White, Harold B., and Stacy Pirro. 2024. "The Complete Genome Sequences of Two Species of Seventeen-Year Cicadas: *Magicicada Septendecim* and *Magicicada Septendecula*." *Biodiversity Genomes*, May. https://doi.org/10.56179/001c.117365.

GENOME SEQUENCING

The complete genome sequences of two species of seventeen-year cicadas: *Magicicada septendecim* and *Magicicada septendecula*

Harold B. White[1a], Stacy Pirro[2]

[1] Department of Chemistry and Biochemistry, University of Delaware, [2] Biodiversity, Iridian Genomes

https://doi.org/10.56179/001c.117365

Biodiversity Genomes

The genus *Magicicada* (Hemiptera: Cicadidae) includes the periodical cicadas of Eastern North America. Spending the majority of their long lives underground, the adult cicadas emerge every 13 or 17 years to spend 4-6 weeks as adult to mate. We present the whole genome sequences of two species of 17-year cicadas, *Magicicada septendecim* and *Magicicada septendecula*. The reads were assembled by a de novo method followed by alignments to related species. Annotation was performed by GeneMark-ES. The raw and assembled data is available via NCBI Short Read Archive and Assembly databases.

Introduction

Periodical cicadas in North America spend 13 or 17 years in the larval stage underground, and emerge in very large numbers for 4–6 weeks to mate and lay eggs. This strategy, known as "predator satiation" is intended to ensure that after all predators have eaten as much as possible, most cicadas will survive. (Williams and Simon 1995). The emergence occurring in prime-numbered years is thought to be a mechanism to avoid competition between species for egg-laying sites and accidental cross-species mating as the emergence of the 13- and 17-year cicadas would only coincide once every 221 years (Tanaka, Yoshimura, Simon, et al. 2009). The length of time spent in the larval stage is thought to be dependent on a single gene, although this has not yet been demonstrated at the genomic level (Cox and Carlton 1991).

Complete genome sequences for these two species will assist with studies on taxonomy, longevity, and the timing of long-term larval development.

Methods

Wild caught specimens of *Magicicada septendecim* and *Magicicada septendecula* from a small premature emergence of Brood X (2017) collected in Newark, Delaware, USA were used in this study. DNA extraction was performed using the Qiagen DNAeasy genomic extraction kit for tissue, using the standard process. A paired-end sequencing library was constructed

---

[a] Harold B. White is now retired from Department of Chemistry and Biochemistry, University of Delaware.

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 RagTag v1.0.0 (Alonge 2020) to make additional contig joins based on conserved regions in related insect species: *Rhopalosiphum maidis* (GCA_003676215), *Euschistus heros* (GCA_003667255), and *Aphis glycines* (GCA_009928515). Default parameters were used for all assembly steps.

Annotation was performed using GeneMark-ES v2.0 (Lomsadze et al. 2005). Annotation was performed fully de novo without a curated training set and using default parameters.

Results

The genome assembly for *Magicicada septendecim* yielded a total sequence length of 1,579,033,894 with an N50 value of 983 kb and 27,124 gene models.

The genome assembly of *Magicicada septendecula* yielded 1,585,977,997 with an N50 value of 281 kb and 28,651 gene models.

Data availability

Raw reads and assembled genomes available from Genbank:

Species	Raw reads	Assembled genome
Magicicada septendecim	SRR6782667	GCA_011326945
Magicicada septendecula	SRR6792649	GCA_011763675

Author information

Harold B. White is now retired from Department of Chemistry and Biochemistry, University of Delaware.

Funding

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

Submitted: May 02, 2024 EDT, Accepted: May 06, 2024 EDT

REFERENCES

Alonge, M. 2020. *Ragtag: Reference-Guided Genome Assembly Correction and Scaffolding*. GitHub archive.

Bankevich, A., S. Nurk, D. Antipov, and et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *J Comput Biol* 19 (5): 455–77. https://doi.org/10.1089/cmb.2012.0021.

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

Cox, R.T., and C.E. Carlton. 1991. "Evidence of Genetic Dominance of the 13-Year Life Cycle in Periodical Cicadas (Homoptera: Cicadidae: Magicicada Spp.)." *Am Midl Nat* 125 (1): 63–74. https://doi.org/10.2307/2426370.

Lomsadze, A., V. Ter-Hovhannisyan, Y.O. Chernoff, and et al. 2005. "Gene Identification in Novel Eukaryotic Genomes by Self-Training Algorithm." *Nucleic Acids Res* 33 (20): 6494–6506. https://doi.org/10.1093/nar/gki937.

Tanaka, Y., J. Yoshimura, C. Simon, et al. 2009. "Allee Effect in the Selection for Prime-Numbered Cycles in Periodical Cicadas." *Proc Natl Acad Sci U S A* 106 (22): 8975–79. https://doi.org/10.1073/pnas.0900215106.

Williams, K. S., and C. Simon. 1995. "The Ecology, Behavior, and Evolution of Periodical Cicadas." *Annu Rev Entomol* 40 (1): 269–95. https://doi.org/10.1146/annurev.en.40.010195.001413.