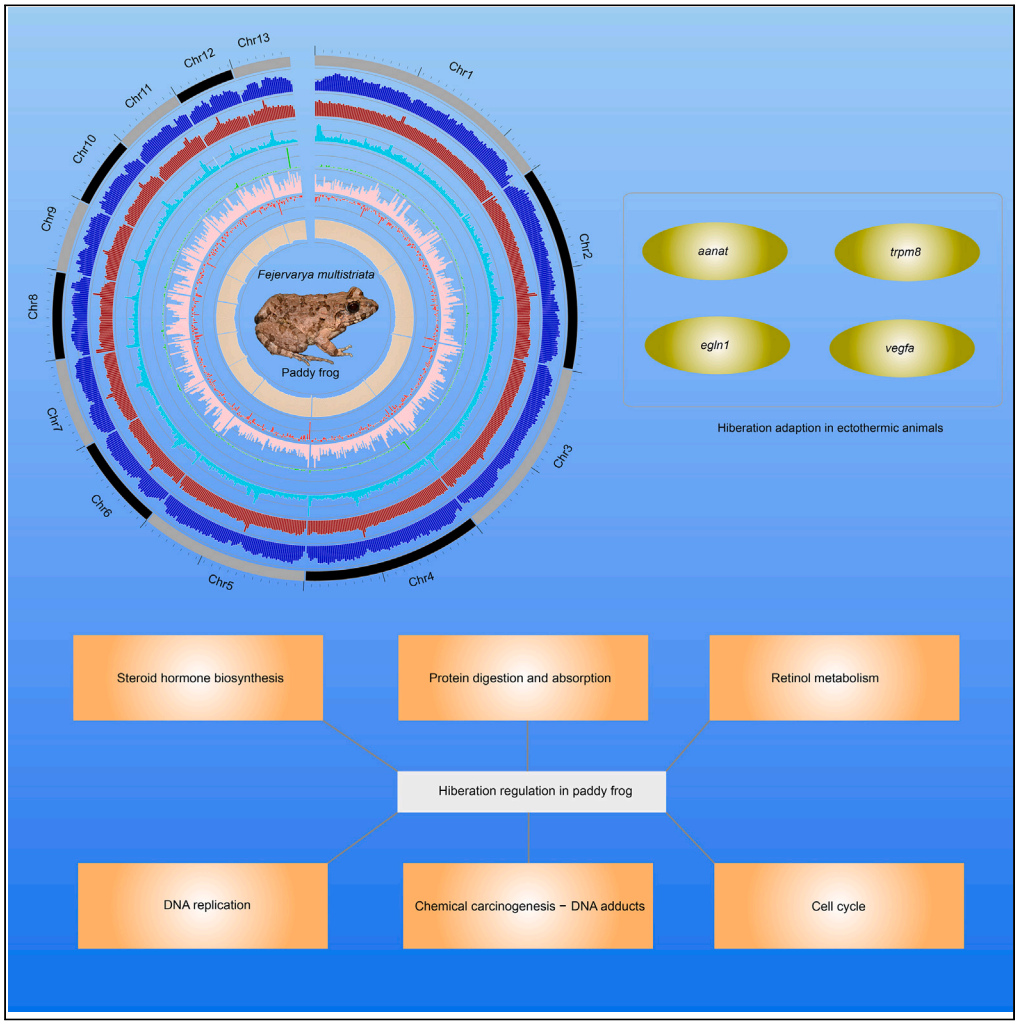
iScience

CellPress
OPEN ACCESS

Article

The paddy frog genome provides insight into the molecular adaptations and regulation of hibernation in ectotherms



Yunyun Lv, Chuan Chen, Chengzhi Yan, Wenbo Liao

liaobo_0_0@126.com

Highlights

De novo chromosome-level assembly of the *Fejervarya multistriata* genome

Identification of genes adapted to hibernation in ectothermic animals

Regulation of molecular pathways related to hibernation in ectotherms

Lv et al., iScience *27*, 108844
February 16, 2024 © 2024
https://doi.org/10.1016/j.isci.2024.108844

iScience

CellPress
OPEN ACCESS

Article

The paddy frog genome provides insight into the molecular adaptations and regulation of hibernation in ectotherms

Yunyun Lv,[1,2] Chuan Chen,[1] Chengzhi Yan,[1] and Wenbo Liao[1,3,4,*]

SUMMARY

Amphibians, like the paddy frog (*Fejervarya multistriata*), have played a critical role in the transition from water to land. Hibernation is a vital survival adaptation in cold environments with limited food resources. We decoded the paddy frog genome to reveal the molecular adaptations linked to hibernation in ectotherms. The genome contained 13 chromosomes, with a significant proportion of repetitive sequences. We identified the key genes encoding the proteins of AANAT, TRPM8, EGLN1, and VEGFA essential for circadian rhythms, thermosensation, and hypoxia during hibernation by comparing the hibernator and non-hibernator genomes. Examining organ changes during hibernation revealed the central regulatory role of the brain. We identified 21 factors contributing to hibernation, involving hormone biosynthesis, protein digestion, DNA replication, and the cell cycle. These findings provide deeper insight into the complex mechanisms of ectothermic hibernation and contribute to our understanding of the broader significance of this evolutionary adaptation.

INTRODUCTION

Amphibians have played a pivotal role in the evolutionary transition from aquatic to terrestrial life, making them captivating subjects for scientific exploration. The ability of amphibians to thrive in diverse environments has been attributed to various adaptations, including hibernation, which allows them to conserve energy and endure unfavorable conditions during the winter.[1,2]

Unlike endotherms, ectotherms do not generate internal body heat. Hibernation in ectotherms, such as amphibians and reptiles, involves regulating the body temperature primarily through external heat sources.[3,4] Additionally, ectotherms rely on environmental cues to adjust their metabolic activity,[5] suggesting that the physiological mechanisms involved in regulating hibernation differ between endotherms and ectotherms.

It is reported that hibernators from endotherms significantly lower their metabolic rate, conserve energy, and rely on stored fat reserves for sustenance during the dormant period.[6–10] A few studies have revealed the physiological status of ectotherms during hibernation. For example, the liver and muscle of ranid frogs serve as energy storage sites.[11] Degradation of skeletal muscle in *Rana sylvatica* leads to the production of glycogen, contributing to energy maintenance during hibernation.[12] Moreover, studies on *Pogona vitticeps* have identified specific genes that are enhanced in neuroprotective pathways within the brain. These genes likely play a role in self-protective functions during hibernation.[13] These findings offer some insight into the physiological adaptations of ectotherms during hibernation, but a deeper understanding of the functions of multiple organs at the same time is unavailable.

In addition to these physiological changes, hibernation in ectotherms is believed to involve significant long-term molecular adaptations.[14] One important molecular process is gene conversion, which refers to the nonreciprocal transfer of genetic information between homologous DNA sequences without an accompanying information exchange.[15] Gene conversion may contribute to changing the genetic material[16] and also play a role in shaping the hibernation response. Despite its potential impact, further investigation focused on elucidating the role of gene conversion in hibernation is currently unavailable, leaving this aspect of research relatively unexplored.

The paddy frog (*Fejervarya multistriata*) is a general example of a wintering hibernator. When faced with challenging environmental conditions, this species adopts a hibernation strategy. Similar to many other amphibians and reptiles, paddy frogs actively seek out sheltered locations like burrows or underwater environments. These safe havens provide protection and help minimize exposure to extreme temperatures. The paddy frog has a broad geographic distribution range, including China, the southern islands of Japan, northern India, and southern Asia, such as Vietnam, Laos, and Thailand.[17] This wide distribution makes it convenient for sampling and breeding. Additionally, this frog is

[1]Key Laboratory of Southwest China Wildlife Resources Conservation (Ministry of Education), China West Normal University, Nanchong, Sichuan 637009, China
[2]Key Laboratory of Sichuan Province for Fishes Conservation and Utilization in the Upper Reaches of the Yangtze River, College of Life Science, Neijiang Normal University, Neijiang 641100, China
[3]College of Panda, China West Normal Univetsity, Nanchong, Sichuan 637009, China
[4]Lead contact
*Correspondence: liaobo_0_0@126.com
https://doi.org/10.1016/j.isci.2024.108844

commonly observed in rice fields, moors, ditches, and other environments near standing water and grass, making it an ideal indicator species for estimating environmental pollution and studying the relationship between the environment and species variability.[18] However, there is currently a paucity of molecular studies on this species, with limited research focused on molecular phylogeny and taxonomy.[19] Establishing a genome resource for the paddy frog would greatly benefit our understanding of the life-history evolution of this species.

This study aimed to decipher the paddy frog genome and conduct comparative genome analyses with other ectotherms, such as toads, frogs, and reptiles. We sought to reveal the unique molecular biology and evolution of hibernators among ectotherms. Additionally, we synthesized the physiological changes related to gene regulation in multiple organs to uncover the main molecular pathways involved in hibernation in the paddy frog. This study will deepen our understanding of the complex mechanisms underlying hibernation in ectotherms and contribute to a broader knowledge of evolutionary adaptations.

RESULTS

Genomic features of the paddy frog

We generated a comprehensive dataset consisting of 240 Gb of next-generation sequencing reads and 389 Gb of PacBio CLR reads in the paddy frog genome. We successfully assembled a high-quality paddy frog genome of 3.86 Gb with GC content of 0.403 (Figure 1; Table S1). The sequencing data provided 163-fold coverage based on the genome size.

Utilizing chromosome-to-chromosome contact information, we updated the assembly of the paddy frog genome at the chromosome-scale level. This updated assembly comprised 13 chromosomes ranging from 139 to 601 Mb, with clear boundaries observed in the chromosome contact map (Figure S1).

The paddy frog genome exhibited a large number of repetitive sequences, accounting for 70.2% of the total genome size (2.71 Gb). Among the transposons investigated (Table S2), DNA transposons and long terminal repeats (LTRs) had the highest ratios compared to long interspersed repetitive elements (LINEs) and short interspersed repetitive elements (SINEs) (Figure 1; Table S1). Moreover, the percent divergence analysis among the types of transposons revealed that the DNA and LTR transposons had lower divergence than the SINEs and LINEs (Figure S2), suggesting recent expansion of the DNA and LTR transposons.

We conducted gene annotation within the genomic regions containing transposons and identified 22,030 protein-encoding genes (Figure 1) in the paddy frog genome. This annotation represented 88.7% completeness using 3,354 conserved single-copy genes (Table S3).

Molecular evolution of hibernating animals

Comparative genomics analysis across hibernating animals, including the paddy frog, common frog (*Rana temporaria*), common toad (*Bufo bufo*), Asiatic toad (*Bufo gargarizans*), common lizard (*Zootoca vivipara*), Burmese python (*Python molurus bivittatus*), Western terrestrial garter snake (*Thamnophis elegans*), and Green anole (*Anolis carolinensis*), revealed the presence of 682 rapidly evolving genes (REGs) and 34 positively selected genes (PSGs) (Figure 2A). The analysis of gene family dynamics identified two clades representing amphibians and reptiles, with 27 expanded and 15 contracted gene families observed in these clades.

We identified 34 PSGs. However, through functional annotation and investigation, we discovered that four of these genes are specifically associated with crucial biological functions related to hibernation in ectothermic animals. These functions include regulating rhythm, sensing temperature, and responding to low oxygen levels. We explored the functions of other genes, but they had little relevance to hibernation physiology. The four PSGs found in hibernating animals encoded the proteins AANAT (Figure 2B), TRPM8 (Figure 2C), EGLN1 (Figure 2D), and VEGFA (Figure 2E). These PSGs formed a functional link within the hibernation process, involving rhythms, thermosensation, and hypoxia; thus, playing a role in adaptation to hibernation (Figures 2B–2E).

Gene expression patterns of hibernating paddy frogs

The hibernating conditions during winter were compared with the active conditions in the spring. As a result, differentially expressed genes (DEGs) were observed among various tissues, including the brain, heart, lungs, liver, kidneys, spleen, muscle, and skin. Notably, the heart and kidneys displayed the highest number of DEGs (Table S4), while the liver exhibited the fewest (Figures 3A and 3B). Notably, the brain, heart, and kidneys exhibited a distinct pattern in which the upregulated DEGs outnumbered the downregulated DEGs, whereas the opposite trend was observed in other tissues with higher numbers of downregulated DEGs. This finding suggests that the activity levels of genes in the brain, heart, and kidneys are higher during hibernation, while gene expression in other tissues is generally suppressed. In addition, an analysis of overlapping DEGs revealed that the heart exhibited a higher degree of overlap with the kidneys, lungs, brain, muscle, and skin (Figure 3B).

Furthermore, our analysis revealed that the upregulated DEGs in the brain and heart were significantly enriched in molecular pathways that play crucial roles in cellular processes. These pathways included regulation of the cell cycle, DNA replication, the p53 signaling pathway, and the Fanconi anemia pathway (Figures S3 and S4), suggesting that there may be shared mechanisms driving important cellular processes in the brain and heart, such as regulation of the cell cycle and repair of genetic material.

In contrast, the downregulated genes were enriched in the heart, liver, lungs, muscle, and skin. The functions of these genes were predominantly related to energy metabolism (Figures S5–S9). Notably, some genes in molecular pathways associated with vascular or heart diseases were downregulated (Figures S5, S7, and S8), such as fluid shear stress, atherosclerosis, viral myocarditis, hypertrophic cardiomyopathy, and dilated cardiomyopathy.

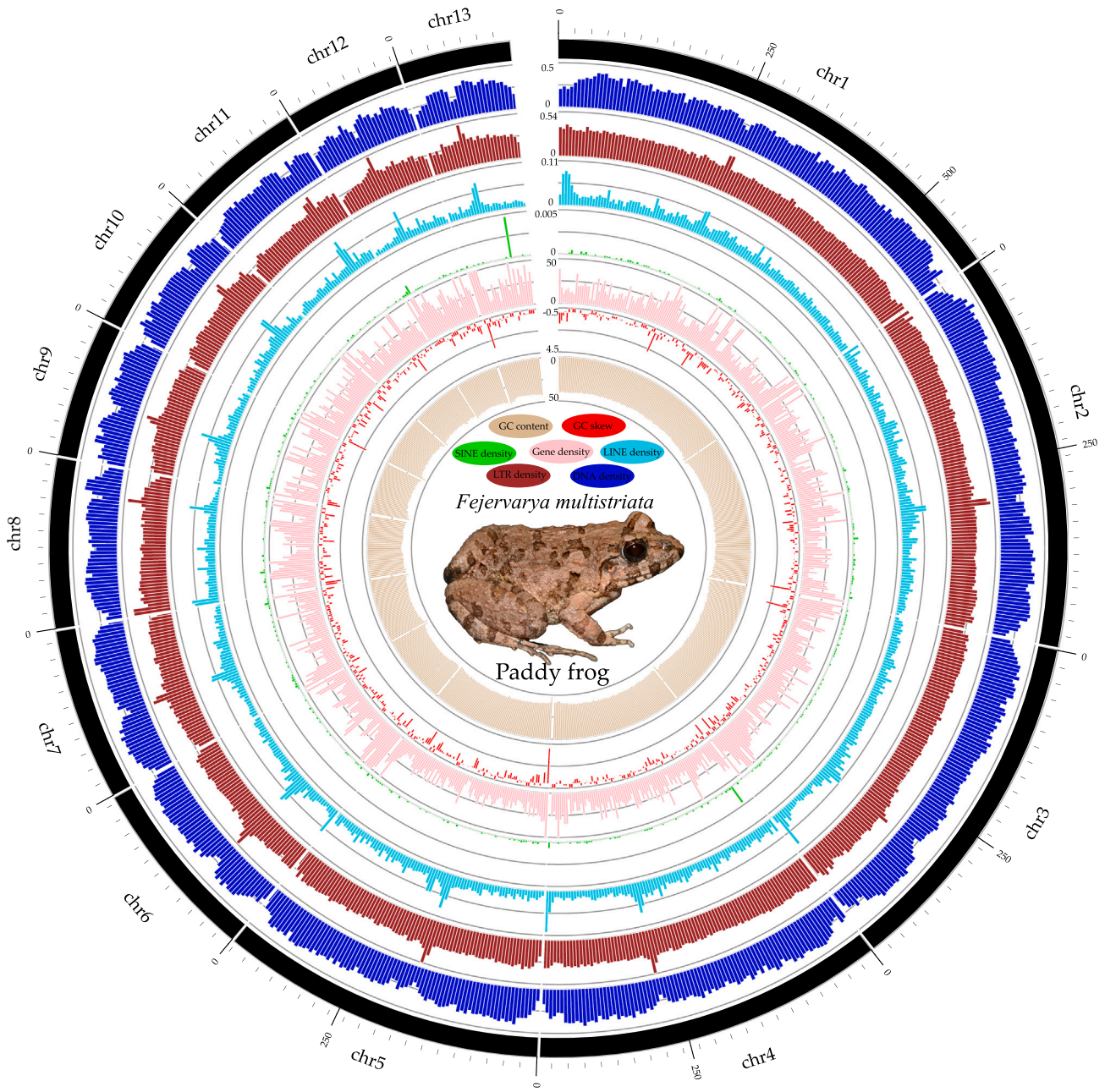


Figure 1. Genome features of the paddy frog: Insight into DNA transposons, LTRs, LINEs, Genes, SINEs, GC Content, and GC Skew

Each ring from the center to the outer edge of the circle represents various characteristics, including GC content, GC skew, SINE density, gene density, LINE density, LTR density, and DNA transposons density. The length window was fixed at 5 Mb.

Gene regulatory networks involving hibernation in paddy frogs

The DEGs were analyzed using gene expression clustering, resulting in the identification of 16 co-expression modules (Figure 4A; Figure S10). The turquoise module exhibited the highest number of genes (Figure 4B; Figure S11) but was not significantly associated with a specific tissue (Figure 4C; Figure S11), suggesting that hibernation may affect the changes in expression across multiple tissues, thereby lacking a distinct tissue-specific link with this module.

In contrast, the blue module (Figures 4C and 4D) had the second-highest number of genes and displayed a significant association with the brain, indicating that the brain likely plays a central role in gene regulation during hibernation. Additionally, the pink module (Figures 4C and 4D) was significantly associated with the brain. The blue and pink modules together exhibited the highest co-expression relationships with the brain, making them the modules with the largest numbers of genes among those showing significant links with other tissues (Figure 4D; Figure S11).

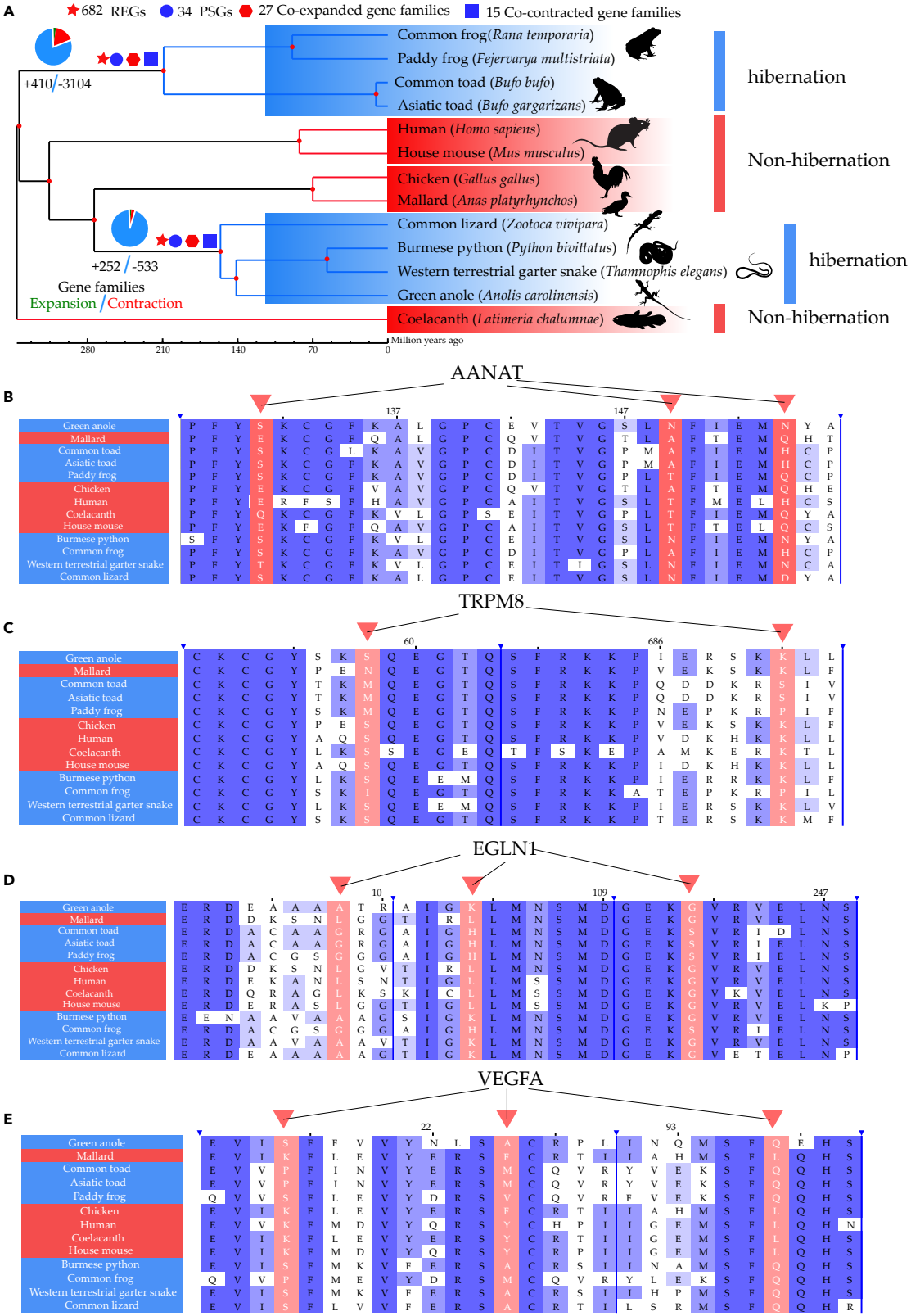

A
682 REGs
34 PSGs
27 Co-expanded gene families
15 Co-contracted gene families
+410/-3104
+252/-533
Gene families
Expansion/Contraction
Common frog(Rana temporaria)
Paddy frog (Fejervarya multistriata)
Common toad (Bufo bufo)
Asiatic toad (Bufo gargarizans)
Human (Homo sapiens)
House mouse (Mus musculus)
Chicken (Gallus gallus)
Mallard (Anas platyrhynchos)
Common lizard (Zootoca vivipara)
Burmese python (Python bivittatus)
Western terrestrial garter snake (Thamnophis elegans)
Green anole (Anolis carolinensis)
Coelacanth (Latimeria chalumnae)
hibernation
Non-hibernation
hibernation
Non-hibernation
280
210
140
70
0
Million years ago
B
AANAT
137
147
Green anole
Mallard
Common toad
Asiatic toad
Paddy frog
Chicken
Human
Coelacanth
House mouse
Burmese python
Common frog
Western terrestrial garter snake
Common lizard
C
TRPM8
60
686
Green anole
Mallard
Common toad
Asiatic toad
Paddy frog
Chicken
Human
Coelacanth
House mouse
Burmese python
Common frog
Western terrestrial garter snake
Common lizard
D
EGLN1
10
109
247
Green anole
Mallard
Common toad
Asiatic toad
Paddy frog
Chicken
Human
Coelacanth
House mouse
Burmese python
Common frog
Western terrestrial garter snake
Common lizard
E
VEGFA
22
93
Green anole
Mallard
Common toad
Asiatic toad
Paddy frog
Chicken
Human
Coelacanth
House mouse
Burmese python
Common frog
Western terrestrial garter snake
Common lizard

Figure 2. Comparative genome analysis reveals adaptations to hibernation in hibernators and non-hibernators

(A) Phylogenetic analysis of species including hibernators and non-hibernators based on single-copy genes. Identification of rapidly evolving genes (REGs), positively selected genes (PSGs), co-expanded gene families, and co-contracted gene families associated with adaptation to hibernation in hibernators.

(B) Amino acid sequence alignment of AANAT. The three red triangles correspond to three positively selected sites identified in the positive selection analysis. These sites are located at positions 131, 149, and 154 in the alignment.

(C) Amino acid alignment of TRPM8. The two red triangles correspond to positions 58 and 691 in the sequence alignment.

(D) Amino acid alignment of EGLN1 and identification of three positively selected sites. These sites correspond to positions 8, 103, and 241 in the sequence alignment.

(E) Amino acid alignment of VEGFA and identification of three positively selected sites located at positions 16, 25, and 97.

We identified the hub genes within each module to further investigate the gene regulatory network (Figure S12). These hub genes are crucial for regulating the overall gene network. Notably, we identified more hub genes in the brain than in other tissues (Figures 4E–4L).

Overall, these findings suggest that hibernation induces common expression changes across various tissues, with the brain playing a central role in gene regulation. The hub genes identified in the brain module highlight their importance in orchestrating the gene network during hibernation.

Molecular pathways related to hibernation

Hibernation is a complex process involving long-term adaptation and gene regulation. Our study explored the evolutionary factors influencing hibernation and examined tissue-specific genes involved in regulation. Overall, we categorized 21 factors that could potentially be associated with adaptation or regulation of hibernation. To gain further insight, we performed enrichment analyses of these genes, which revealed significant enrichment in six molecular pathways, including steroid hormone biosynthesis, protein digestion and absorption, retinol metabolism, DNA replication, chemical carcinogenesis-DNA adducts, and the cell cycle ($p < 0.05$, Figure 5A). We gained a better understanding of the mechanisms underlying adaptation to and regulation of hibernation by considering these factors and pathways.

After identifying the genes in the pathways, we discovered that these genes were associated with various factors related to gene evolution or gene regulation. Specifically, in the steroid hormone biosynthetic pathway, we observed gene expansion in four genes, such as *Sult5a1*, *SULT2B1*, *SDR9C7*, and *CYP2B6*. In contrast, other genes, such as *sult1a4*, *HSD11B1L*, *CYP2g1*, and *CYP2C8*, were downregulated in the lungs. Interestingly, the *sult5a1*, *HSD17B7*, and *HAO2* genes were upregulated in the heart (Figure 5B).

We discovered several genes that were associated with rapid gene evolution in the analysis of the protein digestion and absorption pathway, including *SMC1B*, *SLC9A1*, *SLC16A10*, *SLC15A1*, *COL9A2*, *COL6A3*, *COL6A1*, and *ATP1B3* (Figure 5C). Additionally, we identified the hub genes in the co-expression networks encoding the proteins SLC1A1, SLC15A2, DPP4, COL6A2, COL3A1, COL1A2, and COL1A1. However, some of these genes encoding the proteins COL6A2, COL3A1, COL1A2, and COL1A1 were downregulated in the lungs, suggesting suppressed expression in this organ (Figure 5C).

The genes encoding SDR9C7, CYP2C18, and CYP2B6 in the retinol metabolic pathway were REGs. The ALDH2, ADH7, and ADH4 genes were upregulated in the heart. However, the genes encoding Cyp2g1, CYP2C8, and ALDH2 were downregulated in the lungs. Interestingly, the gene encoding ALDH2 was also identified as a REG (Figure 5D).

Five REGs were detected in the DNA replication pathway encoding RPA2, RFC3, POLA1, PCNA, and MRPS34. Among the tissues, the most prominent regulation of genes was observed in the brain and heart, where the following genes were upregulated (brain: *RFC3*, *PRIM1*, *PCNA*, and *FEN1*; heart: *RPA2*, *PRIM2*, *POLE*, and *MCM3*). This finding indicates that the molecular changes in this pathway are related to adaptation of the brain and heart during hibernation (Figure 5E).

Interestingly, the pathway of chemical carcinogenesis-DNA adducts was enriched. Few genes are evolutionarily related to this pathway. Instead, most of the gene regulation was suppressed, such as downregulated genes in the heart (*CCBL2*, *ADH7*, and *ADH4*), lungs (*NAT1*, *HSD11B1L*, *Cyp2g1*, and *CYP2C8*), and muscle (*ADH7*, *UGT8*, and *SULT1B1*) (Figure 5F).

The cell cycle was similar to the DNA replication pathway (Figure 5G). Many of the genes involved in this pathway were REGs, and they were mainly regulated in the brain and heart, where related genes were upregulated.

In summary, our results combining evolutionary factors and gene regulation explored the relationships between molecular changes and hibernation. These factors collectively affected the molecular changes associated with hibernation. However, some of the changes were controlled primarily by gene regulation, as seen in the chemical carcinogenesis-DNA adducts pathway. Gene expansion or rapid evolution appeared to be the main driver contributing to molecular adaptation during hibernation.

DISCUSSION

Paddy frog genome resource

Our study provides valuable insight into the paddy frog genome, revealing its complexity and relatively large size of approximately 3.86 Gb. The paddy frog genome falls within the intermediate range compared to other sequenced frogs.[20–30] Notably, the paddy frog genome ranks as the fourth largest among all anuran genomes sequenced to date, with repetitive elements accounting for 70.2% of the total. This high proportion of repetitive sequences is consistent with previous findings on amphibian genomes.[20–30]

The accumulation of interspersed repeats, particularly likelihood ratio test (LRT) expansions, primarily contributed to the abundance of repetitive sequences in the paddy frog genome. We identified a substantial presence of DNA transposons, suggesting their role in enlarging

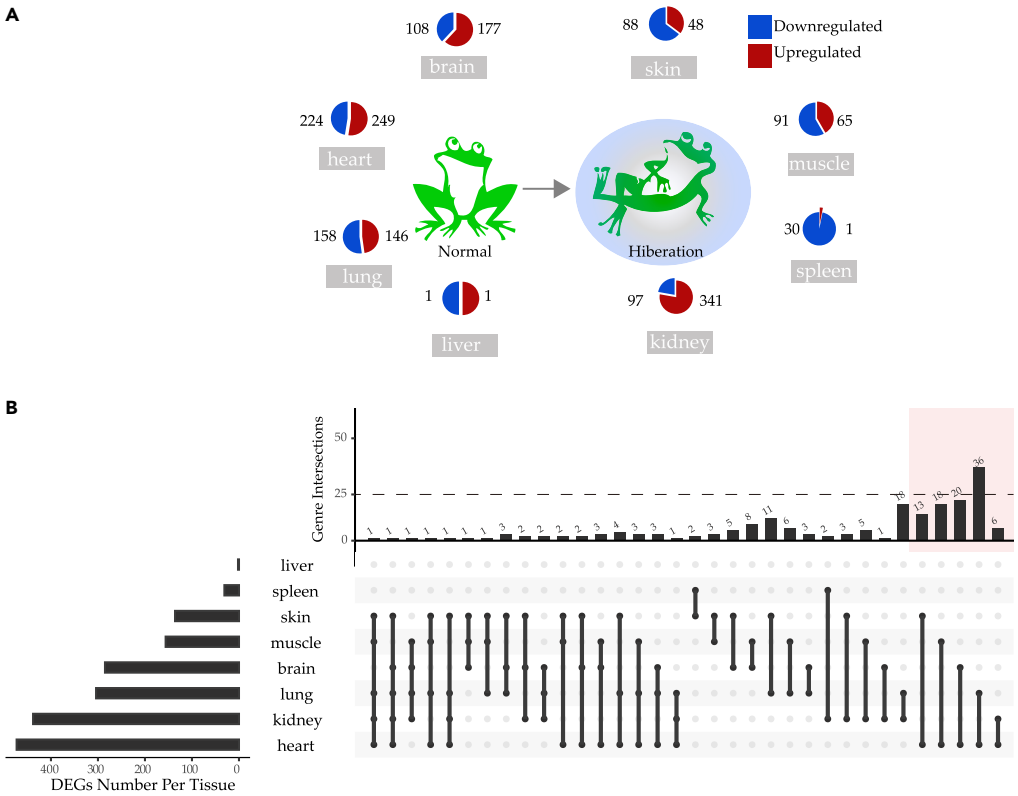


Figure 3. Identification of differentially expressed genes (DEGs) in various organs during hibernation

(A) Differential expression of genes between normal state (Spring Sampling) and hibernation.

(B) Number of DEGs identified in different organs and overlap analysis of DEGs across organs reveals strong association with the heart.

the genome. These findings enhance our understanding of the genetic characteristics of the paddy frog and provide a valuable resource for future studies in comparative genomics, evolutionary biology, and conservation genetics.

Long-term molecular adaption signals to rhythmic hibernation

Research on hibernation in ectotherms is limited, particularly regarding the evolutionary factors driving the adaptation to hibernation. However, as amphibians and reptiles transitioned from water to land, hibernation may have evolved as an adaptation to extremely cold temperatures and limited food resources in the terrestrial environment. Long-term evolutionary processes contribute to the ability of ectotherms to adapt to their environment.[31]

To address these knowledge gaps, we conducted an evolutionary analysis by comparing ectothermic and endothermic hibernator species. Our investigation revealed the conversion signals associated with molecular adaptations to hibernation. Specifically, we identified several genes encoding the proteins including AANAT, TRPM8, EGLN1, and VEGFA that exhibited positive selection, indicating their involvement in the evolutionary adaptation to hibernation. AANAT, which is responsible for melatonin synthesis,[32] displayed high expression and activity during torpor and low expression during arousal, suggesting its role in regulating hibernation patterns. TRPM8 is a cold-sensitive ion channel that is upregulated in specific tissues during hibernation,[33] potentially aiding in temperature adaptation. EGLN1 is involved in the response to low oxygen levels, and exhibits increased expression during hibernation, indicating its role in adapting to hypoxia.[34] VEGFA, which promotes the growth of blood vessels, was downregulated during hibernation,[35] which was likely a strategy to conserve energy. These genes had specific sites that were correlated with their respective adaptations.

We revealed the genetic changes and biological functions contributing to the evolutionary adaptation of hibernating animals in our comparative genomic analysis and investigation of PSGs. These findings deepen our understanding of the rhythmic, temperature, hypoxic, and metabolic adaptations underlying hibernation and provide valuable insight into the genetic basis of this unique physiological state.

Changes in gene regulation across multiple organs during hibernation in paddy frogs

Significant changes in gene expression were observed in the heart and kidney tissues during hibernation, as indicated by the high number of DEGs (Figures 3A and 3B). Although we did not identify any significantly enriched molecular pathways in the kidney, we discovered several hub genes associated with kidney function. These genes belong to the solute carrier family, which includes: (1) *SLC1A1*: involved in the

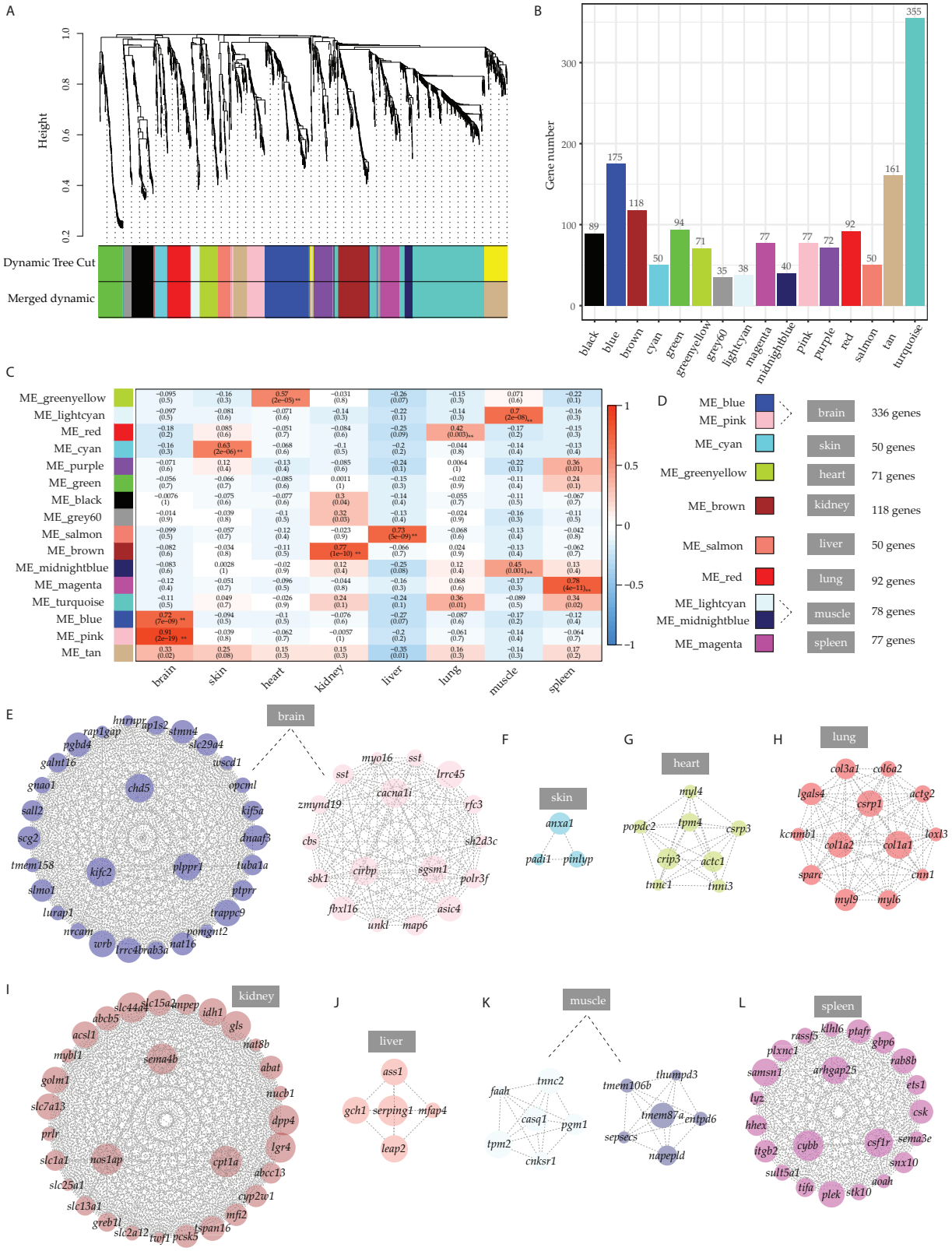

A
Height
Dynamic Tree Cut
Merged dynamic
B
Gene number
black
blue
brown
cyan
green
greenyellow
grey60
lightcyan
magenta
midnightblue
pink
purple
red
salmon
tan
turquoise
C
ME_greenyellow
ME_lightcyan
ME_red
ME_cyan
ME_purple
ME_green
ME_black
ME_grey60
ME_salmon
ME_brown
ME_midnightblue
ME_magenta
ME_turquoise
ME_blue
ME_pink
ME_tan
brain
skin
heart
kidney
liver
lung
muscle
spleen
D
ME_blue
ME_pink
ME_cyan
ME_greenyellow
ME_brown
ME_salmon
ME_red
ME_lightcyan
ME_midnightblue
ME_magenta
brain 336 genes
skin 50 genes
heart 71 genes
kidney 118 genes
liver 50 genes
lung 92 genes
muscle 78 genes
spleen 77 genes
E
brain
F
skin
G
heart
H
lung
I
kidney
J
liver
K
muscle
L
spleen

Figure 4. Exploring the co-expression network and organ-specific modules in paddy frog

(A) Exploring the co-expression network and divisions for different modules.

(B) Gene numbers in the modules (C and D) organ-specific module identification.

(E–L) Hub gene co-expression network across organs.

reabsorption of the neurotransmitter glutamate in the kidneys and central nervous system.[36] (2) *SLC13A1*: plays a role in the reabsorption of sulfate ions.[37] (3) *SLC25A1*: encodes a mitochondrial citrate transporter that is involved in the reabsorption of citrate in the kidneys.[38] (4) *SLC2A12*: may be involved in glucose reabsorption.[39] (5) *SLC15A2*: encodes the PEPT2a peptide transporter, which participates in the reabsorption of di and tripeptides in the kidneys and other tissues.[40] (6) *SLC44A4*: encodes a choline transporter involved in the reabsorption of choline, an essential nutrient, in the kidneys.[41] These genes were upregulated in the kidneys during hibernation, suggesting that they play a significant role in facilitating multiple solute reabsorption processes, allowing hibernating animals to effectively maintain the solute balance in their bodies.

The downregulated genes in the heart were enriched in several pathways related to energy metabolism, including pyruvate metabolism, glycolysis/gluconeogenesis, the TCA cycle, and fatty acid degradation. Additionally, pathways associated with muscular movement, such as cardiac muscle contraction, were also downregulated. These findings suggest that the downregulation of these pathways contributes to energy conservation,[42] allowing paddy frogs to survive the winter.

The brain and heart of paddy frogs revealed upregulated genes primarily involved in the cell cycle and DNA replication processes, suggesting that upregulation may be a result of repair and maintenance mechanisms. Cold temperatures and harsh environmental conditions during winter cause cellular damage. Cells repair DNA damage and maintain optimal functioning by upregulating genes associated with the cell cycle and DNA replication.

The presence of other enriched pathways, such as the p53 signaling pathway and the Fanconi anemia pathway, further supports the idea that the upregulated genes in the brain and heart are involved in genetic repair and maintenance. Additionally, the downregulation of energy metabolism-related genes in paddy frogs suggests that they reduce their metabolic rate resembling both homeothermic animals and ectotherms like *R. sylvatica* during overwintering.[12] Interestingly, many genes were also enriched in heart and vascular diseases, suggesting that downregulation of these genes could potentially lead to various disease-related conditions.

Overall, the findings suggest that paddy frogs have developed mechanisms to protect and repair cells during hibernation, which could have implications for understanding and potentially treating certain human health conditions.

Furthermore, the presence of shared DEGs between the heart and other tissues suggests shared genetic pathways or regulatory mechanisms in these organs/tissues during hibernation. This finding highlights a potential connection between the heart and these other organs/tissues in terms of gene expression patterns.

Joint contribution of gene conversion and regulation in the adaptation to hibernation

We discovered diverse contributions while investigating the integration of evolutionary factors, such as gene conversion and gene family dynamics, during hibernation in ectotherms (Figures 5B–5E). Notably, the expansion of genes related to steroid hormone biosynthesis plays a crucial role in the glucocorticoid and sex hormone pathways, which are vital for regulating metabolism, reproduction, and the stress response.[43] These hormones are primarily synthesized in the adrenal glands and gonads, playing essential roles in normal functioning and survival.[44]

Significant changes in steroid hormone levels occur during hibernation to facilitate the adaptation to hibernation. For example, levels of the glucocorticoid cortisol decrease during hibernation, aiding in metabolic regulation and energy conservation.[45] Additionally, testosterone and estrogen levels decrease, suppressing reproductive functions.[46] The control of steroid hormone biosynthesis during hibernation involves a complex interplay of factors, including environmental cues, neural signaling, and specific molecular pathways.[47] A decrease in ambient temperature triggers the release of particular signals that act on the hypothalamus and pituitary gland, resulting in suppressed steroid hormone production.[48] However, the outcomes in amphibians may differ from the patterns observed in mammals. In amphibians, cholesterol synthesis and steroid production have been detected in the oviduct of *R. dybowskii* before hibernation.[49] Transcriptome analysis revealed high expression of the steroid synthetic pathway during the pre-hibernation period,[49] which differs from mammals. Additionally, boreal toads (*Anaxyrus boreas boreas*) exhibit increased sperm concentrations, a higher percentage of forward sperm movement, and improved quality of forward sperm movement during hibernation compared to non-hibernating toads,[50] suggesting that hibernation plays a significant role in amphibian reproduction.

Previous research has emphasized a lack of understanding regarding the link between retinol metabolism and hibernation. Retinol, commonly known as vitamin A, plays a crucial role in vision, cell differentiation, and immune function. It is acquired through dietary sources and is converted to its active form, retinoic acid, within the body. The changes in the expression levels of genes related to retinol metabolism in the heart and lungs during hibernation raise intriguing questions. It is possible that these genes have functional relevance in these organs, although further investigation is needed to determine the exact mechanisms and implications. The downregulation of these genes, along with decreased enzymatic activity, such as transcription, during hibernation, might indicate a metabolic adaptation to the physiological state of hibernation.

Our study revealed that this metabolic pathway exhibited the most significant downregulation of DEGs in the heart and lungs. This downregulation could potentially explain the scarcity of dietary sources during hibernation and may represent an adaptive mechanism to facilitate hibernation.

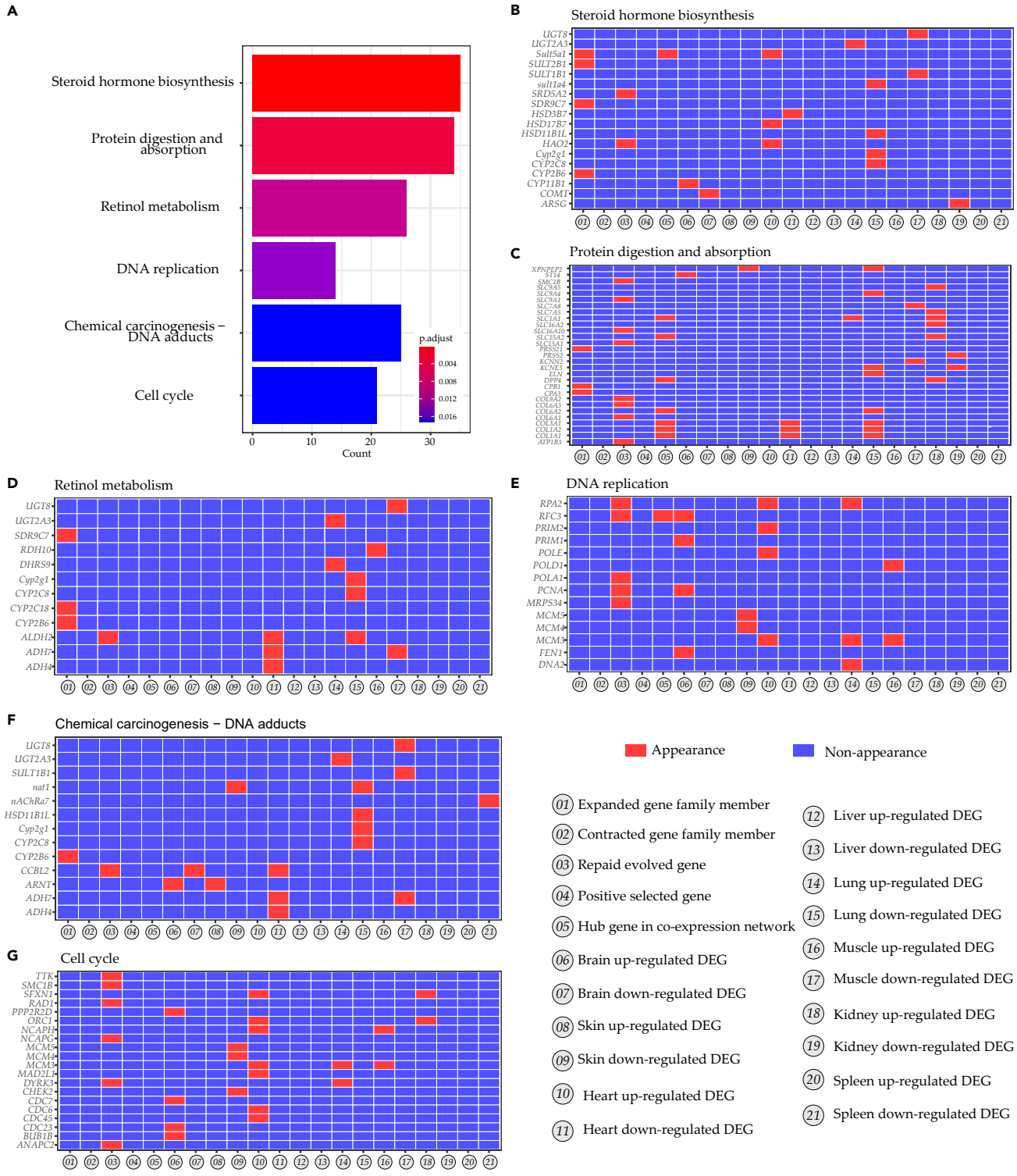


Figure 5. Exploring the enrichment of hibernation-related genes: Insights from evolutionary factors and gene regulation

(A) Revealing the six enriched molecular pathways associated with hibernation (B–G) deciphering the varied contributions of 21 investigated factors to the six molecular pathways of hibernation.

To assess the evolutionary factors affecting the molecular pathways, the expansion of genes and rapid gene evolution significantly contributes to the adaptation to hibernation. These factors are likely the main drivers of the adaption to hibernation. Thus, we propose that gene conversion and gene family dynamics are associated with gene regulation and play a role in the complex ectothermic hibernation process.

In summary, the integration of evolutionary factors, such as gene conversion and gene family dynamics, along with the regulation of steroid hormone biosynthesis and retinol metabolism, play a crucial role in the adaptation of ectotherms to hibernation. Modulating hormones to regulate metabolism, reproduction, and other physiological processes supports survival and adaptation to the hibernating state. Further research is needed to fully understand the intricate mechanisms underlying these processes and their significance in adapting to hibernation.

Limitations of the study

We obtained some molecular clues about hibernation in ectotherms from the paddy frog genome, but the lack of a model hibernation species limits the experimental validation of the identified genes. Further species comparisons and extensive experimental verification are required to identify the key genes regulating hibernation in ectothermic animals.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Animals
- METHOD DETAILS
  - Genome sequencing
  - Genome assembly and annotation
  - Comparative genomic analysis
  - Identification and functional enrichment of the DEGs
  - Gene module and WGCNA analyses
  - Molecular pathway enrichment analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.108844.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (32200347, 32370456) and the Open Fund Project (XNYB23-05) of the Key Laboratory of Southwest China Wildlife Resource Conservation (China West Normal University), the Ministry of Education the Key Project of Natural Sciences Foundation of Sichuan Province (22NSFSC0011) and the Neijiang Normal University Genomic Innovation Research Team (X23B0002). We thank Jiang, K. and Chen, C. for helping with collecting the field samples.

AUTHOR CONTRIBUTIONS

W.L. initiated and designed the study. Y.L., C.C., and C.Y. conducted the work. Y.L. and W.L. wrote the manuscript. W.L. edited the manuscript. All authors have read and approved the final manuscript.

DECLARATION OF INTERESTS

The author(s) declare no conflicts of interest.

Received: September 6, 2023
Revised: October 22, 2023
Accepted: January 3, 2024
Published: January 9, 2024

REFERENCES

1. Costanzo, J.P., do Amaral, M.C.F., Rosendale, A.J., and Lee, R.E., Jr. (2013). Hibernation physiology, freezing adaptation and extreme freeze tolerance in a northern population of the wood frog. J. Exp. Biol. *216*, 3461–3473.
2. Akat Çömden, E., Yenmiş, M., and Çakır, B. (2023). The Complex Bridge between Aquatic and Terrestrial Life: Skin Changes during Development of Amphibians. J. Dev. Biol. *11*, 6.
3. Tattersall, G.J., Sinclair, B.J., Withers, P.C., Fields, P.A., Seebacher, F., Cooper, C.E., and Maloney, S.K. (2012). Coping with thermal challenges: physiological adaptations to environmental temperatures. Compr. Physiol. *2*, 2151–2202.
4. Bogert, C.M. (1949). Thermoregulation in reptiles, a factor in evolution. Evolution *3*, 195–211.
5. Seebacher, F. (2009). Responses to temperature variation: integration of thermoregulation and metabolism in vertebrates. J. Exp. Biol. *212*, 2885–2891.
6. Storey, K.B., and Storey, J.M. (2010). Metabolic rate depression: the biochemistry of mammalian hibernation. Adv. Clin. Chem. *52*, 77–108.
7. Dark, J. (2005). Annual lipid cycles in hibernators: integration of physiology and behavior. Annu. Rev. Nutr. *25*, 469–497.
8. Storey, K.B., and Storey, J.M. (1990). Metabolic rate depression and biochemical adaptation in anaerobiosis, hibernation and estivation. Q. Rev. Biol. *65*, 145–174.
9. Dausmann, K.H., Glos, J., and Heldmaier, G. (2009). Energetics of tropical hibernation. J. Comp. Physiol. B *179*, 345–357.
10. Klug, B.J., and Brigham, R.M. (2015). Changes to Metabolism and Cell Physiology that Enable Mammalian Hibernation. Springer Sci. Rev. *3*, 39–56.
11. Tattersall, G.J., and Ultsch, G.R. (2008). Physiological ecology of aquatic overwintering in ranid frogs. Biol. Rev. *83*, 119–140.
12. Costanzo, J.P. (2019). Overwintering adaptations and extreme freeze tolerance in a subarctic population of the wood frog, Rana sylvatica. J. Comp. Physiol. B *189*, 1–15.
13. Capraro, A., O'Meally, D., Waters, S.A., Patel, H.R., Georges, A., and Waters, P.D. (2019). Waking the sleeping dragon: gene expression profiling reveals adaptive strategies of the hibernating reptile Pogona vitticeps. BMC Genom. *20*, 460–516.
14. Storey, K.B. (1990). Life in a frozen state: adaptive strategies for natural freeze tolerance in amphibians and reptiles. Am. J. Physiol. *258*, R559–R568.
15. Lorenz, A., and Mpaulo, S.J. (2022). Gene conversion: a non-Mendelian process integral to meiotic recombination. Heredity *129*, 56–63.
16. Chen, J.M., Férec, C., and Cooper, D.N. (2010). Gene conversion in human genetic disease. Genes *1*, 550–563.
17. Djong, H.T., Matsui, M., Kuramoto, M., Nishioka, M., and Sumida, M. (2011). A new species of the Fejervarya limnocharis complex from Japan (Anura, Dicroglossidae). Zoolog. Sci. *28*, 922–929.
18. Simon, E., Puky, M., Braun, M., and Tóthmérész, B. (2011). Frogs and toads as biological indicators in environmental assessment. Frogs, 141–150.
19. Jiang, L., Lv, G., Liu, L., Wu, B., Xu, Z., and Li, Y. (2020). Characterization of the complete mitochondrial genome of the paddy frog Fejervarya multistriata (Anura: Dicroglossidae) and its phylogeny. Mitochondrial DNA Part B *5*, 1248–1250.
20. Streicher, J.W.; Wellcome Sanger Institute Tree of Life programmeWellcome Sanger Institute Scientific Operations DNA Pipelines collective; Tree of Life Core Informatics collective; Darwin Tree of Life Consortium (2021). The genome sequence of the common frog, Rana temporaria Linnaeus 1758. Wellcome Open Res. *6*, 286.
21. Chen, W., Chen, H., Liao, J., Tang, M., Qin, H., Zhao, Z., Liu, X., Wu, Y., Jiang, L., Zhang, L., et al. (2023). Chromosome-level genome assembly of a high-altitude-adapted frog (Rana kukunoris) from the Tibetan plateau provides insight into amphibian genome evolution and adaptation. Front. Zool. *20*, 1–12.
22. Lu, B., Jiang, J., Wu, H., Chen, X., Song, X., Liao, W., and Fu, J. (2021). A large genome with chromosome-scale assembly sheds light on the evolutionary success of a true toad (Bufo gargarizans). Mol. Ecol. Resour. *21*, 1256–1273.
23. Li, Y., Ren, Y., Zhang, D., Jiang, H., Wang, Z., Li, X., and Rao, D. (2019). Chromosome-level assembly of the mustache toad genome using third-generation DNA sequencing and Hi-C analysis. GigaScience *8*, giz114.
24. Li, J., Yu, H., Wang, W., Fu, C., Zhang, W., Han, F., and Wu, H. (2019). Genomic and transcriptomic insights into molecular basis of sexually dimorphic nuptial spines in Leptobrachium leishanense. Nat. Commun. *10*, 5551.
25. Rodríguez, A., Mundy, N.I., Ibáñez, R., and Pröhl, H. (2020). Being red, blue and green: the genetic basis of coloration differences in the strawberry poison frog (Oophaga pumilio). BMC Genom. *21*, 301–316.
26. Hammond, S.A., Warren, R.L., Vandervalk, B.P., Kucuk, E., Khan, H., Gibb, E.A., Pandoh, P., Kirk, H., Zhao, Y., Jones, M., et al. (2017). The North American bullfrog draft genome provides insight into hormonal regulation of long noncoding RNA. Nat. Commun. *8*, 1433.
27. Edwards, R.J., Tuipulotu, D.E., Amos, T.G., O'Meally, D., Richardson, M.F., Russell, T.L., Vallinoto, M., Carneiro, M., Ferrand, N., Wilkins, M.R., et al. (2018). Draft genome assembly of the invasive cane toad, Rhinella marina. GigaScience *7*, giy095.
28. Session, A.M., Uno, Y., Kwon, T., Chapman, J.A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., et al. (2016). Genome evolution in the allotetraploid frog Xenopus laevis. Nature *538*, 336–343.
29. Sun, Y.-B., Xiong, Z.-J., Xiang, X.-Y., Liu, S.-P., Zhou, W.-W., Tu, X.-L., Zhong, L., Wang, L., Wu, D.-D., Zhang, B.-L., et al. (2015). Whole-genome sequence of the Tibetan frog Nanorana parkeri and the comparative evolution of tetrapod genomes. Proc. Natl. Acad. Sci. USA *112*, E1257–E1262.
30. Wu, W., Gao, Y.-D., Jiang, D.-C., Lei, J., Ren, J.-L., Liao, W.-B., Deng, C., Wang, Z., Hillis, D.M., Zhang, Y.-P., and Li, J.T. (2022). Genomic adaptations for arboreal locomotion in Asian flying treefrogs. Proc. Natl. Acad. Sci. USA *119*, e2116342119.
31. Rozen-Rechels, D., Dupoué, A., Lourdais, O., Chamaillé-Jammes, S., Meylan, S., Clobert, J., and Le Galliard, J.F. (2019). When water interacts with temperature: Ecological and evolutionary implications of thermo-hydroregulation in terrestrial ectotherms. Ecol. Evol. *9*, 10029–10043.
32. Yu, E.Z., Hallenbeck, J.M., Cai, D., and McCarron, R.M. (2002). Elevated arylalkylamine-N-acetyltransferase (AA-NAT) gene expression in medial habenular and suprachiasmatic nuclei of hibernating ground squirrels. Brain Res. Mol. Brain Res. *102*, 9–17.
33. Mohr, S.M., Bagriantsev, S.N., and Gracheva, E.O. (2020). Cellular, molecular, and physiological adaptations of hibernation: the solution to environmental challenges. Annu. Rev. Cell Dev. Biol. *36*, 315–338.
34. Aragonés, J., Schneider, M., Van Geyte, K., Fraisl, P., Dresselaers, T., Mazzone, M., Dirkx, R., Zacchigna, S., Lemieux, H., Jeoung, N.H., et al. (2008). Deficiency or inhibition of oxygen sensor Phd1 induces hypoxia tolerance by reprogramming basal metabolism. Nat. Genet. *40*, 170–180.
35. Logan, S.M., and Storey, K.B. (2021). MicroRNA expression patterns in the brown fat of hibernating 13-lined ground squirrels. Genomics *113*, 769–781.
36. Bailey, C.G., Ryan, R.M., Thoeng, A.D., Ng, C., King, K., Vanslambrouck, J.M., Auray-Blais, C., Vandenberg, R.J., Bröer, S., and Rasko, J.E.J. (2011). Loss-of-function mutations in the glutamate transporter SLC1A1 cause human dicarboxylic aminoaciduria. J. Clin. Invest. *121*, 446–453.
37. Nakada, T., Zandi-Nejad, K., Kurita, Y., Kudo, H., Broumand, V., Kwon, C.Y., Mercado, A., Mount, D.B., and Hirose, S. (2005). Roles of Slc13a1 and Slc26a1 sulfate transporters of eel kidney in sulfate homeostasis and osmoregulation in freshwater. Am. J. Physiol. Regul. Integr. Comp. Physiol. *289*, R575–R585.
38. Granchi, D., Baldini, N., Ulivieri, F.M., and Caudarella, R. (2019). Role of citrate in pathophysiology and medical management of bone diseases. Nutrients *11*, 2576.
39. Toyoda, Y., Takada, T., Miyata, H., Matsuo, H., Kassai, H., Nakao, K., Nakatochi, M., Kawamura, Y., Shimizu, S., Shinomiya, N., et al. (2020). Identification of GLUT12/SLC2A12 as a urate transporter that regulates the blood urate level in hyperuricemia model mice. Proc. Natl. Acad. Sci. USA *117*, 18175–18177.
40. Smith, D.E., Clémençon, B., and Hediger, M.A. (2013). Proton-coupled oligopeptide transporter family SLC15: physiological, pharmacological and pathological implications. Mol. Aspects Med. *34*, 323–336.
41. Inui, K.-I., Masuda, S., and Saito, H. (2000). Cellular and molecular aspects of drug transport in the kidney. Kidney Int. *58*, 944–958.
42. Li, N., Zhou, J., Wang, H., Mu, C., Shi, C., Liu, L., and Wang, C. (2020). Transcriptome analysis of genes and pathways associated with metabolism in Scylla paramamosain under different light intensities during indoor overwintering. BMC Genom. *21*, 775–815.
43. Manna, P.R., Stetson, C.L., Slominski, A.T., and Pruitt, K. (2016). Role of the steroidogenic acute regulatory protein in health and disease. Endocrine *51*, 7–21.
44. Whirledge, S., and Cidlowski, J.A. (2010). Glucocorticoids, stress, and fertility. Minerva Endocrinol. *35*, 109–125.

45. Frøbert, A.M., Toews, J.N.C., Nielsen, C.G., Brohus, M., Kindberg, J., Jessen, N., Fröbert, O., Hammond, G.L., and Overgaard, M.T. (2022). Differential Changes in Circulating Steroid Hormones in Hibernating Brown Bears: Preliminary Conclusions and Caveats. Physiol. Biochem. Zool. *95*, 365–378.
46. Rochira, V., and Carani, C. (2000). In Estrogens, Male Reproduction and Beyond, K.R.F. Endotext, B. Anawalt, M.R. Chrousos, E. Corpas, K. Dhatariya, K. Dungan, and J. Hofland, et al., eds.
47. Niwa, Y.S., and Niwa, R. (2014). Neural control of steroid hormone biosynthesis during development in the fruit fly Drosophila melanogaster. Genes Genet. Syst. *89*, 27–34.
48. Turnbull, A.V., and Rivier, C.L. (1999). Regulation of the hypothalamic-pituitary-adrenal axis by cytokines: actions and mechanisms of action. Physiol. Rev. *79*, 1–71.
49. Chen, Y., Liu, Y., Wang, Y., Zhang, Y., Xie, W., Zhang, H., Weng, Q., and Xu, M. (2023). The expression of cholesterol synthesis and steroidogenic markers in females of the Chinese brown frog (Rana dybowskii) during pre-spawning and pre-hibernation. Am. J. Physiol. Regul. Integr. Comp. Physiol. *325*, R750–R758.
50. Kouba, A.J., Langhorne, C.J., Willard, S.T., Smith, T., Kouba, C.K., and Clulow, J. (2022). Spermiation response to exogenous hormone therapy in hibernated and non-hibernated boreal toads (Anaxyrus boreas boreas). Reprod. Fertil. Dev. *34*, 453–460.
51. Xiao, C.-L., Chen, Y., Xie, S.-Q., Chen, K.-N., Wang, Y., Han, Y., Luo, F., and Xie, Z. (2017). MECAT: fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. Nat. Methods *14*, 1072–1074.
52. Hu, J., Fan, J., Sun, Z., and Liu, S. (2020). NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics *36*, 2253–2255.
53. Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol. *16*, 259.
54. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell systems *3*, 99–101.
55. Dudchenko, O., Batra, S.S., Omer, A.D., Nyquist, S.K., Hoeger, M., Durand, N.C., Shamim, M.S., Machol, I., Lander, E.S., Aiden, A.P., et al. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. Science *356*, 92–95.
56. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell systems *3*, 95–98.
57. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. USA *117*, 9451–9457.
58. Chen, N. (2004). Using Repeat Masker to identify repetitive elements in genomic sequences. Curr. Protoc. Bioinformatics *5*. 4.10.11-14.10.14.
59. Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. *34*, 435–439.
60. Korf, I. (2004). Gene finding in novel genomes. BMC Bioinf. *5*, 59.
61. Gish, W., and States, D.J. (1993). Identification of protein coding regions by database similarity search. Nat. Genet. *3*, 266–272.
62. Slater, G.S.C., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinf. *6*, 31.
63. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. *8*, 1494–1512.
64. Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. Genome Res. *12*, 656–664.
65. Wu, T.D., and Watanabe, C.K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics *21*, 1859–1875.
66. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. *9*, R7–R22.
67. Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol. Biol. Evol. *38*, 5825–5829.
68. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. *20*, 238–314.
69. 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.
70. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591.
71. 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.
72. 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.
73. 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.
74. Love, M., Anders, S., and Huber, W. (2014). Differential analysis of count data–the DESeq2 package. Genome Biol. *15*, 10–1186.
75. 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*, 284–287.
76. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinf. *9*, 559–613.
77. Geng, X., Li, W., Shang, H., Gou, Q., Zhang, F., Zang, X., Zeng, B., Li, J., Wang, Y., Ma, J., et al. (2017). A reference gene set construction using RNA-seq of multiple tissues of Chinese giant salamander, Andrias davidianus. GigaScience *6*, 1–7.
78. Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mob. DNA *6*, 11–16.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Latimeria chalumnae reference genome	National Center for Biotechnology Information	GCF_000225785.1
Anolis carolinensis reference genome	National Center for Biotechnology Information	GCF_000090745.1
Thamnophis elegans reference genome	National Center for Biotechnology Information	GCF_009769535.1
Python bivittatus reference genome	National Center for Biotechnology Information	GCF_000186305.1
Zootoca vivipara	National Center for Biotechnology Information	GCF_011800845.1
Anas platyrhynchos	National Center for Biotechnology Information	GCF_015476345.1
Gallus gallus	National Center for Biotechnology Information	GCF_016699485.2
Mus musculus	National Center for Biotechnology Information	GCF_000001635.27
Homo sapiens	National Center for Biotechnology Information	GCF_000001405.40
Bufo gargarizans	National Center for Biotechnology Information	GCF_014858855.1
Bufo bufo	National Center for Biotechnology Information	GCF_905171765.1
Rana temporaria	National Center for Biotechnology Information	GCF_905171775.1
Fejervarya multistriata	This paper	CNGBdb:CNA0069583
Illumina short-reads data of Fejervarya multistriata	This paper	CNGBdb:CNR0949994
Illumina short-reads data of Fejervarya multistriata	This paper	CNGBdb:CNR0949995
PacBio reads data of Fejervarya multistriata	This paper	CNGBdb:CNR0949996
PacBio reads data of Fejervarya multistriata	This paper	CNGBdb:CNR0949997
Hi-C data of Fejervarya multistriata	This paper	CNGBdb:CNR0949998
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0949999
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950000
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950001
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950002
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950003
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950004
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950005
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950006
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950007
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950008
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950009
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950010

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950011
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950012
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950013
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950014
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950015
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950016
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950017
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950018
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950019
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950020
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950021
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950022
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950023
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950024
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950025
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950026
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950027
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950028
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950029
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950030
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950031
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950032
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950033
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950034
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950035
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950036
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950037
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950038
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950039
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950040
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950041
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950042
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950043
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950044
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950045
RNA-seq data of Fejervarya multistriata	This paper	CNGBdb:CNR0950046
Software and algorithms		
MECAT2 default version	Xiao et al.[51]	https://github.com/xiaochuanle/MECAT2
NextPolish version 1.4.0	Hu et al.[52]	https://github.com/Nextomics/NextPolish
HiC-Pro version 2.2	Servant et al.[53]	https://github.com/nservant/HiC-Pro
Juicer version 1.5	Durand et al.[54]	https://github.com/aidenlab/juicer
3D-DNA version 180922	Dudchenko et al.[55]	https://github.com/aidenlab/3d-dna
Juicebox	Durand et al.[56]	https://github.com/aidenlab/Juicebox
RepeatModeler version 2.0.1	Flynn et al.[57]	http://www.repeatmasker.org
RepeatMasker version 4.1.1	Chen[58]	http://www.repeatmasker.org

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Augustus version 3.3.3	Stanke et al.[59]	https://github.com/Gaius-Augustus/Augustus
Snap version 2006-07-28	Korf[60]	https://github.com/KorfLab/SNAP
Tblastn version 2.11.0	Gish and States[61]	https://blast.ncbi.nlm.nih.gov/
Exonerate version 2.2.0	Slater and Birney[62]	https://www.ebi.ac.uk/about/vertebrate-genomics/software
Trinity version 2.11.0	Haas et al.[63]	https://github.com/trinityrnaseq/trinityrnaseq
Blat version 36	Kent[64]	https://sourceforge.net/projects/blat/
Gmap version 2017-11-15	Wu and Watanabe[65]	http://www.gene.com/share/gmap
EvidenceModeler version 1.1.1	Haas et al.[66]	https://github.com/EVidenceModeler
eggnog-mapper version 2	Cantalapiedra et al.[67]	https://github.com/eggnogdb/eggnog-mapper
Orthofinder version 2.3.12	Emms and Kelly[68]	https://github.com/davidemms/OrthoFinder
iqtree version 1.6.12	Nguyen et al.[69]	http://www.iqtree.org/
PAML version 4.9	Yang[70]	http://abacus.gene.ucl.ac.uk/software/paml.html
CAFÉ version 2.0	De Bie et al.[71]	https://hahnlab.github.io/CAFE/src_docs/html/index.html
Hisat2 version 2.2.1	Kim et al.[72]	http://daehwankimlab.github.io/hisat2/
Stringtie version 2.1.4	Pertea et al.[73]	https://ccb.jhu.edu/software/stringtie/
DESeq2	Love et al.[74]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
clusterProfiler	Yu et al.[75]	https://guangchuangyu.github.io/software/clusterProfiler/
WGCNA version 1.69	Langfelder and Horvath[76]	https://horvath.genetics.ucla.edu/html/CoexpressionNetwork/Rpackages/WGCNA/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Wenbo Liao (liaobo_0_0@126.com).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The raw reads, chromosomal assembly and gene annotations of the paddy frog were deposited in the Chinese National GeneBank Database (CNGBdb) under project accession number CNP0004609 and were publicly available at the date of publication. Accession numbers of the data files are listed in the key resources table. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animals

Genomic sequencing was conducted on a single specimen (female, sample ID: LJT-LAB2021157) of *F. multistriata* collected in Nanchong City (30.796°N, 106.046°E; elevation: 295 m) in Sichuan Province, Southwestern China. We collected RNA samples for analysis by extracting tissues from a minimum of three gender-neutral individual replicates of *F. multistriata*, including the brain, dorsal skin, heart, kidneys, liver, muscles, lungs, and spleen. The selection of samples was based on similar body lengths, and the samples were collected under different conditions within a consistent time frame (within 2 h). The sampling location in Suining City (30.374°N, 105.382°E; elevation: 301 m) remained the same for both hibernating and non-hibernating frogs. Paddy frogs in hibernation were captured underground during the winter, which represented their hibernating state, while frogs in an active state were captured in the same field during the spring. The frogs were immediately euthanized at the capture site, and representative tissues were dissected for analysis. All tissue samples were promptly flash-frozen in liquid nitrogen until

RNA extraction was performed. The experiments were performed in compliance with the Institutional Research Ethics Committee of China West Normal University (X-17002).

METHOD DETAILS

Genome sequencing

We employed three different types of whole-genome sequencing techniques, including Illumina short-read sequencing, PacBio long-read sequencing, and Hi-C. Libraries with an insert size of 350 were prepared according to the manufacturer's instructions for Illumina short-read sequencing. Sequencing was performed on an Illumina NovaSeq 6000 platform using 150 bp paired-end sequences. Libraries with 20 kb insert sizes were prepared following the manufacturer's protocol for PacBio long-read sequencing. The sequencing was conducted on a PacBio Sequel II platform. The library for Hi-C sequencing was prepared following the manufacturer's instructions. Subsequently, the library was sequenced on the BGI MGISEQ-2000 platform in PE150 mode.

RNA extraction was performed on the brain, dorsal skin, heart, kidneys, liver, lungs, muscle, and spleen utilizing the standard TRIzol method.[77] The extracted RNA samples were evaluated for quantity, purity, and integrity using state-of-the-art techniques. The RNA was quantified using the Qubit 2.0 fluorometer, while the NanoDrop spectrophotometer was employed to assess purity. To ensure the integrity of the RNA samples, the Agilent 2100 Bioanalyzer was utilized, employing industry-standard protocols. RNA-sequencing libraries were prepared to facilitate the subsequent analysis following the quality assessment of the RNA samples. The libraries were subjected to high-throughput sequencing on the NovaSeq 6000 platform, which provided accurate and efficient sequencing capabilities.

Genome assembly and annotation

We generated the *de novo* assembly using the default version of MECAT2.[51] The errors in the assembly generated from the PacBio long reads were fixed by two rounds of self-correction using the TGS reads, and further polished with three iterations of the NGS reads using NextPolish v.1.4.0.[52]

We employed RepeatModeler v.2.0.1[57] to build the *de novo* repeat library. Second, we searched for repeats belonging to the Amphibia lineages, including paddy frog, from the Repbase database[78] as a known library. The two libraries were combined and then searched in the assembled Midas cichlid genome with optimized parameters "-nolow -gff -poly -a -inv -e rmblast" to obtain a complete list of repeats in RepeatMasker v.4.1.1.[58]

We applied three distinct strategies to locate the protein-coding regions in the assembled paddy frog genome, including *de novo*, homolog-based, and transcript-based predictions. Two programs, such as Augustus v.3.3.3[59] and Snap v.2006-07-28,[60] were used for *de novo* prediction. The proteins from three representative species, including common toad (*Bufo bufo*; GCF_905171765.1), Asiatic toad (*Bufo gargarizans*; GCF_014858855.1), and common frog (*Rana temporaria*; GCF_905171775.1) were used for the homolog-based prediction performed with Tblastn v.2.11.0[61] and Exonerate v.2.2.0.[62] We first generated *de novo* assembly of the transcriptome using Trinity v.2.11.0 for the transcripts-based prediction,[63] and then the raw transcripts were mapped onto the assembled genome using Blat v.36[64] and Gmap v.2017-11-15,[65] followed by verification and integration to assemble the spliced alignment pipeline.[66] Finally, we employed EvidenceModeler v.1.1.1[66] to combine these results generated from the three strategies and the false predictions with a stop codon(s) inside the coding regions were removed. We used Eggnog-mapper v.2[67] to perform the functional annotation and understand the biological functions of the deduced proteins.

Comparative genomic analysis

After carefully evaluating the genomic data of these species with a focus on those with exceptional genome assembly and gene annotation quality, we utilized the high-quality genome assemblies of 12 representative species from various taxonomic groups to enhance our analysis. Our evaluation process involved scrutiny of the data to ensure its reliability and accuracy. The species included the paddy frog, which was the focus of this study. Additionally, we incorporated the coelacanth (*Latimeria chalumnae*; GCF_000225785.1) as the outgroup. The amphibian species used comprised the common toad, the Asiatic toad, and the common frog. The reptiles included the Green anole (GCF_000090745.1), the Western terrestrial garter snake (GCF_009769535.1), the Common lizard (GCF_011800845.1), and the Burmese python (GCF_000186305.1). Our avian representatives consisted of the chicken (GCF_016699485.2) and mallard (GCF_015476345.1). Finally, we incorporated two mammalian genomes, including humans (GCF_000001405.40) and the house mouse (GCF_000001635.27).

We classified the amphibian and reptile lineages as hibernating clades due to their shared characteristics associated with hibernation. In contrast, we categorized the bird, mammal, and coelacanth lineages as non-hibernating clades, distinguishing them from the other groups in our analysis.

We investigated the gene families based on sequence similarity and constructed a species tree using single-copy genes. These genes were categorized by gene-family clustering in Orthofinder v.2.3.12[68] among the examined species. A phylogenetic tree was constructed using the inside transversion of 4-fold degenerate sites, employing Iqtree v.1.6.12[69] with 10,000 super-fast bootstraps. We utilized two appropriate time-calibrated points to calibrate the species tree: the human-amphibian split (347–358 million years ago [Mya]) and the frog-toad split (145–160 Mya) estimated by TimeTree. We calibrated the species tree at a timing scale using MCMCtree in PAML v.4.9.[70] A log-normal independent molecular clock model was employed, generating 1,000,000 samples, with the first 25% discarded.

To eliminate potential false results, we purged the gene family clusters that exhibited an excessive number of members in some species while other species had no members. The remaining gene clusters were analyzed for changes in members using CAFÉ v.2.0.[71] We focused on identifying the expanded and contracted gene families that appeared in the ancestral branch of the hibernating lineages, specifically in the paddy frog.

Further analysis involved extracting a subset of single-copy genes between hibernating and non-hibernating animals to identify the PSGs and REGs in the paddy frog genome. The CODEML branch model in PAML was employed to detect the REGs in the paddy frog. The null model assumed that the non-synonymous to synonymous substitution ratio (dN/dS) in each branch was equal, while the alternative model allowed for variations in dN/dS in the ancestral branch of the hibernating clades. p values were calculated using the likelihood ratio test (LRT) comparing the two models. Significant REGs were determined based on a corrected p value <0.05 using the Bonferroni method.

Similarly, the CODEML branch-site model in PAML was used to identify potential PSGs. The dN/dS value of each site on each branch was fixed to 1 under the null hypothesis, whereas the alternative hypothesis allowed for variable dN/dS values at specific sites within the hibernating clades. The p values for the LRT were obtained following the CODEML process. PSGs were defined based on a corrected p value <0.05 and the presence of at least one positively selected site with a posterior probability >0.90, as determined by Bayes empirical Bayes analysis.

Identification and functional enrichment of the DEGs

To explore the gene transcriptional differences between the hibernating and the normal state of the paddy frog, we performed a gene differential expression analysis among the brain, dorsal skin, heart, kidneys, liver, muscle, lungs, and spleen (n = 3). All of the transcriptomes were mapped to our paddy frog genome assembly using Hisat2 v.2.2.1[72] to obtain the coverage depth of every gene. Then, read mapping counts were calculated based on gene annotation using Stringtie v.2.1.4.[73] We used DESeq2[74] to identify the DEGs based the read mapping counts of each gene between the hibernating and the normal state with 0.05 as the false discovery rate cut-off and a log 2-fold change cut-off of 1.5. We divided the upregulated and downregulated DEGs and enriched them to understand functional convergence using the R package 'clusterProfiler'.[75] Furthermore, we emphasized the gene functional discussion based on the enrichment analysis. Our main objective was to investigate the overall variation trend within the data rather than extensively examining the functions of individual genes.

Gene module and WGCNA analyses

The read mapping counts were converted into transcripts per kilobase million (TPM) counts using an in-house Perl script. Only the TPM counts for DEGs previously identified were extracted, and a trait matrix was constructed. In the matrix, DEGs from tissues sampled during hibernation were assigned values of 1 or 2, while these DEGs had a value of 0 in other tissues.

The dataset matrix was integrated to construct a co-expression network using the R package WGCNA v.1.69.[76] The most suitable soft threshold for network building was determined by selecting the best scale-free R2 fit. Modules were defined using a topological overlap dendrogram with a minimum module size of 30 genes, and the threshold cut height for merging modules was 0.25. Genes with similar transcription levels were clustered into co-expression modules based on the trait matrix.

Module eigengenes (MEs), representing the first principal component of the expression matrix for each module, were calculated. Pearson's correlation coefficient analysis was performed to identify the modules with the highest correlations with each tissue. A correlation coefficient ≥ 0.9 between the ME and gene expression in a specific clustered module was indicative of a hub gene within that module, suggesting a high degree of connectivity with other genes in the module.

Molecular pathway enrichment analysis

We performed a gene integration analysis in the paddy frog, focusing on the expanded gene families, the contracted gene families, the REGs, the PSGs, the hub genes in the co-expression networks, and the DEGs that were upregulated or downregulated in the sampled tissues. We enriched these genes in molecular pathways using the Kyoto Encyclopedia of Genes and Genomes database and the R package 'clusterProfiler'.[75] The significance of pathway enrichment was determined using the Bonferroni method to correct the p values. We constructed heatmaps to visualize the relationships between the genes and evolutionary or gene regulatory factors within the enriched molecular pathways.

QUANTIFICATION AND STATISTICAL ANALYSIS

The expanded gene families and the REGs and PSGs were identified with CAFÉ v.2.0[71] and PAML v.4.9[70] respectively. Each tool implements their own method to assess significance, as described in the "method details" section. All DEGs identification and enrichment analyses described in the "method details" section were performed with R package 'DESeq2'[74] and 'clusterProfiler',[75] respectively. This package implements a hypergeometric test. p values for enrichment tests were corrected with the Benjamini-Hochberg method. Gene co-expression analysis were done with the WGCNA package[76] in R. Each tool implements their own method to assess significance.