Draft Genome Sequence of *Pseudomonas koreensis* Strain AB36, Isolated from Gold Mining Soil

Olubukola Oluranti Babalola,[a] Ayansina Segun Ayangbenro[a]

aFood Security and Safety Niche, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa

ABSTRACT Here, we report the draft genome sequence of *Pseudomonas koreensis* strain AB36, isolated from gold mining soil in South Africa. The draft sequence consists of 5,902,614 bp, with a G+C content of 60.1% and 5,242 protein-coding genes. The genome provides insight into its metabolism and the degradation of environmental pollutants.

One of the most diverse bacterial genera is *Pseudomonas*. These bacterial strains have varied metabolic properties and are widely distributed in the ecosystem (1). The genus has been divided into two intrageneric groups (IGs) (IG *P. aeruginosa* and IG *P. fluorescens*) based on multilocus sequence analysis of the 16S rRNA, *gyrB*, *rpoB*, and *rpoD* genes (2). IG *P. fluorescens* is divided into six groups which are represented by the species *P. fluorescens*, *P. syringae*, *P. lutea*, *P. putida*, *P. anguilliseptica*, and *P. straminea*. *Pseudomonas koreensis* belongs to the *P. fluorescens* subdivision (2, 3).

Pseudomonas koreensis strain AB36 was isolated from gold-contaminated soil samples, as described by Babalola et al. (4). The genomic DNA of strain AB36 was extracted from pure culture grown on LB agar using a ZR soil microbe miniprep DNA extraction kit (Zymo Research, CA, USA), according to the kit's protocol. The DNA was quantified with a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific, CA, USA) and checked for quality on a 1% agarose gel stained with ethidium bromide. The genome of strain AB36 was sequenced on an Illumina MiSeq sequencer at Molecular Research (MR DNA), Shallowater, TX. The library was constructed using a Nextera DNA sample preparation kit (Illumina), following the manufacturer's instructions. After the library preparation, the final concentration of the library was measured using the Qubit double-stranded DNA (dsDNA) high-sensitivity (HS) assay kit (Life Technologies, Inc.). The average library size was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies). The library was pooled and diluted to 12.0 pM and paired-end sequenced using a 600-cycle v.3 reagent kit (Illumina), with an average coverage of 50×.

The read quality of the raw sequence was assessed with FastQC (v.1.0.4) (5), and the reads were trimmed to remove adapter sequences, low-quality sequences, and ambiguous reads using Trimmomatic (v.0.36) (6), with default settings. *De novo* assembly was performed using SPAdes (v.3.12.0) (7), with default parameters, to yield 44 contigs. The assembled genome of strain AB36 generates a total size of 5,902,614 bp, with a G+C content of 60.1%. The contigs have N_{50} and L_{50} values of 419,669 bp and 5, respectively. The contigs were annotated with the Rapid Annotations using Subsystems Technology (RAST; v.2.0) server (8) and the NCBI Prokaryotic Genome Annotation Pipeline (PGAP; v.4.7) (9), which identified 5,242 coding genes, 71 RNA genes, 60 tRNA genes, and 159 pseudogenes.

The annotated draft genome showed the presence of genes responsible for metal resistance, such as arsenic resistance protein (ArsH), cobalt-zinc-cadmium resistance protein (CzcD), and chromate transport protein (ChrA). The annotated genome also carries genes for aromatic hydrocarbon (benzoate, phenylalkanoic acid, quinate, and

heterocyclic aromatic hydrocarbon) degradation. The biosynthetic gene cluster (BGC) analysis was carried out using antiSMASH (v.3.0) (10), and seven BGCs were predicted. Most BGCs are predicted to be nonribosomal peptide synthetase clusters, which include known clusters such as pyoverdine, jagaricin, and mangotoxin clusters. Other BGCs are for aryl polyene, bacteriocin, and beta-lactone. The draft genome of *P. koreensis* provides a wealth of information about the aromatic hydrocarbon degradative genes and metal resistance genes that are of environmental importance.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number SEUB00000000. The version described in this paper is the first version, SEUB01000000. The Sequence Read Archive (SRA) accession number is SRR8529805.

REFERENCES

1. Jeong H, Rha E, Kim H, Lee S-G. 2018. Complete genome sequence of the soil bacterium *Pseudomonas kribbensis* strain 46-2ᵀ. Microbiol Resour Announc 7:e01161-18. https://doi.org/10.1128/MRA.01161-18.
2. Mulet M, Lalucat J, García-Valdés E. 2010. DNA sequence-based analysis of the *Pseudomonas* species. Environ Microbiol 12:1513–1530. https://doi.org/10.1111/j.1462-2920.2010.02181.x.
3. Lozano GL, Bravo JI, Handelsman J. 2017. Draft genome sequence of *Pseudomonas koreensis* CI12, a *Bacillus cereus* "hitchhiker" from the soybean rhizosphere. Genome Announc 5:e00570-17. https://doi.org/10.1128/genomeA.00570-17.
4. Babalola OO, Aremu BR, Ayangbenro AS. 2019. Draft genome sequence of heavy metal-resistant *Bacillus cereus* NWUAB01. Microbiol Resour Announc 8:e01706-18. https://doi.org/10.1128/MRA.01706-18.
5. Andrews S. 2011. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
7. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clin-genpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-meta-genomes from chimeric MDA products. J Comput Biol 20:714–737. https://doi.org/10.1089/cmb.2013.0084.
8. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
9. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Res 46:D851–D860. https://doi.org/10.1093/nar/gkx1068.
10. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.