



## OPEN ACCESS

## EDITED BY

Frederic J.J. Chain,  
University of Massachusetts Lowell,  
United States

## REVIEWED BY

Adam Stuckert,  
University of New Hampshire,  
United States  
Shubo Jin,  
Freshwater Fisheries Research Center  
(CAFS), China  
Jennifer Stynoski,  
University of Costa Rica, Costa Rica

## \*CORRESPONDENCE

Yunyun Lv  
lvyunyun@njtc.edu.cn  
Qiong Shi  
shiqiong@genomics.cn

†These authors have contributed  
equally to this work

## SPECIALTY SECTION

This article was submitted to  
Evolutionary and Population Genetics,  
a section of the journal  
Frontiers in Ecology and Evolution

RECEIVED 20 April 2022

ACCEPTED 14 July 2022

PUBLISHED 28 July 2022

## CITATION

Lv Y, Li Y, Wen Z and Shi Q (2022)  
Transcriptomic and gene-family  
dynamic analyses reveal gene  
expression pattern and evolution  
in toxin-producing tissues of Asiatic  
toad (*Bufo gargarizans*).  
*Front. Ecol. Evol.* 10:924248.  
doi: 10.3389/fevo.2022.924248

## COPYRIGHT

© 2022 Lv, Li, Wen and Shi. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Transcriptomic and gene-family dynamic analyses reveal gene expression pattern and evolution in toxin-producing tissues of Asiatic toad (*Bufo gargarizans*)

Yunyun Lv<sup>1\*†</sup>, Yanping Li<sup>1†</sup>, Zhengyong Wen<sup>1</sup> and  
Qiong Shi<sup>1,2\*</sup>

<sup>1</sup>Key Laboratory of Sichuan Province for Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River, College of Life Sciences, Neijiang Normal University, Neijiang, China,

<sup>2</sup>Shenzhen Key Lab of Marine Genomics, Guangdong Provincial Key Lab of Molecular Breeding in Marine Economic Animals, BGI Academy of Marine Sciences, BGI Marine, Shenzhen, China

Comprising a major clade of Anura, toads produce and secrete numerous toxins from both the parotoid glands behind their eyes and their dorsal skin. These toxins, made of various proteins and compounds, possess pharmacological potential to be repurposed to benefit human health. However, the detailed genetic regulation of toad toxin production is still poorly understood. A recent publication uncovering the genome of the representative Asiatic toad (*Bufo gargarizans*) provides a good reference to resolve this issue. In the present study, we sequenced the transcriptomes of parotoid gland, dorsal skin and liver from the Asiatic toad. Combining our data with 35 previously published transcriptomes across eight different tissues from the same species but from different locations, we constructed a comprehensive gene co-expression network of the Asiatic toad with the assistance of the reference genome assembly. We identified 2,701 co-expressed genes in the toxin-producing tissues (including parotoid gland and dorsal skin). By comparative genomic analysis, we identified 599 expanded gene families with 2,720 genes. Through overlapping these co-expressed genes in the toad toxin-producing tissues, we observed that three cytochrome P450 (*Cyp*) family members (*Cyp27a1*, *Cyp2c29*, and *Cyp2c39*) were significantly enriched in pathways related to cholesterol metabolism. Cholesterol is a critical precursor to steroids, and the known main steroidal toxins of bufadienolides are considered as the major bioactive components in the parotoid glands of Asiatic toad. We found 3-hydroxy-methylglutaryl CoA reductase (*hmgcr*), encoding the major rate-limiting enzyme for cholesterol biosynthesis, appears with multiple copies in both Asiatic toad and common toad, possibly originating from a tandem duplication event. The five copies of *hmgcr* genes consistently displayed higher transcription levels in the parotoid

gland when compared with the abdominal skin, suggesting it as a vital candidate gene in the involvement of toad toxin production. Taken together, our current study uncovers transcriptomic and gene-family dynamic evidence to reveal the vital role of both expanded gene copies and gene expression changes for production of toad toxins.

#### KEYWORDS

toxin, transcriptome, *Cyp*, *hmgcr*, gene family

## Introduction

Amphibians are one of the most diverse group among vertebrates that exhibit high importance for human beings, including medical applications such as tissue regeneration, pharmaceutically useful compounds, direct socio-economic benefits, and overall ecosystem services (West, 2018). Many amphibian species are able to use their chemical defenses including toxic alkaloids, cardiogenic steroids, and antimicrobial peptides as weapons to deter their predators and to protect themselves (Shibao et al., 2018; Posso-Terranova and Andrés, 2020).

In true toads (Anura: Bufonidae), the chemical defense secretion is comprised of several components including cardiac glycosides (bufadienolides or bufotoxins), alkaloids (bufotenine) and biogenic amines, and/or antimicrobial peptides (Baldo et al., 2019; Li et al., 2021). Evidence has shown that some of these toxins were effective in the treatment of chronic pain, heart disorders, and cancer (Qi et al., 2018; Li et al., 2021). Bufonidae skin secretion extracts also show inhibitory effects *in vitro* against clinical isolates of bacteria, as well as resistant and standard strains of bacterial, fungal and parasitic pathogens (Rodriguez et al., 2020).

The Asiatic toad (*Bufo gargarizans*), a true toad, is characterized by a pair of well-developed parotoid glands and numerous granular glands in dorsal skin, which give it the capacity to produce bufotoxins (Cao et al., 2019; Firreno et al., 2022). It has a wide distribution in East Asia, and is readily collected for experimental purposes on various studies of physiological and biochemical characteristics of toad toxins (Cao et al., 2019), thus it is a good animal model for understanding of chemical constituents and metabolic pathways relevant to toad poison synthesis. The chemical constituents in parotoid glands of the Asiatic toad includes peptides, steroids, indole alkaloids, bufogargarizanines, organic acids and others (Wang et al., 2011; Tian et al., 2013; Chen et al., 2020). Among them, steroids are the predominant constituents that are mainly comprised of bufadienolides and cholesterol (Wang et al., 2011).

Cholesterol act as a precursor of steroid hormones leading to the formation of multiple steroid hormones such as cortisol,

testosterone, and progesterone as well as vitamin-D3 (Nelson et al., 2008; Rone et al., 2009). Cholesterol are also thought to be the precursors for production of bufadienolides, which are considered as the major bioactive components in the skin of Asiatic toad (Porto and Gros, 1970, 1971; Porto et al., 1972; Flier et al., 1980; Santa Coloma et al., 1984; Garraffo and Gros, 1986). Thus, steroid biosynthesis should be important for toad-toxin production as evidenced by recent investigations of mRNA expression among 10 true toads (Firreno et al., 2022).

Other recent studies have focused on the evolutionary innovation of various genes associated with toxin-producing pathways (Huang et al., 2016; Firreno et al., 2022). However, such studies have not yet identified a complete picture of positively selected genes directly related to toxin production, implying that some other genetic mechanisms such as gene family changes may have evolved for toxin production. In contrast to some amphibians that are unable to synthesize toxins, investigations on gene-family dynamics related to toxin-producing tissues in toads have not yet been performed, providing no evidence to support or reject the hypothesis. In fact, only a few datasets of high-quality amphibian genome assemblies and gene annotations are publicly available. In this case, previous studies using comparative transcriptomes to reveal differential expression patterns of toxin-producing genes were based on a non-reference genome approach (Lan et al., 2022), but surprisingly found only a weak relationship with steroid metabolism that is responsible for production of the primary bioactive constituents. This result was possibly affected by the absence of a high-quality gene annotation, which may lead to inaccurate conclusions. Thus, the key genes potentially responsible for toad toxin production need in-depth explorations, especially those genes in many families with significant changes in copies should be paid more attention, since they may provide new insights into genetic mechanisms for controlling toad toxin-producing.

Recently, the genome of the representative Asiatic toad (*B. gargarizans*) was sequenced and annotated (Lu et al., 2021), providing a good opportunity to reveal the genetic differences between toads and frogs, as well as a substantial pool of genes to investigate gene-family dynamics related to toxin production. In the current study, we performed transcriptome sequencing

of three important tissues, including liver, dorsal skin, and parotoid gland, which were collected from an individual toad in Southern China. In combination with early released 35 transcriptomes, we organized a total dataset including 38 transcriptomes derived from Asiatic toad to align all available transcripts to compare against the assembled Asiatic toad genome. The main aim is to construct a co-expression network for the identification of core genes in the regulation of toad toxin production. We also analyzed some expanded gene families caused by duplication for their potential roles in gene regulation, supporting a genetic basis for toad toxin generation. Our comparative transcriptomic analysis of toad toxin-producing tissues based on the reference genome offers the present study a distinct advantage in contrast to previous transcriptomic studies without this genomic reference, providing not only novel insights into steroid biosynthesis for production of the main bioactive constituents bufotoxins in Asiatic toad, but also some key candidate genes for studies of controlling toxin production in other amphibians.

## Materials and methods

### Sampling, transcriptome sequencing, and quantification

Parotoid gland, dorsal skin and liver were sampled from an adult female Asiatic toad, which was purchased from a local market in Shenzhen, Guangdong, China. These tissues were individually frozen in liquid nitrogen for 10 min, and then stored at  $-80^{\circ}\text{C}$  before use.

Total ribonucleic acid (RNA) was extracted using PureLin RNA Mini Kit (Thermo Fisher Scientific, Waltham, MA, United States) in accordance with the manufacturer's protocol. RNA quantity, purity and integrity were assessed by a Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States), a Nanodrop spectrophotometer (Nanodrop technologies Inc., Wilmington, DE, United States) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, United States), respectively. Those samples with RNA integrity number (RIN)  $\geq 8$  were selected for sequencing. A Taq PCR core kit (Qiagen, Germantown, MD, United States) was used for cDNA library construction and subsequent breaking of RNAs into fragments and PCR amplification. An Illumina HiSeq 4000 platform (San Diego, CA, United States) was applied to generate transcriptome (RNA-seq) data with paired-end reads at a length of 150 bp.

In addition to the three sequenced samples, we searched available transcriptome data from public databases, including National Center for Biotechnology Information (NCBI) and China Nucleotide Sequence Archive (CNSA). These data, obtained from different resources including parotoid glands ( $n = 10$ ), dorsal skin ( $n = 10$ ), liver ( $n = 1$ ), abdomen

skin ( $n = 10$ ), follicle tissue ( $n = 1$ ), heart ( $n = 1$ ), muscle ( $n = 1$ ) and spleen ( $n = 1$ ), permitted us to reach a more reliable consensus conclusion for the regulation of toad toxin production. Meanwhile, the reference genome of the Asiatic toad and related gene annotations were downloaded from the NCBI (GCF\_014858855.1) for RNA-seq alignment. The longest transcripts for the corresponding genes were retained as representatives for calculation of transcription levels.

### Ribonucleic acid quantification and co-expression network construction

For RNA quantification, we aligned each transcriptome dataset to the Asiatic toad reference genome using Hisat2 version 2.2.1 (Kim et al., 2019) with default parameters. The TPM (transcripts per million) index, calculated by Stringtie version 2.1.4 (Pertea et al., 2015) using default parameters setting, was transformed with different resource studies as covariate to decrease batch effect and then used to represent relative transcription level of each gene.

For construction of a co-expression network, we firstly built a trait matrix according to the source of each sample. For the toad toxin-producing tissues (including parotoid gland and dorsal skin) we merged them under one group denominated "toad toxin tissues," marking dorsal skin as "1" and parotoid gland as "2." Meanwhile, other samples were treated as "0" to represent those without toxin-producing qualities.

Secondly, TPM values were integrated to construct a co-expression network using WGCNA version 1.69 in the R package (Langfelder and Horvath, 2008). The best scale-free R2 fit was chosen according to the most suitable threshold in network building. A topological overlap dendrogram was used to define modules with a minimum module size of 120 genes, and the threshold for merging modules was set to 0.25. Genes displaying similar transcription levels were clustered into the same co-expression module according to the trait matrix. Module eigengenes, defined as the first principal component of the expression matrix of the corresponding module, were subsequently calculated. We identified the module with the highest correlations as being toad toxin-related by calculating Pearson's correlation coefficient between module eigengene and toad skin. Genes within the module were extracted and annotated with Gene Ontology (GO) terms (Boyle et al., 2004) using GOATOOLS (Klopfenstein et al., 2018). Process over-representation of certain GO terms was based on a Fisher's exact test. Corrected  $p$ -value was generated by correction of false discovery rate (FDR) through the BH (Benjamini-Hochberg) procedure. These enriched GO terms (corrected  $p$ -value  $< 0.05$ ) were then used to explore functional changes within the parotoid gland and dorsal skin. We defined "enrichment ratio" as the number ratio of GO terms in the studied group compared to those in the population.

## Identification of gene-family expansion

In addition to the Asiatic toad, we employed high-quality genome assemblies of three other amphibians and human to assist our analysis. The amphibian species included common toad (*Bufo bufo*; GCF\_905171765.1), common frog (*Rana temporaria*; GCF\_905171775.1), and high Himalaya frog (*Nanorana parkeri*; GCF\_000935625.1). Thus, we constituted a dataset of five species to compare two toads (the Asiatic toad and common toad) and two frogs (the common frog and Himalaya frog), with humans as the outgroup (**Supplementary Table 2**).

We firstly built a species tree of the five examined species using single-copy genes categorized by gene-family clustering using Orthofinder version 2.3.12 (Emms and Kelly, 2019). The inside transversions of four-fold degenerate (4dTV) sites without influence from natural selection were used to construct a phylogenetic tree by IQ-TREE version 1.6.12 (Nguyen et al., 2015) with 10,000 super-fast bootstraps. With the assistance of two putative time-calibrate points, the human-amphibian split [347~358 million years ago (mya)] and the frog-toad split (145~160 mya) estimated by TimeTree (Kumar et al., 2017), we calibrated the species tree at a time scale by using MCMCtree version 4.9 (Yang, 2007). A log-normal independent molecular clock model was established for calculating the entire collection of 1,000,000 samples, while the former 25% samples were discarded.

We refined the included gene family clusters by removing some possible false results in which some species had more than 200 paralogs whereas other species were blank. The retained gene clusters were used for analyzing their member changes with CAFÉ version 2.0 (De Bie et al., 2006) with a significance level ( $p \leq 0.05$ ). We specifically extracted the expanded gene families that occurred in the common ancestor of common toad and Asiatic toad, and then enriched them to predict their main functions (represented by GO terms). Some expanded families were extracted from the dataset to construct a phylogenetic framework for prediction of gene evolution. Those protein sequences from human were used as the outgroup, and the phylogeny was constructed as above-mentioned species tree. We applied 3 different student *t*-tests to examine the expressional differences of certain genes in paratoid gland vs dorsal skin, paratoid gland vs abdomen skin, and paratoid gland vs abdomen skin, respectively.

## Identification of expanded gene paralogs in association with toad toxin-production, and construction of a protein–protein interaction network

To reveal the expanded gene-family members in association with toad toxin-production, we overlapped these expanded members and the above-mentioned co-expression genes. The

main biological functions of the overlapped genes were enriched. Each gene within the vital GO term was further explored for evolutionary history, genomic localization, and transcriptional changes. We also imported protein sequences encoded by the overlapped genes into Search Tool for the Retrieval of Interacting Genes (STRING<sup>1</sup>) to examine protein relationships (Szklarczyk et al., 2019) with human as the representative species. Enriched pathways were chosen to display their main effects among the overlapped genes.

## Results

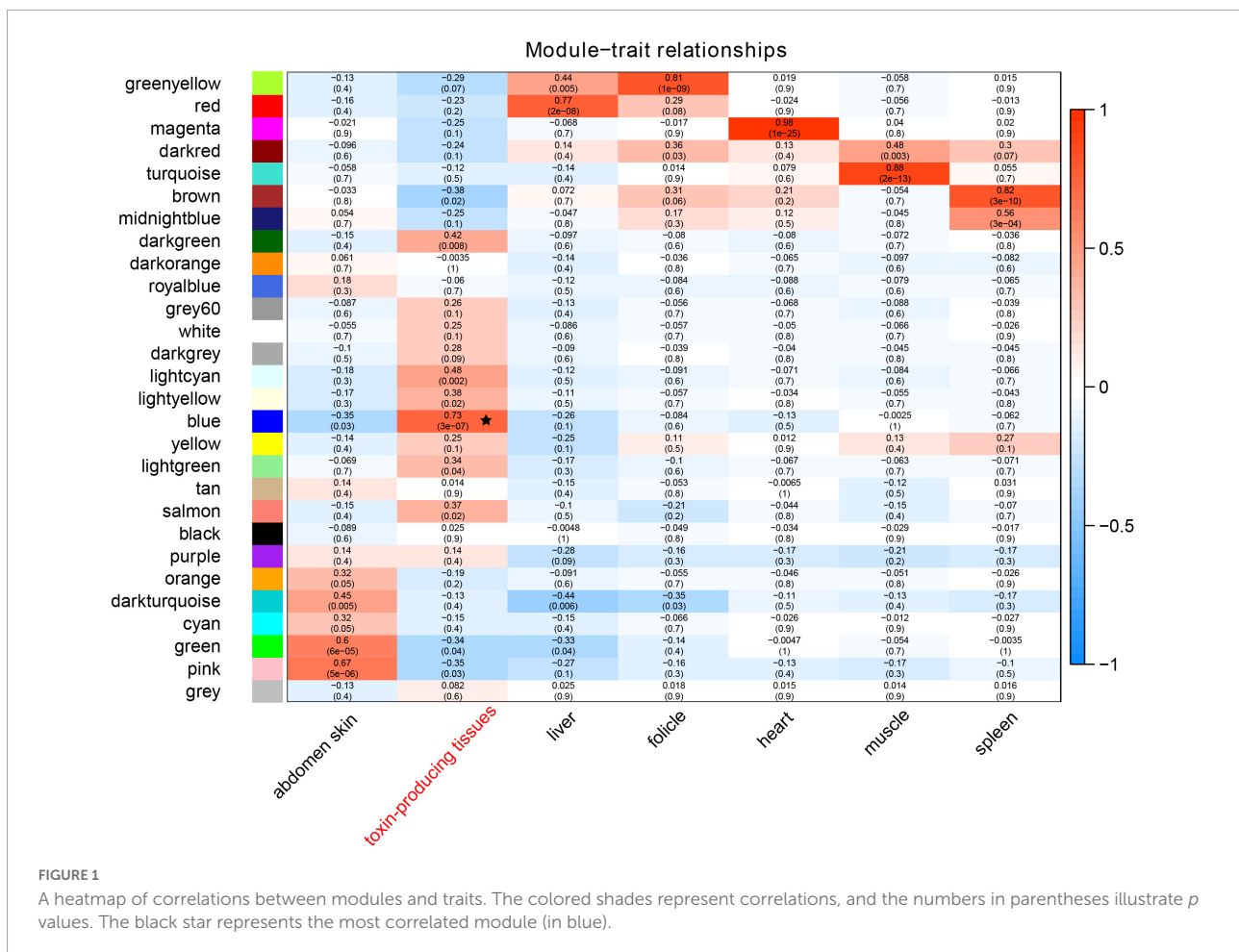
### Transcriptome data collection for construction of the co-expression network

Our transcriptome datasets contained 51,668,570 reads (7,750,285,500 bp) in the paratoid gland, 38,894,474 reads (5,834,171,100 bp) in the dorsal skin, and 44,857,778 reads (6,728,666,700 bp) in the liver, respectively. In addition to the three samples, we searched and downloaded 35 previously published transcriptomes derived from paratoid gland, dorsal skin, abdominal skin, follicle, heart, muscle, liver, and spleen from NCBI and CNSA (See **Supplementary Table 1**). In sum, we organized a total of 38 transcriptomes covering eight different tissues for our in-depth RNA analysis.

Within RNA alignments to the reference genome, a total mapping ratio of 78~89% appeared in each transcriptome dataset. Every expressional matrix (represented by a TMP index) was transformed and integrated for construction of the co-expression network. Gene connectivity in the network exhibited a good performance when the correlation coefficient threshold was set at 0.9 and the best soft-thresholding power was set to 7 (**Supplementary Figure 1A**). Those modules with high correlation ( $R > 0.8$ ) were merged to a single cluster (**Supplementary Figure 1B**). These resulting genes were allocated into 28 co-expression modules, which were independent from each other (**Figure 1**). Interestingly, the blue module (marked by a star) containing 2,701 genes displayed the highest correlation to the toad toxin-producing tissues.

We extracted those genes and further calculated the coefficient of association between gene significance and module membership in the blue module, which was at 0.61 with a *p*-value of  $1e-200$  (**Supplementary Figure 1C**). The main biological functions of these genes related to the processes of intracellular transport, establishment of protein localization, and nitrogen compound transport (left panel of **Figure 2**). The primarily effected cellular components were the vesicle coat, oligosaccharyltransferase complex, and intrinsic component of

<sup>1</sup> <http://string-db.org>



peroxisomal membrane (middle panel of **Figure 2**). The largest effects on molecular functions pertained to protein deacetylase activity, basal RNA polymerase II transcription machinery binding, and RNA polymerase binding (right panel of **Figure 2**).

## Summary of expanded gene families

By gene-family clustering of the Asiatic toad, common toad, common frog and High Himalaya frog, and subsequent purging of possible false clusters, we obtained a total of 17,485 gene families in the final gene-family dynamic analysis. Among them, 599 gene families were expanded in the ancestral branch of toad (**Figure 3A**). These families contained a total of 2,720 genes identified within the Asiatic toad.

Through functional enrichment of these genes, we observed that the top three significant GO terms were related to biological processes, cellular components, and molecular functions, corresponding to “egg coat formation, sertoli cell barrier remodeling and cell death in response to oxidative stress,” “organism cell membrane, nucleosome and *NLRP1* inflammasome complex related to cellular component,” and

“structural constituent of egg coat, oxygen binding and neurotransmitter receptor regulator activity related to molecular function,” respectively (**Figure 3B**).

## Association of expanded candidate genes with regulation of toad toxin-production

The 2,701 gene family members and 2,720 genes associated with toad toxin-producing tissues were overlapped. A total of 218 genes were shared between shared between both bioinformatic processes (**Figure 4A**). They were enriched into several essential metabolic processes, such as lipoxin metabolism, steroid metabolism, and alcohol metabolism (**Figure 4B**). Among the overlapped genes, *Cyp27a1*, *Cyp2c29* and *Cyp2c39* belonging to the *Cyp* family, were significantly enriched in the process of cholesterol metabolism. In addition, we identified a tandem-duplicate gene cluster of 3-hydroxy-methylglutaryl CoA reductase (*hmgcr*) in both the Asiatic toad and the common toad (**Figures 4C,D**), although more

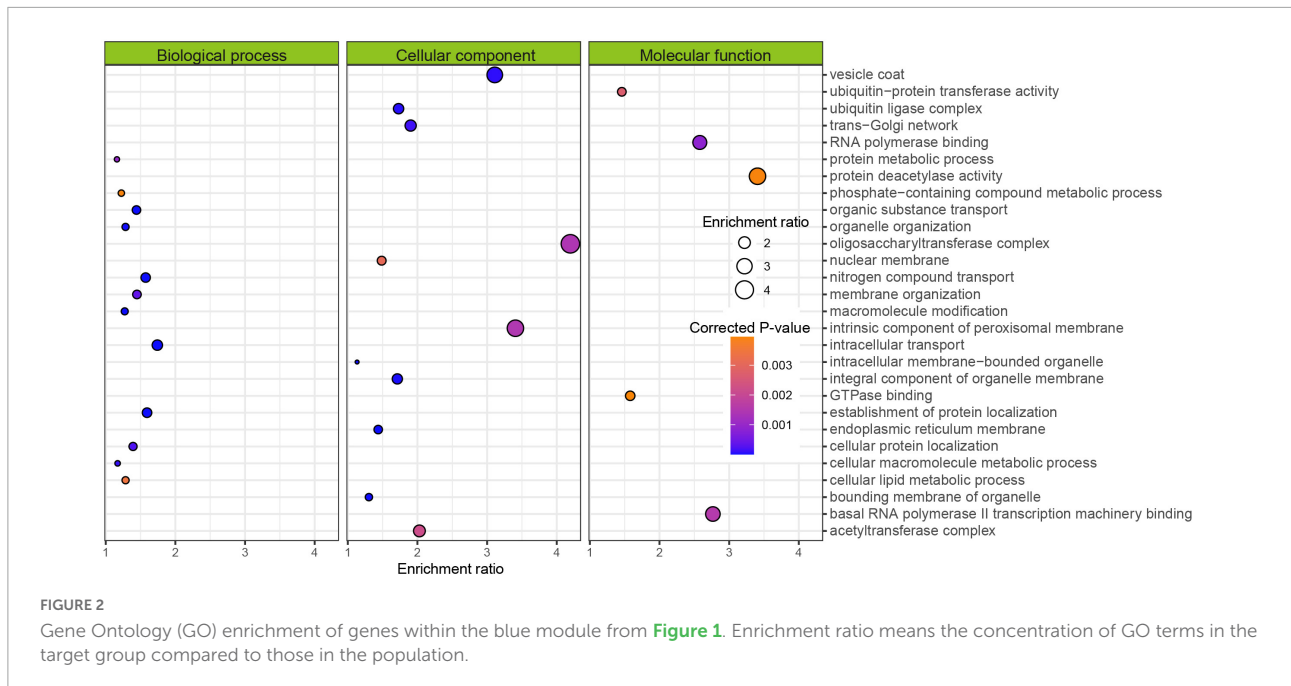


FIGURE 2

Gene Ontology (GO) enrichment of genes within the blue module from Figure 1. Enrichment ratio means the concentration of GO terms in the target group compared to those in the population.

investigations are required to exclude errors of the released assemblies or annotations.

This tandem duplication event may have occurred after the ancestral split because the duplicates from either toad were clustered into an independent clade with high supports in the phylogenetic inference (Figure 4C). Interestingly, we found that the five *hmgcr* copies consistently displayed higher transcription levels in both the parotoid gland and the dorsal skin over the abdominal skin (Figure 4E), and a significantly higher transcription level in the parotoid gland than the abdominal skin. By further searching protein-protein interaction (PPI) relationships, we observed that the *hmgcr*-encoded protein (Hmgcr) and other proteins encoded by the overlapped genes exhibited an obvious interaction, mainly contributing to steroid or cholesterol biosynthesis (Figure 5). For example, Hmgcr, Fdft1, Slc27a2, Ces1d, Akr1c6, Hsd17b12, and Cyp27a1 participate in the steroid biosynthetic process, while Hmgcr, Fdft1, Hsd17b12, Cyp27a1, Soat2, Slc27a2, Sult2b1, Ces1d, and Akr1c6 are involved in cholesterol metabolism.

## Discussion

Recently, increasingly more vertebrate genomes have been sequenced with related assemblies available for academic studies (Rhie et al., 2021). However, publicly available amphibian genomes are still fewer in number than those of other groups (like fishes, birds, and mammals), limiting our understanding of genetic mechanisms of such innovative features as toad toxin-producing tissues. The recently released Asiatic toad genome provides a crucial genetic resource to reveal gene-family

expansions associated with the perception of smell and bitter taste (Lu et al., 2021). In this paper, the *Cyp2c* subfamily that performs a vital role in detoxification was identified as a genetic expansion in the Asiatic toad. Our results in the present study further confirmed the expansion of the *Cyp* gene family, which appeared not only in the Asiatic toad but also in the common toad (Figure 3A).

Additionally, a previous non-reference comparative transcriptomic analysis (Lan et al., 2022) revealed a series of differentially expressed genes (DEGs) in association with toad toxin-production and identified the *Cry2c* subfamily to likely be involved in the regulation of toad toxin generation. In the present study, our co-expression network confirmed that some of the *Cyp* members (such as *Cyp27a1*, *Cyp2c29*, and *Cyp2c39*) were associated with toad toxin-producing tissues (Figure 5). In particular, these genes were expanded with more copies. These additional copies may produce neofunctionalization or subfunctionalization, which could have provided the genetic basis for new functional generation. For example, one primary active component of toad toxins is Bufalin, a cardiotonic steroid (Gao et al., 2017). The *Cyp* family is responsible for steroid hormone biosynthesis; therefore, it is reasonable to speculate that some of the expanded *Cyp* gene members may be crucial for Bufalin biosynthesis, suggesting that their mRNA regulation is essential for toad toxin-producing tissues.

Cholesterol is thought to serve as an essential precursor in the production of several steroid hormones including cortisol, testosterone and progesterone by providing the core ring system in chemical structure (Gesto et al., 2020). Cholesterol synthesis occurs in the cytoplasm and the major locations are liver and intestinal tissues (Hussain et al., 1999; Willnow, 1999).

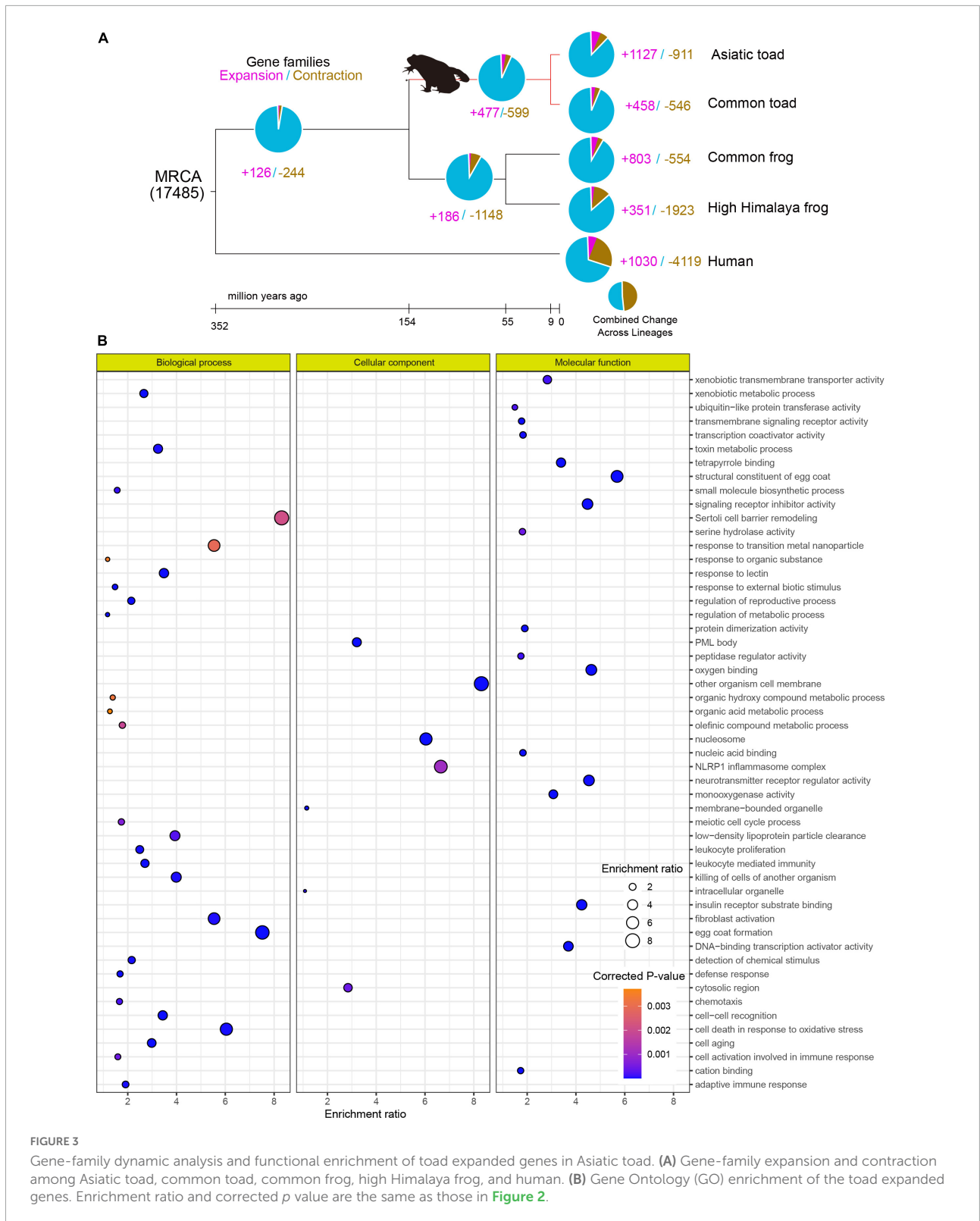


FIGURE 3

Gene-family dynamic analysis and functional enrichment of toad expanded genes in Asiatic toad. (A) Gene-family expansion and contraction among Asiatic toad, common toad, common frog, high Himalaya frog, and human. (B) Gene Ontology (GO) enrichment of the toad expanded genes. Enrichment ratio and corrected *p* value are the same as those in Figure 2.

The process involves several steps, in which the rate-limiting step of the synthesis known as the committed step using Hmgcr (Hmg-CoAR) as a regulatory enzyme

(Sharpe and Brown, 2013; Gesto et al., 2020). As for its importance for cholesterol synthesis, it was a target enzyme for prevention of cardiovascular diseases by using statins as the

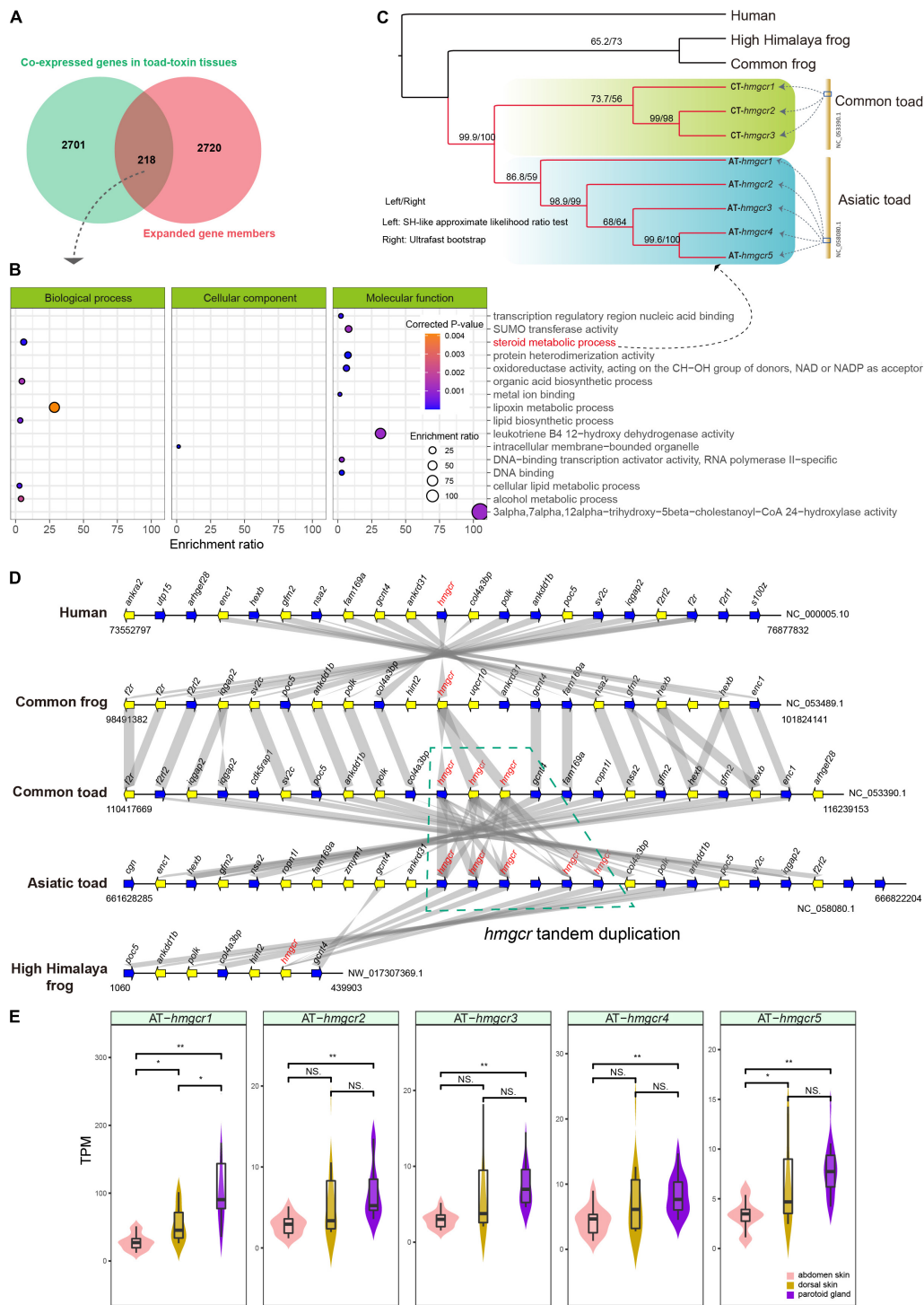
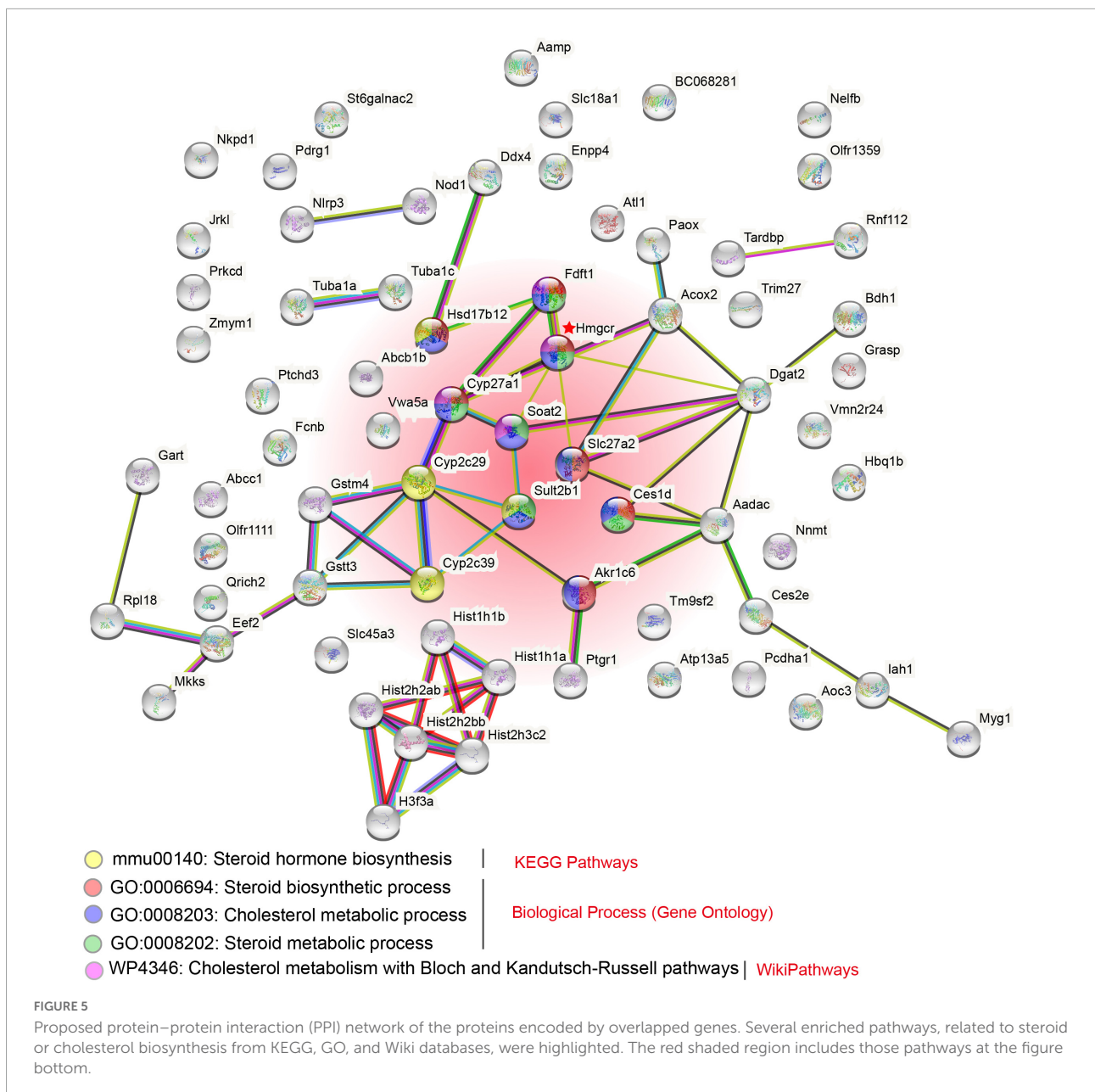


FIGURE 4

Overlapping of co-expressed genes in toad toxin producing tissues and expanded gene members. (A) The 218 overlapped genes. (B) Gene Ontology (GO) enrichment of the overlapped genes. (C) The phylogeny of multiple *hmgcr* copies within the GO terms of steroid metabolic process among Asiatic toad (five copies including AT-*hmgcr1*, AT-*hmgcr2*, AT-*hmgcr3*, AT-*hmgcr4* and AT-*hmgcr5*), common toad (three copies including CT-*hmgcr1*, CT-*hmgcr2* and CT-*hmgcr3*), common frog, high Himalaya frog, and human. (D) Landscape of *hmgcr* copies and their neighbor genes among the five examined species. Gene syntenic relationships were represented by the gray bands; the green polygon means the tandem duplicates within the two toads, whereas other species have only one *hmgcr* copy. (E) Transcriptional comparisons of the five *hmgcr* copies among abdominal skin, dorsal skin and parotoid glands of Asiatic toad based on a student *t*-test. “\*” and “\*\*” represent the significant levels less than 0.05 and 0.01, respectively. “NS” stands for “no significant difference.”





most commonly prescribed drugs because it sterically inhibits cholesterol formation (Colca and Kletzien, 2009; Stormo et al., 2012; Gesto et al., 2020).

In our present study, we identified the tandem-duplicated five and three *hmgcr* genes in the Asiatic toad and the common toad, respectively, whereas other examined species (human, the common frog, and the High Himalaya frog) had only one copy (see Figure 4D). Meanwhile, within the Asiatic toad, the five *hmgcr* genes consistently displayed higher transcription levels in the parotoid gland as compared to the abdominal skin (Figure 4E). In the PPI analysis, *Hmgcr* was predicted to be involved in both steroid biosynthesis and cholesterol metabolism based on the database in combination

to other important proteins such as *Fdft1*, *Slc27a2*, *Ces1d*, *Akr1c6*, *Hsd17b12*, and *Cyp27a1* for the steroid biosynthetic process, and *Fdft1*, *Hsd17b12*, *Cyp27a1*, *Soat2*, *Slc27a2*, *Sult2b1*, *Ces1d*, and *Akr1c6* for the cholesterol metabolic process. These interactions list crucial target genes for understanding the regulation of toad toxin production. Given that the most bioactive bufotoxins are considered to be the steroids, we speculate the tandem duplication of *hmgcr* gene encoding multiple HMGCR proteins enable the toads to enhance cholesterol synthesis, which seems to be an evolutionary innovation for toad toxin production. Its higher expression in the toxin-producing parotoid gland should be relevant to production of those poisons; however, a low expression in

the abdominal skin suggests that this tissue may be unable to produce poisons. Whereas previous studies have not succeeded in identifying genes under positive selection that are associated with toad toxin-producing tissues (Huang et al., 2016; Firreno et al., 2022), our current works emphasize the gene-family changes such as tandem duplication may be responsible for the occurrence of toad toxin production. Taken together, our current study reveals that both evolutionary expansion of gene copies and gene regulation patterns are vital for toad toxin-producing tissues, which are important for the useful functions of new biological features.

## Conclusion

A functional innovation presented in various toads is the production of toxins by the important parotoid gland, holding promises as pharmaceutical agents with the potential to benefit human health; however, it currently remains poorly characterized in the context of its gene regulation. Our current study organized a large dataset of 38 transcriptomes from different resources, including three generated in the present study, and revealed 2,701 general genes related to toad toxin-producing tissues. Some genes, such as *Cyp* members, overlapped with the expanded genes and were significantly enriched into several important biosynthetic pathways, illustrating the importance of these genes for the regulation of toad toxin production and secretion. Interestingly, multiple copies of the *hmgcr* gene were identified in both the common toad and the Asiatic toad, and those copies (5) within the Asiatic toad were transcriptionally higher in toxin-producing tissues than in others, suggesting that they are a vital contributor to the control of toad toxin production. In summary, our present study proposes that both gene-family expansion and gene expression are important for the formation of toad toxin-producing tissues, thereby revealing molecular clues for the regulation of toad toxin-production.

## Data availability statement

The datasets presented in this study can be found in online repositories. The name of the repository and accession number can be found below: National Center for Biotechnology Information (NCBI), <https://www.ncbi.nlm.nih.gov/bioproject/>, PRJNA825169.

## Ethics statement

The animal study was reviewed and approved by Neijiang Normal University.

## Author contributions

YYL and QS conceived and designed the project. YYL, YPL, and ZW analyzed the data. YYL wrote the manuscript. QS revised the manuscript. All authors have read and approved the final manuscript.

## Funding

This study was supported by Shenzhen Science and Technology Innovation Program for International Cooperation (no. GJHZ20190819152407214 to QS), Open Fund of Key Laboratory of Sichuan Province for Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River (NJTCSC07 to YYL), and Neijiang Normal University Innovation Team Fund (2021TD03 to YYL).

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2022.924248/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Co-expression network construction and identification of co-expressed genes in the toad-toxin produced tissues. (A) Selection of the best scale-free R2 fit. (B) The cluster dendrogram of genes. Modules with high similarity were merged. (C) The scatterplot of gene significance (GS) for weight vs Module Membership (MM) in the blue module. Obviously, there is a highly significant correlation between GS and MM.

### SUPPLEMENTARY TABLE 1

Summary of the 38 transcriptomes used for co-expression network construction.

### SUPPLEMENTARY TABLE 2

Genome assemblies used for our comparative genomic analysis.

## References

- Baldo, M. A., Cunha, A. O. S., Godoy, L. D., Liberato, J. L., Yoneda, J. S., Fornari-Baldo, E. C., et al. (2019). Assessment of neuropharmacological potential of low molecular weight components extracted from *Rhinella schneideri* toad poison. *J. Venom. Anim. Toxins Includ. Trop. Dis.* 25:e148418. doi: 10.1590/1678-9199-JVATITD-1484-18
- Boyle, E. I., Weng, S., Gollub, J., Jin, H., Botstein, D., Cherry, J. M., et al. (2004). GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20, 3710–3715. doi: 10.1093/bioinformatics/bth456
- Cao, Y., Cui, K., Pan, H., Wu, J., and Wang, L. (2019). The impact of multiple climatic and geographic factors on the chemical defences of Asian toads (*Bufo gargarizans* Cantor). *Sci. Rep.* 9:17236. doi: 10.1038/s41598-019-52641-4
- Chen, Y.-L., Dai, Y.-H., Wang, A.-D., Zhou, Z.-Y., Lei, M., Liu, J., et al. (2020). Two New Indole Alkaloids from Toad Venom of *Bufo bufo gargarizans*. *Molecules* 25:4511. doi: 10.3390/molecules25194511
- Colca, J. R., and Kletzien, R. F. (2009). Current and emerging strategies for treating dyslipidemia and macrovascular disease. *Adv. Pharmacol.* 57, 237–251. doi: 10.1016/S1054-3589(08)57006-2
- De Bie, T., Cristianini, N., Demuth, J. P., and Hahn, M. W. (2006). CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22, 1269–1271. doi: 10.1093/bioinformatics/btl097
- Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 238. doi: 10.1186/s13059-019-1832-y
- Firreno, T. J., Ramesh, B., Maldonado, J. A., Hernandez-Briones, A. I., Emery, A. H., Roelke, C. E., et al. (2022). Transcriptomic analysis reveals potential candidate pathways and genes involved in toxin biosynthesis in true toads. *J. Hered.* 113, 311–324. doi: 10.1093/jhered/esac015
- Flier, J., Edwards, M. W., Daly, J. W., and Myers, C. W. (1980). Widespread occurrence in frogs and toads of skin compounds interacting with the ouabain site of Na<sup>+</sup>, K<sup>+</sup>-ATPase. *Science* 208, 503–505. doi: 10.1126/science.6245447
- Gao, F., Wang, X., Li, Z., Zhou, A., Tiffany-Castiglioni, E., Xie, L., et al. (2017). Identification of anti-tumor components from toad venom. *Oncol. Lett.* 14, 15–22. doi: 10.3892/ol.2017.6160
- Garraffo, H. M., and Gros, E. G. (1986). Biosynthesis of bufadienolides in toads. VI. experiments with [1, 2-<sup>3</sup>H] cholesterol, [21-<sup>14</sup>C] coprostanol, and 5 $\beta$ -[21-<sup>14</sup>C] pregnanolone in the toad *Bufo arenarum*. *Steroids* 48, 251–257. doi: 10.1016/0039-128X(86)90008-5
- Gesto, D. S., Pereira, C. M., Cerqueira, N. M., and Sousa, S. F. (2020). An atomic-level perspective of HMG-CoA-reductase: the target enzyme to treat hypercholesterolemia. *Molecules* 25:3891. doi: 10.3390/molecules25173891
- Huang, L., Li, J., Anbukaria, H., Luo, Z., Zhao, M., and Wu, H. (2016). Comparative transcriptome analyses of seven anurans reveal functions and adaptations of amphibian skin. *Sci. Rep.* 6:24069. doi: 10.1038/srep24069
- Hussain, M. M., Strickland, D. K., and Bakillah, A. (1999). The mammalian low-density lipoprotein receptor family. *Ann. Rev. Nutr.* 19, 141–172. doi: 10.1146/annurev.nutr.19.1.141
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. doi: 10.1038/s41587-019-0201-4
- Klopfenstein, D., Zhang, L., Pedersen, B. S., Ramírez, F., Warwick Vesztrocy, A., Naldi, A., et al. (2018). GOATOOLS: a python library for gene ontology analyses. *Sci. Rep.* 8:10872. doi: 10.1038/s41598-018-28948-z
- Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819. doi: 10.1093/molbev/msx116
- Lan, Y., He, L., Dong, X., Tang, R., Li, W., Wang, J., et al. (2022). Comparative transcriptomes of three different skin sites for the Asiatic toad (*Bufo gargarizans*). *PeerJ* 10:e12993. doi: 10.7717/peerj.12993
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9:559. doi: 10.1186/1471-2105-9-559
- Li, F. J., Hu, J. H., Ren, X., Zhou, C. M., Liu, Q., and Zhang, Y. Q. (2021). Toad venom: a comprehensive review of chemical constituents, anticancer activities, and mechanisms. *Arch. Der Pharm.* 354:e2100060. doi: 10.1002/ardp.202100060
- Lu, B., Jiang, J., Wu, H., Chen, X., Song, X., Liao, W., et al. (2021). A large genome with chromosome-scale assembly sheds light on the evolutionary success of a true toad (*Bufo gargarizans*). *Mol. Ecol. Res.* 21, 1256–1273. doi: 10.1111/1755-0998.13319
- Nelson, D. L., Lehninger, A. L., and Cox, M. M. (2008). *Lehninger Principles of Biochemistry*. London: Macmillan.
- Nguyen, L.-T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Porto, A., Baralle, F., and Gros, E. (1972). Biosynthesis of bufadienolides in toads: III—Experiments with [2-<sup>14</sup>C] mevalonic acid, [20-<sup>14</sup>C] 3 $\beta$ -hydroxy-5-pregnen-20-one and [20-<sup>14</sup>C] cholesterol. *J. Steroid Biochem.* 3, 11–17. doi: 10.1016/0022-4731(72)90006-4
- Porto, A. M., and Gros, E. (1970). Biosynthesis of animal and plant bufadienolides. Parallel experiments with pregn-5-en-3- $\beta$ -ol-20-one-20-<sup>14</sup>C in *Scilla maritima* and *Bufo paracnemis*. *Experientia* 26:11. doi: 10.1007/BF01900357
- Porto, A. M., and Gros, E. (1971). Biosynthesis of the bufadienolide marinobufagin in toads *Bufo paracnemis* from cholesterol-20-<sup>14</sup>C. *Experientia* 27, 506–506. doi: 10.1007/BF02147562
- Posso-Terranova, A., and Andrés, J. (2020). Skin transcriptional profiles in Oophaga poison frogs. *Genet. Mol. Biol.* 43:e20190401. doi: 10.1590/1678-4685-gmb-2019-0401
- Qi, J., Zufiker, A. H. M., Li, C., Good, D., and Wei, M. Q. (2018). The development of toad toxins as potential therapeutic agents. *Toxins* 10:336. doi: 10.3390/toxins10080336
- Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., et al. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592, 737–746. doi: 10.1038/s41586-021-03451-0
- Rodriguez, C., Ibáñez, R., Rollins-Smith, L. A., Gutiérrez, M., and Durant-Archibold, A. A. (2020). Antimicrobial secretions of toads (Anura, Bufonidae): bioactive extracts and isolated compounds against human pathogens. *Antibiotics* 9:843. doi: 10.3390/antibiotics9120843
- Rone, M. B., Fan, J., and Papadopoulos, V. (2009). Cholesterol transport in steroid biosynthesis: role of protein–protein interactions and implications in disease states. *Biochim. Biophys. Acta* 1791, 646–658. doi: 10.1016/j.bbali.2009.03.001
- Santa Coloma, A., Garraffo, H., Pignataro, O., Charreau, E., and Gros, E. (1984). Biosynthesis of bufadienolides in toads. V. The origin of the cholesterol used by toad parotoid glands for biosynthesis of bufadienolides. *Steroids* 44, 11–22. doi: 10.1016/s0039-128x(84)80012-4
- Sharpe, L. J., and Brown, A. J. (2013). Controlling cholesterol synthesis beyond 3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR). *J. Biol. Chem.* 288, 18707–18715. doi: 10.1074/jbc.R113.479808
- Shibao, P. Y. T., Cologna, C. T., Morandi-Filho, R., Wiesel, G. A., Fujimura, P. T., Ueira-Vieira, C., et al. (2018). Deep sequencing analysis of toad *Rhinella schneideri* skin glands and partial biochemical characterization of its cutaneous secretion. *J. Venom. Anim. Toxins Includ. Trop. Dis.* 24:36. doi: 10.1186/s40409-018-0173-8
- Stormo, C., Kringen, M. K., Grimholt, R. M., Berg, J. P., and Piehler, A. P. (2012). A novel 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGCR) splice variant with an alternative exon 1 potentially encoding an extended N-terminus. *BMC Mol. Biol.* 13:29. doi: 10.1186/1471-2199-13-29
- Szklarczyk, D., Gable, A. L., Lyon, D., Jung, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131
- Tian, H.-Y., Luo, S.-L., Liu, J.-S., Wang, L., Wang, Y., Zhang, D.-M., et al. (2013). C23 steroids from the venom of *Bufo bufo gargarizans*. *J. Nat. Prod.* 76, 1842–1847. doi: 10.1021/np400174f
- Wang, D. L., Qi, F. H., Tang, W., and Wang, F. S. (2011). Chemical constituents and bioactivities of the skin of *Bufo bufo gargarizans* Cantor. *Chem. Biodiver.* 8, 559–567. doi: 10.1002/cbdv.201000283

West, J. (2018). *Importance of amphibians: A synthesis of their environmental functions, benefits to humans, and need for conservation*. InBSU Honors Program Theses and Projects. Item 261. Available online at: [http://vc.bridgew.edu/honors\\_proj/261](http://vc.bridgew.edu/honors_proj/261) (accessed September 5, 2018).

Willnow, T. E. (1999). The low-density lipoprotein receptor gene family: multiple roles in lipid metabolism. *J. Mol. Med.* 77, 306–315. doi: 10.1007/s001090050356

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088