

1  
2 **Complete genome sequences of two *Bacillus thuringiensis* serovar *kurstaki***  
3 **strains isolated from Lebanon and Tunisia, highly toxic against lepidopteran**  
4 **larvae**

5 Nancy FAYAD<sup>a,b</sup>, Rita BARSSOUM<sup>a,c</sup>, Nathalie MARSAUD<sup>c,d</sup>, Rayan NASSEREDINE<sup>a</sup>, Nouha  
6 ABDELMALEK<sup>e</sup>, Souad ROUIS<sup>f</sup>, Marie Ange TESTE<sup>d</sup>, Vincent PAILLER<sup>g</sup>, Veronique GAUTIER<sup>g</sup>,  
7 Elodie BELMONTE<sup>g</sup>, César Arturo ACEVES LARA<sup>c</sup>, Julien CESCUT<sup>h</sup>, Luc FILLAUDEAU<sup>c</sup>, Mireille  
8 KALLASSY AWAD<sup>a#</sup>

- 9 a. Laboratory of Biodiversity and Functional Genomics, UR-EGP, Faculty of Science, Université  
10 Saint-Joseph de Beyrouth, B.P. 1514—Riad El Solh, Beirut 1107 2050, Lebanon  
11 b. Multi-Omics Laboratory, School of Pharmacy, Lebanese American University, Lebanon  
12 c. Toulouse Biotechnology Institute, CNRS UMR5504, INRAE UMR792, INSA Toulouse, 135 Av. de  
13 Rangueil, 31077 Toulouse, France  
14 d. GenoToul GeT-BioPUCE, CNRS UMR5504, INRAE UMR792, INSA Toulouse, 135 Av. de Rangueil,  
15 31077 Toulouse, France  
16 e. Laboratoires Pharmaceutiques MédiS, B.P 206 Nabeul 8000, Tunisia  
17 f. Laboratory of Biopesticides, Centre of Biotechnology of Sfax, University of Sfax, P.O. Box 1177,  
18 3018 Sfax, Tunisia.  
19 g. Gentyane, GDEC, INRAE UMR1095, 5 chemin de beaulieu, 63000 Clermont-Ferrand  
20 h. Toulouse White Biotechnology, INRAE UMS1337, CNRS UMS3582, INSA, 135 Av. de Rangueil,  
21 31077 Toulouse, France

22 **Running title:** *Bacillus thuringiensis* *kurstaki* Lip and BLB1 genomes.

23 # Address correspondence to Mireille Kallassy Awad, [mireille.kallassy@usj.edu.lb](mailto:mireille.kallassy@usj.edu.lb).

## 24 Abstract

25 *Bacillus thuringiensis*-based products are key in the biopesticides market. *Bacillus thuringiensis* *kurstaki*  
26 strains Lip and BLB1 were isolated from Lebanese and Tunisian soils, respectively. These strains  
27 exhibited a high efficiency against lepidopteran larvae *Ephestia kuehniella*. Here, we report Lip and BLB1  
28 complete genomes, including their plasmid and toxin contents.

## 29 Announcement

30 *Bacillus thuringiensis* serovar *kurstaki* (*Btk*) strains Lip and BLB1 were isolated from Lebanese and Tunisia  
31 soil samples, respectively (1-3). Both strains showed an increased toxicity against the lepidopteran  
32 *Ephestia kuehniella* larvae, in comparison to the reference strain HD-1. In the context of IPM-4-Citrus  
33 (MSCA RISE, No. 734921, 2017-2023), a project aiming to optimize the culture of these strains on a

34 wheat bran based medium, a whole genome sequencing (WGS) approach was adopted to elucidate all  
35 genomic aspects of *Btk* Lip and BLB1.

36 For WGS, strains were grown in liquid LB medium for 16 hours at 30°C, after which DNA was extracted  
37 using the Monarch® HMW DNA Extraction Kit as per manufacturer instructions. WGS was conducted  
38 using a PacBio circular consensus sequencing (CCS) with Illumina MiSeq polishing (4-5) at the Gentyane  
39 platform at INRAE, Toulouse – France. On average for both strains, the number of CCS reads was  
40 2,806,648, and the N50 reads was 9,589 nucleotides with over 99.97 and 32.61% of reads larger than  
41 1,000 and 10,000 nucleotides. CCS lengths ranged between 97 and 40,567 nucleotides. After  
42 demultiplexing, 524,791 and 374,689 CCS reads were generated for Lip and BLB1, respectively. Assembly  
43 was done with flye v2.5 (6). Genome completeness was assessed using BUSCO v5.0 (7) and was found to  
44 be at 99.78% for Lip and 96% for BLB1. The GC content of Lip and BLB1 chromosomes was at 35.2 and  
45 35.3%, respectively. That of the plasmids ranged between 29.8 and 40.3 % (Table 1). The completeness  
46 and circularity of chromosomes and relevant plasmids was verified by multiple sequence alignment  
47 comparison with reference *Btk* strains HD-1 (NCBI assembly #: ASM71753v1; 8), YBT-1520 (NCBI  
48 assembly #: ASM74754v1) and HD73 (NCBI assembly #: ASM33875v1; 9).

49 Each strain carried 11 different plasmids, with sizes ranging between 2 and 457 kb for Lip, and 7.7 and  
50 317 kb for BLB1. Plasmid topology was circular, except for pLip15 who was identified as a linear tectiviral  
51 prophage, quite similar to GIL16, a tectivirus also isolated from a *Bacillus thuringiensis* strain (10).

52 Lip and BLB1 complete genomes were annotated with RAST (11) and NCBI's automated annotation  
53 pipeline PGAP (12), the toxin content was mined using a t.BLAST.n approach with the toxin protein  
54 sequences recovered from the Bacterial Pesticidal Protein Resource Center (<https://www.bpprc-db.org>;  
55 13). Cut-off was set at 70% query coverage and 45% identity. Genes encoding toxins from the Cry1 and  
56 Cry2 families were detected in both Lip and BLB1 (Table 1). In Lip, pLip300 and pLip97 were highlighted  
57 as the toxin-carrying plasmids, with five and two *cry* genes, respectively. As for BLB1, the toxin-carrying  
58 plasmids are pBLB1\_317 and pBLB1\_12.

59

60

61 **Table 1.** Sequence features and accession numbers of replicons from *Btk* strains Lip (A) and BLB1 (B).

<b>A. Lip</b>						
Replicon	Bioproject / Biosample	Accession #	Length (bp)	CDS # (total)	# Cry coding genes	GC content (%)
Chromosome		<a href="#">CP116313</a>	5,293,947	5,632		35.2
pLip2189		<a href="#">CP116316</a>	2,189	4		35.3
pLip7635		<a href="#">CP116320</a>	7,635	8		32.2
pLip7911		<a href="#">CP116321</a>	7,911	9		32.3
pLip8513		<a href="#">CP116322</a>	8,513	9		30.8
pLip12	PRJNA924104 / SAMN32746259	<a href="#">CP116314</a>	12,276	21		31.1
pLip15		<a href="#">CP116315</a>	15,008	28		34.9
pLip69		<a href="#">CP116319</a>	69,004	76		32.2
pLip91		<a href="#">CP116323</a>	91,357	102		31.2
pLip97		<a href="#">CP116324</a>	97,437	85	2: Cry1Ab; Cry1Ac	34.5
pLip300		<a href="#">CP116317</a>	300,451	267	5: Cry1Aa; Cry1Ac; Cry1Ia; Cry2Aa; Cry2Ab;	33.1
pLip457		<a href="#">CP116318</a>	457,481	408		32.7
<b>B. BLB1</b>						
Chromosome		<a href="#">CP116325</a>	5,677,911	6,060		35.3
pBLB1_7792		<a href="#">CP116333</a>	7,792	9		32.3
pBLB1_8398		<a href="#">CP116335</a>	8,398	10		29.8
pBLB1_8548		<a href="#">CP116336</a>	8,548	8		30.8
pBLB1_12	PRJNA924104 / SAMN32746260	<a href="#">CP116326</a>	11,521	9	1: Cry1Ab	40.3
pBLB1_14		<a href="#">CP116327</a>	14,174	21		32.5
pBLB1_15		<a href="#">CP116328</a>	15,071	22		31.2
pBLB1_48		<a href="#">CP116330</a>	47,643	73		35.4
pBLB1_56		<a href="#">CP116331</a>	56,399	62		32.2
pBLB1_74		<a href="#">CP116332</a>	73,558	88		30.6
pBLB1_81		<a href="#">CP116334</a>	80,698	86		33.5
pBLB1_317	<a href="#">CP116329</a>	317,321	282	5: Cry1Aa; Cry1Ac; Cry1Ia; Cry2Aa; Cry2Ab;	33.2	

62

### 63 Data availability statement

64 Whole genome sequences for *Bacillus thuringiensis* serovar *kurstaki* strains Lip and BLB1 are available in  
 65 NCBI's GenBank under the bioproject number PRJNA924104 for both strains, and the accession numbers  
 66 CP116313 - CP116324 for Lip and CP116325 - CP116336 for BLB1.

## 67 Acknowledgments

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

## 72 References

- 73 1. Saadaoui I., Rouis S., Jaoua S. 2009. A new Tunisian strain of *Bacillus thuringiensis kurstaki*  
74 having high insecticidal activity and  $\delta$ -endotoxin yield. Arch Microbiol;191(4):341–8  
75 <https://doi.org/10.1007/s00203-009-0458-y>
- 76 2. El khoury M., Azzouz H., Chavanieu A., Abdelmalak N., Chopineau J., Kallassy Awad M. 2014.  
77 Isolation and characterization of a new *Bacillus thuringiensis* strain Lip harboring a new  
78 cry1Aa gene highly toxic to *Ephestia kuehniella* (Lepidoptera: Pyralidae) larvae. Arch  
79 Microbiol.;196(6):435–45. <https://doi.org/10.1007/s00203-014-0981-3>
- 80 3. Fayad, N.; Abboud, J.; Driss, F.; Louka, N.; Kallassy Awad, M. 2022. Optimization of Culture  
81 Conditions and Wheat Bran Class Selection in the Production of *Bacillus thuringiensis*-Based  
82 Biopesticides. Fermentation, 8, 666. <https://doi.org/10.3390/fermentation8120666>
- 83 4. Rhoads A, Au KF. 2015. PacBio Sequencing and Its Applications. Genomics Proteomics  
84 Bioinformatics. Oct;13(5):278-89. doi: 10.1016/j.gpb.2015.08.002. Epub 2015 Nov 2.
- 85 5. Kanwar, N., Blanco, C., Chen, I.A., Seelig, B. 2021 PacBio sequencing output increased through  
86 uniform and directional fivefold concatenation. Sci Rep 11, 18065.  
87 <https://doi.org/10.1038/s41598-021-96829-z>
- 88 6. Kolmogorov, M., Yuan, J., Lin, Y., Pevzner, P.A. 2019. Assembly of long, error-prone reads using  
89 repeat graphs. Nat Biotechnol 37, 540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- 90 7. Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation  
91 completeness. Methods Mol Biol 1962:227–245. [https://doi.org/10.1007/978-1-4939-9173-](https://doi.org/10.1007/978-1-4939-9173-0_14)  
92 [0\\_14](https://doi.org/10.1007/978-1-4939-9173-0_14).
- 93 8. Day M, Ibrahim M, Dyer D, Bulla L Jr. 2014 Genome Sequence of *Bacillus thuringiensis* subsp.  
94 *kurstaki* Strain HD-1. Genome Announc. Jul 17;2(4):e00613-14. doi:  
95 10.1128/genomeA.00613-14.
- 96 9. Liu G, Song L, Shu C, Wang P, Deng C, Peng Q, Lereclus D, Wang X, Huang D, Zhang J, Song F.  
97 2013. Complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73.  
98 Genome Announc. Mar 14;1(2):e0008013. doi: 10.1128/genomeA.00080-13.
- 99 10. Verheust C, Fornelos N, Mahillon J. 2005. GIL16, a new gram-positive tectiviral phage related to  
100 the *Bacillus thuringiensis* GIL01 and the *Bacillus cereus* pBClin15 elements. J Bacteriol.  
101 Mar;187(6):1966-73. doi: 10.1128/JB.187.6.1966-1973.2005.
- 102 11. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch  
103 GD, Shukla M, Thomason JA 3rd, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a  
104 modular and extensible implementation of the RAST algorithm for building custom  
105 annotation pipelines and annotating batches of genomes. Sci Rep. Feb 10;5:8365. doi:  
106 10.1038/srep08365.

- 107 12. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F,  
108 Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J,  
109 Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq:  
110 expanding the Prokaryotic Genome Annotation Pipeline reach with protein family model  
111 curation. *Nucleic Acids Res.* Jan 8;49(D1):D1020-D1028. doi: 10.1093/nar/gkaa1105.
- 112 13. Crickmore N, Berry C, Panneerselvam S, Mishra R, Connor TR, Bonning BC. 2021. A structure-  
113 based nomenclature for *Bacillus thuringiensis* and other bacteria-derived pesticidal proteins.  
114 *J Invertebr Pathol.* Nov;186:107438. doi: 10.1016/j.jip.2020.107438.
- 115