RESEARCH ARTICLE

Highly contiguous reference genome assembly of the endangered Orcés' blue whiptail *Holcosus orcesi*

Gabriela Pozo[1], Diego F. Cisneros-Heredia[2,3], Doménica Barragán-Orbe[1], Juan C. Sánchez-Nivicela[2,3], Ernesto Arbeláez[4], Maria de Lourdes Torres[1]*

1 Laboratorio de Biotecnología Vegetal, Universidad San Francisco de Quito USFQ, Quito, Ecuador, **2** Laboratorio de Zoología Terrestre, Instituto de Biodiversidad Tropical IBIOTROP, Colegio de Ciencias Biológicas y Ambientales, Universidad San Francisco de Quito USFQ, Quito, Ecuador, **3** Instituto Nacional de Biodiversidad (INABIO), Quito, Ecuador, **4** Amaru Bioparque Cuenca, Cuenca, Ecuador

* ltorres@usfq.edu.ec

OPEN ACCESS

Citation: Pozo G, Cisneros-Heredia DF, Barragán-Orbe D, Sánchez-Nivicela JC, Arbeláez E, Torres MdL (2026) Highly contiguous reference genome assembly of the endangered Orcés' blue whiptail *Holcosus orcesi*. PLoS One 21(8): e0350380. https://doi.org/10.1371/journal.pone.0350380

Editor: Luisa Maria Diele-Viegas, University of Mississippi, BRAZIL

Received: May 27, 2026

Accepted: August 6, 2026

Published: August 27, 2026

Data availability statement: The raw sequencing reads generated for this study have been deposited in the NCBI Sequence Read Archive (SRR36082360). The genome assembly has also been uploaded to NCBI under the BioProject accession number

Abstract

Holcosus orcesi, the Orcés' Blue Whiptail, is a Critically Endangered lizard endemic to the upper Jubones River basin in southern Ecuador. Restricted to a narrow elevational range within semi-arid Andean shrublands, it represents one of the few montane members of a predominantly lowland lineage. Here we present the first high-quality reference genome for *H. orcesi*, generated using Oxford Nanopore Technologies long-read sequencing. The assembly spans 1.68 Gb across only 91 contigs, with an N50 of 76.2 Mb and a BUSCO completeness of 96.8%, making it among the most contiguous and complete squamate genomes to date. Structural annotation predicted 25,682 genes, of which 85% showed homology to known proteins and 45% were assigned Gene Ontology terms. Repetitive elements accounted for 46.3% of the genome, with LINEs representing the predominant class. This genome provides a foundational resource for future evolutionary, comparative and conservation-genomic research of *H. orcesi* and other mountain reptiles, enabling studies of population genomics, local adaptation, and genomic erosion in isolated populations. By expanding the genomic representation of tropical montane reptiles, this work helps address longstanding phylogenetic and geographic gaps in global biodiversity genomics and provides a foundation for evidence-based conservation of *H. orcesi* and related taxa.

Introduction

Holcosus orcesi (PETERS, 1964) [1], commonly known as Orcés' Blue Whiptail, is a small teiid lizard endemic to the upper Jubones River basin on the southwestern Andes of Ecuador, where it occupies a narrow elevational range between 1,000 and 1,600 m (Fig 1A and 1B) [1,2]. The species went unrecorded for nearly six decades and was presumed extinct until its rediscovery in the mid-2010s [3,4]. It inhabits xeric montane shrublands and persists in small, isolated patches of native vegetation

PRJNA1153955 and BioSample accession number SAMN53165835.

Funding: This research was supported by the Universidad San Francisco de Quito College of Biological and Environmental Sciences (COCIBA) Grants awarded to GP, MdLT, and DFC-H; the Oxford Nanopore Technologies ORG.one initiative grant awarded to GP; and operational funds from the Institute of Tropical Biodiversity assigned to the Laboratory of Terrestrial Zoology and TUERI Wildlife Hospital. No grant numbers are available for these funding sources. The sponsors and funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

surrounded by pastures and croplands [1,3]. The region is semi-arid, with low annual rainfall (≈400 mm/year), concentrated between January and May, followed by a prolonged dry season, and experiences mild days and cool nights [5–7]. *Holcosus orcesi* is strictly diurnal and heliothermic, active only under full sunlight and retreating when clouds obscure the sun. It forages swiftly among roots, thorny shrubs, and sun-exposed rocky slopes, feeding mainly on beetles and grasshoppers, and relies on vigilance and sprint speed as primary defenses [1].

The genus *Holcosus* extends from southern Mexico to north-western South America, and *H. orcesi* is one of the few montane species within this predominantly lowland radiation [2]. Its apparent narrow thermal niche and dependence on open, sunlit microhabitats likely increase its sensitivity to climatic and habitat changes. Habitat conversion for agriculture, invasive predators and poaching have contributed to population decline, leading the IUCN to list it as Critically Endangered [4]. As a geographically restricted species occupying a biogeographically unusual position within its genus, *H. orcesi* exemplifies evolutionary persistence under severe environmental stress [1,2]. Studying its genome could offer a valuable opportunity to understand how this lineage has adapted to an environmentally demanding, semi-arid montane habitat. However, genomic resources for Andean species remain scarce, limiting comparative work on the evolutionary and ecological diversification of this fauna.

High-quality genomic resources have become indispensable tools for modern biodiversity conservation [8,9]. Reference genomes provide the foundation for investigating species' evolutionary histories, demographic trajectories, and adaptive potential, particularly in lineages with restricted distributions and elevated extinction risk [8,10]. In threatened taxa such as *H. orcesi*, where obtaining large numbers of samples can be challenging, whole-genome resources are particularly valuable because they provide substantially more genomic information per individual than traditional genetic markers, enabling more comprehensive analyses of population structure, inbreeding, and evolutionary history. The availability of a reference genome opens the possibility of identifying genetic signatures of adaptation to its harsh high-altitude environment, including tolerance to hypoxia, UV exposure, and cold temperatures. Genomic approaches have revealed high-altitude adaptations in other vertebrates, including hypoxia-related pathways in yak (*Bos grunniens*) [11], the Tibetan ground tit (*Parus humilis*) [12], and multiple mammals such as the Tibetan sheep and plateau zokor [13,14].

The integration of genomics into conservation has already transformed our ability to monitor population health. Genomic data enable more accurate estimates of effective population size, levels of genetic diversity, and the genomic basis of inbreeding depression, which are crucial to designing effective management plans [10,15]. In particular, in endangered species with small and isolated populations, such as *H. orcesi*, these resources allow detection of deleterious alleles and runs of homozygosity that may otherwise go unnoticed through traditional genetic markers [15–17].

Beyond species-specific applications, genomic resources contribute to comparative and macroevolutionary studies. Initiatives such as the Earth BioGenome Project and related efforts have emphasized the importance of expanding genomic representation across taxonomic and geographic scales, noting significant biases in available

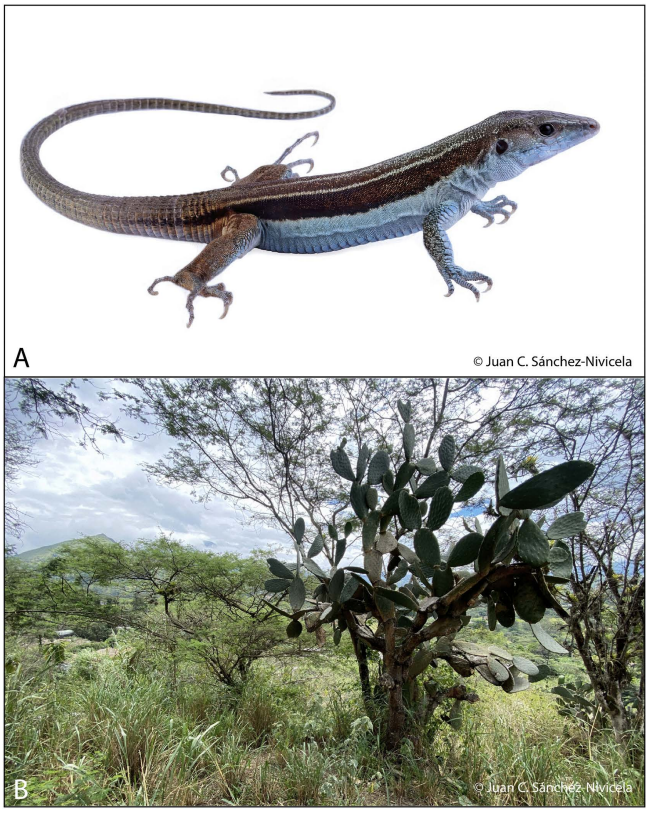

Fig 1. *Holcosus orcesi* and its natural habitat. Representative individual of the Ecuadorian endemic lizard *Holcosus orcesi* and its natural dry forest habitat in southern Ecuador. A) Adult *H. orcesi* individual. B) Typical habitat characterized by xeric vegetation and cacti.

https://doi.org/10.1371/journal.pone.0350380.g001

references toward temperate species [8,18]. By contributing the genome of a narrowly distributed Andean reptile, our study helps to fill this latitudinal and phylogenetic gap, providing a resource for Neotropical squamates that benefits both regional and global biodiversity knowledge.

Ultimately, genomic resources represent one of the most powerful tools available to address the global loss of biodiversity. They inform not only the immediate conservation of threatened taxa, but also broader strategies aimed at preserving adaptive capacity in the face of rapid environmental change [9]. In this context, the genome of *H. orcesi* will serve as a cornerstone for future studies in evolutionary biology, population genomics, and conservation management of Andean reptiles. Here, we present the first reference genome assembly for *Holcosus orcesi*, generated using Oxford Nanopore long-read sequencing, together with structural annotation and repeat characterization. This resource will serve as a cornerstone for future studies in evolutionary biology, population genomics, and conservation management of Andean reptiles, and supports research on *H. orcesi*, *Holcosus*, and other poorly represented tropical montane squamates.

Materials and methods

Sample collection and DNA extraction

A single specimen of *Holcosus orcesi* was collected in February 2025 from the upper Jubones River basin, southwestern Ecuador. The exact locality is withheld to reduce the risk of poaching but is available upon reasonable request from the Museum of Zoology at Universidad San Francisco de Quito USFQ. The individual was captured by hand and handled only for the minimum time necessary. All clinical procedures involving anesthesia, blood collection, and euthanasia were

performed by a certified wildlife veterinarian specialized in wildlife medicine at the TUERI Wildlife Hospital–USFQ. To minimize handling-associated stress and prevent pain or distress during the invasive procedure, the animal was premedicated intramuscularly with ketamine (10 mg kg^{-1}) and medetomidine (0.1 mg kg^{-1}). Once an adequate anesthetic plane had been achieved, sevoflurane was administered at 8% in oxygen through a face mask. Intracardiac blood collection was performed only after a deep plane of anesthesia had been established.

The procedure was terminal, and the animal remained under continuous deep general anesthesia throughout blood collection and euthanasia without recovering consciousness. Euthanasia was performed by maintaining administration of 8% sevoflurane in oxygen through the face mask until anesthetic overdose and cessation of cardiorespiratory function occurred. Death was confirmed by the absence of spontaneous respiration and reflex responses and by the absence of cardiac activity assessed using Doppler ultrasonography. No additional analgesic was administered because all invasive procedures were performed under continuous deep general anesthesia and no recovery or postoperative period occurred. The use of a single individual, minimal handling and restraint, pharmacological premedication, confirmation of an adequate anesthetic plane before intracardiac blood collection, maintenance of deep anesthesia throughout the terminal procedure, and performance of all procedures by an experienced wildlife veterinarian were the principal measures taken to minimize pain, distress, and suffering.

The Ethics Committee for the Use of Animals in Research and Teaching at USFQ formally waived the requirement for study-specific ethics approval, as these procedures were conducted under TUERI's institutional clinical protocols, which had been previously reviewed and approved by the Committee without the assignment of protocol identification numbers. The animal was premedicated intramuscularly with ketamine and medetomidine and subsequently anaesthetized with sevoflurane. Once a deep plane of anesthesia had been achieved, intracardiac blood collection was performed. The animal was then euthanized under deep anesthesia by sevoflurane overdose. The voucher specimen is deposited at the Zoology Museum - USFQ. Whole blood was preserved in K2-EDTA and stored at –80 °C until DNA extraction. All procedures followed the guidelines for the use of live reptiles in field research [19]. Research activities were authorized by the Ecuadorian Ministry of Environment, Water and Ecological Transition under Research Permit No. MAATE-ARSFC-2025-0144 and Framework Contract for Access to Genetic Resources No. MAATE-DBI-CM-2023-0313.

Ultra-high molecular weight genomic DNA was isolated from blood obtained from this same individual using two protocols: the Monarch® HMW DNA Extraction Kit (T3050) for nucleated blood, and the Monarch® Genomic DNA Purification Kit (T3010) for nucleated red blood cells. For HMW extraction, 15 µl of blood was used as input material, while 10 µl was used for genomic DNA extraction. In both cases DNA was eluted in 100 µl. DNA integrity was assessed by electrophoresis on a 1.5% agarose gel, followed by quantification using a Qubit 4 Fluorometer.

Library preparation and sequencing

Two sequencing libraries were prepared using the Ligation Sequencing Kit SQK-LSK114 (Oxford Nanopore Technologies), starting with 1000 ng of input DNA. The library preparation protocol was followed with minor modifications in the DNA repair and end-prep step (extended incubation time of 15 minutes at 20°C and 15 minutes at 65°C and a single ethanol wash), and in the adapter ligation and clean-up step (extended incubation time of 20 minutes at room temperature and a single Long Fragment Buffer wash).

Sequencing was performed on a PromethION 2 Solo device using two R10.4.1 flow cells, each run for ~21 hours. Data acquisition and run control were managed with MinKNOW (v6.5.14), and raw signals were base called with Dorado (v7.9.8) in high-accuracy mode, using a minimum Q score threshold of 7.

Read processing, genome assembly and quality assessment

All passed Oxford Nanopore reads with quality scores ≥7 were concatenated, and sequencing adapters were trimmed using Porechop v0.2.3_seqan2.1.1 [20]. Read quality and length distributions were first assessed with NanoPlot v1.32.1

[21]. Reads shorter than 5,000 bp were discarded with NanoFilt v2.3.0 prior to assembly [21]. The filtered dataset was assembled de novo using Hifiasm v0.25.0-r726 [22] in ONT-optimized mode, and primary contigs were extracted from the assembly graph with gfatools v0.4-r214-dirty [23]. To improve consensus accuracy, raw ONT reads were mapped back to the draft assembly with minimap2 v2.30-r1287 [24], and the resulting alignments were used to polish the genome with Racon v1.4.20 [25]. Two successive rounds of Racon polishing were performed to maximize assembly accuracy.

Assembly completeness was quantified with BUSCO v5.8.3 [26] (mode: genome) using the squamata_odb12 lineage dataset, which surveys 11,294 near-universal single-copy orthologs expected to be present in any squamate genome. To assess contiguity, we ran QUAST v5.0.2 [27] with default settings to report standard metrics, including total assembly length, number of contigs (≥500 bp), largest contig, N50/L50, N90/L90, and GC content.

Genome annotation and identification of repetitive elements

Structural annotation was performed in OmicsBox v3.4.5 (BioBam Bioinformatics, Valencia, Spain) using AUGUSTUS, with *Gallus gallus* specified as the closest available model species. To increase lineage-specific accuracy, all available Identical Protein Groups (IPGs) from members of the family Teiidae were downloaded from NCBI and used as external support for gene model construction. Prior to annotation, the genome was soft masked with the Dfam profile HMM library via HMMER.

Functional annotation of the predicted gene set was carried out in OmicsBox through two complementary approaches. First, InterProScan v5.72-103.0 was used to identify conserved protein domains, families, and motifs, and to assign Gene Ontology (GO) terms [28]. Second, similarity-based annotation was performed with DIAMOND BLAST against the NCBI non-redundant (nr) protein database, enabling functional assignments based on homology to previously characterized proteins [29]. The combined evidence from domain analysis and sequence similarity provided robust functional characterization of the annotated gene models.

Repetitive elements were identified and categorized using RepeatModeler2 (v2.0.7) [30]. First, a custom repeat library was constructed from the polished genome assembly. This library was then used by RepeatMasker (v4.1.9) to annotate and soft-mask repetitive elements in the genome [31].

Results

Sequencing output and assembly statistics

A total of 10.11 million reads were generated using Oxford Nanopore sequencing, yielding 79.03 Gb of long-read data. All passed reads had a quality score greater than 7, with a mean read quality of 14.9. Regarding read size, the average read length was 7.8 kb, with an N50 read length of 17.8 kb. The coverage (~49x) was calculated based on the expected genome size from a closely related species, *Holcosus undulatus*, which is 1.6 Gb [32] (Table 1).

The final assembly of *H. orcesi* generated with Hifiasm spans 1.68 Gb, with an N50 of 76.2 Mb. The assembly is divided into 91 contigs, with 8 contigs representing 50% of the genome (L50), and the largest contig measuring 179.6 Mb. In terms of completeness, BUSCO analysis recovered 96.8% of complete orthologs (10,936 BUSCOs), including 96.5% single-copy and 0.4% duplicated genes. Additionally, 1.4% of the genes were classified as fragmented and 1.7% as missing (Table 2, Fig 2).

Table 1. Sequencing output summary.

Generated bases	Mean read length	Mean read quality	Read length N50	Coverage
79.03 Gb	7.8 kb	14.9	17.8 kb	49x

https://doi.org/10.1371/journal.pone.0350380.t001

Table 2. Assembly metrics for the *Holcosus orcesi* polished genome.

Parameter	Value
Total length (bp)	1,676,717,785
# contigs	91
Largest contig (bp)	179,640,633
N50 (bp)	76,215,922
L50	8
GC (%)	42.39%
Complete BUSCOs	96.8% (10,936)
Single copy	96.5% (10,897)
Duplicated	0.4% (42)
Fragmented BUSCOs	1.4% (159)
Missing BUSCOs	1.7% (196)

https://doi.org/10.1371/journal.pone.0350380.t002

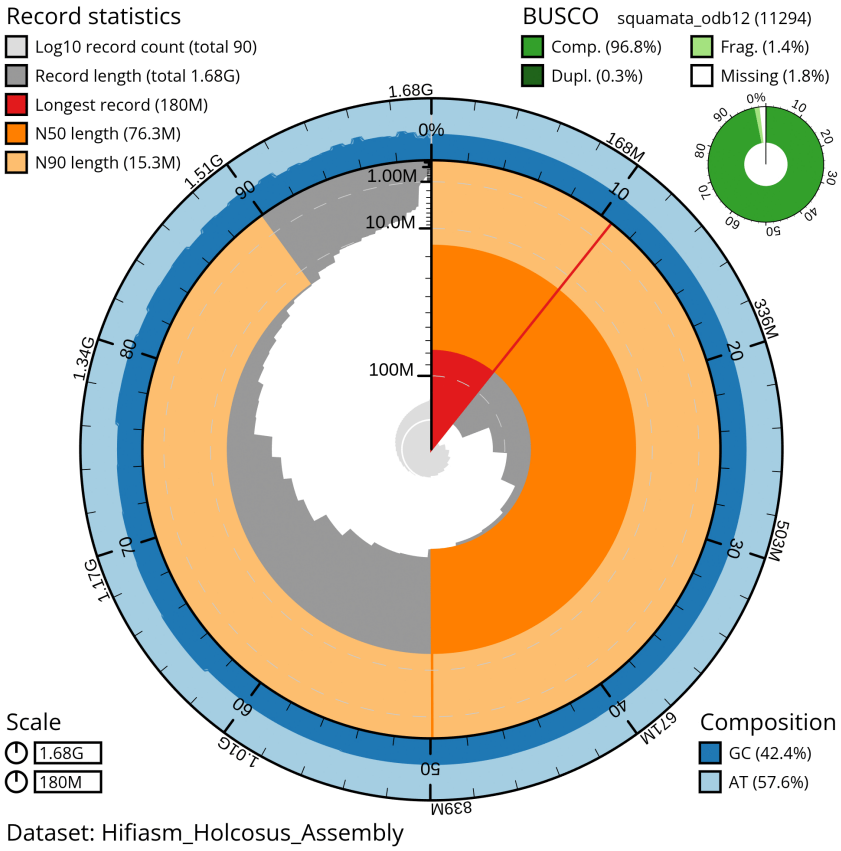

Fig 2. Reference genome assembly statistics for *Holcosus orcesi*. Snailplot summary statistics of the *Holcosus orcesi* reference genome assembly generated using Oxford Nanopore sequencing. The plot summarizes assembly size, contiguity metrics, GC/AT composition, and BUSCO completeness based on the squamata_odb12 dataset.

https://doi.org/10.1371/journal.pone.0350380.g002

Genome annotation and repetitive elements

The final gene set consisted of 25,682 predicted genes, with a total of 152,998 exons and 127,316 introns. On average, each gene encoded a single mRNA, containing 5.9 exons and 4.9 introns. The mean gene length was 18,052 bp, with exons averaging 244 bp and introns 3,347 bp in length (Table 3). In this gene set, 84.92% of proteins showed significant similarity to known proteins in the nr database, 85.09% contained at least one conserved domain, and 44.79% were assigned at least one Gene Ontology (GO) term (Table 4).

InterProScan identified conserved domains and families in 21,853 proteins, representing 85.1% of the predicted gene set. Among PFAM domains, the most frequent were reverse transcriptase (5.1%), integrase catalytic core (3.7%), GPCR family 3 C-terminal (2.4%), zinc finger C2H2-type (1.9%), and protein kinase domains (1.5%). Similarly, the most represented InterPro families were the ribonuclease H superfamily (7.7%), DNA/RNA polymerase superfamily (7.5%), ribonuclease H-like superfamily (5.3%), reverse transcriptase/diguanylate cyclase domain family (4.4%), and the L1 transposable element C-terminal domain (4.4%). These patterns reflect the prevalence of mobile element-related proteins, signaling components, and nucleic acid–binding factors in the *H. orcesi* proteome (Table 5).

Gene Ontology (GO) analysis assigned functional categories to 11,502 proteins (44.8% of the predicted gene set). Within the Biological Process domain, the most represented categories were cellular processes (7.5%), metabolic processes (3.9%), and regulation of biological processes (3.9%), highlighting the prevalence of genes involved in fundamental cellular and regulatory pathways. In the Molecular Function domain, the majority of proteins were associated with binding activities. This included protein, nucleic acid, and small molecule binding (together accounting for more than 10% of annotated proteins), followed by catalytic activity, particularly hydrolase and transferase functions. For the Cellular Component domain, proteins were primarily assigned to intracellular anatomical structures (e.g., organelles, cytoplasm) and to membrane-associated compartments, reflecting the structural and functional organization of the *H. orcesi* proteome (Fig 3).

Table 3. Genome annotation statistics.

Statistic	Value
Number of genes	25,682
Number of exons	152,998
Number of introns in CDS	127,316
Overlapping genes	0
Mean mRNAs per gene	1
Mean exons per mRNA	5.9
Mean introns per mRNA	4.9
Mean gene length (bp)	18,052
Mean exon length (bp)	244
Mean intron length (bp)	3,347

https://doi.org/10.1371/journal.pone.0350380.t003

Table 4. Functional annotation of the predicted gene set.

Annotation step	Metric	Value	Percentage of predicted gene set
InterProScan	Proteins with ≥1 conserved domain	21,853	85.09%
DIAMOND BLAST (nr)	Proteins with significant hit (e-value<1e-5)	21,811	84.92%
Gene Ontology	GO mapped	13,952	54.33%
	GO annotated	11,502	44.79%

https://doi.org/10.1371/journal.pone.0350380.t004

Table 5. Top InterProScan domains and families.

InterProScan Domain	Counts	Percent of annotated genes
Reverse transcriptase, RNA-dependent DNA polymerase (PF07727)	1,106	5.06%
Integrase, catalytic core (PF00665)	812	3.72%
GPCR family 3, C-terminal (PF00003)	516	2.36%
Zinc finger C2H2-type (PF00096)	421	1.93%
Protein kinase domain (PF00069)	330	1.51%
InterProScan Family	**Counts**	**Percent of annotated genes**
Ribonuclease H superfamily	1,685	7.71%
DNA/RNA polymerase superfamily	1,630	7.46%
Ribonuclease H-like superfamily	1,164	5.33%
Reverse transcriptase/Diguanylate cyclase domain	971	4.44%
L1 transposable element, C-terminal domain	953	4.36%

https://doi.org/10.1371/journal.pone.0350380.t005

Analysis of the *Holcosus orcesi* genome assembly revealed that repetitive elements occupy approximately 46.29% of the total genome sequence. Among these, retroelements were the most abundant class, comprising 19.05% of the genome and including SINEs (1.45%), LINEs (13.73%), and LTR elements (3.87%). LINEs were the most prevalent subclass within retroelements, both in terms of number (960,488 elements) and sequence coverage (230,443,544 bp) (Table 6).

DNA transposons were relatively scarce, representing only 0.23% of the genome. In contrast, unclassified repetitive elements occupied the largest proportion of the genome at 25.01%, indicating a significant portion of repeats that could not be assigned to known categories. Other repeat types included small RNAs (0.31%), simple repeats (1.46%), and low complexity regions (0.24%) (Table 6).

Discussion

The *Holcosus orcesi* genome assembly presented in this paper is of very high quality, both in terms of completeness and contiguity. The assembly spans 1.68 Gb in only 91 contigs, with an N50 of 76.2 Mb and a maximum contig length of 179 Mb. BUSCO analysis using the squamata_odb12 dataset (11,294 near-universal single-copy orthologs) recovered 96.8% complete genes, indicating an excellent representation of the expected gene number [33,34]. This combination of high contiguity and completeness provides a robust foundation for downstream analyses, including the identification of structural variants, population genomic studies, and functional analyses of adaptive traits. In particular, the recovery of nearly all conserved orthologs ensures that comparative genomics and annotation pipelines can be performed without the biases introduced by missing gene families. Such metrics indicate that the *H. orcesi* genome meets current standards for reference-quality assemblies in non-model reptiles [26,35].

Within the genus, the only other available reference is for *H. undulatus* (GCA_046270765.1), which spans ~1.6 Gb but is highly fragmented across more than 800,000 scaffolds. This contrast underscores the advances achieved in our study, where contig number was reduced by more than four orders of magnitude while achieving greater contiguity and completeness, placing it within the upper range of current lizard assemblies [34,36]. Compared to other squamate assemblies, our results are on par with recent chromosome-level references for well-studied lizards such as *Anolis carolinensis* and *Sceloporus undulatus*, where N50 values often exceed 10–50 Mb [33,37]. Thus, the *H. orcesi* genome not only represents a substantial improvement within its genus but also provides a genomic resource of comparable quality to leading reptile references, filling a critical gap in representation of high-Andean reptiles.

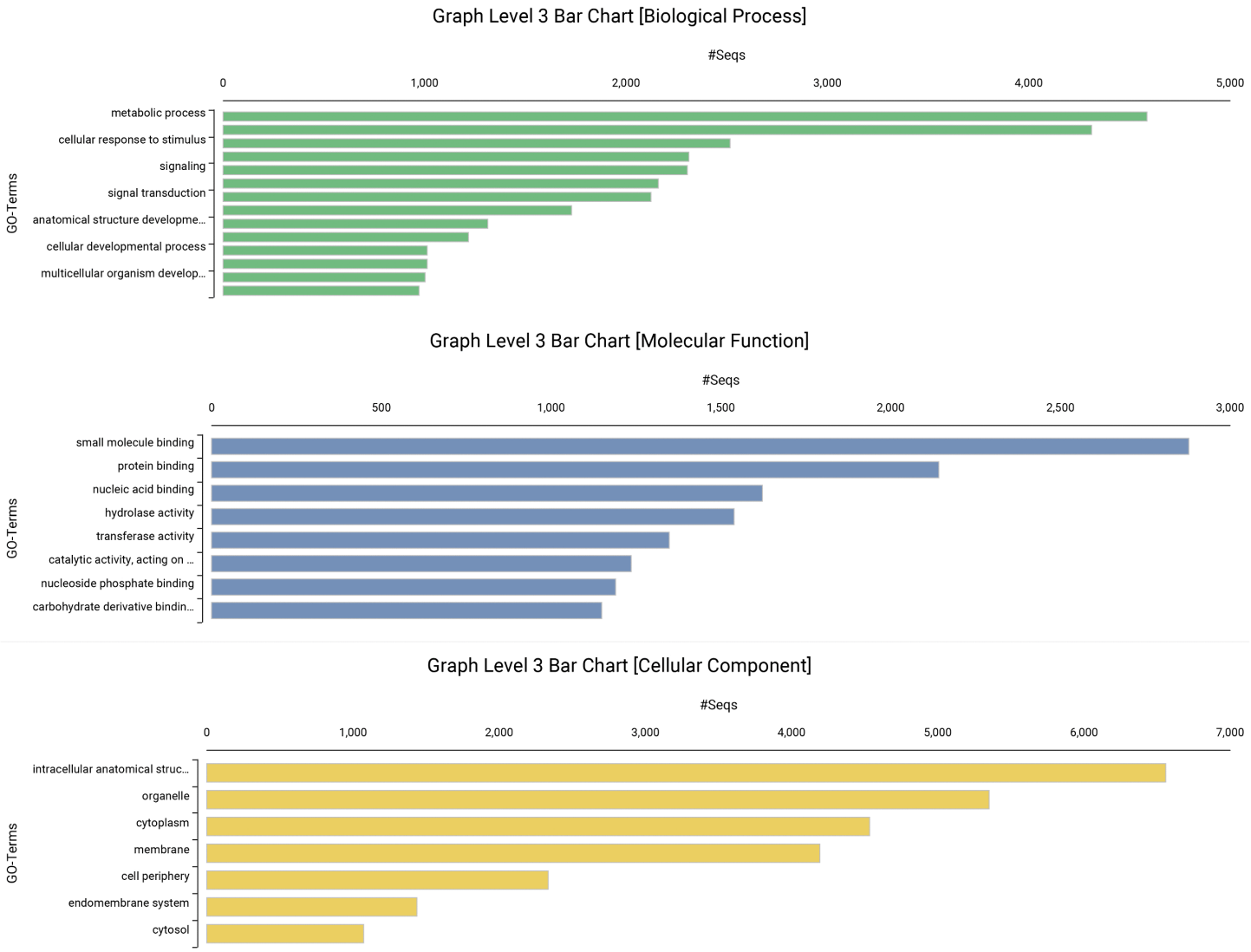

Fig 3. Gene Ontology Level 3 functional annotation of the *Holcosus orcesi* genome. Distribution of Gene Ontology (GO) Level 3 functional annotations for the annotated genes of *Holcosus orcesi*. Functional categories are grouped into the three principal GO domains: Biological Process, Molecular Function, and Cellular Component.

https://doi.org/10.1371/journal.pone.0350380.g003

The structural annotation values (Table 3) are consistent with those reported for other squamate genomes. For example, the green anole (*Anolis carolinensis*) genome contains ~19,000 protein-coding genes [37], while the Argentine black and white tegu (*Salvator merianae*) and other lizard genomes typically report between 20,000 and 26,000 predicted genes [38]. The average exon and intron lengths observed here also fall within the range described for reptiles, where exons are generally short (<300 bp) and introns considerably longer (>2 kb) [39]. Thus, the structural annotation of *Holcosus orcesi* is broadly comparable to other well-assembled squamate genomes, supporting the reliability of the predicted gene models.

The functional annotation revealed that the most enriched domain families in the *H. orcesi* proteome were strongly associated with retroelements, including reverse transcriptase, integrase, and ribonuclease H families. This reflects the

Table 6. Repetitive elements report for the *H. orcesi* genome.

Repetitive elements	Number of elements	Length occupied (bp)	% of sequence
Retroelements	1,411,885	319,681,646	19.05%
SINEs	181,916	24,258,473	1.45%
LINEs	960,488	230,443,544	13.73%
LTR elements	269,481	64,979,629	3.87%
DNA transposons	11,995	3,780,565	0.23%
Unclassified	2,470,382	419,569,037	25.01%
Small RNA	49,987	5,194,979	0.31%
Simple repeats	345,672	24,472,508	1.46%
Low complexity	36,658	4,007,354	0.24%

https://doi.org/10.1371/journal.pone.0350380.t006

pervasive influence of mobile elements on the squamate gene space, as observed in other reptilian annotations [40,41]. Nevertheless, functional categories were broadly distributed across core cellular and metabolic pathways. GO analysis revealed a predominance of genes involved in cellular and metabolic processes, binding functions (protein, nucleic acid, and small molecules), and catalytic activity, mirroring the general functional composition observed in other reptiles [42]. The assignment of ~45% of predicted proteins to GO terms aligns with annotation levels reported in other non-model squamates, where limited experimental evidence and incomplete functional databases constrain full annotation [37,41].

The repetitive landscape of the *H. orcesi* genome is consistent with the high repeat content typically found in squamates, with ~47% of the genome annotated as repetitive, including substantial contributions from LINEs (~14%) and LTR retrotransposons (~4%) [40,43,44]. This proportion falls within the broad range reported across lizards and snakes, where repeat content varies strikingly from ~25% to over 70% despite relatively constrained genome sizes [40,43,44]. Such variation contrasts with the more conserved repeat landscapes often reported for mammals and birds, where genome repeat content and composition tend to be relatively stable across closely related species. The broader range observed in squamates suggests that transposable element dynamics in reptiles may follow evolutionary patterns that differ from those inferred for other vertebrate lineages [40,45]. Notably, our assembly also recovered a large proportion of unclassified repeats (25%), a feature common to other reptile genomes and likely reflecting the rapid evolution of lineage-specific transposable element families and the limited representation of reptiles in repeat libraries [40,46].

Taken together, the functional annotation and repetitive elements landscape underscore both the reliability of the predicted gene set and the strong genomic imprint of repetitive element activity on squamate protein repertoires [40,46]. The *H. orcesi* annotation provides a robust resource for future studies of adaptation and conservation and underscores the need to expand functional databases with reptile-specific data to improve annotation accuracy in this clade.

The reference genome of *H. orcesi* represents a significant step forward for conservation genomics and evolutionary research on Neotropical reptiles [47]. From a conservation perspective, this genomic resource opens the door to evaluating genetic diversity and population structure in one of Ecuador's most threatened reptiles. These data are essential for defining evolutionarily significant units, developing science-based management plans, and monitoring genetic health over time [10]. Integrating genomic evidence into conservation actions will enable the design of strategies to preserve the adaptive potential and long-term viability of this Critically Endangered species [3].

From an evolutionary standpoint, the *H. orcesi* genome provides key insights into the molecular basis of adaptation to high-elevation, semi-arid ecosystems in the Andes. The species persists under marked seasonality, limited water availability, and high diurnal thermal variation, conditions that impose significant physiological and genetic challenges [48]. Genomic analyses will help test hypotheses on how selection has influenced metabolism, thermoregulation, and cellular stress responses in reptiles inhabiting montane dry valleys [49], a biome that remains largely unexplored at the genomic

level. As one of the few high-altitude representatives of a mainly lowland lineage, *H. orcesi* offers a valuable model to investigate the genomic signatures associated with ecological transitions across altitude gradients in tropical lineages.

The genome of *H. orcesi* opens new opportunities for integrative research that combines evolution, ecology, and conservation. A key next priority will be to identify candidate genes and regulatory pathways associated with high-altitude adaptation, particularly those linked to oxygen transport, energy metabolism, and UV damage repair. Comparative analyses with lowland congeners and other teiids will help distinguish lineage-specific adaptations from broader reptilian genomic patterns, providing insight into how environmental pressures have driven diversification within the group [40]. Population genomic and landscape genetic approaches are also needed to evaluate connectivity and demographic trends among remnant populations. These analyses will inform conservation priorities by identifying barriers to gene flow, estimating effective population sizes, and assessing the genetic consequences of habitat loss and other anthropogenic pressures [50]. Integrating genomic, ecological, and spatial data will strengthen our understanding of Andean evolutionary dynamics and guide the development of evidence-based strategies for conserving *H. orcesi* and other reptiles inhabiting fragile montane dry ecosystems.

Conclusions

The genome assembly of *Holcosus orcesi* represents a major advance for the study and conservation of Ecuador's endemic reptiles. Its high contiguity and completeness establish it as a reference-quality resource for future genomic investigations. Structural and functional annotations are consistent with those reported for other squamate genomes, further supporting the reliability of the assembly and the accuracy of the annotations.

This genomic resource opens opportunities to assess genetic diversity, population connectivity, and adaptive capacity in *H. orcesi*, a species with a critically restricted range and high susceptibility to habitat change. It also provides a platform for exploring the molecular mechanisms underlying adaptation to high-altitude and semi-arid Andean environments, which remain underrepresented in vertebrate genomics.

Integrating this reference genome with population-level sequencing will be essential for identifying candidate genes associated with thermoregulation, hypoxia tolerance, and cellular stress response.

Comparative analyses within the Teiidae family will help clarify evolutionary trajectories across altitudinal gradients, while conservation genomics applications will inform management strategies to ensure the long-term persistence of *H. orcesi*. This genome stands as both a scientific milestone and a practical tool for safeguarding Ecuador's unique Andean biodiversity.

Acknowledgments

We thank Wilson Ochoa for his invaluable help in locating the specimen of *Holcosus orcesi*, Renato Morales for his assistance in transporting the specimen from Cuenca to Quito, Carolina Reyes-Puig and Emilia Peñaherrera for their assistance at the Laboratory of Terrestrial Zoology and Museum of Zoology at USFQ, and Carolina Sáenz, TUERI Wildlife Hospital, for collecting the blood sample. We thank the members of the Plant Biotechnology Laboratory (USFQ) for their support in DNA extraction, sequencing, and bioinformatics analyses.

Author contributions

Conceptualization: Gabriela Pozo, Diego F. Cisneros-Heredia, Maria Torres.

Formal analysis: Gabriela Pozo, Doménica Barragán-Orbe.

Funding acquisition: Gabriela Pozo, Diego F. Cisneros-Heredia, Maria Torres.

Investigation: Gabriela Pozo, Diego F. Cisneros-Heredia, Doménica Barragán-Orbe, Juan C. Sánchez-Nivicela, Ernesto Arbeláez.

Methodology: Gabriela Pozo, Maria Torres.

Supervision: Maria Torres.

Writing – original draft: Gabriela Pozo, Diego F. Cisneros-Heredia, Doménica Barragán-Orbe, Juan C. Sánchez-Nivicela, Ernesto Arbeláez, Maria Torres.

Writing – review & editing: Gabriela Pozo, Diego F. Cisneros-Heredia, Maria Torres.

References

1. Peters JA. The lizard genus Ameiva in Ecuador. Bull South Calif Acad Sci. 1964;63(3):113–27.
2. Harvey MB, Ugueto GN, Gutberlet RL. Review of teiid morphology with a revised taxonomy and phylogeny of the teiidae (Lepidosauria: squamata). Zootaxa. 2012;3459(1). https://doi.org/10.11646/zootaxa.3459.1.1
3. Parra D, León J, Arbeláez E, Siavichay F, Arpi J, Juela F. Plan de acción para la conservación de la lagartija coliazul (Teiidae): Holcosus orcesi. Fundación Amaru, Fundación Jocotoco Ecuador, Universidad San Francisco de Quito USFQ, Ministerio de Ambiente, Agua y Transición Ecológica del Ecuador; 2024.
4. IUCN. *Holcosus orcesi*: Cisneros-Heredia, D., Yánez-Muñoz, M., Brito, J. & Sánchez, J.: The IUCN Red List of Threatened Species 2017:e. T49981813A49981970. 2016 [cited 2025 Nov 11]. Available from: https://www.iucnredlist.org/species/49981813/49981970doi:10.2305/IUCN.UK.2017-2.RLTS.T49981813A49981970.en
5. Luna-Romero A, Ramírez I, Sánchez C, Conde J, Agurto L, Villaseñor D. Spatio-temporal distribution of precipitation in the Jubones river basin, Ecuador: 1975-2013. Sci.agropec. 2018;9(1):63–70. https://doi.org/10.17268/sci.agropecu.2018.01.07
6. Ochoa A, Campozano L, Sánchez E, Gualán R, Samaniego E. Evaluation of downscaled estimates of monthly temperature and precipitation for a Southern Ecuador case study. Intl Journal of Climatology. 2015;36(3):1244–55. https://doi.org/10.1002/joc.4418
7. Hasan MM, Wyseure G. Impact of climate change on hydropower generation in Rio Jubones Basin, Ecuador. Water Sci Eng. 2018 Apr;11(2):157–66. https://doi.org/10.1016/j.wse.2018.07.002
8. Formenti G, Theissinger K, Fernandes C, Bista I, Bombarely A, Bleidorn C, et al. The era of reference genomes in conservation genomics. Trends Ecol Evol. 2022;37(3):197–202. https://doi.org/10.1016/j.tree.2021.11.008 PMID: 35086739
9. Theissinger K, Fernandes C, Formenti G, Bista I, Berg PR, Bleidorn C, et al. How genomics can help biodiversity conservation. Trends Genet. 2023;39(7):545–59. https://doi.org/10.1016/j.tig.2023.01.005 PMID: 36801111
10. Supple MA, Shapiro B. Conservation of biodiversity in the genomics era. Genome Biol. 2018;19(1):131. https://doi.org/10.1186/s13059-018-1520-3 PMID: 30205843
11. Qiu Q, Zhang G, Ma T, Qian W, Wang J, Ye Z, et al. The yak genome and adaptation to life at high altitude. Nat Genet. 2012;44(8):946–9. https://doi.org/10.1038/ng.2343 PMID: 22751099
12. Qu Y, Zhao H, Han N, Zhou G, Song G, Gao B, et al. Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. Nat Commun. 2013;4:2071. https://doi.org/10.1038/ncomms3071 PMID: 23817352
13. Zhang T, Chen J, Zhang J, Guo Y-T, Zhou X, Li M-W, et al. Phenotypic and genomic adaptations to the extremely high elevation in plateau zokor (*Myospalax baileyi*). Mol Ecol. 2021;30(22):5765–79. https://doi.org/10.1111/mec.16174 PMID: 34510615
14. Zhang W, Yuan C, An X, Guo T, Wei C, Lu Z, et al. Genomic insights into Tibetan sheep adaptation to different altitude environments. Int J Mol Sci. 2024;25(22):12394. https://doi.org/10.3390/ijms252212394
15. Kardos M, Taylor HR, Ellegren H, Luikart G, Allendorf FW. Genomics advances the study of inbreeding depression in the wild. Evol Appl. 2016;9(10):1205–18. https://doi.org/10.1111/eva.12414 PMID: 27877200
16. Van Oosterhout C, Speak SA, Birley T, Hitchings LW, Bortoluzzi C, Percival-Alwyn L, et al. Genomic erosion in the assessment of species' extinction risk and recovery potential. Genomics. 2022. https://doi.org/10.1101/2022.09.13.507768
17. Kardos M, Åkesson M, Fountain T, Flagstad Ø, Liberg O, Olason P, et al. Genomic consequences of intensive inbreeding in an isolated wolf population. Nat Ecol Evol. 2018;2(1):124–31. https://doi.org/10.1038/s41559-017-0375-4 PMID: 29158554
18. Linck EB, Cadena CD. A latitudinal gradient of reference genomes. Mol Ecol. 2025;34(23):e17551. https://doi.org/10.1111/mec.17551 PMID: 39400919
19. Beaupre SJ, Jacobson ER, Lillywhite HB, Zamudio K. Guidelines for use of live amphibians and reptiles in field and laboratory research. Second ed; 2004.
20. Wick R. Porechop: adapter trimmer for Oxford Nanopore reads. Available from: https://github.com/rrwick/Porechop. 2017.
21. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics. 2018;34(15):2666–9. https://doi.org/10.1093/bioinformatics/bty149 PMID: 29547981
22. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18(2):170–5. https://doi.org/10.1038/s41592-020-01056-5 PMID: 33526886
23. Li H. Gfatools. Available from: https://github.com/lh3/gfatools. 2018.
24. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094–100. https://doi.org/10.1093/bioinformatics/bty191

25. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res. 2017;27(5):737–46. https://doi.org/10.1101/gr.214270.116 PMID: 28100585
26. 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(10):4647–54. https://doi.org/10.1093/molbev/msab199 PMID: 34320186
27. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–5. https://doi.org/10.1093/bioinformatics/btt086 PMID: 23422339
28. 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(9):1236–40. https://doi.org/10.1093/bioinformatics/btu031 PMID: 24451626
29. Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods. 2021;18(4):366–8. https://doi.org/10.1038/s41592-021-01101-x PMID: 33828273
30. Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C. RepeatModeler2 for automated genomic discovery of transposable element families. Proc Natl Acad Sci USA. 2020;117(17):9451–7. https://doi.org/10.1073/pnas.1921046117
31. Smit A, Hubley R, Green P. RepeatMasker Open-4.0. Available from: http://www.repeatmasker.org. 2013.
32. Colston TJ, Pirro S, Pyron RA. The complete genome sequences of 101 species of reptiles. Biodivers Genomes. 2025. https://doi.org/10.56179/001c.129597
33. Westfall AK, Telemeco RS, Grizante MB, Waits DS, Clark AD, Simpson DY, et al. A chromosome-level genome assembly for the eastern fence lizard (*Sceloporus undulatus*), a reptile model for physiological and evolutionary ecology. Gigascience. 2021;10(10):giab066. https://doi.org/10.1093/gigascience/giab066 PMID: 34599334
34. Pinto BJ, Gamble T, Smith CH, Keating SE, Havird JC, Chiari Y. The revised reference genome of the leopard gecko (*Eublepharis macularius*) provides insight into the considerations of genome phasing and assembly. J Hered. 2023;114(5):513–20. https://doi.org/10.1093/jhered/esad016 PMID: 36869788
35. Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, et al. Towards complete and error-free genome assemblies of all vertebrate species. Nature. 2021;592(7856):737–46. https://doi.org/10.1038/s41586-021-03451-0 PMID: 33911273
36. Gable SM, Mendez JM, Bushroe NA, Wilson A, Byars MI, Tollis M. The state of squamate genomics: past, present, and future of genome research in the most speciose terrestrial vertebrate order. Genes (Basel). 2023;14(7):1387. https://doi.org/10.3390/genes14071387 PMID: 37510292
37. Alföldi J, Di Palma F, Grabherr M, Williams C, Kong L, Mauceli E, et al. The genome of the green anole lizard and a comparative analysis with birds and mammals. Nature. 2011;477(7366):587–91. https://doi.org/10.1038/nature10390 PMID: 21881562
38. Roscito JG, Sameith K, Pippel M, Francoijs KJ, Winkler S, Dahl A, et al. The genome of the tegu lizard Salvator merianae: combining Illumina, PacBio, and optical mapping data to generate a highly contiguous assembly. GigaScience. 2018;7(12):giy141. https://doi.org/10.1093/gigascience/giy141
39. Organ CL, Moreno RG, Edwards SV. Three tiers of genome evolution in reptiles. Integr Comp Biol. 2008;48(4):494–504. https://doi.org/10.1093/icb/icn046 PMID: 21669810
40. Pasquesi GIM, Adams RH, Card DC, Schield DR, Corbin AB, Perry BW, et al. Squamate reptiles challenge paradigms of genomic repeat element evolution set by birds and mammals. Nat Commun. 2018;9(1):2774. https://doi.org/10.1038/s41467-018-05279-1 PMID: 30018307
41. Schield DR, Card DC, Hales NR, Perry BW, Pasquesi GM, Blackmon H, et al. The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. Genome Res. 2019;29(4):590–601. https://doi.org/10.1101/gr.240952.118 PMID: 30898880
42. Yin W, Wang Z ji, Li Q ye, Lian J ming, Zhou Y, Lu B zheng, et al. Evolutionary trajectories of snake genes and genomes revealed by comparative analyses of five-pacer viper. Nat Commun. 2016;7(1):13107. https://doi.org/10.1038/ncomms13107
43. Gomez-Garrido J, Cruz F, Alioto TS, Feiner N, Uller T, Gut M, et al. Chromosome-level genome assembly of Lilford's wall lizard, Podarcis lilfordi (Günther, 1874) from the Balearic Islands (Spain). DNA Research. 2023;30(3). https://doi.org/10.1093/dnares/dsad008
44. Geneva AJ, Park S, Bock DG, de Mello PLH, Sarigol F, Tollis M, et al. Chromosome-scale genome assembly of the brown anole (Anolis sagrei), an emerging model species. Commun Biol. 2022;5(1):1126. https://doi.org/10.1038/s42003-022-04074-5 PMID: 36284162
45. Kapusta A, Suh A, Feschotte C. Dynamics of genome size evolution in birds and mammals. Proc Natl Acad Sci USA. 2017;114(8). https://doi.org/10.1073/pnas.1616702114
46. Chalopin D, Naville M, Plard F, Galiana D, Volff J-N. Comparative analysis of transposable elements highlights mobilome diversity and evolution in vertebrates. Genome Biol Evol. 2015;7(2):567–80. https://doi.org/10.1093/gbe/evv005 PMID: 25577199
47. Shaffer HB, Gidiş M, McCartney-Melstad E, Neal KM, Oyamaguchi HM, Tellez M, et al. Conservation genetics and genomics of amphibians and reptiles. Annu Rev Anim Biosci. 2015;3:113–38. https://doi.org/10.1146/annurev-animal-022114-110920 PMID: 25580719
48. Navas CA. Herpetological diversity along Andean elevational gradients: links with physiological ecology and evolutionary physiology. Comp Biochem Physiol A Mol Integr Physiol. 2002;133(3):469–85. https://doi.org/10.1016/s1095-6433(02)00207-6 PMID: 12443907
49. Cheviron ZA, Brumfield RT. Genomic insights into adaptation to high-altitude environments. Heredity (Edinb). 2012;108(4):354–61. https://doi.org/10.1038/hdy.2011.85 PMID: 21934702
50. Hohenlohe PA, Funk WC, Rajora OP. Population genomics for wildlife conservation and management. Mol Ecol. 2021;30(1):62–82. https://doi.org/10.1111/mec.15720 PMID: 33145846