AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Draft Genome Sequences of Spacecraft-Associated Microbes Isolated from Six NASA Missions

**Michelle T. Tran,[a]* Arman Seuylemezian,[b] Alonna Wright,[c] David Coil,[c] Jonathan Eisen,[c] Lisa Guan[b]**

[a]California State Polytechnic University, Pomona, California, USA
[b]Biotechnology and Planetary Protection, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[c]Genome Center, University of California, Davis, California, USA

**ABSTRACT** Whole-genome sequencing can be used to better understand and assess the functional abilities of microorganisms isolated from spacecraft hardware and associated surfaces for planetary protection (PP) purposes. We sequenced 191 isolates from 6 spaceflight missions with PP requirements and identified them using Illumina-based sequencing methods and matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry.

The microbial archive at the Jet Propulsion Laboratory (JPL) contains over 8,000 bacterial strains previously isolated from missions with planetary protection (PP) requirements. The Genome Encyclopedia of Spacecraft Associated Microbes (GESAM) project has set out to perform a high-throughput identification of bacterial isolates spanning 6 past spaceflight missions from 1975 to 2012 (Viking, Mars Pathfinder, Mars Odyssey, Mars Exploration Rover [MER], Phoenix, and Mars Science Laboratory [MSL]). These PP-sensitive missions must minimize biological contamination on hardware surfaces to preserve the scientific integrity of mission objectives. The NASA standard assay was used throughout the mission timelines to assess hardware cleanliness. The assay involved heat-shocking samples at 80°C for 15 min, plating them on Trypticase soy agar (TSA), and incubating them for 72 h at 32°C (1). The resulting colonies were subcultured and stored long term in the JPL microbial archive following previously established procedures (1). In 2019, 1,418 stored isolates were revived from three replicate glycerol stocks and incubated for 72 h on TSA at 32°C. Of those, 1,249 strains displayed growth within 72 h and were inspected for colony morphology and visual purity across multiple stocks.

These 1,249 pure strains were screened using matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) and 16S rRNA sequencing for rapid, comprehensive identification filtering for the most abundant and potentially novel isolates. MALDI-TOF MS is a rapid and systematic identification method that uses the main spectrum profiles of isolates, composed of at least 10 spectra each. If MALDI-TOF MS did not match the isolate to a custom in-house MALDI-TOF MS database (2) (with at least 3 replicates with a log score of >2.2 for species level identification and >2.0 for genus level, for stringency and quality control), DNA was extracted, and the 16S rRNA gene was sequenced. DNA extraction was performed using the Maxwell rapid sample concentrator (RSC) automated DNA extraction instrument and the Maxwell RSC blood DNA extraction kit (Promega, USA). DNA was eluted in 10 mM Tris-HCl (pH 8.0) and quantified using the Quant-iT double-stranded DNA (dsDNA) high-sensitivity (HS) kit and Qubit fluorometer (Thermo Fisher, USA). The DNA was sent to MicrobesNG (UK) for library preparation and sequencing. 16S amplicon libraries were prepared using the Nextera XT library prep kit (Illumina, USA) using the manufacturer's protocol with the following modifications: (i) the input DNA was increased 2-fold, and (ii) the PCR elongation time was increased to

**Editor** Irene L. G. Newton, Indiana University, Bloomington
This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.
Address correspondence to Lisa Guan, lisa.guan@jpl.nasa.gov.
*Present address: Michelle T. Tran, University of California, Davis, California, USA.
The authors declare no conflict of interest.
**Received** 24 September 2022
**Accepted** 3 February 2023
**Published** 22 February 2023

**TABLE 1** Genome statistics for the 191 draft genomes isolated from 6 NASA mission spacecraft and associated surfaces[a]

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 41.1 | *Brevibacillus invocatus* | 2008-05-14 | JAMATA000000000 | SRX16346898 | 48.77 | 195 | 2,047,596 | 110,673 | 4,771 | 5, 0, 0 | 4,876,904 | 456 |
| 68.1 | *Bacillus safensis* | 2008-10-03 | JAMATB000000000 | SRX16346899 | 41.77 | 52 | 870,943 | 953,608 | 3,699 | 5, 1, 0 | 3,696,265 | 150 |
| 154.2 | *Bacillus subtilis* | 2008-08-08 | JAMATC000000000 | SRX16346910 | 44.73 | 427 | 1,356,939 | 1,071,018 | 4,462 | 8, 1, 0 | 4,283,513 | 348 |
| 183.1 | *Caldibacillus thermoamylovorans* | 2008-08-08 | JAMATD000000000 | SRX16346921 | 37.46 | 156 | 1,698,559 | 107,952 | 3,529 | 8, 0, 0 | 3,833,088 | 492 |
| 197.1 | *Staphylococcus warneri* | 2008-12-12 | JAMATE000000000 | SRX16346932 | 32.57 | 57 | 909,941 | 451,487 | 2,440 | 5, 1, 0 | 2,575,646 | 396 |
| 251.1 | *Bacillus altitudinis* | 2009-01-05 | JAMATH000000000 | SRX16346933 | 41.33 | 46 | 1,245,875 | 844,862 | 3,718 | 5, 1, 1 | 3,714,328 | 371 |
| 128.1.2 | *Mesobacillus maritimus* | 2009-01-15 | JAMBNU000000000 | SRX16346934 | 38.92 | 186 | 1,591,579 | 205,680 | 4,914 | 21, 0, 0 | 5,283,993 | 333 |
| 206.2.2 | *Bacillus amyloliquefaciens* | 2008-12-27 | JAMATF000000000 | SRX16346935 | 46.24 | 69 | 1,223,409 | 875,626 | 3,703 | 8, 0, 0 | 3,759,069 | 361 |
| 214.1.1 | *Alkalihalobacillus oceani* | 2009-01-03 | JAMBOL000000000 | SRX16346936 | 46.43 | 124 | 1,484,379 | 165,889 | 4,264 | 9, 0, 0 | 4,474,054 | 367 |
| 236.1.1 | *Bacillus altitudinis* | 2009-01-03 | JAMATG000000000 | SRX16346937 | 41.42 | 75 | 1,134,683 | 844,862 | 3,754 | 5, 1, 0 | 3,727,933 | 338 |
| 258.1A | *Lederbergia lenta* | 2009-01-05 | JAMATI000000000 | SRX16346900 | 36.47 | 47 | 1,268,550 | 1,157,153 | 4,411 | 9, 0, 0 | 4,617,171 | 307 |
| AMY 1.1.1 | *Bhargavaea ginsengi* | 2009-01-05 | JAMATJ000000000 | SRX16346901 | 53.68 | 108 | 614,536 | 106,446 | 3,125 | 7, 0, 0 | 3,187,319 | 207 |
| AMY 11.1.2 | *Staphylococcus warneri* | 2007-05-31 | JAMATN000000000 | SRX16346902 | 32.74 | 48 | 814,382 | 550,861 | 2,372 | 4, 0, 0 | 2,485,735 | 357 |
| AMY 13.1.2 | *Priestia megaterium* | 2007-06-01 | JAMATO000000000 | SRX16346903 | 37.65 | 105 | 1,886,191 | 2,128,716 | 5,721 | 16, 0, 0 | 5,722,597 | 368 |
| AMY 15.2 | *Cytobacillus* sp. | 2007-06-01 | JAMATP000000000 | SRX16346904 | 41.25 | 174 | 1,663,153 | 155,819 | 5,147 | 8, 0, 1 | 5,257,607 | 350 |
| AMY 19.1.2 | *Fictibacillus phosphorivorans* | 2007-06-05 | JAMBOR000000000 | SRX16346905 | 39.99 | 101 | 1,272,624 | 131,198 | 3,643 | 5, 0, 0 | 3,689,124 | 386 |
| AMY 19.1.2 vial 1 | *Fictibacillus phosphorivorans* | 2007-06-05 | JAMBOU000000000 | SRX16346906 | 40.02 | 559 | 1,365,050 | 131,198 | 3,675 | 5, 0, 0 | 3,903,370 | 390 |
| AMY 2.1.4 | *Staphylococcus equorum* | 2007-08-31 | JAMATK000000000 | SRX16346907 | 32.96 | 40 | 604,049 | 1,097,483 | 2,726 | 7, 1, 0 | 2,876,286 | 229 |
| AMY 24.1.1 | *Bacillus safensis* | 2007-06-01 | JAMATQ000000000 | SRX16346908 | 41.37 | 63 | 982,797 | 652,330 | 4,022 | 4, 0, 1 | 3,934,836 | 276 |
| AMY 28.1.2 | *Bacillus safensis* | 2007-06-01 | JAMATR000000000 | SRX16346909 | 41.67 | 34 | 830,701 | 1,900,424 | 3,654 | 5, 1, 0 | 3,673,368 | 251 |
| AMY 31.2 | *Bacillus pumilus* | 2007-06-01 | JAMATS000000000 | SRX16346911 | 41.66 | 67 | 874,120 | 446,334 | 3,768 | 5, 0, 1 | 3,777,130 | 257 |
| AMY 5.1.1 | *Cytobacillus firmus* | 2007-05-11 | JAMBOJ000000000 | SRX16346912 | 41.17 | 212 | 1,049,930 | 66,879 | 4,822 | 9, 0, 0 | 4,922,695 | 236 |
| AMY 52 | *Oceanobacillus luteolus* | 2007-05-11 | JAMBON000000000 | SRX16346913 | 38.89 | 207 | 1,216,340 | 89,594 | 3,891 | 8, 0, 0 | 4,009,121 | 338 |
| AMY 6.1.1 | *Bacillus licheniformis* | 2007-05-20 | JAMATL000000000 | SRX16346914 | 46.29 | 62 | 1,376,333 | 387,437 | 4,063 | 7, 0, 1 | 4,098,880 | 368 |
| AMY 7.1 | *Priestia flexa* | 2007-05-19 | JAMATM000000000 | SRX16346915 | 37.96 | 255 | 2,202,843 | 62,929 | 3,852 | 10, 1, 0 | 3,867,371 | 637 |
| CFPSW 5.3 | *Cytobacillus oceanisediminis* | 2008-04-10 | JAMATT000000000 | SRX16346916 | 40.94 | 106 | 1,436,564 | 289,172 | 5,683 | 10, 0, 0 | 5,770,654 | 274 |
| FAIRING 10M-2.2 | *Bacillus cytotoxicus* | 2004-04-21 | JAMBOP000000000 | SRX16346917 | 36.22 | 159 | 1,793,913 | 119,518 | 4,382 | 8, 1, 0 | 4,644,912 | 417 |
| FAIRING 12A-4 | *Heyndrickxia oleronia* | 2007-01-23 | JAMATW000000000 | SRX16346918 | 34.88 | 95 | 2,158,617 | 634,364 | 5,028 | 12, 0, 0 | 5,270,304 | 448 |
| FAIRING 19B-1.2 | *Staphylococcus warneri* | 2007-01-27 | JAMATX000000000 | SRX16346919 | 32.8 | 60 | 913,332 | 426,680 | 2,345 | 6, 1, 1 | 2,487,207 | 410 |
| FAIRING 3B-1.2 | *Bacillus velezensis* | 2013-08-19 | JAMATU000000000 | SRX16346920 | 46.18 | 438 | 769,918 | 811,981 | 4,235 | 7, 0, 1 | 4,212,269 | 204 |
| FAIRING 4G-1.1 | *Staphylococcus epidermidis* | 2013-08-20 | JAMATV000000000 | SRX16346922 | 31.94 | 74 | 1,100,052 | 167,417 | 2,385 | 6, 0, 1 | 2,585,673 | 474 |
| FAIRING W8B-1 | *Stenotrophomonas maltophilia* | 2013-08-19 | JAMATY000000000 | SRX16346923 | 66.54 | 73 | 535,590 | 175,432 | 4,439 | 3, 0, 1 | 4,898,338 | 118 |
| KSC 114 | *Bacillus altitudinis* | 2013-10 | JAMATZ000000000 | SRX16346924 | 41.3 | 33 | 702,844 | 483,649 | 3,738 | 5, 0, 1 | 3,712,377 | 206 |
| KSC 155 | *Bacillus safensis* | 2013-10 | JAMAUA000000000 | SRX16346925 | 41 | 82 | 1,190,920 | 672,012 | 3,809 | 5, 1, 0 | 3,813,861 | 344 |
| KSC 283 | *Priestia megaterium* | 2013-10 | JAMAUB000000000 | SRX16346926 | 38.25 | 146 | 1,524,780 | 497,735 | 5,282 | 14, 1, 0 | 5,197,219 | 322 |

(Continued on next page)

**TABLE 1** (Continued)

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| KSC 339 | *Staphylococcus saprophyticus* | 2013-10 | JAMAUC000000000 | SRX16346927 | 33.16 | 67 | 668,787 | 957,638 | 2,611 | 7, 1, 0 | 2,713,654 | 275 |
| KSC 351 | *Priestia aryabhattai* | 2013-10 | JAMAUD000000000 | SRX16346928 | 38.01 | 129 | 1,818,110 | 944,494 | 5,491 | 13, 0, 1 | 5,394,993 | 375 |
| KSC 386 | *Bacillus subtilis* | 2013-10 | JAMAUE000000000 | SRX16346929 | 45.48 | 846 | 1,510,313 | 1,048,650 | 5,014 | 8, 1, 0 | 4,623,969 | 362 |
| KSC 418 | *Paenibacillus cellulositrophicus* | 2013-10 | JAMAUF000000000 | SRX16346930 | 53.16 | 47 | 1,881,832 | 4,455,811 | 6,501 | 8, 0, 0 | 7,260,279 | 284 |
| KSC 432 | *Weizmannia ginsengihumi* | 2013-10 | JAMAUG000000000 | SRX16346931 | 36.29 | 212 | 1,172,830 | 90,727 | 3,518 | 11, 0, 0 | 3,778,211 | 343 |
| KSC 591 | *Bacillus subtilis* | 2013-10 | JAMAUH000000000 | SRX16342414 | 44.64 | 381 | 860,329 | 502,380 | 4,295 | 4, 0, 1 | 4,187,511 | 229 |
| KSC 640 | *Bacillus pumilus* | 2013-10 | JAMAUI000000000 | SRX16342415 | 41.44 | 147 | 2,196,861 | 277,292 | 3,730 | 4, 1, 1 | 3,690,879 | 657 |
| KSC 657 | *Niallia circulans* | 2013-10 | JAMAUJ000000000 | SRX16342452 | 35.56 | 79 | 2,442,071 | 338,051 | 4,663 | 10, 1, 0 | 5,008,815 | 541 |
| MER 10 | *Bacillus safensis* | 2003-04-09 | JAMAUL000000000 | SRX16342463 | 41.71 | 53 | 1,164,678 | 941,139 | 3,638 | 5, 0, 1 | 3,646,425 | 354 |
| MER 100 | *Niallia taxi* | 2004-04-20 | JAMAVA000000000 | SRX16342525 | 37.7 | 70 | 2,556,818 | 874,164 | 5,281 | 9, 1, 1 | 5,558,241 | 515 |
| MER 101 | *Paenibacillus illinoisensis* | 2004-04-20 | JAMAVB000000000 | SRX16342536 | 47.35 | 91 | 2,268,682 | 1,002,168 | 5,789 | 9, 0, 0 | 6,488,901 | 394 |
| MER 107 | *Neobacillus mesonae* | 2004-04-20 | JAMBOW000000000 | SRX16342499 | 44.31 | 93 | 2,412,451 | 889,947 | 5,466 | 8, 0, 0 | 5,956,811 | 450 |
| MER 108 | *Terribacillus saccharophilus* | 2004-04-20 | JAMAVC000000000 | SRX16342510 | 42.52 | 54 | 987,784 | 675,549 | 3,665 | 8, 0, 1 | 3,639,914 | 305 |
| MER 110 | *Priestia megaterium* | 2004-04-20 | JAMAVD000000000 | SRX16342545 | 37.84 | 121 | 1,668,830 | 1,289,863 | 5,772 | 13, 1, 0 | 5,687,929 | 328 |
| MER 112 | *Bacillus altitudinis* | 2004-04-20 | JAMAVE000000000 | SRX16342556 | 41.37 | 79 | 855,101 | 535,620 | 3,738 | 5, 0, 1 | 3,705,254 | 255 |
| MER 116 | *Bacillus cereus* | 2004-04-20 | JAMAVF000000000 | SRX16342416 | 35.13 | 65 | 1,895,614 | 847,946 | 5,308 | 11, 1, 0 | 5,433,442 | 372 |
| MER 118 | *Bacillus cereus* | 2004-04-20 | JAMAVG000000000 | SRX16342427 | 34.72 | 114 | 1,943,681 | 430,491 | 5,875 | 12, 1, 0 | 6,085,090 | 357 |
| MER 128 | *Priestia megaterium* | 2004-05-28 | JAMAVH000000000 | SRX16342438 | 37.79 | 156 | 1,508,285 | 938,338 | 5,866 | 13, 0, 0 | 5,783,390 | 292 |
| MER 13 | *Priestia koreensis* | 2004-04-09 | JAMAUM000000000 | SRX16342476 | 39.99 | 88 | 1,451,686 | 279,082 | 4,785 | 11, 0, 0 | 4,785,760 | 339 |
| MER 131 | *Paenibacillus lautus* | 2004-05-28 | JAMAVI000000000 | SRX16342487 | 51 | 57 | 1,605,299 | 4,447,497 | 6,290 | 4, 0, 0 | 7,140,748 | 249 |
| MER 132 | *Bacillus cereus* | 2004-05-26 | JAMAVJ000000000 | SRX16342447 | 34.99 | 55 | 1,300,785 | 1,104,589 | 5,504 | 11, 0, 0 | 5,651,989 | 256 |
| MER 135A | *Bacillus safensis* | 2004-05-26 | JAMAWQ000000000 | SRX16342448 | 41.27 | 72 | 1,261,935 | 697,687 | 3,982 | 5, 1, 0 | 3,915,961 | 359 |
| MER 145 | *Priestia aryabhattai* | 2004-05-28 | JAMBOQ000000000 | SRX16342449 | 37.63 | 132 | 1,341,366 | 1,075,173 | 5,539 | 13, 0, 0 | 5,481,448 | 272 |
| MER 153 | *Priestia aryabhattai* | 2004-05-28 | JAMAVK000000000 | SRX16342450 | 37.66 | 126 | 1,514,592 | 992,532 | 5,650 | 13, 1, 0 | 5,632,345 | 301 |
| MER 156 | *Bacillus subtilis* | 2004-05-26 | JAMAVL000000000 | SRX16342451 | 43.56 | 118 | 1,544,947 | 606,528 | 4,377 | 6, 0, 1 | 4,296,488 | 401 |
| MER 157 | *Paenibacillus elgii* | 2004-06-15 | JAMAVM000000000 | SRX16342453 | 52.88 | 334 | 2,640,790 | 237,149 | 7,374 | 10, 1, 0 | 8,123,320 | 359 |
| MER 161 | *Bacillus licheniformis* | 2004-06-15 | JAMAVN000000000 | SRX16342454 | 46.07 | 55 | 974,474 | 2,240,388 | 4,314 | 7, 0, 0 | 4,282,357 | 245 |
| MER 165 | *Bacillus licheniformis* | 2004-06-15 | JAMAVO000000000 | SRX16342455 | 46.03 | 29 | 1,157,486 | 2,240,383 | 4,292 | 7, 1, 0 | 4,271,713 | 293 |
| MER 166 | *Bacillus velezensis* | 2004-06-15 | JAMAVP000000000 | SRX16342456 | 45.94 | 120 | 971,171 | 367,388 | 4,369 | 9, 0, 1 | 4,412,102 | 242 |
| MER 169 | *Bacillus velezensis* | 2004-06-15 | JAMAVQ000000000 | SRX16342457 | 46.48 | 367 | 856,709 | 385,121 | 4,627 | 9, 0, 1 | 4,544,742 | 205 |
| MER 170 | *Priestia megaterium* | 2004-06-15 | JAMAVR000000000 | SRX16342458 | 37.67 | 126 | 1,343,731 | 1,011,460 | 5,663 | 15, 0, 0 | 5,639,457 | 265 |
| MER 172A | *Streptomyces pseudogriseolus* | 2004-06-28 | JAMAVS000000000 | SRX16342459 | 72.43 | 132 | 536,706 | 166,114 | 6,596 | 5, 1, 0 | 7,637,093 | 73 |
| MER 180 | *Paenibacillus* sp. | 2004-06-28 | JAMAVT000000000 | SRX16342460 | 46.52 | 129 | 2,045,322 | 344,519 | 5,798 | 7, 0, 0 | 6,744,104 | 339 |
| MER 189 | *Staphylococcus capitis* | 2004-06-28 | JAMAVU000000000 | SRX16342461 | 32.81 | 72 | 747,197 | 673,926 | 2,293 | 7, 1, 0 | 2,449,216 | 339 |
| MER 193 | *Exiguobacterium* sp. | 2004-06-28 | JAMAVV000000000 | SRX16342462 | 48.07 | 149 | 1,280,301 | 96,253 | 3,019 | 9, 0, 0 | 2,936,244 | 489 |
| MER 196A | *Streptomyces thermoviolaceus* | 2004-03-30 | JAMAVW000000000 | SRX16342464 | 72.39 | 177 | 1,120,250 | 243,441 | 4,710 | 4, 0, 0 | 5,572,416 | 216 |
| MER 20 | *Sporosarcina aquimarina* | 2003-04-09 | JAMBOM000000000 | SRX16342516 | 41.63 | 38 | 969,673 | 474,897 | 3,308 | 8, 1, 0 | 3,423,265 | 311 |

(Continued on next page)

**TABLE 1** (Continued)

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MER 26 | *Paenibacillus polysaccharolyticus* | 2003-04-09 | JAMAUN000000000 | SRX16342517 | 45.8 | 95 | 1,876,503 | 390,937 | 5,511 | 8, 0, 0 | 6,431,905 | 323 |
| MER 33 | *Mesobacillus* sp. | 2003-04-09 | JAMAUO000000000 | SRX16342518 | 42.95 | 83 | 1,784,978 | 377,099 | 4,616 | 8, 0, 0 | 4,787,833 | 414 |
| MER 36 | *Paenibacillus motobuensis* | 2003-04-10 | JAMAUP000000000 | SRX16342519 | 46.37 | 84 | 2,000,963 | 877,394 | 5,101 | 6, 0, 0 | 5,858,058 | 381 |
| MER 37 | *Paenibacillus pasadenensis* | 2003-04-10 | JAMBOT000000000 | SRX16342520 | 53 | 101 | 1,108,751 | 490,536 | 4,673 | 4, 0, 0 | 5,301,299 | 228 |
| MER 46 | *Metabacillus litoralis* | 2003-04-09 | JAMAUQ000000000 | SRX16342521 | 35.5 | 115 | 2,326,433 | 243,754 | 5,324 | 15, 0, 0 | 5,544,653 | 457 |
| MER 47 | *Sporosarcina luteola* | 2003-04-09 | JAMBOK000000000 | SRX16342522 | 44.89 | 31 | 541,805 | 2,093,321 | 3,191 | 7, 0, 1 | 3,306,674 | 182 |
| MER 48 | *Mesobacillus* sp. | 2003-04-09 | JAMAUR000000000 | SRX16342523 | 42.95 | 65 | 1,745,171 | 733,545 | 4,616 | 8, 0, 0 | 4,784,164 | 405 |
| MER 50.2 | *Paenibacillus macerans* | 2003-04-09 | JAMBOE000000000 | SRX16342524 | 52.02 | 137 | 2,502,116 | 363,305 | 6,312 | 4, 0, 0 | 7,289,640 | 322 |
| MER 51 | *Brevibacillus* sp. | 2003-04-09 | JAMAUS000000000 | SRX16342526 | 47.35 | 105 | 1,847,069 | 384,568 | 5,926 | 5, 0, 0 | 6,300,854 | 327 |
| MER 53-1 | *Neobacillus niacini* | 2003-04-09 | JAMBOX000000000 | SRX16342527 | 42.32 | 132 | 2,759,808 | 249,238 | 5,533 | 13, 0, 0 | 5,696,635 | 543 |
| MER 6 | *Niallia* sp. | 2003-04-09 | JAMAUK000000000 | SRX16342528 | 38.07 | 157 | 1,281,869 | 160,136 | 5,457 | 9, 1, 1 | 5,627,590 | 254 |
| MER 62 | *Hydrogenophaga intermedia* | 2003-03-05 | JAMBNO000000000 | SRX16342529 | 66.86 | 150 | 1,197,190 | 189,322 | 4,435 | 2, 2, 2 | 4,651,607 | 281 |
| MER 65 | *Priestia megaterium* | 2003-03-13 | JAMAUT000000000 | SRX16342530 | 37.92 | 110 | 1,577,240 | 1,914,888 | 5,483 | 15, 1, 0 | 5,407,622 | 326 |
| MER 73 | *Peribacillus frigoritolerans* | 2003-03-13 | JAMAUU000000000 | SRX16342531 | 40.26 | 154 | 1,521,411 | 261,737 | 5,391 | 9, 0, 0 | 5,714,701 | 293 |
| MER 74 | *Neobacillus* sp. | 2003-03-11 | JAMAUV000000000 | SRX16342532 | 38.82 | 147 | 1,010,466 | 241,461 | 5,230 | 19, 0, 0 | 5,541,816 | 204 |
| MER 78 | *Paenibacillus* sp. | 2003-03-11 | JAMAUW000000000 | SRX16342533 | 44.73 | 87 | 1,899,338 | 561,118 | 5,008 | 8, 0, 0 | 5,519,166 | 384 |
| MER 82 | *Cytobacillus horneckiae* | 2003-04-17 | JAMAUX000000000 | SRX16342534 | 37.8 | 90 | 1,253,825 | 715,509 | 5,041 | 9, 1, 0 | 5,219,193 | 270 |
| MER 89 | *Bacillus subtilis* | 2003-04-03 | JAMAUY000000000 | SRX16342535 | 43.71 | 75 | 669,189 | 323,094 | 4,014 | 5, 1, 0 | 4,049,231 | 185 |
| MER 99-2 | *Paenibacillus* sp. | 2004-04-21 | JAMAUZ000000000 | SRX16342537 | 44.85 | 77 | 2,016,740 | 916,723 | 5,679 | 9, 0, 0 | 6,526,312 | 346 |
| MER TA 106 | *Cytobacillus oceanisediminis* | 2013-07-30 | JAMAWE000000000 | SRX16342538 | 41.3 | 183 | 1,123,154 | 219,663 | 5,444 | 11, 1, 0 | 5,516,092 | 226 |
| MER TA 107 | *Neobacillus mesonae* | 2013-08-09 | JAMBNQ000000000 | SRX16342539 | 38.75 | 182 | 2,115,476 | 110,087 | 5,086 | 15, 0, 0 | 5,410,113 | 434 |
| MER TA 111 | *Fictibacillus nanhaiensis* | 2013-07-17 | JAMBOD000000000 | SRX16342492 | 39.49 | 78 | 1,398,058 | 385,213 | 4,208 | 6, 0, 0 | 4,217,690 | 369 |
| MER TA 112 | *Streptococcus oralis* | 2013-08-15 | JAMAWF000000000 | SRX16342493 | 41.35 | 51 | 442,268 | 298,081 | 1,900 | 3, 1, 0 | 2,019,356 | 243 |
| MER TA 114 | *Cytobacillus kochii* | 2013-08-01 | JAMAWG000000000 | SRX16342494 | 36.67 | 83 | 2,220,982 | 759,874 | 4,544 | 9, 0, 0 | 4,661,354 | 527 |
| MER TA 13 | *Rummeliibacillus stabekisii* | 2013-06-27 | JAMAVX000000000 | SRX16342495 | 37.41 | 68 | 1,020,678 | 953,447 | 3,251 | 12, 0, 1 | 3,349,733 | 340 |
| MER TA 136-3-2 | *Bacillus halotolerans* | 2013-07-02 | JAMAWH000000000 | SRX16342496 | 44.05 | 178 | 1,693,236 | 2,123,097 | 4,136 | 10, 1, 0 | 4,165,036 | 455 |
| MER TA 137-5 | *Bacillus subtilis* | 2013-07-03 | JAMAWI000000000 | SRX16342497 | 43.41 | 156 | 1,567,322 | 110,716 | 4,178 | 7, 1, 0 | 4,124,121 | 424 |
| MER TA 138-1 | *Alkalihalobacillus rhizosphaerae* | 2013-06-27 | JAMAWJ000000000 | SRX16342498 | 44.74 | 68 | 1,561,971 | 413,356 | 4,035 | 7, 0, 0 | 4,189,294 | 414 |
| MER TA 138-2 | *Cytobacillus oceanisediminis* | 2013-07-17 | JAMAWK000000000 | SRX16342500 | 41.04 | 133 | 1,892,273 | 153,567 | 5,559 | 9, 0, 0 | 5,681,886 | 373 |
| MER TA 139-2 | *Oceanobacillus profundus* | 2013-07-18 | JAMAWL000000000 | SRX16342501 | 37.64 | 90 | 803,015 | 262,861 | 4,334 | 8, 0, 0 | 4,561,942 | 197 |
| MER TA 14 | *Kocuria palustris* | 2013-07-23 | JAMAVY000000000 | SRX16342502 | 70.54 | 71 | 1,418,798 | 307,465 | 2,466 | 1, 1, 1 | 2,858,109 | 542 |
| MER TA 152 | *Mesobacillus maritimus* | 2013-07-17 | JAMBNR000000000 | SRX16342503 | 39.02 | 160 | 3,151,095 | 261,852 | 4,878 | 19, 0, 0 | 5,260,991 | 654 |
| MER TA 17 | *Psychrobacillus* sp. | 2013-08-09 | JAMAVZ000000000 | SRX16342504 | 44.91 | 121 | 1,923,583 | 343,608 | 4,245 | 7, 0, 0 | 4,324,992 | 496 |
| MER TA 170 | *Mesobacillus subterraneus* | 2013-07-18 | JAMBOF000000000 | SRX16342505 | 42.04 | 77 | 796,197 | 299,944 | 4,599 | 11, 1, 0 | 4,670,006 | 188 |

(Continued on next page)

**TABLE 1** (Continued)

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MER TA 171 | *Psychrobacillus* sp. | 2013-07-18 | JAMAWN000000000 | SRX16342506 | 37.17 | 114 | 1,399,241 | 146,555 | 3,976 | 9, 1, 0 | 4,064,209 | 381 |
| MER TA 176 | *Ureibacillus chungkukjangi* | 2013-07-17 | JAMAWO000000000 | SRX16342507 | 36.18 | 172 | 829,210 | 109,322 | 4,264 | 11, 1, 0 | 4,459,889 | 208 |
| MER TA 18 | *Staphylococcus warneri* | 2013-06-27 | JAMAWA000000000 | SRX16342508 | 32.65 | 27 | 566,180 | 920,533 | 2,266 | 7, 1, 1 | 2,424,208 | 256 |
| MER TA 181 | *Metabacillus halosaccharovorans* | 2013-07-17 | JAMAWP000000000 | SRX16342509 | 35.74 | 137 | 2,797,792 | 202,899 | 5,666 | 10, 0, 0 | 6,017,055 | 520 |
| MER TA 32b | *Planococcus* sp. | 2013-08-12 | JAMCAN000000000 | SRX17618098 | 46.17 | 94 | 1,585,957 | 2,020,338 | 3,209 | 5, 1, 0 | 3,245,854 | 546 |
| MER TA 38 | *Paenibacillus glycanilyticus* | 2013-07-24 | JAMBNP000000000 | SRX16342511 | 50.3 | 105 | 1,743,564 | 537,285 | 5,968 | 6, 0, 0 | 6,724,235 | 288 |
| MER TA 48 | *Sporosarcina luteola* | 2013-07-23 | JAMBNV000000000 | SRX16342512 | 44.5 | 47 | 1,529,788 | 2,309,686 | 3,444 | 8, 0, 0 | 3,552,921 | 482 |
| MER TA 81-3 | *Paenibacillus* sp. | 2013-07-23 | JAMAWB000000000 | SRX16342513 | 48.69 | 173 | 1,832,882 | 307,130 | 5,899 | 8, 0, 0 | 6,872,147 | 296 |
| MER TA 82 | *Cytobacillus kochii* | 2014-07-23 | JAMAWC000000000 | SRX16342514 | 36.97 | 197 | 2,780,397 | 759,874 | 4,654 | 9, 0, 0 | 4,712,805 | 659 |
| MER TA 86 | *Metabacillus litoralis* | 2013-07-30 | JAMBNW000000000 | SRX16342515 | 36.5 | 211 | 2,110,902 | 227,801 | 5,337 | 15, 0, 0 | 5,990,291 | 390 |
| MER TA 87 | *Neobacillus niacini* | 2013-07-30 | JAMBOB000000000 | SRX16342540 | 38.14 | 111 | 2,406,523 | 495,822 | 5,107 | 13, 0, 0 | 5,316,554 | 504 |
| MER TA 97 | *Staphylococcus capitis* | 2013-08-09 | JAMAWD000000000 | SRX16342541 | 33.42 | 150 | 860,100 | 469,804 | 2,337 | 7, 0, 0 | 2,456,498 | 390 |
| MER 9 | *Sporosarcina luteola* | 2003-04-09 | JAMBOI000000000 | SRX16342542 | 44.52 | 92 | 871,704 | 443,563 | 3,551 | 7, 0, 0 | 3,602,600 | 272 |
| MER TA 168 | *Niallia* sp. | 2013-07-18 | JAMAWM000000000 | SRX16342543 | 35.05 | 199 | 1,450,973 | 100,872 | 4,786 | 10, 0, 0 | 5,178,238 | 306 |
| MER TA 154 | *Solibacillus isronensis* | 2013-07-17 | JAMBOC000000000 | SRX16342544 | 38.62 | 65 | 1,166,522 | 2,361,702 | 3,684 | 11, 0, 1 | 3,808,720 | 337 |
| MPF 19 | *Bacillus velezensis* | 2004-05-01 | JAMAWS000000000 | SRX16342546 | 48.09 | 699 | 366,662 | 155,510 | 4,354 | 8, 6, 6 | 4,203,241 | 98 |
| MPF 2 | *Bacillus licheniformis* | 2004-05-01 | JAMAWY000000000 | SRX16342547 | 46.09 | 142 | 1,271,867 | 612,391 | 4,505 | 7, 0, 1 | 4,398,042 | 319 |
| MPF 24 | *Bacillus subtilis* | 2004-05-06 | JAMAWT000000000 | SRX16342548 | 43.86 | 168 | 1,310,983 | 1,057,136 | 4,260 | 7, 0, 0 | 4,199,270 | 349 |
| MPF 38 | *Bacillus paralicheniformis* | 2004-05-06 | JAMAWU000000000 | SRX16342549 | 45.91 | 55 | 1,027,464 | 1,116,529 | 4,287 | 7, 0, 0 | 4,395,789 | 259 |
| MPF 57 | *Bacillus intestinalis* | 2004-05-19 | JAMAWV000000000 | SRX16342550 | 44.03 | 115 | 1,404,213 | 325,227 | 3,975 | 9, 0, 0 | 4,032,203 | 383 |
| MPF 67 | *Bacillus licheniformis* | 2004-05-11 | JAMAWW000000000 | SRX16342551 | 45.97 | 82 | 1,062,046 | 583,595 | 4,430 | 7, 0, 1 | 4,369,009 | 268 |
| MPF 76A | *Bacillus licheniformis* | 2004-05-11 | JAMAWX000000000 | SRX16342552 | 48.09 | 1093 | 2,123,917 | 491,455 | 5,461 | 7, 0, 1 | 4,892,212 | 484 |
| MPF 8 | *Bacillus licheniformis* | 2004-05-01 | JAMAWR000000000 | SRX16342553 | 46.54 | 400 | 1,453,562 | 553,893 | 4,732 | 7, 0, 1 | 4,521,756 | 359 |
| MSL 004.1.2 | *Rothia dentocariosa* | 2008-03-17 | JAMAWZ000000000 | SRX16342554 | 53.94 | 76 | 616,186 | 389,714 | 2,098 | 3, 1, 0, 1 | 2,448,390 | 280 |
| MSL 016.1 | *Staphylococcus epidermidis* | 2008-03-27 | JAMAXA000000000 | SRX16342555 | 32.08 | 72 | 983,790 | 187,847 | 2,236 | 7, 1, 0 | 2,462,434 | 445 |
| MSL 036.1 | *Brevibacillus invocatus* | 2008-05-14 | JAMAXB000000000 | SRX16342557 | 48.69 | 214 | 1,135,485 | 93,236 | 4,810 | 5, 0, 0 | 4,940,093 | 258 |
| MSL 047.1 | *Sphingopyxis alaskensis* | 2008-05-14 | JAMAXC000000000 | SRX16342558 | 65.11 | 75 | 1,639,175 | 145,245 | 3,615 | 1, 1, 1 | 3,777,004 | 475 |
| MSL 047.2 | *Sphingomonas paucimobilis* | 2008-05-14 | JAMAXD000000000 | SRX16342559 | 65.48 | 71 | 2,658,702 | 593,513 | 4,195 | 1, 1, 1 | 4,553,552 | 644 |
| MSL 058.1.2 | *Metabacillus litoralis* | 2008-07-21 | JAMAXE000000000 | SRX16342560 | 35.6 | 122 | 2,116,836 | 349,633 | 5,388 | 15, 0, 0 | 5,577,721 | 417 |
| MSL 060.1.1 | *Bacillus atrophaeus* | 2008-07-21 | JAMAXF000000000 | SRX16342561 | 43.2 | 26 | 1,049,730 | 1,047,913 | 3,912 | 7, 1, 1 | 4,086,418 | 273 |
| MSL 107 | *Staphylococcus epidermidis* | 2008-10-08 | JAMAXG000000000 | SRX16342562 | 32.13 | 64 | 1,243,709 | 183,913 | 2,276 | 6, 0, 0 | 2,512,070 | 552 |
| MSL 160.2 | *Bacillus velezensis* | 2008-11-12 | JAMAXI000000000 | SRX16342563 | 46.43 | 79 | 1,304,922 | 382,423 | 3,776 | 7, 0, 0 | 4,006,624 | 364 |
| MSL 172.1.2 | *Bacillus safensis* | 2008-12-02 | JAMAXJ000000000 | SRX16342465 | 41.67 | 32 | 585,827 | 946,465 | 3,680 | 4, 0, 0 | 3,671,895 | 177 |
| MSL 173.2.2 | *Peribacillus simplex* | 2008-12-02 | JAMAXK000000000 | SRX16342466 | 39.84 | 227 | 1,856,358 | 102,406 | 5,323 | 9, 0, 0 | 5,725,607 | 361 |
| MSL 179.1 | *Heyndrickxia oleronia* | 2008-12-04 | JAMAXL000000000 | SRX16342467 | 34.83 | 120 | 2,771,284 | 152,837 | 5,134 | 11, 0, 0 | 5,286,683 | 579 |
| MSL 185.1 | *Caldibacillus thermoamylovorans* | 2008-12-08 | JAMAXM000000000 | SRX16342417 | 37.36 | 111 | 911,571 | 277,299 | 3,456 | 9, 0, 0 | 3,809,017 | 259 |
| MSL 187.1 | *Paenibacillus lactis* | 2008-12-08 | JAMAXN000000000 | SRX16342418 | 52.01 | 96 | 1,872,690 | 556,132 | 5,905 | 6, 1, 0 | 6,686,450 | 308 |

(Continued on next page)

**TABLE 1** (Continued)

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MSL 200.1 | *Priestia megaterium* | 2008-12-04 | JAMAXO000000000 | SRX16342419 | 37.98 | 112 | 2,287,220 | 1,185,537 | 5,366 | 12, 0, 0 | 5,295,778 | 482 |
| MSL 225.1.2 | *Domibacillus indicus* | 2009-01-15 | JAMBOO000000000 | SRX16342420 | 43.49 | 181 | 1,271,139 | 161,986 | 4,858 | 23, 0, 0 | 5,083,418 | 278 |
| MSL 259.1 | *Staphylococcus lugdunensis* | 2008-04-06 | JAMAXP000000000 | SRX16342421 | 35.88 | 455 | 737,693 | 302,134 | 2,803 | 5, 0, 0 | 2,789,329 | 297 |
| MSL 3003 | *Staphylococcus xylosus* | 2014-07-23 | JAMAXW000000000 | SRX16342422 | 32.7 | 78 | 732,430 | 291,655 | 2,679 | 7, 1, 1 | 2,864,935 | 285 |
| MSL 314.1 | *Staphylococcus hominis* | 2010-06-28 | JAMAXQ000000000 | SRX16342423 | 31.51 | 70 | 882,042 | 362,212 | 2,130 | 6, 0, 0 | 2,202,842 | 445 |
| MSL 316.2 | *Staphylococcus warneri* | 2010-06-28 | JAMAXR000000000 | SRX16342424 | 32.48 | 37 | 809,071 | 431,070 | 2,409 | 5, 1, 0 | 2,550,295 | 356 |
| MSL 321.1 | *Micrococcus luteus* | 2010-06-28 | JAMAXS000000000 | SRX16342425 | 72.9 | 126 | 1,098,795 | 174,685 | 2,311 | 1, 0, 0 | 2,546,255 | 468 |
| MSL 336.2 | *Ralstonia pickettii* | 2010-06-28 | JAMAXT000000000 | SRX16342426 | 63.35 | 44 | 1,779,304 | 460,861 | 5,262 | 1, 1, 1 | 5,563,662 | 352 |
| MSL 348 | *Cupriavidus pauculus* | 2010-04-20 | JAMAXU000000000 | SRX16342428 | 64.08 | 82 | 1,395,623 | 425,000 | 3,760 | 1, 1, 1 | 6,426,329 | 237 |
| MSL 359 | *Kocuria rosea* | 2010-05-06 | JAMAXV000000000 | SRX16342429 | 72.48 | 129 | 959,164 | 265,844 | 3,612 | 1, 1, 0 | 3,983,493 | 259 |
| MSL 140.1 | *Brevibacillus borstelensis* | 2008-10-08 | JAMAXH000000000 | SRX16342430 | 51.36 | 209 | 2,022,985 | 103,406 | 5,067 | 3, 1, 0 | 5,382,697 | 419 |
| ODYSSEY 48 V2 | *Curtobacterium* sp. | 2007-01-29 | JAMAXX000000000 | SRX16342431 | 71.88 | 81 | 901,252 | 346,312 | 3,217 | 2, 1, 0 | 3,485,755 | 286 |
| P10 | *Microbacterium oleivorans* | 2003-12 | JAMBOV000000000 | SRX16342432 | 69.95 | 31 | 1,265,533 | 641,097 | 2,993 | 3, 0, 1 | 3,171,571 | 441 |
| P100 | *Rothia* sp. | 2003-12 | JAMAYG000000000 | SRX16342433 | 56.18 | 20 | 646,400 | 515,480 | 2,107 | 3, 1, 1 | 2,404,879 | 297 |
| P106 | *Alkalihalobacillus oceani* | 2003-12 | JAMBOS000000000 | SRX16342434 | 46.36 | 146 | 2,330,371 | 138,530 | 4,178 | 8, 0, 1 | 4,381,113 | 589 |
| P107 | *Priestia aryabhattai* | 2003-12 | JAMAYH000000000 | SRX16342435 | 37.64 | 98 | 1,469,448 | 1,101,254 | 5,663 | 13, 0, 0 | 5,597,164 | 289 |
| P112 | *Cellulosimicrobium funkei* | 2003-12 | JAMAYI000000000 | SRX16342436 | 74.71 | 65 | 721,097 | 514,644 | 3,743 | 1, 1, 1 | 4,266,893 | 176 |
| P121 | *Priestia megaterium* | 2003-12 | JAMAYJ000000000 | SRX16342437 | 37.76 | 92 | 1,223,494 | 752,388 | 5,608 | 13, 0, 1 | 5,578,758 | 242 |
| P18 | *Mesobacillus subterraneus* | 2003-12 | JAMBNN000000000 | SRX16342439 | 42.33 | 106 | 2,401,577 | 200,889 | 4,984 | 6, 1, 0 | 5,105,592 | 523 |
| P25 | *Microbacterium enclense* | 2003-12 | JAMAXY000000000 | SRX16342440 | 70.46 | 26 | 1,223,278 | 360,559 | 3,132 | 1, 1, 1 | 3,486,885 | 376 |
| P26 | *Microbacterium* sp. | 2003-12 | JAMAXZ000000000 | SRX16342468 | 70.34 | 43 | 2,257,191 | 588,147 | 3,244 | 1, 0, 1 | 3,517,995 | 699 |
| P30 | *Microbacterium hydrocarbonoxydans* | 2003-12 | JAMBOY000000000 | SRX16342469 | 69.02 | 27 | 769,560 | 838,703 | 3,288 | 2, 0, 1 | 3,483,679 | 239 |
| P42 | *Alkalihalophilus marmarensis* | 2003-12 | JAMAYA000000000 | SRX16342470 | 40 | 82 | 949,855 | 206,984 | 4,125 | 7, 0, 0 | 4,103,950 | 257 |
| P67 | *Sutcliffiella horikoshii* | 2003-12 | JAMAYB000000000 | SRX16342471 | 41.05 | 68 | 1,547,227 | 410,581 | 4,695 | 9, 0, 0 | 4,708,311 | 368 |
| P7 | *Georgenia satyanarayanai* | 2003-12 | JAMBNZ000000000 | SRX16342472 | 72.21 | 196 | 2,481,880 | 304,175 | 3,395 | 1, 1, 0 | 3,723,384 | 718 |
| P75 | *Staphylococcus capitis* | 2003-12 | JAMAYC000000000 | SRX16342473 | 32.83 | 72 | 562,317 | 503,111 | 2,463 | 7, 0, 0 | 2,597,991 | 240 |
| P83 | *Staphylococcus capitis* | 2003-12 | JAMAYD000000000 | SRX16342474 | 34.11 | 272 | 551,867 | 672,612 | 2,542 | 7, 0, 0 | 2,583,208 | 240 |
| P86 | *Nocardioides* sp. | 2003-12 | JAMAYE000000000 | SRX16342475 | 74.08 | 64 | 2,061,229 | 1,297,617 | 3,697 | 3, 1, 1 | 4,031,233 | 553 |
| P97 | *Curtobacterium* sp. | 2003-12 | JAMAYF000000000 | SRX16342477 | 71.82 | 248 | 382,234 | 219,850 | 3,469 | 3, 1, 0 | 3,616,588 | 117 |
| PF24B.2 | *Robertmurraya korlensis* | 2007-08-20 | JAMBNS000000000 | SRX16342478 | 38.22 | 191 | 1,836,561 | 290,106 | 4,733 | 13, 0, 0 | 4,761,956 | 428 |
| PF3F.2 | *Priestia endophytica* | 2007-08-20 | JAMAYK000000000 | SRX16342479 | 36.58 | 51 | 1,159,426 | 3,768,234 | 5,003 | 9, 0, 0 | 5,145,492 | 248 |
| PF4F.2.1 | *Cytobacillus oceanisediminis* | 2007-08-20 | JAMAYL000000000 | SRX16342480 | 41.45 | 94 | 2,220,633 | 703,746 | 5,136 | 10, 1, 0 | 5,254,108 | 470 |
| RA 14A.10 | *Agromyces mediolanus* | 2007-04-06 | JAMBNX000000000 | SRX16342481 | 71.47 | 90 | 853,226 | 376,941 | 3,757 | 1, 1, 1 | 4,038,592 | 225 |
| TA 121-4 | *Bacillus subtilis* | 2004-03-08 | JAMAYN000000000 | SRX16342482 | 73.04 | 103 | 733,140 | 195,078 | 2,335 | 1, 1, 0 | 2,559,014 | 306 |
| TA 127 | *Micrococcus luteus* | 2004-03-08 | JAMAYO000000000 | SRX16342483 | 73.08 | 106 | 282,980 | 228,776 | 2,225 | 2, 0, 0 | 2,454,472 | 122 |
| TA 149 | *Micrococcus luteus* | 2004-03-30 | JAMAYP000000000 | SRX16342484 | 44.57 | 137 | 1,361,878 | 202,381 | 4,315 | 7, 0, 1 | 4,492,163 | 339 |
| TA 170 | *Alkalihalobacillus clausii* | 2004-03-30 | JAMAYQ000000000 | SRX16342485 | 42.07 | 118 | 2,852,326 | 299,670 | 4,599 | 11, 1, 0 | 4,682,835 | 682 |

(Continued on next page)

**TABLE 1** (Continued)

| Strain name | Taxonomic identification | Date of isolation (yr/mo/day) | GenBank accession no. | SRA accession no. for Illumina reads | G+C content (%) | No. of contigs | No. of Illumina reads | Genome assembly $N_{50}$ (bp) | No. of protein-coding genes | No. of complete rRNAs (5S, 16S, 23S) | Genome size (bp) | Genome coverage (×) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TA 172 | *Mesobacillus subterraneus* | 2004-03-30 | JAMBNT000000000 | SRX16342486 | 40.51 | 93 | 2,255,199 | 533,846 | 4,951 | 8, 1, 0 | 5,059,287 | 489 |
| TA 28 | *Metabacillus idriensis* | 2004-02-25 | JAMAYR000000000 | SRX16342488 | 73.14 | 73 | 465,310 | 158,923 | 3,269 | 1, 1, 1 | 3,635,835 | 133 |
| TA 29 | *Kocuria rosea* | 2012-07-30 | JAMBOG000000000 | SRX16342489 | 42.96 | 112 | 2,081,125 | 265,106 | 4,374 | 5, 0, 0 | 4,885,757 | 464 |
| TA 33-2 | *Paenibacillus camelliae* | 2013-07-23 | JAMBNY000000000 | SRX16342490 | 35 | 53 | 694,264 | 526,229 | 2,340 | 4, 1, 1 | 2,516,602 | 303 |
| TA 76 | *Carnobacterium inhibens* | 2004-02-17 | JAMAYM000000000 | SRX16342491 | 38.92 | 203 | 2,196,655 | 209,477 | 5,964 | 16, 0, 0 | 6,523,156 | 374 |
| TA 104 | *Neobacillus cucumis* | 2004-02-06 | JAMBOH000000000 | SRX16342441 | 43.83 | 78 | 1,210,078 | 473,068 | 4,003 | 5, 0, 1 | 4,048,329 | 328 |
| TPS 10-6.1 | *Brevibacillus borstelensis* | 2006-08-24 | JAMAYT000000000 | SRX16342442 | 51.39 | 257 | 1,013,852 | 81,183 | 5,155 | 2, 0, 0 | 5,437,219 | 207 |
| TPS 11-9.1 | *Caldibacillus thermoamylovorans* | 2006-08-24 | JAMAYU000000000 | SRX16342443 | 37.43 | 291 | 678,229 | 68,324 | 3,523 | 10, 0, 0 | 3,812,260 | 182 |
| TPS 14-3.1 | *Brevibacillus borstelensis* | 2006-04-26 | JAMAYV000000000 | SRX16342444 | 51.57 | 231 | 835,682 | 96,347 | 5,088 | 3, 1, 0 | 5,368,369 | 171 |
| TPS 8-13.1 | *Brevibacillus borstelensis* | 2006-09-20 | JAMAYS000000000 | SRX16342445 | 51.36 | 237 | 1,999,409 | 77,807 | 5,096 | 3, 1, 0 | 5,389,870 | 414 |
| V41-32 | *Janibacter melonis* | 2013-02-26 | JAMAYW000000000 | SRX16342446 | 72.4 | 98 | 783,955 | 415,084 | 3,439 | 1, 0, 0 | 3,628,094 | 231 |

[a] The taxonomic identification was based off the species identification of GTDB-Tk, with adjustments made during submission to NCBI. The number of complete rRNAs does not take into account that all strains contain partial 16S rRNA.

45 s. Library preparation and quantification of DNA were executed on a Microlab STAR automated liquid handling system (Hamilton Bonaduz AG, Switzerland). The pooled libraries were then quantified using the KAPA library quantification kit for Illumina (Kapa Biosystems, USA). The libraries were sequenced using the Illumina HiSeq platform, with a 250-bp paired-end protocol. Adapters were trimmed from the raw reads using Trimmomatic 0.30, with a sliding cutoff of Q15 (3). SPAdes 2.7 (4) was used to perform *de novo* assembly of the samples, and the contigs were annotated using Prokka 1.11 (5). The assembled sequences were queried against the SILVA LTP 16S rRNA gene type strain database. Sequences with >98.7% sequence similarity to SILVA reference sequence were assigned a species level taxonomic identification. This screening process down-selected 191 isolates, which were sent for whole-genome sequencing (WGS). All 191 isolates had 16S sequences with greater than 98.7% similarity to a SILVA reference sequence. For whole-genome sequencing, DNA was also sent to MicrobesNG for library preparation and sequencing. All library preparation, sequencing, read trimming, quality control, and assembly were performed using the same procedures as described for 16S sequencing. The read quality was assessed using in-house scripts in combination with the following software: SAMtools (6), BEDTools (7), and BWA-MEM (8). The assembly quality was assessed using QUAST (9), and CheckM (10) was used to assess the read quality. All 191 sequences were >90% complete, with <5% contamination. The Genome Taxonomy Database Toolkit (GTDB-Tk) (11) was used to provide the final species level identification that was used for isolate identification. No new genus or species designations were assigned to any of the strains.

The assembled draft genomes were submitted for annotation using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The assembly statistics for all 191 strains are provided in Table 1.

**Data availability.** The 16S rRNA sequences, whole-genome sequences, and Sequence Read Archive (SRA) submission data were submitted to the National Center for Biotechnology Information (NCBI) under the BioProject accession number PRJNA832800 and SRA accession number SRP387068. The draft genome sequences of all 191 isolates have been deposited at NCBI under the accession numbers provided in Table 1.

## ACKNOWLEDGMENT

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

## REFERENCES

1. National Aeronautics and Space Administration. 2010. Handbook for the microbial examination of space hardware. NASA-HDBK-6022. National Aeronautics and Space Administration, Washington, DC. https://explorers.larc.nasa.gov/2019APSMEX/SMEX/pdf_files/NASA-HDBK-6022b.pdf.
2. Seuylemezian A, Aronson HS, Tan J, Lin M, Schubert W, Vaishampayan P. 2018. Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. Front Microbiol 9:780. https://doi.org/10.3389/fmicb.2018.00780.
3. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
4. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
5. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
6. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience 10:giab008. https://doi.org/10.1093/gigascience/giab008.
7. Quinlan AR, Ira MH. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842. https://doi.org/10.1093/bioinformatics/btq033.
8. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997v2 [q-bio.GN]. https://doi.org/10.48550/arXiv.1303.3997.
9. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
10. 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. https://doi.org/10.1101/gr.186072.114.
11. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.