

Updated draft genome sequence of *Streptomyces* sp. isolate H28 from the Meycauayan River, Philippines

Hanna Hasmin T. Estadilla,¹ Aroin B. Verdera,² Arizaldo E. Castro,³ Kohsuke Honda,⁴ Marie Christine M. Obusan¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report an updated draft genome sequence of *Streptomyces* sp. isolate H28, a melanin-producing bacterium recovered from the sediments of Meycauayan River, Philippines. The improvement in sequencing methodology is reflected in the reported statistics. The new sequence data is highly useful for future exploration given the melanogenesis of the bacterium.

KEYWORDS draft genome sequence, *Streptomyces*, melanin

Melanin-producing isolate H28 was recovered from the sediments of the Meycauayan tributary of the Marilao-Meycauayan-Obando River System (MMORS) (14°44'14.29"N, 120°57'26.11"E) in Bulacan, Philippines—one of the world's most polluted water systems (1, 2). Melanin production is likely the bacterium's stress response to the environment (3–5).

For bacterial isolation, standard dilutions of 9:1 normal saline solution and one gram of collected sediments were prepared and inoculated onto growth media. Preliminary screening showed that isolate H28 can produce melanin, as evidenced by the occurrence of dark pigmentation when inoculated onto tyrosine agar media. A pure culture of isolate H28 was maintained in Trypticase Soy Agar media (30°C, 7 days incubation) (6, 7).

Genomic DNA was extracted using the Vivantis GF-1 Bacterial DNA Extraction Kit. The Sequencing library was prepared using the Illumina TruSeq Nano DNA library (350 bp insert). Sequencing was performed on a NovaSeq platform with a 100 bp paired-end setting. Initial assessment revealed a total of 15,560,742 sequence reads comprising 2,349,672,042 bases from the whole genome sequencing. Raw reads were processed using fastp (0.19.5) (8) with default quality control parameters, including adapter trimming, quality filtering, and removal of reads with low-quality bases ($Q < 15$). After filtering, there were a total of 15,291,866 reads and 2,308,275,818 bases (97% of remaining bases are $\geq Q20$). Filtered reads were used for *de novo* assembly using the Unicycler pipeline (0.5.1) (9) with SPAdes as the assembly algorithm (3.15.3) (10). A k-mer size of 120 and a minimum contig length of 20,000 were used as assembly parameters. The resulting assembly was analyzed using Quast (5.3.0) (11). The whole genome assembly has a final length of 6,791,600 bp and a coverage of 340x. It has 20 contigs, with the largest being 1,237,597 bp in length. The assembly has an N_{50} of 517,880 bp, L_{50} of 4 contigs, and a GC content of 72.97%.

Average Nucleotide Identity (ANI) based on mumMER (4.0) (12) identified the bacterium to be most closely related to *Streptomyces viridosporus* NRRL 2414 (88.31% ANI) (GCA_002078235.1). BLAST+ (2.16.0) (13) revealed *Streptomyces griseomycini* JCM 4382 (GCF_014649595.1) as the closest identity (85.83% ANI).

The assembled genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (14). A total of 6,075 genes were annotated—5,924 of which are protein-coding genes. Genes coding for putative tyrosinase and tyrosinase cofactor that are associated with melanin production were found. Melanin is formed through the

Editor Irene L. G. Newton, Indiana University, Bloomington, Bloomington, Indiana, USA

Address correspondence to Marie Christine M. Obusan, mmobusan@up.edu.ph.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 18 February 2025

Accepted 26 September 2025

Published 29 October 2025

Copyright © 2025 Estadilla et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

TABLE 1 Putative BGCs and number of regions found in the updated genome assembly of *Streptomyces* sp. isolate H28

BGCs found	No. of regions
nonribosomal peptides (NRPS)	4
furan	1
lanthipeptide class I	1
butyrolactone	1
melanin	2
type II polyketide synthase (T2PKS)	1
terpene	6
thiopeptide	1
leukocyte alkaline phosphatase (LAP)	1
NRPS-independent (NI)-siderophore	2
ribosomally synthesized and post-translationally modified peptides (RiPP)-like	3
RiPP recognition element (RRE)-containing	1
type I polyketide synthase (T1PKS)	2
indole	1
trans-acyltransferase polyketide synthases (transAT-PKS)	1
highly reducing type II PKS (HR-T2PKS)	1
tripeptide	1
ectoine	1

action of tyrosinase, which converts tyrosine to melanin via a series of enzymatic and non-enzymatic reactions (6, 7). Melanin is one of the 18 secondary metabolite types that are encoded by the 31 biosynthetic gene clusters (BGCs) discovered in the genome (antiSMASH (7.1.0) [15]) (Table 1).

The resequencing and reassembly were conducted to improve assembly quality reflected in the improved statistics as reported. We report higher coverage (26× to 340×), improved genome contiguity (1,181 contigs to 20 contigs), higher N50 (9,931 bp to 517,880 bp), and lower L50 (four contigs) compared to the previous draft assembly (1). Such improvement allowed a more accurate identification and annotation of BGCs, specifically those involved in the melanin metabolic pathway. The new assembly is deemed more appropriate for future use and exploration of the bacterium, given its melanogenesis and potential novelty.

ACKNOWLEDGMENTS

This study was funded by the University of the Philippines (UP) Office of the Vice President for Academic Affairs - Enhanced Creative Work and Research Grant (OVPAA-ECWRG) (Identifier Number: ECWRG-2020-2-13R).

AUTHOR AFFILIATIONS

¹Institute of Biology, College of Science, University of the Philippines Diliman, Quezon City, Philippines

²Department of Biology, College of Arts and Sciences, University of the Philippines, Manila, Philippines

³Biological Sciences Department, College of Science, De La Salle University-Dasmariñas, Dasmariñas City, Philippines

⁴Department of Biotechnology, Graduate School of Engineering, Osaka University Suita Campus, Suita, Japan

PRESENT ADDRESS

Marie Christine M. Obusan, Microbial Ecology of Terrestrial and Aquatic Systems Laboratory, Institute of Biology, College of Science, University of the Philippines Diliman, Philippines

AUTHOR ORCIDs

Hanna Hasmin T. Estadilla  <http://orcid.org/0009-0003-8624-7038>

Aroin B. Verdera  <http://orcid.org/0000-0001-9929-6611>

Arizaldo E. Castro  <http://orcid.org/0000-0001-6929-5407>

Kohsuke Honda  <https://orcid.org/0000-0001-9069-7574>

Marie Christine M. Obusan  <http://orcid.org/0000-0002-5453-1881>

FUNDING

Funder	Grant(s)	Author(s)
University of the Philippines (UP) Office of the Vice President for Academic Affairs-Enhanced Creative Work and Research Grant (OVPAA-ECWRG)	ECWRG-2020-2-13R	Marie Christine M. Obusan

AUTHOR CONTRIBUTIONS

Hanna Hasmin T. Estadilla, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Aroin B. Verdera, Investigation, Methodology, Resources, Software, Supervision, Writing – review and editing | Arizaldo E. Castro, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Kohsuke Honda, Supervision, Validation, Writing – review and editing | Marie Christine M. Obusan, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

This whole genome sequencing and assembly project was deposited at GenBank under the accession numbers [JBPVSX000000001](#) to [JBPVSX010000020](#), BioSample number [SAMN46394803](#), and Sequencing Read Archive (SRA) number [SRR32141758](#).

REFERENCES

- Verdera AB, Montecillo AD, Obusan MCM. 2021. Draft genome sequence of *Streptomyces* sp. isolate H28 from the Meycauayan river, Philippines. Microbiol Resour Announc 10:10–128. <https://doi.org/10.1128/MRA.01269-20>
- Pleto JV. 2015. Assessment of different bioremediation strategies on the environmental quality of the aquaculture ponds of Brgy.[Village] Nagbalon, Marilao and Brgy. Liputan, Meycauayan, Bulacan, Philippines.
- Belo LP, Orbecido AH, Beltran AB, Vallar EA, Galvez MC, Eusebio RC, Ledesma NA, Deocaris CC. 2009. Water quality assessment of Meycauayan River, Bulacan, Philippines (Research Note). Sylvatrop: the technical journal of Philippine Ecosystems and Natural Resources 28:123–140.
- ADB. 2009. Pilot and demonstration activity: Philippines. Reduction of mercury and heavy metals contamination resulting from artisanal gold refining in Meycauayan, Bulacan River system. Asian Development Bank Water for All - Final Report
- Cruz GM, Mandap CM. 2012. Bioaccumulation of heavy metals in selected fishes and shellfishes along the Marilao-Meycauayan-Obando (MMO) River system, province of Bulacan. 4th International Education Conference. p 11
- El-Naggar NE-A, El-Ewasy SM. 2017. Bioproduction, characterization, anticancer and antioxidant activities of extracellular melanin pigment produced by newly isolated microbial cell factories *Streptomyces glaucescens* NEAE-H. Sci Rep 7:42129. <https://doi.org/10.1038/srep42129>
- Marino SM, Fogal S, Bisaglia M, Moro S, Scartabelli G, De Gioia L, Spada A, Monzani E, Casella L, Mammi S, Bubacco L. 2011. Investigation of *Streptomyces* antibiotic tyrosinase reactivity toward chlorophenols. Arch Biochem Biophys 505:67–74. <https://doi.org/10.1016/j.abb.2010.09.019>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell

- sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
11. 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>
 12. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol* 14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>
 13. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>
 14. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>
 15. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJM, van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res* 51:W46–W50. <https://doi.org/10.1093/nar/gkad344>