



Transcriptome Analysis and Identification of the Cholesterol Side Chain Cleavage Enzyme BbgCYP11A1 From *Bufo bufo gargarizans*

Guangli Li^{1†}, Tianyue An^{2†}, Yu Li¹, Jinyang Yue¹, Ruoshi Huang¹, Jia Huang¹, Jincui Liang¹, Wei Yao¹, Liufang Huang¹, Yidu Chen¹, Rongrong Zhang¹, Aijia Ji^{1*} and Lixin Duan^{1*}

¹Joint Laboratory for Translational Cancer Research of Chinese Medicine of the Ministry of Education of the People's Republic of China, International Institute for Translational Chinese Medicine, School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China, ²School of Integrated Traditional Chinese and Western Medicine, Binzhou Medical University, Yanta, China

OPEN ACCESS

Edited by:

Shikai Liu,
Ocean University of China, China

Reviewed by:

D. Fernando Estrada,
University at Buffalo, United States
Huapu Chen,
Guangdong Ocean University, China

*Correspondence:

Lixin Duan
nlzn@gzucm.edu.cn
Aijia Ji
ajji@gzucm.edu.cn

[†]These authors have contributed equally to this work and share first authorship

Specialty section:

This article was submitted to
Livestock Genomics,
a section of the journal
Frontiers in Genetics

Received: 04 December 2021

Accepted: 16 February 2022

Published: 05 April 2022

Citation:

Li G, An T, Li Y, Yue J, Huang R, Huang J, Liang J, Yao W, Huang L, Chen Y, Zhang R, Ji A and Duan L (2022) Transcriptome Analysis and Identification of the Cholesterol Side Chain Cleavage Enzyme BbgCYP11A1 From *Bufo bufo gargarizans*. *Front. Genet.* 13:828877. doi: 10.3389/fgene.2022.828877

Bufo bufo gargarizans Cantor are precious medicinal animals in traditional Chinese medicine (TCM). Bufadienolides as the major pharmacological components are generated from the venomous glands of *B. bufo gargarizans*. Bufadienolides are one type of cardiac aglycone with a six-member lactone ring and have properties of antitumor, cardiogenic, tonsillitis, and anti-inflammatory. The biosynthesis of bufadienolides is complex and unclear. This study explored the transcriptome of three different tissues (skin glands, venom glands, and muscles) of *B. bufo gargarizans* by high-throughput sequencing. According to the gene tissue-specific expression profile, 389 candidate genes were predicted possibly participating in the bufadienolides biosynthesis pathway. Then, BbgCYP11A1 was identified as a cholesterol side chain cleaving the enzyme in engineering yeast producing cholesterol. Furthermore, the catalytic activity of BbgCYP11A1 was studied with various redox partners. Interestingly, a plant NADPH-cytochrome P450 reductase (CPR) from *Anemarrhena asphodeloides* showed notably higher production than BbgAdx-2A-BbgAdR from *B. bufo gargarizans*. These results will provide certainly molecular research to reveal the bufadienolides biosynthesis pathway in *B. bufo gargarizans*.

Keywords: *Bufo bufo gargarizans*, transcriptome, cholesterol side chain cleavage enzyme, bufadienolides biosynthesis, *Saccharomyces cerevisiae*

INTRODUCTION

Bufo bufo gargarizans Cantor (Bufonidae) are Chinese national protected wild animals with important economic and scientific research value. Chan'Su is a class of dried secretion in the venom gland or skin gland of *Bufo bufo gargarizans* Cantor and is used as an animal medicinal material with a long history. Bufadienolides, one kind of cardiac aglycone with a six-membered lactone ring, are the important bioactive ingredients of Chan'Su. Recent biological studies confirm the significant anticancer activities of bufadienolides *in vitro* and *in vivo* (Nogawa et al., 2001; Watanabe et al., 2003; Han et al., 2007). Cinobufagin, a hydrophobic compound, shows a great ability to inhibit the tumor cell by a pilot study and is allowed to be used for injection in a clinic that is approved by the Chinese Food and Drug Administration (CFDA) (Meng et al., 2009). Bufadienolides

is mainly extracted from medicinal toads but with a high economic cost. In addition, the medical resources of bufadienolides are in short supply. Recently, synthetic biology has played an increasingly important role in valuable natural products.

At present, the biosynthesis of bufadienolides is complex and unclear. Depending on the isotope labeling results, cholesterol is the starting compound and bile acid is the important intermediate involved in endogenous bufadienolides biosynthesis of mammals, and *Bufo* was researched in the 1970s (Porto and Gros, 1970, 1971; Porto et al., 1972). Recently, two important enzymes, 3β -hydroxysteroid dehydrogenase (3β -HSD) and steroid 5β -reductase (5β -POR), were characterized from *Bufo* species and are possibly involved in the bufadienolides biosynthetic pathway. 3β -HSD, which belongs to the short-chain dehydrogenase/reductase (SDR), can catalyze two sequential reactions to generate progesterone, including converting the 3β -hydroxysteroid to 3-oxo intermediate and then isomerizing the Δ^5 form to the Δ^4 form and the enzyme is investigated to be accumulated nearly 20-fold in venom glands than in liver (Xu et al., 2018). Then, the 5β -POR encodes the progesterone 5β -reductase, which is also confirmed *in vitro* (Zhang et al., 2021). The biosynthetic pathway of bufadienolides is so complicated that the relating key enzymes still need to be deeply illustrated by the method of transcriptomics.

Another essential enzyme, P450scc, was investigated to be possibly involved in the cardiac glycosides biosynthesis pathway in plants such as *Scilla* and *Urginea* (Sauer et al., 1967; Stache et al., 1969), *Calotropis procera* (Pandey et al., 2016). It is interesting to explore the correlation of biosynthetic pathways and bufadienolides compounds between the animal and plant sources. It is noteworthy that abnormal steroidogenesis could cause serious disease in the human body (Zhang et al., 2012). The demand for steroid precursors seems to continuously enlarge due to the increasing usage of steroid drugs, such as glucocorticoids, mineralocorticoids, and sex hormones. According to a biosynthesis study of steroid hormone compounds in animals, P450scc, as a key enzyme, can catalyze cholesterol to produce the biologically active compound pregnenolone (Hochberg et al., 1974; Duque et al., 1978; Efimova et al., 2019), which plays an important role in all classes of steroidal drug biosynthetic pathways. This reaction first contains two hydroxylation steps on the side chain of cholesterol at the C22 and C20 positions to produce 22R-hydroxycholesterol and 20 α , 22R-dihydroxycholesterol, and then scission of the side chain between C20-C22 (Strushkevich et al., 2011). The identification of BbgP450scc enriched the species diversity of this enzyme and was of significance to study the characteristics of BbgCYP11A1, such as substrate specificity, enzyme kinetics.

In addition, a complete transcriptomic analysis of different tissues (skins, venom glands, and muscles) from the *Bufo* species is needed. In this study, we systematically investigated the transcriptome analysis of *B. bufo gargarizans* in three different tissues to understand the molecular basis of bufadienolides biosynthesis. Through the transcriptome sequencing of *B. bufo gargarizans*, candidate genes of the P450scc system were isolated

with a homologous gene from *G. rugose*, *X. tropicalis*, *X. laevis*, and *N. parkeri* as probes. The biochemical characterization of the BbgP450scc system successfully catalyzed cholesterol into pregnenolone by heterologous expression in an engineering yeast producing cholesterol. These data provide a comprehensive resource to speculate on the biosynthesis of *B. bufo gargarizans*.

MATERIALS AND METHODS

Animal Materials and RNA Extraction

Three *B. bufo gargarizans* were obtained from Foshan, Guangdong Province in China. Animal treatment and experimental protocols were conducted in accordance with the Animal Care Guidelines of Guangzhou University of Chinese Medicine. After being sacrificed, three tissues, including muscles, skins (containing glands), and venom glands from three toads, were collected, respectively. All samples were frozen immediately in liquid nitrogen and then stored in -80°C for further processing. Total RNA was extracted using a Quick RNA Isolation Kit (Huayueyang Biotechnology) according to the manufacturer's protocol. The quality and quantity of total RNA were assessed using 1% agarose gels and a Qubit 2.0 Fluorometer (Invitrogen, United States).

cDNA Library Construction for Sequencing and *de novo* Assembly

The mRNA was purified from the total RNA using mRNA capture beads and subsequently sheared short fragments under the condition of divalent cations and high temperatures. Further, these fragments were reverse-transcribed into first-strand cDNA by random hexamers. Then, second-strand cDNA was synthesized in the presence of DNA Polymerase I and RNase H. Clean cDNA was produced using $1 \times$ Ampure XP DNA clean beads after a series of operations, including end repair, addition of "A" base and ligation of Illumina adapters. For the enrichment of adapter ligated fragments, 15 cycles of PCR reaction were carried out to amplify cDNA libraries. Nine cDNA libraries were sequenced on the Illumina HiSeq Nova Sequencing platform, and 150 bp paired-end reads were generated. Quality control of raw reads was performed by FastQC. Reads with unknown nucleotides and low-quality reads less than 20 bp were trimmed using Prinseq software (Bolger et al., 2014) and sequence contamination was assessed by blast+ in NCBI (Altschul et al., 1997). Clean reads were assembled into the unigenes by Trinity software v2.4.0 (Haas et al., 2013).

Sequence Annotation and Functional Classification

Functional annotation was performed by BLASTP against the protein databases NR (NCBI nonredundant proteins sequences), PFAM (Protein family), CDD (Conserved Domain Database), KOG (eukaryotic Ortholog Groups), GO (Gene Ontology), KEGG (Kyoto Encyclopedia of Genes and Genomes), Swiss-

Prot (a manually annotated and reviewed protein sequence database), TrEMBL, and NT (NCBI nucleotide sequences). GO terms were mainly divided into three parts, including biological processes, molecular functions, and cellular components. GO enrichment analysis was implemented by the BLAST result from the UniProt database. The *B. bufo gargarizans* gene set was also mapped to the KEGG pathway database to identify the best match for each gene (Gerlich and Neumann, 2000).

Quantification of the Differentially Expressed Gene Expression Levels

Gene expression levels were calculated using Salmon (Patro et al., 2017), and the read count for each gene in various organs was normalized into a transcripts per million (TPM) value. Genes with TPM value >3 were selected for further differentially expressed gene (DEG) analysis. DEGs among muscles, skins, and venom glands, were performed by DESeq software. The unigenes with a |Fold Change| > 2 and an adjusted *p*-value (*q*-value) < .05 were considered as DEGs. GO enrichment analysis of DEGs was carried out by topGO, and KEGG enrichment analysis of DEGs was implemented by clusterProfiler (Yu et al., 2012). DEGs from GO terms and KEGG terms with *q*-value < 0.05 were considered to be considerably enriched.

Phylogenetic Analyses and Multiple Alignment of P450scc System

To identify the nucleotide and deduced amino acid sequences of the cloned cDNA sequence, a neighbor-joining tree was generated using MEGA7 software with default parameters based on 1,000 bootstrap replications (Sudhir et al., 2016). Evolview (<https://www.evolgenius.info/evolview/>) was utilized to edit the phylogenetic trees. Multiple sequence alignment was implemented by DNAMAN software (version 6.0).

Validation of P450scc System by Quantitative Real-Time PCR

Total RNA of nine samples from three tissues were isolated by a Quick RNA Isolation Kit (Huayueyang Biotechnology), and the cDNA was synthesized using the Evo M-MLV One Step RT-PCR Kit (Accurate Biotechnology). The sequence encoding BbgP450scc system, including BbgCYP11A1, BbgAdx (adrenodoxin) and BbgAdR (adrenodoxin reductase), was obtained from the transcriptome data. Here, the housekeeping gene β -actin from *B. bufo gargarizans* was used as an internal reference gene (Wu et al., 2017). Real-time PCR reactions were performed using SYBR[®] Green Premix ProTaq HS qPCR Kit (Accurate, Hunan, China) on the QuantStudio[™] 5 Real-Time PCR Instrument (384-Well Block) (Thermo Fisher Scientific, United States). Each reaction was performed in a final volume of 10 μ L, containing 8.2 μ L SYBR Green PCR Master Mix, 10 mM of each primer, and 1 μ L cDNA template. The amplification program was as follows: 50°C for 20 s, 90°C for 30 s, followed by 40 cycles of 95°C for 15 s and 60°C for 34 s. All reactions were performed with three biological replicates. The relative

expression level of genes was calculated by the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2002). Primers used in the qRT-PCR detection were listed in **Supplementary Table S4**.

Cloning of P450scc System Genes

The cDNA sample of *B. bufo gargarizans* was used for gene cloning. The full-length cDNA sequences of BbgAdx (MZ 384243) and BbgAdR (MZ 384244) were obtained by PCR reaction. The nucleotide sequences of redox partners from *Mus musculus* were downloaded from NCBI database and named as MmAdx (NM 007996.2) and MmAdR (NM 007997.1). A 2A peptide linker was used to fuse the redox partners (Ryan and Drew, 1994). The fused proteins were derived by overlapping PCR and named as CTL-1 (BbgAdx-2A-BbgAdR) and CTL-2 (MmAdx-2A-MmAdR). Both of them were cloned into the pESC-Leu vector (Agilent Technologies, Inc.) to obtain plasmids pGL01 and pGL02 using the ClonExpress[®] II One Step Cloning Kit (Vazyme, Nanjing, China).

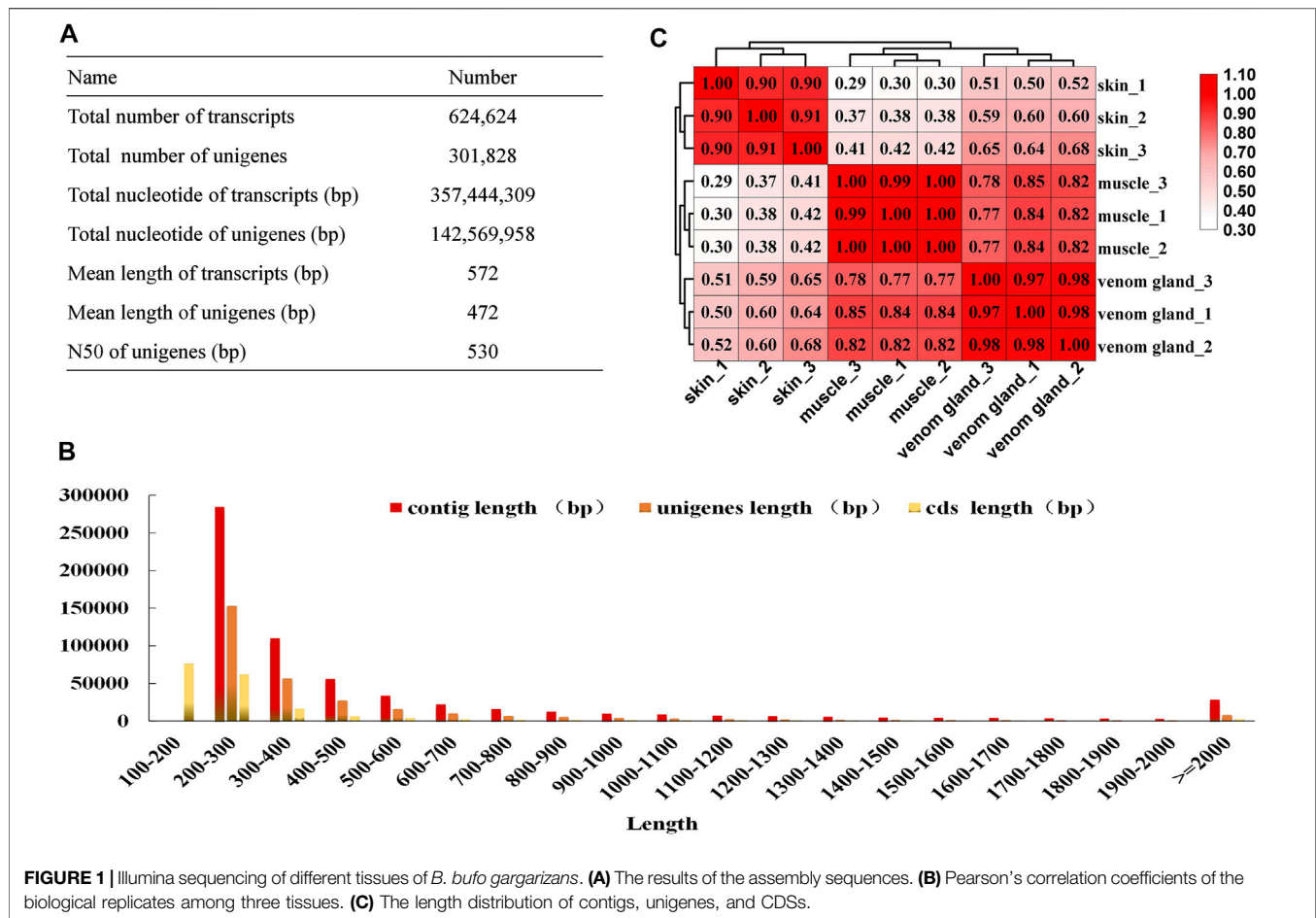
The full-length cDNA sequence of BbgCYP11A1 (MZ 384245) was amplified by using RACE PCR methods and then cloned into plasmids pESC-Leu, pESC-Leu-AaCPR (constructed previously from our laboratory), pGL01 and pGL02, respectively, to obtain vector pGL03, pGL04, pGL05 and pGL06. The primers and sequences used in this section are shown in **Supplementary Table S5** and **Supplementary Table S6**, and detailed information about the plasmids is listed in **Supplementary Table S7**.

Heterologous Expression of P450scc System in *Saccharomyces cerevisiae*

A cholesterol producing strain, RH6829 (*MATa ura3 leu2 his3 trp1 can1 bar1 erg5Δ:: HIS5-TDH3-DHCR24 erg6Δ:: TRP1-TDH3-DHCR7*), was kindly gifted by Professor Howard Riezman (Souza et al., 2011). The expression plasmids, pGL03, pGL04, pGL05, and pGL06, were transformed into these strains to get strains GL01, GL02, GL03, and GL04, respectively. The transformants were selected on SD-Trp-His-Leu medium containing 2% glucose and further identified by the PCR method. Positive colonies of these strains were cultured in 50 ml SD-Trp-His-Leu medium with 18% galactose and 2% glucose for 7 days with shaking at 220 rpm and 30°C.

Extraction and GC-MS Analysis of Steroid Compounds

The cultures of each strain were centrifuged at 4000 rpm for 5 min, and the cells were harvested and then lysed in 10 ml saponification reagent [20% (w/v) KOH in 50% (v/v) ethanol]. After being extracted twice by equal volumes of *n*-hexane, the organic phase extracts were collected and dried by a centrifugal vacuum evaporator. The crude extract was derived by using 30 μ L N-Methyl-N-(trimethylsilyl) trifluoroacetamide (Sigma-Aldrich, China). Then, the reactant was analyzed using a Shimadzu QP2010SE instrument under the electronic impact of 70 eV with a Rxi-5Sil MS capillary column (0.25 mm \times 30 m) (Shimadzu, China). The sample analysis on Gas chromatography mass



spectrometry (GS-MS) was performed according to the previous report (Zhang et al., 2019).

RESULTS

Transcriptome Sequencing and *de novo* Assembly

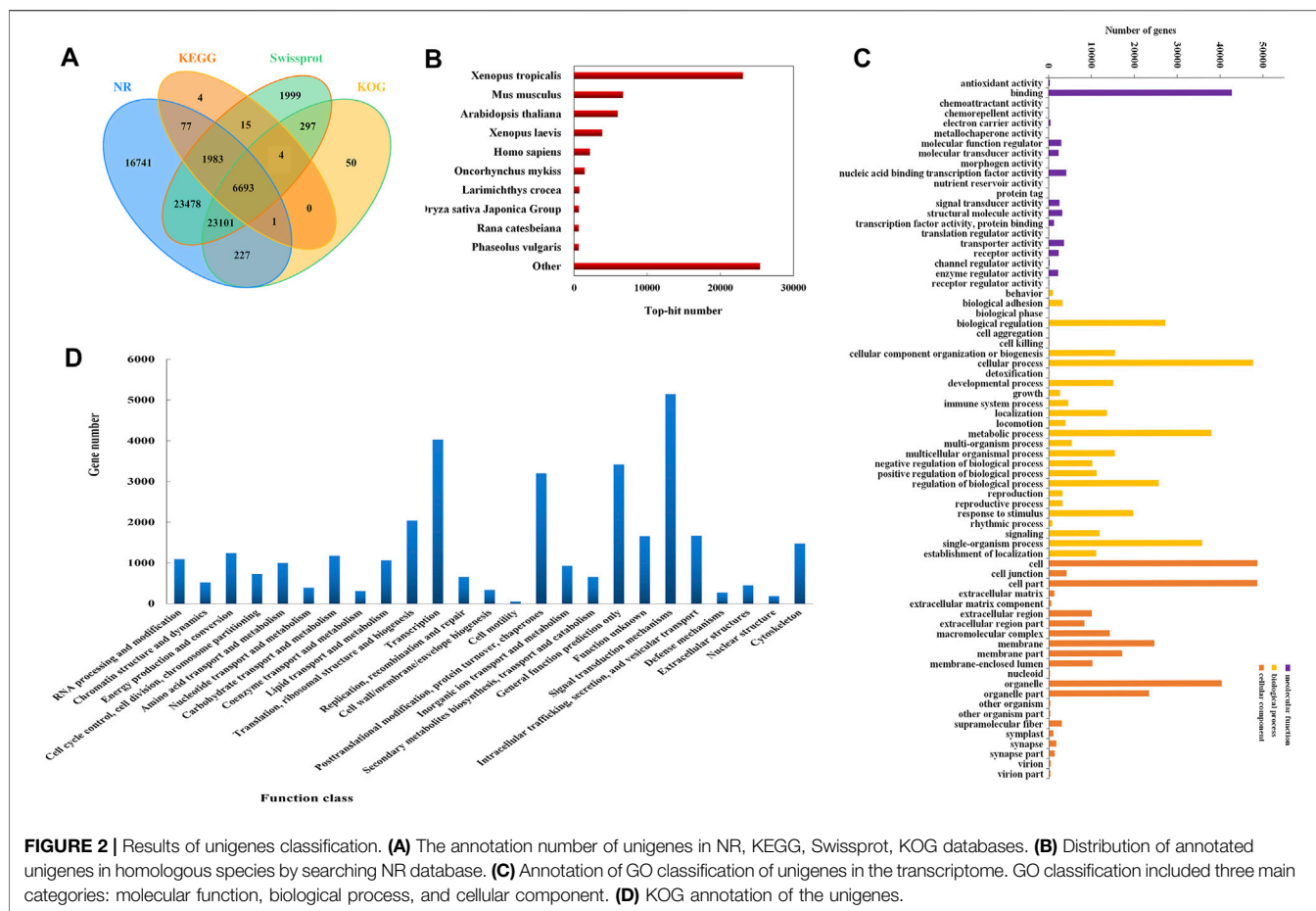
To obtain the key genes of the bufadienolides biosynthetic pathway, RNA-seq was applied using three tissues (muscles, skins, and venom glands) based on the bufadienolides accumulation pattern in *B. bufo gargarizans*. Nine cDNA libraries were constructed and sequenced on the Illumina HiSeq Nova Sequencing platform with a pair-end length of 150 bp. Approximately 493.8 million raw reads were obtained, and after filtering low-quality reads, adaptor, and short reads, 443.9 million clean reads were generated with the Q30 > 96% (Supplementary Table S1). All the clean reads were *de novo* assembled into 624,624 transcripts, and after redundancy elimination, 301,828 unigenes were identified with a mean length of 472 bp and an N50 length of 530 bp (Figure 1A). The length distribution of transcripts, unigenes, and CDSs is presented in Figure 1B. For unigenes, 2.56% of the unigenes were >2000 bp in length, and most sequence lengths were between 200 and 600 bp. Pearson's correlation coefficients of the biological replicates of each

tissue were more than 0.90, indicating their high correlation (Figure 1C). These results reveal that the sequencing data has high quality and can be used in the subsequent analysis.

Functional Annotation and Classification of Unigenes

Total 301,828 unigenes were annotated against nine publicly available databases comprising Conserved Domain Database (CDD), Kyoto Encyclopedia of Genes and Genomes (KEGG), NCBI non-redundant protein sequences (NR), NCBI nucleotide sequences (NT), Protein family (PFAM), a manually annotated and reviewed protein sequence database (Swiss-Prot), TrEMBL, Gene Ontology (GO) and eukaryotic Ortholog Groups (KOG), and 86,783 unigenes had significant hits in at least one database (Figure 2A, Supplementary Table S2). Most unigenes were annotated in the NR database (23.95%), and the annotated sequences were blasted against other species. The result of species distribution showed 23,096 homologous unigenes between *B. bufo gargarizans* and *Xenopus tropicalis*, indicating their high homology (Figure 2B).

A total of 652,248 unigenes were classified into three main GO functional categories, corresponding to molecular function (93,600), biological process (325,157), and cellular components



(258,683). The three categories were further divided into 21, 27, and 22 terms, respectively. The majority of unigenes in the biological process belong to the cellular process (47,642), then followed by the metabolic process (37,904). In the category of molecular function, ATP binding was the main term, which was of 42,741 unigenes. In addition, most unigenes in the cellular component were assigned to cell (48,584) and cell part (48,537) (Figure 2C).

33,696 unigenes were assigned into 25 different groups in KOG analysis, and the largest group was “signal transduction mechanisms”, which consists of 5141 members. Moreover, there were 657 unigenes clustered into “secondary metabolites biosynthesis, transport, and catabolism”, and the unigenes between them may be involved in the bufadienolides biosynthetic pathway (Figure 2D).

Differential Expression Genes Among Tissues

To investigate the DEGs, the DEGs were identified based on the values of TPM among different tissues. The figure showed 168,316 genes expressed in all three tissues, whereas there were 32,484 genes specifically expressed in the skins, followed by 6419 in muscles and 6393 in the venom glands (Figure 3A). Compared with the muscles, there are 3697 upregulated genes and 2328 downregulated genes were identified in the skins, whereas in the venom glands, 4980

upregulated genes and 1629 downregulated genes were detected (Figure 3B). Compared with the venom glands, there were 2608 upregulated and 3216 downregulated genes in the skins (Figure 3B, Supplementary Sheet S1). The expression patterns of all DEGs among three tissues were hierarchically clustered in the heat map (Figure 3C). The DEGs of venom glands were clustered together with the DEGs of skins in a subclade. However, genes in the muscles were clustered into a different clade that indicates their significant discrepancy.

KEGG enrichment analysis was performed to explore the potential DEGs in different metabolism pathways. Compared with the muscles, there were 1485 upregulated genes in the skin and 1694 upregulated genes in the venom glands. These DEGs were enriched in several pathways, including terpenoid backbone, steroid, steroid hormone, and primary bile acid biosynthesis (Supplementary Figure S1, Supplementary Figure S2 and Supplementary Figure S3).

Candidate Genes Involved in Putative Bufadienolides Biosynthesis

Candidate genes were screened via homology-based searches of the transcriptome. In total, 389 unigenes were predicted to be involved in the bufadienolides metabolic pathway, which can be divided into three parts: upstream terpenoid backbone biosynthetic, midstream

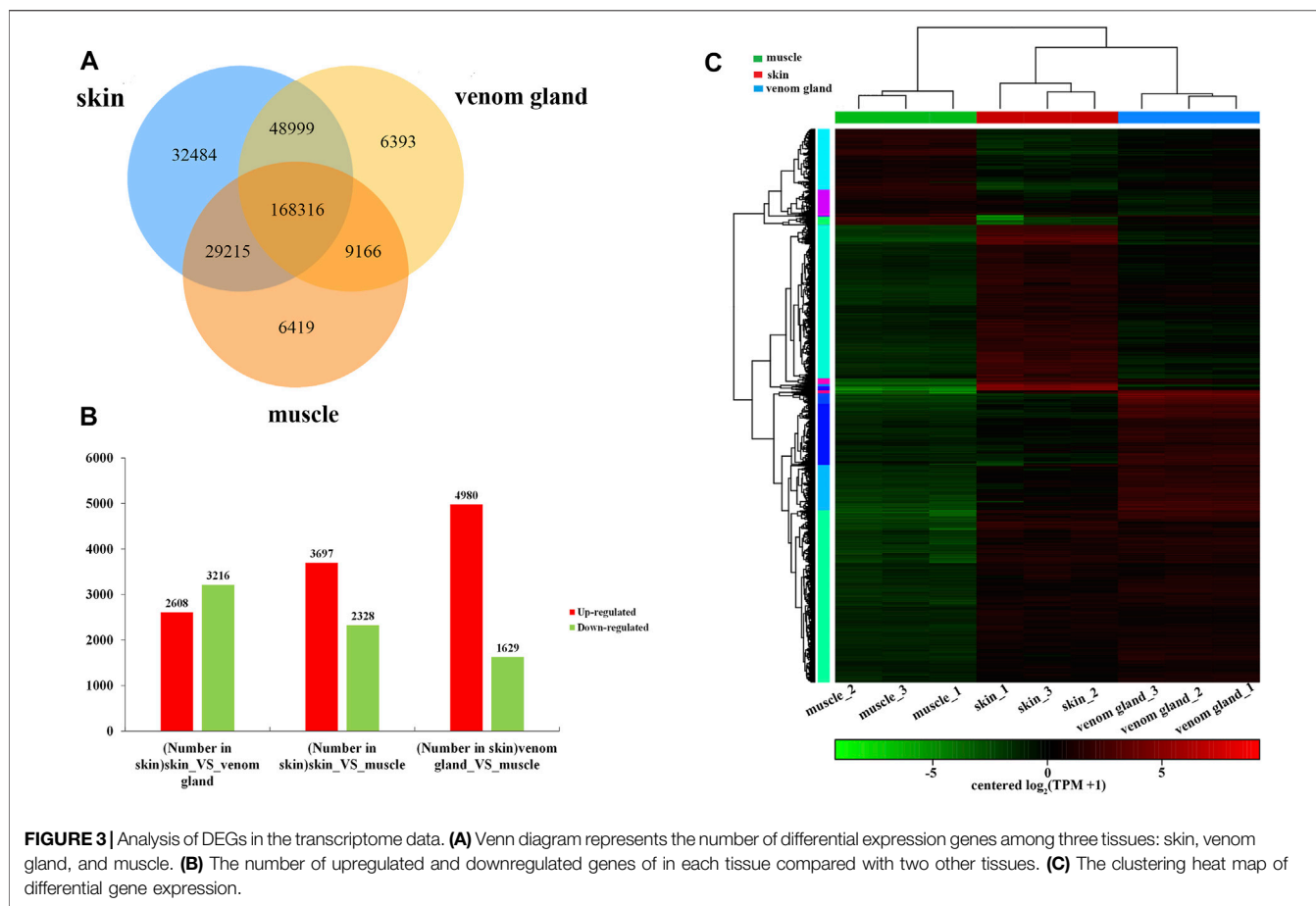


FIGURE 3 | Analysis of DEGs in the transcriptome data. **(A)** Venn diagram represents the number of differential expression genes among three tissues: skin, venom gland, and muscle. **(B)** The number of upregulated and downregulated genes in each tissue compared with two other tissues. **(C)** The clustering heat map of differential gene expression.

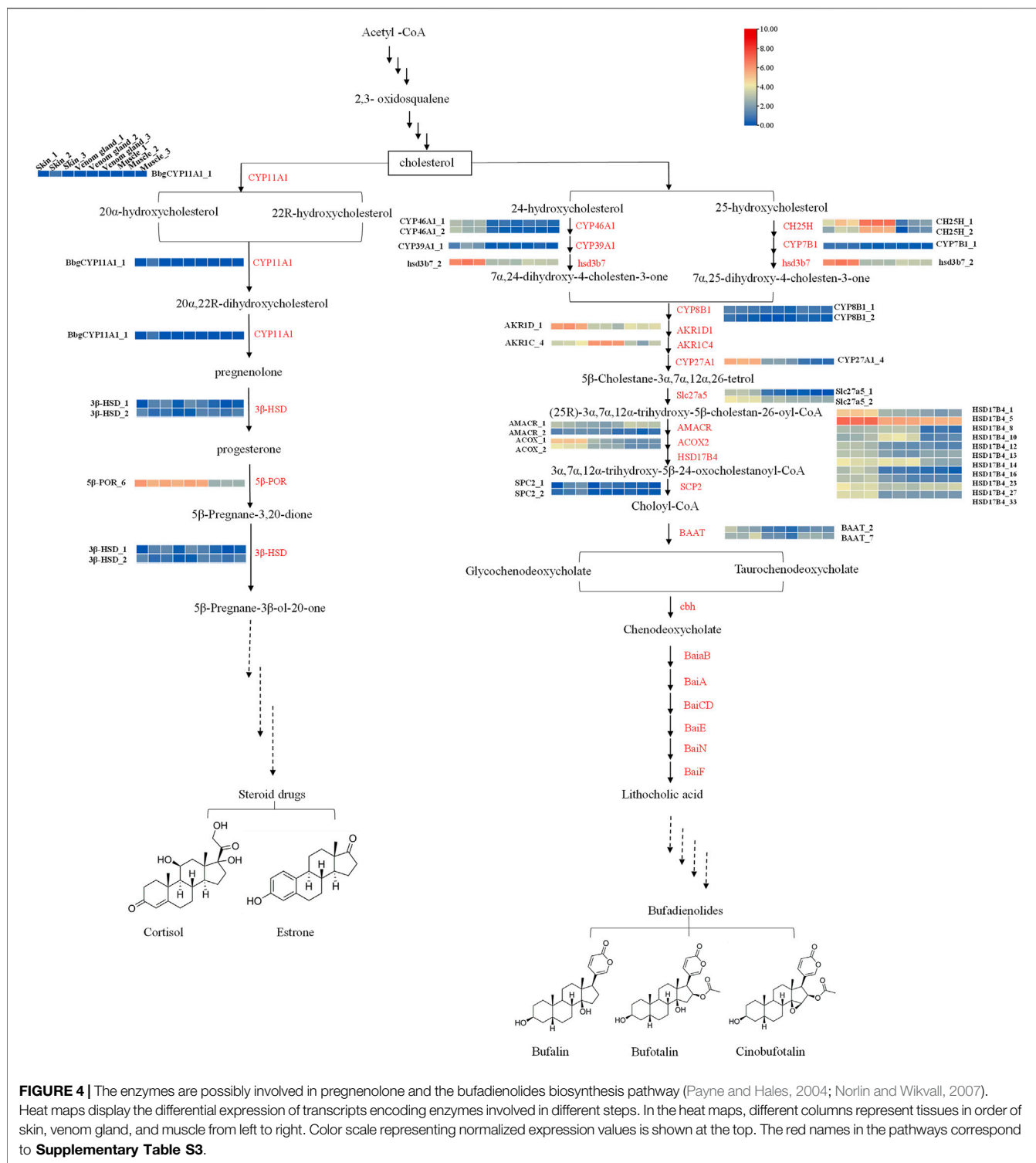
cholesterol biosynthetic, and downstream bufadienolides biosynthetic pathways (**Figure 4**, **Supplementary Figure S4**). In terpenoid backbone biosynthesis, 42 candidate genes encoding seven enzymes were identified as follows: AACT, Acetyl-CoA C-acetyltransferase (16 unigenes); HMGS, hydroxymethylglutaryl-CoA synthase (4 unigenes); HMGR, hydroxymethylglutaryl-CoA reductase (NADPH) (13 unigenes); MK, mevalonate kinase (2 unigenes); PKM, phosphomevalonate kinase (2 unigenes); MVD, diphosphomevalonate decarboxylase (1 unigene); FDPS, farnesyl diphosphate synthase (4 unigenes). Through the expression pattern analysis, seven unigenes encoding six enzymes were identified as DEGs in the upstream pathway. According to the heat map as shown, the enzymes AACT_4 and FDPS1_1 showed higher expression in the skin than in the venom glands (**Figure 5A**, **Supplementary Sheet S2**).

In the midstream pathway, there were 53 candidate genes encoding 12 enzymes that may involve the cholesterol biosynthesis pathway (e.g., SQS, squalene synthase; SQE, squalene monooxygenase; Lss, lanosterol synthase; tm7sf2, Delta14-sterol reductase; msmo1, methyl sterol monooxygenase, etc.). In addition, 11 unigenes encoding seven enzymes were selected as DEGs. The enzyme SQS showed the highest expression in the skins and the lowest expression in muscles. The enzymes dhcr24 and tm7sf2 also have the similar expression patterns with the SQS (**Figure 5B**, **Supplementary Sheet S2**).

The downstream pathway may relate to the bile acid biosynthesis pathway in animals (Fedorova et al., 2015). Bile acid is a common production in the liver in animals (Reuben, 1987; Russell and Setchell, 1992). The enzymes related in the metabolic pathway of **Figure 4** were predicted according to bile acid and steroid hormone biosynthesis from the KEGG pathway and those previous studied (Payne and Hales 2004; Norlin and Wikvall 2007). There were 145 unigenes encoding 19 enzymes (CYP46A1, CYP39A1, HSD3B7, CH25H, etc.) that may relate to the bile acid biosynthesis pathway (**Figure 4**). For the downstream pathway, 294 candidate genes encoding 27 enzymes, including 52 unigenes of cytochromes P450 that may involve the pregnenolone and bufadienolides biosynthesis pathways, were predicted. The heat map showed that most of DEGs had relatively high expression in the skin (Chen et al., 2020) (**Figure 5C**, **Supplementary Sheet S2**).

Phylogenetic Analyses and Multiple Alignment of BbgP450scc System

To employ the relationship of BbgP450scc and the biosynthesis of bufadienolides, BbgP450scc gene was screened from transcriptome data. P450scc is also an important first step enzyme participating in steroid drugs biosynthesis and can convert cholesterol to pregnenolone, which is the key precursor to synthesize all classes



of steroidal drugs. Two short sequences TRINITY_DN118184_c0_g1 (BbgCYP11A1_1) and TRINITY_DN130445_c0_g1 (BbgCYP11A1_2) were screened from the transcriptome data by BLASTP. BbgCYP11A1_1, and BbgCYP11A1_2 encoded 73 and 49 amino acids, respectively. Only the sequence of BbgCYP11A1_1 could be derived by RACE-PCR for the subsequent research. The

complete sequence BbgCYP11A1_1 encoding 512 amino acids was isolated from the cDNA library. At the same time, two unigenes (BbgAdR_1, BbgAdR_2) were annotated as adrenodoxin oxidoreductase and six unigenes (BbgAdx_1, BbgAdx_2, BbgAdx_3, BbgAdx_4, BbgAdx_5, BbgAdx_6) were annotated as adrenodoxin. Full ORF sequences are found in BbgAdx_1 and

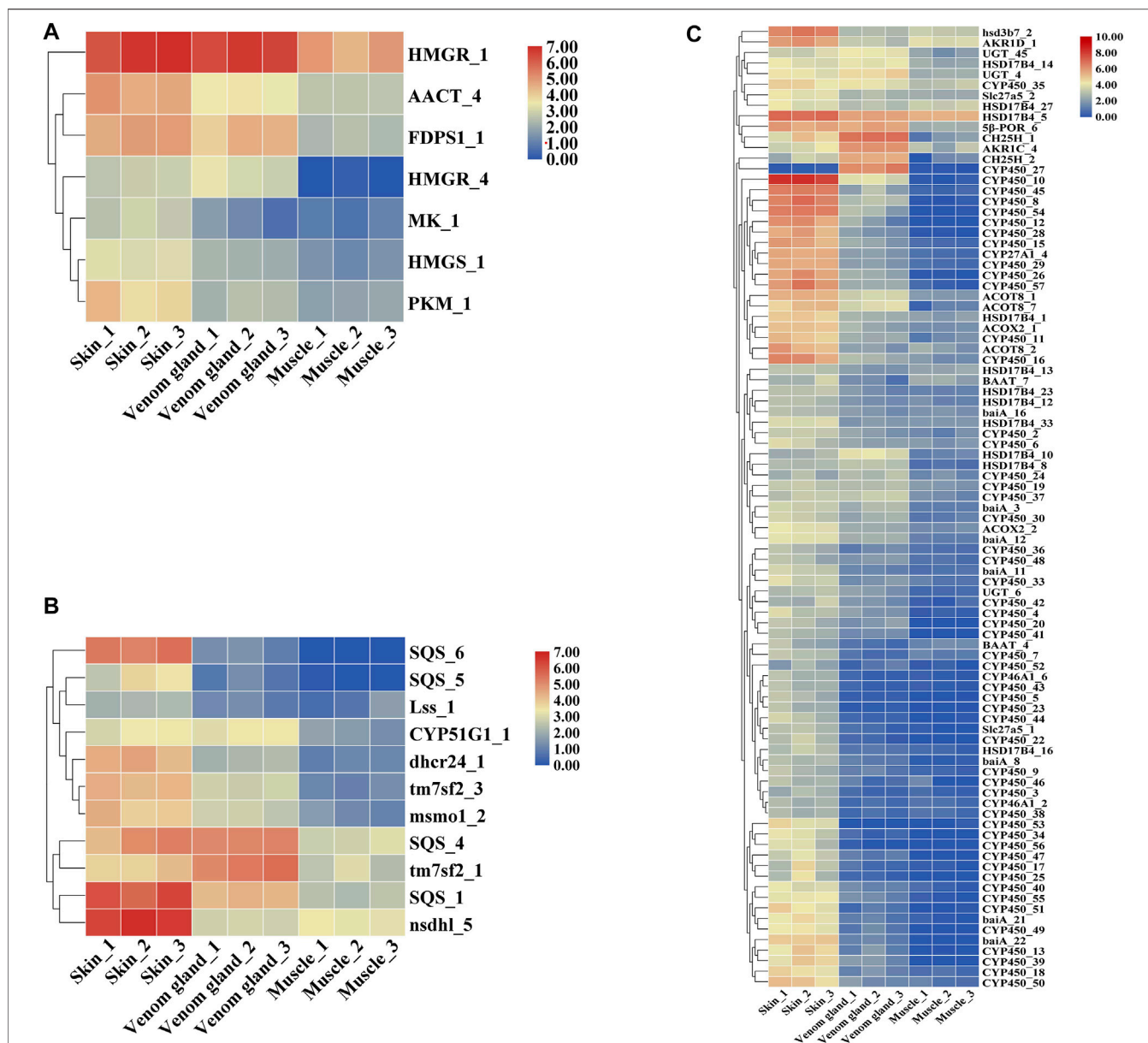


FIGURE 5 | The heat map of the expression pattern of genes involved in pregnenolone and bufadienolides biosynthetic pathway. **(A)** Upstream terpenoid backbone biosynthetic pathway, **(B)** Midstream cholesterol biosynthetic pathway, **(C)** Downstream pregnenolone and bufadienolides biosynthetic pathway.

BbgAdR_1, encoding 167 and 494 amino acids, respectively. A phylogenetic tree was constructed using the protein sequences with CYP11A1, Adx, and AdR from the species of mammals, birds, amphibians, and bony fish, etc. (Figure 6). As the multiple sequence alignment result shows, the deduced amino acid sequences BbgCYP11A1 had more than 78% identity with Amphibian CYP11A1 (Supplementary Figure S5).

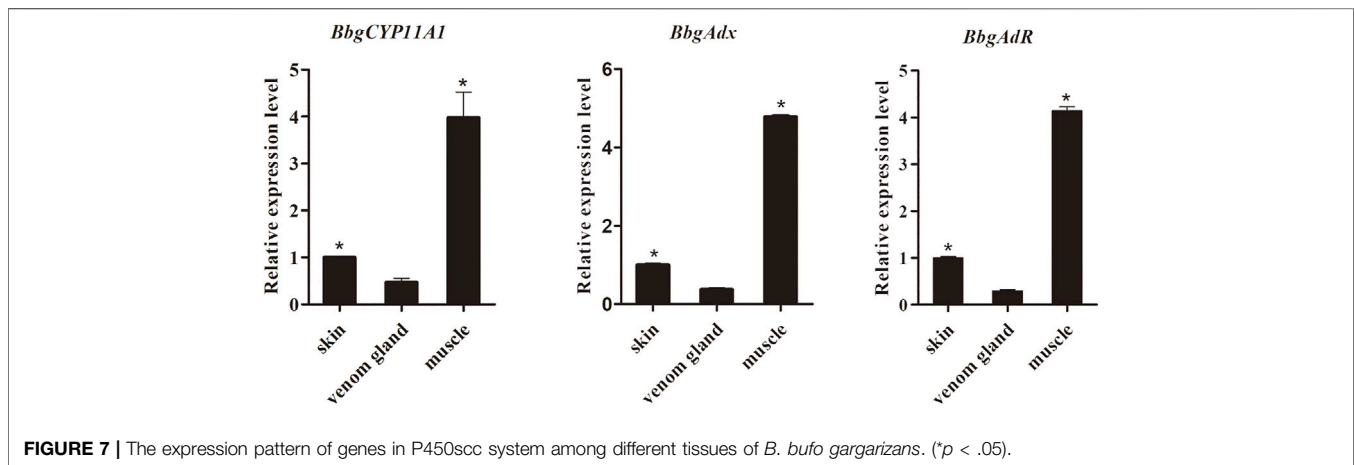
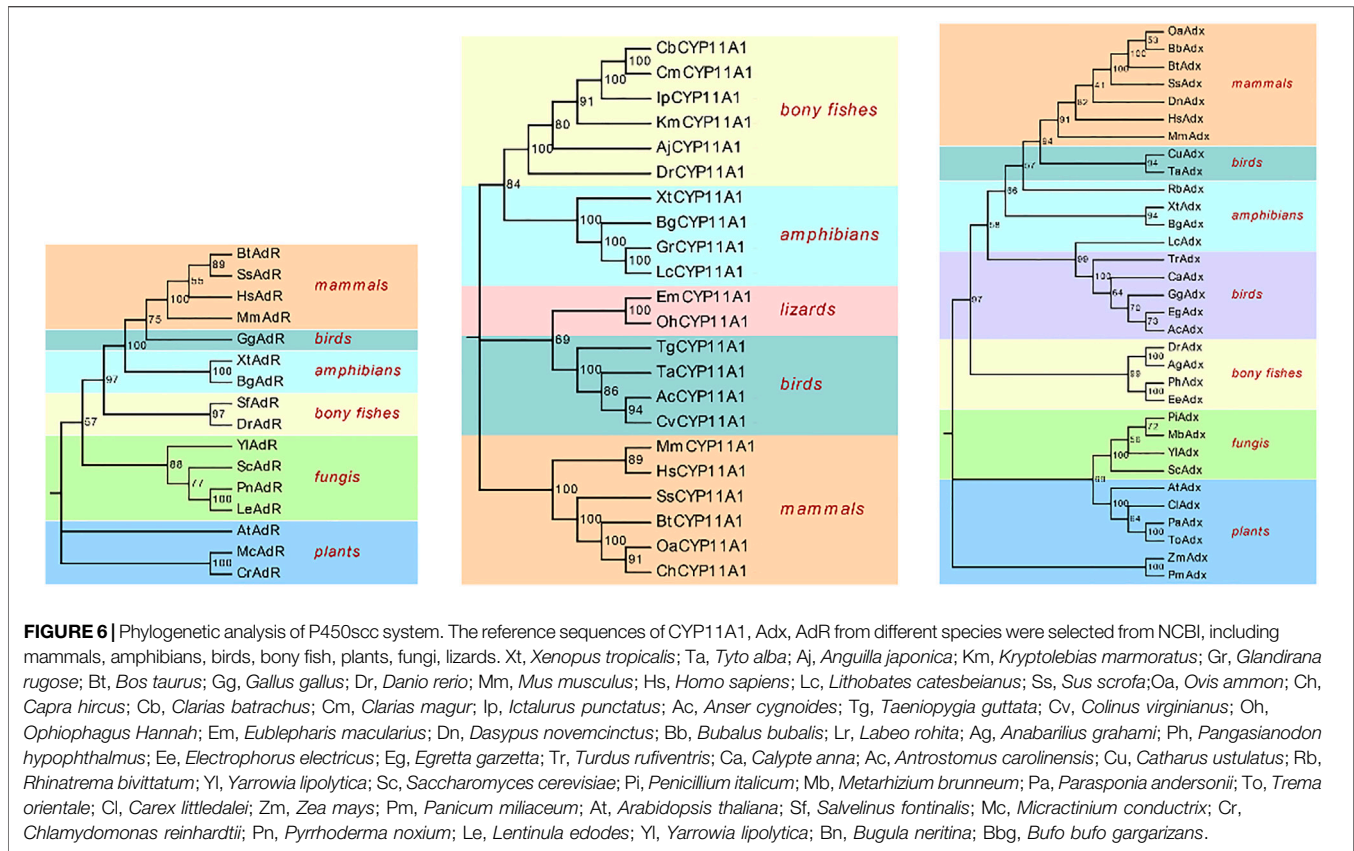
Isolation and Tissue Expression Profile of BbgP450scc System Genes

Aimed to verify the biochemical function of BbgP450scc system genes, including BbgCYP11A1 and its redox partners BbgAdx

and BbgAdR, the complete sequences of them were acquired and qRT-PCR was carried out to detect the expression patterns of BbgP450scc system genes. BbgP450scc system genes had similar expression profiles showing the higher accumulation in muscles than in venom glands and in skins (Figure 7).

Functional Characterization of BbgP450scc System in Yeast

P450scc can convert cholesterol, campesterol, 7-reductase ergosterol, β -sitosterol, and lumisterol into pregnenolone. As the best substrate, cholesterol can be fully converted to pregnenolone by P450scc (Tuckey et al., 2014; Gerber et al.,



2015). To characterize the catalytic function of BbgCYP11A1, the heterologous vector pGL05 containing BbgCYP11A1 was transformed into the cholesterol-producing strain RH6829 (Souza et al., 2011). The catalytic product was identified as pregnenolone by comparing the MS spectrum and retention time with the standard (Figure 8A and Figure 8C). In addition, the catalytic efficiency of P450scc is influenced by

the redox partners, adrenodoxin reductase (AdR), and adrenodoxin (Adx) (McLean et al., 2005). BbgAdx-2A-BbgAdR (CTL1) from *B. bufo gargarizans*, MmAdx-2A-MmAdR (CTL2) from *M. musculus*, and AaCPR from plant *Anemarrhena asphodeloides* Bunge were employed here to evaluate the catalytic capability of BbgCYP11A1 in yeast. Results show that all the redox partners can work with

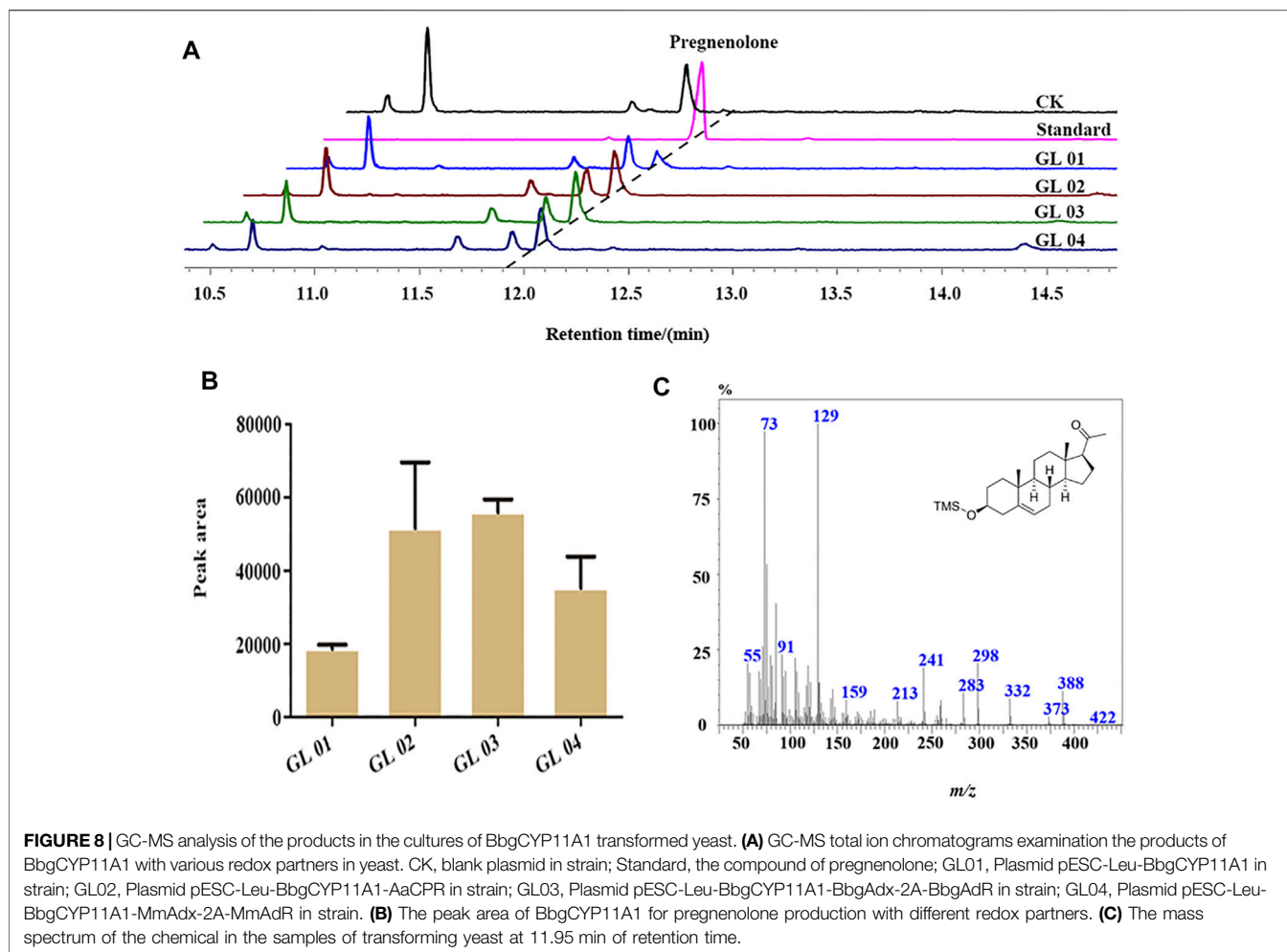


FIGURE 8 | GC-MS analysis of the products in the cultures of BbgCYP11A1 transformed yeast. **(A)** GC-MS total ion chromatograms examination the products of BbgCYP11A1 with various redox partners in yeast. CK, blank plasmid in strain; Standard, the compound of pregnenolone; GL01, Plasmid pESC-Leu-BbgCYP11A1 in strain; GL02, Plasmid pESC-Leu-BbgCYP11A1-AaCPR in strain; GL03, Plasmid pESC-Leu-BbgCYP11A1-BbgAdx-2A-BbgAdR in strain; GL04, Plasmid pESC-Leu-BbgCYP11A1-MmAdx-2A-MmAdR in strain. **(B)** The peak area of BbgCYP11A1 for pregnenolone production with different redox partners. **(C)** The mass spectrum of the chemical in the samples of transforming yeast at 11.95 min of retention time.

BbgCYP11A1, and the strain GL02 and GL03 displayed higher production than GL01 and GL04, showing the great function in yeast (Figure 8B).

DISCUSSION

Cardiac glycosides widely exist in the plant kingdom and in animals (Hollman, 1985). Especially in *Bufo* species, their body is capable of producing bufadienolides, which have similar pharmacologic actions with cardenolides generating in plants. The putative biosynthetic pathway leading to the cardenolides is basically deduced from the radiolabeled precursors. The cholesterol side chain cleaving enzyme is important for the formation of cardenolides, which is still unknown in plants (Sauer et al., 1967). Here, we characterized the function of cholesterol side chain cleaving enzyme in *B. bufo gargarizans*, and the expression level analysis of BbgP450scc system was done by RT-PCR. Our study provides more data for deciphering the

process of biosynthesis pathway of bufadienolides in *B. bufo gargarizans*.

As the biosynthesis pathway of bufadienolides in *B. bufo gargarizans* has not been clarified yet, to explore the candidate genes of the bufadienolides biosynthetic pathway is needed. According to the result of the transcriptome sequencing and assembly, we analyzed the expression patterns of all the transcripts involved in the putative biosynthesis pathway of bufadienolides and steroid hormones. The transcripts relating to the bile acid biosynthesis, such as cholesterol 24-hydroxylase, cholestanetriol 26-monooxygenase, showed their higher expression levels in skin and venom glands than in muscles. Cholesterol is proved as the precursor of bufadienolides, which are a group of polyhydroxy C-24 steroids with a six-membered lactone (apyrone) ring at the C-17 position. The heat map in Figure 8C shows that CYPs (8, 10, 12, 15, 28, 45, and 54) has significant differential expression in skin, which demonstrates these P450s enzymes might participate in the modification of its parent nucleus. These

results indicate that these DEGs are probably involved in bufadienolides biosynthesis generating in skins. This work also provides the foundation for the identification of genes in the biosynthesis pathway.

Steroid compounds are one of the most routinely used chemicals in disease treatment with extensive pharmacological effects (Kuhl, 2011). Nowadays, the most typical marketed products of steroid drugs are cortisol, progesterone, and sex hormone for the treatment of anti-inflammatory, anti-allergic, and contraceptive agents. Using the approach of genetic engineering to improve the production of precursor pregnenolone is an important step in steroidogenesis producing (Duport et al., 1998; Szczebara et al., 2003; Zhang et al., 2019). Then, different redox partners were used to explore the enzymatic activities of BbgCYP11A1, and the results demonstrate that the confusion of BbgAdx and BbgAdR exhibited the relatively higher catalytic efficiency compared with the confusion of the MmAdx and MmAdR. A redox partner AaCPR from the plant also enhanced the enzymatic activity of BbgCYP11A1. Surprisingly, the hybrid combination of BbgP450scc with plant redox partner AaCPR led to a comparable increase in the production of pregnenolone than other combinations. This result provides a cross-species combinatorial biosynthesis approach for pregnenolone metabolic engineering. Moreover, the endogenous redox partners in yeast also can interact with BbgCYP11A1, further suggesting the existence of cross-species redox chain for animal cytochrome P450s.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

REFERENCES

- Altschul, S., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a New Generation of Protein Database Search Programs. *Nucleic Acids Res.* 25 (17), 3389–3402. doi:10.1093/nar/25.17.3389
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* 30 (15), 2114–2120. doi:10.1093/bioinformatics/btu170
- Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: an Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* 13 (8), 1194–1202. doi:10.1016/j.molp.2020.06.009
- Duport, C., Spagnoli, R., Degryse, E., and Pompon, D. (1998). Self-sufficient Biosynthesis of Pregnenolone and Progesterone in Engineered Yeast. *Nat. Biotechnol.* 16 (2), 186–189. doi:10.1038/nbt0298-186
- Duque, C., Morisaki, M., Ikekawa, N., Shikita, M., and Tamaoki, B.-i. (1978). The Final Step of Side-Chain Cleavage of Cholesterol by Adrenocortical Cytochrome P-450(scc) Studied with [22-¹⁸O] 20, 22-dihydroxycholesterols, [¹⁸O] Isocaproaldehyde, [¹⁸O] Water and Atmospheric [¹⁸O] Oxygen. *Biochem. Biophysical Res. Commun.* 85 (1), 317–325. doi:10.1016/S0006-291X(78)80045-X

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Care Guidelines of Guangzhou University of Chinese Medicine.

AUTHOR CONTRIBUTIONS

LD and AJ designed the project. LD, AJ, and TA participated in manuscript revision. GL wrote the original manuscript. GL, TA, YL, JY, and RH participated in figure preparation. JH collected the samples. WY did the RT-PCR. LH, YC, and RZ performed the data analysis.

FUNDING

This work was supported by the National Natural Science Foundation of China (No. 81874333) and the Guangdong Basic and Applied Basic Research Foundation (No. 2020A1515010926) and the Guangdong Key Laboratory for translational Cancer research of Chinese Medicine (No. 2018B030322011).

ACKNOWLEDGMENTS

We are grateful to Howard Riezman of the Department of Biochemistry at the University of Geneva for kindly providing the yeast strain RH6829.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.828877/full#supplementary-material>

- Efimova, V. S., Isaeva, L. V., Rubtsov, M. A., and Novikova, L. A. (2019). Analysis of *In Vivo* Activity of the Bovine Cholesterol Hydroxylase/lyase System Proteins Expressed in *Escherichia coli*. *Mol. Biotechnol.* 61, 261–273. doi:10.1007/s12033-019-00158-6
- Fedorova, O. V., Zernetkina, V. I., Shilova, V. Y., Grigороva, Y. N., Juhasz, O., Wei, W., et al. (2015). Synthesis of an Endogenous Steroidal Na Pump Inhibitor Marinobufagenin, Implicated in Human Cardiovascular Diseases, Is Initiated by CYP27A1 via Bile Acid Pathway. *Circ. Cardiovasc. Genet.* 8, 745–736. doi:10.1161/CIRCGENETICS.115.001217
- Gerber, A., Kleser, M., Biedendieck, R., Bernhardt, R., and Hannemann, F. (2015). Functionalized PHB Granules Provide the Basis for the Efficient Side-Chain Cleavage of Cholesterol and Analogs in Recombinant *Bacillus Megaterium*. *Microb. Cel Fact.* 14 (1), 107. doi:10.1186/s12934-015-0300-y
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De Novo transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis. *Nat. Protoc.* 8 (8), 1494–1512. doi:10.1038/nprot.2013.084
- Han, K.-Q., Huang, G., Gu, W., Su, Y.-H., Huang, X.-Q., and Ling, C.-Q. (2007). Anti-tumor Activities and Apoptosis-Regulated Mechanisms of Bufalin on the Orthotopic Transplantation Tumor Model of Human Hepatocellular Carcinoma in Nude Mice. *Wjg* 13 (24), 3374–3379. doi:10.3748/wjg.v13.i24.3374

- Hochberg, R. B., Ladany, S., Welch, M., and Lieberman, S. (1974). Cholesterol and Cholesterol Sulfate as Substrates for the Adrenal Side-Chain Cleavage Enzyme. *Biochemistry* 13 (9), 1938–1945. doi:10.1021/bi00706a025
- Hollman, A. (1985). Plants and Cardiac Glycosides. *Br. Heart J.* 54 (3), 258–261. doi:10.1136/hrt.54.3.258
- Kanehisa, M., and Neumann, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 28 (1), 27–30. doi:10.1093/nar/28.1.27
- Kuhl, H. (2011). Pharmacology of Progestogens. *J. Reproduktionsmed Endokrinol* 8 (1), 157–177. doi:10.1007/978-3-319-14385-9_2
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi:10.1093/molbev/msw054
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262
- McLean, K. J., Sabri, M., Marshall, K. R., Lawson, R. J., Lewis, D. G., Clift, D., et al. (2005). Biodiversity of Cytochrome P450 Redox Systems. *Biochem. Soc. Trans.* 33 (4), 796–801. doi:10.1042/BST0330796
- Meng, Z., Yang, P., Shen, Y., Bei, W., Zhang, Y., Ge, Y., et al. (2009). Pilot Study of Huachansu in Patients with Hepatocellular Carcinoma, Non-small-Cell Lung Cancer, or Pancreatic Cancer. *Cancer* 115 (22), 5309–5318. doi:10.1002/cncr.24602
- Nogawa, T., Kamano, Y., Yamashita, A., and Pettit, G. R. (2001). Isolation and Structure of Five New Cancer Cell Growth Inhibitory Bufadienolides from the Chinese Traditional Drug Ch'an Su. *J. Nat. Prod.* 64 (9), 1148–1152. doi:10.1021/np0101088
- Norlin, M., and Wikvall, K. (2007). Enzymes in the Conversion of Cholesterol into Bile Acids. *Cmm* 7 (2), 199–218. doi:10.2174/156652407780059168
- Pandey, A., Swarnkar, V., Pandey, T., Srivastava, P., Kanojija, S., Mishra, D. K., et al. (2016). Transcriptome and Metabolite Analysis Reveal Candidate Genes of the Cardiac Glycoside Biosynthetic Pathway from *Calotropis Procera*. *Sci. Rep.* 6, 34464. doi:10.1038/srep34464
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression. *Nat. Methods* 14 (4), 417–419. doi:10.1038/nmeth.4197
- Payne, A. H., and Hales, D. B. (2004). Overview of Steroidogenic Enzymes in the Pathway From Cholesterol to Active Steroid Hormones. *Endocr. Rev.* 25 (6), 947–970. doi:10.1210/er.2003-0030
- Porto, A. M., Baralle, F. E., and Gros, E. G. (1972). Biosynthesis of Bufadienolides in Toads. *J. Steroid Biochem.* 3 (1), 11–17. doi:10.1016/0022-4731(72)90006-4
- Porto, A. M., and Gros, E. G. (1970). Biosynthesis of Animal and Plant Bufadienolides. Parallel Experiments with pregn-5-en-3- β -ol-20-one-20- $_{14}C$ in *Scilla Maritima* and *Bufo Paracnemis*. *Experientia* 26 (1), 11–11. doi:10.1007/BF01900357
- Porto, A. M., and Gros, E. G. (1971). Biosynthesis of the Bufadienolide Marinobufagin in toads *Bufo Paracnemis* from cholesterol-20- $_{14}C$. *Experientia* 27 (5), 506. doi:10.1007/BF02147562
- Reuben, A. (1987). Sterols and Bile Acids. *Gastroenterology* 93 (1), 214–215. doi:10.1016/0016-5085(87)90349-0
- Russell, D. W., and Setchell, K. D. R. (1992). Bile Acid Biosynthesis. *Biochemistry* 31 (20), 4737–4749. doi:10.1021/bi00135a001
- Ryan, M. D., and Drew, J. (1994). Foot-and-mouth Disease Virus 2A Oligopeptide Mediated Cleavage of an Artificial Polyprotein. *EMBO J.* 13 (4), 928–933. doi:10.1002/j.1460-2075.1994.tb06337.x
- Sauer, H. H., Bennett, R. D., and Heftmann, E. (1967). Pregnenolone Metabolism in *Digitalis Lanata*. *Phytochemistry* 6 (11), 1521–1526. doi:10.1016/S0031-9422(00)82945-3
- Souza, C. M., Schwabe, T. M. E., Pichler, H., Ploier, B., Leitner, E., Guan, X. L., et al. (2011). A Stable Yeast Strain Efficiently Producing Cholesterol Instead of Ergosterol Is Functional for Tryptophan Uptake, but Not Weak Organic Acid Resistance. *Metab. Eng.* 13 (5), 555–569. doi:10.1016/j.ymben.2011.06.006
- Stache, U., Radsch, K., Fritsch, W., Kohl, H., Haede, W., and Ruschig, H. (1969). Synthesen von bufadienoliden synthese von scillarenin. *Tetrahedron Lett.* 10 (35), 3033–3038. doi:10.1016/S0040-4039(01)88337-4
- Strushkevich, N., MacKenzie, F., Cherkosova, T., Grabovec, I., Usanov, S., and Park, H.-W. (2011). Structural Basis for Pregnenolone Biosynthesis by the Mitochondrial Monooxygenase System. *Proc. Natl. Acad. Sci.* 108 (25), 10139–10143. doi:10.1073/pnas.1019441108
- Szcebara, F. M., Chandelier, C., Villeret, C., Masurel, A., Bourot, S., Duport, C., et al. (2003). Total Biosynthesis of Hydrocortisone from a Simple Carbon Source in Yeast. *Nat. Biotechnol.* 21 (2), 143–149. doi:10.1038/nbt775
- Tuckey, R. C., Slominski, A. T., Cheng, C. Y. S., Chen, J., Kim, T.-K., Xiao, M., et al. (2014). Lumisterol Is Metabolized by CYP11A1: Discovery of a New Pathway. *Int. J. Biochem. Cel Biol.* 55, 24–34. doi:10.1016/j.biocel.2014.08.004
- Watanabe, K., Mimaki, Y., Sakagami, H., and Sashida, Y. (2003). Bufadienolide and Spirostanol Glycosides from the Rhizomes of *Helleborus Orientalis*. *J. Nat. Prod.* 66 (2), 236–241. doi:10.1021/np0203638
- Wu, C., Zhang, Y., Chai, L., and Wang, H. (2017). Oxidative Stress, Endocrine Disruption, and Malformation of *Bufo gargarizans* Embryo Exposed to Sub-lethal Cadmium Concentrations. *Environ. Toxicol. Pharmacol.* 49, 97–104. doi:10.1016/j.etap.2016.12.005
- Xu, D., Wu, M., Li, X., Xia, M., Liu, D., Dai, Y., et al. (2018). Cloning, Prokaryotic Expression and Function of the *Bufo bufo* gargarizans 3β -Hydroxysteroid Dehydrogenase (3β HSD) Gene. *Int. J. Biol. Macromolecules* 120, 673–682. doi:10.1016/j.ijbiomac.2018.08.165
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zhang, C.-W., Zhang, X.-L., Xia, Y.-J., Cao, Y.-X., Wang, W.-J., Xu, P., et al. (2012). Association between Polymorphisms of the CYP11A1 Gene and Polycystic Ovary Syndrome in Chinese Women. *Mol. Biol. Rep.* 39 (8), 8379–8385. doi:10.1007/s11033-012-1688-7
- Zhang, R., Zhang, Y., Wang, Y., Yao, M., Zhang, J., Liu, H., et al. (2019). Pregnenolone Overproduction in *Yarrowia Lipolytica* by Integrative Components Pairing of the Cytochrome P450sc System. *ACS Synth. Biol.* 8 (12), 2666–2678. doi:10.1021/acssynbio.9b00018
- Zhang, Y., Li, X., Xu, D., Wu, M., Dai, Y., Xia, M., et al. (2021). Cloning and Characterization of Steroid 5β -Reductase from the Venom Gland of *Bufo bufo* gargarizans. *Int. J. Biol. Macromolecules* 175, 67–78. doi:10.1016/j.ijbiomac.2021.01.211

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

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, An, Li, Yue, Huang, Huang, Liang, Yao, Huang, Chen, Zhang, Ji and Duan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.