Morrill, Benson H., H. Patrick MacKnight, Ari R. Flagle, and Stacy Pirro. 2022. "The Complete Genome Sequence of the *Simalia Boeleni*, the Boelen's Python." *Biodiversity Genomes*, September.

GENOME SEQUENCING

The Complete Genome Sequence of the *Simalia boeleni*, the Boelen's Python

Benson H. Morrill[1], H. Patrick MacKnight[2], Ari R. Flagle[3], Stacy Pirro[4]

[1] Rare Genetics Inc., [2] MacKnight Stables, [3] ProjectBlackPython.org, [4] Iridian Genomes

Keywords: genome, reptile, snake, squamata

https://doi.org/10.56179/001c.38128

Biodiversity Genomes

The Boelen's Python (*Simalia boeleni*) is a python endemic to the mountains of New Guinea. We present the whole genome sequence of this species. Illumina sequencing was performed on a genetic sample from a single individual. The reads were assembled using a de novo method followed by a series of references from related species for finishing. The raw and assembled data is publicly available via Genbank: Sequence Read Archive (SRR19167501) and assembled genome (JANKYG000000000).

Introduction

The Boelen's Python is found in Indonesia (Western New Guinea in the Paniai Lakes region) and Papua New Guinea (the provinces of Eastern Highlands, Central and Morobe, and Goodenough Island). receives the highest legal protection possible in Papua New Guinea. It is currently listed on CITES Appendix II.

In adults, the dorsal pattern is dark bluish-black or purplish-black, and the anterior part of the underside is white to pale yellow. The white extends up the flanks as a series of streaks. The upper and lower lips are also patterned with pale or whitish labial scales. The black portions are commonly iridescent with an oil-slick-like sheen. The body is stocky and the head large. Neonates are predominantly red upon emerging from the egg. Gradual black pigmentation presents itself as the neonate grows and sheds. Ontogenetic color charge begins as the juvenile snake approaches 1 m (39 in) in length, usually around 2 years of age. Adults reach an average length of 3 m (9.8 ft), including the tail.

Methods

The genetic sample was obtained from a naturally shed skin from single long term wild caught individual..

DNA extraction was performed using the Qiagen DNAeasy genomic extraction kit using the standard process. A paired-end sequencing library was constructed using the Illumina TruSeq kit according to the manufacturer's instructions. The library was sequenced on an Illumina Hi-Seq platform in paired-end, 2 × 150bp format. The resulting fastq files were trimmed of adapter/primer sequence and low-quality regions with Trimmomatic v0.33 (Bolger, Lohse, and Usadel 2014). The trimmed sequence was assembled by SPAdes v2.5 (Bankevich et al. 2012) followed by a finishing step using Zanfona (Kieras, O'Neill, and Pirro 2021).

Simalia boeleni, female, ~8'. Illugua West Papua, Indonesia.

Image by Ari Flagle.

Results

The genome assembly yielded a total sequence length of 1,471,293,121 bp over 81,708 scaffolds > 500 bp.

……

Data availability

Raw and assembled data is publicly available via GenBank:

RAW GENOME DATA

https://trace.ncbi.nlm.nih.gov/Traces/sra/?run=SRR19167501

ASSEMBLED GENOME

https://www.ncbi.nlm.nih.gov/nucleotide/JANKYG000000000

Funding

Funding was provided by Iridian Genomes, grant# IRGEN_RG_2021-1345 Genomic Studies of Eukaryotic Taxa.

Special thank you to Todd Goodman and Josh Hansen for generously supporting projectblackpython.com

“Boelen’s python conservation and research”
http://www.projectblackpython.org
817-449-3096

Submitted: September 01, 2022 EDT, Accepted: September 06, 2022 EDT

REFERENCES

Bankevich, Anton, Sergey Nurk, Dmitry Antipov, Alexey A. Gurevich, Mikhail Dvorkin, Alexander S. Kulikov, Valery M. Lesin, et al. 2012. “SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing.” *Journal of Computational Biology* 19 (5): 455–77. https://doi.org/10.1089/cmb.2012.0021.

Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. https://doi.org/10.1093/bioinformatics/btu170.

Kieras, M., K. O’Neill, and S. Pirro. 2021. *Zanfona, a genome assembly finishing tool for paired-end Illumina reads*. https://github.com/zanfona734/zanfona.