

RESEARCH

Open Access



Polyploidization of *Indotyphlops braminus*: evidence from isoform-sequencing

Fei Zhu^{1*†}, Jing Lu^{1†}, Ke Sun¹, Cao Deng² and Yu Xu¹

Abstract

Background *Indotyphlops braminus*, the only known triploid parthenogenetic snake, is a compelling species for revealing the mechanism of polyplloid emergence in vertebrates.

Methods In this study, we applied PacBio isoform sequencing technology to generate the first full-length transcriptome of *I. braminus*, aiming to improve the understanding of the molecular characteristics of this species.

Results A total of 51,849 nonredundant full-length transcript assemblies (with an N50 length of 2980 bp) from *I. braminus* were generated and fully annotated using various gene function databases. Our analysis provides preliminary evidence supporting a recent genome duplication event in *I. braminus*. Phylogenetic analysis indicated that the divergence of *I. braminus* subgenomes occurred approximately 11.5~15 million years ago (Mya). The full-length transcript resource generated as part of this research will facilitate transcriptome analysis and genomic evolution studies in the future.

Keywords *Indotyphlops braminus*, Polyploidization, Transcriptome, Sequencing, Evolution

Background

Indotyphlops braminus, previously known as *Ramphotyphlops braminus* [1–3], is classified in the genus *Indotyphlops* of the family Typhlopidae. It is one of the smallest snake species, with a body length of 7–17 cm. Although its origin was speculated to be southern or eastern Asia, *I. braminus* is now identified as a globally

invasive species. It has been documented in tropical and subtropical regions worldwide, except South America [1, 4, 5]. The species is predominantly located in provinces south of the Yangtze River in China [6]. The extensive invasion of *I. braminus* can be partly attributed to the potted plant trade and its ability for parthenogenetic reproduction [3, 7]. Parthenogenesis, a reproductive strategy where a female can reproduce without male involvement to create an entire population [8, 9], has been widely studied for its long-term consequences [10, 11]. As the only truly parthenogenetic vertebrates, reptiles provide critical insights into the persistence of sexual reproduction [11, 12]. Furthermore, *I. braminus* is an allotriploid species, that results from hybridization [2, 3, 5]. Allopolyploids potentially benefit from heterosis, by harbouring multiple gene copies that can evolve new or varied functions, facilitating niche expansion and adaptation to environmental changes [13]. McDowell's [14] initial proposal that *I. braminus* is an all-female species

*Fei Zhu and Jing Lu contributed equally to this work.

This work was supported by the Joint Fund of the National Natural Science Foundation of China and the Karst Science Research Center of Guizhou Province (grant number U1812401) and the new seedling plans of Guizhou Normal University ([2021]B03).

*Correspondence:

Fei Zhu

defish@163.com

¹School of Life Sciences, Guizhou Normal University, 550025 Guiyang, Guizhou, China

²Department of Bioinformatics, DNA Stories Bioinformatics Center, 610000 Chengdu, China



was later confirmed by Nussbaum [1]. Nevertheless, limited information is available regarding the reproductive characteristics of this diminutive snake [7]. Wynn et al. [2] and Ota et al. [3] demonstrated that *I. braminus* as a triploid asexual species according to karyotyping. While karyotyping offers initial evidence, additional molecular biology research is necessary. Elucidating the timing and mechanism of polyploidy in this snake species will yield insights crucial for future research into polyploid vertebrates. A recent publication presented *I. braminus* draft genome, which exhibited a total length of 1.86 Gbp and an N50 scaffold size of 1.25 Mbp, indicative of a potentially chimeric single haplotype [5]. Acquiring high-quality genomic or transcriptomic data is vital for advancing related research.

Full-length transcriptome data enhance the understanding of gene content and refine genome annotation, facilitating detailed analysis of gene structure and transcriptional information [15, 16]. Polyploids, which have multiple chromosome sets, typically exhibit more intricate transcriptomes than diploids [17, 18]. In polyploids, duplicated genes may result in redundancies, potentially introducing new functions [19]. Additionally, a link might exist between an increase in gene numbers or genomes and phenotype complexity [20]. Assembling transcripts from complex polyploid genomes accurately can be challenging when using short-read sequencing technologies. Short-read sequencing risks errors, such as merging similar gene copies into one contig [21]. Single-molecule long-read sequencing, represented by isoform sequencing (Iso-seq) by Pacific Biosciences (PacBio), excels in accurately analysing transcript structural information [22, 23]. This approach is particularly advantageous for polyploid species due to its ability to differentiate homeologs [24, 25]. Long-read sequencing, which covers the entire transcript, can resolve complex repeats and provide additional information on transcript isoforms [16, 26]. A limitation of the PacBio platform is its elevated sequence error rate; however, PacBio SMRT software can enhance sequencing data accuracy with the reads of insert (ROIs) algorithm, which generates a circular consensus sequence (CCS), thereby reducing sequencing errors and improving data quality [26, 27]. This sequencing technology is currently effective for full-length transcriptome profiling across various organisms [28, 29], such as plants (*Saccharum officinarum*), invertebrates (*Litopenaeus vannamei*), and vertebrates (*Misgurnus anguillicaudatus* and *Bungarus multicinctus*) [30–33]. Advances in single-molecule sequencing and functional analysis technologies have enabled a growing body of research into genome replication mechanisms [34, 35].

In this study, we generated the first full-length transcriptome of *I. braminus* using Iso-seq technology. The main goals were threefold: (1) establish a reference

full-length transcriptome; (2) utilize this reference full-length transcriptome for phylogenetic analysis; and (3) explore a recent polyploidization event in *I. braminus*. Specifically, the aim of this study was to determine the timing of the genome duplication event. This study provides a comprehensive set of coding genes, offering preliminary evidence for the polyploidization of *I. braminus* at the molecular level. These data constitute an essential genetic resource that will facilitate further research in this field.

Methods

Sample collection and RNA extraction

Specimens were collected from Wenshan (Yunnan) and Wangmo (Guizhou) between 2019 and 2023. All individuals were maintained on moist soil in the laboratory before RNA extraction and chromosome preparation (animal handling and experiments were approved by the Ethics Committee at Guizhou Normal University, Permission number: 20,230,300,015). The experiments adhered to the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines [36]. All methods were conducted in compliance with relevant guidelines and regulations.

The specimens were euthanized using ethyl acetate and then preserved in 70% alcohol. For maximum mRNA extraction, five representative organs (brain, heart, liver, skin, and muscle) were collected and mixed from two healthy adult females. Total RNA was extracted from the mixed tissues using TRIzol reagent (Invitrogen, MA, USA) on dry ice, following the manufacturer's instructions. DNA was removed using TURBO DNase I (Promega, Beijing, China). RNA degradation and contamination were assessed by 1% agarose gel electrophoresis. RNA purity was determined using a NanoDrop 2000 microspectrophotometer (Thermo Scientific, USA; NanoDrop 2000 detection blank reference: DEPC water). The RNA integrity (RIN) was accurately measured with an Agilent 4200 (Agilent Technologies). Only RNA samples with a RIN ≥ 8 were considered suitable for cDNA library construction.

Library construction and PacBio sequencing

PolyA-containing mRNAs were enriched with oligo (dT) bead primers. The enriched mRNAs were reverse transcribed into cDNA using a Clontech SMARTer™ PCR cDNA Synthesis Kit (Clontech, CA, USA). Subsequently, the synthesized full-length cDNA was amplified via PCR. The cDNA fragments were purified using Pronex beads (Promega), with ratios varying according to transcript size. Purified cDNA was subjected to DNA damage repair, end repair, and ligation with SMRT dumbbell-type sequencing adapters. Following library construction, the Qubit 2.0 system (Life Technologies) was used

for quantification, and the Agilent 2100 system (Agilent Technologies) was used to verify library insert size. The SMRTbell template was annealed with a sequencing primer, bound to polymerase, and sequenced using the PacBio Sequel II platform for data acquisition.

Data processing and transcriptome assembly

High-quality CCSs were produced using the IsoSeq3 pipeline's CCS command (<https://github.com/Pacific-Biosciences/IsoSeq3>), with the following parameters: min_predicted_accuracy 0.9 --min_passes 1 --top_passes 100 --min_length 200 --max_length 100,000. The construction of full-length transcripts involved four steps: (1) obtaining full-length reads by primer removal and demultiplexing using lima (v2.2.0, <https://lima.how/>); (2) classifying CCS reads into full-length nonchimeric (FLNC) and non-full-length (nFL) reads based on splice primer and chimaera presence; (3) further refining FLNC reads using IsoSeq3's refine function (v3.4.0, parameter: --require_polyA), involving polyA tail identification and removal; and (4) deriving the final full-length transcripts by clustering sequences with the IsoSeq3 clustering function (v3.4.0, parameter: --use_qvs). To ensure no possible contamination from other organisms, 1,000 random reads were aligned against the NCBI NT database using BLASTN [37] (e-value $\leq 1e^{-5}$). The completeness of the transcriptome assembly was evaluated using the benchmarking universal single-copy orthologue (BUSCO, v5.2.2) [38].

Gene annotation

Gene structures were annotated using a homologous protein-based method with Genewise (v2.4.1) [39] (parameter: --tfor_sum_genesf_Gff_subs 0.01 --indel 0.01 --trans_pseudo) and GeMoMa (v1.6.3) [40]. The reference protein sets used were obtained from *Python molurus bivittatus* (NCBI RefSeq GCF_000186305.1), *Deinagkistrodon acutus* (GigaDB, <https://doi.org/10.5524/100196>), and *Protobothrops mucrosquamatus* (NCBI RefSeq GCF_001527695.2). The gene structures from GeneWise and GeMoMa were merged and the longest protein at each locus was selected for final annotation. The remaining unannotated genes were further predicted de novo using TransDecoder.

Representative protein sequences were annotated using the following five functional databases. BLASTP (v2.7.1, e-value $\leq 1e^{-5}$, identity $\geq 30\%$ and subject coverage $\geq 30\%$) [37] was utilized to perform searches against the NCBI NonRedundant Protein (NR, <http://www.ncbi.nlm.nih.gov>) and SwissProt [41] protein databases. Kofam (v1.3.0, e-value $\leq 1e^{-5}$) [42] for KO annotation in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [43]. Gene Ontology (GO) terms [44] and protein domain

(ProDom) [45] were predicted by InterProScan (v5.2) [46].

Identification of gene families

The protein sequences of 13 vertebrate species (including mammals, birds, amphibians, and reptiles) were downloaded from GigaDB or the NCBI (Supplementary Table S1). For gene loci with alternative splicing variants, only the longest transcript was selected. Genes with fewer than 50 amino acids were removed. Several closely related snake species and representative mammals, birds, amphibians, and reptiles were selected to ensure coverage of the major evolutionary clades. Self-to-self alignment of pooled protein sequences from species with available genomes was conducted using BLASTP (E-value of $1e^{-5}$) [37], with low-quality hits being removed (identity $< 30\%$ and coverage $< 30\%$) [47]. Orthologous groups were constructed from the filtered BLASTP results using OrthoFinder2 [48].

Evolution analysis and divergence time estimation

Single-copy gene families were extracted from the OrthoFinder2 results for the 13 species. The protein sequences of *I. braminus* were aligned with those of the obtained single-copy gene families to extract the reciprocal best hits (RBHs). Subsequently, single-copy gene families for these 14 species were generated. Protein alignment for each single-copy family was conducted using MUSCLE (v3.8.31) [49]. The corresponding coding sequence (CDS) alignments were back-translated from the corresponding protein alignments using PAL2NAL [50]. Gblocks [51] was used to extract the conserved CDS alignments. For phylogenetic tree construction, a supermatrix was created by concatenating the CDS alignments of single-copy families. Maximum likelihood (ML) trees for supergenes constructed from full-length and 4DTv sites, were generated using the GTR+I+Γ model with RaxML [52]. The three codon positions in the concatenated supermatrix were treated as separate partitions due to significant differences in evolutionary rates, corresponding to the 1st, 2nd and 3rd codon sites of the CDS. Divergence dates were estimated using a relaxed clock model via the PAML4.7 package [53]. The “independent rates model (clock=2)” and “JC69” models in the MCMCTREE program were used, running six million MCMC iterations after two million burn-in iterations [47]. For consistency, each data type was run twice through the program. In the first run, the chronogram was generated using FigTree (v1.4.0, <http://tree.bio.ed.ac.uk/>).

Polyplody analysis

Gene families encompassing 14 species were constructed using OrthoFinder2. In each family, all paralogues from *I. braminus* were retained, while only the longest parologue

was retained for the other 13 species. Proteins in each family were aligned with MUSCLE (v3.8.31) [49] using default parameters, and CDS alignments were generated from these using PAL2NAL [50]. The ML phylogenetic tree was constructed using RAxML [52] with the GTR+I+Γ model. Gene trees that conflicted with the species tree were filtered out. The divergence times in the gene tree were estimated using the MCMCTREE program in the PAML4.7 package [53]. MCMCTREE operated similarly to the above description, with the exception that CDS alignments were not partitioned [47].

The synonymous mutation rate (K_s) distribution of paralogues is commonly used to infer whole-genome duplications (WGDs) [54, 55]. K_s distributions for *I. braminus*, *P. bivittatus*, *Xenopus laevis* ($2n=4x=38$), and *Xenopus tropicalis* ($2n=2x=18$) were obtained using the WGDdetector [55]. The WGDdetector pipeline integrates gene family construction and K_s value estimation for paralogues pairs, plotting the distribution using an R script [55]. K_s plots indicate of past duplications, while karyotyping analysis directly infers contemporary ploidy. Chromosomes were prepared according to a standardized procedure [56]. For this experiment, a healthy adult female snake underwent an intraperitoneal injection of 0.1% colchicine for 8 h. After the snakes were euthanized with ethyl ether vapour, the digestive tract was dissected and immersed in 0.6% normal saline (NS). Intestinal samples were sectioned into small pieces and treated with hypotonic KCl (0.075 M). The cells were fixed in fresh cool Carnoy's fixative (glacial acetic acid/methanol, 1:3). Chromosome suspensions were prepared by dropping them onto clean slides, followed by staining and banding after drying. Conventional staining was achieved using 20% Giemsa solution for 3–5 min. After staining, the slides were rinsed thoroughly with running

water and dried before imaging or observation. Twenty optimal metaphase plates were selected for photographic analysis using a 100x objective microscope (E100, Nikon). Chromosomes were classified according to Levan et al. [57], and karyotype measurements were conducted using ImageJ software [58].

Results

Data summary

A total of 96 Gb of raw sequencing data was obtained from Iso-seq using the PacBio SMRT sequencing method. Following initial quality control, which involved the removal of adapter sequences and subreads < 50 bp in length, a total of 43,131,390 subreads with an average length of 2,226 bp were generated (Table 1 and Supplementary Fig. S1a). To assess accuracy, 1,000 random sequences were aligned to the NCBI NT database; 98% of the sequences were identified, most of which were similar to closely related reptiles. Thereafter, all subreads underwent CCS analysis, which produced 1,356,270 CCSs averaging 2,513 bp in length (Table 1 and Supplementary Fig. S1b). The FLNC reads were clustered using the cluster function of IsoSeq3 to correct errors in the third-generation sequencing data. Ultimately, 51,849 full-length transcriptomes were obtained, averaging 2,719 bp in length, with a maximum of 9,319 bp, an N90 of 1,791 bp, and an N50 of 2,980 bp (Fig. 1a). The completeness of the assembly was 89.4% as determined by BUSCO, with 19% as complete and single-copy BUSCOs, 66.4% as complete and duplicated BUSCOs, and 4% as fragmental BUSCOs (Fig. 1b).

Coding sequence prediction and gene annotation

Identifying CDSs is crucial for gene annotation, aiding preliminary gene structure analysis and providing valuable insights for functional annotation and evolutionary analysis [59, 60]. In this study, the full-length transcripts were annotated based on protein sequence information from homologous species. A total of 46,660 (89.99%) transcripts were successfully annotated, via the use of three software tools. The CDS lengths ranged from 36 to 7,875 bp, with an average length of 1,383 bp (Fig. 2). Among these sequences, 35,481 (68.43%) CDS transcripts matched the reference protein sequence, and 34,633 (66.8%) not only aligned with the reference but also had only one terminator.

Full-length nonredundant transcripts were annotated using five databases, and 46,406 (99.46%) transcripts were successfully identified. Functional annotation revealed 45,997 (98.58%), 43,368 (92.82%), 40,368 (86.52%), 35,055 (75.12%), and 30,590 (65.56%) transcripts annotated in the Iprscan, Nr, Swiss-Prot, GO, and KEGG databases, respectively (Fig. 3). According to the NR database, most transcripts were annotated to the Pythonidae and

Table 1 Assembly statistics for *I. braminus* transcriptome

Subreads	
Number of reads	43,131,390
Number of bases sequenced (bp)	96,030,005,322
Average length (bp)	2,226
CCS	
Number of reads	1,356,270
Number of CCS bases (bp)	3,407,675,187
Average length (bp)	2,513
Full-length transcriptome	
Total number	5,1849
Total length (bp)	141,014,862
Average length (bp)	2,719
Max length (bp)	9,319
Min length (bp)	87
N50 length (bp)	2,980
N90 length (bp)	1,791
GC content (%)	47.9

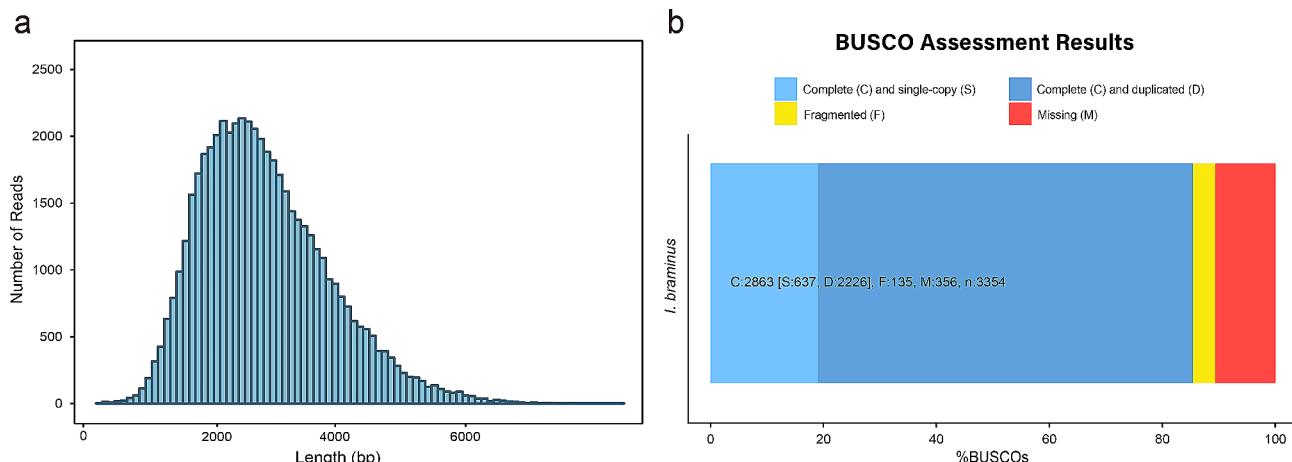


Fig. 1 Length distribution (a) and integrity assessment (b) of *I. braminus* transcripts

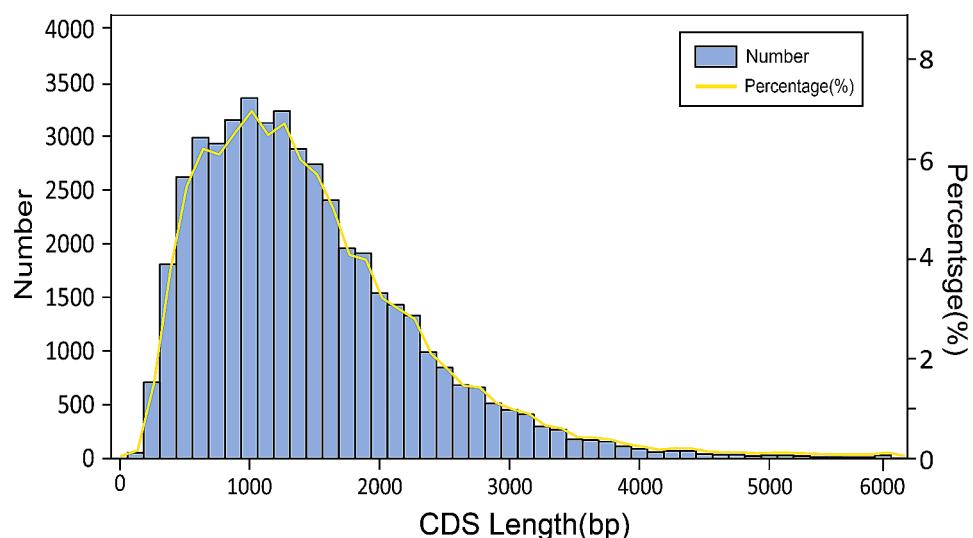


Fig. 2 Number, percentage and length distributions of coding sequences of *I. braminus* transcripts

Viperidae families. The 5 most common annotated species were *P. bivittatus* (19,383, 44.76%), *Probothrops mucrosquamatus* (7,274, 16.8%), *Thamnophis sirtalis* (3,409, 7.87%), *Pogona vitticeps* (3,136, 7.24%), and *Anolis carolinensis* (2,509, 5.79%) (Fig. 4a). With regard to GO annotation, the most enriched terms in the biological processes category were cellular process (13,498, 26.35%) and metabolic process (12,176, 23.77%). Within the molecular function category, the most enriched GO terms were binding (23,037, 58.83%) and catalytic activity (11,110, 28.37%). In the cellular component category, the most abundant GO terms were cell (7,456, 21.39%) and cell part (7,456, 21.39%) (Fig. 4b). According to the KEGG pathway annotation, 30,590 isoforms were annotated and assigned to 43 biological pathways (Fig. 4c). Numerous annotated isoforms were classified into pathways related to environmental information processing, organismal systems, and metabolism processing. In

particular, 10,196 (12.94%) isoforms were associated with “signal transduction”, indicating the importance of signal transduction-related genes in *I. braminus*.

Gene duplication

A large proportion of duplicate BUSCOs were identified in the integrity assessment of the full-length transcriptome. Of the 2,998 (89.4%) BUSCO groups in the transcriptome, 2,226 (66.4%) were duplicated, indicating that the gene duplication possibly resulted from WGD (Fig. 1b). We used WGDdetector [55] to estimate *Ks* values for four species (*I. braminus*, *P. bivittatus*, *X. laevis* (tetraploid), and *X. tropicalis*), and their *Ks* distributions were plotted. The *Ks* plots showed a clear *Ks* peak for *X. laevis* and *I. braminus* (Fig. 5). This finding suggested a recent gene duplication event in *I. braminus*.

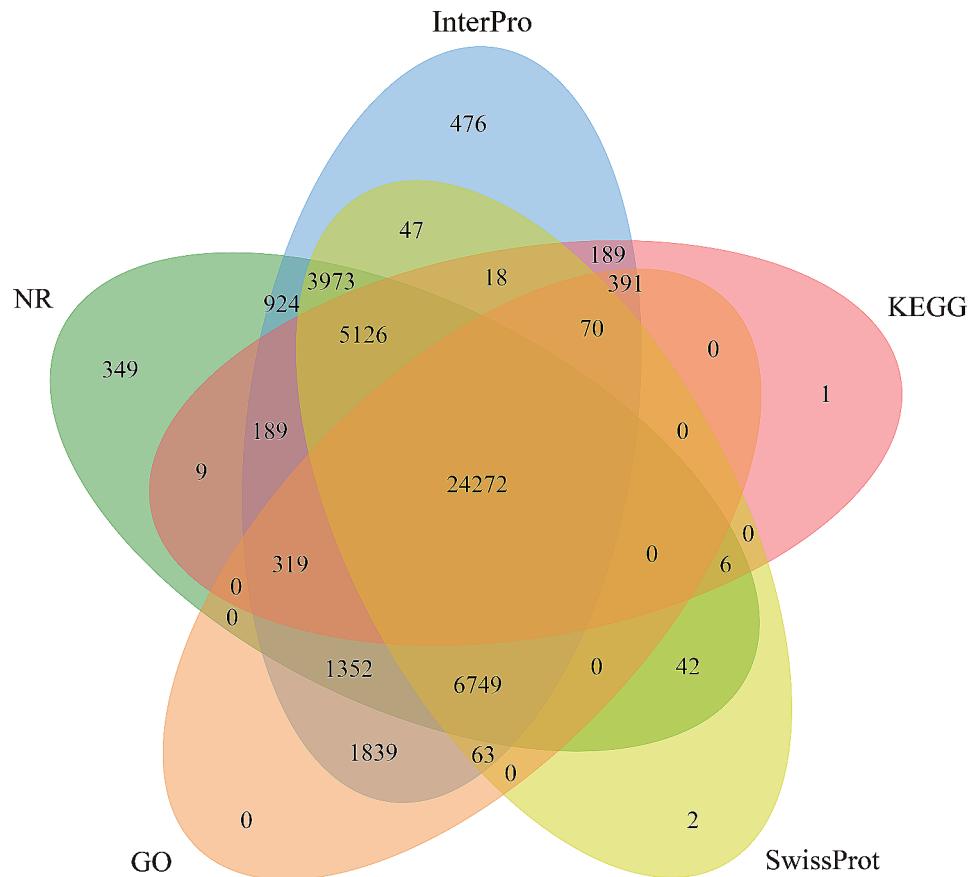


Fig. 3 Venn diagram of the annotations between the InterPro, NR, GO, KEGG, and Swiss-Prot databases

Karyological analysis

The examined female specimens of *I. braminus* had karyotypes of $3n=42$ chromosomes, with 8 macrochromosome triplets and 6 microchromosome triplets. This alignment was consistent with findings from previous studies by Ota et al. [3] and Patawang et al. [61]. Among the macrochromosomes, the first four pairs were larger and metacentric, while the other four included two metacentric pairs (pairs 5 and 8) and two submeta-subtelocentric pairs (pairs 6 and 7), as shown in Fig. 6. The fundamental number (NF, number of chromosome arms) was 60, and the karyotypic formula was as follows: $3n(42)=L^m_{12}+S^m_6+S^m_4+S^st_2+18$ microchromosomes.

Phylogenetic analyses and divergence time estimation

The OrthoFinder2 results revealed 3,249 single-copy gene families across 13 species (Supplementary Table S2). The protein sequences of *I. braminus* were aligned with those of single-copy gene families to extract RBHs, identifying 1,826 single-copy gene families across 14 species. Phylogenetic trees were constructed using the ML method in RaxML. The phylogenetic tree aligns with the snake sub-order estimates provided by Yan et al. [62] and Liu [63] based on mitochondrial genomes (Fig. 7a). The bootstrap

support value for each branch was 100 (Supplementary Fig. S2). The phylogenetic tree showed that nine snake species formed a monophyletic clade, with *I. braminus* diverging the earliest. The position of *I. braminus*, as the sister lineage to the other eight snakes, suggested its more ancient position in the evolutionary history of snakes.

Divergence dates were estimated under a relaxed clock model using the MCMCTREE program in the PAML4.7 package. Time calibration of the estimated tree was also conducted (Supplementary Table S3). Divergence dates for the 14 species were determined to be within a certain range (noted by green bars in Fig. 7a). In each gene family, all *I. braminus* paralogues were retained, while only the longest parologue was retained in the other 13 species. The gene divergence time was estimated according to previous methods, except for CDS alignment partitioning, which was not performed. The results revealed that the divergence between *I. braminus* and *P. bivittatus* occurred ~98.15 Mya (Fig. 7a), and the divergence of *I. braminus* subgenomes occurred more recently, approximately 11.5~15 Mya (Fig. 7b).

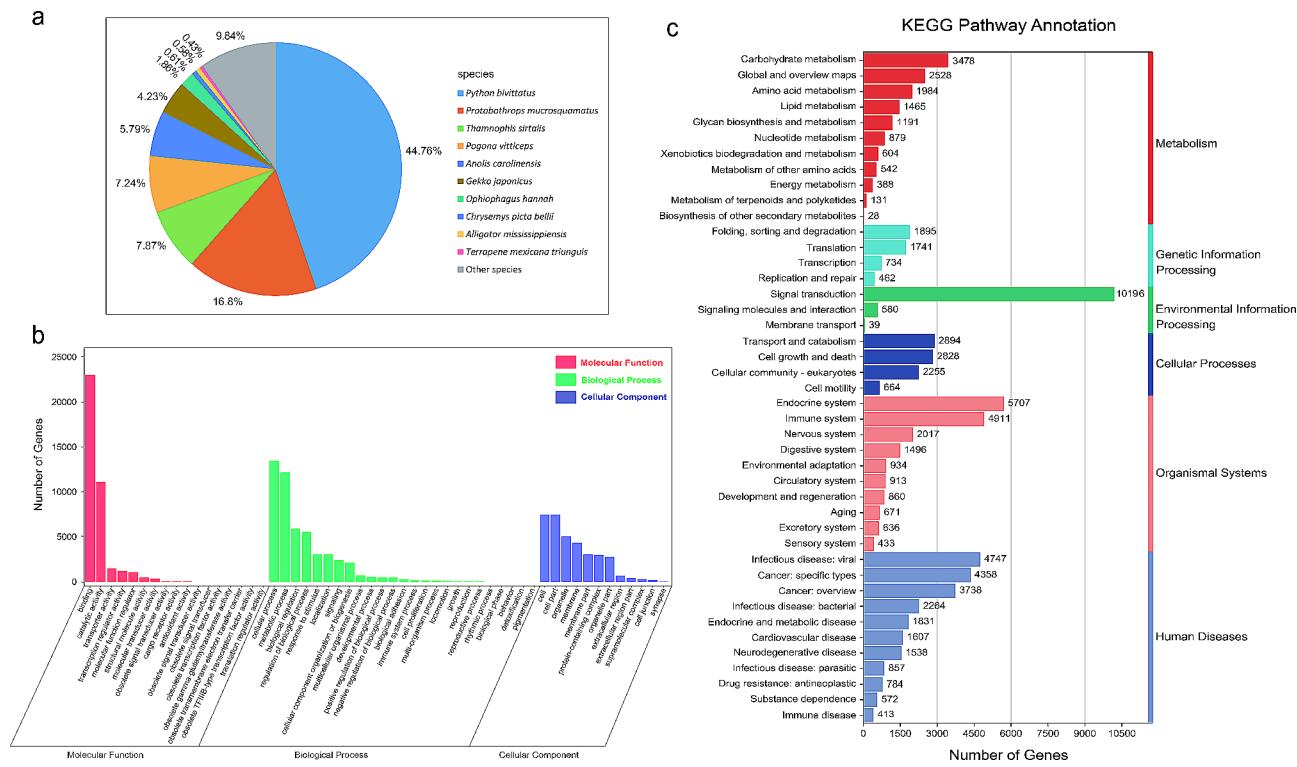


Fig. 4 Gene functional annotations in the public databases. **a** Distribution of homologous species annotated in the NR database (the first three species belonging to snakes); **b** Distribution of functional classifications based on GO terms; **c** Distribution of pathway classifications based on the KEGG database

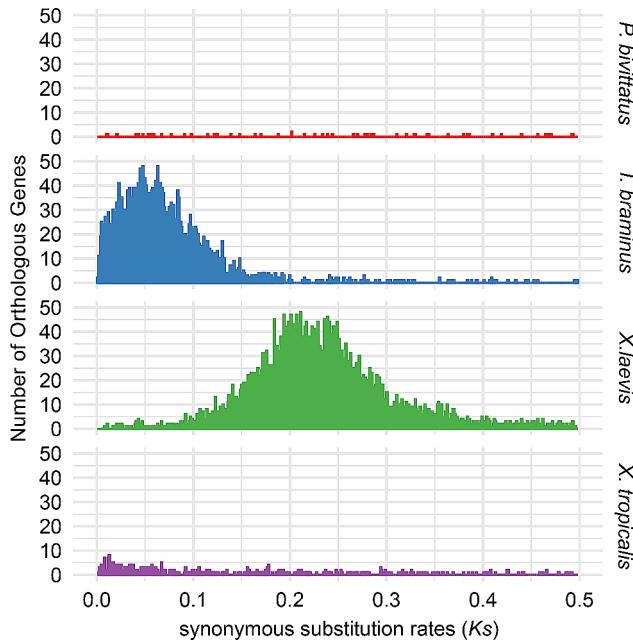


Fig. 5 Distributions of synonymous substitution rates (K_s) between paralogous genes of *P. bivittatus*, *I. braminus*, *X. laevis*, and *X. tropicalis*

Discussion

Transcriptome analysis provides crucial insights into genomic characteristics, including genome duplication [64]. Iso-seq, a third-generation sequencing technology,

has emerged as a potent tool in transcriptomics due to its single-molecule sequencing and long read capabilities [65]. In recent years, Iso-seq has greatly enhanced our understanding of the complex nature of the transcriptome. In this study, the Iso-seq platform was used to sequence and analyse the full-length transcriptome of *I. braminus*. Through assembly and splicing, 51,849 transcripts were ultimately obtained. The majority of the reads exhibited high accuracy, with most having a Phred quality score above 20 (indicating an error probability of 1%) and some above 60 (indicating an even lower error probability), emphasizing the reliability of the full-length transcript data. Subsequent transcript annotation, via multiple databases, provided deeper insights into the structure and function of the transcripts. According to the NR database, the species most closely related to *I. braminus* was *P. bivittatus*. Among the full-length transcriptome of *I. braminus*, 27,707 transcripts were annotated in the GO database; these genes were associated predominantly with biological processes, followed by molecular functions and cellular components. A total of 47,197 *I. braminus* transcripts were annotated in 43 KEGG pathways, with the top four pathways being involved in signal transduction, the endocrine system, the immune system, and infectious disease (viral). These findings highlight the importance of our PacBio transcript data as valuable resources and references for future

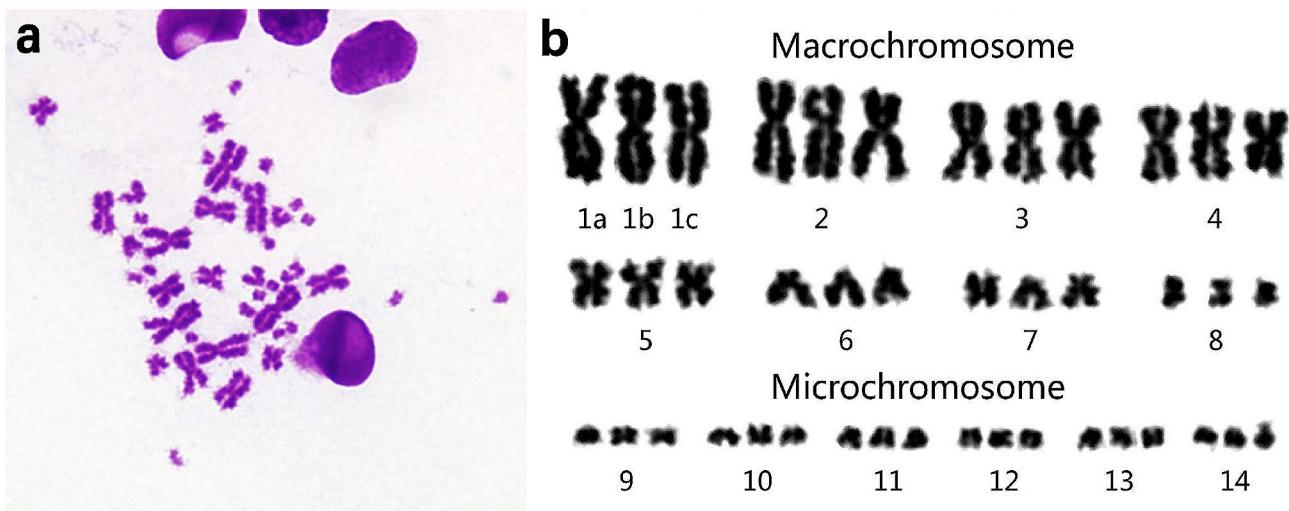


Fig. 6 Metaphase chromosome plates (a) and standardized karyotypes (b) of *I. braminus* according to conventional staining

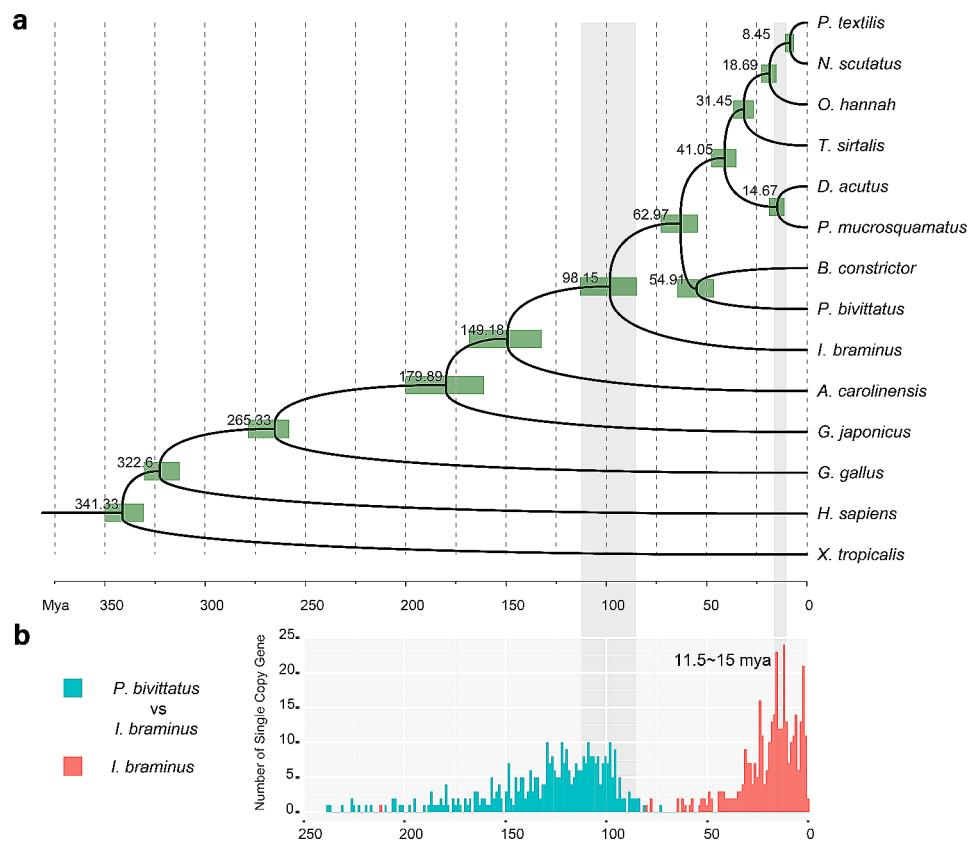


Fig. 7 Phylogenetic tree and divergence dates of 14 species. **a** Phylogenetic tree highlighting the phylogenetic position of *I. braminus*. The green bars are the time ranges of the divergence dates. The grey boxes correspond to the divergence date of *I. braminus*-*P. bivittatus*; **b** Divergence date of *I. braminus*-*P. bivittatus* (blue); The estimated divergence dates of *I. braminus* subgenomes (red). The total number of single-copy genes for *I. braminus*-*P. bivittatus* was 1826

studies, particularly in annotating reptile gene structures, conducting functional analysis, and performing pathway research. Additionally, our results enrich the genetic knowledge of *I. braminus*, aiding future research into

genes related to snake development, reproduction, and evolution.

Polyploidization, or WGD, is a typical feature of eukaryotic evolution, thought to confer selective benefits to polyploids and play a key role in speciation and

eukaryotic development [66, 67]. For instance, Wang et al. [68] showed that, compared with *Danio rerio*, *Cyprinus carpio* has experienced an additional WGD event, resulting in the divergence of common carp as an independent species from its common ancestor. Multiple genome duplication events occurred during the evolution of chordates, with some occurring near the origin of vertebrates [69, 70]. Our study revealed the presence of many duplicated genes in the *I. braminus* transcriptome, based on BUSCO assessments (Fig. 1b) and the *Ks* peak detected by the WGDdetector (Fig. 5). Furthermore, the gene count in *I. braminus* (46,660) surpassed that of other snakes, possibly due to genome duplication. Dating analysis using MCMCTREE suggested at least two distinct subgenomes in *I. braminus* (Fig. 7b). Integrating these findings with the karyotyping results, it can be concluded that *I. braminus* is triploid. *Ks* plots provide evidence of past duplication, while karyotyping results help identify contemporary polyploidy. Although each of these approaches has limitations, we considered putative polyploidization to be supported when these results were consistent. The most reliable evidence for WGD requires synteny-based analysis with high-quality whole genomes (three haplotypes) [71]. Therefore, further genomic studies are necessary to fully elucidate the mechanisms driving these gene duplication events.

Polyplody occurs more frequently in plants than in animals. It is observed in only a few species of insects, bony fish, amphibians, and reptiles [66, 67]. The reason for the scarcity of animal polyploidy was first proposed by Muller [72], who proposed that changes in chromosomes may impact reproductive mechanisms or sex determination. Consequently, polyplody is generally perceived as an evolutionary blind alley, primarily due to its association with unisexual reproduction [73, 74]. However, extensive research on polyploidy has shown that many animals exist as stable polyploids [66, 75]. Polyploids such as *I. braminus* thrived in terms of survival and reproduction, and even this unisexual polyploid species has existed for millions of years. In the book "*The Evolution of the Genome*," Gregory et al. [76] emphasized the potential advantages of polyploidy, including increased adaptability to harsh conditions and wider geographic ranges, increasing resistance to extinction and facilitating genealogical selection. The divergence of *I. braminus* subgenomes took place during the middle Miocene (11~17 Mya), coinciding with significant climatic events such as the Miocene Climatic Optimum (MCO) and a subsequent sudden cooling and Antarctic ice-sheet expansion phase, called the Middle Miocene Climate Transition (MMCT) [77, 78]. Global climatic changes during the Miocene period are hypothesized to have influenced the evolutionary trajectory of *I. braminus*, but the precise mechanisms underlying this phenomenon

remain to be elucidated. Fundamental questions about *I. braminus* persist, including the mechanism of polyploidization, the consequences of genome duplication, and gene interactions. Thus, there is a notable research gap regarding the genetic aspects of *I. braminus* polyploidy; revealing its evolutionary history and genomic characteristics in the postgenomic era is urgent. Notably, the recent publication of the draft genome sequence of *I. braminus* in Scientific Data offers a valuable resource for future research endeavours. Future efforts will focus on mapping transcripts to this genome assembly and establishing gene models related to the draft genome. In the context of rapid growth in molecular technology, using the transcriptome or genome sequence as an entry point for analysis may lead to additional insights.

Conclusions

In this study, we successfully obtained the full-length transcriptome of *I. braminus* using Iso-seq high-throughput sequencing technology, thereby providing a novel perspective for confirming the polyploidization of *I. braminus*. Our analysis provides preliminary evidence supporting a genome duplication event in *I. braminus*, with an estimated divergence date of its subgenomes between 11.5 and 15 Mya. These results provide valuable insights for future research into snake transcriptomes and genomes, aiding the exploration of other polyploid vertebrates. Additionally, this study has the potential to broaden the application of PacBio sequencing in vertebrate transcriptome research.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12863-024-01208-y>.

Supplementary Material 1

Acknowledgements

This work was financially supported by the Joint Fund of the National Natural Science Foundation of China and the Karst Science Research Center of Guizhou Province (grant number U1812401) and new seedling plans from Guizhou Normal University ([2021]B03). We would like to express our gratitude for the support provided by the aforementioned funding projects. We acknowledge the DNA Stories Bioinformatics Center for contributing to the transcriptome sequencing and for providing technical assistance with the data analyses.

Author contributions

FZ and JL conceived the study. FZ and KS carried out the sample collection. Data collection and analysis were performed by FZ, JL and CD. The first draft of the manuscript was written by FZ and JL. CD and YX read the article and modified it. YX acquired the funding. All the authors read and approved the final manuscript.

Funding

The research design and data collection were financially supported by the Joint Fund of the National Natural Science Foundation of China and the Karst Science Research Center of Guizhou Province (Grant number U1812401) and new seedling plans from Guizhou Normal University ([2021]B03). The data

analysis and interpretation were supported by research conducted by the DNA Stories Bioinformatics Center.

Data availability

The transcript sequences from the PacBio Iso-seq transcriptome are available at the NCBI Sequence Read Archive (SRA accession SRR24061511 <https://www.ncbi.nlm.nih.gov/sra/SRR24061511>).

Declarations

Ethics approval and consent to participate

All samples were collected following Chinese regulations for the Implementation of the Protection of Terrestrial Wild Animals (State Council Decree (1992) No. 13). All the experiments and methods were performed in accordance with the Animal Research: Reporting of In Vivo Experiments (ARRIVE) guidelines. All the methods were carried out in accordance with relevant guidelines and regulations. Animal handling and experimental procedures were approved by the Ethics Committee of Guizhou Normal University (Permission number: 20230300015).

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 3 March 2023 / Accepted: 14 February 2024

Published online: 26 February 2024

References

1. Nussbaum RA. The brahminy blind snake (*Ramphotyphlops braminus*) in the Seychelles Archipelago: distribution, variation, and further evidence for parthenogenesis. *Herpetologica*. 1980;215–21.
2. Wynn AH, Cole CJ, Gardner AL. Apparent triploidy in the unisexual brahminy blind snake, *Ramphotyphlops braminus*. *Am Mus Novit*. 1987;2868:1–7.
3. Ota H, Hikida T, Matsui M, Mori A, Wynn AH. Morphological variation, karyotype and reproduction of the parthenogenetic blind snake, *Ramphotyphlops braminus*, from the insular region of East Asia and Saipan. *Amphibia-Reptilia*. 1991;12:181–93. <https://doi.org/10.1163/156853891x00158>.
4. Wallach V. *Ramphotyphlops braminus* (Daudin): a synopsis of morphology, taxonomy, nomenclature and distribution (Serpentes: Typhlopidae). *Hama-dryad*. 2009;34:34–61.
5. Khedkar G, Kambayashi C, Tabata H, Takemura I, Minei R, Ogura A, et al. The draft genome sequence of the Brahma blind snake *Indotyphlops braminus*. *Sci Data*. 2022;9:1–7. <https://doi.org/10.1038/s41597-022-01530-z>.
6. Tian W, Jiang Y, Wu G, Hu Q, Zhao E, Huang Q. A checklist of Chinese reptiles and amphibians. Beijing: Scientific; 1986.
7. Booth W, Schuett GW. The emerging phylogenetic pattern of parthenogenesis in snakes. *Biol J Linn Soc*. 2016;118:172–86. <https://doi.org/10.1111/bij.12744>.
8. Lampert K. Facultative parthenogenesis in vertebrates: reproductive error or chance? *Sex Dev*. 2008;2:290–301. <https://doi.org/10.1159/000195678>.
9. Sinclair EA, Pramuk JB, Bezy RL, Crandall KA, Sites Jr JW. DNA evidence for nonhybrid origins of parthenogenesis in natural populations of vertebrates. *Evol: J Org Evol*. 2010;64:1346–57. <https://doi.org/10.1111/j.1558-5646.2009.00893.x>
10. Stelzer CP. Does the avoidance of sexual costs increase fitness in asexual invaders? *P Natl Acad Sci Usa*. 2015;112:8851–8. <https://doi.org/10.1073/pnas.1501726112>.
11. Kearney M, Fujita MK, Ridenour J. Lost sex in the reptiles: constraints and correlations. In: Lost sex. Edited by Schön I, Martens K, Dijk P eds. Springer, Netherlands. 2009;447–74.
12. Fujita MK, Singhal S, Brunes TO, Maldonado JA. Evolutionary dynamics and consequences of parthenogenesis in vertebrates. *Annu Rev Ecol Evol S*. 2020;51:191–214. <https://doi.org/10.1146/annurev-ecolsys-011720-114900>.
13. Madlung A. Polyploidy and its effect on evolutionary success: old questions revisited with new tools. *Heredity*. 2012;110:99–104. <https://doi.org/10.1038/hdy.2012.79>.
14. McDowell SB. A catalogue of the snakes of New Guinea and the Solomons, with special reference to those in the Bernice P. Bishop Museum, Part I. *Scole-cophidia*. *J Herpetol*. 1974;8:1–57. <https://doi.org/10.2307/1563076>.
15. Dong L, Liu H, Zhang J, Yang S, Kong G, Chu JSC, et al. Single-molecule real-time transcript sequencing facilitates common wheat genome annotation and grain transcriptome research. *BMC Genomics*. 2015;16:1–13. <https://doi.org/10.1186/s12864-015-2257-y>.
16. Piriyapongsa J, Kaewprommal P, Vaiwiri S, Anuntakaran S, Wirojsirasak W, Punpee P, et al. Uncovering full-length transcript isoforms of sugarcane cultivar Khon Kaen 3 using single-molecule long-read sequencing. *PeerJ*. 2018;6:e5818. <https://doi.org/10.7717/peerj.5818>.
17. Osborn TC, Pires JC, Birchler JA, Auger DL, Chen ZJ, Lee H-S, et al. Understanding mechanisms of novel gene expression in polyploids. *Trends Genet*. 2003;19:141–7. [https://doi.org/10.1016/S0168-9525\(03\)00015-5](https://doi.org/10.1016/S0168-9525(03)00015-5).
18. Wang M, Wang P, Liang F, Ye Z, Li J, Shen C, et al. A global survey of alternative splicing in allopolyploid cotton: landscape, complexity and regulation. *New Phytol*. 2017;217:163–78. <https://doi.org/10.1111/nph.14762>.
19. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*. 1999;151:1531–45. <https://doi.org/10.1093/genetics/151.4.1531>.
20. Sidow A. Gen(om)e duplications in the evolution of early vertebrates. *Curr Opin Genet Dev*. 1996;6:715–22. [https://doi.org/10.1016/s0959-437x\(96\)80026-8](https://doi.org/10.1016/s0959-437x(96)80026-8).
21. Wang B, Kumar V, Olson A, Ware D. Reviving the Transcriptome studies: an insight into the emergence of single-molecule transcriptome sequencing. *Front Genet*. 2019;10:384. <https://doi.org/10.3389/fgene.2019.00384>.
22. Kuo RI, Tseng E, Eory L, Paton IR, Archibald AL, Burt DW. Normalized long read RNA sequencing in chicken reveals transcriptome complexity similar to human. *BMC Genomics*. 2017;18:1–19. <https://doi.org/10.1186/s12864-017-3691-9>.
23. Workman RE, Myrka AM, Wong GW, Tseng E, Welch Jr KC, Timp W. Single-molecule, full-length transcript sequencing provides insight into the extreme metabolism of the ruby-throated hummingbird *Archilochus colubris*. *Gigascience*. 2018;7:1–12. <https://doi.org/10.1093/gigascience/giy009>.
24. Cheng B, Furtado A, Henry RJ. Long-read sequencing of the coffee bean transcriptome reveals the diversity of full-length transcripts. *Gigascience*. 2017;6:1–13. <https://doi.org/10.1093/gigascience/gix086>.
25. Sahlin K, Tomaszewicz M, Makova KD, Medvedev P. Deciphering highly similar multigene family transcripts from Iso-Seq data with IsoCon. *Nat Commun*. 2018;9:1–12. <https://doi.org/10.1038/s41467-018-06910-x>.
26. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, et al. Hybrid error correction and de novo assembly of single-molecule sequencing reads. *Nat Biotechnol*. 2012;30:693–700. <https://doi.org/10.1038/nbt.2280>.
27. Liu X, Mei W, Soltis PS, Soltis DE, Barbazuk WB. Detecting alternatively spliced transcript isoforms from single-molecule long-read sequences without a reference genome. *Mol Ecol Resour*. 2017;17:1243–56. <https://doi.org/10.1111/1755-0998.12670>.
28. Rhoads A, Au KF. PacBio sequencing and its applications. *Genom Proteom Bioinf*. 2015;13:278–89. <https://doi.org/10.1016/j.gpb.2015.08.002>.
29. Wang L, Zhu P, Mo Q, Luo W, Du Z, Jiang J, et al. Comprehensive analysis of full-length transcriptomes of *Schizothorax prenanti* by single-molecule long-read sequencing. *Genomics*. 2022;114:456–64. <https://doi.org/10.1016/j.ygeno.2021.01.009>.
30. Hoang NV, Furtado A, Mason PJ, Marquardt A, Kasirajan L, Thirugnanasambandam PP, et al. A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing. *BMC Genomics*. 2017;18:1–22. <https://doi.org/10.1186/s12864-017-3757-8>.
31. Zhang X, Li G, Jiang H, Li L, Ma J, Li H, et al. Full-length transcriptome analysis of *Litopenaeus vannamei* reveals transcript variants involved in the innate immune system. *Fish Shellfish Immun*. 2019;87:346–59. <https://doi.org/10.1016/j.fsi.2019.01.023>.
32. Yi S, Zhou X, Li J, Zhang M, Luo S. Full-length transcriptome of *Misgurnus anguillicaudatus* provides insights into evolution of genus *Misgurnus*. *Sci Rep-Uk*. 2018;8:1–9. <https://doi.org/10.1038/s41598-018-29991-6>.
33. Yin X, Guo S, Gao J, Luo L, Liao X, Li M, et al. Kinetic analysis of effects of temperature and time on the regulation of venom expression in *Bungarus multicinctus*. *Sci Rep-Uk*. 2020;10:14142.
34. Kyriakidou M, Tai HH, Anglin NL, Ellis D, Strömvik MV. Current strategies of Polyploid Plant Genome Sequence Assembly. *Front Plant Sci*. 2018;9:1660. <https://doi.org/10.3389/fpls.2018.01660>.

35. Yuan H, Yu H, Huang T, Shen X, Xia J, Pang F, et al. The complexity of the *Fragaria x ananassa* (octoploid) transcriptome by single-molecule long-read sequencing. *Hortic Res-England*. 2019;6:46. <https://doi.org/10.1038/s41438-019-0126-6>.
36. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol*. 2010;8:e1000412.
37. Camacho C, Coulouris G, Avagyan V, Ning M, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:1–9. <https://doi.org/10.1186/1471-2105-10-421>.
38. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO Update: Novel and Streamlined Workflows along with broader and deeper phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38:4647–54. <https://doi.org/10.1093/molbev/msab199>.
39. Birney E, Clamp M, Durbin R. GeneWise and Genomewise. *Genome Res*. 2004;14:988–95. <https://doi.org/10.1101/gr.1865504>
40. Keilwagen J, Hartung F, Grau J. GeMoMa: Homology-Based Gene Prediction Utilizing Intron Position Conservation and RNA-seq Data. *Methods Mol Biol*. 2019;1962:161–77. https://doi.org/10.1007/978-1-4939-9173-0_9
41. Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A, Gasteiger E, et al. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res*. 2003;31:365–70. <https://doi.org/10.1093/nar/gkg095>.
42. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics*. 2020;36:2251–2.
43. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res*. 2012;40:D109–114. <https://doi.org/10.1093/nar/gkr988>.
44. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25:25–9. <https://doi.org/10.1038/75556>.
45. Bru C, Courcelle E, Carrère S, Beausse Y, Dalmar S, Kahn D. The ProDom database of protein domain families: more emphasis on 3D. *Nucleic Acids Res*. 2005;33:D212–215. <https://doi.org/10.1093/nar/gki034>.
46. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40. <https://doi.org/10.1093/bioinformatics/btu031>.
47. He Y, Peng F, Deng C, Xiong L, Huang Z-y, Zhang R-q, et al. Building an octoploid genome and transcriptome of the medicinal plant *Pogostemon cablin* from Lamiales. *Sci Data*. 2018;5:1–11. <https://doi.org/10.1038/sdata.2018.274>.
48. Emms D, Kelly S. OrthoFinder2: fast and accurate phylogenomic orthology analysis from gene sequences. *BioRxiv*. 2018;466201. <https://doi.org/10.1101/466201>.
49. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792–7. <https://doi.org/10.1093/nar/gkh340>.
50. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res*. 2006;34:W609–612. <https://doi.org/10.1093/nar/gkl315>.
51. Talavera G, Castresana J. Improvement of Phylogenies after removing Divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 2007;56:564–77. <https://doi.org/10.1080/10635150701472164>.
52. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 2014;30:1312–3. <https://doi.org/10.1093/bioinformatics/btu033>
53. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24:1586–91. <https://doi.org/10.1093/molbev/msm088>.
54. Li M, Zheng Z, Liu J, Yang Y, Ren G, Ru D, et al. Evolutionary origin of a tetraploid *Allium* species on the Qinghai-Tibet Plateau. *Mol Ecol*. 2021;30:5780–95.
55. Yang Y, Li Y, Chen Q, Sun Y, Lu Z. WGDdetector: a pipeline for detecting whole genome duplication events using the genome or transcriptome annotations. *BMC Bioinformatics*. 2019;20:1–6. <https://doi.org/10.1186/s12859-019-2670-3>.
56. Ota H. Karyotype of Pareas iwasakii The First Chromosomal description of a pareatine snake (Colubridae). *Jpn J Herpetol*. 1999;18:16–8. https://doi.org/10.5358/hjsj1972.18.1_16.
57. Levan A, Fredga K, Sandberg AA. Nomenclature for centromeric position on chromosomes. *Hereditas*. 1964;52:201–20.
58. Sakamoto Y, Zacaro A. LEVAN, an ImageJ plugin for morphological cytogenetic analysis of mitotic and meiotic chromosomes. Initial version An Open Source Java Plugin Distributed Over the Internet. 2009. from: <http://rsbweb.nih.gov/ij>.
59. Furuno M, Kasukawa T, Saito R, Adachi J, Suzuki H, Baldarelli R, et al. CDS annotation in full-length cDNA sequence. *Genome Res*. 2003;13:1478–87. <https://doi.org/10.1101/gr.1060303>.
60. Yang H, Hu J, Wang Z, Xu D, Zhuo Z. Using PacBio Iso-Seq to determine the transcriptome of *Rhynchophorus Ferrugineus*. *Chin J Appl Entomol*. 2021;58:655–63. (In Chinese).
61. Patawang I, Tanomtong A, Kaewmad P, Chuaynkern Y, Duengkae P. New record on karyological analysis and first study of NOR localization of parthenogenetic brahminy blind snake, *Ramphotyphlops braminus* (Squamata, Typhlopidae) in Thailand. *Nucleus*. 2016;59:61–6. <https://doi.org/10.1007/s13237-015-0154-z>.
62. Yan J, Li H, Zhou K. Evolution of the mitochondrial genome in snakes: gene rearrangements and phylogenetic relationships. *BMC Genomics*. 2008;9:1–7. <https://doi.org/10.1186/1471-2164-9-569>.
63. Liu Y. Complete mitochondrial genome and phylogenetic analysis of the Burmese python (*Python bivittatus*) and genetic differentiation between two local populations. University of Hainan; 2014.
64. Blanc G, Wolfe KH. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. *Plant Cell*. 2004;16:1667–78. <https://doi.org/10.1105/tpc.021345>.
65. Zhong W, Zhang X, Zhao Q, Ma D, Tang H. The application of three generation PacBio sequencing in the study of transcriptome. *J Fujian Agric Univ (Nat Sci Ed)*. 2018;47:524–9. (In Chinese).
66. Otto SP, Whitten J. Polyploid incidence and evolution. *Annu Rev Genet*. 2000;34:401–37.
67. Novikova PY, Brennan IG, Booker W, Mahony M, Doughty P, Lemmon AR, et al. Polyploidy breaks speciation barriers in Australian burrowing frogs *Neobatrachus*. *Plos Genet*. 2020;16:e1008769. <https://doi.org/10.1371/journal.pgen.1008769>.
68. Wang JT, Li JT, Zhang XF, Sun XW. Transcriptome analysis reveals the time of the fourth round of genome duplication in common carp (*Cyprinus carpio*). *BMC Genomics*. 2012;13:1–10. <https://doi.org/10.1186/1471-2164-13-96>.
69. Holland PW, Garcia-Fernández J, Williams NA, Sidow A. Gene duplications and the origins of vertebrate development. *Development*. 1994;125–33. <https://doi.org/10.1242/dev.1994.Supplement.125>.
70. Meyer A, Schartl M. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol*. 1999;11:699–704. [https://doi.org/10.1016/s0955-0674\(99\)00039-3](https://doi.org/10.1016/s0955-0674(99)00039-3).
71. Chen CL, Zhang L, Li JL, Mao XX, Zhang LS, Hu QJ, et al. Phylogenomic analysis reveals extensive gene duplication in the subtribe Gentianinae (Gentianaceae). *J Syst Evol*. 2021;59:1198–208.
72. Muller HJ. Why polyploidy is rarer in animals than in plants. *Am Nat*. 1925;59:346–53. <http://www.jstor.org/stable/2456335>.
73. Mayrose I, Zhan SH, Rothfels CJ, Arrigo N, Barker MS, Rieseberg LH, et al. Methods for studying polyploid diversification and the dead end hypothesis: a reply to Soltis (2014). *New Phytol*. 2015;206:27–35. <https://doi.org/10.1111/nph.13192>.
74. Van de Peer Y, Mizrachi E, Marchal K. The evolutionary significance of polyploidy. *Nat Rev Genet*. 2017;18:411–24. <https://doi.org/10.1038/nrg.2017.26>.
75. Song C, Liu S, Xiao J, He W, Zhou Y, Qin Q, et al. Polyploid organisms. *Sci China Life Sci*. 2012;55:301–11. <https://doi.org/10.1007/s11427-012-4310-2>.
76. Gregory TR, Mable BK. Polyploidy in animals. In: *The evolution of the genome*. Edited by Gregory TR ed. London: Elsevier; 2005:427–517.
77. Badger MPS, Lear CH, Pancost RD, Foster GL, Bailey TR, Leng MJ, et al. CO₂ drawdown following the middle Miocene expansion of the Antarctic ice sheet. *Paleoceanogr Paleocl*. 2013;28:42–53. <https://doi.org/10.1002/palo.20015>.
78. Super JR, Thomas E, Pagani M, Huber M, O'Brien C, Hull PM. North Atlantic temperature and pCO₂ coupling in the early-middle Miocene. *Geology*. 2018;46:519–22. <https://doi.org/10.1130/G40228.1>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.