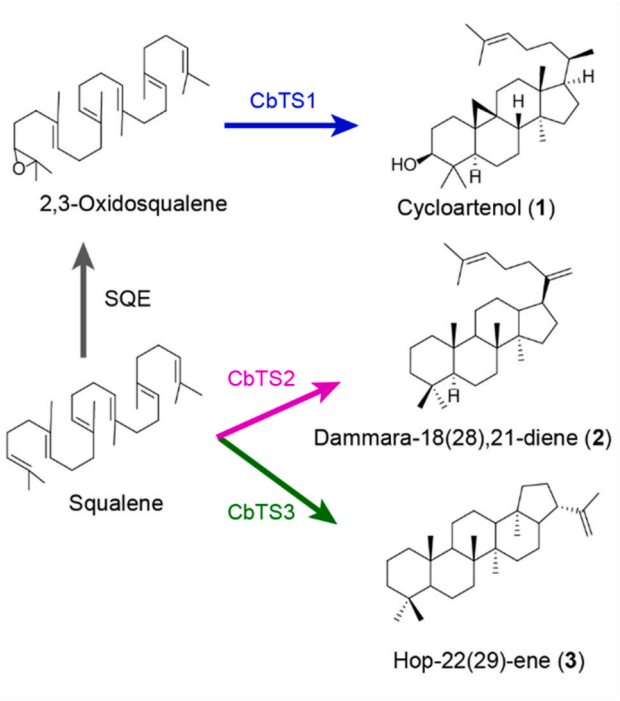

Fig. 5. The synthesis of final products catalyzed by CbTS1, CbTS2 and CbTS3.

sequence was cloned and inserted into a pESC-URA expression vector and then transformed into the engineering yeast strain p24 with the plasmid. The extracts, including the products from transgenic yeasts expressing *CbTSs*, were derivatized and analyzed by GC/MS. The total ion chromatogram (TIC) of CbTS1 showed one different peak1 at 16.9 min compared with the empty vector (Fig. 4A). The mass spectra presented a typical molecular ion peak at *m/z* 498 for trimethylsilation (Fig. 4B). The retention time and the mass fragmentation patterns of peak1 were well matched with cycloartenol standard (Fig. S4). So CbTS1 was characterized as cycloartenol synthase, consistent with the predicted function according to its feature motifs (Fig. 3). Cycloartenol yield of 33.03 ± 0.60 mg/L in p24 strain.

The TIC of CbTS2 revealed another new product peak2 at 14.7 min. The mass spectra presented a typical molecular ion peak at *m/z* 410, the molecular ion peak for hopane triterpenes. Because of the lack of a hydroxy at C-3, the molecules cannot be derivatized by silylating agents (Fig. 4B). Other intense ions at *m/z* 191, *m/z* 109, and *m/z* 69 are well known as characteristic fragmentation ions of the dammarane skeleton with two double bonds in its side-chain [22]. Hopanes and migrated hopanes are widely distributed in most ferns. Various hopane triterpenes have been identified in nature, such as pentacyclic hopanes, oleananes, tetracyclic dammaranes, tricyclic malabaricanes, and bicyclic polypodanes [23]. In transgenic yeast expressing CbTS3, a single product peak3 was obtained at 17.3 min. The spectra of peak3 showed similar mass fragmentation patterns with peak2. The molecular ion peak *m/z* 410 indicated peak3 was another hopane (Fig. 4B). Unfortunately, no product peak was found in the TIC of CbTS4, even though we extracted the characteristic ions such as *m/z* 498 or 410. It is possibly dues to the lack of codon optimization which prevented its translation and expression in yeast.

3.4. Structure identification of products for CbTSs in S. cerevisiae

To identify the structures of CbTS2 and CbTS3 products, the extract of a fed-batch fermentation (5L) was isolated by silica column and RP-HPLC. The purified products were analyzed in deuterated chloroform by nuclear magnetic resonance (NMR) spectroscopy—no hydroxyl group signal was presented in the ^{1}H NMR spectrum of the CbTS2 product. In total, five tertiary methyl and two olefinic methyl signals were observed (Fig. S5). The carbon spectrum signal of peak2 was very consistent with reference [24], and the structure of product of CbTS2 was determined as dammara-18(28),21-diene (Fig. S6, Table S7). CbTS2 was characterized as tetracyclic dammara-18(28),21-diene synthase (Fig. 5). Dammara-18(28),21-diene yield of 9.67 ± 0.43 mg/L in p24 strain.

The ^{1}H NMR spectrum of peak3 showed six tertiary methyl group signals and one olefinic methyl signal. Methyl proton signals assigned to H-23, H-24, H-25, H-26, H-27, H-28, and H-30 gave appropriate values for a hopene. Two proton signals appeared as a single peak at δ_H 4.78, the characteristic alkene for hop-22(29)-ene, because the two H-29 have almost the same chemical environment (Fig. S7). The carbon spectrum of peak3 showed 30 carbon signals, including two olefin carbon signals (Fig. S8). The total carbon signals were very consistent with hop-22(29)-ene reference [25] (Table S8). CbTS3 was characterized as pentacyclic hop-22(29)-ene synthase (Fig. 5). Hop-22(29)-ene yield of 35.18 ± 0.56 mg/L in p24 strain.

3.5. Phylogenetic analysis of CbTSs

The phylogenetic tree of triterpene synthases was generated to predict the possible function of CbTS1-4 and reflected a taxonomic

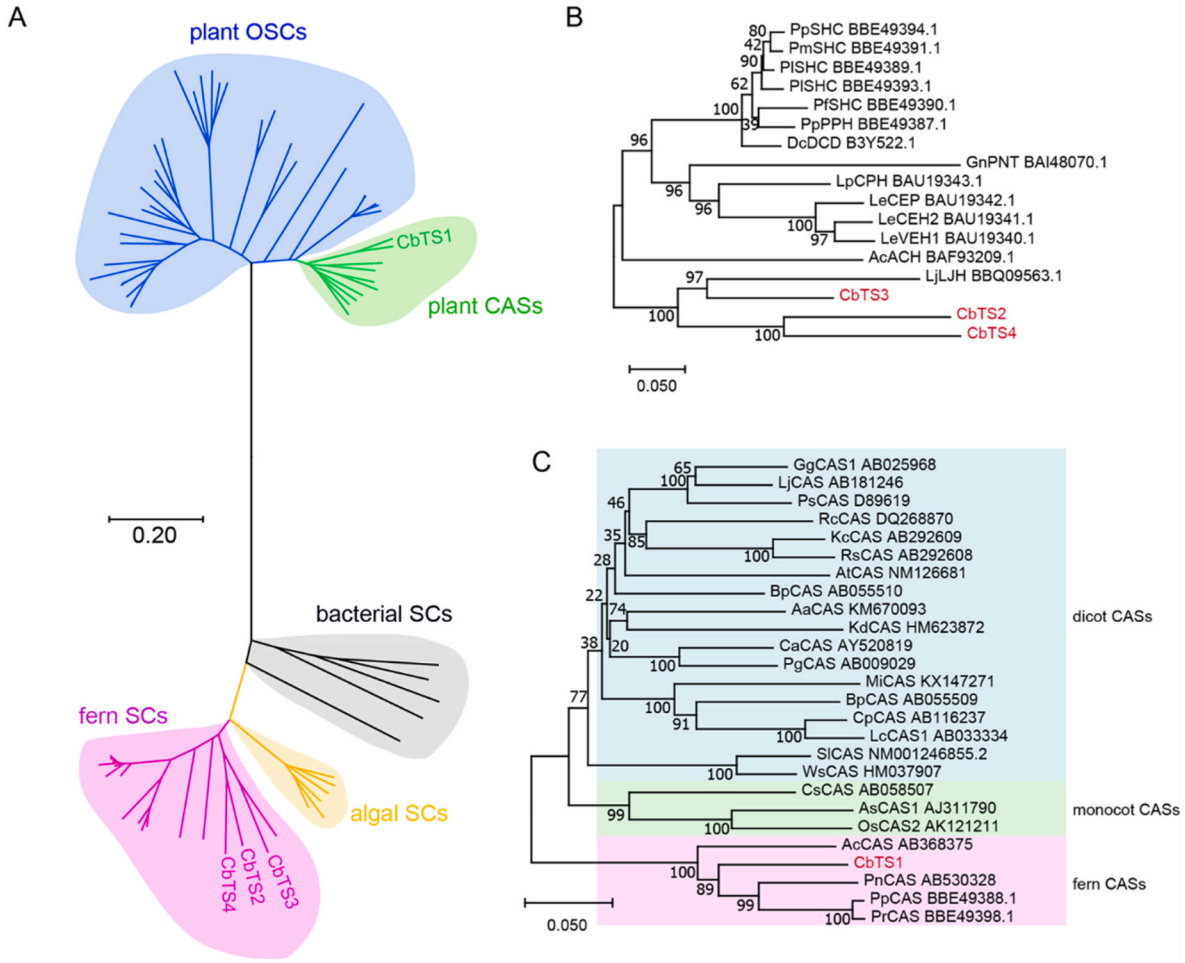

Fig. 6. Phylogenetic analysis of the CbTSs (A) Unrooted neighbor-joining phylogenetic tree of OSC and SC homologs. Plant OSCs and Cycloartenol synthases (CASs) clades are in blue and green, respectively. Fern, algal and bacterial SC clades are in magenta, yellow and gray, respectively. For clarity, the branches have no taxon labels. The tree showing all taxon names is presented in Supplementary Fig. 7. The scale represents 0.2 amino acid substitution per site. (B) Expanded fern SCs clade from the phylogenetic tree in Fig. 6A. Bootstrap values with 1000 replicates are shown at the nodal branches (cutoff value = 50%). (C) Expanded plant CASs clade from the phylogenetic tree in Fig. 6A. Bootstrap values with 1000 replicates are shown at the nodal branches (cutoff value = 50%).

relationship (Fig. 6A). All CbTSs along with other plant triterpene synthases listed in Table S9, were utilized for phylogenetic analysis. Phylogenetic analysis with all taxon names showed in Fig. S9.

In the phylogenetic tree, OSCs and SCs were clearly divided into two branches accompanied by two-type catalytic characteristics. Among the SCs branch, the fern SCs clade was obviously separated from those for bacteria and algal. From the perspective of phylogenetic relationship, the fern SCs are closer to the algal SCs than bacteria SCs. The three CbTSs and *Lygodium japonicum* hopene synthase (LjLJH) formed a new cluster (Fig. 6B). The OSCs branches are basically from plants and were resolved as a monophyletic family separating into CAS and the other OSC clusters. CbTS1 grouped with fern CASs, which clearly separated from those of higher plant origin, and closely related to CASs from *Polypodiodes* (PnCAS) and *Polystichum* (PpCAS and PrCAS) but not to AcCAS (*Adiantum capillus*, Fig. 6C). Since the coexistence of bacterial-type SCs and higher-plant-type OSCs in ferns, ferns SCs was suggested to be acquired by horizontal transfer from bacteria [7]. Interestingly, ferns prefer to evolve various bacterial SCs rather than plant OSCs.

4. Conclusion

Triterpene synthases play a crucial role in the biosynthesis of triterpenes, which are essential compounds found in many organisms. Enzymes for the biosynthesis of triterpenes/saponins in *C. barometz* were not previously identified and biochemically characterized. In this study, we pursued a combined approach of transcriptomics and enzyme biochemical characterization. We identified a suite of triterpene synthases that catalyzed the critical scaffold diversifying reactions in the biosynthesis of squalene, produced cycloartenol (CbTS1), dammara-18 (28),21-diene (CbTS2) and hop-22(29)-ene (CbTS3), respectively. CbTS2 and CbTS3 were first characterized to be responsible for the tetracyclic hopane-type triterpene and pentacyclic hopane-type triterpene from *C. barometz*. These will enrich the plant triterpene cyclase gene diversity and benefit deciphering the unusual hopanoids skeleton biosynthesis pathway. These unique triterpene products highlight the importance of continued exploration and discovery of the diversity of triterpene synthases in different species. This could lead to the development of new natural products with valuable biological activities. To our knowledge, this study presented functional characterization of the triterpene saponin pathway enzymes of *C. barometz* for the first time.

CRediT authorship contribution statement

Zhongju Ji: Conceptualization, Formal analysis, Investigation, Visualization. **Baolian Fan:** Data curation, Formal analysis, Investigation, Writing – original draft. **Yidu Chen:** Visualization, Data curation. **Jingyang Yue:** Formal analysis, Data curation. **Jiabo Chen:** Data curation. **Rongrong Zhang:** Formal analysis, Data curation. **Yi Tong:**

Plant identification. **Zhongqiu Liu:** Funding acquisition, Project administration. **Jincai Liang:** Funding acquisition, Project administration. **Lixin Duan:** Conceptualization, Funding acquisition, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (No. 81874333), the Key Laboratory of Guangdong Drug Administration (2021ZDB03), and the Guangdong Basic and Applied Basic Research Foundation(No. 2020B1515130005).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.synbio.2023.06.005.

References

[1] He J, et al. Therapeutic anabolic and anticatabolic benefits of natural Chinese medicines for the treatment of osteoporosis. Front Pharmacol 2019;10:1344. https://doi.org/10.3389/fphar.2019.01344.
[2] Rape M. Plant biology informs drug discovery. Nat Rev Mol Cell Biol 2014;15(8):501. https://doi.org/10.1038/nrm3842.
[3] Huang D, et al. Two novel polysaccharides from rhizomes of Cibotium barometz promote bone formation via activating the BMP2/SMAD1 signaling pathway in MC3T3-E1 cells. Carbohydr Polym 2020;231:115732. https://doi.org/10.1016/j.carbpol.2019.115732.
[4] Cao H, et al. Phytochemicals from fern species: potential for medicine applications. Phytochemistry Rev 2017;16(3):379–440. https://doi.org/10.1007/s11101-016-9488-7.
[5] Heng YW, et al. Biological activities and phytochemical content of the rhizome hairs of Cibotium barometz (Cibotiaceae). Ind Crop Prod 2020;153:112612. https://doi.org/10.1016/j.indcrop.2020.112612.
[6] Fan Z, et al. Identification of a novel multifunctional oxidosqualene cyclase from Zea mays sheds light on the biosynthetic pathway of three pentacyclic triterpenoids. Synth Syst Biotechnol 2022;7(4):1167–72. https://doi.org/10.1016/j.synbio.2022.08.004.
[7] Belin BJ, et al. Hopanoid lipids: from membranes to plant-bacteria interactions. Nat Rev Microbiol 2018;16(5):304–15. https://doi.org/10.1038/nrmicro.2017.173.
[8] Kannenberg EL, Poralla K. Hopanoid biosynthesis and function in bacteria. Naturwissenschaften 1999;86(4):168–76. https://doi.org/10.1007/s001140050592.
[9] Li Y, et al. Natural products of pentacyclic triterpenoids: from discovery to heterologous biosynthesis. Nat Prod Rep 2022. https://doi.org/10.1039/d2np00063f.
[10] De La Pena R, et al. Complex scaffold remodeling in plant triterpene biosynthesis. Science 2023;379(6630):361–8. https://doi.org/10.1126/science.adf1017.
[11] Shinozaki J, Nakene T, Takano A. Squalene cyclases and cycloartenol synthases from Polystichum polyblepharum and six allied ferns. Molecules 2018;23(8). https://doi.org/10.3390/molecules23081843.
[12] Abe I, et al. Cyclization of (3S)29-Methylidene-2,3-oxidosqualene by bacterial squalene:hopene cyclase: irreversible enzyme inactivation and isolation of an unnatural dammarenoid. J Am Chem Soc 1997;119(46):11333–4. https://doi.org/10.1021/ja972770p.
[13] Shinozaki J, et al. Squalene cyclase and oxidosqualene cyclase from a fern. FEBS Lett 2008;582(2):310–8. https://doi.org/10.1016/j.febslet.2007.12.023.
[14] Hu Z-F, et al. Construction and optimization of microbial cell factories for sustainable production of bioactive dammarenediol-II glucosides. Green Chem 2019;21(12):3286–99. https://doi.org/10.1039/C8GC04066D.
[15] Yu Y, et al. Engineering Saccharomyces cerevisiae for high yield production of alpha-amyrin via synergistic remodeling of alpha-amyrin synthase and expanding the storage pool. Metab Eng 2020;62:72–83. https://doi.org/10.1016/j.ymben.2020.08.010.
[16] Wang P, et al. Synthesizing ginsenoside Rh2 in Saccharomyces cerevisiae cell factory at high-efficiency. Cell Discov 2019;5:5. https://doi.org/10.1038/s41421-018-0075-5.
[17] Heigwer F, Kerr G, Boutros M. E-CRISP: fast CRISPR target site identification. Nat Methods 2014;11(2):122–3. https://doi.org/10.1038/nmeth.2812.
[18] Yu L, et al. Efficient genome editing in Claviceps purpurea using a CRISPR/Cas9 ribonucleoprotein method. Synth Syst Biotechnol 2022;7(2):664–70. https://doi.org/10.1016/j.synbio.2022.09.002.
[19] Hoshino T, Sato T. Squalene-hopene cyclase: catalytic mechanism and substrate recognition. Chem Commun 2002;(4):291–301. https://doi.org/10.1039/B108995C.
[20] Wu S, et al. An unexpected oxidosqualene cyclase active site architecture in the Iris tectorum multifunctional α-amyrin synthase. ACS Catal 2020;10(16):9515–20. https://doi.org/10.1021/acscatal.0c03231.
[21] Kushiro T, et al. Mutational studies on triterpene synthases: engineering lupeol synthase into â-amyrin synthase. J Am Chem Soc 2000;122:6816–24. https://doi.org/10.1021/ja0010709.
[22] Shinozaki J, et al. Dammaradiene synthase, a squalene cyclase, from Dryopteris crassirhizoma Nakai. Phytochemistry 2008;69(14):2559–64. https://doi.org/10.1016/j.phytochem.2008.07.017.
[23] Shinozaki J, et al. Molecular evolution of fern squalene cyclases. Chembiochem 2010;11(3):426–33. https://doi.org/10.1016/j.febslet.2007.12.023.
[24] Yamashita H, et al. Dammarane triterpenoids from rhizomes of Pyrrosia lingua. Phytochemistry 1998;49(8):2461–6. https://doi.org/10.1016/S0031-9422(98)00303-3.
[25] Hoshino T, et al. Squalene-hopene cyclase: final deprotonation reaction, conformational analysis for the cyclization of (3R,S)-2,3-oxidosqualene and further evidence for the requirement of an isopropylidene moiety both for initiation of the polycyclization cascade and for the formation of the 5-membered E-ring. Org Biomol Chem 2004;2(10):1456–70. https://doi.org/10.1039/B401172D.