



# Aerobiology from an Inactive Pyrite Mine: the Genome Sequence of the Airborne *Pseudomonas* sp. Strain L5B5

 Jose L. Gonzalez-Pimentel,<sup>a</sup> Irene Domínguez-Moñino,<sup>b</sup> Valme Jurado,<sup>b</sup> Ana Teresa Caldeira,<sup>a</sup> Cesareo Saiz-Jimenez<sup>b</sup>

<sup>a</sup>HERCULES Laboratory, Évora University, Évora, Portugal

<sup>b</sup>Instituto de Recursos Naturales y Agrobiología de Sevilla (IRNAS-CSIC), Seville, Spain

**ABSTRACT** *Pseudomonas* sp. strain L5B5 is an antimicrobial-producing bacterium isolated from an air sample collected in a pyrite mine in Lousal, Portugal. Genomic analyses predicted genes involved in virulence factors. Here, we report the complete genome sequence of this bacterium, which consists of a circular chromosome with a length of 6,811,662 bp.

The genus *Pseudomonas* is a very abundant group found in antagonistic environments; some interact with plants, some are useful as growth-promoting and biocontrol agents, and some are parasites, such as *Pseudomonas aeruginosa* and *Pseudomonas syringae*, two opportunistic pathogens for plants and animals (1). *Pseudomonas* are also found in subterranean environments, including mines, where some of them are able to remove heavy metals, which is useful for bioremediation (2, 3).

Mining is one of the oldest ways of producing and extracting minerals, and it has been intrinsically linked to the economic and social evolution of human beings. Lousal Mine (Grândola, Portugal) was closed in 1988 because the mined ores were no longer viable economically. Afterwards, within the environmental, social, economic, and heritage context, the mining area was remediated and rehabilitated in 2010, a geo-tourism center created, and a mining gallery opened to visitors (4).

The airborne *Pseudomonas* sp. strain L5B5 was isolated using a surface air system (Duo SAS, model Super 360). The sample was managed as described by Porca et al. (5). The culture medium was Trypticase-soy-agar (TSA; BD), and the culture plate was incubated at 28°C after picking a single colony. DNA extraction was performed using the Canvax HigherPurity bacterial genomic DNA isolation kit (Córdoba, Spain) with RNase treatment.

Genomic DNA was sequenced using both the NovaSeq 6000 (Illumina, USA) and SMRT RS II (PacBio, Menlo Park, CA) platforms. A TruSeq DNA PCR-free library was constructed using 150-bp paired-end short reads (Illumina); an SMRT library (insert size, 20 kb), constructed following the instructions for the Pacific Biosciences SMRTbell prep kit, was used for *de novo* sequencing. Default parameters were used for all software unless otherwise specified. The long reads were assembled using Canu v2.1.1 (6) and Circlator v1.5.5 (7) to circularize and orientate the chromosome. Pilon v1.24 (8) was employed for assembly improvement. Annotation was carried out using the NCBI Prokaryotic Genome Annotation Pipeline (9) and Sma3s v2 (10) with the “uniprot” flag. antiSMASH v6.0 (11), with the detection parameters in “strict” mode and all extra features on, was used to predict and annotate the secondary metabolite biosynthesis gene clusters. The Virulence Factors of Pathogenic Bacteria Database (VFDB), through Vfanalyzer, was used for virulence factors prediction (12). The Comprehensive Antibiotic Resistance Database (CARD), through Resistance Gene Identifier (RGI) software, was launched for prediction of the resistome (13). The closest relatives of L5B5 were identified using the JSpecies Web

**Editor** David A. Baltrus, University of Arizona

**Copyright** © 2021 Gonzalez-Pimentel et al.

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

Address correspondence to Jose L. Gonzalez-Pimentel, [pimentel@irnas.csic.es](mailto:pimentel@irnas.csic.es).

The authors declare no conflict of interest.

**Received** 19 October 2021

**Accepted** 16 November 2021

**Published** 2 December 2021

**TABLE 1** Genome features and comparison between strain L5B5 and related bacteria

Strain	GenBank accession no.	GC content (%)	Genome size (bp)	Total no. of genes	No. of antiSMASH-predicted clusters	No. of VFDB mechanisms	No. of CARD hits
L5B5	<a href="#">CP084742</a>	63.2	6,811,662	6,265	20	185	7 (strict)
<i>P. aestus</i> CMAA1215	<a href="#">AVOY00000000.1</a>	63.8	6,658,235	6,380	19	169	4 (strict)
<i>P. protegens</i> Cab57	<a href="#">AP014522</a>	63.34	6,827,892	6,236	15	175	4 (strict)
<i>P. protegens</i> Pf-5	<a href="#">CP000076</a>	63.30	7,074,893	6,392	16	178	4 (strict)
<i>P. protegens</i> CHA0 <sup>T</sup>	<a href="#">CP003190</a>	63.39	6,867,980	6,252	15	178	4 (strict)

tool (14), and the average nucleotide identity was calculated using the BLAST (ANIb) and MUMmer (ANIm) algorithms.

The PacBio sequencing generated a total of 140,791 subreads, averaging 10,080 bp long, for a genome coverage of >208×, whereas the Illumina sequencing produced a total of 19,767,134 reads for a genome coverage of >435×. A tetra correlation search against the JSpecies database linked the L5B5 genome with the genomes of the type species *Pseudomonas aestus* CMM1215 and *P. protegens* CHA0 and with the biotechnologically useful strains *P. protegens* Cab57 and *P. protegens* Pf-5. The ANIb and ANIm values were below the 95% threshold suggested for species differentiation.

Table 1 provides a comparison of the relationship between L5B5 and related strains, as well as their functional characterization. L5B5 showed a more dynamic secondary metabolism, as well as an increased presence of virulence factors and genes involved in the resistome, with respect to the related genomes.

**Data availability.** The whole-genome shotgun project for *Pseudomonas* sp. L5B5 has been deposited at DDBJ/ENA/GenBank under the accession number listed in Table 1. The version described in this paper is the first version. The BioProject accession number is [PRJNA769239](#), and the raw data have been deposited in the Sequence Read Archive (SRA) under accession numbers [SRR16308509](#) and [SRR16308510](#), for the long and short reads, respectively.

## ACKNOWLEDGMENT

Financial support was obtained through project 0483\_PROBIOMA\_5\_E, cofinanced by the European Regional Development Fund within the framework of the INTERREG VA Spain-Portugal program (POCTEP), 2014 to 2020.

## REFERENCES

- Sitaraman R. 2015. *Pseudomonas* spp. as models for plant-microbe interactions. *Front Plant Sci* 6:787. <https://doi.org/10.3389/fpls.2015.00787>.
- Saxena D, Joshi N, Srivastav S. 2002. Mechanism of copper resistance in a copper mine isolate *Pseudomonas putida* strain S4. *Curr Microbiol* 45: 410–414. <https://doi.org/10.1007/s00284-002-3787-5>.
- Palanivel TM, Sivakumar N, Al-Ansari A, Victor R. 2020. Bioremediation of copper by active cells of *Pseudomonas stutzeri* LA3 isolated from an abandoned copper mine soil. *J Environ Manage* 253:109706. <https://doi.org/10.1016/j.jenvman.2019.109706>.
- Relvas JMRS, Pinto A, Fernandes C, Matos JX, Vieira A, Mendonça A, Malha C, Albuquerque F, Alegre L, Abrunhosa M, Pinheiro M, Oliveira M, Alves M, Ferreira M, Rufino R, Pratas S, Ferreira T. 2014. Lousal: an old mine, a recent dream, a new reality. *Com Geol* 101:1345–1347.
- Porca E, Jurado V, Martin-Sanchez PM, Hermosin B, Bastian F, Alabouvette C, Saiz-Jimenez C. 2011. Aerobiology: an ecological indicator for early detection and control of fungal outbreaks in caves. *Ecol Indic* 11:1594–1598. <https://doi.org/10.1016/j.ecolind.2011.04.003>.
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
- Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SM. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
- 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>.
- 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>.
- Muñoz-Mérida A, Viguera E, Claros MG, Trelles O, Pérez-Pulido AJ. 2014. Sma3s: a three-step modular annotator for large sequence datasets. *DNA Res* 21:341–353. <https://doi.org/10.1093/dnares/dsu001>.
- Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87. <https://doi.org/10.1093/nar/gkz310>.
- Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive Web interface. *Nucleic Acids Res* 47:D687–D692. <https://doi.org/10.1093/nar/gky1080>.

13. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen AV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, McArthur AG. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>.
14. Richter M, Rosselló-Mora R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>.