

Whole-genome sequencing of an axenic *Achromobacter* sp. strain PD1 reveals genomic potential for cyanophycin production

Yingfen Zhao,¹ Zhen Jia,¹ Dan Li,^{1,2} Keith Tyo,³ George F. Wells¹

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT This study reports the whole-genome sequence of *Achromobacter* sp. strain PD1, isolated from a wastewater treatment bioreactor. Using Illumina and Nanopore sequencing, a complete genome of 6,994,232 bp was obtained. This genome contains genes predicted to encode a complete pathway for cyanophycin synthesis, a candidate for nitrogen recovery from wastewater.

KEYWORDS *Achromobacter* sp., wastewater, nitrogen recovery, biopolymer, cyanophycin

Cyanophycin, a biopolymer composed of the amino acids arginine and aspartate, offers a valuable approach to nitrogen (N) recovery from N-rich wastewater streams. Although many studies have investigated cyanophycin production in cyanobacteria (1–3), the isolation and identification of strains with cyanophycin-producing potential from wastewater treatment systems remains limited (4). In this work, we isolated the strain *Achromobacter* sp. PD1 from activated sludge and identified its genomic potential to produce cyanophycin via whole-genome sequencing.

The strain was isolated from a mixed-culture lab-scale sequencing batch reactor performing biological nitrogen and phosphorus removal from wastewater. Microbial biomass was withdrawn from the bioreactor and subjected to serial dilution prior to spread plating onto agar plates prepared using ATCC 450 T2 medium supplemented with 2% agar. Resulting colonies were further enriched in ATCC 450 T2 liquid medium at 37°C in an incubator at 150 rpm to test their ability to consume nitrogenous substrates. Promising candidates were further purified by repeated single colony streaking on Luria-Bertani (LB) agar plates, and the resulting isolate was propagated for DNA extraction and genomic sequencing.

Genomic DNA was extracted with the FastDNA SPIN kit for soil (MPBio, Santa Ana, CA, USA) according to the manufacturer's protocol and quantified using a Qubit fluorometer (Invitrogen, San Diego, CA, USA). Both Illumina short-read sequencing and Nanopore long-read sequencing were performed at SeqCenter (Pittsburgh, PA, USA) for the isolated strain, followed by a hybrid *de novo* assembly strategy to ensure a high-quality and comprehensive whole-genome assembly.

The Illumina sequencing library was prepared using the tagmentation- and PCR-based Illumina DNA prep kit (Illumina) with custom IDT 10 bp unique dual indices and a target insert size of 280 bp (IDT, Coralville, IA, USA). The sequencing was performed on an Illumina NovaSeq X Plus sequencer, producing 2 × 151 bp paired-end reads. Adapter trimming and quality control were performed with bcl-convert (5) (v4.2.4). The Nanopore sequencing library was generated using the Oxford Nanopore Technologies (ONT) Ligation Sequencing Kit (SQK-NBD114.24). The sequencing was performed on an Oxford Nanopore MinION Mk1B sequencer, followed by super-accurate base calling

Editor André O. Hudson, Rochester Institute of Technology, Rochester, New York, USA

Address correspondence to George F. Wells, george.wells@northwestern.edu.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 14 July 2025

Accepted 5 August 2025

Published 25 August 2025

Copyright © 2025 Zhao et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

TABLE 1 Summary of sequencing, assembly, and annotation results of the sequenced genome

Sequencing parameters	Illumina	Nanopore
Number of reads	7,274,764	694,980
Size (bp)	1,075,530,433	3,910,343,322
Total bp >Q30 (20) ^a	982,902,161	3,551,491,299
% bp >Q30 (20) ^a	91	91
Genome parameters	Hybrid-assembly results	
Assembly size (bp)	6,994,232	
Contig	1	
N ₅₀ (bp)	6,994,232	
G + C content (%)	65.97%	
No. of coding genes	6,454	
Total coding sequence length (bp)	6,336,711	
Average gene length (bp)	982	
Genome coverage	702x	
Number of rRNA	13	
Number of tRNA	61	
No. of ncRNAs	7	

^aTotal bp >Q20* for Nanopore sequencing results.

and adapter trimming with Guppy (6) (v6.5.7). Residual adapters that were missed during basecalling were further removed by Porechop (7) (v0.2.4). Genome assembly was performed using Flye (8) (v2.9.2, --asm-coverage 50 --genome-size 6000000) for ONT high-quality reads. The assembled genome was further polished with Pilon (9) (v1.24) using Illumina sequencing reads to correct errors introduced by low-quality Nanopore reads. The assembly quality was evaluated using QUAST (10) (v5.2.0), and genome completeness was assessed with BUSCO (v5.4.7) based on the bacteria_odb10 lineage

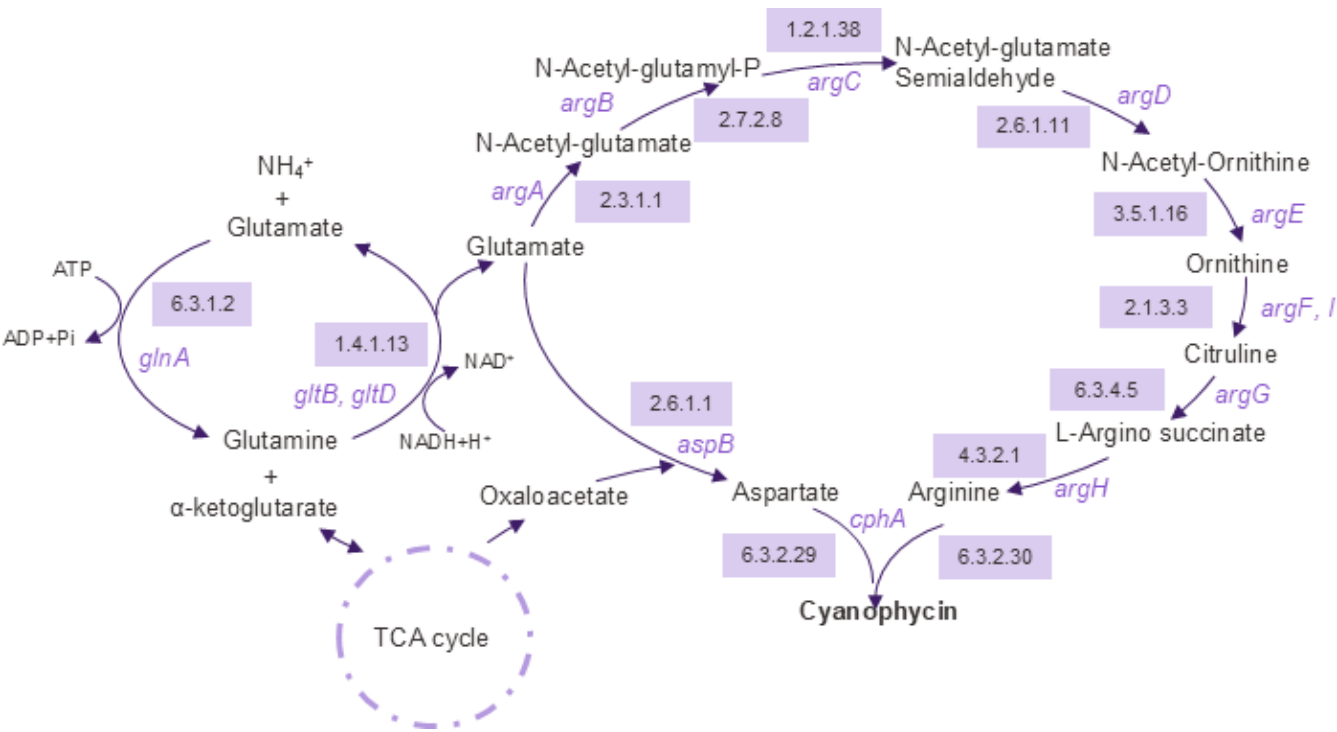


FIG 1 Cyanophycin synthesis pathway of *Achromobacter* sp. PD1 reconstructed from whole-genome sequencing results. Enzyme Commission (EC) numbers in purple boxes represent the specific enzymes catalyzing each step in the pathway. Gene names (in purple text) correspond to the encoding genes for these enzymes.

data set. Annotation of the assembled genome was performed using Bakta (11) (v1.8.1, db version 5.0). Default parameters were used for all software unless otherwise specified. The summary results of the sequencing, assembly, and annotation are presented in Table 1.

Whole-genome sequencing results were analyzed using the genome analysis tool–METABOLIC to explore the potential metabolic pathways of *Achromobacter* sp. PD1 (12). METABOLIC-G pipeline intaking individual genome sequence was used to obtain KEGG ortholog (KO) identifiers with KO IDs and the corresponding gene hits (12). The KO identifiers were subsequently mapped to metabolic pathways using the KEGG Mapper tool (13), which allowed reconstruction of the cyanophycin synthesis pathway (Fig. 1).

ACKNOWLEDGMENTS

This study was supported by the U.S. National Science Foundation under award number 2033793.

AUTHOR AFFILIATIONS

¹Department of Civil and Environmental Engineering, Northwestern University, Evanston, Illinois, USA

²School of Environment, Tsinghua University, Beijing, China

³Chemical and Biological Engineering, Northwestern University, Evanston, Illinois, USA

AUTHOR ORCID*s*

Yingfen Zhao  <http://orcid.org/0000-0002-1144-7620>

George F. Wells  <http://orcid.org/0000-0002-9100-1628>

FUNDING

Funder	Grant(s)	Author(s)
National Science Foundation	2033793	George F. Wells

AUTHOR CONTRIBUTIONS

Yingfen Zhao, Conceptualization, Methodology, Software, Writing – original draft | Zhen Jia, Methodology, Resources | Dan Li, Data curation, Methodology | Keith Tyo, Funding acquisition, Resources, Validation, Writing – review and editing | George F. Wells, Conceptualization, Funding acquisition, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The genome data of *Achromobacter* sp. strain PD1 have been deposited in NCBI GenBank under BioProject accession number [PRJNA1221536](https://doi.org/10.1007/s10811-021-02575-0). The genome accession number is [CP183349](https://doi.org/10.1007/s10811-021-02575-0). The version described in this paper is the first version. Raw sequence reads have been submitted to the NCBI Sequence Read Archive under Bioproject number [PRJNA1221325](https://doi.org/10.1007/s10811-021-02575-0). The accession numbers are [SRR32280757](https://doi.org/10.1007/s10811-021-02575-0) for Illumina sequencing reads and [SRR32280756](https://doi.org/10.1007/s10811-021-02575-0) for Nanopore sequencing reads.

REFERENCES

- Canizales S, Sliwczinka M, Russo A, Bentvelzen S, Temmink H, Verschoor AM, Wijffels RH, Janssen M. 2021. Cyanobacterial growth and cyanophycin production with urea and ammonium as nitrogen source. *J Appl Phycol* 33:3565–3577. <https://doi.org/10.1007/s10811-021-02575-0>
- Jyoti J, Khattar JIS, Gulati A, Singh DP. 2019. Optimization of conditions and partial characterization of cyanophycin synthetase from a thermophilic cyanobacterium *Chlorogloeopsis fritschii*. *Biocatal Agric Biotechnol* 17:339–346. <https://doi.org/10.1016/j.bcab.2018.12.011>
- Lu Z, Ye J, Chen Z, Xiao L, Lei L, Han B-P, Paerl HW. 2022. Cyanophycin accumulated under nitrogen-fluctuating and high-nitrogen conditions facilitates the persistent dominance and blooms of *Raphidiopsis raciborskii* in tropical waters. *Water Res* 214:118215. <https://doi.org/10.1016/j.watres.2022.118215>
- Farmer M, Rajasabhai R, Tarpeh W, Tyo K, Wells G. 2023. Meta-omic profiling reveals ubiquity of genes encoding for the nitrogen-rich biopolymer cyanophycin in activated sludge microbiomes. *Front Microbiol* 14:1287491. <https://doi.org/10.3389/fmicb.2023.1287491>

5. Illumina. 2021. BCL convert: a proprietary Illumina software for the conversion of BCL files to basecalls
6. Oxford Nanopore Technologies. 2023. Guppy protocol. Available from: <https://nanoporetech.com/document/Guppy-protocol>
7. WickRR. 2017. Porechop: adapter trimmer for Oxford Nanopore reads. <https://github.com/rrwick/Porechop>.
8. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>
10. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
11. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb Genom* 7:000685. <https://doi.org/10.1099/mgen.0.000685>
12. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, Cowley ES, Karaoz U, Anantharaman K. 2022. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. *Microbiome* 10:33. <https://doi.org/10.1186/s40168-021-01213-8>
13. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* 51:D587–D592. <https://doi.org/10.1093/nar/gkac963>