Draft Genome Sequences of *Klebsiella* Species Isolated from the International Space Station

Samuel A. Solomon[1], Achintya R. Bharadwaj[1], Nitin K. Singh[1], Jason M. Wood[1], Marilyne Debieu[2], Niamh B. O'Hara[2,3], Christopher E. Mason[4], and Kasthuri Venkateswaran[1*]

[1]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA*
[2]*Biotia, New York, NY, USA*
[3]*The Department Cell Biology/College of Medicine, SUNY Downstate Health Sciences University, Brooklyn, NY*
[4]*The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA*

*Corresponding author: Kasthuri Venkateswaran
California Institute of Technology, Jet Propulsion Laboratory
Biotechnology and Planetary Protection Group; M/S 89-2
4800 Oak Grove Dr., Pasadena, CA 91109
Tel: (818) 393-1481; Fax: (818) 3934176
E-mail: kjvenkat@jpl.nasa.gov

E-mail addresses:
SAS: solomons@caltech.edu
ARB: achintya.r.bharadwaj@jpl.nasa.gov
NKS: nitin.k.singh@jpl.nasa.gov
JMW: jason.m.wood@jpl.nasa.gov
MD: debieu@biotia.io
NBO: niamh.ohara@cornell.edu
CEM: chm2042@med.cornell.edu
KV: kjvenkat@jpl.nasa.gov

ABSTRACT

Isolated across four locations aboard the International Space Station (ISS), 10 bacterial strains were compared using a whole-genome sequence (WGS) analysis and phylogenetically identified as *Klebsiella*. The WGS will aid in comparative genomics studies of ISS Klebsiella strains with Earth counterparts to gain insight into their adaptation to space conditions.

ANNOUNCEMENT

The genus *Klebsiella* was discovered by Carl Friedlander in 1882 from the lungs of patients who died of pneumonia [1]. In infected individuals, *Klebsiella* species can populate the gastrointestinal tract and nasopharynx, surviving on mucosal surfaces, and are known for being highly virulent and resistant to antibiotics [2,3]. When found on Earth, this genus of bacterial pathogens has varying degrees of pathogenicity, which can lead to severe breathing problems requiring a ventilator and on-site rapid treatment [4]. However, when exposed to space conditions, *Klebsiella* species might potentially become immunogenic and therefore pose a risk to already immunocompromised astronauts aboard the ISS [5,6]. Since microgravity and radiation in space are reported to induce multiple genetic adaptations within microbial species, such as structural modifications of the cell membrane, that can subsequently alter their virulence [5], the strains identified here might pose a problem to the health of astronauts. A genetic comparison is hence imperative to illuminate more details on the survival of *Klebsiella* species, which have gained more attention as classical *Klebsiella pneumoniae* (cKp) and its hypervirulent pathotype (hvKp) are becoming increasingly more resistant to various antibiotics such as carbapenems [3,4,7]. The pathogenicity and resistance of the members of the genus *Klebsiella* might potentially create a problem for space travel, specifically for the safety of astronauts.

Several strains of *Klebsiella* species belonging to *K. aerogenes* (n=1), *K. pneumoniae* (n=1), and *K. quasipneumoniae* (n=8) were isolated from various locations on ISS environmental surfaces [8]. The flight number, location, and other sampling characteristics of the ISS *Klebsiella* isolates are detailed in Table 1. Briefly, the environmental samples collected from the ISS and subsequently brought down to Earth at room temperature were aseptically handled as per the established procedures [8], where 100 µl aliquots of concentrated samples were spread onto either R2A (25˚C; 7 days) or blood agar (37˚C; 2 days) media for isolating microorganisms. After morphological observation, pure colonies were archived at minus 80˚C until further analyses. Cultures of the 10 *Klebsiella* strains were grown overnight on tryptic soy agar at 25˚C until harvested before DNA was extracted using the ZymoBIOMICS DNA Magbead Kit.

Genomes were sequenced using the Illumina (San Diego, CA) Nextera Flex protocol for library preparation, and a NovaSeq 6000 S4 flowcell PE 2x150 was used for paired-end sequencing. The quality was assessed with FastQC (v0.11.7) [9]. Adapter trimming and quality filtering were then carried out with fastp (v0.20.0) [10]. After quality control, the sequences were assembled using SPAdes (v3.11.1) [11]. To assess the quality of the final sequences, a QUAST (v5.0.2) analysis [12] was performed to check the N50 values, number of contigs, and the total genome length (Table 1). The GC contents are *K. aerogenes* 54.96%, *K. pneumoniae* 57.25%, and *K. quasipneumoniae* 58.11-58.13%. The 16S rRNA gene sequences of the *Klebsiella* strains were compared to find out the nearest neighbor and phylogenic characterization was determined by comparing the Average Nucleotide Identity using EZBioCloud calculator [13] with respective type strains (*K. aerogenes* ATCC 13048^{T}, *K. pneumoniae* ATCC 13883^{T}, and *K. quasipneumoniae* 01A030^{T}). Default parameters were used with all software.

Data availability. This WGS project has been deposited at GenBank and the GenBank and raw read accession numbers are given in Table 1. The BioProject accession numbers are: PRJNA635942, PRJNA640688, and PRJNA640693. WGS are also deposited in NASA GeneLab (GLDS-302, GLDS-309, and GLDS-311). The version described in this paper is the first version.

ACKNOWLEDGEMENTS

The research described in the article above was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA. We would like to thank astronauts Captain Terry Virts and Commander Jeffrey Williams for collecting samples aboard the ISS, and the Implementation Team at NASA Ames Research Center (Fathi Karouia) for coordinating this effort. Authors thanks Ryan Kemp, Zymo Corporation for extracting DNA and Dan Butler, Cornell Medicine in generating shotgun sequencing using NovaSeq. We would also like to acknowledge the JPL supercomputing facility staff, notably Narendra J. Patel (Jimmy) and Edward Villanueva, for their continuous support in providing the best possible infrastructure for BIG-DATA analysis. This research was funded by a 2012 Space Biology NNH12ZTT001N grant nos. 19-12829-26 under Task Order NNN13D111T award to KV, which also funded post-doctoral fellowship for JMW and a subcontract to Biotia, Inc.

REFERENCES

1. Friedlaender, C. 1882. Ueber die Schizomyceten bei der acuten fibrösen Pneumonie. Archiv f. pathol. Anat. 87, 319–324. DOI:10.1007/BF01880516
2. Martin, R. M., & Bachman, M. A. 2018. Colonization, Infection, and the Accessory Genome of *Klebsiella pneumoniae*. *Frontiers in cellular and infection microbiology*, *8*, 4. https://doi.org/10.3389/fcimb.2018.00004
3. Garbati, M. A., & Al Godhair, A. I. 2013. The growing resistance of *Klebsiella pneumoniae*; the need to expand our antibiogram: case report and review of the literature. African journal of infectious diseases, *7*(1), 8–10. DOI:10.4314/ajid.v7i1.2
4. Jorge Domínguez-Andrés, Marc Eleveld, Georgios Renieris, Thomas J. Boltje, Rob J. Mesman, Laura van Niftrik, Sam J. Moons, Petra Rettberg, Jos W.M. van der Meer, Evangelos J. Giamarellos-Bourboulis, Huub J.M. Op den Camp, Marien I. de Jonge, and Mihai G. Netea. 2020. Growth on Carbohydrates from Carbonaceous Meteorites Alters the Immunogenicity of Environment-Derived Bacterial Pathogens. Astrobiology. Ahead of print. DOI:10.1089/ast.2019.2173
5. Taylor PW. 2015. Impact of space flight on bacterial virulence and antibiotic susceptibility. *Infect Drug Resist*. 2015;8:249-262. DOI:10.2147/IDR.S67275
6. Russo TA, Marr CM. 2019. Hypervirulent *Klebsiella pneumoniae*. Clin Microbiol Rev. 2019;32(3):e00001-19. DOI:10.1128/CMR.00001-19
7. 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. DOI:10.1186/s12866-018-1325-2

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(1):50 doi:10.1186/s40168-019-0666-x
9. Andrews S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: 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. DOI:/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. DOI: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. DOI: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. DOI:10.1007/s10482-017-0844-4

Table 1: Metadata and genome statistics of *Klebsiella* strains isolated from various ISS environmental surfaces during Microbial Tracking-1 flight project

Sample Name	Nearest identity[1]	ANI[2]	GenBank accession	Raw sequence accession	Flight / Location	Location Description	No. of Contigs	Genome Size (bp)	N50 (bp)	Median Coverage	Raw Read Counts (x10^6)	QC Reads (x10^6)
IIIF7SW-P1	*K. aerogenes* ATCC 13048^T	98.66	JACBPC000000000	SRR12071884	F3-7	Lab 3	25	5,155,046	502,580	346.88	21.6	21.5
F3-2P(2*)	*K. pneumoniae* ATCC 13883^T	99.01	JACAUF000000000	SRR12068826	F3-2	WHC	75	5,489,009	332,420	293.62	19.0	18.9
IF1SW-B2	*K. quasipneumoniae* 01A030^T	96.54	JABWPD000000000	SRR11884995	F1-1	Cupola	24	5,192,853	601,624	533.04	31.4	31.2
IF1SW-P3	*K. quasipneumoniae* 01A030^T	96.60	JABWOZ000000000	SRR11885008	F1-1	Cupola	28	5,192,422	600,611	467.41	28.2	28.0
IF1SW-P4	*K. quasipneumoniae* 01A030^T	96.55	JABWPA000000000	SRR11885009	F1-1	Cupola	29	5,192,468	343,111	297.32	18.6	18.6
IF2SW-B3	*K. quasipneumoniae* 01A030^T	96.54	JABWPC000000000	SRR11885011	F1-2	WHC	24	5,192,354	600,611	712.50	39.8	39.6
IF2SW-P1	*K. quasipneumoniae* 01A030^T	96.61	JABWPB000000000	SRR11885010	F1-2	WHC	27	5,192,385	601,139	480.80	30.2	30.1
IIIF3SW-P1	*K. quasipneumoniae* 01A030^T	96.63	JABXWM000000000	SRR12070037	F3-3	ARED	31	5,154,415	766,557	286.96	18.0	17.8
F3-6P(1)	*K. quasipneumoniae* 01A030^T	96.57	JABXWL000000000	SRR12070038	F3-6	PMM	28	5,196,291	1,009,008	371.62	23.2	22.9
F3-6P(2)	*K. quasipneumoniae* 01A030^T	96.60	JABXWK000000000	SRR12070039	F3-6	PMM	31	5,196,372	1,009,008	367.78	23.4	23.1

Abbreviations: F1: Flight 1; F3: Flight 3; ARED: Advanced Resistive Exercise Device; WHC: Waste and Hygiene Compartment; PMM: Permanent Multipurpose Module Port 1; ANI: Average Nucleotide Identity

[1]The 16S rRNA gene sequences were retrieved from the WGS of the queried genome and BLAST analyzed against type strains of all 16S rRNA sequences in the NCBI database. The bacterial species identity was determined when the queried sequence was showing >97.5% similarity with the 16S rRNA gene sequences of the type strain. The WGS of the nearest neighbor listed was further selected for ANI evaluation.

[2]ANI calculations were carried out as per EZBioCloud ANI Calculator (https://www.ezbiocloud.net/tools/ani) by comparing with the listed type strain.