

## Complete Genome Sequence of *Bacillus thuringiensis* Strain RC340, Isolated from a Temperate Forest Soil Sample in New England

Brendan Sullivan<sup>a</sup>, Claire E. Kitzmiller<sup>a</sup>, Wyatt C. Tran<sup>a</sup>, Mallory Choudoir<sup>a,b</sup>, Rachel Simoes<sup>a</sup>, Nipuni Dayarathne<sup>a</sup>, Kristen M. DeAngelis<sup>a</sup>

<sup>a</sup> Department of Microbiology, University of Massachusetts, Amherst, Massachusetts, USA

<sup>b</sup> Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA

### ABSTRACT

The complete genome sequence of *Bacillus thuringiensis* strain RC340, isolated from an environmental microbiology experiment soil sample is presented here. *B. thuringiensis* strain RC340 sequenced by GridION consists of a single genome consisting of 5.86 million bases, 8,152 predicted genes, and 0.23% contamination.

### ANNOUNCEMENT

Soil microenvironments are highly competitive, and bacteria challenge one another for the limited resources available. *B. thuringiensis* has several mechanisms to gain a competitive advantage including bacteriocin production and antimicrobial resistance (AMR) proteins. Soil warming may induce the prevalence of AMR [\(1\)](#), so these gene signatures were the focus of our genomics analysis.

Soils from the Harvard Forest soil warming experiment were experimentally heated 5°C above ambient temperature with unheated controls [\(2\)](#). RC340 was isolated in 2022 from a heated soil plot, then grown on actinobacteria isolation agar (AIA) [\(3\)](#) with 100 mg of L-1 cycloheximide. The isolate was then grown on 10% tryptic soy broth at 30°C with shaking at 150 rpm until cultures reached an OD of 0.5, then spun at 4000 rpm for 15 minutes. Genomic DNA was extracted from cell pellets using the CTAB method [\(4\)](#). 16S rRNA gene was sequenced as previously described (REF) to estimate target coverage for the genome assembly. Genomes were sequenced by SeqCenter (Pittsburgh, PA) using GridION sequencing with nanopore R9 flow cells (R9.4.1) and GuPPy v5.0.16 for base calling (Sherathiya et al. 2021). DNA was not sheared or size-selected for ONT library preparation.

The genome was assembled, annotated, and analyzed as part of the Bioinformatics Lab (MICROBIO 590B) course at University of Massachusetts Amherst [\(5\)](#). All software was deployed using default parameters unless otherwise specified. Reads were filtered by FiltLong v0.2.1 [\(6\)](#) to retain ~40X coverage (228,000,000 bp), >85% mean quality score, and 1,000 base pairs minimum read length. Reads were assembled using the *de novo* Flye assembler v2.9.2 [\(7\)](#) producing 216,153 reads, then aligned and mapped with Minimap2 v2.26 [\(8\)](#). The draft assembly was corrected with Racon v1.4.3 [\(9\)](#) and polished with Medaka v1.8.1 [\(10\)](#) to generate a final consensus sequence assembly.

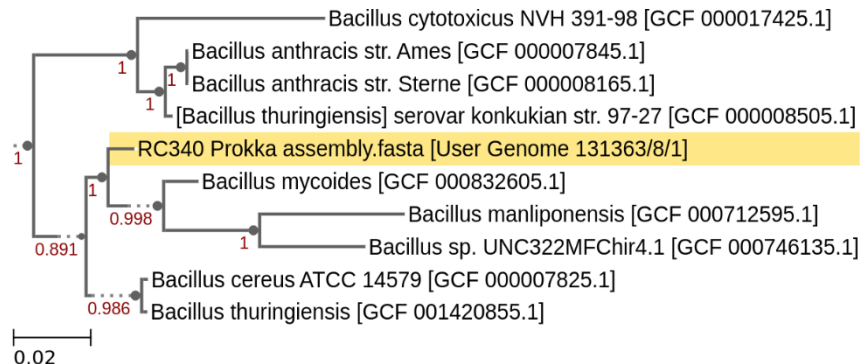
The final assembly was uploaded into KBase (11) and QUAST v4.4 (12) and CheckM v1.0.18 (13) determined the completeness and the contamination, concluding the assembly is high quality (14) (Table 1).

Features	<i>RC340</i>
Total base pairs in the assembly	229,161,969
Assembled genome size (bp)	5,868,817
Fold-coverage	39.05
Number of contigs	3
Largest contig (bp)	5,469,113
Assembly N50 (bp)	5,469,113
Assembly L50	1
G+C content (%)	34.99%
Completion (%)	95.91%
Contamination (%)	0.23%

**Table 1:** Genome assembly quality assessment values shown, data obtained through QUAST and CheckM. Quality assessment metrics are largest contig, completion, and contamination.

The final RC340 assembly file was annotated in KBase using Prokka v1.14.5 (15). A phylogenetic tree was constructed by Insert Genome Into SpeciesTree v2.2.0 (16) with a neighbor public genome count of 9 (Figure 1).

RC340\_Output\_New\_Tree: Species Tree generated by Species Tree Builder



**Figure 1:** Insert Genome Into SpeciesTree v2.2.0 program created tree with Prokka file as input and a neighbor public genome count of 9, closest phylogenetic neighbors shown above. *B. mycoides*, *B. cereus*, and *B. thuringiensis* are closest neighbors.

The taxonomy for RC340 from Classify Microbes with GTDB-Tk v1.7.0 (17) is Domain Bacteria, Phylum Bacillota, Class Bacilli, Order Bacillales, Family Bacillaceae, Genus *Bacillus*, Species *thuringiensis*. Compute ANI with FastANI v0.1.3 (18) was utilized to quantify the average nucleotide identity between RC340 and its nearest neighbor *B. thuringiensis*. The ANI estimate was 95.865% which classifies RC340 genome as the same species.

To assess the presence of bacitracin resistance genes in the RC340 genome and the 9 closest related neighbors, we used Compare Genomes from Pangenomes v0.0.7 (19). The pangenome had 6 total bacitracin resistance genes, 5 of which were only present in the RC340 genome. These were bacitracin export, resistance, and transport proteins. The *Bt* strain RC340 bacitracin resistance genes may provide a competitive advantage in the battle for nutrients in nutrient-depleted soils due to climate change stress (20, 21).

#### **Data availability.**

The 16S rRNA GenBank sequence accession number is [OQ534872](#). The raw sequence reads are available in GenBank under the BioProject accession number [PRJNA949990](#). The Sequence Read Archive (SRA) accession number is [SRX19731073](#), and the genome assembly accession number is [JARXOY000000000](#).

#### **ACKNOWLEDGMENTS**

This work was supported by the National Science Foundation CAREER award program under contract DEB-1749206 (K.M.D), as well as the NSF Long-Term Ecological Research (LTER) program grant 1237491 to Harvard Forest. Special thanks to Ashley Eng for her continued support during the Bioinformatics 590B course and throughout the entirety of this project.

#### **REFERENCES**

1. [MacFadden DR, McGough SF, Fisman D, Santillana M, Brownstein JS. 2018. Antibiotic resistance increases with local temperature. 6. Nat Clim Change 8:510–514.](#)
2. [Melillo JM, Frey SD, DeAngelis KM, Werner WJ, Bernard MJ, Bowles FP, Pold G, Knorr MA, Grandy AS. 2017. Long-term pattern and magnitude of soil carbon feedback to the climate system in a warming world. Science 358:101–105.](#)
3. [Olsen. 1960. Personal Communication.](#)
4. [William: Bacterial genomic DNA isolation using CTAB - Google Scholar.](#)  
[https://scholar.google.com/scholar\\_lookup?title=Bacterial+genomic+DNA+isolation+using+CTAB&publication\\_year=2012](https://scholar.google.com/scholar_lookup?title=Bacterial+genomic+DNA+isolation+using+CTAB&publication_year=2012). Retrieved 6 April 2023.

5. DeAngelis KM, Choudoir M, Eng A, Morrow MA. 2023. Bioinformatics Lab: A Course-Based Undergraduate Research Experience Curriculum <https://doi.org/10.7275/g0fg-5939>.
6. Wick R. 2023. rwick/Filtlong. C++.
7. Kolmogorov M. 2023. Flye assembler. C.
8. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100.
9. Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746.
10. 2023. Medaka. Python. Oxford Nanopore Technologies.
11. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. 7. Nat Biotechnol 36:566–569.
12. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.
13. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055.

14. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloie-Fadrosch EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu W-T, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Murat Eren A, Schriml L, Banfield JF, Hugenholtz P, Woyke T. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. 8. Nat Biotechnol 35:725–731.
15. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
16. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE 5:e9490.
17. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927.
18. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114.
19. Dejongh M. Compare Genomes from Pangenome | App Catalog | KBase. [https://narrative.kbase.us/#catalog/apps/GenomeComparisonSDK/compare\\_genomes/d1a44adac7a87e700b65df0c209aa8b9183cacf](https://narrative.kbase.us/#catalog/apps/GenomeComparisonSDK/compare_genomes/d1a44adac7a87e700b65df0c209aa8b9183cacf). Retrieved 24 February 2023.
20. Scofield V, Jacques SMS, Guimarães JRD, Farjalla VF. 2015. Potential changes in bacterial metabolism associated with increased water temperature and nutrient inputs in tropical humic lagoons. Front Microbiol 6:310.
21. Guoju X, Qiang Z, Jiangtao B, Fengju Z, Chengke L. 2020. The Relationship Between Winter Temperature Rise and Soil Fertility properties. Air Soil Water Res 5.