Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Draft genome sequences and functional annotation data for two Pseudoalteromonas isolates from coastal seawater off Jeju Island, Republic of Korea

Jong Woo Hyeon
J2K Bio Co., Ltd
Ji-Hyun Kim
J2K Bio Co., Ltd
Yong Han Cho
J2K Bio Co., Ltd
Jong Bin Yoon
J2K Bio Co., Ltd
Jun-Tae Bae
J2K Bio Co., Ltd
Man-Young Jung
myjung@jejunu.ac.kr

Jeju National University

Data Note

Keywords: Pseudoalteromonas, draft genome, genome mining, pigment-associated biosynthetic gene cluster, vioABCDE-like gene cluster, functional genome analysis

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10519154/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Draft genome sequences and functional annotation data for two *Pseudoalteromonas* isolates from coastal seawater off Jeju Island, Republic of Korea

Jong Woo Hyeon[1,2], Ji-Hyun Kim[1,3], Yong Han Cho[1], Jong Bin Yoon[1], Jun-Tae Bae[1], Man-Young Jung[2,4,5,*]

[1]J2K Bio Co., Ltd., 50-3, Dureungyuri-ro, Ochang-eup, Cheongju-si, Chungcheongbuk-do 28104, Republic of Korea

[2]Interdisciplinary Graduate Program in Advanced Convergence Technology and Science, Jeju National University, Jeju 63243, Republic of Korea

[3]Department of Chemistry and Cosmetics, Jeju National University, Jeju 63243, Republic of Korea

[4]Department of Biology Education, Jeju National University, Jeju 63243, Republic of Korea

[5]Jeju Microbiome Center, Jeju National University, Jeju, 63243, Republic of Korea

*** Corresponding author:** Man-Young Jung, E-mail: myjung@jejunu.ac.kr; Tel: +82-64-754-3282

Abstract

Objectives: Two marine *Pseudoalteromonas* strains, J2K-J031 and J2K-J979, were isolated from coastal seawater off Jeju Island, Republic of Korea, and showed distinct colony pigmentation: cream-white in J2K-J031 and purple in J2K-J979. To characterize these isolates at the genome level, we generated draft genome sequences, functional annotation profiles, and biosynthetic gene cluster predictions for both strains.

Data description: The draft genomes were 4.46 Mb for J2K-J031 and 4.42 Mb for J2K-J979, with 100% estimated completeness and low contamination. Carbohydrate-active enzyme profiles and functional classifications based on Clusters of Orthologous Groups and Gene Ontology were broadly comparable between the two strains. In the biosynthetic gene cluster analysis, antiSMASH detected an indole-type cluster containing a putative *vioABCDE*-like gene set in J2K-J979 but not in J2K-J031. In a maximum-likelihood tree based on concatenated VioA–VioE sequences, the J2K-J979 locus was placed within the *Pseudoalteromonas* clade, close to homologs from *Pseudoalteromonas tunicata* and *Pseudoalteromonas ulvae*. These datasets are provided to support further investigation of pigment-associated biosynthesis and comparative genomics in the genus *Pseudoalteromonas*.

Keywords: *Pseudoalteromonas*; draft genome; genome mining; pigment-associated biosynthetic gene cluster; *vioABCDE*-like gene cluster; functional genome analysis

Objective

Members of the genus *Pseudoalteromonas* are widely distributed in marine environments [1, 2] and are recognized for their metabolic versatility and ability to produce diverse extracellular products and bioactive metabolites, including hydrolytic enzymes, antimicrobial compounds, exopolysaccharides, and pigments [3-5]. Several species produce pigments such as violacein, prodiginines, alterochromides, tambjamines, and pyomelanin, which have attracted interest for potential biotechnological, pharmaceutical, and cosmetic applications [3, 6-8]. In particular, violacein is a bisindole pigment biosynthesized from L-tryptophan through the five-gene *vioABCDE* cluster and is known for its antimicrobial, antioxidant, anti-inflammatory, and antitumor activities [9].

This study generated draft genome data and compared their genomic features for two marine *Pseudoalteromonas* isolates, J2K-J031 and J2K-J979, recovered from coastal seawater off Jeju Island, Republic of Korea. The isolates showed distinct colony pigmentation, with J2K-J031 forming cream-white colonies and J2K-J979 forming purple colonies. Genome assembly characteristics, carbohydrate-active enzyme (CAZyme) profiles, Clusters of Orthologous Groups (COG) classifications, Gene Ontology (GO) categories, and biosynthetic gene clusters (BGCs) were compared between the two strains. Particular attention was given to a putative *vioABCDE*-like gene cluster detected in J2K-J979 but not in J2K-J031. To provide evolutionary context for this locus, the predicted VioA–VioE amino acid sequences of J2K-J979 were compared with homologous sequences from publicly available bacterial genomes. The resulting genome assemblies, functional annotation datasets, biosynthetic gene cluster predictions, sequence alignments, and phylogenetic data are provided to support further comparative genomic and pigment-associated biosynthetic studies.

Data description

The data package is summarized in Table 1, and detailed experimental and bioinformatic methods are provided in Data file 1. Strains J2K-J031 and J2K-J979 were isolated from surface seawater collected off Seogwipo-si and Jeju-si, Jeju Island, Republic of Korea, respectively. Both isolates were Gram-stain-negative, rod-shaped, and facultatively anaerobic. J2K-J031 formed cream-white colonies, whereas J2K-J979 formed purple colonies under the same cultivation conditions. Sample metadata, accession numbers, and genome statistics are compiled in Data file 2. The 16S rRNA genes of J2K-J031 and J2K-J979 were amplified using primers 27F and 1492R, and sequenced using primers 340F, 518R, and 805F. The resulting 16S rRNA gene sequences are provided as Data sets 1 and 2. EzBioCloud analysis [17] showed the highest sequence similarities to *Pseudoalteromonas tunicata* D2^T for J2K-J031 and *Pseudoalteromonas ulvae* UL12^T for J2K-J979, with similarities of 98.43% and 98.74%, respectively.

Whole-genome sequencing was performed using the Illumina NovaSeq X platform. The raw sequencing reads are available as Data sets 3 and 4. Reads were evaluated using FastQC v0.11.9 [18], trimmed using fastp v0.20.1 [19], and de novo assembled using SPAdes v4.1.0 [20]. Assembly quality was evaluated using QUAST v5.3.0 [21] and CheckM2 v1.1.0 [22]. The J2K-J031 assembly comprised 4,460,068 bp in 70 contigs, whereas the J2K-J979 assembly comprised 4,415,625 bp in 71 contigs. The corresponding draft genome assemblies are available as Data sets 5 and 6. Mean sequencing coverages were 737× and 786×, G+C contents were 40.3% and 41.1%, and estimated completeness was 100% for both assemblies, with contamination estimates of 0.1% and 0%, respectively. Average nucleotide identity (ANI) values calculated using FastANI v1.34 [23] between each strain and its most closely related type strains were 79.6–85.8%, supporting the designation of both isolates as *Pseudoalteromonas* sp.

Genome annotation using Bakta v1.11.4 [24] predicted 3,934 coding sequences (CDSs), 70

tRNAs, and 6 rRNAs in J2K-J031, and 3,885 CDSs, 61 tRNAs, and 4 rRNAs in J2K-J979. Circular genome maps generated using Proksee [25] are provided in Data file 3.

Functional annotation using dbCAN3 [26] and eggNOG-mapper v2.1.12 [27] is provided in Data file 4, with corresponding CAZyme, COG, and GO graphical summaries in Data file 5. The overall functional profiles were broadly comparable between the two strains. By contrast, BGCs were predicted using antiSMASH v8.0.4 [28], with the results provided in Data file 6. An indole-type *vioABCDE*-like gene cluster was detected in J2K-J979 but not in J2K-J031.

To place this locus in an evolutionary context, the predicted VioA–VioE amino acid sequences of J2K-J979 were compared with homologs from the bacterial genomes listed in Data file 7. Individual and concatenated VioA–VioE amino acid sequence alignments are provided in Data file 8. Data file 9 shows the genomic organization of the *vioABCDE*-like gene cluster identified in strain J2K-J979, together with the maximum-likelihood phylogenetic tree constructed from concatenated VioA–VioE amino acid sequences to determine the phylogenetic placement. The J2K-J979 VioA–VioE sequence set was placed within a distinct *Pseudoalteromonas* clade, separate from the clades formed by the other bacterial genera included in the analysis.

Limitations

The genome sequences of J2K-J031 and J2K-J979 were generated using short-read sequencing and assembled as draft, contig-level genomes. Therefore, unresolved repetitive regions and incomplete genomic contexts may remain. Functional annotations and biosynthetic gene cluster assignments were based on computational predictions. Although a *vioABCDE*-like cluster was detected in J2K-J979, the chemical identity of the purple pigment and the activity of the predicted genes were not experimentally validated.

Abbreviations

CAZyme: Carbohydrate-active enzyme

COG: Clusters of Orthologous Groups

GO: Gene Ontology

BGCs: Biosynthetic gene clusters

ANI: Average nucleotide identity

CDSs: Coding sequences

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

Strains J2K-J031 and J2K-J979 were deposited in the Korean Collection for Type Cultures (KCTC, Republic of Korea) under accession numbers KCTC 18638 and KCTC 18639, respectively. The supporting data files are available in Figshare [10]. The 16S rRNA gene sequences are available in GenBank under accession numbers PZ311729 and PZ311730 [11, 12]. The raw sequencing reads were deposited in the NCBI Sequence Read Archive under accession numbers SRR39712427 and SRR39712428 [13, 14]. The draft genome assemblies are available in DDBJ/ENA/GenBank under accession numbers JBYQUK000000000 and JBYQQY000000000 [15, 16]. Please see Table 1 for details and links to the data.

Competing interests

The authors declare no competing interests.

Funding

This research was financially supported by the Ministry of Trade, Industry and Energy, Korea, under the "Regional Innovation Cluster Development Program (Non-R&D, P0031131)" supervised by the Korea Institute for Advancement of Technology (KIAT).

Authors' contributions

Jong Woo Hyeon: Software, Data curation, Writing original draft. **Ji-Hyun Kim:** Data curation, Writing original draft. **Yong Han Cho:** Investigation, Formal analysis. **Jong Bin Yoon:** Investigation, Formal analysis. **Jun-Tae Bae:** Resources, Investigation. **Man-Young Jung:** Supervision, Conceptualization, Writing, review, and editing.

Acknowledgements

Not applicable.

References

1. Park S, Jung YT, Park DS, Yoon JH. *Pseudoalteromonas aestuariivivens* sp. nov., isolated from a tidal flat. Int J Syst Evol Microbiol. 2016;66:2078–83. https://doi.org/10.1099/ijsem.0.000995.
2. Ivanova EP, Shevchenko LS, Sawabe T, Lysenko AM, Svetashev VI, Gorshkova NM, Satomi M, Christen R, Mikhailov VV. *Pseudoalteromonas maricaloris* sp. nov., isolated from an

Australian sponge, and reclassification of [*Pseudoalteromonas aurantia*] NCIMB 2033 as *Pseudoalteromonas flavipulchra* sp. nov. Int J Syst Evol Microbiol. 2002;52:263–71. https://doi.org/10.1099/00207713-52-1-263.

3. Bowman JP. Bioactive compound synthetic capacity and ecological significance of marine bacterial genus *Pseudoalteromonas*. Mar Drugs. 2007;5:220–41. https://doi.org/10.3390/md504220.
4. Paulsen SS, Strube ML, Bech PK, Gram L, Sonnenschein EC. Marine chitinolytic *Pseudoalteromonas* represents an untapped reservoir of bioactive potential. mSystems. 2019;4:e00060-19. https://doi.org/10.1128/mSystems.00060-19.
5. Holmström C, Kjelleberg S. Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. FEMS Microbiol Ecol. 1999;30:285–93. https://doi.org/10.1111/j.1574-6941.1999.tb00656.x.
6. Wu YH, Cheng H, Xu L, Jin XB, Wang CS, Xu XW. Physiological and genomic features of a novel violacein-producing bacterium isolated from surface seawater. PLoS One. 2017;12:e0179997. https://doi.org/10.1371/journal.pone.0179997.
7. Zeng Z, Guo XP, Cai X, Wang P, Li B, Yang JL, Wang X. Pyomelanin from *Pseudoalteromonas lipolytica* reduces biofouling. Microb Biotechnol. 2017;10:1718–31. https://doi.org/10.1111/1751-7915.12773.
8. Sakai-Kawada FE, Ip CG, Hagiwara KA, Awaya JD. Biosynthesis and bioactivity of prodiginine analogs in marine bacteria, *Pseudoalteromonas*: A mini review. Front Microbiol. 2019;10:1715. https://doi.org/10.3389/fmicb.2019.01715.
9. Sun H, Zhao D, Xiong B, Zhang C, Bi C. Engineering *Corynebacterium glutamicum* for violacein hyper production. Microb Cell Fact. 2016;15:148. https://doi.org/10.1186/s12934-016-0545-0.

10. Hyeon JW, Kim JH, Cho YH, Yoon JB, Bae JT, Jung MY. Supporting data for functional genome analysis of two Pseudoalteromonas isolates from coastal seawater off Jeju Island, Republic of Korea. Figshare. 2026. https://doi.org/10.6084/m9.figshare.33098717.
11. NCBI. 16S rRNA gene sequence of *Pseudoalteromonas* sp. J2K-J031. 2026. https://identifiers.org/ncbi/nucleotide:PZ311729.
12. NCBI. 16S rRNA gene sequence of *Pseudoalteromonas* sp. J2K-J979. 2026. https://identifiers.org/ncbi/nucleotide:PZ311730.
13. NCBI. Paired-end Illumina raw reads of *Pseudoalteromonas* sp. J2K-J031. 2026. https://identifiers.org/ncbi/insdc.sra:SRR39712427.
14. NCBI. Paired-end Illumina raw reads of *Pseudoalteromonas* sp. J2K-J979. 2026. https://identifiers.org/ncbi/insdc.sra:SRR39712428.
15. NCBI. Draft genome sequence of *Pseudoalteromonas* sp. J2K-J031. 2026. https://identifiers.org/ncbi/nucleotide:JBYQUK000000000.
16. NCBI. Draft genome sequence of *Pseudoalteromonas* sp. J2K-J979. 2026. https://identifiers.org/ncbi/nucleotide:JBYQQY000000000.
17. Yoon SH, Ha SM, Kwon S, Lim J, Kim Y, Seo H, Chun J. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol. 2017;67:1613–17. https://doi.org/10.1099/ijsem.0.001755.
18. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics; 2010.
19. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90. https://doi.org/10.1093/bioinformatics/bty560.
20. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, et al. SPAdes: a new genome assembly algorithm and

its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. https://doi.org/10.1089/cmb.2012.0021.

21. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–75. https://doi.org/10.1093/bioinformatics/btt086.
22. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. Nat Methods. 2023;20:1203–12. https://doi.org/10.1038/s41592-023-01940-w.
23. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9:5114. https://doi.org/10.1038/s41467-018-07641-9.
24. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom. 2021;7:000685. https://doi.org/10.1099/mgen.0.000685.
25. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen CY, Graham M, Van Domselaar G, Stothard P. Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res. 2023;51:W484–92. https://doi.org/10.1093/nar/gkad326.
26. Zheng J, Ge Q, Yan Y, Zhang X, Huang L, Yin Y. dbCAN3: automated carbohydrate-active enzyme and substrate annotation. Nucleic Acids Res. 2023;51:W115–21. https://doi.org/10.1093/nar/gkad328.
27. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021;38:5825–29. https://doi.org/10.1093/molbev/msab293.
28. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, Cediel-Becerra JD, de Crécy-Lagard V, Koetsier RA, Williams SE, et al. antiSMASH 8.0: extended gene cluster detection

capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Res. 2025;53:W32–38. https://doi.org/10.1093/nar/gkaf334.

232 **Table 1.** Overview of data files/data sets

Label	Name of data file/data set	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1	Detailed methods	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 2	Sample and genome information	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 3	Circular genome maps	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 4	Functional annotation data	MS Excel file (.xlsx)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 5	Functional annotation figures	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 6	antiSMASH results	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 7	Strains and sequences used for VioABCDE phylogenetic analysis	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 8	VioABCDE sequence alignments	Compressed amino acid FASTA files (.zip; .faa)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data file 9	*vioABCDE*-like gene cluster and phylogeny	PDF file (.pdf)	Figshare: https://doi.org/10.6084/m9.figshare.33098717 [10]
Data set 1	16S rRNA gene sequence of *Pseudoalteromonas* sp. J2K-J031	Sequence file (.fasta)	NCBI GenBank: https://identifiers.org/ncbi/nucleotide:PZ311729 [11]
Data set 2	16S rRNA gene sequence of *Pseudoalteromonas* sp. J2K-J979	Sequence file (.fasta)	NCBI GenBank: https://identifiers.org/ncbi/nucleotide:PZ311730 [12]
Data set 3	Paired-end Illumina raw reads of *Pseudoalteromonas* sp. J2K-J031	Sequence file (.fastq.gz)	NCBI Sequence Read Archive (SRA): https://identifiers.org/ncbi/insdc.sra:SRR39712427 [13]
Data set 4	Paired-end Illumina raw reads of *Pseudoalteromonas* sp. J2K-J979	Sequence file (.fastq.gz)	NCBI Sequence Read Archive (SRA): https://identifiers.org/ncbi/insdc.sra:SRR39712428 [14]
Data set 5	Draft genome sequence of *Pseudoalteromonas* sp. J2K-J031	Sequence file (.fasta)	NCBI GenBank: https://identifiers.org/ncbi/nucleotide:JBYQUK000000000 [15]
Data set 6	Draft genome sequence of *Pseudoalteromonas* sp. J2K-J979	Sequence file (.fasta)	NCBI GenBank: https://identifiers.org/ncbi/nucleotide:JBYQQY000000000 [16]

233