DATA NOTE

A chromosome-level reference genome for the critically endangered Southern Corroboree frog (*Pseudophryne corroboree*)

[version 1; peer review: 1 approved]

Tiffany A. Kosch[1], Lee Berger[1], Andrew J. Crawford[2], Deon Gilbert[3], Damian Goodall[3], David Hunter[4], Michael McFadden[5], Linelle Abueg[6], Giulio Formenti[6], Joanna Collins[7], Tatiana Tilley[6], Nivesh Jain[6], Brian O'Toole[6], Conor Whelan[6], Ira Cooke[8], Matthieu Muffato[7], Bethan Yates[7], Olivier Fedrigo[6,9], Jennifer Balacco[6], Erich D. Jarvis[6], Lee F. Skerratt[1]

[1]Melbourne Veterinary School, Faculty of Science, The University of Melbourne, Melbourne, Victoria, Australia
[2]Departamento de Ciencias Biológicas, Universidad de Los Andes, Bogotá, Colombia
[3]Wildlife Conservation and Science, Zoos Victoria, Parkville, Victoria, Australia
[4]NSW Department of Climate Change, Energy, Environment, and Water, Parramatta, NSW, Australia
[5]Taronga Institute of Science and Learning, Taronga Conservation Society Australia, Mosman, New South Wales, Australia
[6]The Vertebrate Genome Laboratory, The Rockefeller University, New York, New York, USA
[7]Wellcome Sanger Institute, Hinxton, England, UK
[8]Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, Townsville City, Queensland, Australia
[9]Colossal Bioscienes, Dallas, Texas, USA

v1 **First published:** 30 Apr 2025, **10**:228
https://doi.org/10.12688/wellcomeopenres.23820.1
Latest published: 30 Apr 2025, **10**:228
https://doi.org/10.12688/wellcomeopenres.23820.1

Abstract

The Southern Corroboree frog (*Pseudophryne corroboree*; Anura; Myobatrachidae) is a Critically Endangered amphibian, according to the IUCN, and is endemic to the Snowy Mountains region of Kosciuszko National Park in New South Wales, Australia. This species has been driven to functional extinction by the introduction of the fungal disease, chytridiomycosis. Here we provide the first reference genome for *P. corroboree*. Using PacBio HiFi sequencing, Arima Hi-C, and Bionano optical mapping, we produced a chromosome-level genome assembly. Additionally, we generated a reference transcriptome based on multiple tissues from both male and female individuals to support genome annotation. The resulting genome spans 8.87 Gb across 12 chromosomes, with a contig N50 of 6.8 Mb. This research provides a phased, annotated genome assembly along with transcriptomic resources to facilitate future conservation genomic studies of *P. corroboree*. Furthermore, the genome offers an

Open Peer Review

Approval Status ✓

	1
version 1 30 Apr 2025	✓ view

1. **Yanbo Sun**, Yunnan University, Kunming, China

Any reports and responses or comments on the article can be found at the end of the article.

invaluable resource for taxonomic and evolutionary research, particularly given the nearest available chromosome-level reference genome is from *Mixophyes fleayi*, a species that last shared a common ancestor with *P. corroboree* 80 million years ago.

Keywords
Genome assembly, reference genome, Anura, Critically Endangered, Myobatrachidae, conservation breeding

Corresponding author: Tiffany A. Kosch (tiffany.kosch@gmail.com)

Author roles: Kosch TA: Conceptualization, Data Curation, Project Administration, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Berger L**: Conceptualization, Investigation, Resources, Writing – Review & Editing; **Crawford AJ**: Conceptualization, Writing – Review & Editing; **Gilbert D**: Conceptualization, Writing – Review & Editing; **Goodall D**: Resources, Writing – Review & Editing; **Hunter D**: Conceptualization, Writing – Review & Editing; **McFadden M**: Conceptualization, Writing – Review & Editing; **Abueg L**: Formal Analysis, Writing – Review & Editing; **Formenti G**: Formal Analysis, Writing – Review & Editing; **Collins J**: Formal Analysis, Writing – Review & Editing; **Tilley T**: Investigation, Writing – Review & Editing; **Jain N**: Formal Analysis, Investigation, Writing – Review & Editing; **O'Toole B**: Investigation, Writing – Review & Editing; **Whelan C**: Investigation, Writing – Review & Editing; **Cooke I**: Formal Analysis, Writing – Review & Editing; **Muffato M**: Investigation, Visualization, Writing – Review & Editing; **Yates B**: Investigation, Visualization; **Fedrigo O**: Conceptualization, Formal Analysis, Project Administration, Writing – Review & Editing; **Balacco J**: Formal Analysis, Project Administration, Writing – Review & Editing; **Jarvis ED**: Conceptualization, Funding Acquisition, Project Administration, Resources, Writing – Review & Editing; **Skerratt LF**: Conceptualization, Funding Acquisition, Project Administration, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The post assembly genome QC work and the genome note template formatting was supported by Wellcome through core funding to the Wellcome Sanger Institute [206194, https://doi.org/10.35802/206194] and the Darwin Tree of Life Discretionary Award [218328, https://doi.org/10.35802/218328]. VGL work was supported by a Howard Hughes Medical Institute grant awarded to E.D.J., and T.A.K., L.F.S, and L.B. were supported by Australian Research Council grants [FT190100462 and LP200301370] awarded to L.F.S.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

How to cite this article: Kosch TA, Berger L, Crawford AJ *et al.* **A chromosome-level reference genome for the critically endangered Southern Corroboree frog (*Pseudophryne corroboree*) [version 1; peer review: 1 approved]** Wellcome Open Research 2025, **10**:228 https://doi.org/10.12688/wellcomeopenres.23820.1

First published: 30 Apr 2025, **10**:228 https://doi.org/10.12688/wellcomeopenres.23820.1

Species taxonomy

Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amphibia; Batrachia; Anura; Neobatrachia; Myobatrachoidea; Myobatrachidae; Myobatrachinae; Pseudophryne *Pseudophryne corroboree* Moore, 1953 (NCBI:txid495146).

Background

The genome of a Southern Corroboree frog (*Pseudophryne corroboree*), referred to as "Gyack" by Australian Indigenous people[1], was sequenced as part of the Vertebrate Genomes Project (VGP), a collaborative effort aiming to produce high-quality reference genomes of all named vertebrate species[2,3]. Here we present a chromosome-level complete genome assembly for *P. corroboree*, from a male captive-bred specimen (BioSample SAMN32631236) provided by Melbourne Zoo (Victoria, Australia).

The Southern Corroboree frog is one of Australia's most threatened species of amphibians[4]. Driven to functional extinction in the wild after the introduction of the fungal pathogen, *Batrachochytrium dendrobatidis*, into Australia in the late 1970s, this species is now reliant on *ex situ* conservation breeding for its continued existence[5,6].

Pseudophryne corroboree is a sub-alpine frog restricted to sphagnum peat bogs from 1300 and 1760 meters in elevation in Kosciuszko National Park in the Snowy Mountains region of New South Wales (NSW), Australia[7]. The female lays terrestrial eggs in mid- to late summer that are attended by the male throughout development[8,9]. Eggs hatch into tadpoles after flooding by winter rains[8]. *Pseudophryne corroboree* have bright yellow and black coloration. Unlike many other poisonous frogs (e.g., Dendrobatidae, Mantellidae), they are able to synthesize toxins in addition to obtaining them from their diets[10]. The sister species, *P. pengilleyi*, has declined more slowly[4,11], which may be due to greater resistance to chytridiomycosis in this species.

Various efforts are underway to restore *P. corroboree* to the wild including conservation breeding and reintroduction, *in situ* enclosures, conservation genomics analysis, and targeted genetic intervention[6,12–15]. Here we report a high-quality reference genome and transcriptome to serve as a key resource for these efforts by facilitating informed genetic management of the species and efforts to increase resistance to the chytrid pathogen[12,16].

Methods

Sample acquisition

An adult captive-bred male *P. corroboree* (Zoo Record ID: B50597), obtained from the conservation breeding program at Melbourne Zoo (Victoria, Australia), was used to generate the reference genome. Two additional captive animals: 1 male (B40300; the sire of the focal animal, wild-collected from unknown location in 2007) and 1 female (B40182; wild-collected from Upper Jagumba, NSW in 2007) were also sequenced using short reads. All frogs were humanely euthanized at the University of Melbourne (Victoria, Australia) using a buffered solution of 0.2% tricaine mesylate (MS-222) and decapitation (University of Melbourne Animal Ethics permit #10267). Using sterile techniques, multiple tissues including heart, liver, kidney, muscle, gonads, brain, and whole body were collected from each animal and flash frozen in liquid nitrogen. The tissues were immediately transferred to -80°C until they were sent for genome sequencing. Tissues were shipped on dry ice to the Vertebrate Genomics Laboratory (Rockefeller University, New York, USA) for nucleotide extraction and sequencing (Australian wildlife trade export permit #PWS2020-AU-001530).

Nucleic acid extraction and sequencing

For PacBio HiFi sequencing, high molecular weight DNA (HMW DNA) was extracted from a flash frozen kidney using the MagAttract HMW DNA Kit (Qiagen 67563). The tissue was stored at -80°C and kept on dry ice until homogenized with the Qiagen TissueRuptor II (Cat. No. 9002755). The DNA was quantified with the Qubit 3 fluorometer (Invitrogen Qubit dsDNA Broad Range Assay cat no. Q32850) and fragment size was assessed with the Agilent Fragment Analyzer.

After isolation, DNA was sheared using the Megaruptor 3 (Diagenode, Denville, NJ, USA) to attain a 15–20 Kb PacBio library insert size. The PacBio HiFi library was prepared with the SMRTbell Express Template Prep Kit 2.0 (PN 100-938-900) following the manufacturer's procedure (PN 101-853-100 Version 03) and then size-selected with Pippin HT (Sage Science, Beverly, MA, USA). The PacBio library was sequenced on a Sequel IIe instrument with 8M SMRT cells and Sequencing Plate 2.0 (PN 101-820-200).

For Hi-C sequencing, libraries were prepared from liver using the Arima-HiC 2.0 kit (Arima Genomics, Carlsbad, CA, USA) following the manufacturer's protocol. The library was then sequenced with the Illumina NovaSeq 6000 platform with 2×150 bp read length at Psomagen, Inc. (Rockville, MD, USA).

For Bionano optical mapping, HMW DNA was extracted from the kidney with the Circulomics Nanobind Tissue Big DNA Kit and fragment size was evaluated with a pulsed field gel electrophoresis (Pippin Pulse, SAGE Science, Beverly, MA). The DNA was labelled using direct labelling enzyme (DLE1) and Bionano Prep Direct Label and Stain (DLS) protocol (document number 30206) and then labels sequenced on a Bionano Saphyr instrument.

Total RNA was extracted from five tissues from the focal male (B50597: testes, brain, muscle, liver, and whole body) and four tissues from the female (B40182: ovary, muscle, liver, and whole body) using a QIAGEN RNAeasy Protect kit (cat. no. 74124). RNA quantity was determined using a Qubit 3 fluorometer (Invitrogen Qubit RNA High Sensitivity (HS) Kit (cat. no. Q32852) and RNA integrity (RIN) score determined using an Agilent Fragment Analyzer. RNA paired-end sequencing was performed on an Illumina NovaSeq 6000 machine.

Short read Illumina WGS sequencing of the presumed parents was performed using HMW DNA extracted from kidney tissue from the presumed parents (B40300, B40182) of the focal frog. The library was sequenced with the Illumina NovaSeq

6000 platform with 2×150 bp read length at Psomagen, Inc. (Rockville, MD, USA).

Genome assembly, curation, evaluation, and annotation

Assembly

The assembly was performed using the VGP Pipeline 2.0 with Hi-C phasing[3]. First, HiFi reads were screened for PacBio adapters using cutadapt 4.0 to remove any matching reads before assembly. Contig assembly was then performed using hifiasm version 0.16.1 in Hi-C phased mode to obtain two phased haplotypes. Contigs of each haplotype were then scaffolded individually using Bionano Solve 3.7.0 followed by YaHS 1.2a.

Assembly curation

Haplotype 2 had higher assembly metrics and thus was selected for manual curation and to be used as the main reference. The assembly was decontaminated using the Assembly Screen for Cobionts and Contaminants (ASCC) pipeline (https://pipelines.tol.sanger.ac.uk/ascc; Aunin *et al.* in prep). Flat files and Hi-C contact maps used in curation were generated via the TreeVal pipeline[17]. Manual curation was conducted using the rapid curation pipeline documented at (https://gitlab.com/wtsi-grit/rapid-curation; Wood *et al.* in prep) primarily using PretextView and HiGlass[18] with additional insights provided by JBrowse2[19]. Scaffolds were visually inspected, and assembly errors corrected as described in Howe *et al.*[20]. Any identified contamination, missed joins, and mis-joins were amended, and duplicate sequences were tagged and removed.

Evaluation of the final assembly

The final assembly was post-processed and evaluated with Nextflow[21]. DSL2 pipelines "sanger-tol/readmapping"[22], "sanger-tol/genomenote"[23], and "sanger-tol/blobtoolkit"[24]. The pipeline sanger-tol/readmapping aligns the Hi-C reads with bwa-mem2[25] and combines the alignment files with SAMtools[26]. The sanger-tol/genomenote pipeline transforms the Hi-C alignments into a contact map with BEDTools[27] and the Cooler tool suite[28], which is then visualised with HiGlass[18]. It also provides statistics about the assembly with the NCBI datasets report[29], computes *k*-mer completeness and QV consensus quality values with FastK and MERQURY.FK, and provides a completeness assessment with BUSCO[30].

The sanger-tol/blobtoolkit pipeline is a Nextflow port of the previous Snakemake Blobtoolkit pipeline[31]. It aligns the PacBio reads with SAMtools and minimap2[32] and generates coverage tracks for regions of fixed size. In parallel, it queries the GoaT database[33] to identify all matching BUSCO lineages to run BUSCO[30]. For the three domain-level BUSCO lineages, the pipeline aligns the BUSCO genes to the Uniprot Reference Proteomes database[34] with DIAMOND blastp[35]. The genome was split into chunks according to the density of the BUSCO genes from the closest taxonomically lineage, and each chunk was aligned to the Uniprot Reference Proteomes database with DIAMOND blastx. Genome sequences that had no hit were then chunked with seqtk and aligned to the NT database with blastn[36]. All outputs were then combined with the blobtools suite into a blobdir for visualisation.

The genome assembly and evaluation pipelines were developed using the nf-core tooling[37], using MultiQC[38], and making extensive use of the Conda package manager, the Bioconda initiative[39], the Biocontainers infrastructure[40], and the Docker[41] and Singularity[42] containerisation solutions.

Only one of the presumed parents, the male frog (B40300), was confirmed to be an actual parent. As a result, we could not utilise parent WGS data for haplotype phasing.

The curated assemblies were submitted to NCBI (BioProject PRJNA928730), and haplotype 2 was annotated by the NCBI Eukaryotic Genome Annotation Pipeline[43] using RNA-Seq data from the kidney and liver from the male frog (SAMN32631236) and brain, ovaries, and whole body from the female frog (SAMN39610159). This annotated RefSeq version of this assembly has the accession number GCF_028390025.1 (PRJNA1082331).

Repeats were *de novo* modelled with RepeatModeler (Apptainer v. 1.2.3)[44] and then annotated using RepeatMasker (v. 4.1.2-p1)[45] with a concatenated library of genome-specific repeats generated from RepeatModeler and the Dfam amphibian repeat library (v. Dfam.h5)[46].

Results

Genome sequence report

The assembly of a male *Pseudophryne corroboree* (Figure 1) resulted in a genome that was 8.87 Gb in length across 12 chromosomes. The genome was sequenced using PacBio HiFi reads, generating a total of 230 Gb from 17,266,474 reads, providing approximately 26.9-fold coverage. Primary assembly contigs were scaffolded with chromosome conformation Hi-C data, which produced 381 Gb from 1,266,207,409 reads, yielding an approximate coverage of 44.5-fold. Specimen and sequencing information are summarised in Table 1.

Manual assembly curation corrected 280 missed joins or mis-joins and 1 haplotypic duplication, reducing the assembly length by 0.79% and the scaffold number by 11%, and increasing the scaffold N50 by 24.6%. The final assembly has a total of 3,127 scaffolds, with 2,699 gaps, and a large scaffold N50 of 846.9 Mb and a contig N50 of 6.8 Mb (Table 2), well surpassing the VGP minimum metrics of 10 and 1 Mb, respectively[2]. The snail plot in Figure 2 provides a summary of the assembly statistics, while the distribution of assembly

Figure 1. Photograph of a captive-bred *Pseudophryne corroboree*. Photo by C. Doughty.

Table 1. **Specimen and sequencing data for *Pseudophryne corroboree*.**

Project information			
Study title	Pseudophryne corroboree overview		
Umbrella BioProject	PRJNA921726		
Species	*Pseudophryne corroboree*		
BioSample	SAMN32631236		
NCBI taxonomy ID	495146		
Specimen information			
Technology	**Specimen ID (sex)**	**BioSample accession**	**Organism part**
PacBio long read sequencing	B50597 (male)	SAMN32631236	Kidney, liver
Hi-C sequencing	B50597 (male)	SAMN32631236	Liver
RNA sequencing	B50597 (male)	SAMN32631236	Brain, muscle, testes, liver, whole body
RNA sequencing	B40182 (female)	SAMN39610159	Ovary, muscle, liver, whole body
Illumina DNA sequencing	B40182 (female)	SAMN39610159	Kidney
Illumina DNA sequencing	B40300 (male)	SAMN40858189	Kidney

Table 2. **Genome assembly data for *Pseudophryne corroboree*, aPseCor3.hap2.**

Genome assembly		
Assembly name	aPseCor3.hap2	
Assembly accession haplotype	GCA_028390025.1 haplotype 2; GCA_028390055.1 haplotype 1	
Assembly length	8872758793	
Chromosome assembly length	8190274449	
Number of contigs	5826	
Contig N50 length (Mb)	6.8	
Number of scaffolds	3127	
Scaffold N50 length (Mb)	846.9	
Longest scaffold (Mb)	947.25	
Assembly metrics*		***Benchmark***
Consensus quality (QV)	58.5	*≥ 50*
k-mer completeness	97.66%	*≥ 95%*
BUSCO**	C:89.8% [S:88.3%, D:1.5%], F:3.0%, M:7.2%, n:5310	*C ≥ 95%*
Percentage of assembly mapped to chromosomes (N=12)	92.3%	*≥ 95%*
Genome annotation of assembly GCF_028390025.1-RS_2024_02 at RefSeq		
Number of protein-coding genes	24591	
Number of non-coding genes	128683	
Number of gene transcripts	183540	

Genome repeat content		
Repeat element	**Number of elements**	**% of genome**
DNA transposons	1749727	11.7
LINEs	821317	5.75
SINEs	134337	1.04
LTRs	1435160	27.15
Simple	1353528	0.82
Unclassified	10152528	28.08

* Assembly metric benchmarks are adapted from column VGP-2020 of "Table 1: Proposed standards and metrics for defining genome assembly quality" from Rhie *et al.* (2021).

** BUSCO scores based on the tetrapoda_odb10 BUSCO reference set using version 5.4.3. C = complete [S = single copy, D = duplicated], F = fragmented, M = missing, n = number of orthologues in comparison. A full set of BUSCO scores is available at https://blobtoolkit.genomehubs.org/view/GCA_028390025.1/dataset/GCA_028390025.1/busco.

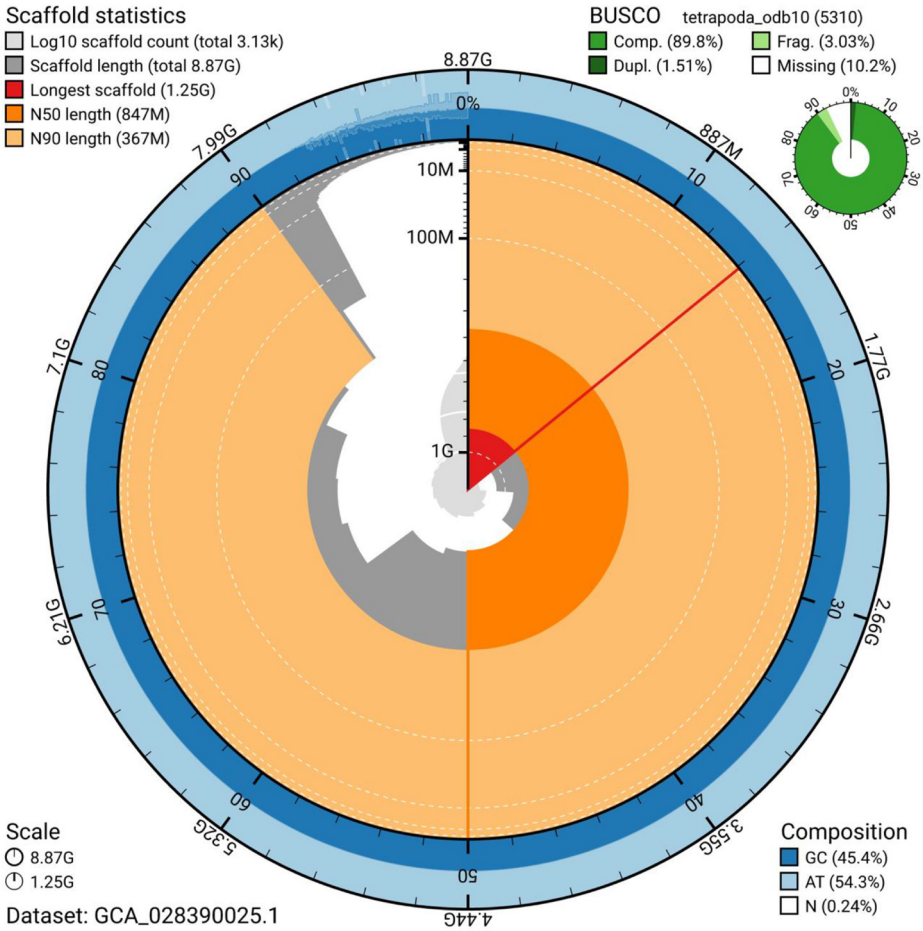

Figure 2. Genome assembly metrics for *Pseudophryne corroboree*, aPseCor3.hap2. The BlobToolKit snail plot shows N50 scaffold metrics and BUSCO gene completeness. An interactive version of this figure is available at: https://blobtoolkit.genomehubs.org/view/GCA_028390025.1/dataset/GCA_028390025.1/snail.

scaffolds by GC proportion and coverage is shown in Figure 3. The cumulative assembly plot in Figure 4 shows curves for subsets of scaffolds assigned to different phyla. Most (92.3%) of the assembly sequence was assigned to 12 chromosomal-level scaffolds. These chromosome-scale scaffolds were confirmed by the Hi-C data and are named in order of size (Figure 5; Table 3).

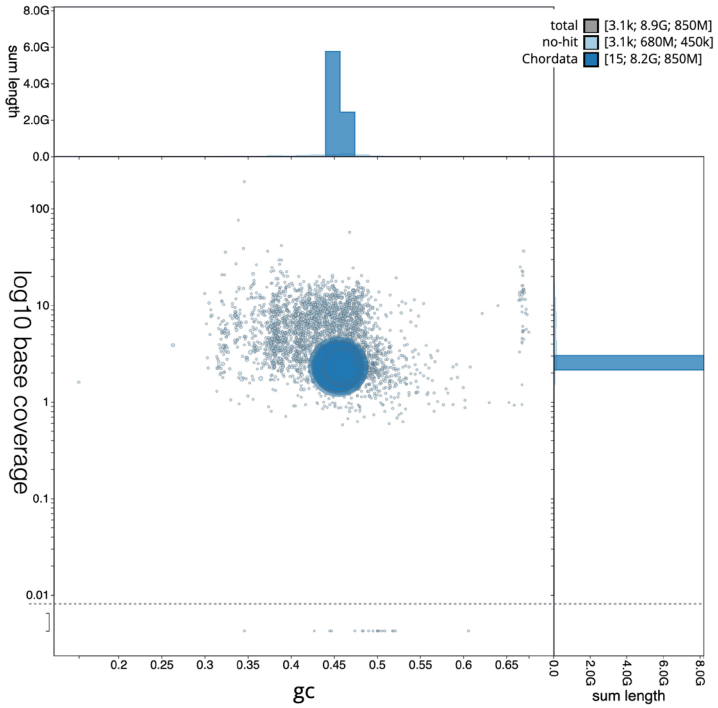

Figure 3. BlobToolKit plot of base against GC proportion for genome assembly *Pseudophryne corroboree*, aPseCor3.hap2. Circles are sized in proportion to sequence length. Histograms show the distribution of sequence length sum along each axis. Base coverage values are log10 transformed. An interactive version of this figure is available at: https://blobtoolkit.genomehubs.org/view/GCA_028390025.1/dataset/GCA_028390025.1/blob?plotShape=circle.

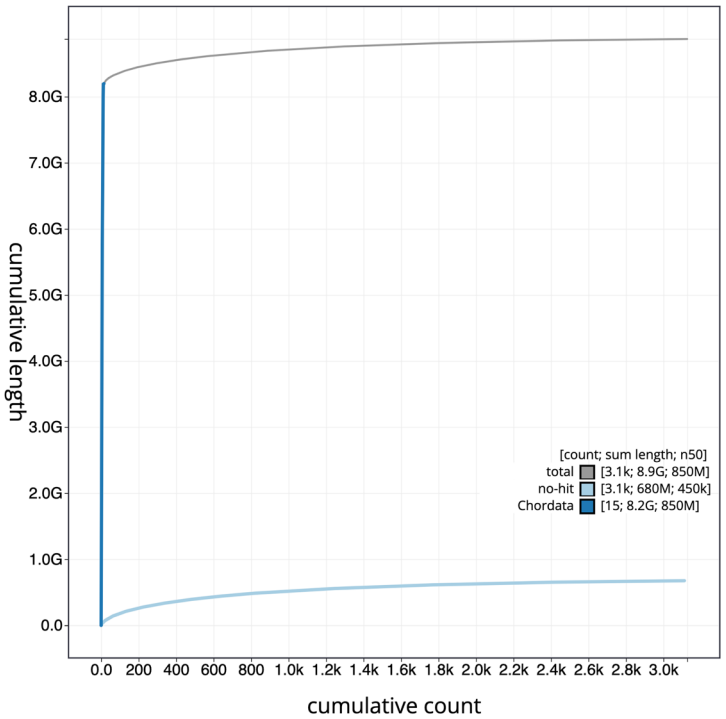

Figure 4. Genome assembly of *Pseudophryne corroboree*, aPseCor3.hap2: BlobToolKit cumulative sequence plot. The grey line shows cumulative length for all sequences. Coloured lines show cumulative lengths of sequences assigned to each phylum using the buscogenes taxrule. An interactive version of this figure is available at: https://blobtoolkit.genomehubs.org/view/GCA_028390025.1/dataset/GCA_028390025.1/cumulative.

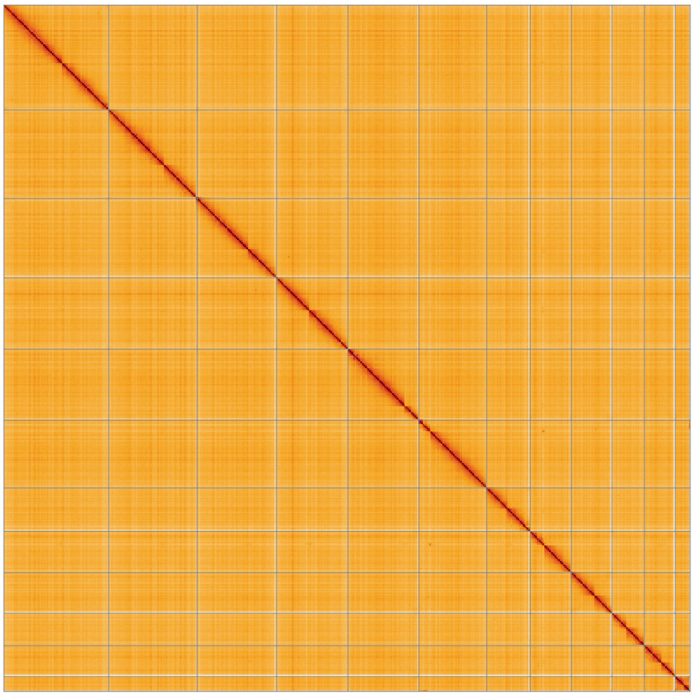

Figure 5. Hi-C contact map of the curated aPseCor3.hap2 assembly, visualised using HiGlass. Chromosomes are shown in order of size from left to right and top to bottom. An interactive version of this figure may be viewed at https://genome-note-higlass.tol.sanger.ac.uk/l/?d=RTD6MMzAT-mgaBhXdbrKig.

Table 3. Chromosomal pseudomolecules in the genome assembly of *Pseudophryne corroboree*.

GenBank accession	Name	Length (Mb)	GC%
CM051801.1	1	1251.52	45.5
CM051802.1	2	1055.99	46
CM051803.1	3	808.54	46
CM051804.1	4	947.25	45.5
CM051805.1	5	852.35	45.5
CM051806.1	6	846.89	45.5
CM051807.1	7	518.98	46
CM051808.1	8	493.02	46
CM051809.1	9	483.89	46
CM051810.1	10	387.42	46
CM051811.1	11	366.57	46
CM051812.1	12	177.84	46.5

The estimated Quality Value (QV) of the final assembly is 58.5 (no more than 1 error per ~1Mb) and BUSCO v5.5.0 completeness of 89.8% (single = 88.3%, duplicated = 1.5%), using the tetrapoda_odb10 reference set (n = 5,310). *K-mer* completeness for the combined haplotypes was 97.66% complete (aPseCor3.hap1 = 91.91% and aPseCor3.hap2 = 91.96%).

A considerable portion (76%) of the *P. corroboree* genome consisted of repeats, with the majority of these (27%) classified as Long Terminal Repeats (LTRs; Table 2).

Ethics and consent

Frogs were humanely euthanised following University of Melbourne (Victoria, Australia) Animal Ethics permit #10267. Tissue samples were exported for sequencing at Rockefeller University (New York, USA) under Australian wildlife trade export permit (#PWS2020-AU-001530).

Data availability

NCBI Archive: *Pseudophryne corroboree* overview. The genome sequence is released openly for reuse. The *Pseudophryne corroboree* genome sequencing initiative is part of the Vertebrate Genomes Project (VGP). Further, raw data and assembly accession identifiers are reported in Table 1 and Table 2. Relevant software tool versions and sources are listed in Table 4. Interactive genome figures are available at https://blobtoolkit.genomehubs.org/.

Table 4. Software tools: versions and sources.

Software tool	Version	Source
BEDTools	2.30.0	https://github.com/arq5x/bedtools2
BLAST	2.14.0	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
BlobToolKit	4.3.7	https://github.com/blobtoolkit/blobtoolkit
BUSCO	5.4.3 and 5.5.0	https://gitlab.com/ezlab/busco
bwa-mem2	2.2.1	https://github.com/bwa-mem2/bwa-mem2
Cooler	0.8.11	https://github.com/open2c/cooler
DIAMOND	2.1.8	https://github.com/bbuchfink/diamond
fasta_windows	0.2.4	https://github.com/tolkit/fasta_windows
FastK	427104ea91c78c3b8b8b49f1a7d6bbeaa869ba1c	https://github.com/thegenemyers/FASTK
Gfastats	1.3.6	https://github.com/vgl-hub/gfastats
GoaT CLI	0.2.5	https://github.com/genomehubs/goat-cli
Hifiasm	0.19.8-r603	https://github.com/chhylp123/hifiasm
HiGlass	44086069ee7d4d3f6f3f0012569789ec138f42b84aa44357826c0b6753eb28de	https://github.com/higlass/higlass
Merqury.FK	d00d98157618f4e8d1a9190026b19b471055b22e	https://github.com/thegenemyers/MERQURY.FK
MitoHiFi	3	https://github.com/marcelauliano/MitoHiFi
MultiQC	1.14, 1.17, and 1.18	https://github.com/MultiQC/MultiQC
NCBI Datasets	15.12.0	https://github.com/ncbi/datasets
Nextflow	23.04.0-5857	https://github.com/nextflow-io/nextflow
PretextView	0.2	https://github.com/sanger-tol/PretextView
purge_dups	1.2.5	https://github.com/dfguan/purge_dups
samtools	1.16.1, 1.17, and 1.18	https://github.com/samtools/samtools
repeatmodeler	Apptainer v. 1.2.3	https://github.com/Dfam-consortium/RepeatModeler/blob/master/RepeatModeler
repeatmasker	v. 4.1.2-p1	https://github.com/Dfam-consortium/RepeatMasker
sanger-tol/ascc	-	https://github.com/sanger-tol/ascc
sanger-tol/genomenote	1.1.1	https://github.com/sanger-tol/genomenote
sanger-tol/readmapping	1.2.1	https://github.com/sanger-tol/readmapping
Seqtk	1.3	https://github.com/lh3/seqtk
Singularity	3.9.0	https://github.com/sylabs/singularity
TreeVal	1.0.0	https://github.com/sanger-tol/treeval
YaHS	1.2a.2	https://github.com/c-zhou/yahs

Acknowledgements

We thank Zoos Victoria (Melbourne, Victoria, Australia) for providing the three frogs used for genome sequencing and Laura Brannelly for logistical help.

References

1. Williams S, Connolly D, Williams A: **The recognition of cultural water requirements in the montane rivers of the Snowy Mountains, Australia.** *Australas J Environ Manage.* 2019; **26**(3): 255–272.
Publisher Full Text

2. Rhie A, McCarthy SA, Fedrigo O, *et al.*: **Towards complete and error-free genome assemblies of all vertebrate species.** *Nature.* 2021; **592**(7856): 737–746.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Larivière D, Abueg L, Brajuka N, *et al.*: **Scalable, accessible and reproducible reference genome assembly and evaluation in Galaxy.** *Nat Biotechnol.* 2024; **42**(3): 367–370.
PubMed Abstract | Publisher Full Text | Free Full Text

4. Geyle HM, Hoskin CJ, Bower DS, *et al.*: **Red hot frogs: identifying the Australian frogs most at risk of extinction.** *Pac Conserv Biol.* 2021; **28**(3): 211–223.
Publisher Full Text

5. Kosch TA, Silva CNS, Brannelly LA, *et al.*: **Genetic potential for disease resistance in critically endangered amphibians decimated by chytridiomycosis.** *Anim Conserv.* 2019; **22**(3): 238–250.
Publisher Full Text

6. Hunter D, Clemann N, Coote D, *et al.*: **Frog declines and associated management response in South-Eastern Mainland Australia and Tasmania.** Australia: CSIRO Publ., 2018; 38–58.
Reference Source

7. Moore J: **A new species of *Pseudophryne* from Victoria.** *Proc Linn Soc NSW.* 1953; **78**(3–4): 179–180.
Reference Source

8. Hunter D, Osborne W, Smith M, *et al.*: **Breeding habitat use and the future management of the critically endangered Southern Corroboree frog.** *Ecol Manage Restor.* 2009; **10**(s1): S103–S109.
Publisher Full Text

9. Pengilley R: **Breeding biology of some species of *Pseudophryne* (Anura: Leptodactylidae) of the Southern Highlands.** *Aust Zool.* 1973; **18**(1): 15–30.
Reference Source

10. Daly JW, Garraffo HM, Pannell LK, *et al.*: **Alkaloids from Australian frogs (Myobatrachidae): pseudophrynamines and pumiliotoxins.** *J Nat Prod.* 1990; **53**(2): 407–21.
PubMed Abstract | Publisher Full Text

11. Hunter DA, Speare R, Marantelli G, *et al.*: **Presence of the amphibian chytrid fungus *Batrachochytrium dendrobatidis* in threatened corroboree frog populations in the Australian Alps.** *Dis Aquat Organ.* 2010; **92**(2–3): 209–16.
PubMed Abstract | Publisher Full Text

12. Kosch TA, Waddle AW, Cooper CA, *et al.*: **Genetic approaches for increasing fitness in endangered species.** *Trends Ecol Evol.* 2022; **37**(4): 332–345.
PubMed Abstract | Publisher Full Text

13. Berger L, Skerratt LF, Kosch TA, *et al.*: **Advances in managing chytridiomycosis for Australian frogs: *Gradarius Firmus Victoria*.** *Annu Rev Anim Biosci.* 2024; **12**(1): 113–133.
PubMed Abstract | Publisher Full Text

14. Hunter D, Marantelli G, McFadden M, *et al.*: **Assessment of re-introduction methods for the Southern corroboree frog in the Snowy Mountains region of Australia.** Global re-introduction perspectives: additional case-studies from around the globe IUCN/SSC Reintroduction Specialist Group, Abu Dhabi, 2010; 72–6.
Reference Source

15. Lees C, McFadden M, Hunter D: **Genetic management of Southern corroboree frogs: workshop report and plan.** IUCN Conservation Breeding Specialist Group, Apple Valley, MN, 2013.
Reference Source

16. Davidson MJ, Zenger K, Keogh JS, *et al.*: **Design and application of a genome-wide SNP array to improve conservation outcomes in the critically endangered Southern Corroboree frog.** Authorea, 2024.
Publisher Full Text

17. Pointon D, Eagles W, Sims Y, *et al.*: **sanger-tol/treeval v1. 0.0–Ancient Atlantis**. 2023.
Publisher Full Text

18. Kerpedjiev P, Abdennur N, Lekschas F, *et al.*: **HiGlass: web-based visual exploration and analysis of genome interaction maps.** *Genome Biol.* 2018; **19**(1): 125.
PubMed Abstract | Publisher Full Text | Free Full Text

19. Diesh C, Stevens GJ, Xie P, *et al.*: **JBrowse 2: a modular genome browser with views of synteny and structural variation.** *Genome Biol.* 2023; **24**(1): 74.
PubMed Abstract | Publisher Full Text | Free Full Text

20. Howe K, Chow W, Collins J, *et al.*: **Significantly improving the quality of genome assemblies through curation.** *GigaScience.* 2021; **10**(1): giaa153.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Di Tommaso P, Chatzou M, Floden EW, *et al.*: **Nextflow enables reproducible computational workflows.** *Nat Biotechnol.* 2017; **35**(4): 316–319.
PubMed Abstract | Publisher Full Text

22. Surana P, Muffato M, Qi G: **sanger-tol/readmapping: sanger-tol/readmapping v1.1.0 - Hebridean Black (1.1.0).** *Zenodo.* 2023.
Publisher Full Text

23. Muffato M, Sadasivan Baby C, Surana P: **sanger-tol/genomenote (v1.0.dev).** *Zenodo.* 2023.
Publisher Full Text

24. Muffato M, Butt Z, Challis R: **sanger-tol/blobtoolkit: v0. 3.0–Poliwag.** 2024.

25. Vasimuddin M, Misra S, Li H, *et al.*: **Efficient architecture-aware acceleration of BWA-MEM for multicore systems.** In: *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS).* IEEE, 2019.
Publisher Full Text

26. Danecek P, Bonfield JK, Liddle J, *et al.*: **Twelve years of SAMtools and BCFtools.** *GigaScience.* 2021; **10**(2): giab008.
PubMed Abstract | Publisher Full Text | Free Full Text

27. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics.* 2010; **26**(6): 841–842.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Abdennur N, Mirny LA: **Cooler: scalable storage for Hi-C data and other genomically labeled arrays.** *Bioinformatics.* 2020; **36**(1): 311–316.
PubMed Abstract | Publisher Full Text | Free Full Text

29. Sayers EW, Cavanaugh M, Clark K, *et al.*: **GenBank 2024 update.** *Nucleic Acids Res.* 2024; **52**(D1): D134–D137.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Manni M, Berkeley MR, Seppey M, *et al.*: **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): 4647–4654.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Challis R, Richards E, Rajan J, *et al.*: **BlobToolKit - interactive quality assessment of genome assemblies.** *G3 (Bethesda).* 2020; **10**(4): 1361–1374.
PubMed Abstract | Publisher Full Text | Free Full Text

32. Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
PubMed Abstract | Publisher Full Text | Free Full Text

33. Challis R, Kumar S, Sotero-Caio C, *et al.*: **Genomes on a Tree (GoaT): a versatile, scalable search engine for genomic and sequencing project metadata across the eukaryotic Tree of Life [version 1; peer review: 2 approved].** *Wellcome Open Res.* 2023; **8**: 24.
PubMed Abstract | Publisher Full Text | Free Full Text

34. Bateman A, Martin MJ, Orchard S, *et al.*: **UniProt: the universal protein knowledgebase in 2023.** *Nucleic Acids Res.* 2023; **51**(D1): D523–D531.
PubMed Abstract | Publisher Full Text | Free Full Text

35. Buchfink B, Reuter K, Drost HG: **Sensitive protein alignments at tree-of-life scale using DIAMOND.** *Nat Methods.* 2021; **18**(4): 366–8.
PubMed Abstract | Publisher Full Text | Free Full Text

36. Altschul SF, Gish W, Miller W, *et al.*: **Basic Local Alignment Search Tool.** *J Mol Biol.* 1990; **215**(3): 403–410.
PubMed Abstract | Publisher Full Text

37. Ewels PA, Peltzer A, Fillinger S, *et al.*: **The nf-core framework for community-curated bioinformatics pipelines.** *Nat Biotechnol.* 2020; **38**(3): 276–278.
PubMed Abstract | Publisher Full Text

38. Ewels P, Magnusson M, Lundin S, *et al.*: **MultiQC: summarize analysis results for multiple tools and samples in a single report.** *Bioinformatics.* 2016; **32**(19): 3047–3048.
PubMed Abstract | Publisher Full Text | Free Full Text

39. Grüning B, Dale R, Sjödin A, *et al.*: **Bioconda: sustainable and comprehensive software distribution for the life sciences.** *Nat Methods.* 2018; **15**(7): 475–476.
PubMed Abstract | Publisher Full Text | Free Full Text

40. da Veiga Leprevost F, Grüning BA, Alves Aflitos S, *et al.*: **BioContainers: an open-source and community-driven framework for software standardization.** *Bioinformatics.* 2017; **33**(16): 2580–2582.
PubMed Abstract | Publisher Full Text | Free Full Text

41. Merkel D: **Docker: lightweight Linux containers for consistent development and deployment.** *Linux J.* 2014; **2014**(239): 2, [Accessed 2 April 2024].
Reference Source

42. Kurtzer GM, Sochat V, Bauer MW: **Singularity: scientific containers for mobility of compute.** *PLoS One.* 2017; **12**(5): e0177459.
PubMed Abstract | Publisher Full Text | Free Full Text

43. Thibaud-Nissen F, Souvorov A, Murphy T, *et al.*: **Eukaryotic genome annotation pipeline.** *The NCBI handbook.* 2013; **2**.
Reference Source

44. Flynn JM, Hubley R, Goubert C, *et al.*: **RepeatModeler2 for automated genomic discovery of transposable element families.** *Proc Natl Acad Sci U S A.* 2020; **117**(17): 9451–7.
PubMed Abstract | Publisher Full Text | Free Full Text

45. Smit A, Hubley R, Green P: **RepeatMasker Open-4.0. 2013-2015.** 2013.

46. Storer J, Hubley R, Rosen J, *et al.*: **The Dfam community resource of transposable element families, sequence models, and genome annotations.** *Mob DNA.* 2021; **12**(1): 2.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔

Version 1

Reviewer Report 06 June 2025

https://doi.org/10.21956/wellcomeopenres.26272.r123629

✔ **Yanbo Sun**
Yunnan University, Kunming, China

The study provides a technically robust and well-motivated chromosome-level genome assembly for P. corroboree. To enhance reproducibility and accessibility:
Table Enhancement: Include SRA accessions for all sequencing datasets (e.g., Hi-C: SAXXX; RNA-Seq: SAXXX).
Supplemental Material: Share Snakemake/Nextflow workflow scripts via GitHub or Zenodo.
Annotation Methods: Briefly describe evidence sources (e.g., ab initio prediction, transcript alignment).
Synteny or collinearity studies with closer relatives (e.g., Pseudophryne pengilleyi) are needed

Is the rationale for creating the dataset(s) clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of methods and materials provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Evolutionary genomics of amphibians and reptiles

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.