Hains, Taylor, Stacy Pirro, John Bates, and Shannon Hackett. 2022. "The Complete Genome Sequence of *Cyanopsitta Spixii*, the Spix's Macaw." *Biodiversity Genomes*, August.

GENOME SEQUENCING

The Complete Genome Sequence of *Cyanopsitta spixii,* the Spix's Macaw

Taylor Hains[1], Stacy Pirro, John Bates[1], Shannon Hackett[1]

[1] Negaunee Integrative Research Center, Field Museum of Natural History

Keywords: genome, parrot, endangered, conservation

https://doi.org/10.56179/001c.37839

Biodiversity Genomes

The Spix's Macaw (*Cyanopsitta spixii*) is a critically endangered parrot that was once endemic to Brazil. We present the whole genome sequence of this species. Illumina sequencing was performed on a genetic sample from a single captive 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 (SRR15037507) and Assembly (GCA_024336845).

Introduction

The Spix's Macaw is a medium-sized parrot with grayish-blue plumage. Males are slightly larger than females but otherwise identical in appearance.

The species was endemic to riparian Caraibeira (*Tabebuia aurea*) gallery woodlands in the drainage basin of the Rio São Francisco within the Caatinga of northeastern Brazil. It had a very restricted natural habitat along several small rivers. It fed primarily on seeds of two species of Euphorbiaceae, the dominant vegetation of the Caatinga (Barros et al. 2012). Due to deforestation in its limited range and specialized habitat, the bird was rare in the wild throughout the twentieth century. The IUCN conducted several years of survey attempting to locate living wild populations, but it was officially declared extinct in the wild in 2019.

Methods

The genetic sample was provided from San Diego Zoo Wildlife Alliance's Wildlife Biodiversity Bank by the Conservation Genetics Team.

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).

Results

The genome assembly yielded a total sequence length of 1,092,480,392 bp over 53,615 scaffolds.

Data availability

Raw and assembled data is publicly available via GenBank:

RAW GENOME DATA

https://trace.ncbi.nlm.nih.gov/Traces/sra/?run=SRR15037507

ASSEMBLED GENOME

https://www.ncbi.nlm.nih.gov/assembly/GCA_024336845

Acknowledgements

The authors are grateful to the San Diego Zoo Wildlife Alliance's Conservation Genetics team for supplying the genetic sample.

Funding

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

Submitted: August 20, 2022 EDT, Accepted: August 20, 2022 EDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-SA-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-sa/4.0 and legal code at https://creativecommons.org/licenses/by-sa/4.0/legalcode for more information.

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.

Barros, Y., Y. Soye, C. Miyaki, R. Watson, L. Crosta, and C. Lugarini. 2012. *Plano de ação nacional para a conservação da ararinha-azul: Cyanopsitta spixii*. Brasília: Chico Mendes Institute of Biodiversity Conservation, ICMBio.

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.