AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Biotechnology | Announcement

Complete genome sequence of *Pseudomonas* sp. PP3, a dehalogenase-producing bacterium, confirms the unusual mobile genetic element *DEH*

Gordon Webster,[1] Amy J. Baldwin,[1] Edward Cunningham-Oakes,[1] Alex J. Mullins,[1,2] Rachel Dodds,[1] Katja E. Hill,[1] Li Ling Lee,[1] Mark J. Leggett,[1] Andrew W. Topping,[1] Andrew J. Weightman[1]

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT *Pseudomonas* sp. PP3, originally isolated from contaminated soil enriched in a chemostat culture on 2,2-dichloropropionic acid, has a 6.42 Mb genome, most closely related to *P. reinekei*. This well-characterized organism continues to provide key insights into adaptive dehalogenase-mediated bioremediation of halogenated organic pollutants.

KEYWORDS dehalogenase, biodegradation, mobile genetic elements, *Pseudomonas*, bioremediation

Halo-organic compounds are used as herbicides, pesticides, preservatives, solvents, and other applications. The persistence and toxicity of many of these compounds raise serious environmental concerns regarding their use (1). Biodegradation of such compounds, including haloalkanoic acids, by *Pseudomonas* species is well known (2).

Pseudomonas sp. strain PP3 was isolated from a soil microbial community in a chemostat culture on the herbicide 2,2-dichloropropionic acid as the sole carbon and energy source as described (3–5). For long-term storage, the strain was kept in 40% (vol/vol) glycerol stocks at −80°C and for routine laboratory culture grown aerobically in standard basal salts medium [SBS; (4)] containing 5 mM halogenated substrate at 30°C, shaking at 150 rpm.

For genome sequencing, PP3 was grown in SBS with 10 mM 2-chloropropionic acid for 24 h. Cells were pelleted by centrifugation (4,000 rpm, 10 min, ALC-PK120 centrifuge), and genomic DNA was extracted using the Wizard genomic DNA purification kit (Promega) according to the manufacturer's protocol. Following QC, DNA was sheared and size-selected (~10 kb) using Covaris g-TUBE and sequenced by Novogene (UK). Sequencing libraries were prepared with the SMRTbell template prep kit 1.0 (PacBio) and NEBNext DNA library prep kit for Illumina. Genome sequencing was performed on a PacBio Sequel SMRT Cell 1M and an Illumina NovaSeq 6000 (paired-end, 2 × 150 bp). In total, 208,470 high-quality PacBio subreads (average length = 5,218 bp; N50 = 7,983 bp) and 19,253,220 Illumina raw reads were obtained and quality checked with FastQC v0.11.8. The genome (620× coverage) was assembled *de novo* from the PacBio subreads into one contig using Flye v2.8 (6) and polished with both the PacBio reads using Arrow (via pbmm2 v1.4.0, GCpp v2.0.0 tools; https://github.com/PacificBiosciences), and Illumina reads with Pilon v1.23 (7) using default settings. The assembled genome was reoriented with Circlator v1.5.5 (8) at the *dnaA* gene start position.

Genome size and other metrics for the assembly are as follows: 6,421,237 bp and 59.17%G + C, similar to other *Pseudomonas* species (Table 1; 9), 5,745 coding DNA sequences (CDS), 19 rRNAs, 71 tRNAs, and four ncRNAs identified using PGAP v6.7 (10) and CGView (11) (Fig. 1A). The PP3 genome contains two 2-haloalkanoate dehalogenase

Editor David A. Baltrus, The University of Arizona, Tucson, Arizona, USA

Address correspondence to Gordon Webster, Websterg@cardiff.ac.uk, or Andrew J. Weightman, Weightman@cardiff.ac.uk.

The authors declare no conflict of interest.

See the funding table on p. 4.

Received 31 January 2025
Accepted 25 March 2025
Published 21 April 2025

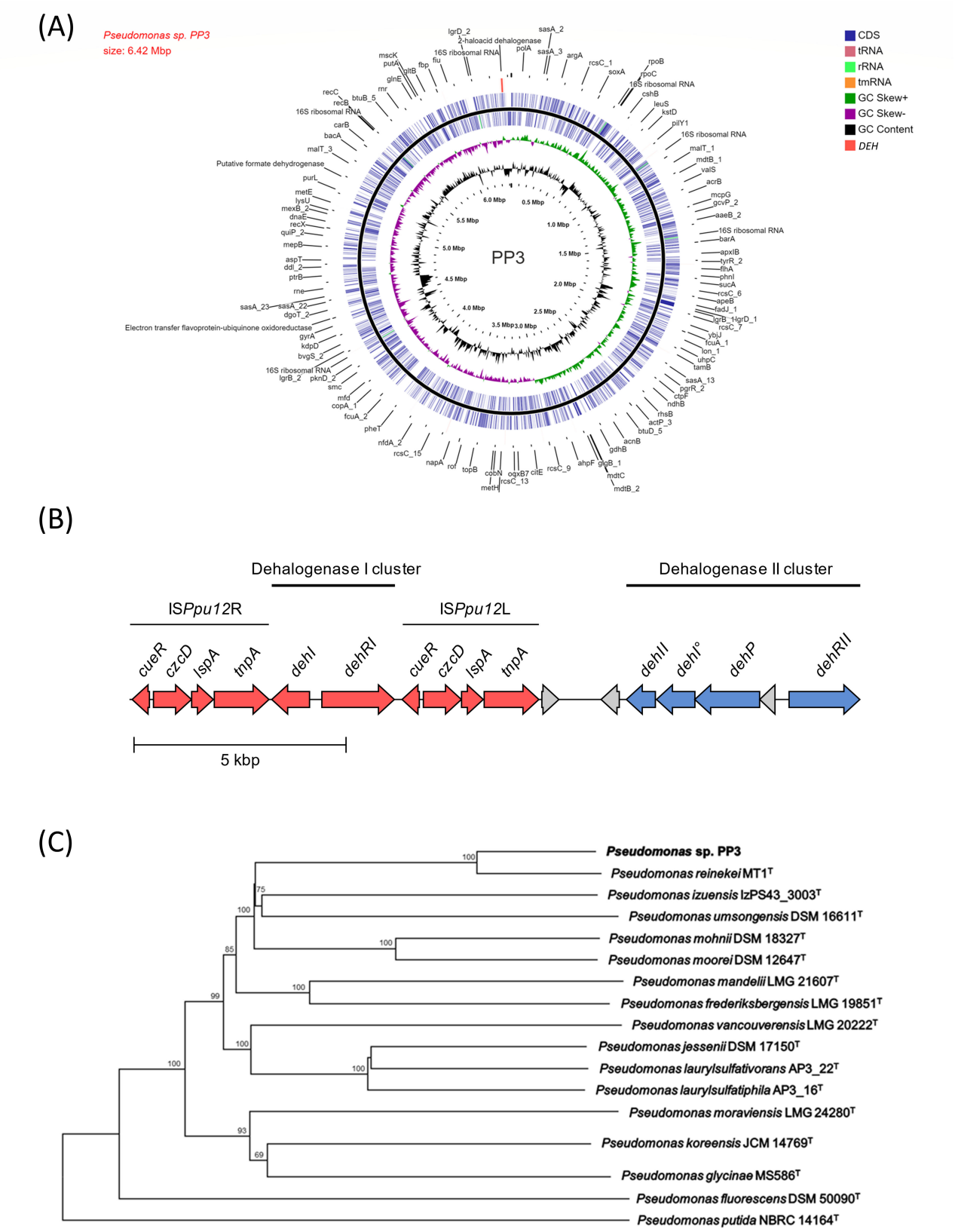

FIG 1 Visual representation and phylogenetic classification of the *Pseudomonas* sp. PP3 genome and *DEH* mobile element. (A) *Pseudomonas* sp. PP3 chromosome map of 5,744 predicted proteins. From outer circle to the center: select view of gene identification; dehalogenase gene clusters (within and around *DEH*); all CDSs (blue) and RNAs (crimson/green/orange) genes on forward strand; all CDSs (blue) and RNAs (crimson/green/orange) genes on reverse strand; GC skew

(Continued on next page)

Fig 1 (Continued)

(positive GC skew values are plotted in green, and negative values are in purple); GC content (black); scale bar. The map was generated using the Circular Genome Viewer (CGView). (B) Close-up of the dehalogenase gene clusters (*dehI*) and the upstream *dehII* region highlighted in (A), generated using Clinker v0.0.27 (14). The gene designations for the dehalogenase I cluster within the *DEH* element are as follows: *dehI*, dehalogenase I family gene (12); *dehRI*, σ^{54}-dependent activator; *tnpA*, putative ISL3-family transposase; *lspA*, putative lipoprotein signal peptidase; *czcD*, putative heavy-metal-associated efflux transporter; *cueR*, putative heavy-metal-associated responsive transcriptional regulator. The positions of the IS*Ppu12* insertion sequence (independently mobile) flanking regions are also indicated (13). The gene designations for the dehalogenase II cluster are as follows: *dehII*, dehalogenase II family gene; *dehI°*, cryptic dehalogenase I family gene; *dehP*, putative permease transporter; *dehRII*, σ^{54}-dependent activator. (C) Phylogenetic classification of *Pseudomonas* sp. PP3. The genomes of the 14 closest related type strains along with type strains *P. putida* and *P. fluorescens* were used for phylogenetic analysis as described by TYGS (15). The tree was inferred with FastME 2.1.6.1 (16) using Genome BLAST Distance Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d_5. The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of 86.1%. The tree was rooted at the midpoint.

genes (representing *dehI* and *dehII* gene families) previously described (12), and one silent *dehI* gene. The *DEH* mobile genetic element in PP3 (13) is confirmed to contain a *dehI* family gene (12) and its regulatory gene, *dehRI,* flanked by two almost identical insertion sequences (IS*Ppu12*). The *DEH* element is located close to a separate putative dehalogenase operon (Fig. 1B). Several other ORFs encoding putative dehalogenases and enzymes associated with halo-organic catabolism are also evident in the genome.

Average nucleotide identity (ANI) analysis with pyani v0.2.10 (18) and genome comparisons using the Type (Strain) Genome Server [TYGS; (15)] inferred that PP3 groups phylogenomically with *P. reinekei* MT1^T (Table 1; Fig. 1C), with an ANI value >95% proposed as the same species (17). However, dDDH analysis produced a value below the 70% species threshold (17) compared with *P. reinekei* and other *Pseudomonas* genomes (Table 1). The ANI and dDDH values together suggest that PP3 may represent a new species closely related to *P. reinekei*.

TABLE 1 Pairwise ANI and digital DDH between *Pseudomonas* sp. PP3 and other *Pseudomonas* species, with their respective genome sizes and %G + C content

Genome (accession number)	Genome size (Mbp)	G + C content (%)	*Pseudomonas* sp. PP3	
			Pairwise ANI (%)[a]	Pairwise DDH (%)[b]
Pseudomonas sp. PP3	6.42	59.2	100	100
Pseudomonas reinekei MT1^T (GCA_001945365)	6.25	59.2	95.6	63.3
Pseudomonas izuensis IzPS43_3003^T (GCA_009861505)	6.86	59.6	88.5	33.6
Pseudomonas umsongensis DSM 16611^T (GCA_002236105)	6.70	59.7	88.2	33.0
Pseudomonas mohnii DSM 18327^T (GCA_900105115)	6.59	59.6	88.3	33.5
Pseudomonas moorei DSM 12647^T (GCA_900102045)	6.55	59.7	88.5	34.0
Pseudomonas putida NBRC 14164^T (GCA_000412675)	6.16	62.3	84.3	22.4
Pseudomonas fluorescens DSM 50090^T (GCA_001269845)	6.39	60.2	85.6	24.7

[a]Average nucleotide identity (ANI) values <95% indicate different species.
[b]*In silico* DNA-DNA hybridization (DDH) values <70% indicate different species (17).

ACKNOWLEDGMENTS

This study was supported by various PhD studentships, funded by UKRI and independently. A.J.B acknowledges support from a UKRI Innovate UK Knowledge Transfer Partnership (KTP010966) with Volac International Ltd., G.W. from a Biotechnology and Biological Sciences Research Council (BBSRC) grant BB/S007652/1, and A.J.M. and E.C.-O. from a BBSRC Southwest doctoral training partnership (BB/M009122/1). All analysis was done using the Medical Research Council (MRC) Cloud Infrastructure for Microbial Bioinformatics (MR/L015080/1). *Pseudomonas* sp. strain PP3 is available on request from the corresponding author.

AUTHOR AFFILIATIONS

[1]Microbiomes, Microbes and Informatics Group, Organisms and Environment Division, School of Biosciences, Cardiff University, Cardiff, Wales, United Kingdom
[2]Department of Chemistry, University of Warwick, Coventry, England, United Kingdom

PRESENT ADDRESS

Edward Cunningham-Oakes, Department of Infection Biology and Microbiomes, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, England, United Kingdom
Rachel Dodds, All Wales Medical Genomics Service, Wales Genomic Health Centre, Cardiff Edge Business Park, Cardiff, Wales, United Kingdom
Katja E. Hill, School of Dentistry, Cardiff University, Cardiff, Wales, United Kingdom
Mark J. Leggett, Volac International Ltd., Port Talbot, Wales, United Kingdom
Andrew W. Topping, FUJIFILM Diosynth Biotechnologies, Billingham, England, United Kingdom

AUTHOR ORCIDs

Gordon Webster http://orcid.org/0000-0002-9530-7835
Amy J. Baldwin http://orcid.org/0000-0002-2162-3771
Edward Cunningham-Oakes http://orcid.org/0000-0003-0260-5508
Alex J. Mullins http://orcid.org/0000-0001-5804-9008
Katja E. Hill http://orcid.org/0000-0002-8590-0117
Mark J. Leggett http://orcid.org/0009-0009-6689-1284
Andrew J. Weightman http://orcid.org/0000-0002-6671-2209

FUNDING

Funder	Grant(s)	Author(s)
Biotechnology and Biological Sciences Research Council	BB/S007652/1	Gordon Webster
UK Research and Innovation	KTP010966	Amy J. Baldwin
Biotechnology and Biological Sciences Research Council	BB/M009122/1	Edward Cunningham-Oakes Alex J. Mullins

AUTHOR CONTRIBUTIONS

Gordon Webster, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft | Amy J. Baldwin, Investigation, Methodology, Software, Validation, Writing – review and editing | Edward Cunningham-Oakes, Investigation, Methodology, Software, Validation, Writing – review and editing | Alex J. Mullins, Investigation, Methodology, Software, Validation, Visualization, Writing – review and editing | Rachel Dodds, Investigation, Methodology, Validation, Writing – review and editing | Katja E. Hill, Investigation, Validation | Li Ling Lee,

Investigation, Validation | Mark J. Leggett, Investigation, Validation, Writing – review and editing | Andrew W. Topping, Investigation, Validation | Andrew J. Weightman, Conceptualization, Project administration, Resources, Supervision, Validation, Writing – original draft

DATA AVAILABILITY

The genome sequences and Illumina raw sequence reads were deposited via the European Nucleotide Archive (ENA) under the ENA project/study number PRJEB43414. The accession numbers for the PP3 genome assembly and raw reads are GCA_905336995 and ERX5224659, ERX5224704, ERX5225337 and ERX5225391, respectively.

REFERENCES

1. Chaudhry GR, Chapalamadugu S. 1991. Biodegradation of halogenated organic compounds. Microbiol Rev 55:59–79. https://doi.org/10.1128/mr.55.1.59-79.1991
2. Pieper DH, Reineke W. 2004. Degradation of chloroaromatics by Pseudomona(d)s, p 509–574. In Ramos JL (ed), *Pseudomonas*. Vol. 3. Biosynthesis of macromolecules and molecular metabolism. Springer US, Boston, MA.
3. Senior E, Bull AT, Slater JH. 1976. Enzyme evolution in a microbial community growing on the herbicide Dalapon. Nature 263:476–479. https://doi.org/10.1038/263476a0
4. Slater JH, Lovatt D, Weightman AJ, Senior E, Bull AT. 1979. The growth of *Pseudomonas putida* on chlorinated aliphatic acids and its dehalogenase activity. J Gen Microbiol 114:125–136. https://doi.org/10.1099/00221287-114-1-125
5. Weightman AJ, Slater JH, Bull AT. 1979. The partial purification of two dehalogenases from *Pseudomonas putida* PP3. FEMS Microbiol Lett 6:231–234. https://doi.org/10.1016/0378-1097(79)90067-3
6. 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
7. 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
8. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0
9. Hesse C, Schulz F, Bull CT, Shaffer BT, Yan Q, Shapiro N, Hassan KA, Varghese N, Elbourne LDH, Paulsen IT, Kyrpides N, Woyke T, Loper JE. 2018. Genome-based evolutionary history of *Pseudomonas* spp. Environ Microbiol 20:2142–2159. https://doi.org/10.1111/1462-2920.14130
10. 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
11. Stothard P, Grant JR, Van Domselaar G. 2019. Visualizing and comparing circular genomes using the CGView family of tools. Brief Bioinformatics 20:1576–1582. https://doi.org/10.1093/bib/bbx081
12. Hill KE, Marchesi JR, Weightman AJ. 1999. Investigation of two evolutionarily unrelated halocarboxylic acid dehalogenase gene families. J Bacteriol 181:2535–2547. https://doi.org/10.1128/JB.181.8.2535-2547.1999
13. Weightman AJ, Topping AW, Hill KE, Lee LL, Sakai K, Slater JH, Thomas AW. 2002. Transposition of DEH, a broad-host-range transposon flanked by ISPpu12, in *Pseudomonas putida* is associated with genomic rearrangements and dehalogenase gene silencing. J Bacteriol 184:6581–6591. https://doi.org/10.1128/JB.184.23.6581-6591.2002
14. Gilchrist CLM, Chooi YH. 2021. Clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics 37:2473–2475. https://doi.org/10.1093/bioinformatics/btab007
15. 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
16. Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. Mol Biol Evol 32:2798–2800. https://doi.org/10.1093/molbev/msv150
17. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0
18. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods 8:12–24. https://doi.org/10.1039/C5AY02550H