AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of *Enterobacteriales* Strains Isolated from the International Space Station

Achintya R. Bharadwaj,[a] Robert Daudu,[a] Nitin K. Singh,[a] Jason M. Wood,[a] Marilyne Debieu,[b] Niamh B. O'Hara,[b,c] Fathi Karouia,[d,e] Christopher E. Mason,[f,g] Kasthuri Venkateswaran[a]

[a]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
[b]Biotia, New York, New York, USA
[c]Department of Cell Biology, College of Medicine, SUNY Downstate Health Sciences University, Brooklyn, New York, USA
[d]Blue Marble Space Institute of Science, Exobiology Branch, NASA Ames Research Center, Moffett Field, California, USA
[e]Space Research Within Reach, San Francisco, California, USA
[f]WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, New York, USA
[g]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, New York, USA

ABSTRACT The whole-genome sequences of 26 strains isolated from the International Space Station were generated, and the strains were identified as being members of the order *Enterobacteriales*. Characterization of these whole-genome sequences might enable the identification of potential pathogenic bacteria that have been adapting to the space environment.

Members of the order *Enterobacteriales* have been found to exhibit human pathogenicity and therefore pose a health risk for people on Earth and for astronauts aboard the International Space Station (ISS) (1, 2). The latter is of concern for long-duration missions, as astronauts have been shown to be immunocompromised (3). Importantly, these bacteria are able to adapt to extreme conditions such as microgravity and radiation and thus persist, necessitating the development of appropriate countermeasures to control them. Members of the order *Enterobacteriales* that were found on ISS surfaces were *Pantoea brenneri*, *Pantoea agglomerans*, *Kalamiella piersonii*, and *Enterobacter bugandensis* (4–6). On Earth, *P. agglomerans* and *P. brenneri* were reported to have been isolated from human infections (4). *K. piersonii* is a member of a novel genus in the family *Erwiniaceae* that has exhibited resistance to multiple clinical drugs, such as penicillin and vancomycin, allowing it to be an emerging pathogen (5). *E. bugandensis* was documented from blood as a causative agent of septicemia in various geological locations (7). Analyses of draft genome assemblies for these species might pave the way to identify the genetic processes responsible for potential pathogenicity, as previously reported for some of these strains (5, 6).

The strains used for whole-genome sequencing (WGS) were isolated from four different locations in the ISS across three flights and are detailed in Table 1 (8). The ISS surface samples collected and brought back to Earth were aseptically handled, suitable aliquots of the sample concentrate (100 μl) were plated onto Reasoner's 2A (R2A) medium and incubated at 25°C for 7 days, and a single well-isolated colony was archived at −80°C until DNA extraction. DNA was extracted from cultures grown in R2A medium using the ZymoBIOMICS DNA MagBead kit according to the manufacturer's instructions.

WGS of 26 bacterial isolates from the ISS was performed using the Illumina Nextera Flex protocol for library preparation, as used in similar studies (6). The NovaSeq 6000 system with an S4 flow cell (paired-end 2 × 150-bp reads) was used to execute paired-end sequencing. FastQC (v0.11.7) was used to validate the quality of the raw

Citation Bharadwaj AR, Daudu R, Singh NK, Wood JM, Debieu M, O'Hara NB, Karouia F, Mason CE, Venkateswaran K. 2020. Draft genome sequences of *Enterobacteriales* strains isolated from the International Space Station. Microbiol Resour Announc 9:e00817-20. https://doi.org/10.1128/MRA.00817-20.

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Kasthuri Venkateswaran, kjvenkat@jpl.nasa.gov.

Received 15 July 2020
Accepted 13 August 2020
Published 10 September 2020

TABLE 1 Metadata and genome statistics for *Enterobacter*, *Kalamiella*, and *Pantoea* strains isolated from various ISS environmental surfaces during the Microbial Tracking 1 flight project

Sample name	Nearest species identity[a]	GenBank accession no.	Raw sequence accession no.	Flight(s) or facility[b]	Sampling location[c]	No. of contigs[d]	Genome size (bp)	N_{50} (bp)	Median coverage (×)	No. of quality-controlled reads	No. of raw reads (×10^6)	G+C content (%)
IF2SW-B4	*E. bugandensis*	JABWOY000000000	SRR11885007	F1-2	WHC	36	4,892,220	511,556	657.59	34,787,506	17.5	55.9
IFACSW-B2	*E. bugandensis*	JABWOX000000000	SRR11885006	F1	FC	40	4,892,163	481,191	662.95	35,107,666	17.6	55.9
IFACSW-B4	*E. bugandensis*	JABWOW000000000	SRR11885005	F1	FC	35	4,892,584	808,304	606.70	31,965,988	16.0	55.9
IFACSW-B5	*E. bugandensis*	JABWOV000000000	SRR11885004	F1	FC	37	4,891,741	481,191	672.32	35,130,172	17.7	55.9
IFACSW-P1	*E. bugandensis*	JABWOU000000000	SRR11885003	F1	FC	36	4,891,763	808,252	640.18	35,538,578	17.9	55.9
IF2SW-F2	*E. bugandensis*	JACBPD000000000	SRR12071883	F1-2	WHC	25	4,892,159	511,419	547.77	29,062,046	14.5	55.9
IF2SW-F3	*E. bugandensis*	JACBPE000000000	SRR12071879	F1-2	WHC	22	4,892,298	808,302	518.30	27,421,928	13.7	55.9
F3-6B(4)	*K. piersonii*	JACBPM000000000	SRR12071882	F3-6	PMM	39	4,850,268	503,530	313.39	17,045,882	8.5	57.1
F3-6B(5)	*K. piersonii*	JACBPN000000000	SRR12071881	F3-6	PMM	50	4,850,704	310,993	411.16	24,040,576	12.0	57.1
IIIF_BACT_A	*K. piersonii*	JACBPO000000000	SRR12071880	F3-6	PMM	42	4,849,373	503,411	441.96	22,790,000	11.4	57.1
FJII-L5-SW-P2	*P. agglomerans*	JACBPL000000000	SRR12071872	JPL SAF II	Cleanroom floor	26	4,861,660	445,707	413.84	23,353,202	11.7	55.1
IF5SW-B1	*P. brenneri*	JABWPM000000000	SRR11885013	F1-5	N1-O4	108	5,022,545	216,403	487.50	27,782,700	14.0	55.9
IF5SW-B2	*P. brenneri*	JABWPL000000000	SRR11885012	F1-5	N1-O4	107	5,023,215	216,403	372.32	20,371,906	10.2	55.9
IFACSW-B3	*P. brenneri*	JABWPK000000000	SRR11885002	F1	FC	108	5,023,154	216,403	549.11	31,932,976	16.0	55.9
IF5SW-P1	*P. brenneri*	JABWPJ000000000	SRR11885001	F1-5	N1-O4	106	5,023,034	216,403	534.38	32,864,806	16.5	55.9
IF5SW-P2	*P. brenneri*	JABWPI000000000	SRR11885000	F1-5	N1-O4	106	5,023,268	216,131	586.61	37,186,174	18.7	55.9
IFACSW-P2	*P. brenneri*	JABWPH000000000	SRR11884999	F1	FC	108	5,023,383	216,403	553.13	34,845,798	17.5	55.9
IIF5SW-B1	*P. brenneri*	JABWPG000000000	SRR11884998	F1-5	N1-O4	106	5,022,674	216,137	436.61	26,223,526	13.2	55.9
IIF5SW-B2	*P. brenneri*	JABWPF000000000	SRR11884997	F1-5	N1-O4	106	5,023,080	216,403	460.71	27,740,334	13.9	55.9
IIF5SW-B5	*P. brenneri*	JABWPE000000000	SRR11884996	F1-5	N1-O4	111	5,021,866	176,626	366.96	22,097,600	11.1	55.9
IIF5SW-P1	*P. brenneri*	JACBPF000000000	SRR12071878	F2-5	N1-O4	75	5,020,903	176,626	444.64	26,701,110	13.4	55.9
IIF5SW-P2	*P. brenneri*	JACBPG000000000	SRR12071877	F2-5	N1-O4	75	5,022,945	216,403	451.34	26,675,346	13.4	55.9
IIF5SW-P3	*P. brenneri*	JACBPH000000000	SRR12071876	F2-5	N1-O4	75	5,021,950	216,403	459.38	26,806,520	13.5	55.9
IIF5SW-P4	*P. brenneri*	JACBPI000000000	SRR12071875	F2-5	N1-O4	75	5,023,133	216,403	586.61	35,389,420	17.8	55.9
IIF5SW-P5	*P. brenneri*	JACBPJ000000000	SRR12071874	F2-5	N1-O4	75	5,023,087	216,131	510.27	29,693,196	14.9	55.9
IIFCSG-B1	*P. brenneri*	JACBPK000000000	SRR12071873	CRV2	CRV-FC	74	5,022,524	216,403	345.54	20,467,790	10.3	55.9

[a] The 16S rRNA gene sequences were retrieved from the whole-genome sequence of the queried genome and analyzed with BLAST against type strains for all 16S rRNA sequences in the NCBI database. The bacterial species identity was determined when the queried sequence showed >97.5% similarity to the 16S rRNA gene sequence of the type strain (*E. bugandensis* DSM 29888^T, *K. piersonii* DSM 108198^T, *P. agglomerans* DSM 3493^T, or *P. brenneri* DSM 24232^T). The whole-genome sequence of the nearest neighbor was further selected for ANI evaluation. The ANI value for all strain comparisons was 99%.

[b] Hyphenated designations indicate the flight number followed by the location; for example, F1-2 indicates flight 1 and location 2. JPL, Jet Propulsion Laboratory; SAF, Spacecraft Assembly Facility; CRV, commercial resupply vehicle.

[c] WHC, waste and hygiene compartment; PMM, permanent multipurpose module; FC, field control (a sampling wipe was exposed to the air for 120 s at the center of node 2); N1-O4, node 1 overhead 4.

[d] Contigs that were less than 200 nucleotides long were not analyzed.

sequencing data (9). Adapter trimming and quality filtering were carried out using the software fastp (v0.20.0) to perform quality control (10). The cleaned sequences were assembled using SPAdes (v3.11.1) (11). The N_{50} values, numbers of contigs, and total genome lengths were generated using QUAST (v5.0.2) and used to assess the quality of the final assembly (12). The average nucleotide identity (ANI) values were calculated by comparing all strains to their respective type strains, and their taxonomic affiliations and genome statistics are given in Table 1 (13). The NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (v.4.11 and v.4.12) was used for genome annotation. Default parameters were used for all software.

Data availability. This WGS project was deposited in DDBJ/ENA/GenBank (accession numbers are given in Table 1 [BioProject accession no. PRJNA635942]) and also deposited in the NASA GeneLab database (accession no. GLDS-302 and GLDS-311). The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

We thank astronauts Captain Terry Virts and Commander Jeffrey Williams for collecting samples aboard the ISS and the Implementation Team at NASA Ames Research Center for coordinating this effort. We thank Ryan Kemp (Zymo Corp.) for extracting DNA and Dan Butler (Weill Cornell Medicine) for shotgun sequencing using the NovaSeq platform.

Part of the research described in this publication was carried out at the Jet Propulsion Laboratory (JPL), California Institute of Technology, under a contract with NASA. Government sponsorship is acknowledged. This research was funded by 2012 Space Biology project NNH12ZTT001N grant 19-12829-26 under task order NNN13D111T awarded to K.V.; this also funded a postdoctoral fellowship for J.M.W., a JPL graduate fellowship for R.D., and a subcontract to Biotia, Inc.

REFERENCES

1. Christoff AP, Sereia AFR, Cruz GNF, Bastiani DC, Silva VL, Hernandes C, Nascente APM, Reis AA, Viessi RG, Marques ASP, Braga BS, Raduan TPL, Martino MDV, Menezes FG, Oliveira LFV. 2020. One year cross-sectional study in adult and neonatal intensive care units reveals the bacterial and antimicrobial resistance genes profiles in patients and hospital surfaces. PLoS One 15:e0234127. https://doi.org/10.1371/journal.pone.0234127.
2. Wong WC, Oubre C, Mehta SK, Ott CM, Pierson DL. 2017. Preventing infectious diseases in spacecraft and space habitats, p 3–17. *In* Hurst CJ (ed), Modeling the transmission and prevention of infectious disease. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-60616-3_1.
3. Taylor PW. 2015. Impact of space flight on bacterial virulence and antibiotic susceptibility. Infect Drug Resist 8:249–262. https://doi.org/10.2147/IDR.S67275.
4. Cruz AT, Cazacu AC, Allen CH. 2007. *Pantoea agglomerans*, a plant pathogen causing human disease. J Clin Microbiol 45:1989–1992. https://doi.org/10.1128/JCM.00632-07.
5. Singh NK, Wood JM, Mhatre SS, Venkateswaran K. 2019. Metagenome to phenome approach enables isolation and genomics characterization of *Kalamiella piersonii* gen. nov., sp. nov. from the International Space Station. Appl Microbiol Biotechnol 103:4483–4497. https://doi.org/10.1007/s00253-019-09813-z.
6. Singh NK, Bezdan D, Checinska Sielaff A, Wheeler K, Mason CE, Venkateswaran K. 2018. Multi-drug resistant *Enterobacter bugandensis* species isolated from the International Space Station and comparative genomic analyses with human pathogenic strains. BMC Microbiol 18:175. https://doi.org/10.1186/s12866-018-1325-2.
7. Doijad S, Imirzalioglu C, Yao Y, Pati NB, Falgenhauer L, Hain T, Foesel BU, Abt B, Overmann J, Mirambo MM, Mshana SE, Chakraborty T. 2016. *Enterobacter bugandensis* sp. nov., isolated from neonatal blood. Int J Syst Evol Microbiol 66:968–974. https://doi.org/10.1099/ijsem.0.000821.
8. Checinska Sielaff A, Urbaniak C, Mohan GBM, Stepanov VG, Tran Q, Wood JM, Minich J, McDonald D, Mayer T, Knight R, Karouia F, Fox GE, Venkateswaran K. 2019. Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. Microbiome 7:50. https://doi.org/10.1186/s40168-019-0666-x.
9. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
10. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
11. 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.
12. 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.
13. Yoon S-H, Ha S, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4.