

GENOME SEQUENCING

Whole-Genome sequencing of the schizomid arachnid *Rowlandius potiguar* Santos, Ferreira and Buzzato, 2013 from the caves of the Brazilian semiarid

Origilene Bezerra Dantas^{1,2}, Matheus Arthur Lúcio da Rocha³, Rodrigo Lopes Ferreira¹, Diego de Medeiros Bento⁴, Sergio Maia Queiroz Lima¹, Stacy Pirro⁵, Waldir M. Berbel-Filho⁶

¹ Universidade Federal de Lavras (UFLA), Centro de Estudos em Biologia Subterrânea (CEBS), Lavras, MG, Brazil, ² Programa de Pós-Graduação em Ecologia Aplicada, Universidade Federal de Lavras (UFLA), Lavras, MG, Brazil, ³ Universidade Federal do Rio Grande do Norte (UFRN), Centro de Biociências (CB), Departamento de Botânica e Zoologia (DBZ), Laboratório de Ictiologia Sistemática e Evolutiva (LISE), Natal, RN, Brazil, ⁴ Centro Nacional de Pesquisa e Conservação de Cavernas (CECAV), Instituto Chico Mendes de conservação da biodiversidade, Natal, RN, Brazil, ⁵ Iridian Genomes, ⁶ Department of Biology, University of West Florida, Pensacola, USA

<https://doi.org/10.56179/001c.164269>

Biodiversity Genomes

Short-tailed whip scorpions (Arachnida: Schizomida) are a poorly studied group of arachnids, particularly in terms of genomic resources, despite their ecological and evolutionary relevance in subterranean environments. In this study, we present the first draft genome assembly of *Rowlandius potiguar*, a cave-associated species endemic to karst systems of northeastern Brazil. Genomic DNA was sequenced using Illumina paired-end technology (2 × 150 bp), and reads were assembled using a de novo assembly pipeline. The resulting genome assembly has a total length of 1.8 Gb, distributed across 1,505,104 scaffolds and 1,510,944 contigs, indicating a highly fragmented assembly. The scaffold N50 and contig N50 values were both 3.2 kb, and the GC content was 32.5%. Genome completeness was assessed using BUSCO with Arthropoda and Metazoa datasets. The Arthropoda analysis recovered 20.5% complete BUSCOs (19.1% single-copy and 1.5% duplicated), 53.0% fragmented, and 26.5% missing genes. Similarly, the Metazoa dataset identified 24.7% complete BUSCOs (23.4% single-copy and 1.4% duplicated), with 54.7% fragmented and 20.5% missing genes. The low BUSCO completeness and high proportion of fragmented genes reflect the highly fragmented nature of the assembly, likely resulting from the use of short-read sequencing technologies and the presence of repetitive genomic regions. Despite these limitations, this genome represents an important first genomic resource for Schizomida and provides a foundation for future studies on genome evolution, adaptation to subterranean environments, and the genetic basis of troglomorphic traits in arachnids.

Introduction

Schizomida Petrunkevitch, 1945 arachnids are commonly known as short-tailed whip scorpions, and comprise a small-bodied group (3-5 mm) with cryptic habits and a strong association with humid environments, frequently found in leaf litter and caves (Reddell and Cokendolpher 1995; Santos et al. 2013). Within this group, the genus *Rowlandius* Reddell and Cokendolpher 1995 includes more than 63 described species (*World Schizomida Catalog* 2022), with species distributed from the Caribbean islands to northeastern Brazil, occurring in tropical forests and semiarid environments (Reddell and Cokendolpher 1995; Armas 2002; Teruel et al. 2012; Santos et al. 2013;

Giupponi et al. 2016). Despite their broad distribution across tropical regions, genomic resources for this group remain virtually nonexistent, limiting our understanding of their evolution, genomic architecture, and adaptive processes.

In Brazil, *Rowlandius potiguar* Santos, Ferreira & Buzzato, 2013 stands out as the first record of schizomid in caves in the country (Santos et al. 2013). The species was originally described from caves in the Jandaíra Formation, in northeastern Brazil, and currently is known in 29 caves in the semiarid Caatinga (Santos et al. 2013; Bento et al. 2021; Dantas et al. 2026). In the same geological formation, *R. potiguar* co-occurs with another congeneric species, *R. tybykyra* Silva, Dantas, Carvalho, Ferreira, Bento & Lima, 2026; however, the two species have never been recorded in syntopy, suggesting spatial segregation or niche differentiation (Dantas et al. 2026). Its distribution in the semiarid might be shaped by paleoclimatic changes throughout the Quaternary (Conti and Furlan 1996) and suggests a dynamic history of habitat expansion and contraction, making this species a promising model for investigating eco-evolutionary patterns in subterranean environments.

The advent of next-generation sequencing technologies has enabled the generation of draft genomes for non-model organisms, greatly expanding the ability to investigate genome structure, gene content, and evolutionary patterns across diverse taxa (Ekblom and Galindo 2011). For neglected taxa, such as schizomids, these data are essential to refine phylogenetic hypotheses and support broader comparative analyses within Arachnida. In this study, we present the first draft genome of *R. potiguar*. We characterize the main assembly metrics and assess genome completeness using BUSCO analyses based on Arthropoda and Metazoa datasets. By providing this resource, we contribute to advancing the understanding of Schizomida genomic biology and to increasing the representation of Neotropical subterranean invertebrates in genomic studies.

Material and methods

Specimens of *R. potiguar* used in this study were collected at its type locality (Santos et al. 2013), Carrapateira Cave (05°33'38.4"S, 37°39'50.2"W), inserted in limestones of the Jandaíra Formation, in the municipality of Felipe Guerra, Rio Grande do Norte State, northeastern Brazil ([Fig. 1](#)).

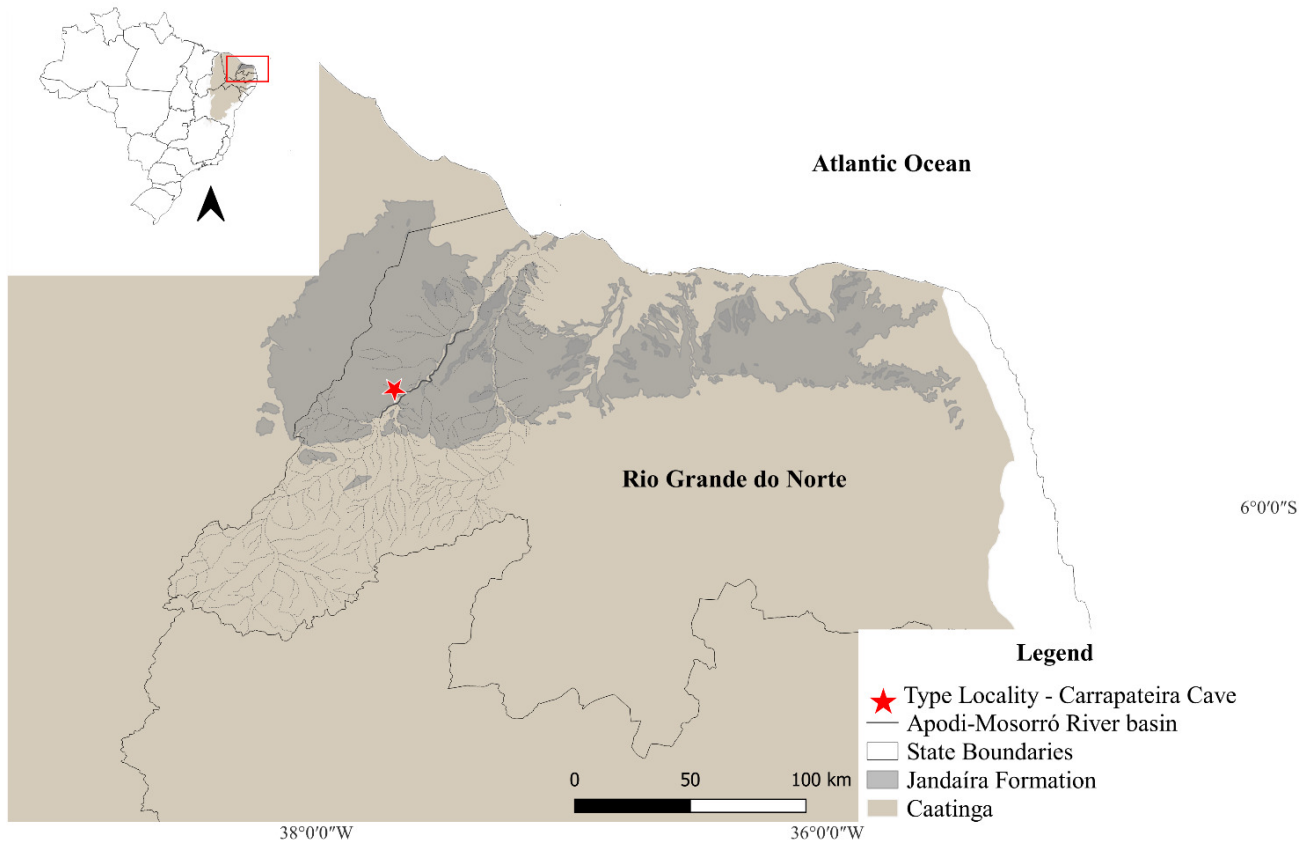


Figure 1. Map of the type-locality of *Rowlandius potiguar* in the Carrapateira Cave in the Jandaíra karstic Formation (in gray), Rio Grande do Norte State, northeastern Brazil.

Individuals were manually collected using fine brushes through active visual searching, primarily in humid microhabitats within the cave. After collection, specimens were immediately fixed in absolute ethanol (P.A.) and stored in thermal containers during transport from the field to the laboratory, where they were subsequently kept under refrigeration. Additional specimens from the same sampling event were deposited in the Coleção de Invertebrados Subterrâneos de Lavras (ISLA) under accession numbers ISLA-118278 and ISLA-118291. Sampling was conducted under permit issued by SISBIO/ICMBio (No. 54334-11, 2023). Access to Brazilian genetic heritage was registered in the Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado (SISGEN) under registration number A06F341.

DNA was extracted using the Qiagen DNeasy genomic extraction kit, following the standard protocol. A paired-end sequencing library was then prepared using the Illumina TruSeq kit, according to the manufacturer's instructions. Sequencing was performed on an Illumina HiSeq platform using a paired-end format of 2×150 bp. The resulting FASTQ files were processed to remove adapter/primer sequences as well as low-quality regions using Trimmomatic v0.33 (Bolger et al. 2014). The cleaned reads were assembled using SPAdes v2.5 (Bankevich et al. 2012), followed by a final

assembly step using Zanfona (Kieras et al. 2021). Genome assembly completeness was assessed using BUSCO v5 (Simão et al. 2015) and CEGMA 2.4 (Parra et al. 2007).

Results

Genome assembly statistics and completeness

The draft genome assembly of *Rowlandius potiguar* resulted in a total length of approximately 1.8 Gb, distributed across 1,505,104 scaffolds and 1,510,944 contigs (Table 1). The assembly is highly fragmented, as indicated by the low scaffold and contig N50 values (3.2 kb) and the exceptionally high number of scaffolds. The L50 values (158,314 scaffolds and 160,202 contigs) further highlight the low contiguity of the assembly. The GC content was estimated at 32.5%, which is within the range reported for other arachnids (Su et al. 2025).

Table 1. Genome assembly statistics for *Rowlandius potiguar*.

Parameter	Value
Assembly accession	GCA_048605005.1
Number of scaffolds	1,505,104
Number of contigs	1,510,944
Genome size (bp)	1,800,000,000
Scaffold N50 (bp)	3,200
Scaffold L50	158,314
Contig N50 (bp)	3,200
Contig L50	160,202
GC content (%)	32.5
Genome coverage	60×
Assembly level	Scaffold

Genome completeness assessed using BUSCO revealed low recovery of conserved genes (Table 2). Using the Arthropoda dataset, only 20.5% of BUSCO genes were identified as complete (19.1% single-copy and 1.5% duplicated), while a high proportion was classified as fragmented (53.0%) or missing (26.5%). Similarly, analysis with the Metazoa dataset recovered 24.7% complete BUSCOs (23.4% single-copy and 1.4% duplicated), with 54.7% fragmented and 20.5% missing genes.

Table 2. BUSCO completeness assessment for *Rowlandius potiguar*.

Category	Arthropoda (%)	Arthropoda (n)	Metazoa (%)	Metazoa (n)
Complete (C)	20.5%	208	24.7%	236
Single-copy (S)	19.1%	193	23.4%	223
Duplicated (D)	1.5%	15	1.4%	13
Fragmented (F)	53.0%	537	54.7%	522
Missing (M)	26.5%	268	20.5%	196
Total BUSCO groups	—	1,013	—	954

The predominance of fragmented BUSCOs, together with the low proportion of complete genes, reflects the highly fragmented nature of the assembly. The extremely low N50 values and the large number of scaffolds indicate limited genome contiguity, which likely compromises gene prediction accuracy and contributes to the observed patterns of incomplete BUSCO recovery (Simão et al. 2015). Additionally, approximately 25% of complete BUSCO genes contained internal stop codons, suggesting gene prediction artifacts and/or fragmentation of coding regions (Simão et al. 2015). Such characteristics are typical of draft genome assemblies generated using short-read sequencing technologies, particularly in taxa with complex genomic architecture or high repeat content (Ekblom and Wolf 2014). Despite its large genome size (~1.8 Gb), the assembly shows limited structural continuity, which may hinder downstream analyses such as gene annotation and structural genomic studies.

Nevertheless, the genome of *R. potiguar* represents an important genomic resource for a poorly studied subterranean arachnid lineage. Given its occurrence in cave environments, this dataset provides a foundation for future studies on genome evolution, adaptation to subterranean habitats, and the genetic basis of troglomorphic traits. Improvements through long-read sequencing or hybrid assembly approaches are expected to substantially enhance genome continuity and completeness.

Financing

Funding for sequencing was provided by Iridian Genomes (grant IRGEN_RG_2021-1345). LSM and OBD were supported by CAPES through PhD scholarships (LSM: #88887.483610/2020-00; OBD #88887.954829/2024-00). SML, MP, and RLF were supported by CNPq (SML #312066/2021-0, MP #310688/2019-1, RLF #302925/2022-8). This study was partially supported by the Speleological Compensation Commitment Agreements (TCCE 01/2018 and 01/2022), established between the Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio) and Vale S.A. Financial resources were managed by the Instituto Brasileiro de Desenvolvimento e Sustentabilidade (IABS).

Submitted: June 29, 2026 EDT. Accepted: June 29, 2026 EDT. Published: June 29, 2026 EDT.

REFERENCES

- Armas, L. D. 2002. "Nuevas especies de *Rowlandius* Reddell & Cokendolpher, 1995 (Schizomida: Hubbardiidae) de Cuba." *Revista Ibérica de Aracnología* 6: 149–67.
- Bankevich, A., S. Nurk, D. Antipov, 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>.
- Bento, D. M., M. Souza-Silva, A. Vasconcellos, C. B. Bellini, X. Prous, and R. Ferreira. 2021. "Subterranean 'Oasis' in the Brazilian Semiarid Region: Neglected Sources of Biodiversity." *Biodiversity and Conservation* 30 (13): 3837–57. <https://doi.org/10.1007/s10531-021-02277-6>.
- 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>.
- Conti, J. B., and S. A. Furlan. 1996. *Geoecologia: o clima, os solos e a biota*. Geografia do Brasil.
- Dantas, O. B., I. S. S. Silva, L. S. Carvalho, R. L. Ferreira, D. D. M. Bento, and S. M. Q. Lima. 2026. "Unravelling the Hidden Diversity of Schizomids (Arachnida: Schizomida: Hubbardiidae) in Caves of the Brazilian Semiarid." *Zoological Journal of the Linnean Society* 206 (3): zlag010. <https://doi.org/10.1093/zoolinnean/zlag010>.
- Ekblom, R., and J. Galindo. 2011. "Applications of next Generation Sequencing in Molecular Ecology of Non-Model Organisms." *Heredity* 107 (1): 1–15. <https://doi.org/10.1038/hdy.2010.152>.
- Ekblom, R., and J. B. Wolf. 2014. "A Field Guide to Whole-Genome Sequencing, Assembly and Annotation." *Evolutionary Applications* 7 (9): 1026–42. <https://doi.org/10.1111/eva.12178>.
- Giupponi, A. P., G. S. de Miranda, and O. M. Villarreal. 2016. "*Rowlandius dumitrescoae* Species Group: New Diagnosis, Key and Description of New Cave-Dwelling Species from Brazil (Schizomida, Hubbardiidae)." *ZooKeys* 632: 13. <https://doi.org/10.3897/zookeys.632.9337>.
- Kieras, M., K. O'Neill, and S. Pirro. 2021. *Zanfona, a Genome Assembly Finishing Tool for Paired-End Illumina Reads*.
- Parra, G., K. Bradnam, and I. Korf. 2007. "CEGMA: A Pipeline to Accurately Annotate Core Genes in Eukaryotic Genomes." *Bioinformatics* 23 (9): 1061–67. <https://doi.org/10.1093/bioinformatics/btm071>.
- Reddell, J. R., and J. C. Cokendolpher. 1995. *Catalogue, Bibliography, and Generic Revision of the Order Schizomida (Arachnida)*. Texas Memorial Museum, College of Natural Sciences, University of Texas at Austin.
- Santos, A. J., R. L. Ferreira, and B. A. Buzatto. 2013. "Two New Cave-Dwelling Species of the Short-Tailed Whipscorpion Genus *Rowlandius* (Arachnida: Schizomida: Hubbardiidae) from Northeastern Brazil, with Comments on Male Dimorphism." *PLoS One* 8 (5): e63616. <https://doi.org/10.1371/journal.pone.0063616>.
- Simão, F. A., R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, and E. M. Zdobnov. 2015. "BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs." *Bioinformatics* 31 (19): 3210–12. <https://doi.org/10.1093/bioinformatics/btv351>.
- Su, Y. K., M. H. Xiu, H. Y. Yang, and C. M. Shi. 2025. "A Chromosome-Level Genome Assembly for the Desert Scorpion *Mesobuthus przewalskii* from Asian Drylands." *Journal of Heredity* 116 (4): 532–39. <https://doi.org/10.1093/jhered/esae059>.
- Teruel, R., L. F. de Armas, and T. M. Rodríguez. 2012. "De cuatro nuevos *Rowlandius* Reddell & Cokendolpher 1995 (schizomida: hubbardiidae)." *Revista Ibérica de Aracnología* 21: 97–112.

World Schizomida Catalog. 2022. Natural History Museum Bern. <http://wac.nmbe.ch>.