

MRA-2 Sino

by Har Woko

Submission date: 27-Apr-2023 07:58AM (UTC+0700)

Submission ID: 2076656769

File name: mra.01268-22.pdf (507.21K)

Word count: 2165

Character count: 11886



Draft Genome Sequence of *Sinomicrobium* sp. Strain PAP.21, Isolated from a Coast Sample of Papua, Indonesia

Riyanti,^a Celine M. Zumkeller,^b Marius Spohn,^{b,e} Sanja Mihajlovic,^b Oliver Schwengers,^{c,d} Alexander Goesmann,^{c,d} Riviani Riviani,^a Maria D. N. Meinita,^a Till F. Schäberle,^{b,d,e} Harwoko Harwoko^f

^aFaculty of Fisheries and Marine Science, Jenderal Soedirman University, Purwokerto, Indonesia

^bBranch for Bioresources, Fraunhofer Institute for Molecular Biology and Applied Ecology (IME), Giessen, Germany

^cBioinformatics and Systems Biology, Justus Liebig University Giessen, Giessen, Germany

^dGerman Center for Infection Research (DZIF), Partner Site Giessen-Marburg-Langen, Giessen, Germany

^eInstitute for Biotechnology, Justus Liebig University Giessen, Giessen, Germany

^fDepartment of Pharmacy, Faculty of Health Sciences, Jenderal Soedirman University, Purwokerto, Indonesia

Riyanti and Celine M. Zumkeller contributed equally to this work. Author order was determined both alphabetically and in order of increasing seniority.

ABSTRACT *Sinomicrobium* sp. strain PAP.21 (EXT111902) was isolated from the coast of Cenderawasih Bay National Park in West Papua, Indonesia. Its genome was assembled into 151 contigs with a total size of 5.439 Mbp, enabling the prediction of its specialized metabolite production capacity.

The phylum *Bacteroidetes* is a promising, although underexplored, bioresource for natural product discovery (1, 2). *Bacteroidetes* are among the most abundant bacteria within marine ecosystems (3), and their known bioactive natural products represent a remarkable diversity. This diversity is, e.g., exemplified by the lanthipeptide pinensin from *Chitinophaga pinensis* exhibiting antifungal activity (4). Expecting a positive correlation between the *Bacteroidetes* diversity and their produced chemical diversity, we aimed at accessing new strains.

In this study, *Sinomicrobium* sp. PAP.21 was isolated from marine sediment of the upper layer (5 to 10 cm) collected at a coast area of Cenderawasih Bay National Park, Papua, Indonesia (2°23'06.8"S 134°57'53.5"E). The strain was deposited in the Fraunhofer strain collection (5) under its identifier EXT111902.

The marine sediment sample was stored in a sterile plastic bag before being plated onto artificial seawater (1.5% agar, 0.01% KBr, 2.3% NaCl, 1.1% MgCl₂ · 6 H₂O, 0.1% CaCl₂ · 2H₂O, 0.1% KCl, 0.004% SrCl₂ · 6 H₂O, 0.4% Na₂SO₄, 0.02% NaHCO₃, and 0.003% H₃BO₃ in H₂O), including *Escherichia coli* prey, and incubated at 30°C. Grown colonies were purified on marine agar (Carl Roth GmbH; product no. CP73.1 + 1.5% agar), and a sequencing sample was prepared by growing EXT111902 aerobically in marine broth (30°C for 24 h; Carl Roth GmbH; product no. CP73.1). The cell pellet was resuspended in ATL buffer (Qiagen) containing RNase A. BashingBead lysis tubes (Zymo Research) were used for cell disruptions. DNA was isolated using QIAmp 96 DNA Acube high-throughput (HT) kits with the addition of proteinase K (Qiagen). Libraries for short-read sequencing were prepared using the Illumina DNA prep tagmentation kit with 500 ng DNA input and 5 cycles of indexing PCR. Library quality was evaluated (Agilent 2100 bioanalyzer) and sequenced on an Illumina NovaSeq instrument using a NovaSeq 6000 SP v1 sequencing kit with a 2 × 150 bp read length and a depth of 4.0 to 5.0 Mio reads. For sequence processing and analysis, software tools were run with default settings unless otherwise stated. The sequencing was demultiplexed (Illumina bcl2fastq, v2.19.0.316), quality checked (Fastp [6] v0.20.1), and visualized (MultiQC [7] v1.7). A total of 11.98 million paired-end reads were quality filtered (Fastp [6] v0.20.1; additional parameter “-detect_adapter_for_pe -cut_by_quality5 -cut_by_quality3

Editor Frank J. Stewart, Montana State University

Copyright © 2023 Riyanti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Till F. Schäberle, till.schaeberle@ime.fraunhofer.de, or Harwoko Harwoko, harwoko.unsoed@gmail.com. The authors declare no conflict of interest.

Received 6 December 2022

Accepted 22 February 2023

Published 21 March 2023

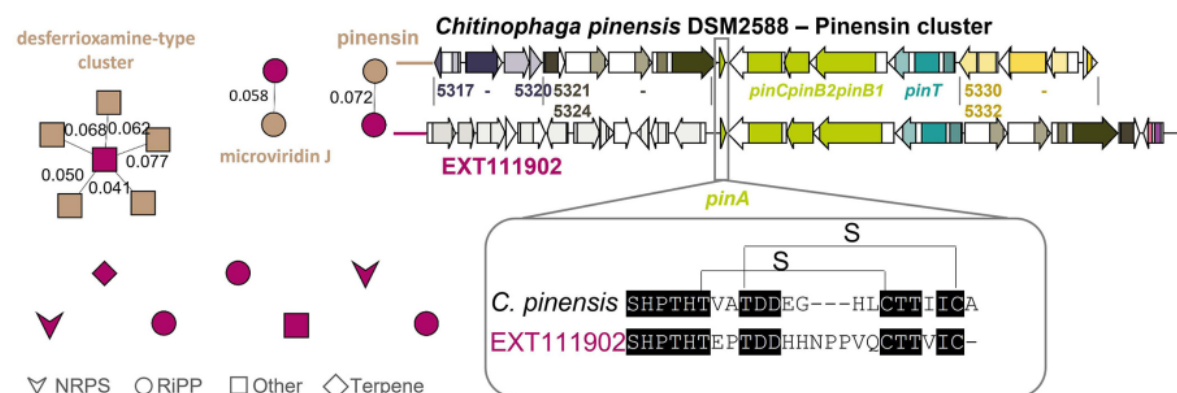


FIG 1 BiG-SCAPE network (16) (left) of strain EXT111902 (magenta) (manually curated). EXT111902 carries five ribosomally synthesized and posttranslationally modified peptides (RiPPs; circle), two nonribosomal peptides (NRPSs; arrow), one terpene (diamond), and two other BGCs (rectangle). By networking with the MIBiG reference clusters (brown), a desferrioxamine-like NRPS-independent siderophore BGC (similarity index up to 0.077, BGC0001478), a microviridin-like BGC (similarity index of 0.058, BGC0000593), and a pinensin-like BGC (similarity index of 0.072) were annotated. Alignment of the pinensin-like cluster of EXT111902 to the pinensin BGC0001392 of *C. pinensis* DSM2588 (right) shows the presence of homologous genes to the biosynthetic core genes *pinA* to *C* (light green) and *pinT* (turquoise), including the split dehydratase genes *pinB1* and *pinB2*. Homologous genes to Cpin5321 to 5324 (dark green) of BGC0001392, suggested to be involved in the perception and import of pinensin, are present. Homologs to Cpin5330 to 5332 (yellow), proposed to be responsible for pinensin export, instead are not detected in the EXT111902 BGC. Alignment of the pinensin core peptide (bottom right) suggests a sequence of 24 amino acids for EXT111902, while the pinensin A peptide contains 22 amino acids.

–low_complexity_filter –length_required 21 –correction”), assembled (Unicycler [8] v0.4.8), and annotated (Bakta [9] v1.5.1). The genome consists of 5,438,544 bp (coverage, 320×; N_{50} 141,686 bp; L_{50} 13) in 20 contigs and has a GC content of 44.4%. Using CheckM (v1.0.18) (10), the degree of genome completeness was determined 19 99.34% with 3.9% contamination. The genome encodes 4,679 protein-coding genes, 47 tRNAs, 1 transfer-messenger RNA (tmRNA), 3 rRNAs and 6 noncoding RNAs (ncRNAs). The taxonomical rank was established using the Type Strain Genome Server (11). This process revealed *Sinomicrobium oceani* CGMCC 1.12145 (12) as the closest related type strain. Digital DNA-DNA hybridization (dDDH) values 22 exceed the species delineation threshold of 70% (76.9% [d0], 92.1% [d4], and 82.3% [d6]). An average nucleotide identity (ANI) (13) value of 98.89% supports the affiliation of EXT111902 to the species *S. oceani*.

antiSMASH v6.0 (14) was employed to predict the biosynthetic gene clusters (BGCs). BGC annotation was achieved by their clustering with MIBiG (15) reference clusters into gene cluster families (GCFs) using BiG-SCAPE (16) and setting a cutoff value of 0.6. EXT111902 carries one BGC clustering with BGC0001478, encoding the synthesis of the siderophore desferrioxamine E (17) and another BGC was annotated to BGC0000593, encoding microviridin J (18). The third database annotation refers to pinensin (BGC0001392) (4). Pinensin-like BGCs were detected previously in the genomes of certain *Chitinophaga*, *Chryseobacterium*, *Elizabethkingia*, *Pedobacter*, and *Sinomicrobium* strains (19). Variations in the amino acid sequence of the core peptides indicate a yet undiscovered structural diversity within this lantipeptide type 4. Such alterations toward the known pinensins are also predicted for EXT111902 (Fig. 1).

Data availability. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under BioProject PRJNA904543 with accession number SAMN31846359. Raw data can be obtained from the Sequence Read Archive (SRP421602). The draft genome sequence has been deposited in GenBank under the accession number JAPPSR000000000.

ACKNOWLEDGMENTS

We acknowledge the financial support of the International Research Collaboration Grant, Jendral Soedjono University (grant number T/478/UN23.18/PT.01.03/2022); the Hessian State Ministry of Higher Education, Research and the Arts (HMWK) via the LOEWE Center for Insect Biotechnology and Bioresources; and the Federal Ministry of Education and Research (BMBF) via a German Centre for Infection Research (DZIF) sequencing grant.

Sequencing was performed by the Institute for Medical Microbiology (part of the NGS Competence Center NCCT, Tübingen, Germany) while data management, including data storage of raw data for this project, was done by the Quantitative Biology Center (QbiC; Tübingen, Germany). We acknowledge provision of computing resources and technical support by the Bioinformatics Core Facility (BCF) at Justus Liebig University Giessen.

REFERENCES

- Brinkmann S, Kurz M, Patras MA, Hartwig C, Marner M, Leis B, Billion A, Kleiner Y, Bauer A, Toti L, Pöwerlein C, Hammann PE, Vilcinskis A, Glaeser J, Spohn MS, Schäberle TF. 2021. Genomic and chemical decryption of the Bacteroidetes phylum for its potential to biosynthesize natural products. *Microbiol Spectr* 10:e0247921. <https://doi.org/10.1128/spectrum.02479-21>.
- Brinkmann S, Spohn MS, Schäberle TF. 2022. Bioactive natural products from Bacteroidetes. *Nat Prod Rep* 39:1045–1065. <https://doi.org/10.1039/d1np00072a>.
- Fernández-Gómez B, Richter M, Schüler M, Pinhassi J, Acinas SG, González JM, Pedrós-Alió C. 2013. Ecology of marine Bacteroidetes: a comparative genomics approach. *ISME J* 7:1026–1037. <https://doi.org/10.1038/ismej.2012.169>.
- Mohr KI, Volz C, Jansen R, Wray V, Hoffmann J, Bernecker S, Wink J, Gerth K, Stadler M, Müller R. 2015. Pinensins: the first antifungal lantibiotics. *Angew Chem Int Ed Engl* 54:11254–11258. <https://doi.org/10.1002/anie.201500927>.
- Fox JL. 2014. Fraunhofer to mine Sanofi microbial collection. *Nat Biotechnol* 32:305. <https://doi.org/10.1038/nbt0414-305a>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048. <https://doi.org/10.1093/bioinformatics/btw354>.
- 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>.
- 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. *Microbial Genomics* 7:e000685. <https://doi.org/10.1099/mgen.0.000685>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>.
- Xu Y, Tian X-P, Liu Y-J, Li J, Kim C-J, Yin H, Li W-J, Zhang S. 2013. *Sinomicrobium oceanicum* gen. nov., sp. nov., a member of the family Flavobacteriaceae isolated from marine sediment. *Int J Syst Evol Microbiol* 63:1045–1050. <https://doi.org/10.1099/ijso.0.041889-0>.
- Lee I, Ouk Kim Y, Park S-C, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>.
- Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 49:W29–W35. <https://doi.org/10.1093/nar/gkab335>.
- Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hoof JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, Selem-Mojica N, Alanjary M, Robinson SL, Lund G, Epstein SC, Sisto AC, Charkoudian LK, Collemare J, Linington RG, Weber T, Medema MH. 2020. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res* 48:D454–D458. <https://doi.org/10.1093/nar/gkz882>.
- Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappellini LTD, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH. 2020. A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16:60–68. <https://doi.org/10.1038/s41589-019-0400-9>.
- Barona-Gómez F, Wong U, Giannakopoulos AE, Derrick PJ, Challis GL. 2004. Identification of a cluster of genes that directs desferrioxamine biosynthesis in *Streptomyces coelicolor* M145. *J Am Chem Soc* 126:16282–16283. <https://doi.org/10.1021/ja045774k>.
- Ziemert N, Ishida K, Liaimer A, Hertweck C, Dittmann E. 2008. Ribosomal synthesis of tricyclic depsipeptides in bloom-forming cyanobacteria. *Angew Chem Int Ed Engl* 47:7756–7759. <https://doi.org/10.1002/anie.200802730>.
- Caetano T, van der Donk W, Mendo S. 2020. Bacteroidetes can be a rich source of novel lanthipeptides: the case study of *Pedobacter lusitanus*. *Microbiol Res* 235:126441. <https://doi.org/10.1016/j.micres.2020.126441>.

MRA-2 Sino

ORIGINALITY REPORT

19%

SIMILARITY INDEX

15%

INTERNET SOURCES

16%

PUBLICATIONS

8%

STUDENT PAPERS

PRIMARY SOURCES

- | | | |
|---|--|-----|
| 1 | www.pure.ed.ac.uk
Internet Source | 1 % |
| 2 | Domonkos Sváb, Linda Falgenhauer, Manfred Rohde, Judit Szabó, Trinad Chakraborty, István Tóth. "Identification and Characterization of T5-Like Bacteriophages Representing Two Novel Subgroups from Food Products", <i>Frontiers in Microbiology</i> , 2018
Publication | 1 % |
| 3 | Will, Torsten, and Andreas Vilcinskas. "The structural sheath protein of aphids is required for phloem feeding", <i>Insect Biochemistry and Molecular Biology</i> , 2015.
Publication | 1 % |
| 4 | www.hindawi.com
Internet Source | 1 % |
| 5 | academic.oup.com
Internet Source | 1 % |
| 6 | opus4.kobv.de
Internet Source | 1 % |

7	www.researchsquare.com Internet Source	1 %
8	boris.unibe.ch Internet Source	1 %
9	M Trenggono, R R Hidayat, T N Cahyo, M D Mahardiono, A D Destrianty. "An assessment Indonesia's Ocean Thermal Energy Conversion (OTEC) as an electrical energy resource", IOP Conference Series: Earth and Environmental Science, 2021 Publication	1 %
10	Nirmal Singh Mahar, Kartik Aiyer, Juwayria, Lucinda Elizabeth Doyle, Ishaan Gupta. " Draft Genome Sequence of a sp., a Member of an Electroactive Community Enriched from Wastewater from the Indian Institute of Technology Delhi, India ", Microbiology Resource Announcements, 2023 Publication	1 %
11	www.mdpi.com Internet Source	1 %
12	www.ojs.jmolekul.com Internet Source	1 %
13	Alexey M. Afonin, Emma S. Gribchenko, Anton S. Sulima, Gulnar A. Akhtemova, Vladimir A. Zhukov. "Draft Genome Sequence of	1 %

Rhizobium ruizarguesonis (Rhizobium leguminosarum) Strain 1TK341", Microbiology Resource Announcements, 2022

Publication

14

Anna Serroni, Sara Traini, Mariangela Iorio, Iolanda Mangone et al. "Whole-Genome Sequence and Assembly of Eight Africa Horse Sickness Virus Strains Collected in Namibia and South Africa", Microbiology Resource Announcements, 2023

Publication

1 %

15

Nathaniel J. Ennis, Dhanasekaran Dharumadurai, Louis S. Tisa. "Draft Genome Sequences of 6 Actinobacterial Strains Isolated from Rock Surfaces Obtained from Indian Stone Ruins in Tamil Nadu, India, and Rocks from New England, United States", Microbiology Resource Announcements, 2022

Publication

1 %

16

eprints.soton.ac.uk

Internet Source

1 %

17

hdl.handle.net

Internet Source

1 %

18

researchmap.jp

Internet Source

1 %

19

Amit Kumar, Malyaj R. Prajapati, Surendra Upadhyay, Anamika Bhordia et al. "Complete

1 %

Genome Sequence of Brucella abortus 2308, Isolated from an Abortion Storm on a Dairy Farm in India", Microbiology Resource Announcements, 2022

Publication

-
- 20 Jaka Widada, Ema Damayanti, Mustofa. "Complete Genome Sequence of Bacillus velezensis GMEKP1, Isolated from a Natural Bamboo Hive of Stingless Bees", Microbiology Resource Announcements, 2021 1 %

Publication

-
- 21 edepot.wur.nl 1 %
Internet Source

-
- 22 mdpi-res.com 1 %
Internet Source

-
- 23 pure.rug.nl 1 %
Internet Source

-
- 24 www.journaltocs.ac.uk 1 %
Internet Source

Exclude quotes On

Exclude matches < 1%

Exclude bibliography On