

Laboratory Misidentifications Resulting from Taxonomic Changes to *Bacillus cereus* Group Species, 2018–2022

Laura M. Carroll, Itumeleng Matle, Jasna Kovac, Rachel A. Cheng, Martin Wiedmann

Whole-genome sequencing (WGS) is being applied increasingly to *Bacillus cereus* group species; however, misinterpretation of WGS results may have severe consequences. We report 3 cases, 1 of which was an outbreak, in which misinterpretation of *B. cereus* group WGS results hindered communication within public health and industrial laboratories.

The *Bacillus cereus* group of bacteria is a complex of closely related species with varying pathogenic potential (1). Notable members, as defined in the US Food and Drug Administration's Bacteriological Analytical Manual, include *B. anthracis*, an etiologic agent of anthrax, and *B. cereus*, which has been associated with numerous illnesses, including foodborne emetic intoxication, diarrheal toxicoinfection, anthrax-like illness, and nongastrointestinal infections (1–5).

Whole-genome sequencing (WGS) is used increasingly in clinical and industrial microbiology laboratories to characterize *B. cereus* group strains (6). However, interpreting WGS results from these organisms is challenging; insights derived from WGS may conflict with information provided by traditional microbiologic assays (6–8). Previously, we hypothesized that results from some WGS-based species classification methods can be easily misinterpreted when applied to the *B. cereus* group (7). Here, we show that this scenario is not hypothetical: we report 3 recent cases among public health and industrial laboratories in which misinterpretation of WGS results directly hindered public health and food safety investigations or responses.

Author affiliations: EMBL, Heidelberg, Germany (L.M. Carroll); Onderstepoort Veterinary Research, Pretoria, South Africa (I. Matle); The Pennsylvania State University, University Park, Pennsylvania, USA (J. Kovac); Cornell University, Ithaca, New York, USA (R.A. Cheng, M. Wiedmann)

DOI: <https://doi.org/10.3201/eid2809.220293>

The Study

We report 3 cases of WGS-based *B. cereus* group species assignment misinterpretations in 3 continents: Europe, North America, and Africa (Table 1; Appendix, <https://wwwnc.cdc.gov/EID/article/28/9/22-0293-App1.pdf>). Two cases (cases 2 and 3) occurred within the last 6 months and involved regional and national public health laboratories; 1 case (case 1) involved an industrial laboratory (Table 1). All cases involved strains known colloquially as group III *B. cereus*, a phylogenetic lineage within the *B. cereus* group that was identified using pantoate- β -alanine ligase gene (*panC*) sequencing (9).

Case 1 occurred in November 2018 at an industrial microbiology laboratory in Europe (Table 1). The inquiring party isolated *B. cereus* group strains from a food processing facility, then characterized them by WGS (protocols unknown). Each *B. cereus* group strain was assigned to a species by comparing its genome to the genomes of all *B. cereus* group species type strains and identifying the most similar species type strain by average nucleotide identity (7). Two strains were classified as *B. anthracis* using this approach (Table 1), and the inquiring party was concerned that the strains represented a potential anthrax threat (because of their *B. anthracis* label). However, subsequent investigation by M.W. and L.M.C. revealed that, although the strains most closely resembled *B. anthracis*, neither strain belonged to the historical, clonal *B. anthracis* lineage typically associated with anthrax toxin production (6), and neither strain possessed anthrax toxin- or capsule-encoding genes (Table 1; Figure 1). Thus, the strains were deemed incapable of causing anthrax, despite being assigned to *B. anthracis* by WGS. The authors noted that historically, these strains would be known colloquially as group III *B. cereus*, using microbiologic methods and *panC* phylogenetic group assignment (Figure 2, panel A).

Table 1. Summary of cases of laboratory misidentifications caused by taxonomic changes to *Bacillus cereus* group species, 2018–2022*

Case	Date	Location	Inquiring party	WGS-assigned species of inquiry	Case summary†
1	November 2018	Europe	Industrial laboratory	<i>B. anthracis</i>	Two <i>B. cereus</i> strains isolated from a food processing facility were assigned to the <i>B. anthracis</i> species but were not closely related to the <i>B. anthracis</i> lineage most commonly responsible for anthrax illness and did not possess anthrax encoding genes or represent an anthrax threat. They would historically be referred to as <i>B. cereus</i> or group III <i>B. cereus</i> .
2	October 2021	North America (USA)	Government laboratory	<i>B. paranthracis</i>	A <i>B. cereus</i> strain isolated from a food product responsible for a foodborne outbreak was assigned to the <i>B. paranthracis</i> species using WGS-based methods. <i>B. paranthracis</i> was historically referred to as <i>B. cereus</i> or group III <i>B. cereus</i> and encompasses <i>B. cereus</i> group strains capable of causing emetic and/or diarrheal foodborne illness.
3	January 2022	Africa (South Africa)	Government laboratory	<i>B. anthracis</i>	Two <i>B. cereus</i> strains isolated during routine surveillance of meat products were classified using multiple WGS-based methods; they were assigned to the <i>B. anthracis</i> species but did not represent an anthrax threat. They would historically be referred to as <i>B. cereus</i> or group III <i>B. cereus</i> .

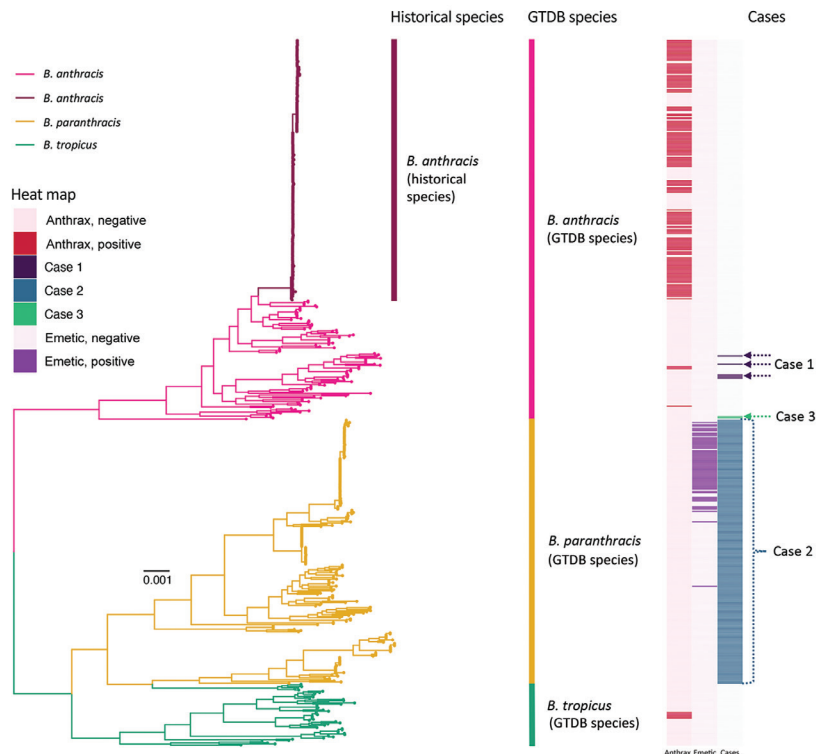
*WGS, whole genome sequencing; ANI, average nucleotide identity; MLST, multilocus sequence typing; ST, sequence type; Group III, *panC* phylogenetic Group III; PubMLST, <https://pubmlst.org>; GTDB, Genome Taxonomy Database Releases R95 and R202, <https://gtdb.ecogenomic.org>.

†*B. cereus* refers to the historical and/or colloquial species definition assigned using traditional microbiological methods, as outlined in the US Food and Drug Administration's Bacteriological Analytical Manual (FDA BAM) (5).

Case 2 occurred in October 2021 at a regional public health microbiology laboratory in the United States (Table 1). The inquiring party was responding to a foodborne outbreak that occurred at a correctional facility in Maryland in mid-September 2021. During the outbreak investigation, a *B. cereus* group

strain was isolated from rehydrated dehydrated potatoes using standard protocols (5). The strain underwent WGS and was classified as *B. paranthracis* (protocols unknown) (Table 1). The inquiring party had never heard of *B. paranthracis* before and conducted a literature search, noting that the species was first

Figure 1. Maximum-likelihood phylogeny constructed using core genes detected among 605 genomes assigned to the (GTDB) *B. anthracis*, *B. paranthracis*, and *B. tropicus* species. Branch colors and clade labels denote GTDB species assignments or, for *B. anthracis*, historical species assignments. The heatmap to the right of the phylogeny shows whether a strain possessed anthrax toxin-encoding genes or not (anthrax); whether a strain possessed cereulide synthetase (emetic toxin)-encoding genes or not (emetic); and the strains or lineages discussed in the cases we detailed here (cases) (Table 1). For case 1, the actual genomes were not publicly available; thus, genomes assigned to the same sequence types (STs, via 7-gene multilocus sequence typing) are highlighted. For case 2, the only information provided to the authors was that the genome in question belonged to species *B. paranthracis*; thus, all genomes assigned to GTDB's *B. paranthracis* species are highlighted. For case 3, the actual strain genomes associated with the case are highlighted. The phylogeny was rooted using *panC* Group II *B. cereus* group strain FSL W8-0169 as an outgroup (National Center for Biotechnology Information RefSeq Assembly accession no. GCF_001583695.1; omitted for readability). Branch lengths are reported in substitutions per site. GTDB, Genome Taxonomy Database.



described in 2017 (10); because of limited documented history of *B. paranthracis*, the inquiring party contacted M.W., J.K., and L.M.C. for assistance. We informed the inquiring party that *B. paranthracis* has historically been identified as group III *B. cereus* on the basis of microbiologic methods and *panC* phylogenetic group assignment. We also noted that *B. paranthracis* encompasses all strains known colloquially as emetic *B. cereus* (for their ability to produce cereulide, an emetic toxin) and some group III *B. cereus* strains capable of causing diarrheal foodborne illness (Figure 2, panel A) (11). We suggested that the inquiring party use multilocus sequence typing and virulence factor detection to determine if the strain belonged to a lineage previously associated with foodborne illness.

Case 3 occurred in January 2022 at a national veterinary public health microbiology laboratory in South Africa (Table 1). I.M. and collaborators isolated *B. cereus* group strains during routine surveillance of meat products (12). WGS was conducted on some strains (13). I.M. and L.M.C. assigned *B. cereus* group strains to species using multiple WGS-based methods (13); one method relied on the Genome Taxonomy Database (GTDB), a popular contemporary microbial species classification framework (14). GTDB releases R95 and R202 classified 2 strains as *B. anthracis* (Table 1); however, neither strain belonged to the historical, clonal *B. anthracis* lineage (6), and neither possessed anthrax toxin- or capsule-encoding genes (Table 1; Figure 1). Nevertheless, an inquiring party was concerned that the strains represented an anthrax threat because of the GTDB *B. anthracis* label (Table 1). We informed the inquiring party that neither possessed anthrax toxin-encoding genes. We noted that historically these strains would be known as group III *B. cereus*, using microbiologic methods and *panC* phylogenetic group assignment (Figure 2, panel A).

Conclusions

The growing popularity of WGS offers tremendous potential for improving *B. cereus* group surveillance, source tracking, and outbreak investigations. However, taxonomic issues in the *B. cereus* group have become more pronounced as researchers grapple with historical and WGS-based species definitions.

Here, we detailed 3 cases in which misinterpretation of *B. cereus* group WGS results directly hindered public health and food safety efforts. Two cases (cases 1 and 3) represented false-positive scenarios, in which group III *B. cereus* strains incapable of causing anthrax were incorrectly assumed to be anthrax-causing agents (Table 1). As noted previously,

strains that lack anthrax toxin-encoding genes but are assigned to *B. anthracis* using WGS-based methods are not uncommon (Table 2); these strains have been isolated from diverse environments (e.g., meat, milk, spices, egg whites, baby wipes) on 6 continents and the International Space Station, and although some may cause illness, they cannot cause anthrax (6). One way of denoting that a *B. cereus* group strain may produce anthrax toxin is to append the term “biovar Anthracis” to the genus/species name (Figure 2, panel B) (2).

The remaining case (case 2) represented a worst-case, false-negative scenario, in which a WGS-assigned species label with limited clinical interpretability or previous associations to foodborne illness

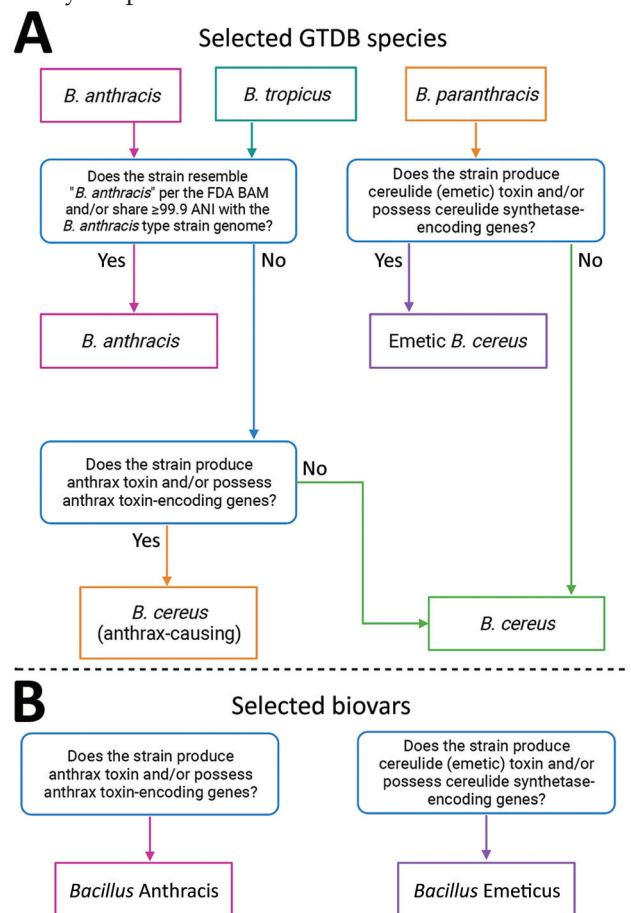


Figure 2. Flowcharts of *Bacillus* species and biovar name assignments. A) Flowchart depicting how 3 *Bacillus* species names assigned using GTDB releases R95 and R202 can be translated to historically important or colloquial names for *B. cereus* group species, as outlined in the US FDA’s BAM (5). B) Chart depicting how anthrax and cereulide (emetic) toxin-producing strains can be referred to using a previously proposed standardized collection of *B. cereus* group biovar terms (6). Figure was created using BioRender.com. ANI, average nucleotide identity; BAM, Bacteriological Analytical Manual; FDA, Food and Drug Administration; GTDB, Genome Taxonomy Database.

Table 2. Selected GTDB *Bacillus* species names and the clinically important strains they encompass*

GTDB species name	Encompasses strains which:			Notes†
	Can cause anthrax illness	Can cause emetic illness	Cannot cause anthrax or emetic illness	
<i>B. anthracis</i>	Yes	No	Yes	Encompasses all anthrax-causing <i>B. anthracis</i> strains, some anthrax-causing <i>B. cereus</i> strains, and many <i>B. cereus</i> strains that are incapable of causing anthrax illness but are commonly isolated from environmental and food sources (6,7).
<i>B. paranthracis</i>	No	Yes	Yes	Encompasses all cereulide-producing <i>B. cereus</i> strains known colloquially as emetic <i>B. cereus</i> , including the high-risk ST26 lineage; also encompasses many environmental and food isolates that are incapable of causing emetic illness (7,11).
<i>B. tropicus</i>	Yes	No	Yes	Encompasses some anthrax-causing <i>B. cereus</i> strains, as well as <i>B. cereus</i> strains that are incapable of causing anthrax illness (6,7).

*Obtained using GTDB Releases R95 and R202, but is applicable to any taxonomic framework, in which species names are assigned relative to *B. cereus* group species type strain genomes, e.g., by a species threshold of 95–96 average nucleotide identity or species threshold of 70% *in silico* DNA-DNA hybridization (7). GTDB, Genome Taxonomy Database; ST, sequence type assigned using the PubMLST 7-gene multilocus sequence typing scheme for *B. cereus* (<https://pubmlst.org>).

†*B. anthracis* and *B. cereus* refer to historical and/or colloquial species definitions assigned using traditional microbiological methods, as outlined in the US Food and Drug Administration's Bacteriological Analytical Manual (5).

(*B. paranthracis*) was assigned to an established pathogen (group III *B. cereus*) and directly hindered an outbreak investigation (Table 1). We anticipate that similar problems may arise with anthrax-causing *B. cereus*, because WGS-based methods assign some of these strains to *B. tropicus*, a species proposed in 2017 (Table 2) (7). We encourage readers to be mindful of this potential issue (Table 2). Overall, we hope that the cases we described can serve as cautionary tales for those who are transitioning to WGS for *B. cereus* group strain characterization.

This work was supported by United States Department of Agriculture National Institute of Food and Agriculture (USDA NIFA) Hatch Appropriations (project no. PEN04646; accession no. 1015787) and USDA NIFA grants (no. GRANT12686965 and 2019-67017-29591).

About the Author

Dr. Carroll is a computational biologist at the European Molecular Biology Laboratory in Heidelberg, Germany. Her research focuses on developing and utilizing bioinformatic approaches to study the evolution and transmission dynamics of foodborne and zoonotic pathogens. She has specific expertise in the genomics and taxonomy of the *Bacillus cereus* group.

References

1. Stenfors Arnesen LP, Fagerlund A, Granum PE. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev.* 2008;32:579–606. <https://doi.org/10.1111/j.1574-6976.2008.00112.x>
2. Baldwin VM. You can't *B. cereus*—a review of *Bacillus cereus* strains that cause anthrax-like disease. *Front Microbiol.* 2020;11:1731. <https://doi.org/10.3389/fmicb.2020.01731>
3. Bottone EJ. *Bacillus cereus*, a volatile human pathogen. *Clin Microbiol Rev.* 2010;23:382–98. <https://doi.org/10.1128/CMR.00073-09>
4. Jovanovic J, Ornelis VFM, Madder A, Rajkovic A. *Bacillus cereus* food intoxication and toxicoinfection. *Compr Rev Food Sci Food Saf.* 2021;20:3719–61. <https://doi.org/10.1111/1541-4337.12785>
5. Tallent SM, Knolhoff A, Rhodehamel EJ, Harmon SM, Bennett RW. *Bacillus cereus*. In: *Bacteriological analytical manual (BAM)*. 8th ed. Silver Spring (MD): Food and Drug Administration; 2019.
6. Carroll LM, Wiedmann M, Kovac J. Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. *MBio.* 2020;11:e00034–20. <https://doi.org/10.1128/mBio.00034-20>
7. Carroll LM, Cheng RA, Wiedmann M, Kovac J. Keeping up with the *Bacillus cereus* group: taxonomy through the genomics era and beyond. *Crit Rev Food Sci Nutr.* 2021;1–26. <https://doi.org/10.1080/10408398.2021.1916735>
8. Balloux F, Brønstad Brynildsrud O, van Dorp L, Shaw LP, Chen H, Harris KA, et al. From theory to practice: translating whole-genome sequencing (WGS) into the clinic. *Trends Microbiol.* 2018;26:1035–48. <https://doi.org/10.1016/j.tim.2018.08.004>
9. Guinebretière MH, Velge P, Couvert O, Carlin F, Debuyser ML, Nguyen-The C. Ability of *Bacillus cereus* group strains to cause food poisoning varies according to phylogenetic affiliation (groups I to VII) rather than species affiliation. *J Clin Microbiol.* 2010;48:3388–91. <https://doi.org/10.1128/JCM.00921-10>
10. Liu Y, Du J, Lai Q, Zeng R, Ye D, Xu J, et al. Proposal of nine novel species of the *Bacillus cereus* group. *Int J Syst Evol Microbiol.* 2017;67:2499–508. <https://doi.org/10.1099/ijsem.0.001821>
11. Carroll LM, Wiedmann M. Cereulide synthetase acquisition and loss events within the evolutionary history of group III *Bacillus cereus sensu lato* facilitate the transition between emetic and diarrheal foodborne pathogens. *MBio.* 2020;11:e01263-20. <https://doi.org/10.1128/mBio.01263-20>
12. Madoroba E, Magwedere K, Chaoira NS, Matle I, Muchadeyi F, Mathole MA, et al. Microbial communities of meat and meat products: an exploratory analysis of the

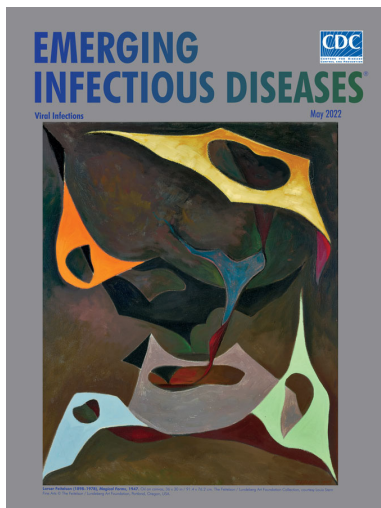
- product quality and safety at selected enterprises in South Africa. *Microorganisms*. 2021;9:507. <https://doi.org/10.3390/microorganisms9030507>
13. Carroll LM, Pierneef R, Mathole A, Atanda A, Matle I. Genomic surveillance of *Bacillus cereus sensu lato* strains isolated from meat and poultry products in South Africa enables inter- and intra-national surveillance and source tracking. *Microbiology Spectrum*. 2022;10:e00700-22. <https://doi.org/10.1128/spectrum.00700-22>
14. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. 2019;btz848. <https://doi.org/10.1093/bioinformatics/btz848>

Address for correspondence: Laura M. Carroll, EMBL, Meyerhofstraße 1, 69117 Heidelberg, Germany; email: lmc297@cornell.edu

May 2022

Viral Infections

- Invasive Group A *Streptococcus* Outbreaks Associated with Home Healthcare, England, 2018–2019
- Genomic Epidemiology of Global Carbapenemase-Producing *Escherichia coli*, 2015–2017
- Risk for Asymptomatic Household Transmission of *Clostridioides difficile* Infection Associated with Recently Hospitalized Family Members
- Estimating Relative Abundance of 2 SARS-CoV-2 Variants through Wastewater Surveillance at 2 Large Metropolitan Sites, United States
- Effectiveness of BNT162b2 Vaccine Booster against SARS-CoV-2 Infection and Breakthrough Complications, Israel
- Effects of Tick-Control Interventions on Tick Abundance, Human Encounters with Ticks, and Incidence of Tickborne Diseases in Residential Neighborhoods, New York, USA
- Pertactin-Deficient *Bordetella pertussis* with Unusual Mechanism of Pertactin Disruption, Spain, 1986–2018
- Determining Existing Human Population Immunity as Part of Assessing Influenza Pandemic Risk
- Disparities in First Dose COVID-19 Vaccination Coverage among Children 5–11 Years of Age, United States
- Multisystem Inflammatory Syndrome in Children after SARS-CoV-2 Vaccination
- Severe Multisystem Inflammatory Symptoms in 2 Adults after Short Interval between COVID-19 and Subsequent Vaccination



- Evidence of Prolonged Crimean-Congo Hemorrhagic Fever Virus Endemicity by Retrospective Serosurvey, Eastern Spain
- Lack of Evidence for Crimean–Congo Hemorrhagic Fever Virus in Ticks Collected from Animals, Corsica, France
- Highly Pathogenic Avian Influenza A(H5N8) Clade 2.3.4.4b Viruses in Satellite-Tracked Wild Ducks, Ningxia, China, 2020
- Novel Hendra Virus Variant Circulating in Black Flying Foxes and Grey-Headed Flying Foxes, Australia
- Increased COVID-19 Severity among Pregnant Patients Infected with SARS-CoV-2 Delta Variant, France
- Mathematical Modeling for Removing Border Entry and Quarantine Requirements for COVID-19, Vanuatu
- SARS-CoV-2 Seroprevalence after Third Wave of Infections, South Africa
- *Angiostrongylus cantonensis* in a Red Ruffed Lemur at a Zoo, Louisiana, USA
- Breast Milk as Route of Tick-Borne Encephalitis Virus Transmission from Mother to Infant
- *atpE* Mutation in *Mycobacterium tuberculosis* Not Always Predictive of Bedaquiline Treatment Failure
- Emerging Novel Reassortant Influenza A(H5N6) Viruses in Poultry and Humans, China, 2021
- *Mycobacterium lepromatosis* as Cause of Leprosy, Colombia
- Rare Case of Rickettsiosis Caused by *Rickettsia monacensis*, Portugal, 2021
- Pathogens that Cause Illness Clinically Indistinguishable from Lassa Fever, Nigeria, 2018
- Duration of Infectious Virus Shedding by SARS-CoV-2 Omicron Variant–Infected Vaccinees
- Imported Monkeypox from International Traveler, Maryland, USA, 2021
- Intercontinental Movement of Highly Pathogenic Avian Influenza A(H5N1) Clade 2.3.4.4 Virus to the United States, 2021
- Rapid Replacement of SARS-CoV-2 Variants by Delta and Subsequent Arrival of Omicron, Uganda, 2021
- SARS-CoV-2 Antibody Prevalence and Population-Based Death Rates, Greater Omdurman, Sudan
- Cross-Variant Neutralizing Serum Activity after SARS-CoV-2 Breakthrough Infections

**EMERGING
INFECTIOUS DISEASES**

To revisit the May 2022 issue, go to:
<https://wwwnc.cdc.gov/eid/articles/issue/28/5/table-of-contents>

Laboratory Misidentifications Resulting from Taxonomic Changes to *Bacillus cereus* Group Species, 2018–2022

Appendix

Phylogeny Construction for Figure 1

Genomic data acquisition and taxonomic assignment. A total of 2,887 *Bacillus cereus* group genomes were downloaded and pre-processed as described previously (1). Briefly, after assessing the quality of each genome, Prokka v1.14.6 (2) was used to annotate each genome. Each genome was additionally assigned to a species within the Genome Taxonomy Database (GTDB) using the classify_wf workflow in GTDB-Tk v1.3.0 and GTDB Release 05-RS95 (GTDB R95) (3–5).

Identification of the clonal *B. anthracis* lineage most commonly associated with anthrax toxin production. FastANI v1.31 (6) was used to calculate average nucleotide identity (ANI) values between each of the 2,887 *B. cereus* group genomes and the genome of *B. anthracis* str. Ames (NCBI RefSeq accession no. GCF_000007845.1); *B. cereus* group genomes that shared ≥ 99.9 ANI with the *B. anthracis* str. Ames genome were considered to belong to the classic clonal *B. anthracis* lineage most commonly associated with anthrax toxin production, as described previously (6,7). Within a recently proposed standardized nomenclatural framework for the *B. cereus* group (7) (referred to as the 2020 Genomospecies/Subspecies/Biovar or 2020 GSB taxonomy in a recent review of *B. cereus* group taxonomy) (8), this lineage is referred to as *B. mosaicus* subspecies *anthracis* (full notation) or *B. anthracis* (using shortened subspecies notation) (7–9).

***In silico* virulence factor detection, multilocus sequence typing, and *panC* phylogenetic group assignment.** BTyper3 v3.1.1 (9) was used to query each genome for (i) anthrax toxin encoding genes *cya*, *lef*, and *pagA* and (ii) cereulide (emetic toxin) synthetase-encoding genes *cesABCD* using BLAST v2.9.0 (10) and default settings. BTyper3 was

additionally used to perform (i) *in silico* seven-gene multilocus sequence typing (MLST) using the PubMLST *B. cereus* database (accessed 25 October 2020); and (ii) *panC* phylogenetic group assignment, using an adjusted eight-group (Group I-VIII) framework (9).

Phylogeny construction. GFF files produced by Prokka associated with all genomes assigned to GTDB's (i) *B. anthracis*, (ii) *B. paranthracis*, and (iii) *B. tropicus* species (n = 325, 226, and 53 genomes, respectively; see section "Genomic data acquisition and taxonomic assignment" above), plus (iv) *panC* Group II *B. cereus* group strain FSL W8–0169 (used as an outgroup; NCBI RefSeq Assembly Accession GCF_001583695.1) (11), were supplied as input to Panaroo v1.2.8 (12). Panaroo was used to identify "core" genes among all 605 genomes, using the following parameters (all other parameters were set to their default values): (i) "strict" mode (–clean-mode strict), (ii) core genome alignment using MAFFT (–a core–aligner mafft), (iii) a core genome sample threshold of 95% (–core_threshold 0.95). The resulting nucleotide alignment was queried using snp-sites v2.5.1 (13), which was used to identify (i) core single nucleotide polymorphisms (SNPs) and (ii) constant sites among all 605 genomes. The resulting core SNP alignment was supplied as input to IQ-TREE v1.5.4 (14), which was used to construct a maximum likelihood (ML) phylogeny using the General Time-Reversible (GTR) nucleotide substitution model (15), one thousand replicates of the ultrafast bootstrap approximation (16), and an ascertainment bias correction obtained using constant sites output by snp-sites. The resulting phylogeny was annotated and displayed using the bactaxR package (7) in R v4.1.2 (17).

Data availability. Metadata and quality information for all genomes displayed in Figure 1 of this study are available in Supplemental Tables S1–S4 of Carroll, et al., 2022 (1).

References

1. Carroll LM, Pierneef R, Mathole A, Atanda A, Matle I. Genomic surveillance of *Bacillus cereus sensu lato* strains isolated from meat and poultry products in South Africa enables inter- and intra-national surveillance and source tracking. *Microbiology Spectrum*. 2022;10: e00700--22 <https://doi.org/10.1128/spectrum.00700-22>
2. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30:2068–9. [PubMed <https://doi.org/10.1093/bioinformatics/btu153>](https://doi.org/10.1093/bioinformatics/btu153)

3. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*. 2019;btz848. [PubMed](#)
<https://doi.org/10.1093/bioinformatics/btz848>
4. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. 2018;36:996–1004. [PubMed](#) <https://doi.org/10.1038/nbt.4229>
5. Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol*. 2020;38:1079–86. [PubMed](#)
<https://doi.org/10.1038/s41587-020-0501-8>
6. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9:5114. [PubMed](#)
<https://doi.org/10.1038/s41467-018-07641-9>
7. Carroll LM, Wiedmann M, Kovac J. Proposal of a taxonomic nomenclature for the *Bacillus cereus* group which reconciles genomic definitions of bacterial species with clinical and industrial phenotypes. *MBio*. 2020;11:e00034–20. [PubMed](#) <https://doi.org/10.1128/mBio.00034-20>
8. Carroll LM, Cheng RA, Wiedmann M, Kovac J. Keeping up with the *Bacillus cereus* group: taxonomy through the genomics era and beyond. *Crit Rev Food Sci Nutr*. 2021;••:1–26. [PubMed](#)
<https://doi.org/10.1080/10408398.2021.1916735>
9. Carroll LM, Cheng RA, Kovac J. No assembly required: using BTyp3 to assess the congruency of a proposed taxonomic framework for the *Bacillus cereus* group with historical typing methods. *Frontiers in Microbiology*. 2020;11. <https://doi.org/10.3389/fmicb.2020.580691>
10. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421. [PubMed](#) <https://doi.org/10.1186/1471-2105-10-421>
11. Miller RA, Beno SM, Kent DJ, Carroll LM, Martin NH, Boor KJ, et al. *Bacillus wiedmannii* sp. nov., a psychrotolerant and cytotoxic *Bacillus cereus* group species isolated from dairy foods and dairy environments. *Int J Syst Evol Microbiol*. 2016;66:4744–53. [PubMed](#)
<https://doi.org/10.1099/ijsem.0.001421>
12. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol*. 2020;21:180. [PubMed](#)
<https://doi.org/10.1186/s13059-020-02090-4>

13. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, et al. *SNP-sites*: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom.* 2016;2:e000056. [PubMed https://doi.org/10.1099/mgen.0.000056](https://doi.org/10.1099/mgen.0.000056)
14. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32:268–74. [PubMed https://doi.org/10.1093/molbev/msu300](https://doi.org/10.1093/molbev/msu300)
15. Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lect Math Life Sci.* 1986;17:57–86.
16. Minh BQ, Nguyen MA, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 2013;30:1188–95. [PubMed https://doi.org/10.1093/molbev/mst024](https://doi.org/10.1093/molbev/mst024)
17. R Core Team. R: a language and environment for statistical computing. 4.1.2 ed. Vienna, Austria: R Foundation for Statistical Computing; 2019.