Strain-Level Identification of Spacecraft-Associated Microorganisms Using Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry

Heidi Aronson[1], Jennifer Jacobs[2], Arianna Jefferson[3], James Tan[4], Wayne Schubert[5] and Parag Vaishampayan[5]*

[1] Occidental College, Los Angeles, CA, [2]Chicago State University, Chicago, IL, [3]Bridgewater State University, Bridgewater, MA, [4]University of Michigan, Ann Arbor, MI, [5] Biotechnology and Planetary Protection Group, California Institute of Technology, Jet Propulsion Laboratory, Pasadena, CA . *contact: vaishamp@jpl.nasa.gov

Abstract

The microbial communities found in a spacecraft-assembly cleanroom environment are often limited to specific, non-diverse group of microorganisms. Standard methods of identification (16S rRNA gene sequencing) are time-consuming, costly, and often cannot fully represent the strain-level variation present in these communities. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been used as a rapid, high throughput, accurate and sensitive method for bacterial identification; however, available databases provided by Bruker Daltonics do not include many of the specific isolates found in cleanroom environments. An in-house database was created using bacterial isolates with known 16S rRNA identifications collected from Phoenix Mars lander hardware and was used to validate the correlation between 16S rRNA-based identifications and MALDI-TOF identifications. 53 unique species identifications were produced from 147 Phoenix isolates based on 16S rRNA sequencing. MALDI-TOF MS was used to correctly identify 73.76% of Phoenix isolates and was able to resolve sub-species variation that was not detected by 16S sequencing. More sampling must be performed and more database entries need to be created in order to capture the full diversity of the cleanroom microbial community.

Introduction

JPL's planetary Protection group has archived the largest collection of the microbes associated with spacecraft surfaces for more than three decades. Though considered as 'gold standard', 16S rRNA gene sequencing is time-consuming, and often cannot fully resolve strain-level variations. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been used as a rapid, high throughput, accurate and sensitive method for bacterial identification using characteristic spectral patterns of ribosomal proteins. The MALDI Biotyper (Bruker Corporation) instrument provides a feasible path forward for the analysis of spacecraft microbial isolates; however, Bruker's proprietary databases do not include many of the specific isolates found in the cleanroom environment. An in-house database was created using bacterial isolates with known 16S rRNA identifications collected from Phoenix Mars Lander and Mars Science Laboratory hardware and was used to validate the correlation between 16S rRNA-based identifications and MALDI-TOF identifications.

Rationale Behind the Research and Objectives

The databases provided by Bruker Corporation are not comprehensive enough to encompass specific bacterial species found on the spacecraft and associate surfaces. In absence of such a database, we were unable to accurately identify isolates and got false positive identifications.

Mx	Detected Species	Log(Score)
	Bacillus muralis DSM 16388T DSM	1.579
	Staphylococcus simiae DSM 17639 DSM	1.345
	Flavobacterium flevense DSM 1076T HAM	1.313
	Erysipelothrix rhusiopathiae EDQM Serotyp2 FLR	1.297
	Lactobacillus alimentarius DSM 20249T DSM	1.248
	Staphylococcus simiae DSM 17638 DSM	1.216
	Bacillus mojavensis DSM 9205T DSM	1.208
	Bacillus atrophaeus DSM 675 DSM	1.202
	Arthrobacter qandavensis DSM 15046T DSM	1.179
	Lactobacillus gastricus DSM 16045T DSM	1.176

Figure 1. The spacecraft-associated microbe did not identify with any entries in the Bruker database (log score <2.0)

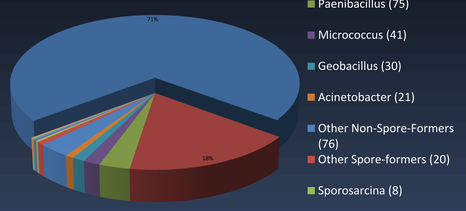

Figure 2. The Venn diagram depicts the abundance of the most dominant genera as identified by 16S rRNA gene sequencing.

OBJECTIVES:

1. Create an in-house database of Mass Spectrometry Profiles (MSP) for MALDI-TOF identification.
2. Real Time Identification (RTC) of space-craft associated microbes from the Phoenix Mars Lander mission using MALDI-TOF to validate the correlation between 16s RNA and MALDI based identifications.

Testing Approach

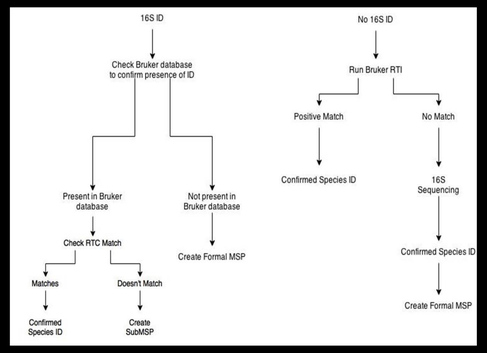

Figure 5. Flow chart for individual isolates.

- If 16S rRNA gene sequence is not available, isolate will be identified using RTC.
- If no identification, the isolate will be sent for 16S rRNA gene sequencing
- MSP will be created after confirming taxonomic affiliation using 16S rRNA gene sequence
- For ideal Bruker-based identification, comprehensive database is necessary

Methods

Bioinformatics Workflow

1. Align sequences using multiple sequence alignment (Clustal Omega)
2. Cluster sequences at 99% 16s rRNA similarity using a furthest-neighbor algorithm to form operational taxonomic units- OTUs (mothur)
3. Choose representative isolate, assign taxonomic identification using a curated database (EzTaxon) and create MSP.
4. Run RTCs for other members of the same OTUs cluster, to verify that they match to their representative MSP.

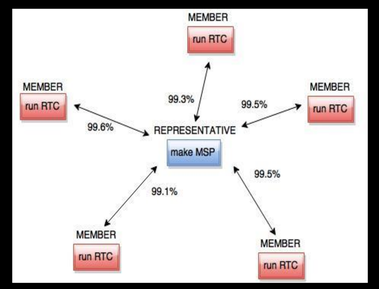

Figure 6. MALDI verification of OTU members

MALDI-TOF Workflow

1. Isolates are stored in -80^{O}C freezer.
2. Isolates are revived from a glycerol or cryobead stock.
3. Isolates are grown overnight in a 32^{O}C incubator.
4. Direct transfer of colony onto MALDI target is performed.
5. Samples prepared for MALDI-TOF target (HCCA matrix).
6. Run samples using software dependent on whether the isolate is for RTC or MSP.
7. Analyze the Results

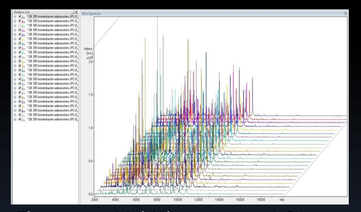

Figure 7. Multiple spectra (48) are obtained for each isolate that is then processed into a consensus spectra to form an MSP.

A1

PF3-1.5
2015-07-17T10:49:28.314
BDAL, Phoenix DT

Matched Pattern	Score Value
Bacillus oceanisediminis PF4F_2.2	2.365
Bacillus firmus DSM 12T DSM	1.597

Figure 8. Example of an isolate matching to the MSP created for the in-house database using spacecraft-associated microbes. The isolate had a confident identification with the in-house database (log score >2.0 = minimum criteria to confirm a match), but did not match well to Bruker's database (log score <2.0).

Results

Phoenix Results

- Total Phoenix isolates: 168
- 53 MSPs created; 26 single-member OTUs
- 143 RTCs run: 73.43% correctly identified (105); 3.5% incorrectly identified (5); 13.29% no identification (19); 9.79% flatline spectrum (14).

MSL Results

- Total MSL isolates: 46
- 14 MSPs created
- 44 RTCs run: 61.36% correctly identified (27); 15.91% incorrectly identified (7); 20.45% no identification (9); 2.27% flatline spectrum (1).

Strain-Level Variation

- Strain-level variation caused certain strains not to be identified with representative MSP
- Created sub-MSP profiles with more distant strains to improve identification

Discussion and Future Work

Discussion

- MALDI-TOF is very sensitive to strain-level variations in protein profile that 16S rRNA sequence identification cannot resolve.
- No identification is the result of non-comprehensive database, which will decrease as the database improves.
- Percentage of incorrect identifications are very low (3.5%).

Future Work

- Update in-house database with MSPs from remaining OTUs. From 1127 available sequences from all missions, 533 OTUs created from 99% sequence similarity.
- Classify new isolates from future missions using MALDI-TOF MS.

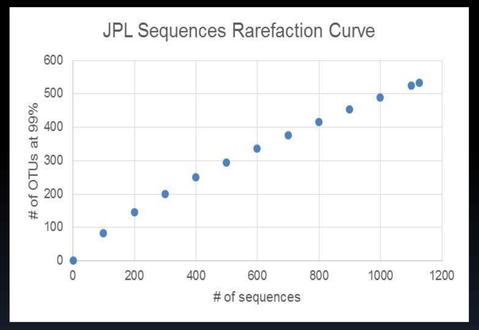

Figure 9. Rarefaction curve of number of sequences vs. number of OTUs. Ideally should plateau, but more isolate collection and sequencing should be done to capture full microbial diversity of spacecraft surfaces.

Acknowledgements

We would like to acknowledge the Planetary Protection Group members, Melissa Jones (group supervisor), and Laura Newlin (acting group supervisor) for their support for this research. We would like to thank Mars Exploration Program Office for their financial support. The research described here was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We thank Kaveri Barros from Bruker Daltonics for her technical assistance

Selected References

1. Bizzini, A. et al. (2011). Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry as an Alternative to 16S rRNA Gene Sequencing for Identification of Difficult-To-Identify Bacterial Strains. *Journal of Clinical Microbiology.* 49(2): 693-696.
2. Cherkaoui, Abdessalam, et al. (2010) Comparison of Two Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry Methods with Conventional Phenotypic Identification for Routine Identification of Bacteria to the Species Level. *Journal of Clinical Microbiology*. 48(4): 1169-1175.
3. Mellmann, A. et al. (2008). Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight Mass Spectrometry in Comparison to 16S rRNA Gene Sequencing for Species Identification of Nonfermenting Bacteria. *Journal of Clinical Microbiology.* 46(6): 1946-1954.
4. NASA Office of Planetary Protection, 22 Dec. 2014. "Methods and Implementation." *National Aeronautics and Space Administration.* http://planetaryprotection.nasa.gov/methods (12 Feb. 2015).
5. Williams, TL, et al. (2003). Experimental Factors Affecting the Quality and Reproducibility of MALDI TOF Mass Spectra Obtained from Whole Bacteria Cells. *Journal of the American Society for Mass Spectrometry.* 14(4): 342-351.
6. Wunschel, SC, et al. (2005). Bacterial Analysis by MALDI-TOF Mass Spectrometry: An Inter-Laboratory Comparison. *Journal of the American Society for Mass Spectrometry.* 16(4): 456-462.