



**HAL**  
open science

## Complete genome sequences of two *Bacillus thuringiensis* serovar *kurstaki* strains isolated from Lebanon and Tunisia, highly toxic against lepidopteran larvae

Nancy Fayad, Rita Barssoum, Nathalie Marsaud, Rayan Nasserredine, Nouha Abdelmalek, Souad Rouis, Marie Ange Teste, Vincent Pailler, Veronique Gautier, Elodie Belmonte, et al.

### ► To cite this version:

Nancy Fayad, Rita Barssoum, Nathalie Marsaud, Rayan Nasserredine, Nouha Abdelmalek, et al.. Complete genome sequences of two *Bacillus thuringiensis* serovar *kurstaki* strains isolated from Lebanon and Tunisia, highly toxic against lepidopteran larvae. *Microbiology Resource Announcements*, 2023, 12 (9), pp.00060-23. 10.1128/MRA.00060-23 . hal-04287518

**HAL Id: hal-04287518**

**<https://hal.science/hal-04287518>**

Submitted on 17 May 2024

**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 Attribution 4.0 International License

# Complete genome sequences of two *Bacillus thuringiensis* serovar *kurstaki* strains isolated from Lebanon and Tunisia, highly toxic against lepidopteran larvae

Nancy Fayad,<sup>1,2</sup> Rita Barssoum,<sup>1,3</sup> Nathalie Marsaud,<sup>3,4</sup> Rayan Nasserredine,<sup>1</sup> Nouha Abdelmalek,<sup>5</sup> Souad Rouis,<sup>6</sup> Marie Ange Teste,<sup>4</sup> Vincent Pailier,<sup>7</sup> Veronique Gautier,<sup>7</sup> Elodie Belmonte,<sup>7</sup> César Arturo Aceves Lara,<sup>3</sup> Julien Cescut,<sup>8</sup> Luc Fillaudeau,<sup>3</sup> Mireille Kallassy Awad<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 3.

**ABSTRACT** *Bacillus thuringiensis*-based products are key in the biopesticides market. *Bacillus thuringiensis* *kurstaki* strains Lip and BLB1 were isolated from Lebanese and Tunisian soils, respectively. These strains are highly toxic against lepidopteran larvae, *Ephestia kuehniella*. Here, we report Lip and BLB1 complete genomes, including their plasmid and toxin contents.

**KEYWORDS** *Bacillus thuringiensis*, genomes, whole-genome sequencing, PacBio

*Bacillus thuringiensis* serovar *kurstaki* (*Btk*) strains Lip and BLB1 were isolated from Lebanese and Tunisia soil samples, respectively (1–3). Both strains showed an increased toxicity against the lepidopteran *Ephestia kuehniella* larvae in comparison to the reference strain HD-1. Their entomopathogenic potential was evaluated by *in vivo* toxicity assays and showed a lethal concentration of 50% of the larvae (LC50) of 33.27 and 70 ng of toxin per milligram of flour for Lip and BLB1, respectively (1–3). In the context of IPM-4-Citrus (MSCA RISE, No. 734921, 2017–2023), a project aiming to optimize the culture of these strains on a wheat bran-based medium, a whole-genome sequencing (WGS) approach was adopted to elucidate all genomic aspects of *Btk* Lip and BLB1.

For WGS, strains were grown in liquid Luria-Bertani (LB) medium for 16 h at 30°C, after which DNA was extracted using the Monarch HMW DNA Extraction Kit as per manufacturer instructions. WGS was conducted using a PacBio Sequel II Sequencer (Pacific Biosciences, Menlo Park, CA, USA) (4, 5) at the Gentyane platform (Clermont-Ferrand, France). For PacBio sequencing, library was prepared using a SMRTbell prep kit3, following manufacturer instructions. Genomic DNA was sheared, cleaned of single-strand overhangs, repaired for damage and A A tailed. Then, barcoded overhang adapters were ligated to generate the SMRTBell templates. Fragments above 5 kb were then size selected with 35% AMPure PB Beads. A Fragment Analyzer (Agilent Technologies) and a Qubit fluorimeter (Life Technologies) allowed quantity and quality checks. A ready-to-sequence SMRTBell Polymerase Complex was created using a Binding Kit 3.2 (PacBio) and the Sequel II primer 3.2. A consensus sequencing (CCS) mode was adopted, and reads were later demultiplexed with Bam2fastx software (Bioconda) under default parameters.

The number of CCS reads was 524,791 for Lip and 374,689 for BLB1 (Table 1). Assembly was done with flye v2.5, under default parameters (6). Overlaps were manually removed. The final coverage was 417.3x and 746.6x for Lip and BLB1, respectively. Genome completeness was assessed by BUSCO v5.0, using default parameters (7), and was found to be at 99.78% for Lip and 96% for BLB1. Circular contigs were

**Editor** J. Cameron Thrash, University of Southern California, Los Angeles, California, USA

Address correspondence to Mireille Kallassy Awad, mireille.kallassy@usj.edu.lb.

The authors declare no conflict of interest.

See the funding table on p. 3.

**Received** 3 February 2023

**Accepted** 22 June 2023

**Published** 8 August 2023

Copyright © 2023 Fayad 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/).

TABLE 1 Sequence features and accession numbers of replicons from *Btk* strains Lip and BLB1

	Replicon	Bioproject/biosample/NCBI SRA accession number	Accession number	Length (bp)	CDS # (total)	# Cry coding genes	GC content (%)	Number of PacBio CCS reads/reads N50
Lip	Chromosome		CP116313	5,293,947	5,632		35.2	524,791/9,650
	pLip2189		CP116316	2,189	4		35.3	
	pLip7635		CP116320	7,635	8		32.2	
	pLip7911		CP116321	7,911	9		32.3	
	pLip8513		CP116322	8,513	9		30.8	
	pLip12		CP116314	12,276	21		31.1	
	pLip15		CP116315	15,008	28		34.9	
	pLip69	PRJNA924104/ SAMN32746260/	CP116319	69,004	76		32.2	
	pLip91	SRX20261752	CP116323	91,357	102		31.2	
	pLip97		CP116324	97,437	85	2: Cry1Ab; Cry1Ac	34.5	
	pLip300		CP116317	300,451	267	5: Cry1Aa; Cry1Ac; Cry1Ia; Cry2Aa; Cry2Ab;	33.1	
	pLip457		CP116318	457,481	408		32.7	
BLB1	Chromosome	PRJNA924104/	CP116325	5,677,911	6,060		35.3	374,689/8,967
	pBLB1_7792	SAMN32746260/	CP116333	7,792	9		32.3	
	pBLB1_8398	SRX20261751	CP116335	8,398	10		29.8	
	pBLB1_8548		CP116336	8,548	8		30.8	
	pBLB1_12		CP116326	11,521	9	1: Cry1Ab	40.3	
	pBLB1_14		CP116327	14,174	21		32.5	
	pBLB1_15		CP116328	15,071	22		31.2	
	pBLB1_48		CP116330	47,643	73		35.4	
	pBLB1_56		CP116331	56,399	62		32.2	
	pBLB1_74		CP116332	73,558	88		30.6	
	pBLB1_81		CP116334	80,698	86		33.5	
	pBLB1_317		CP116329	317,321	282	5: Cry1Aa; Cry1Ac; Cry1Ia; Cry2Aa; Cry2Ab;	33.2	

first highlighted by the Flye assembler repeat graphs and then further identified as plasmids via a database similarity search using a nucleotide BLAST+ (2.12.0+ [8]) executable *blastn* command. This allowed to search for identified plasmids against the complete *Bacillus thuringiensis* genomes from NCBI assembly (NCBI:txid1428). Additional checks of completeness, identity, and circularity of plasmids were done by multiple sequence alignment comparison with several reference *Btk* strains: HD-1 (ASM71753v1 [9]), YBT-1520 (ASM74754v1), and HD73 (ASM33875v1 [10]).

Each strain carried 11 different plasmids ranging between 2 and 457 kb. pLip15 was identified as a linear tectiviral prophage, quite similar to GIL16, a tectivirus also isolated from a *Bacillus thuringiensis* strain (11).

Genome annotation was done with NCBI's automated annotation pipeline Prokaryotic Genome Annotation Pipeline (PGAP) (12). Toxin content was mined using t.BLAST.n with the toxin protein sequences recovered from the Bacterial Pesticidal Protein Resource Center (<https://www.bpprc-db.org> [13]). Genes encoding toxins from the Cry1 and Cry2 families were detected in Lip on plasmids pLip300 and pLip97, and BLB1 on pBLB1\_317 and pBLB1\_12 (Table 1).

## ACKNOWLEDGMENTS

This research was funded by the research council of Saint-Joseph University of Beirut CR-USJ under grant number FS65 and by the European Union's Horizon 2020 funds: H2020-MSCA-RISE-2016 #734921. M.K.A. mobility (Lebanon–France) was funded by Allocation de perfectionnement à la formation-recherche, November 2021, from the Agence Universitaire de la Francophonie.

## AUTHOR AFFILIATIONS

<sup>1</sup>Laboratory of Biodiversity and Functional Genomics, Université Saint-Joseph de Beyrouth, Beirut, Lebanon

<sup>2</sup>Multi-Omics Laboratory, School of Pharmacy, Lebanese American University, Byblos, Lebanon

<sup>3</sup>Toulouse Biotechnology Institute, Toulouse, France

<sup>4</sup>GenoToul GeT-BioPUCE, Toulouse, France

<sup>5</sup>Laboratoires Pharmaceutiques MédiS, Tunis, Tunisia

<sup>6</sup>Laboratory of Biopesticides, Centre of Biotechnology of Sfax, University of Sfax, Sfax, Tunisia

<sup>7</sup>Gentyane, GDEC, INRAE, Clermont-Ferrand, France

<sup>8</sup>Toulouse White Biotechnology, Toulouse, France

## AUTHOR ORCID*s*

Nancy Fayad  <http://orcid.org/0000-0002-8791-490X>

Mireille Kallassy Awad  <http://orcid.org/0000-0003-1489-8686>

## FUNDING

Funder	Grant(s)	Author(s)
<a href="#">Saint Joseph University (USJ)</a>	FS65	Mireille Kallassy Awad
<a href="#">EC   Horizon 2020 Framework Programme (H2020)</a>	H2020-MSCA-RISE-2016 #734921	Rita Barssoum Rayan Nasserredine Nouha Abdelmalek Souad Rouis César Arturo Aceves Lara Julien Cescut Luc Fillaudeau Mireille Kallassy Awad
<a href="#">Agence Universitaire de la Francophonie (AUF)</a>	Allocation de perfectionnement à la formation-recherche November 2021	Mireille Kallassy Awad

## AUTHOR CONTRIBUTIONS

Nancy Fayad, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Rita Barssoum, Formal analysis, Methodology, Writing – review and editing | Nathalie Marsaud, Formal analysis, Methodology, Writing – review and editing | Rayan Nasserredine, Methodology, Writing – review and editing, Resources | Nouha Abdelmalek, Investigation, Methodology, Writing – review and editing | Souad Rouis, Writing – review and editing, Resources, Funding acquisition | Marie Ange Teste, Investigation, Writing – review and editing | Vincent Pailler, Methodology, Software, Writing – review and editing, Resources | Veronique Gautier, Methodology, Writing – review and editing, Resources | Elodie Belmonte, Investigation, Writing – review and editing, Resources | César Arturo Aceves Lara, Investigation, Writing – review and editing, Funding acquisition | Julien

Cescut, Investigation, Writing – review and editing, Funding acquisition | Luc Fillaudeau, Writing – review and editing, Resources, Funding acquisition, Project administration | Mireille Kallassy Awad, Conceptualization, Investigation, Writing – review and editing, Resources, Funding acquisition, Project administration, Supervision

## DATA AVAILABILITY

Whole-genome sequences for *Bacillus thuringiensis* serovar *kurstaki* strains Lip and BLB1 are available in NCBI's GenBank under the bioproject number [PRJNA924104](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA924104) for both strains and the accession numbers [CP116313](https://www.ncbi.nlm.nih.gov/nuccore/CP116313) to [CP116324](https://www.ncbi.nlm.nih.gov/nuccore/CP116324) for Lip and [CP116325](https://www.ncbi.nlm.nih.gov/nuccore/CP116325) to [CP116336](https://www.ncbi.nlm.nih.gov/nuccore/CP116336) for BLB1. Raw reads are also available on the NCBI SRA database under the bioproject number [PRJNA924104](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA924104).

## REFERENCES

1. Saadaoui I, Rouis S, Jaoua S. 2009. A new Tunisian strain of *Bacillus thuringiensis kurstaki* having high Insecticidal activity and  $\delta$ -endotoxin yield. Arch Microbiol 191:341–348. <https://doi.org/10.1007/s00203-009-0458-y>
2. El Khoury M, Azzouz H, Chavanieu A, Abdelmalak N, Chopineau J, Awad MK. 2014. Isolation and characterization of a new *Bacillus thuringiensis* strain lip harboring a new cry1Aa gene highly toxic to *Ephesia kuehniella* (Lepidoptera: Pyralidae) larvae. Arch Microbiol 196:435–444. <https://doi.org/10.1007/s00203-014-0981-3>
3. Fayad N, Abboud J, Driss F, Louka N, Kallassy Awad M. 2022. Optimization of culture conditions and wheat bran class selection in the production of *Bacillus thuringiensis*-based biopesticides. Fermentation 8:666. <https://doi.org/10.3390/fermentation8120666>
4. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. Genomics Proteomics Bioinformatics 13:278–289. <https://doi.org/10.1016/j.gpb.2015.08.002>
5. Kanwar N, Blanco C, Chen IA, Seelig B. 2021. PacBio sequencing output increased through uniform and directional fivefold concatenation. Sci Rep 11:18065. <https://doi.org/10.1038/s41598-021-96829-z>
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. Seppely M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. Methods Mol Biol 1962:227–245. [https://doi.org/10.1007/978-1-4939-9173-0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14)
8. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. <https://doi.org/10.1186/1471-2105-10-421>
9. Day M, Ibrahim M, Dyer D, Bulla L. 2014. Genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD-1. Genome Announc 2:e00613-14. <https://doi.org/10.1128/genomeA.00613-14>
10. Liu G, Song L, Shu C, Wang P, Deng C, Peng Q, Lereclus D, Wang X, Huang D, Zhang J, Song F. 2013. Complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73. Genome Announc 1:e0008013. <https://doi.org/10.1128/genomeA.00080-13>
11. Verheust C, Fornelos N, Mahillon J. 2005. GIL16, a new gram-positive tectiviral phage related to the *Bacillus thuringiensis* GIL01 and the *Bacillus cereus* pBClin15 elements. J Bacteriol 187:1966–1973. <https://doi.org/10.1128/JB.187.6.1966-1973.2005>
12. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. Nucleic Acids Res 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>
13. Crickmore N, Berry C, Panneerselvam S, Mishra R, Connor TR, Bonning BC. 2021. A structure-based nomenclature for *Bacillus thuringiensis* and other bacteria-derived pesticidal proteins. J Invertebr Pathol 186:107438. <https://doi.org/10.1016/j.jip.2020.107438>