8 July 2026

Metagenomics in Spaceflight: Establishing an Implementation Roadmap

Jared T. Broddrick[1], Jonathan E. Allen[2], Jyothi Basapathi Raghavendra[3], Christopher E. Carr[4,5], Sarah L. Castro-Wallace[6], Rodrigo Coutinho De Almeida[7], Christina L. Davis[8], Todd DeSantis[9], Emily Eloe-Fadrosh[10], Noah Fierer[11], Samrawit Gebre[1], Stefan J. Green[12], David Grinspoon[13], Alison L. Halpin[14], Alexis M. Hamilton[15], Erica M. Hartmann[16,17,18], Christina K. Johnson[1], Shunta Kimura[19], Nanako Kuze[20], Jennie H. Kwon[21], Erin Lalime[22], Kristopher M. Locken[23], Alexander Mahnert[24], Chiranjit Mukherjee[25], Anna Maria Niewiadomska[26], Barbara Novak[27], Karen Olsson-Francis[28], Dominic O'Neil[29], Sandra Ortega Ugalde[30], Bernhard O. Palsson[31], Petra Rettberg[32], Lorna Richardson[33], Sophie A. Simon[34], David J. Smith[13], Sarah Stahl-Rommel[35], Yohey Suzuki[36], Shino Suzuki[37], Scott W. Tighe[38], Parag Vaishampayan[1], Kasthuri Venkateswaran[39], Lyle Whyte[8], Maria-Paz Zorzano[40], Takashi Yamazaki[41], Silvio Sinibaldi[42], J. Nick Benardini[22]

1. Space Biosciences, NASA Ames Research Center
2. Lawrence Livermore National Laboratory
3. Northumbria University
4. School of Aerospace Engineering, Georgia Institute of Technology
5. School of Earth and Atmospheric Sciences, Georgia Institute of Technology
6. Biomedical Research and Environmental Sciences Division, NASA Johnson Space Center
7. Telespazio Belgium SRL for the European Space Agency, ESTEC
8. Department of Natural Resource Sciences, McGill University
9. Resilient Biotics
10. Environmental Genomics Group, Lawrence Berkeley National Laboratory
11. University of Colorado Boulder
12. Genomics and Microbiome Core Facility, Rush University Medical Center
13. Planetary Science Division, NASA Headquarters
14. Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention
15. Department of Food Science & Technology, Virginia Tech
16. Center for Synthetic Biology, Northwestern University
17. Department of Civil & Environmental Engineering, Northwestern University
18. Division of Pulmonary Medicine, Northwestern University
19. Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency
20. Massachusetts Institute of Technology

 — https://doi.org/10.22541/essoar.15005823/v1

21. Department of Medicine, Northwestern University Feinberg School of Medicine
22. Office of Safety and Mission Assurance, NASA Headquarters
23. Zymo Research Corporation
24. Diagnostic & Research Institute of Hygiene, Microbiology and Environmental Medicine, Medical University of Graz
25. Battelle Memorial Institute
26. Oxford Nanopore Technologies Inc.
27. Blue Marble Space Institute for Science, NASA Ames Research Center
28. AstrobiologyOU Open University
29. QIAGEN GmbH
30. Life Support and Physical Sciences Instrumentation Section, European Space Agency
31. Department of Bioengineering, University of California
32. Institute of Aerospace Medicine, German Aerospace Center
33. European Bioinformatics Institute (EMBL-EBI
34. Faculty of Chemistry, University of Duisburg-Essen
35. JES Tech
36. Department of Earth and Planetary Science, University of Tokyo
37. Geobiology and Astrobiology Laboratory, RIKEN
38. Environmental Microbiome Engineering Research Group, University of Vermont
39. Dept of Space Studies, University of North Dakota
40. Centro de Astrobiología (CAB, CSIC-INTA
41. Safety and Mission Assurance Department, Japan Aerospace Exploration Agency
42. Independent Safety Office, European Space Agency

Abstract

As space agencies prepare for complex missions such as human exploration of Mars and life detection, Planetary Protection (PP) risk assessments would benefit from extending beyond traditional microbial monitoring methods like the Standard Spore Assay (SSA). Metagenomics offers a powerful complementary approach, enabling detection of non-culturable organisms and providing functional insights essential for risk-informed decision-making. However, operational best practices would include rigorous validation, standardization, and integration into PP frameworks. In November 2024, NASA Ames Research Center hosted a workshop with experts from academia, government, and industry to assess the readiness of metagenomic methods for PP and outline a roadmap for implementation. Discussions addressed six key areas: (1) safetycritical decision making, (2) microbial dark matter, (3) ultra-low biomass metagenomics, (4) bioinformatics and databases, (5) technology development and automation, and (6) human health and the built environment. Cross-cutting recommendations included

adopting tiered testing frameworks, developing shared reference standards, standardizing protocols and metadata, building probabilistic risk models, and investing in automation and contamination-aware technologies. These actions aim to enable metagenomics as a robust tool for Planetary Protection in future space missions.

Metagenomics in Spaceflight: Establishing an Implementation Roadmap

Authors

Broddrick, Jared T.[1], Jonathan E. Allen[2], Jyothi Basapathi Raghavendra[3], Christopher E. Carr[4,5], Sarah L. Castro-Wallace[6], Rodrigo Coutinho De Almeida[7], Christina L. Davis[8], Todd DeSantis[9], Emily Eloe-Fadrosh[10], Noah Fierer[11], Samrawit Gebre[1], Stefan J. Green[12], David Grinspoon[13], Alison L. Halpin[14], Alexis M. Hamilton[15], Erica M. Hartmann[16,17,18], Christina K. Johnson[1], Shunta Kimura[19], Nanako Kuze[20], Jennie H. Kwon[21], Erin Lalime[22], Kristopher M. Locken[23], Alexander Mahnert[24], Chiranjit Mukherjee[25], Anna Maria Niewiadomska[26], Barbara Novak[27], Karen Olsson-Francis[28], Dominic O'Neil[29], Sandra Ortega Ugalde[30], Bernhard O. Palsson[31], Petra Rettberg[32], Lorna Richardson[33], Sophie A. Simon[34], David J. Smith[13], Sarah Stahl-Rommel[35], Yohey Suzuki[36], Shino Suzuki[37], Scott W. Tighe[38], Parag Vaishampayan[1], Kasthuri Venkateswaran[39], Lyle Whyte[8], Maria-Paz Zorzano[40], Takashi Yamazaki[41], Silvio Sinibaldi[42], J. Nick Benardini[22]

1: Space Biosciences, NASA Ames Research Center, Moffett Field, CA, USA
2: Lawrence Livermore National Laboratory, Livermore, CA, USA
3: Northumbria University, New Castle Upon Tyne, England, UK
4: School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA
5: School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA
6: Biomedical Research and Environmental Sciences Division, NASA Johnson Space Center, Houston, TX, USA
7: Telespazio Belgium SRL for the European Space Agency, ESTEC, Noordwijk, The Netherlands
8: Department of Natural Resource Sciences, McGill University, Quebec, Canada
9: Resilient Biotics, Durham, NC, USA
10: Environmental Genomics Group, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
11: University of Colorado Boulder, Boulder, CO, USA
12: Genomics and Microbiome Core Facility, Rush University Medical Center, Chicago, Illinois, USA
13: Planetary Science Division, NASA Headquarters, Washington D.C., USA
14: Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, GA, USA
15: Department of Food Science & Technology, Virginia Tech, Blacksburg, Virginia, USA
16: Center for Synthetic Biology, Northwestern University, Evanston, IL, USA
17: Department of Civil & Environmental Engineering, Northwestern University, Evanston, IL, USA

This document was created with assistance from MS Copilot to summarize conference transcripts. The content has been reviewed and edited by all authors. More information on the extent and nature of AI usage is included in the Methods or Acknowledgements sections of the document.

18: Department of Medicine, Division of Pulmonary Medicine, Northwestern University, Chicago, IL, USA
19: Institute of Space and Astronautical Science, Japan Aerospace Exploration Agency, Kanagawa, Japan
20: Massachusetts Institute of Technology, Cambridge, MA, USA
21: Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
22: Office of Safety and Mission Assurance, NASA Headquarters, Washington D.C., USA
23: Zymo Research Corporation, Irvine, CA, USA
24: Diagnostic & Research Institute of Hygiene, Microbiology and Environmental Medicine, Medical University of Graz, Graz, Austria
25: Battelle Memorial Institute, Columbus, Ohio, USA
26: Oxford Nanopore Technologies Inc., New York, NY, USA
27: Blue Marble Space Institute for Science, Space Biosciences, NASA Ames Research Center, Moffett Field, CA, USA
28: AstrobiologyOU Open University, Milton Keynes, UK
29: QIAGEN GmbH, Germany
30: Life Support and Physical Sciences Instrumentation Section, Directorate of Technology Engineering and Quality, European Space Agency, Noordwijk, The Netherlands
31: Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA
32: Institute of Aerospace Medicine, German Aerospace Center, Cologne, Germany
33: European Bioinformatics Institute (EMBL-EBI), Cambridgeshire, UK
34: Environmental Metagenomics, Research Center One Health Ruhr of the University Alliance Ruhr, Faculty of Chemistry, University of Duisburg-Essen, Essen, Germany
35: JES Tech, Houston, TX, USA
36: Department of Earth and Planetary Science, University of Tokyo, Tokyo Japan
37: Geobiology and Astrobiology Laboratory, RIKEN, Saitama, Japan
38: Environmental Microbiome Engineering Research Group, University of Vermont, Burlington, Vermont, USA
39: Dept of Space Studies, University of North Dakota, Grand Forks, ND, USA
40: Centro de Astrobiología (CAB), CSIC-INTA, 28850 Torrejón de Ardoz, Madrid, Spain
41: Safety and Mission Assurance Department, Japan Aerospace Exploration Agency, Ibaraki, Japan
42: Independent Safety Office, European Space Agency, Noordwijk, The Netherlands

Abstract
As space agencies prepare for complex missions such as human exploration of Mars and life detection, Planetary Protection (PP) risk assessments would benefit from extending beyond traditional microbial monitoring methods like the Standard Spore Assay (SSA). Metagenomics offers a powerful complementary approach, enabling detection of non-culturable organisms and providing functional insights essential for risk-informed decision-making. However, operational best practices would include rigorous validation, standardization, and integration into PP frameworks. In November 2024, NASA Ames Research Center hosted a workshop with experts from academia, government, and industry to assess the readiness of metagenomic methods for PP and outline a roadmap for implementation. Discussions addressed six key areas: (1) safety-critical decision making, (2) microbial dark matter, (3) ultra-low biomass metagenomics, (4) bioinformatics and databases, (5) technology development and automation, and (6) human health and the built environment. Cross-cutting recommendations included adopting tiered testing frameworks, developing shared reference standards, standardizing protocols and metadata, building probabilistic risk models, and investing in automation and contamination-aware technologies. These actions aim to enable metagenomics as a robust tool for Planetary Protection in future space missions.

Introduction

The expanding exploration mission sets for space agencies worldwide motivates the need to modernize Planetary Protection (PP) technologies, protocols, and policies. Life detection missions to Mars and the outer solar system require a new level of rigor to avoid contamination of potentially habitable environments with Earth microorganisms (forward contamination). Additionally, pending sample return missions to Mars and its moons open the risk of bringing back bioactive molecules to Earth (backwards contamination). Perhaps the most pressing challenge is the human exploration of Mars, where crew health, and both forward and backwards contamination all interplay into a complex risk assessment across multiple mission areas. Recognizing this challenge, the National Aeronautics and Space Administration's (NASA's) Moon to Mars Architecture includes PP Technologies for Human Exploration as a technology gap [1]. Space agencies across the globe recognize that traditional culture-based microbial monitoring and PP methods alone are no longer sufficient due to recognition that they measure a limited subset of viable organisms present, many of which are not relevant for protection of life detection targets such as Mars and icy moons, and do not quantify non-viable organic contaminants.

The Planetary Protection Metagenomics in Spaceflight Workshop was hosted in November 2024 (https://sma.nasa.gov/sma-disciplines/planetary-protection/planetary-protection-workshop-ii/home). The workshop introduction, presented by the Planetary Protection Officers from NASA and the European Space Agency (ESA), framed the gathering as a response to this evolving landscape, with particular focus on the fact that the challenges in implementing metagenomics in PP are also present in its application across a variety of space exploration disciplines (e.g., astrobiology, human health, and

space biology). At the same time, an effective implementation roadmap would support safety-critical decision-making across multiple mission areas.

Effective implementation of -omics data can accelerate decisions about contamination risk, astronaut health, and environmental monitoring, allowing these methodologies to support operational mission execution. Many international space agencies are moving toward performance-based, risk-informed frameworks, and metagenomics is a key enabler. However, implementation requires consensus on standards, validation, and integration across domains. This workshop extended years of foundational work and was designed to accelerate progress by identifying shared needs and actionable paths forward.

Recent workshops served as this foundational work. In 2022 NASA held the Multi-Mission Metagenomics Technology Development workshop [2], which was the first to formally assess the potential of metagenomics to replace or supplement the Standard Spore Assay [SSA, also called the NASA Standard Assay (NSA)] for bioburden assessment. The workshop concluded that while metagenomics offers high information content, especially for detecting non-culturable yet functionally relevant organisms, it cannot yet fully replace existing methods. Instead, a hybrid approach was recommended: combining metagenomics with quantitative tools like digital PCR to achieve both breadth and precision.

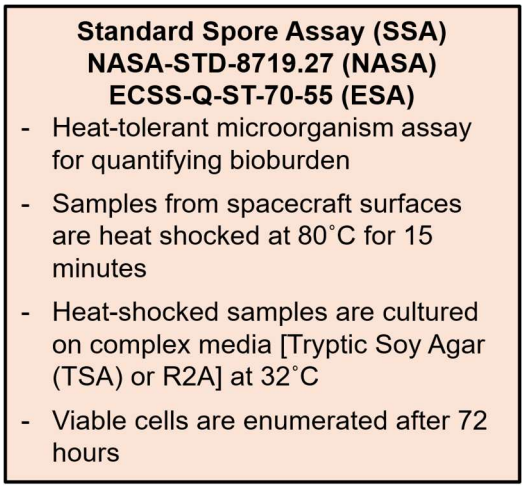

Figure 1. The Standard Spore Assay traditionally used by Space Agencies for detecting and quantifying microbial bioburden on spacecraft and spacecraft cleanrooms.

Key gaps identified included the need for improved protocols for ultra-low biomass samples, better bioinformatics pipelines, and curated reference databases. The workshop emphasized that metagenomics must be experimentally validated for spaceflight conditions and that commercial partnerships are essential for developing reagents and workflows to overcome the technological challenges relevant to PP needs. The desired end-state is a framework where metagenomics informs assay development and risk models, creating a feedback loop between discovery and decision-making.

In 2023, ESA held a workshop “Assessing Metagenomics Methods for Their Inclusion in ESA Standards” (https://atpi.eventsair.com/23m40---ppnse2023/). The workshop brought together international experts from academia, industry, and multiple space agencies. Again, this workshop was motivated by the knowledge that the Standard Spore Assay 1) is not representative of the biodiversity present, 2) identifies less than 0.1% of contaminants, 3) likely misses PP-relevant organisms for different

mission targets, and 4) forces decision making and risk assessments on incomplete information.

Key outcomes included the recognition that ESA's decades of 16S rRNA sequencing data and spacecraft assembly facility cleanroom isolates had not yet been fully leveraged. The workshop called for building a genome catalog of cleanroom organisms to support functional risk modeling and to move beyond taxonomy alone. Additionally, the workshop supported the systematic integration of metagenomics and quantitative methods alongside the SSA. All of this must then be integrated with contamination models informed with representative data of PP-relevant organisms that includes increased complexity to include more realistic, multi-stressor-informed models. Like NASA, ESA emphasized the importance of international consensus, quantitative methods, and policy evolution to support a transition toward DNA-based PP frameworks.

To this end, this workshop selected the following topics for discussion and development:

Topic	Definition and Relevance
1. Safety-Critical Decision Making	Operationalizing metagenomics data towards mission-relevant decision making and risk assessments.
2. Microbial Dark Matter	Novel and uncharacterized microorganisms and metagenomic content complicate PP risk assessments.
3. Low Biomass	Hyper-clean spacecraft cleanrooms and surfaces reduce the available microbial bioburden to near or below current limits of detection.
4. Bioinformatics and Databases	Analyzing, characterizing, modeling, cataloging, and disseminating metagenomics and its derived information.
5. Technology Needs	Technological solutions to overcome the above challenges.
6. Human Health and the Built Environment	Co-developing metagenomic solutions to support crew health and PP risk assessments.

Figure 2. Topics and rationale for the Planetary Protection gaps included in the workshop agenda.

Each session included a series of relevant presentations from experts in academia, government, and industry followed by a breakout session that was open to the public via hybrid attendance. Each session produced a series of gaps and emerging capabilities which were aggregated into an implementation roadmap.

Workshop Results

1. Safety-Critical Decision Making

The topic of Safety-Critical Decision Making was motivated by the urgent need to operationalize metagenomics as a tool for informing probabilistic risk models in PP. For an exploration roadmap emphasizing human exploration and life detection, traditional binary assays are no longer sufficient for risk-informed decisions. To move towards operationalizing metagenomics, this session explored how metagenomics is already being used in safety-critical industries like food production, where it enables early

detection and functional profiling to guide risk assessments. Case studies from ESA's JUICE (JUpiter ICy moons Explorer) and the Japan Aerospace Exploration Agency's (JAXA's) MMX (Martian Moons eXploration) missions further underscored the shift toward probabilistic frameworks in PP, with initial efforts to employ metagenomics to inform decision-making for these missions. These examples complement probabilistic frameworks that have emerged for recent exploration missions (e.g. Europa Clipper [3], [4]) and serve as valuable test beds for developing and validating realistic, consequence-based models. This session thus laid the groundwork for integrating metagenomic data into mission planning and contamination control, not as a replacement for existing tools, but as a critical layer of functional and contextual insight.

1.1 Presentations

1.1.1 Metagenomics in Food Industry Risk Assessments

The food industry has rapidly adopted omics technologies to enhance food safety [5], [6]. These tools allow for detailed characterization of microbial communities in food and production environments, enabling earlier detection of pathogens and contamination risks. This shift toward metagenomic-informed monitoring supports a broader transition to risk-based decision making in food safety, aligning with regulatory mandates. By enabling earlier detection of emerging threats and reducing reliance on destructive testing, these tools help ensure product safety while maintaining supply. The food industry's success in integrating -omics methods offers a compelling model for other domains where understanding microbial dynamics in complex environments is critical for making safety-critical decisions.

This presentation illustrated that metagenomic tools enable proactive, risk-based decision making in food safety by 1) detecting pathogens earlier and more accurately than traditional methods, 2) monitoring microbiome shifts that may indicate emerging safety threats, and 3) supporting regulatory compliance with data-driven safety assessments. These approaches mirror NASA's needs in PP and crew health monitoring, where understanding microbial presence and behavior in complex environments is critical. The food industry's rapid adaptation of omics tools offers a valuable model for similar risk-based frameworks in space exploration.

1.1.2 Advancements in molecular PP methods at NASA

Next, to employ digital PCR and metagenomics in PP assessments requires robust methodologies with the accuracy and reproducibility of the NSA. A presentation of work out of the NASA Jet Propulsion Laboratory provided an overview of the Metagenomics and Digital PCR (MDP) project, a NASA-funded effort to develop and validate molecular tools for PP. The project builds on prior workshops and decades of feasibility studies, aiming to augment the Standard Spore Assay (SSA) with a combined approach that includes shotgun metagenomics for functional insight and digital PCR for quantitative precision. The goal is to create a standardized, system-level framework that can be validated for spacecraft environments.

A key challenge addressed was the difficulty of working with ultra-low biomass samples typical of cleanroom environments. Development and testing of improved

sampling and extraction protocols, including a NASA Small Business Innovation Research SBIR-funded “squeegee” method for surface collection and optimized DNA extraction workflows recovered more DNA and enabled better detection of rare or stress-resistant organisms [7]. However, even with deep sequencing, a large portion of reads remain unclassified, highlighting the persistent issue of microbial dark matter [8]. The work emphasized that rare, potentially high-risk microbes are often missed by metagenomic pipelines due to their low abundance and DNA extraction inefficiencies. Case studies from the Phoenix and Mars 2020 missions showed that many novel strains isolated through culturing were undetectable in metagenomic datasets. This underscores the need to retain culture-based methods alongside molecular tools and to improve reference databases. The path forward includes refining DNA extraction for extremophiles, developing targeted assays for rare and PP-relevant organisms, and integrating functional gene analysis to better assess PP risks.

1.1.3 Probabilistic Risk Assessments for Recent Space Missions

Finally, there were two presentations on recent and upcoming missions led out of JAXA and ESA. These case studies highlight the emerging probabilistic-based risk assessments and entry points for metagenomic analysis in PP assessments.

The Martian Moons eXploration (MMX) mission, led by JAXA, is a sample return mission targeting Phobos, with a planned launch in 2026 and return to Earth by 2031 [9]. The mission aims to investigate the origin of Mars’ moons, the transport of materials across the early solar system, and the evolution of the Martian environment. MMX is expected to collect samples from Phobos that may contain Martian ejecta, offering a unique opportunity to study prebiotic compounds and the potential for life-related chemistry without directly sampling Mars. The Earth return component creates PP challenges. Mars is categorized as Category III, Phobos as Category II, and the return phase as Category V (unrestricted Earth return). To meet COSPAR guidelines (Committee on Space Research; https://cosparhq.cnes.fr), the mission incorporates trajectory biasing, cleanroom protocols, and probabilistic modeling to ensure the risk of backward contamination remains below the 10^{-6} threshold. Detailed simulations of radiation sterilization, impact dynamics, and sampling depth (2-8 cm) support the conclusion that viable Martian microbes are unlikely to be returned. The mission’s approach reflects a rigorous, quantifiable engineering process to assess and mitigate contamination risks.

From ESA, a presentation explored how metagenomic methods can support PP decision-making using the JUICE (JUpiter ICy moons Explorer), which launched in 2023, as a case study. The team conducted deep metagenomic sequencing of spacecraft cleanrooms, generating hundreds of millions of reads per sample and assembling high-quality metagenome-assembled genomes (MAGs). These MAGs were annotated for functional traits relevant to PP, such as sporulation, desiccation resistance, and halotolerance. The goal was to feed these data into a probabilistic framework that could assess microbial survival and proliferation risks across mission phases. To interpret the functional potential of the MAGs, the team incorporated tools to estimate traits like growth rate, growth conditions, environmental tolerance, and

metabolic capabilities, even for incomplete genomes. This systems-level approach aims to move beyond taxonomy alone, offering a more nuanced understanding of microbial risk in cleanroom environments. However, it was acknowledged that MAGs are biased toward abundant organisms, and low-abundance but high-risk microbes may be missed without complementary gene-centric and culture-based methods.

1.2 Breakout Session

The breakout session on Safety-Critical Decision-Making explored how metagenomic data can be integrated into compliance and risk-based frameworks for PP, life detection, and spacecraft cleanliness. Participants emphasized that while metagenomics offers rich information, it cannot yet serve as a standalone decision-making tool. Instead, it should be used to inform and complement existing assays like the SSA, ATP assays, and digital PCR. A recurring theme was the need to shift from prescriptive thresholds (e.g., spore counts) to probabilistic models that can accommodate evolving technologies and mission-specific risks.

Key challenges discussed included the lack of standardized protocols for sample collection, DNA extraction, and quality control across agencies and labs. Participants stressed the importance of developing shared microbial reference standards, metadata reporting practices, and validation frameworks. There was strong support for building international consensus around a tiered approach that uses metagenomics to identify organisms and functional traits of concern, then validating those findings with targeted assays and cultivation. The group also discussed the need for practical implementation strategies, including leveraging ongoing missions and cleanroom studies to build baseline data and test protocols in real-world conditions.

2. Microbial Dark Matter

A primary motivator for moving to a metagenomics-based approach to meet PP requirements is leveraging non-taxonomic information for decision making. Considering forward contamination, a microorganism must survive cleanroom strict cleaning protocols, maintain viability during stasis, and persist in the harsh abiotic environment characteristic of essentially all planetary bodies, with the exception of Earth. The ability to persist through these stressors is not a taxonomic trait, rather it is enabled by genetically encoded functions, which metagenomic analysis can elucidate if a molecular function can be ascribed to the genomic content. A persistent challenge is the fact that much of the genetic content of all organisms is uncharacterized. This “microbial dark matter,” which includes up to 35% of the genetic content even in the well-studied bacterial species *E. coli* [10], presents a significant barrier to metagenomic-based PP risk assessments. If acquired sequence data cannot be connected to a function, it cannot be leveraged for safety-critical decision making.

2.1 Presentations

2.1.1 Multi-condition Transcriptomics Reveals Molecular Function

Functional characterization brings microbial dark matter into the light. Experimental validation of a gene or genetic element’s molecular function is critical; however, most

methods are resource and time intensive and scale poorly. Recent systems biology approaches leverage a data-driven, machine learning framework for uncovering gene function in bacteria. The methodology relies on independently modulated gene modules, iModulons, derived from transcriptomic data of isolated microorganisms. These modules capture regulatory signals, the transcription factor and corresponding regulated genetic elements, that reflect specific cellular functions. Many uncharacterized genes are co-regulated with known pathways. By embedding these genes within functional modules, researchers can assign a functional annotation that enables targeted characterization. By analyzing iModulon activity across diverse conditions, researchers can infer the roles of unannotated genes, offering a scalable approach to understanding microbial dark matter. Examples from several microbial species showed iModulons revealed novel transporters, stress responses, and antibiotic resistance mechanisms [11], [12], [13]. While somewhat constrained to laboratory cultures, functional annotation via this method can be extended into other taxa, to include those with relevance to PP, via sequence and structural homology. This approach reframes dark matter as a tractable frontier, enabling systematic exploration through modular inference.

2.1.2 Emerging Bioinformatic Structures Elucidate Molecular Function

Next, work shared from Lawrence Livermore National Lab discussed pathogen-agnostic metagenomic analysis across clinical, environmental, and cleanroom samples. The presented pipeline integrates *de novo* assembly, whole-genome search, and functional profiling to detect known and novel organisms, including engineered or evolutionarily distant microbes. The presentation highlighted challenges in scaling whole-genome search and functional modeling with an emphasis on the issues surrounding current searchable databases. The presentation introduced a new, decontaminated database where incorrectly labeled sequences have been corrected or removed [14], reducing spurious classification of metagenomic datasets. Additionally, extending search to structural homology was discussed along with a case study where a structural modeling pipeline, coupled with molecular dynamics, was used to assess biotoxins and predict interactions with human ion channels.

2.1.3 Durable Naming Systems to Track Trace Genomic Information

The final presentation addressed the challenge of interpreting unclassified DNA. Using mucosal and environmental samples as case studies, the StrainSelect database [15] assigned durable, provisional naming systems to track recurring but unnamed sequences, especially in low-biomass or complex environments like cleanrooms and livestock microbiomes. This approach may resolve inconsistencies in public genome data and includes provisional identifiers for unnamed but meaningful lineages and improves interpretability in applied settings. Microbial dark matter includes biologically significant sequences that are often overlooked [15], and by assigning stable identifiers, even without formal taxonomy, researchers can retain and analyze these sequences across studies.

2.2 Breakout

The breakout session focused on the persistent challenge of Microbial Dark Matter in PP and metagenomic surveillance. Participants emphasized that while bacterial genomes are increasingly well-characterized, large gaps remain in our understanding of microeukaryotes, viruses, plasmids, and other non-bacterial components, especially in cleanroom and built environments. These gaps limit our ability to interpret metagenomic data fully, even when sequencing depth is high. The group acknowledged that current reference databases are insufficient and that even with extensive sequencing, a large proportion of reads remain unclassified. This underscores the need for new strategies to track, catalog, and interpret unknown sequences.

Discussion also highlighted the need for pragmatic, risk-informed approaches. While complete annotation of all sequences is unlikely in the near term, participants advocated for building environment-specific reference catalogs and probabilistic models to assess risk and guide decision-making. There was strong support for coordinated, multi-agency investment in deep sequencing of cleanrooms and other relevant environments to establish baselines. Participants also discussed the importance of tracking total DNA composition, not just viable organisms, and recognized that some level of uncertainty and risk will always remain. The goal is to manage that uncertainty transparently and systematically.

3. The Low Biomass Challenge

As planetary missions increasingly target environments with high astrobiological potential, the need to detect and characterize microbial life in ultra-low biomass contexts has become a central challenge for PP [7]. This session opened with a discussion about the operational and policy-driven context of international PP standards. As missions approach potentially habitable worlds like Mars, Europa, and Enceladus, the stringency of bioburden control must increase, yet current standards still rely heavily on cultivation methods, including spore-based assays such as the NSA. The absence of a biological equivalent to particle counters in cleanroom environments, particularly for non-culturable organisms, positions metagenomics as a promising but as-yet unstandardized solution. The session focused on the operationalization of metagenomics in spacecraft assembly and contamination control, where biomass is scarce, sensitivity requirements are extreme, and decisions must be made under uncertainty.

3.1 Presentations

3.1.1 Current Challenges at the Metagenomic Signal Baseline

Metagenomic analysis of ultra-low biomass has specific and unique challenges. Based on experiences with the Extreme Microbiome Project [16] and PP collaborations, ultra-low biomass was loosely defined as DNA yields of less than 1 nanogram (ng). A central challenge is the signal to noise for these samples, emphasizing the need for DNA-free reagents, high-efficiency extraction methods, and rigorous controls. Unique tools for DNA extraction, such as metapolyzyme and exopolyzyme, and amplification [multiple

displacement amplification (MDA) [17], primary template-directed amplification (PTA) [18], and PicoPLEX/ThruPlex [19]], show promise for metagenomic sequencing of low biomass samples. Of particular emphasis was the need for standards and controls such as the development of DNA-free certified reagents, enumerated whole-cell positive controls, amplification protocols using deoxyuridine triphosphate/uracil-DNA glycosylase (dUTP/UNG) chemistry to prevent lab contamination, and preventing overamplification.

3.1.2 Case Studies on the Current State of Low Biomass Genomics

The following presentations consisted of a series of case studies and investigations where low biomass metagenomic sequencing was central to the scientific objectives.

Nanopore-based metagenomics in ultra-low biomass environments, specifically targeting uncultivated archaea from deep terrestrial subsurface aquifers, faced DNA yields consistently below Oxford Nanopore's recommended input levels [20]. Using a Zymo mock community to evaluate nanopore sequencing performance across a range of DNA inputs (1000 to 1 ng), all species in the mock community were detectable down to 1 ng. High-quality MAGs and even closed genomes were recovered at inputs as low as 10 nanograms.

A case study from Antarctica's Shackleton Glacier region investigated microbial life in hyper-arid, high-elevation soils [21]. These environments, among the most Mars-analogous on Earth, yielded extremely low biomass (20% of samples failed to produce any detectable DNA). Using a combination of 16S rRNA gene amplicon sequencing, metagenomics, qPCR, ATP assays, and isotopic labeling the absence of detectable microbial life in these samples was confirmed, highlighting the importance of multi-modal validation in low-biomass studies. Echoing previous recommendations, the talk emphasized sequencing blanks, checking for cross-contamination, and avoiding over-reliance on post hoc bioinformatic filtering.

Moving to the hyper-arid Atacama Desert, a case study outlined how a combination of cell-first extraction, concentration, and sub-nanogram library preparation resulted in successful metagenomes from samples with as little as 100 picograms of DNA [22]. Highlighting planetary environmental impacts, results of nanopore sequencing experiments under Mars-like conditions demonstrated successful detection of *Bacillus subtilis* DNA at inputs as low as 2 picograms. The presentation also called for experiments on the Moon to validate and evolve current PP policies as a known abiotic control.

The final presentation explored nanopore sequencing at the picogram scale, with a focus on amplification-free detection of microbial DNA in Mars-relevant and Earth-based low-biomass environments [23]. Working in a custom-built ISO Class 5 cleanroom, successful sequencing of 10 pg of monoculture DNA and 2 pg of mixed community DNA using Oxford Nanopore's MinION sequencing platform was demonstrated. Again, the importance of amplification as a source of false positives and overrepresentation of contaminants was emphasized.

3.2 Breakout

The breakout session on low biomass metagenomics focused on identifying practical barriers and community-driven solutions for working with ultra-low biomass samples, particularly in the context of PP and spaceflight. Participants emphasized that current DNA extraction kits are not optimized for low-volume elution, often requiring 50–100 µL, which dilutes already scarce DNA. This inefficiency, combined with DNA adsorption to plastics and the risk of contamination during concentration steps, makes it difficult to recover and quantify DNA at the femtogram level. Participants also discussed the limitations of current amplification methods, which can introduce noise and overrepresent contaminants, and the need for better strategies to distinguish DNA originating from viable and non-viable cells.

A recurring theme was the need for standardized spike-in controls and mock communities to normalize workflows and benchmark performance across labs. Suggestions included using synthetic DNA, Lambda phage, or enumerated whole-cell standards, with adaptive Nanopore sequencing proposed as a way to enrich for target reads. Participants also called for the development of a shared cleanroom metagenomic database and a PP analysis working group to coordinate international efforts. The group acknowledged that metagenomics should augment, not replace, traditional spore assays and other efforts to cultivate and isolate microorganisms, and emphasized the importance of integrating metagenomic data into probabilistic risk models to support consequence-based contamination assessments. Such data also have great importance for life detection, assessing the risk of false positive after a putative positive detection of life.

4. Bioinformatics and Databases

This session showcased recent advancements in microbial data accessibility, analysis, and application across spaceflight and terrestrial research. Presenters highlighted how NASA's Open Science Data Repository (OSDR) and the Space Microbial Culture Collection (SMCC) have transformed microbial research by providing standardized workflows, rich metadata, and access to physical isolates. These resources supported PP objectives, enabled discovery in extreme environments, and facilitated cross-disciplinary collaboration. EMBL-EBI's MGnify platform demonstrated global best practices in microbiome data sharing and analysis, offering large-scale tools and datasets for biodiscovery. Additionally, strategies for reducing bioinformatic noise in low-biomass samples addressed critical challenges in accurately interpreting microbial signatures under high-contamination-risk or low-yield conditions. Together, these talks emphasized the impact of open, high-quality microbial data in advancing science, safeguarding planetary environments, and deepening our understanding of health in space contexts.

4.1 Presentations

4.1.1 NASA's Open Science Data Repository

The presentations began with an overview of the NASA Open Science Data Repository (OSDR) [24] with a focus on microbial datasets and bioinformatics workflows. OSDR supports PP goals by providing curated microbial data from spaceflight and ground

studies, along with open, reproducible pipelines for amplicon and metagenomic analyses. The talk emphasized the value of transparent, accessible tools and data to advance microbiome research in spaceflight-relevant environments.

Next, was an introduction to the Space Microbial Culture Collection (SMCC), a growing NASA repository that preserves and provides access to microbial isolates from spaceflight and spacecraft-associated environments. SMCC has a dual role: providing physical access to microbial samples and offering genomic, phenotypic, and rich metadata to researchers. SMCC was positioned as a critical asset for PP, enabling deeper understanding of microbial diversity in mission environments and helping assess contamination risks for planetary missions.

4.1.2 MGnify: EMBL-EBI's Open-Access Microbiome Platform

A showcase of MGnify followed, EMBL-EBI's open-access microbiome platform, as a model for scalable, accessible bioinformatic infrastructure (https://www.ebi.ac.uk/metagenomics). The talk discussed the breadth of MGnify's taxonomic and functional profiling capabilities, its extensive protein and genome catalogs, and the importance of reproducible pipelines for global microbiome research. The talk illustrated how similar principles and tools could be adapted for PP applications, particularly in handling complex, environmental microbial data.

4.1.3 Detection Challenges in Low Biomass Samples

Finally, a presentation addressed the unique challenges of analyzing low-biomass samples typical in spacecraft-associated environments. Strategies for reducing contamination noise and improving signal detection in sequencing data, including improved negative controls, background subtraction, and advanced filtering techniques were discussed. The presentation underscored the need for specialized workflows in PP to ensure high-confidence microbial detection in low-biomass contexts.

4.2 Breakout

This session focused on defining the foundational needs for a PP metagenomic genome database and associated bioinformatics workflows tailored to the unique challenges of low-biomass, spaceflight-related microbiome samples. Attendees evaluated database priorities, discussed metadata and data integration, and outlined ideal features for PP workflows.

The group emphasized the importance of leveraging existing resources such as the SMCC, OSDR, and international platforms like MGnify. Discussions highlighted the need for a curated, space-specific genome database that includes kitome profiles, resistance traits, and metadata relevant to spacecraft environments. In parallel, the session explored the development of standardized, reproducible bioinformatics workflows using platforms like Nextflow, with containerized tools and clear input/output expectations. The goal is to support both scientific discovery and operational decision-making in PP.

5. Technology Needs

The integration of metagenomics into PP hinges not only on scientific insight but also on the availability of robust, scalable, and contamination-aware technologies. This session was designed to surface the critical technology gaps that limit the deployment of metagenomic tools in spaceflight and cleanroom environments. Presentations showcased advances in sequencing platforms, ultra-clean reagents, and low-input library preparation, while also revealing persistent challenges in quantification, automation, and standardization. The session's premise was that without targeted investment in enabling technologies, especially those tailored for low-biomass and important but low abundance taxa, metagenomics cannot fulfill its potential as a tool in PP risk assessment.

5.1 Presentations

5.1.1 Oxford Nanopore Sequencing Technology for Metagenomics

Presentations began with an overview of Oxford Nanopore's latest advances in sequencing technology, emphasizing its unique suitability for spaceflight applications. The platform's core advantages were discussed to include real-time sequencing, short and long-read capability, and tailorable sample preparation, which make it ideal for rapid, *in situ* microbial monitoring [25]. The technology's ability to detect base modifications and sequence native DNA without PCR amplification offers a powerful tool for capturing epigenetic and functional information but typically requires higher DNA input compared to other platforms. Improvements in basecalling accuracy demonstrated how long-read metagenomics can yield more complete and contiguous MAGs compared to short-read platforms. Resolution of long homopolymer regions remains a challenge; however, nanopore-only bacterial genomes assemblies are approaching the quality of short-read assemblies [26]. Case studies highlighted successful sequencing from picogram-level DNA inputs and cleanroom samples and airborne viral surveillance [27]. These developments support the broader goal of integrating metagenomics into safety-critical decision-making by enabling metagenomics, even in remote or resource-limited environments like spacecraft.

5.1.2 Open, Modular Metagenomics Pipeline at the Joint Genome Institute

Next the Joint Genome Institute (JGI) provided a comprehensive overview of JGI's capabilities in high-throughput metagenomic sequencing and its relevance to low-biomass environments. Highlights included JGI's dual-path library preparation workflows one for low input and one for ultra-low input samples, emphasis their experience with thousands of metagenomic samples annually, and JGI's rigorous contamination tracking system, which catalogs known contaminants across projects and could serve as a foundation for a broader community reference database. Case studies touched on an investigation involving ancient ice core samples, where JGI successfully recovered meaningful viral metagenomic data from undetectable DNA inputs [28] and using primary template-directed amplification (PTA) to get near complete genomes from single cells [29]. From a technology standpoint, JGI's work exemplifies scalable, quality-controlled metagenomic processing pipelines that can be adapted for spaceflight. These innovations are crucial for extending metagenomic capabilities to extreme

environments, where biomass is scarce and functional insights are essential for risk assessment and microbial monitoring.

5.1.3 Emerging Commercial Solutions for Planetary Protection Technology Gaps

The final two presentations came from commercial entities designing reagents and workflows to support the technical challenges associated with metagenomics for PP. QIAGEN focused on the critical issue of contamination control and DNA recovery in low-biomass metagenomic workflows, highlighting their Ultra Clean Production (UCP) line of reagents and kits, which are specifically engineered to minimize background DNA contamination. QIAGEN evaluates and certifies its kits through rigorous quality control, including mock extractions and quantification of residual DNA. A case study from a European food safety project highlighted a custom low-biomass extraction protocol that enabled successful metagenomic sequencing from stainless steel surfaces in food processing environments [30].

Next, Zymo Research presented a suite of technologies aimed at improving metagenomic analysis in low-input and low-biomass contexts. Zymo has developed products for end-to-end workflows from sample collection and preservation to DNA extraction and sequencing library creation. Spike-in standards for assessing extraction efficiency and detection limits were part of a workflow that performed shotgun metagenomics on as little as 100 femtograms of DNA. This analysis enabled cross-kingdom profiling (bacteria, viruses, and eukaryotes) and strain-level resolution, enabled by their optimized library prep and sequencing protocols. The company offered to collaborate on pilot studies using Zymo's kits and sequencing services.

Both talks underscored the foundational importance of clean, consistent, and contamination-aware workflows for reliable metagenomic analysis in ultra-low biomass environments. From Zymo's integrated approach to QIAGEN's emphasis on whole-system contamination control, each highlighted the critical need for validated reagents and standardized protocols. The lack of community-wide benchmarks for acceptable background DNA levels emerged as a shared concern, with both speakers calling for collaborative efforts to define and enforce such standards. These insights are directly applicable to spaceflight and PP, where sample scarcity and contamination risk demand exceptionally robust and reproducible methods.

5.2 Breakout

The Technology Needs breakout session focused on identifying critical gaps and opportunities in the development of tools and workflows for metagenomics in PP and spaceflight. Participants emphasized the need for improved DNA quantification methods at ultra-low biomass levels, as current tools like Qubit and Nanodrop™ (Thermo Fisher Scientific) lack the sensitivity and reliability required for picogram-scale detection. Digital PCR and synthetic spike-in standards were discussed as promising alternatives, though cost and complexity remain barriers. The group also highlighted the importance of developing standardized, DNA-free reagents and extraction kits tailored for low-volume, high-efficiency recovery, especially for spacecraft-associated samples.

Automation emerged as a major theme, with participants noting the need for fully integrated, hands-off systems capable of sample processing, DNA extraction, and sequencing in spaceflight environments. While current robotic systems can perform specific tasks well, their lack of adaptability and high maintenance requirements limit their utility. Technologies like Nanopore sequencing were praised for their portability and potential for *in situ* use, but concerns remain about library preparation complexity, pore longevity, and contamination from amplification products (when such workflows are used). The group also discussed the need for better sampling strategies, including high-efficiency swabs, air and liquid concentration devices, and proxy surfaces like fallout coupons to improve biomass recovery.

Finally, the session addressed the importance of developing PP-specific standards and controls, including mock communities, spike-ins, and cleanroom metagenomic databases. Participants stressed the need for international collaboration and coordination with industry to ensure that new technologies are accessible, scalable, and aligned with evolving mission requirements. The conversation also touched on the need to balance scientific rigor with practical constraints, especially for commercial and international partners.

6. Human Health and the Built Environment

This session examined metagenomics potential for both astronaut health and PP, highlighting its role in proactive microbial monitoring in closed environments. NASA's work on *in situ* sequencing aboard the International Space Station (ISS) demonstrated a major shift from culture-based methods to real-time, autonomous microbial risk assessment, with applications extending to Gateway and future missions. Presentations from the U.S. Centers for Disease Control and Prevention (CDC) and academic partners emphasized lessons from public health, including antimicrobial resistance surveillance, community-level profiling, and predictive modeling, well aligned with spaceflight needs. Case studies underscored the persistence of resistant organisms in built environments and the importance of viability assays, quantification standards, and functional inference for actionable insights. Together, these talks framed metagenomics as a critical tool for integrated health and contamination control strategies, requiring shared standards, robust workflows, and interdisciplinary collaboration to translate data into operational decision-making.

6.1 Presentations

6.1.1 Microbial Monitoring for Human Spaceflight

The presentations began with a talk from Johnson Space Center on NASA's progress in developing in situ microbial monitoring capabilities for spaceflight, highlighting the transition from traditional culture-based methods to real-time, culture-independent sequencing using Oxford Nanopore's MinION [31]. Over the past eight years, the team has developed and validated a targeted 16S rRNA gene amplicon "swab-to-sequencer" workflow that enables astronauts to collect, process, and sequence microbial samples directly aboard the ISS. This method has been successfully applied to surfaces, potable water, and wastewater, revealing a broader and more nuanced microbial landscape

than culture methods alone. Notably, the system has been baselined for use on Gateway, marking a major milestone in operationalizing metagenomics for long-duration missions.

From a technology and risk-based decision-making perspective, the work addresses a critical need: enabling autonomous, real-time microbial risk assessment in environments where sample return is not feasible. The sequencing platform can provide richer taxonomic resolution and functional insights, including detection of antibiotic resistance genes, which are essential for both crew health and PP. Current efforts include adapting the system for shotgun metagenomic sequencing of low-biomass samples. The integration of this technology into NASA's operational toolkit represents a significant step toward performance-based microbial monitoring and opportunity for cross-disciplinary applications, including food safety, environmental surveillance, and forward contamination control.

The next talk offered a public health perspective on the use of sequencing and metagenomics to combat antimicrobial resistance (AMR) and improve infection prevention. Drawing from CDC's experience with whole-genome sequencing in outbreak investigations, the talk illustrated how genomic tools can identify emerging resistance mechanisms, trace transmission pathways, and inform targeted interventions. Emphasis was placed on the importance of understanding microbial communities, not just individual pathogens, as reservoirs of resistance and infection risk. The concept of "pathogen reduction" (i.e., reduction of colonizing pathogens to reduce infection and transmission risk) was proposed as a more realistic and actionable goal than full decolonization [32].

These insights are highly relevant to spaceflight, where immune dysregulation and microbiome shifts increase the risk of opportunistic infections. The emphasis on community-level surveillance, functional profiling, and risk stratification aligns with the goals of metagenomics in space: to move beyond detection toward prediction and prevention. The CDC's framework for integrating sequencing into public health decision-making through diagnostics, surveillance, and intervention design offers a valuable model for NASA and other space agencies [33], [34]. It also underscores the need for shared standards, reference databases, and cross-sector collaboration to translate metagenomic data into actionable health and safety outcomes.

Two compelling case studies followed that underscore the dual role of the human microbiome and the built environment as reservoirs for antimicrobial resistance. In the first, it was shown even short courses of antibiotics can cause long-term disruptions in the gut microbiome, increasing the burden and diversity of resistance genes [35]. In the second, a longitudinal study of a hospital ICU showed sink drains were found to harbor persistent, transmissible, and clinically relevant resistant organisms [36]. These findings reinforce the need for proactive microbial monitoring in built environments like spacecraft, where limited crew turnover and stringent cleaning protocols can foster microbial persistence.

Finally, a discussion highlighted the development of a metagenomics-based workflow for environmental surveillance in healthcare settings, with clear parallels to spacecraft environments. The effort focused on familiar key challenges: low biomass, live/dead discrimination, quantification, and sequencing resource prediction. A robust, adaptable pipeline was developed that used propidium monoazide (PMA) to distinguish viable cells, internal standards for absolute quantification, and machine learning to estimate sequencing needs [37]. Emphasis on viability, quantification, and functional inference aligns with the needs of spaceflight microbial monitoring, where understanding not just who is present but what they can do is critical. A case study on chloroquine degradation in aquarium systems [38] further illustrates how metagenomics can inform targeted interventions in complex, closed-loop environments, analogous to life support systems in space.

6.2 Breakout

The breakout session on Human Health and the Built Environment explored the complex intersection of astronaut health, microbial monitoring, and PP. A central theme was the challenge of defining a “healthy” human microbiome in space, given the variability between individuals and the dynamic changes that occur during long-duration missions. Participants emphasized the need for personalized baseline microbiome profiles for each crew member, with longitudinal monitoring to detect deviations that might signal health risks or microbial shifts relevant to PP. The group also discussed the convergence of microbiomes in closed environments and the difficulty of distinguishing between natural adaptation and potential infection.

Another major focus was the potential of metagenomic data for both crew health and PP. While crew health will always take priority, the group agreed that shared tools, standards, and databases could serve both domains. Concerns were raised about antimicrobial resistance, the persistence of extremophiles, and the potential for microbial transfer between humans and spacecraft surfaces. Participants also highlighted the need for predictive models, functional gene analysis, and integration of metabolomics or biomarker-based diagnostics to complement DNA-based methods. The session concluded with a call for interdisciplinary collaboration, shared mock communities, and a roadmap that balances operational constraints with scientific rigor.

7. Recommendations

Session 1: Safety-Critical Decision Making and Foundational Frameworks

- **Adopt a tiered testing framework**: Use ATP assays for rapid screening, NSA for standardized spore abundance assessments, digital PCR for microbial load quantification, and metagenomics for functional and taxonomic insight.
- **Develop shared microbial reference standards**: Create a curated genome catalog of cleanroom isolates across NASA, ESA, and JAXA and continue culture-based assessments to expand current knowledge and isolate collections.

- **Standardize quality control protocols**: Define minimum QC requirements (e.g., internal standards, extraction controls) for metagenomic workflows.
- **Incorporate viability assessments**: Evaluate PMA, nucleases, or RNA-based methods to distinguish viable from relic DNA.
- **Use metagenomics to inform—not replace—quantitative assays**: Identify targets for digital PCR or cultivation, especially rare or stress-resistant organisms.
- **Build probabilistic risk models**: Integrate metagenomic data into mission-specific models that account for microbial survival and environmental conditions.
- **Coordinate international pilot studies**: Align sampling and analysis protocols across agencies to validate methods and build a shared roadmap.

Session 2: Microbial Dark Matter and Functional Inference

- **Develop environment-specific reference catalogs**: Focus on cleanrooms and other low-biomass environments to improve classification and risk assessment.
- **Invest in deep, longitudinal sequencing**: Establish microbial baselines and track changes over time.
- **Adopt provisional labeling systems**: Enable consistent tracking of unclassified but reproducible sequences.
- **Expand focus beyond bacteria**: Include archaea, microeukaryotes, viruses, and plasmids in reference databases.
- **Incorporate probabilistic models for unannotated sequences**: Assess functional and ecological relevance despite incomplete annotation.
- **Coordinate multi-agency investment**: Avoid duplication and accelerate progress through shared funding and tasking.
- **Acknowledge and manage residual risk**: Accept that decisions must be made under uncertainty due to incomplete annotation.

Session 3: Ultra-Low Biomass Metagenomics

- **Develop DNA extraction kits optimized for ultra-low biomass**: Include low-volume elution (≤10 µL) and DNA-free certification.
- **Validate standardized spike-in controls**: Use synthetic DNA, Lambda phage, or whole-cell standards for normalization and contamination tracking.
- **Establish shared mock community standards**: Include bacteria, fungi, and archaea for benchmarking workflows.
- **Create a cleanroom metagenomic database**: Include standardized metadata and contamination profiles.
- **Form a PP metagenomics working group (AWG)**: Define and develop protocols for extraction of nucleic acids, library preparation and sequencing, and bioinformatics.

- **Explore adaptive nanopore sequencing**: Improve sequencing efficiency at picogram input levels using carrier DNA or pore management.

Session 4: Bioinformatics and Database Infrastructure

- **Prioritize curated, space-specific microbial genomes**: Include cleanroom organisms, kitome profiles, and SMCC contributions.
- **Align metadata with existing standards**: Extend MIxS, EnvO, MIGS, MISAG/MIMAG to include space-relevant fields (e.g., radiation exposure) in accorfance with the National Microbiome Data Collaborative (NMDC).
- **Build a federated or linked-access database**: Integrate raw reads, assemblies, amplicon profiles, and qPCR/ddPCR data.
- **Develop containerized, reproducible pipelines**: Use platforms like Nextflow with fixed tool versions and parameters.
- **Define clear input/output expectations**: Include quality control metrics and contaminant detection.
- **Incorporate decision-support features**: Link outputs to go/no-go thresholds and the PP genome database.
- **Encourage community coordination**: Use working groups and shared frameworks to ensure interoperability and sustainability.

Session 5: Technology Development and Automation

- **Develop ultra-sensitive DNA quantification tools**: Use digital PCR or synthetic spike-ins for picogram-level detection.
- **Collaborate with industry on DNA-free, low-input kits**: Tailor reagents for PP applications.
- **Invest in automation for in situ workflows**: Focus on robustness, miniaturization, and contamination control.
- **Improve sampling technologies**: Use high-efficiency swabs, air/liquid concentrators, and proxy surfaces (e.g., fallout coupons).
- **Standardize mock communities and spike-in controls**: Support benchmarking and cross-lab comparisons.
- **Explore pore management strategies**: Enhance Nanopore sequencing efficiency at low input levels.
- **Promote international collaboration and open science**: Ensure data sharing and long-term sustainability.

Session 6: Human Health and Built Environment

- **Define and monitor personalized baseline microbiomes**: Use longitudinal sampling to detect health-relevant changes.
- **Develop shared mock communities and standards**: Support both crew health and PP assessments.

- **Establish a curated database of medically relevant microbes**: Include functional gene markers relevant to health and contamination risk.
- **Prioritize predictive models**: Integrate metagenomic, functional, and environmental data to flag deviations from expected baselines.
- **Explore metabolomics and biomarkers**: Use complementary tools for early detection of health issues or microbial shifts.
- **Design spacecraft architecture with microbial risk in mind**: Consider hotspots and airflow patterns in cleaning protocols.
- **Coordinate across disciplines**: Align sampling strategies, data interpretation, and mission planning.

Implementation Roadmap

The workshop generated a robust set of recommendations across technical, operational, and policy domains. From these recommendations, a roadmap was developed that reflects a synthesis of recurring themes and shared priorities that emerged across sessions. It is a structured path forward for how to begin translating metagenomic methods into operational PP frameworks. The topics, which range from international standard operating procedures and computational tool development to cross-disciplinary collaboration and probabilistic risk modeling, represent the foundational infrastructure needed to support future implementation. Each topic builds on the workshop's core insight: that metagenomics, while not yet ready to replace traditional assays, is essential for enabling risk-informed, performance-based decision-making in spaceflight. The roadmap is intended to guide coordinated, international efforts to validate, standardize, and integrate metagenomic tools into PP and related mission areas.

Topic	Lead Organizations	2025		2026	
International SOP for Metagenomics					
Develop a standard community	NASA/ESA/UKSA/JAXA/CSA				
Collaborative testing of cleanrooms	NASA/ESA				
Test spiking DNA, establish controls, and QC	NASA/ESA/UKSA/JAXA/CAS				
Establish an analytical working group for SOP	NASA				
Publication of a standard approach	NASA/ESA/UKSA/JAXA/CSA				
Develop Computational Tools for Metagenomics					
Sequence cleanroom isolates	NASA/ESA				
Integrate/extend/develop genomic databases	NASA/ESA				
Adopt emerging techniques and methods from adjacent fields	NASA/ESA				
Establish a planetary protection analysis working group	NASA				
Develop International and Cross Discipline Collaboration					
COSPAR conferences	NASA/ESA/UKSA/JAXA/CSA				
Follow up workshop to assess progress	NASA				
White paper on low biomass challenges and best practices	Gov/Academia/Industry working group				
Consistent engagement with commercial partners	NASA/ESA/UKSA/JAXA/CSA				
Metagenomics for Probablistic Risk Assessments					
Expand functional annotation of metagenomic datasets	NASA/ESA				
Method development for experimental procedures	NASA/UKSA				
Bayesian approaches for integrated risk assessments	NASA/ESA/UKSA				

1. NASA: National Aeronautics and Space Administration (USA)
2. ESA: European Space Agency
3. UKSA: United Kingdom Space Agency
4. JAXA: Japan Aerospace eXploration Agency
5. CSA: Canadian Space Agency

Figure 3. Implementation roadmap established by the Workshop participants towards closing the identified gaps in Planetary Protection research, policy, and operations.

Acknowledgements

The authors thank Alison Laufer Halpin for providing subject matter expertise on the use of sequencing and metagenomics to combat antimicrobial resistance and improve infection prevention, including her workshop presentation, “The Role of Next Generation Sequencing for Healthcare-Associated Infection and Public Health Applications.” Microsoft Copilot was used to take transcripts of the workshop presentations and summarize the content into paragraph summaries. Those summaries were used by the authors to generate the report text, which was reviewed for accuracy by all authors, to include those whose presentation was summarized by AI.

References

[1] National Aeronautics and Space Administration, “Moon to Mars Architecture Definition Document,” NASA Exploration Systems Development Mission Directorate, Washington, D.C., Technical Report NASA/TP–20240015571, Apr. 2025. [Online]. Available: https://ntrs.nasa.gov/citations/20240015571

[2] S. J. Green *et al.*, “Metagenomic Methods for Addressing NASA’s Planetary Protection Policy Requirements on Future Missions: A Workshop Report,” *Astrobiology*, vol. 23, no. 8, pp. 897–907, Aug. 2023, doi: 10.1089/ast.2022.0044.

[3] A. L. Smith and R. C. Hendrickson, “Protecting ocean worlds: Europa Clipper planetary protection inputs to a probabilistic risk-based approach,” *International Journal of Astrobiology*, vol. 21, no. 6, pp. 470–483, Dec. 2022, doi: 10.1017/S1473550422000192.

[4] K. J. McCoy, M. DiNicola, C. Everline, H. Burgoyne, K. Reinholtz, and B. Clement, “Europa Clipper planetary protection probabilistic risk assessment summary,” *Planetary and Space Science*, vol. 196, p. 105139, Feb. 2021, doi: 10.1016/j.pss.2020.105139.

[5] K. L. Beck *et al.*, “Monitoring the microbiome for food safety and quality using deep shotgun sequencing,” *npj Sci Food*, vol. 5, no. 1, p. 3, Feb. 2021, doi: 10.1038/s41538-020-00083-y.
[6] A. W. Pightling, J. B. Pettengill, Y. Luo, J. D. Baugher, H. Rand, and E. Strain, “Interpreting Whole-Genome Sequence Analyses of Foodborne Bacteria for Regulatory Applications and Outbreak Investigations,” *Front. Microbiol.*, vol. 9, Jul. 2018, doi: 10.3389/fmicb.2018.01482.
[7] A. C. Simpson *et al.*, “Analysis of Microbiomes from Ultra-Low Biomass Surfaces Using Novel Surface Sampling and Nanopore Sequencing,” *J Biomol Tech*, vol. 34, no. 3, p. 3fc1f5fe.bac4a5b3, doi: 10.7171/3fc1f5fe.bac4a5b3.
[8] T. Zamkovaya, J. S. Foster, V. de Crécy-Lagard, and A. Conesa, “A network approach to elucidate and prioritize microbial dark matter in microbial communities,” *ISME J*, vol. 15, no. 1, Art. no. 1, Jan. 2021, doi: 10.1038/s41396-020-00777-x.
[9] Y. Kawakatsu *et al.*, “Preliminary design of Martian Moons eXploration (MMX),” *Acta Astronautica*, vol. 202, pp. 715–728, Jan. 2023, doi: 10.1016/j.actaastro.2022.09.009.
[10] S. Ghatak, Z. A. King, A. Sastry, and B. O. Palsson, “The y-ome defines the 35% of Escherichia coli genes that lack experimental evidence of function,” *Nucleic Acids Res*, vol. 47, no. 5, pp. 2446–2454, Mar. 2019, doi: 10.1093/nar/gkz030.
[11] A. V. Sastry *et al.*, “The Escherichia coli transcriptome mostly consists of independently regulated modules,” *Nat Commun*, vol. 10, Dec. 2019, doi: 10.1038/s41467-019-13483-w.
[12] J. Shin, K. Rychel, and B. O. Palsson, “Systems biology of competency in Vibrio natriegens is revealed by applying novel data analytics to the transcriptome,” *Cell Rep*, vol. 42, no. 6, p. 112619, Jun. 2023, doi: 10.1016/j.celrep.2023.112619.
[13] N. D. Menon *et al.*, “Independent component analysis reveals 49 independently modulated gene sets within the global transcriptional regulatory architecture of multidrug-resistant Acinetobacter baumannii,” *mSystems*, vol. 9, no. 2, p. e0060623, Feb. 2024, doi: 10.1128/msystems.00606-23.
[14] J. M. Martí *et al.*, “Addressing the dynamic nature of reference data: a new nucleotide database for robust metagenomic classification,” *mSystems*, vol. 10, no. 4, pp. e01239-24, Mar. 2025, doi: 10.1128/msystems.01239-24.
[15] T. Z. DeSantis *et al.*, “StrainSelect: A novel microbiome reference database that disambiguates all bacterial strains, genome assemblies and extant cultures worldwide,” *Heliyon*, vol. 9, no. 2, p. e13314, Feb. 2023, doi: 10.1016/j.heliyon.2023.e13314.
[16] S. Tighe *et al.*, “Genomic Methods and Microbiological Technologies for Profiling Novel and Extreme Environments for the Extreme Microbiome Project (XMP),” *J Biomol Tech*, vol. 28, no. 1, pp. 31–39, Apr. 2017, doi: 10.7171/jbt.17-2801-004.
[17] R. S. Lasken, “Genomic DNA amplification by the multiple displacement amplification (MDA) method,” *Biochem Soc Trans*, vol. 37, no. 2, pp. 450–453, Mar. 2009, doi: 10.1042/BST0370450.
[18] V. Gonzalez-Pena *et al.*, “Accurate genomic variant detection in single cells with primary template-directed amplification,” *Proceedings of the National Academy of*

Sciences, vol. 118, no. 24, p. e2024176118, Jun. 2021, doi: 10.1073/pnas.2024176118.
[19] E. Kamberov, T. Tesmer, M. Mastronardi, and J. Langmore, “Two Methods for High-Throughput NGS Template Preparation for Small and Degraded Clinical Samples Without Automation,” *J Biomol Tech*, vol. 23, no. Suppl, p. S36, 2012.
[20] S. A. Simon, K. Schmidt, L. Griesdorn, A. R. Soares, T. L. V. Bornemann, and A. J. Probst, “Dancing the Nanopore limbo – Nanopore metagenomics from small DNA quantities for bacterial genome reconstruction,” *BMC Genomics*, vol. 24, no. 1, p. 727, Dec. 2023, doi: 10.1186/s12864-023-09853-w.
[21] N. B. Dragone *et al.*, “Elevational Constraints on the Composition and Genomic Attributes of Microbial Communities in Antarctic Soils,” *mSystems*, vol. 7, no. 1, pp. e01330-21, Jan. 2022, doi: 10.1128/msystems.01330-21.
[22] A. Mojarro *et al.*, “Nucleic Acid Extraction and Sequencing from Low-Biomass Synthetic Mars Analog Soils for In Situ Life Detection,” *Astrobiology*, vol. 19, no. 9, pp. 1139–1152, Sep. 2019, doi: 10.1089/ast.2018.1929.
[23] J. Basapathi Raghavendra, M.-P. Zorzano, D. Kumaresan, and J. Martin-Torres, “DNA sequencing at the picogram level to investigate life on Mars and Earth,” *Sci Rep*, vol. 13, no. 1, p. 15277, Sep. 2023, doi: 10.1038/s41598-023-42170-6.
[24] S. G. Gebre, R. T. Scott, A. M. Saravia-Butler, D. K. Lopez, L. M. Sanders, and S. V. Costes, “NASA open science data repository: open science for life in space,” *Nucleic Acids Res*, vol. 53, no. D1, pp. D1697–D1710, Jan. 2025, doi: 10.1093/nar/gkae1116.
[25] “Nanopore long-read-only metagenomics enables complete and high-quality genome reconstruction from mock and complex metagenomes | Microbiome.” Accessed: Aug. 29, 2025. [Online]. Available: https://link.springer.com/article/10.1186/s40168-022-01415-8
[26] D. Nagy *et al.*, “Nanopore long-read-only genome assembly of clinical Enterobacterales isolates is complete and accurate,” *Microbial Genomics*, vol. 12, no. 2, p. 001631, 2026, doi: 10.1099/mgen.0.001631.
[27] N. R. Minor *et al.*, “Metagenomic sequencing detects human respiratory and enteric viruses in air samples collected from congregate settings,” *Sci Rep*, vol. 13, no. 1, p. 21398, Dec. 2023, doi: 10.1038/s41598-023-48352-6.
[28] “Glacier ice archives nearly 15,000-year-old microbes and phages | Microbiome.” Accessed: Aug. 29, 2025. [Online]. Available: https://link.springer.com/article/10.1186/s40168-021-01106-w
[29] R. M. Bowers *et al.*, “scMicrobe PTA: near complete genomes from single bacterial cells,” *ISME Commun*, vol. 4, no. 1, p. ycae085, Jan. 2024, doi: 10.1093/ismeco/ycae085.
[30] C. Barcenilla *et al.*, “Improved sampling and DNA extraction procedures for microbiome analysis in food-processing environments,” *Nat Protoc*, vol. 19, no. 5, pp. 1291–1310, May 2024, doi: 10.1038/s41596-023-00949-x.
[31] S. Stahl-Rommel *et al.*, “Real-Time Culture-Independent Microbial Profiling Onboard the International Space Station Using Nanopore Sequencing,” *Genes*, vol. 12, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/genes12010106.
[32] M. R. Mangalea, A. L. Halpin, M. Haile, C. A. Elkins, and L. C. McDonald, “Decolonization and Pathogen Reduction Approaches to Prevent Antimicrobial

Resistance and Healthcare-Associated Infections," *Emerg Infect Dis*, vol. 30, no. 6, pp. 1069–1076, Jun. 2024, doi: 10.3201/eid3006.231338.
[33] A. L. Halpin *et al.*, "A framework towards implementation of sequencing for antimicrobial-resistant and other health-care-associated pathogens," *The Lancet Infectious Diseases*, vol. 25, no. 4, pp. e235–e244, Apr. 2025, doi: 10.1016/S1473-3099(24)00729-1.
[34] A. L. Halpin, L. C. McDonald, and C. A. Elkins, "Framing Bacterial Genomics for Public Health (Care)," *Journal of Clinical Microbiology*, vol. 59, no. 12, p. 10.1128/jcm.00135-21, Nov. 2021, doi: 10.1128/jcm.00135-21.
[35] W. E. Anthony *et al.*, "Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults," *Cell Reports*, vol. 39, no. 2, Apr. 2022, doi: 10.1016/j.celrep.2022.110649.
[36] "Spatiotemporal dynamics of multidrug resistant bacteria on intensive care unit surfaces | Nature Communications." Accessed: Aug. 29, 2025. [Online]. Available: https://www.nature.com/articles/s41467-019-12563-1
[37] J. Shen *et al.*, "An improved workflow for accurate and robust healthcare environmental surveillance using metagenomics," *Microbiome*, vol. 10, no. 1, p. 206, Dec. 2022, doi: 10.1186/s40168-022-01412-x.
[38] J. Hu *et al.*, "Towards understanding microbial degradation of chloroquine in large saltwater systems," *Science of The Total Environment*, vol. 807, p. 150532, Feb. 2022, doi: 10.1016/j.scitotenv.2021.150532.