



Published in final edited form as:

Biodivers Genomes. 2022 ; 2022: . doi:10.56179/001c.66227.

The Complete Genome Sequence of *Conuropsis carolinensis*, the Carolina Parakeet

Taylor Hains^{1,2}, Stacy Pirro³, John M Bates¹, Shannon J Hackett¹

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

²Committee on Evolutionary Biology, University of Chicago,

³Biodiversity, Iridian Genomes

Abstract

The Carolina Parakeet (*Conuropsis carolinensis*) is an extinct species of parrot that was native to the eastern, midwest, and plains regions of the United States. 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 ([SRR21023482](https://www.ncbi.nlm.nih.gov/sra/SRR21023482)) and assembled genome ([JAOPYI000000000](https://www.ncbi.nlm.nih.gov/genbank/JAOPYI000000000)).

Keywords

genome; bird; parrot; extinct

Introduction

The Carolina Parakeet was a small green parakeet neotropical parrot with a bright yellow head, reddish orange face and pale beak. It lived in old-growth forests along rivers and in swamps in the eastern, midwest, and plains regions of the United States. Though formerly prevalent within its range, the bird had become rare by the middle of the 19th century. The last known specimen perished in captivity at the Cincinnati Zoo in 1918, and the species was declared extinct in 1939 (Burgio, Carlson, and Tingley 2017).

Methods

The genetic sample was provided by the Field Museum of Natural History, catalogue number FMNH 124024.

DNA extraction was performed using the Qiagen DNeasy 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 and Data Availability

Raw and assembled data is publicly available via GenBank:

Raw genome data

<https://trace.ncbi.nlm.nih.gov/Traces/sra/?run=SRR21023482>

Assembled genome

<https://www.ncbi.nlm.nih.gov/nucore/JAOBYI000000000>

Acknowledgements

The authors are grateful to the Field Museum for supplying the tissue sample.

Funding

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

REFERENCES

- Bankevich Anton, Nurk Sergey, Antipov Dmitry, Gurevich Alexey A., Dvorkin Mikhail, Kulikov Alexander S., Lesin Valery M., et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology* 19 (5): 455–77. 10.1089/cmb.2012.0021. [PubMed: 22506599]
- Bolger Anthony M., Lohse Marc, and Usadel Bjoern. 2014. "Trimmomatic: A Flexible Trimmer for Illumina Sequence Data." *Bioinformatics* 30 (15): 2114–20. 10.1093/bioinformatics/btu170. [PubMed: 24695404]
- Burgio Kevin R., Carlson Colin J., and Tingley Morgan W.. 2017. "Lazarus Ecology: Recovering the Distribution and Migratory Patterns of the Extinct Carolina Parakeet." *Ecology and Evolution* 7 (14): 5467–75. 10.1002/ece3.3135. [PubMed: 28770082]
- Kieras M, O'Neill K, and Pirro S. 2021. "Zanfona, a Genome Assembly Finishing Tool for Paired-End Illumina Reads." 2021. <https://github.com/zanfona734/zanfona>.