DOES COLLECTION TIME BIAS THE ECOLOGY OF CLEANROOM AIR SAMPLES?

A. B. Regberg1, C.P. Cohen2, C.L. Castro3 R.E. Davis4, C.G. Kmiecik2 and F. Mazhari2, 1Astromaterials Research and Exploration Science Division, NASA Johnson Space Center (aaron.b.regberg@nasa.gov), 2Jacobs, JETS Contract, NASA Johnson Space Center, 3JES Tech, NASA Johnson Space Center 4Texas State, JETS Contract, NASA Johnson Space Center

Introduction

Microbial monitoring of astromaterials collections has taken on increased importance with the return of biologically sensitive samples from the asteroids Ryugu and Bennu and the initiation of the Mars Sample Return Program. Terrestrial bacteria and fungi can alter the mineralogy and organic composition of our collections causing irreversible contamination of pristine samples and increasing the risk of false positives for life detection measurements. NASA has conducted routine microbial monitoring of its existing collections since 2018[1]. Initial monitoring focused on surface samples collected with foam swabs. Although, airborne microbiology is often decoupled from surface microbiology in the built environment[2] culture-based air sampling techniques like impactors were not compliant with existing contamination control requirements. Bringing organic rich media, gelatin or liquids into curation cleanrooms presents an unacceptable risk to pristine samples. In 2022 NASA purchased a materials complaint air sampler and began collecting air samples from the cleanrooms in addition to surface samples[3]. The new instrument uses an electret filter to collect samples that are suitable for cultivating organisms or for direct DNA sequencing. Preliminary DNA sequencing results appeared to indicate that longer sampling times biased the microbial community in favor of hearty, spore-forming bacteria[3]. We present the results of a study comparing overnight sampling (17 hours) to short (1 hour) sampling of unoccupied curation cleanrooms. The results will help us optimize our monitoring protocols and develop a more detailed inventory of the ecology of astromaterials curation cleanrooms.

Methods: We analyzed 72 paired air samples from six different cleanrooms including the meteorite processing lab (ISO 7 equivalent, 16 samples), the lunar lab (ISO 6 equivalent, 10 samples), the stardust lab (ISO 5 equivalent 14 samples), the OSIRIS-REx lab (ISO 5 equivalent, 12 samples), the Hayabusa2 lab (ISO 5 equivalent, 14 samples), and the Genesis lab (ISO 4 equivalent, 6 samples). All the samples were collected with an InnovaPrep Bobcat air sampler operating at a sampling rate of 200 L/min. The sampler operates for 5 minutes out of every 20 minute period. Half of the samples were collected by filtering 3,000L (15 min. of active sampling) of air across an electret filter for one hour. The rest of the samples were collected by filtering approximately 51,000 L air across the filter overnight (~17 hours, 255 min. of active sampling).

Cells were eluted from the filter using 6-7 ml of pressurized 0.15% tween 20 in PBS (phosphate buffered saline). This liquid was used to cultivate bacteria according to previously published methods[1,4,5] and for DNA extraction and next generation sequencing. DNA was extracted with a Qiagen MagAttract PowerMicrobiome kit[6]. To identify bacteria and archaea, the 16S rRNA gene was amplified using Earth Microbiome primers for the V4 region [7]. The amplified DNA was sequenced on an Illumina MiSeq using a V3 reagent kit. The resulting sequences were processed using DADA2 and QIIME2 as implemented on the EDGE bioinformatics platform[8–10].

Results: Only two of the 72 samples had no amplifiable DNA. Amplified DNA concentrations ranged from 2.67 – 0.272 ng/µl. The median concentration of amplified DNA for the 1 hour samples was 0.770 ± 0.368 ng/µl. The median concentration of amplified DNA for the overnight samples was 0.877 ± 0.434 ng/µl. On average the overnight samples had slightly more sequences (58,960 vs. 59,456) and ASV's (amplicon sequence variants) (60 vs 64.5) than the one hour samples, but these differences are not statistically significant. The most abundant ASV in every sample mapped to the genus *Cupravidus.* ASV's mapping to the genuses *Bacillus, Schlegelella, Thermus, and Staphylococcus* were also common.

Discussion and Future Work: Alpha diversity statistics like Shannon Entropy and Faith Phylogenetic Diversity are used to describe the diversity of organisms in a single sample. If a longer sampling time was biasing the data, we would expect to see a change in these diversity statistics vs. sample time. However, we did not observe this in our data. The median Shannon entropy was slightly higher for the overnight samples (3.773 vs 3.611) as was the Faith Phylogenetic Diversity (4.042 vs 3.596), but both values were within a standard deviation of each other for the two sampling times (Fig. 1). It is unlikely, that the longer sampling time is introducing bias into our data. We do observe a significant decrease in diversity when comparing the air samples by lab. The Genesis lab (ISO 4 equivalent) has a lower median number of ASV's (45.5) than the

other labs (62). Median values for Shannon Entropy (3.717 vs. 3.430) and Faith Phylogenetic Diversity (3.796 vs. 3.548) are also lower for Genesis, but those values are with one standard deviation of each other for the different sampling times. This is consistent with previous culture-based results suggesting that the environment in cleanrooms tends to select for a core group of organisms capable of surviving under dry, low nutrient, conditions.

The presence of the ASV's mapping to *Cupravidus* and *Thermus* in our sequencing blanks and controls suggests that several of the most common organisms in our samples represent contaminants from the reagents used to perform the DNA extractions and sequencing. Further work is needed to identify these contaminants, remove them from our data and recalculate the diversity statistics. This is a systematic error. Therefore, we do not expect removing the sequencing contaminants to change our conclusions. Longer air sample collection times appear to result in slightly higher diversity and do not bias the results towards "hardy" bacteria like spore-formers. Based on these preliminary results we conclude that sampling at least 3,000 liters of air is sufficient to capture the microbial diversity of cleanrooms, and that air samples can also be collected overnight without negatively impacting diversity. These results allow us to be flexible when designing microbial monitoring plans so that they do not interfere with routine lab activity.

References: 1. Regberg, A. B. *et al. 49th Lunar and Planetary Science Conference* (2018). 2. The United States Pharmacopeial Convention. USP General Chapter <1116> (2013). 3.Regberg, A. B., *et al. 54th Lunar and Planetary Science Conference* (2023). 4. Regberg, A. B. *et al. 53rd Lunar and Planetary Science Conference* (2022). 5. Davis, R. E.,*et al. 50th Lunar and Planetary Science Conference* (2019). 6. Qiagen. MagAttract® PowerMicrobiome® DNA/RNA EP Kit Handbook. (2018). 7. Walters, W. *et al. mSystems* **1**, (2015). 8. Callahan, B. J. *et al. Nat. Methods* **13**, 581–583 (2016). 9. Hall, M. & Beiko, R. G. *Microbiome Analysis: Methods and Protocols*113–129 (Springer, 2018). 10. Philipson, C. *et al. Bio-Protoc.* **7**, e2622 (2017).

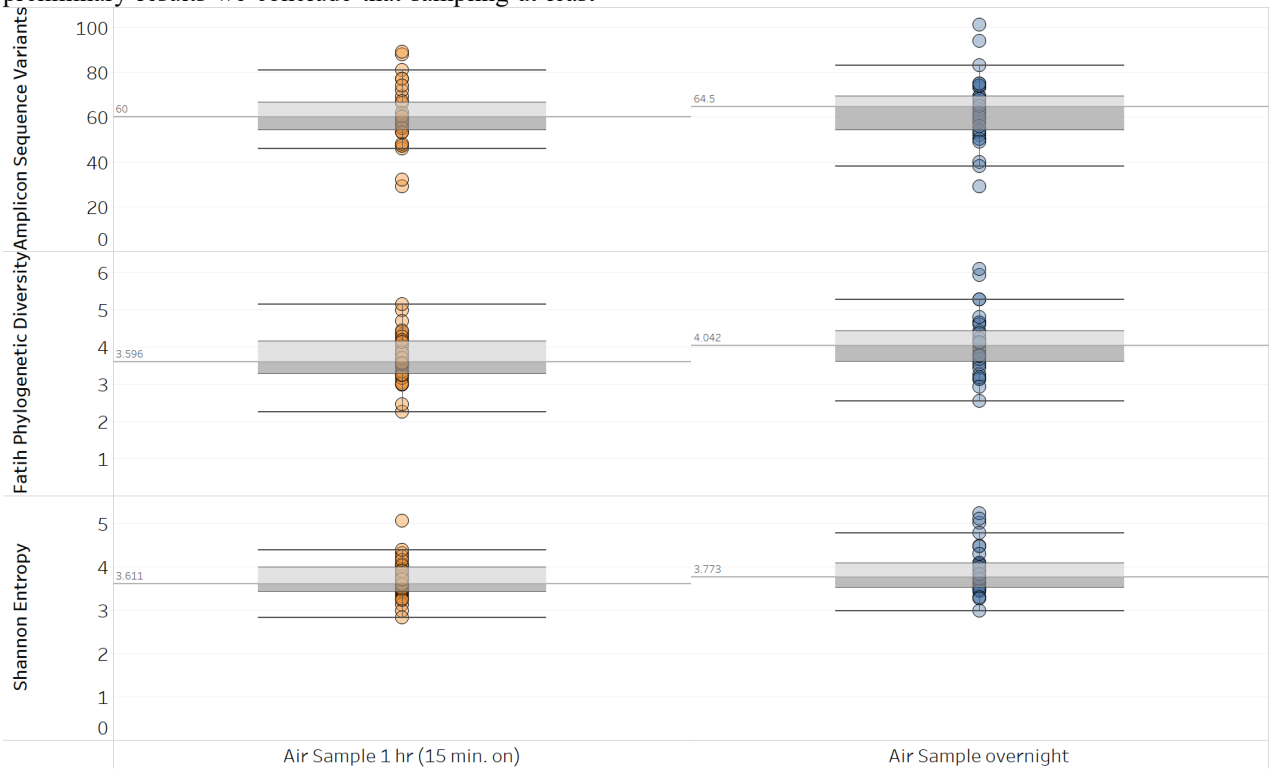

Figure 1: Alpha diversity statistics do not show a statistically significant difference based on sample collection time. Light gray lines mark the median value for each data set. Shaded areas delineate the interquartile range. Whiskers extend to 1.5 times the interquartile range. This result suggests that longer sampling times do not bias community composition towards certain strains of bacteria.