

Development of Computational Environmental Microbiome Workflows for the Laboratory and the International Space Station

G. Marie Sharp¹, Sarah Stahl-Rommel², Hang N. Nguyen², Christian L. Castro², Miten Jain³, and Sarah L. Castro-Wallace⁴

¹KBR, Houston, TX 77058, USA, ²JES Tech, Houston, TX 77058, USA, ³UCSC Genomics Institute, University of California, Santa Cruz, CA 95064, USA, ⁴Biomedical Research and Environmental Sciences Division, NASA Johnson Space Center, Houston, TX 77058, USA

Identification of microorganisms in the spaceflight environment is critical for crew health risk assessment on the International Space Station (ISS). Since 2017, nanopore sequencing technology has been used to support the *in situ* identification of microbial species during spaceflight. Beginning in 2018, a culture-independent, swab-to-sequencer method was implemented onboard the ISS to provide a more thorough insight of the ISS microbiome. Eliminating microbial culture enables identification of difficult-to-culture organisms, reduces risks associated with potentially pathogenic cultures, and could significantly reduce the time from sample-to-answer. However, this molecular-based approach generates large metagenomic data sets that require substantial computational resources for analysis.

To process nanopore-generated sequencing data, the JSC Microbiology Laboratory established a bioinformatics workflow on Amazon EC2 under the security guidance of the NASA Science Managed Cloud Environment (SMCE). This resource allows for the development, testing, and accessing of computational tools for processing large and complex data sets. The work described here will address the downlinking of data from the ISS, the automated pipeline developed to identify targeted bacterial and fungal organisms, and the time from sampling onboard to microbial identification. The pipelines have been enhanced to address high and low biomass samples using optimization based on sample source (air, water, or surface) and type of collection (filter, colony, or swab). The resulting microbiome data can be assessed beyond microbial identifications to gain understanding toward population changes over time, potential selective environmental pressures, and evaluating correlations with a wide range of additional data sets. Metagenome analysis pipelines in development could allow for simultaneous identification of microbial species, gene function, and gene pathways present in the environment.

Beyond the ground processing, the developed analysis pipeline is currently deployed onboard the ISS to allow for near real-time assessments of the ISS microbiome. This study serves as a critical foundation for exploration missions, where rapid microbiome analyses will be required.