



**HAL**  
open science

## **Complete Genome Sequence of the Polymyxin E (Colistin)-Producing *Paenibacillus* sp. Strain B-LR**

Caroline Salsano, Sandrine Didelot, Fatoumata Tambaou, Thibault Caradec, Valérie Sopena, Cyrille Barthélémy, Valérie Thiéry, Romain Chevrot

### ► **To cite this version:**

Caroline Salsano, Sandrine Didelot, Fatoumata Tambaou, Thibault Caradec, Valérie Sopena, et al.. Complete Genome Sequence of the Polymyxin E (Colistin)-Producing *Paenibacillus* sp. Strain B-LR. *Microbiology Resource Announcements*, 2018, 7 (23), <10.1128/MRA.01013-18>. <hal-02383554>

**HAL Id: hal-02383554**

**<https://hal.science/hal-02383554v1>**

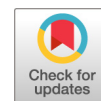
Submitted on 21 Jul 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



# Complete Genome Sequence of the Polymyxin E (Colistin)-Producing *Paenibacillus* sp. Strain B-LR

Caroline Salsano,<sup>a</sup> Sandrine Didelot,<sup>a</sup> Fatoumata Tambadou,<sup>a</sup> Thibault Caradec,<sup>b</sup> Valérie Sopéna,<sup>a</sup> Cyrille Barthélémy,<sup>a</sup> Valérie Thiéry,<sup>a</sup> Romain Chevrot<sup>a</sup>

<sup>a</sup>Université de La Rochelle, Laboratoire Littoral Environnement et Sociétés (LIENSs)–UMR 7266–CNRS, La Rochelle, France

<sup>b</sup>Institut Pasteur de Lille, Lille, France

**ABSTRACT** *Paenibacillus* bacteria are recovered from varied niches, including human lung, rhizosphere, marine sediments, and hemolymph. *Paenibacilli* can have plant growth-promoting activities and be antibiotic producers. They can produce exopolysaccharides and enzymes of industrial interest. Illumina and PacBio reads were used to produce a complete genome sequence of the colistin producer *Paenibacillus* sp. strain B-LR.

*Paenibacillus* bacteria are well-known producers of antimicrobial compounds, such as peptides, nonvolatile organic compounds, and enzymes (1). *Paenibacilli* can produce nonribosomal peptide (NRP) antibiotics, such as polymyxins, fusaricidins, and depsipeptides (2, 3). *Paenibacillus polymyxa* strains are used by pharmaceutical corporations for their bioproduction of polymyxins B and E (colistin), targeting *Pseudomonas aeruginosa* infections. *Paenibacillus* sp. strain B-LR was isolated from an environmental sample during a screening of antagonistic bacteria with antagonism against *P. aeruginosa* PA01 (3). Curiously, even though *Paenibacillus* sp. strain B-LR is related to *Paenibacillus alvei* and not to *P. polymyxa*, it is also a polymyxin E producer.

Preliminary sequencing work was done using an Illumina MiSeq (2 × 150-bp) platform (Eurofins Genomics, Germany). Genomic DNA was extracted using the NucleoSpin tissue kit (Macherey-Nagel, Germany) on overnight cultures grown at 30°C in Luria-Bertani broth (Sigma, France) (3). A total of 2,829,200 paired-end plus 3,042,337 singleton reads with an average of 103 bp were obtained using a long jumping distance library with an insert size of 8 kbp. Unfortunately, the obtained *de novo* genome assembly presented issues, especially within the gene cluster responsible for the production of the polymyxin E that we had already sequenced (3). The repetitive regions within the nonribosomal peptide synthetase (NRPS) genes can be misassembled when using short-read sequencing. Therefore, a second sequencing run was performed using PacBio RS II (Pacific Biosciences, USA) by Eurofins Genomics, Germany. Genomic DNA was extracted as described for the Illumina sequencing. Quality controls of the DNA were obtained from a Quantus fluorometer (Promega, USA) combined with NanoDrop 2000 measurements (Thermo Fisher Scientific, USA) and low-melt agarose gel electrophoresis. SMRTbell libraries were prepared with Blue Pippin size selection. A total of 121,905 reads with a mean length of 8,255 bp were obtained. With a genome size of 6.7 Mb, PacBio single-molecule real-time (SMRT) sequencing provided approximately 150× coverage. They were *de novo* assembled into a single contig using Hierarchical Genome Assembly Process 4 (HGAP4) (4). The MiSeq reads were mapped onto the PacBio contig using Mira v4.0 with default parameters in order to correct sequence errors (5, 6). The genome sequence was annotated using the Microbial Genome Annotation and Analysis Platform (MAGE) (7). The *Paenibacillus* sp. strain B-LR circular genome is 6,694,765 bp long with a 46.76% GC content and contains 6,129

Received 3 September 2018 Accepted 15 November 2018 Published 13 December 2018

**Citation** Salsano C, Didelot S, Tambadou F, Caradec T, Sopéna V, Barthélémy C, Thiéry V, Chevrot R. 2018. Complete genome sequence of the polymyxin E (colistin)-producing *Paenibacillus* sp. strain B-LR. Microbiol Resour Announc 7:e01013-18. <https://doi.org/10.1128/MRA.01013-18>.

**Editor** Catherine Putonti, Loyola University Chicago

**Copyright** © 2018 Salsano et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Romain Chevrot, [romain.chevrot01@univ-lr.fr](mailto:romain.chevrot01@univ-lr.fr).

C.S. and S.D. contributed equally to the work.

coding sequences (CDSs), 81 tRNAs, and 30 rRNAs. The sequence was submitted to antiSMASH analysis of secondary metabolite biosynthesis gene clusters (8). Among the 42 clusters identified, 18 were only putative, whereas 25 were assigned to functions. Eight were involved in NRP biosynthesis, whereas four were involved in hybrid polyketides/NRP biosynthesis. The other clusters were involved in the biosynthesis of 6 saccharides, 4 fatty acids, bacteriocin, lasso peptide, and resorcinol. Among the NRPS clusters, the gene cluster responsible for biosynthesis of polymyxin E was recovered and shown to be fully identical to the one we had previously described (3). The other NRPS clusters shared low gene similarity percentages with those found in the antiSMASH database, which could suggest their novelty.

**Data availability.** This complete genome sequence project has been deposited at GenBank/ENA/DBJ under the accession number [LS992241](https://doi.org/10.1093/nar/gky379). The raw sequencing reads are deposited at GenBank under the accession number [SRR8135148](https://doi.org/10.1093/nar/gky379).

## ACKNOWLEDGMENTS

This work was supported by L'Association Vaincre la Mucoviscidose. The Ph.D. grant to C.S. was given by the Nouvelle Aquitaine region.

## REFERENCES

1. Grady EN, MacDonald J, Liu L, Richman A, Yuan ZC. 2016. Current knowledge and perspectives of *Paenibacillus*: a review. *Microb Cell Fact* 15: 8–18. <https://doi.org/10.1186/s12934-016-0603-7>.
2. Chevrot R, Didelot S, Van den Bossche L, Tambaou F, Caradec T, Marchand P, Izquierdo E, Sopéna V, Caillon J, Barthélémy C, Van Schepdael A, Hoogmartens J, Rosenfeld E. 2013. A novel depsipeptide produced by *Paenibacillus alvei* 32 isolated from a cystic fibrosis patient. *Probiotics Antimicrob Proteins* 5:18–25. <https://doi.org/10.1007/s12602-012-9121-z>.
3. Tambaou F, Caradec T, Gagez AL, Bonnet A, Sopéna V, Bridiau N, Thiéry V, Didelot S, Barthélémy C, Chevrot R. 2015. Characterization of the colistin (polymyxin E1 and E2) biosynthetic gene cluster. *Arch Microbiol* 197:521–532. <https://doi.org/10.1007/s00203-015-1084-5>.
4. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
5. Chevreux B, Wetter T, Suhai S. 1999. Genome sequence assembly using trace signals and additional sequence information. *J Comput Sci Syst Biol* 99:45–46.
6. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, Chilton J, Clements D, Coraor N, Grüning B, Guerler A, Hillman-Jackson J, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 46:537–544. <https://doi.org/10.1093/nar/gky379>.
7. Vallenet D, Engelen S, Mornico D, Cruveiller S, Fleury L, Lajus A, Rouy Z, Roche D, Salvignol G, Scarpelli C, Médigue C. 2009. MicroScope: a platform for microbial genome annotation and comparative genomics. *Database* 2009:bap021. <https://doi.org/10.1093/database/bap021>.
8. Weber T, Blin K, Duddela S, Krug D, Kim H, Brucocoleri R, Lee S, 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* 43:237–243. <https://doi.org/10.1093/nar/gkv437>.