



**MICRONBRANE
MEDICAL**

White Paper

Reducing Contamination, Complexity, and Cost in Metagenomic Sequencing



Introduction

Metagenomic Next-Generation Sequencing (mNGS) is a transformative technology enabling comprehensive microbial profiling across diverse applications, including infectious disease research, clinical diagnostics, biopharmaceutical development, and industrial microbiology. By simultaneously sequencing all nucleic acids in a sample, mNGS provides an unbiased approach to detecting bacteria, viruses, fungi, and parasites.

Despite its advantages, several challenges continue to hinder widespread adoption. Host nucleic acids and low microbial biomass interfere with sample processing, reducing pathogen detection sensitivity. Contaminants from the environment, reagents, and workflows introduce background noise that complicates accurate identification. Data analysis is further constrained by computational demands and difficulty separating true signals from artifacts. Sequencing costs also remain a barrier. Collectively, these issues undermine research integrity, compromise diagnostic accuracy, and increase operational costs.^{1,2}

A Complete Solution to mNGS Workflow Challenges

In this white paper, we present innovations incorporated into the Pathogen Real-Time Identification by Sequencing (PaRTI-Seq™) assay, a commercially available, complete metagenomic sequencing workflow. Developed by Micronbrane Medical, PaRTI-Seq is used in research and clinical laboratories, providing institutions with the ability to perform high-quality, mNGS in-house without relying on external service providers.

Unlike protocols adapted from genomic or transcriptomic applications, PaRTI-Seq is purpose-built to address the technical demands of microbial profiling across complex mixtures of bacterial, viral, fungal, and parasitic nucleic acids.

The assay incorporates patented zwitterionic host depletion, mNGS-grade reagents, ultra-low input library preparation, and a curated bioinformatics pipeline. Together, these components reduce contamination, workflow complexity, and cost while improving taxonomic resolution, consistency, and data quality across diverse sample types.

PaRTI-Seq supports both manual and automated workflows, making it suitable for implementation in labs of any size—from small academic research groups to high-throughput clinical environments.

PaRTI-Seq Workflow

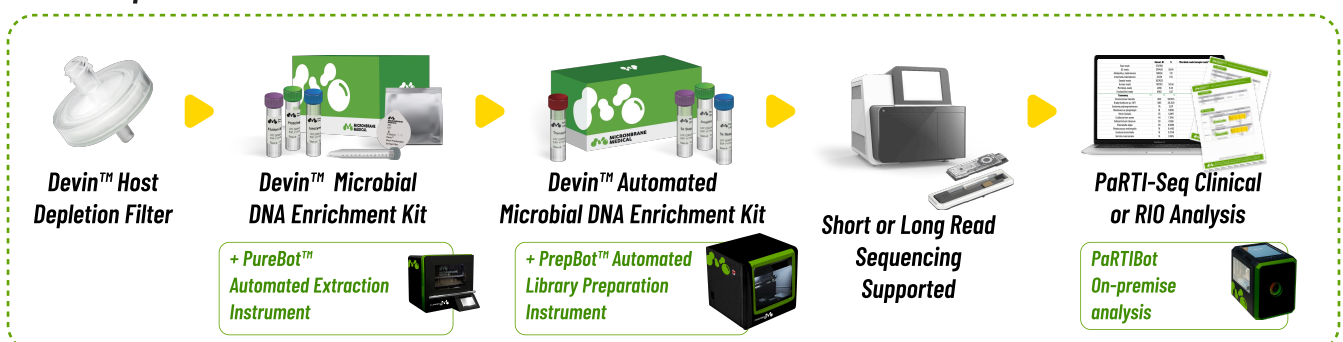


Figure 1. Pathogen Real-Time Identification by Sequencing (PaRTI-Seq) assay components and supporting automation instrumentation

Host Nucleic Acid Predominance as a Quantitative Impediment to Microbial Detection

Microbial genetic material constitutes only a tiny fraction of the total nucleic acids in most biological samples, a phenomenon that significantly limits detection sensitivity. In clinical samples, host-derived nucleic acids account for more than 99% of total genetic content, creating substantial barriers to microbial detection in mNGS.³ The resulting imbalance reduces both taxonomic resolution and functional characterization, often leading to underestimation of species.⁴

Multiple approaches have been developed to address host interference, each employing distinct mechanisms with varying effectiveness dependent on sample type and organisms.

- 1. Physical and Size-based Separation Methods** are pre-extraction approaches that utilize physical differences between host and microbial cells. These approaches offer minimal processing complexity with rapid turnaround (<15 minutes) but demonstrate limited effectiveness with microbial read percentages, well below 5%. Size-exclusion filters provide fast processing with low host depletion efficiency and microbial read percentages lower than 5%.⁵
- 2. Chemical and Enzymatic Differential Lysis** approaches exploit differences in host and microbial cell susceptibility, with effectiveness varying by sample type and protocol.⁶
 - HostZERO, (Zymo Research), claims to achieve ≥90% host DNA depletion through selective lysis in 30 minutes. However, independent studies report a 45.5% reduction in host DNA in respiratory samples.
 - MolYsis (Molzym GmbH & Co. KG) utilizes selective lysis and host DNA degradation to recover microbial DNA, achieving up to 69.6% host DNA reduction in sputum samples; however, it has been associated with high bias in microbial community representation.
 - Other types include Benzonase (nuclease-based), which achieves 6.3–50% host depletion with a 30-minute incubation at 37°C; lyPMA (osmotic lysis), which has minimal effectiveness despite requiring 40–100 minutes processing time.
- 3. Bead-based Selection Methods** use paramagnetic beads that bind differentially to nucleic acids. These methods achieve 65–90% host depletion but may alter microbial integrity during processing. Bead-based selection requires moderate processing times (90–120 minutes), adds some complexity to the workflow, and can introduce potential taxonomic bias through differential binding properties.⁷
- 4. Hybridization-based Technologies** employ biotinylated probes to selectively capture target sequences through complementary binding, achieves 65–90% host depletion but requires 90–120 minutes of processing time with high workflow complexity.⁸ Hybridization methods may alter microbial integrity and introduce potential taxonomic bias through sequence-specific selection, remain vulnerable to off-target effects, plus require significant a priori knowledge of target sequences.
- 5. DNA Modification-based Methods** that exploit methylation differences between host and microbial DNA are effective but complex approaches. Methyl-dependent restriction enzyme (MDRE) methodologies achieve high host depletion (94–96%) and increases in microbial read abundance.⁹ However, these techniques require extremely lengthy processing times (16 hours) and are highly complex. Related approaches include CpG-methylation-based techniques achieving 50-fold host depletion with 8–11.5 fold increases in microbial reads in just 30 minutes.¹⁰ CRISPR-Cas technologies represent another promising strategy, using programmable nucleases for sequence-specific cleavage of host DNA.¹¹

Technical Implications of Current Host Depletion Strategies

1. Impact on Microbial Populations

Most depletion strategies unintentionally remove microbial nucleic acids alongside host DNA, disproportionately affecting gram-negative bacteria due to their fragile cell walls. This results in taxonomic bias, with differential losses across microbial communities, impacting the accuracy of downstream analysis.¹²

2. Variable Efficiency Across Sample Types

No single approach performs optimally across all biological matrices. Protocols optimized for one sample type (e.g., blood) often exhibit reduced efficacy in others (e.g., cerebrospinal fluid, respiratory specimens), necessitating customized workflows for each study.⁵

3. High Library Preparation Failure Rates

Certain depletion techniques are associated with increased failure rates during library preparation, particularly when applied to low-biomass samples or preserved specimens. Inefficiencies in host nucleic acid removal can lead to excessive host-derived reads, rendering microbial read detection infeasible.⁵

4. Lengthy Processing Times

Many depletion workflows introduce significant time constraints, requiring 30 minutes to 16 hours for completion. This limits their suitability for high-throughput sequencing applications or time-sensitive clinical and research settings.⁷

5. Complex Workflows with High Technical Demands

Hybrid capture and enzymatic depletion require specialized expertise to optimize conditions, increasing the likelihood of procedural errors. Variability in reagent performance further complicates reproducibility, making automation difficult in certain laboratory environments.⁵

6. Dependence on A Priori Knowledge

Hybridization-based approaches necessitate detailed prior knowledge of host sequences, restricting their applicability to well-characterized genomes. This limitation impedes the detection of unknown pathogens or novel microbial species.⁸

7. Incompatibility with Preserved Specimens

Several depletion strategies exhibit reduced performance with frozen or chemically preserved samples, limiting their utility for retrospective studies using archived biorepository specimens. Degradation of nucleic acids during storage exacerbates the problem.⁶

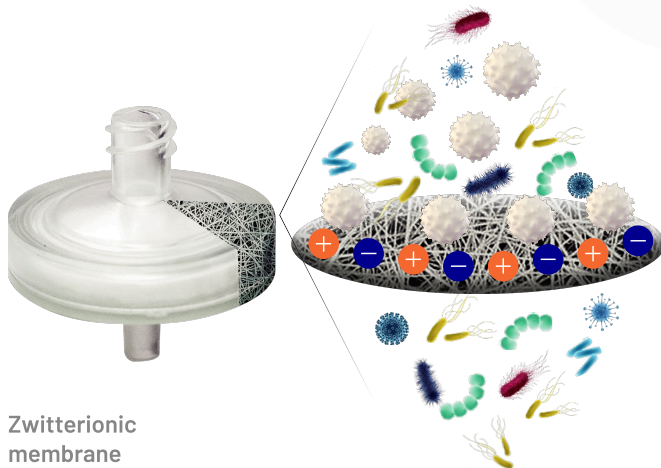
8. Cost Inefficiency

Many commercial host depletion kits require substantial per-sample investment, with variable efficacy. Inefficient depletion increases the proportion of host-derived reads, leading to sequencing cost inflation of 300–500% compared to targeted amplicon-based approaches. Additionally, excess host interference complicates bioinformatic analysis, extending computation time and storage requirements.⁷

Selecting an appropriate host depletion strategy in mNGS requires balancing efficiency, sample type compatibility, microbial preservation, and resource constraints. Ideally, host depletion improvements should focus on minimizing microbial loss while enhancing scalability and cost-effectiveness.

Resolving Efficiency-Integrity-Speed Tradeoffs with Zwitterionic Host Depletion

PaRTI-Seq incorporates the Devin™ Host Depletion Filter, a patented zwitterionic membrane that selectively captures host cells through electrophysical interactions. The reagent-free filtration step removes up to 99% of nucleated host cells in under two minutes without relying on size exclusion, accelerating sample preparation while preserving microbial genomic DNA (gDNA) integrity for downstream analysis (Figure 2).

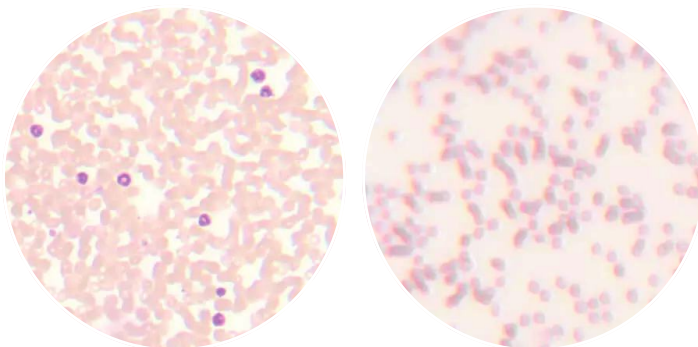


Zwitterionic
membrane

Figure 2. Devin Host Depletion Filter enables selective removal of host cells while preserving microbial gDNA integrity.

Zwitterionic Charge-Based Capture of Host Cells

The Devin filter employs Zwitterionic Interface ultra-Self-assembly Coating (ZISC) technology, featuring spatially organized cationic and anionic groups that interact with the surface charges of eukaryotic membranes. Nucleated host cells, characterized by distinct phospholipid bilayer charge profiles, form multiple non-covalent interactions with the membrane surface, including electrostatic attraction, hydrogen bonding, and dipole-dipole interactions (Figure 3).



Before Host Depletion

After Host Depletion

Figure 3. Blood smear microscopy before and after host depletion. Host leukocytes are efficiently removed post-filtration, with erythrocytes predominantly retained.

The resulting interaction energy is sufficient to retain host cells. Microbial structures such as peptidoglycan layers, viral capsids, fungal chitin, and parasitic glycoproteins lack the necessary charge complementarity and pass through unbound. Reliable across a wide range of biological sample types, the filtration mechanism ensures consistent and selective depletion.

Anti-Fouling Hydration Layer Enhances Recovery

Zwitterionic polymers form a tightly bound hydration shell, minimizing nonspecific adsorption of host DNA, proteins, and other biomolecules. Reducing host-derived background and preserving microbial content improves recovery efficiency and supports batch-to-batch consistency essential for high-resolution metagenomic applications.

Advanced Host Depletion for Metagenomic Applications

The Devin filter transforms host depletion from a blunt instrument into a precise, selective process tailored for metagenomics. Zwitterionic interactions enable targeted capture of nucleated host cells, while anti-fouling properties minimize contamination. Together, these features improve the accuracy of downstream molecular analyses, including NGS and PCR-based detection.

Case Study: Real-World PaRTI-Seq Evidence

Background

The Devin filter transforms host depletion from a blunt instrument into a precise, selective process tailored for metagenomics. Zwitterionic interactions enable targeted capture of nucleated host cells, while anti-fouling properties minimize contamination. Together, these features improve the accuracy of downstream molecular analyses, including NGS and PCR-based detection.

Clinical Evaluation and Results

To evaluate the clinical utility of the PaRTI-Seq workflow, blood samples from patients with suspected sepsis were analyzed at two major Taiwanese institutions: China Medical University Hospital (CMUH) and Taipei Veterans General Hospital (TVGH). Patient samples were processed using PaRTI-Seq and compared against a reference sample analyzed via conventional mNGS methods, based on data from a study conducted at National Taiwan University Hospital (NTUH).¹³

- **CMUH samples (n=87)** processed immediately using PaRTI-Seq showed high microbial recovery, with average microbial reads per million (RPM) values exceeding 16,000.
- **CMUH samples stored at 4°C for several days (n=51)** demonstrated a substantial reduction in microbial reads, with an average RPM of only 2,820, emphasizing the importance of timely processing.
- **TVGH samples (patient n=24, healthy n=24)** also demonstrated strong performance with PaRTI-Seq. One patient sample achieved 54,198 RPM, while others ranged from 1,200 to 4,770 RPM, confirming consistent microbial enrichment across cases.
- In contrast, the **reference sample (n=24)** processed using conventional mNGS protocols (without host depletion) showed microbial RPMs ranging from 66 to 148, highlighting the limited recovery of microbial reads without optimized workflows.
- Based on the **P-value analysis** ($P < 0.05$), all tested sample groups exhibit statistically significantly higher microbial RPM compared to the reference sample, indicating a more substantial microbial presence in these groups.

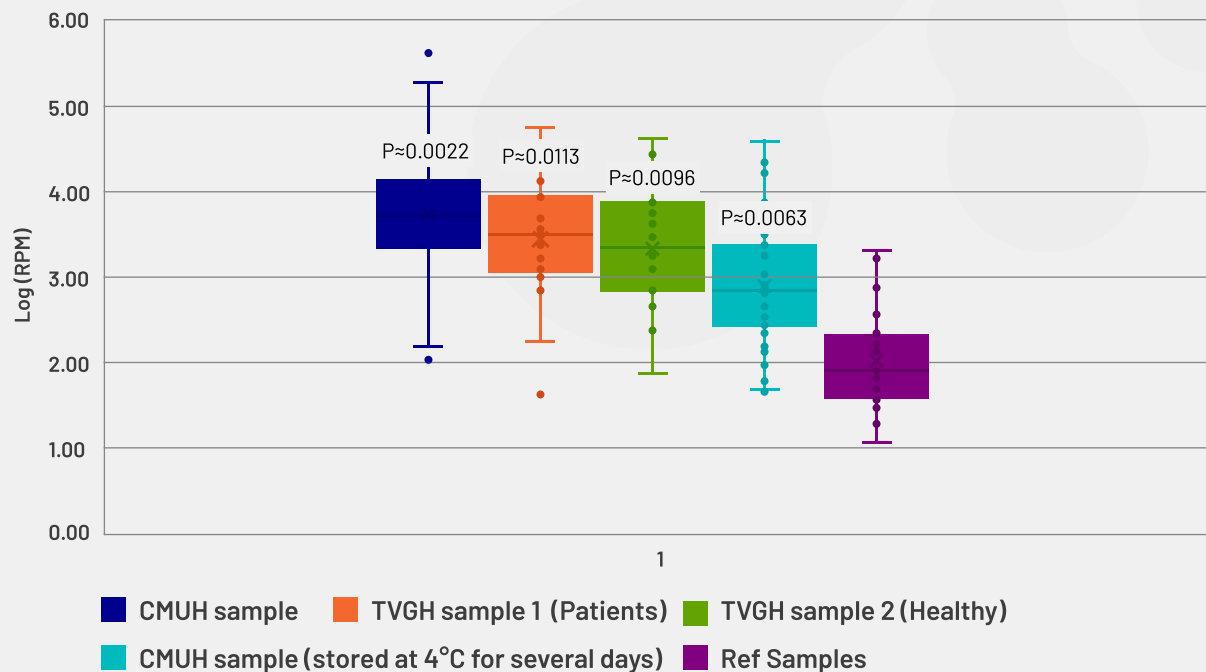


Figure 4. Comparison of microbial RPM (log scale) across different patient samples from CMUH, TVGH and reference.¹³

The P-values were obtained using Welch's t-test, comparing the microbial RPM of each sample group against that of the reference sample group.

Case Study Conclusion

The evaluation highlights the effectiveness of the PaRTI-Seq workflow in recovering microbial reads from blood samples in patients with hematologic malignancies and suspected sepsis. Compared to conventional mNGS methods, PaRTI-Seq demonstrated significantly higher microbial read recovery—even at ultralow input levels—while maintaining consistent performance across multiple hospital sites. The observed reduction in microbial reads in stored samples further underscores the importance of rapid sample processing. Collectively, these findings support the integration of PaRTI-Seq into clinical mNGS workflows to improve the sensitivity and reliability of pathogen detection in diagnostically challenging specimens.

Optimized Extraction: Why mNGS-Grade Reagents and Genomic DNA Output Matter

As metagenomic sequencing gains value in infectious disease research and clinical pathogen detection, upstream DNA extraction remains a critical determinant of assay sensitivity and specificity. Limited microbial biomass and high host background frequently impair performance, while contamination introduced during extraction can generate misleading microbial profiles.^{14,15}



The Devin Microbial DNA Enrichment Kit maximizes microbial DNA recovery using mNGS-grade reagents and a gDNA-based output. Its configuration supports robust microbial read yield, reduces the risk of false positives, and enhances reliability in downstream bioinformatic analysis.

Mitigating Reagent-Derived Contamination in Microbial mNGS

Reagent contamination remains a persistent and often underestimated source of error in mNGS. Numerous studies have demonstrated that DNA extraction kits, PCR-grade enzymes, and molecular biology buffers frequently harbor trace levels of microbial DNA introduced during reagent formulation, packaging, or storage.^{16,17} These exogenous nucleic acids can substantially interfere with sequencing-based microbial profiling, particularly in metagenomic applications where microbial biomass is limited relative to host material.

The consequences of reagent contamination are multifaceted and well-documented:

- **Misclassification of contaminant sequences** as authentic microbial reads, leading to false-positive identifications.
- **Lot-to-lot and brand variability** in background profiles, which impairs reproducibility across studies and laboratories.
- **Suppression of true pathogen reads** due to overwhelming background noise, particularly when contaminant species dominate the sequence pool.
- **Distorted microbial diversity metrics**, which may artificially inflate richness and skew community composition analyses.
- **Introduction of environmental or lab-specific microbes**, such as *Ralstonia*, *Acinetobacter*, and *Cutibacterium spp.*, that frequently appear in negative controls and may complicate interpretation of results.

To mitigate the risks, the Devin Microbial DNA Enrichment Kit is formulated with reagents specifically qualified for mNGS. Each production lot undergoes microbial DNA quantification, and a Certificate of Analysis (CoA) is issued verifying that exogenous DNA content of the most common pathogens such as *Escherichia coli*, *Klebsiella pneumoniae* remains below 100 genome copies per milliliter or the limit of detection of the test.

Common reagent contaminants are proactively excluded from formulations to minimize spurious alignments in taxonomic classification. In addition, the PaRTI-Seq workflow incorporates blank extraction controls and integrates background-aware bioinformatic filtering to reduce the influence of residual contamination.

These quality safeguards collectively reduce the contribution of reagent-derived sequences to metagenomic datasets and enhance the specificity of microbial detection. Figure 5 illustrates a representative CoA, documenting reagent background for one production lot of the Devin Microbial DNA Enrichment Kit.

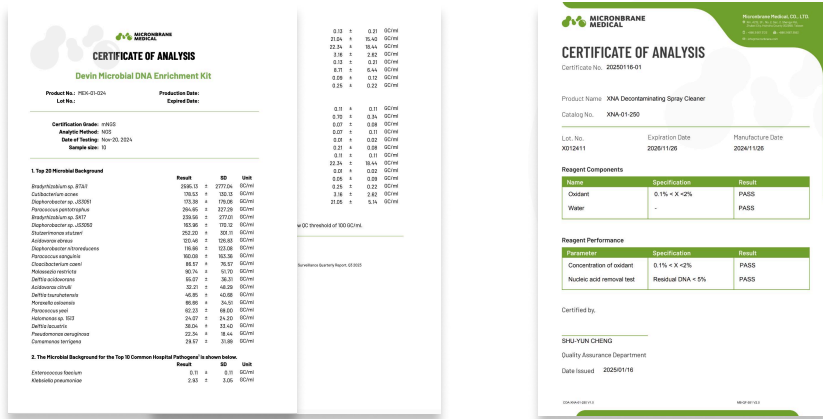


Figure 5. Certificate of Analysis documenting background microbial DNA levels in the Devin Microbial DNA Enrichment Kit.

Genomic DNA Enables Higher Microbial Read Yield

While cell-free DNA-based (cfDNA) approaches are common in oncology and noninvasive testing, their utility in microbial sequencing is limited. cfDNA derived from plasma is highly fragmented, predominantly human in origin, and includes minimal microbial content. In contrast, genomic DNA (gDNA) extracted from the cellular pellet—especially when paired with effective host depletion—preserves intact microbial genomes, resulting in higher microbial read counts and improved taxonomic resolution.¹⁸

Chen, YC, et. al., conducted a comparative evaluation using matched blood samples from patients with culture-confirmed infections. Each sample was processed for both cfDNA (from supernatant) and gDNA (from the cellular pellet). The resulting data, shown in Table 1, demonstrate that gDNA consistently yields significantly more species-specific microbial reads than cfDNA.¹⁹

Pellet gDNA

→

Supernatant cfDNA

Devin™ filter
n/f – no filter

Sample ID	Clinical Diagnosis	BC1	BC2	Species specific gDNA Reads	Species specific cfDNA Reads
MB-210811-1	<i>Escherichia coli</i>	+	+	Devin™ filter 1440 n/f 47	Devin™ filter 119 n/f 133
MB-210909-1	<i>Proteus mirabilis</i>	+	+	Devin™ filter 6234 n/f 1189	Devin™ filter 0 n/f 0
MB-210930-1	<i>Klebsiella pneumoniae</i>	+	+	Devin™ filter 7771 n/f 645	Devin™ filter 1885 n/f 669
MB-211124-1	MRSA	+	+	Devin™ filter 55 n/f 0	Devin™ filter 0 n/f 20
MB-220506-1	<i>Escherichia coli</i> / <i>Candida albicans</i>	+	+	Devin™ filter 525 n/f 66	Devin™ filter 31 n/f 20
MB-211205-1	<i>Pseudomonas aeruginosa</i>	+	–	Devin™ filter 378 n/f 54	Devin™ filter 18 n/f 38
MB-211222-1	<i>Achromobacter species</i>	+	–	Devin™ filter 9407 n/f 8168	Devin™ filter 1424 n/f 26
MB-220116-1	<i>Escherichia coli</i>	+	–	Devin™ filter 1253 n/f 16	Devin™ filter 59 n/f 63

Table 1. Host-depleted gDNA produces greater microbial reads than cfDNA.

In eight clinical samples, gDNA extracted using the Devin filter produced a mean of 2,359 microbial reads per million (RPM), whereas cfDNA yielded only 95 RPM on average. Several cfDNA samples failed to detect the pathogen despite blood culture confirmation. These results suggest that microbial reads in cfDNA samples often approach background levels, increasing the risk of false negatives and reducing the overall utility of cfDNA in metagenomic pathogen detection.

Feature / Capability	cfDNA Extraction Kits	Devin Microbial DNA Enrichment Kit
Primary Purpose	Extract fragmented human cfDNA from plasma or serum	Extract microbial gDNA from liquid biopsy with host cell depletion
Target Analyte	Short human-derived cfDNA fragments (often 150–200 bp)	Intact microbial genomic DNA
Microbial DNA Yield	Low—microbial reads are typically <1% of total yield	High—preserves microbial integrity with minimal contamination
Suitability for mNGS of Infections	Low—requires ultra-deep sequencing or additional depletion steps	High—designed specifically for pathogen detection
Enrichment Strategy	No enrichment—non-specific recovery of all free-floating DNA	Physical host cell depletion + targeted microbial lysis and capture
DNA Fragment Size	Fragmented cfDNA (150–200 bp)	Intact microbial genomes (hundreds of kb)
Reagent Format	Varies; often cold-chain dependent	Pre-filled, room-temp stable
Sample Compatibility	Primarily plasma or serum (1–5 mL)	ALL liquid biopsies (e.g. whole blood, plasma, CSF, BAL up to 10 mL)
Integration with mNGS Workflows	Requires ultra-deep NGS	Validated as part of PaRTI-Seq

The use of mNGS-grade reagents, with each lot verified to contain minimal background DNA, helps ensure that microbial reads reflect true biological content rather than contamination from reagents or the environment. For research and clinical applications that require accurate microbial profiling, purpose-built extraction chemistries like Devin offer a validated and reliable foundation.

Moreover, genomic DNA-based extraction with the Devin Microbial DNA Enrichment Kit improves both the sensitivity and specificity of mNGS. Compared to cfDNA-based approaches, gDNA yields a more comprehensive and intact representation of microbial communities, increasing the likelihood of accurate pathogen identification while reducing the risk of false positives or missed infections.

Streamlined Library Preparation Across the Biomass Spectrum

Library preparation remains a frequent failure point in mNGS, particularly when working with samples that vary widely in biomass. Low-biomass specimens such as cerebrospinal fluid (CSF), plasma, and tissue biopsies often contain microbial DNA at levels below 1% of total nucleic acid content.²⁰

The imbalance can lead to inefficient amplification, reduced detection sensitivity, and increased sequencing requirements. In contrast, high-biomass samples commonly used in microbiome studies, such as stool, present different challenges. These include amplification bias and diminished taxonomic resolution due to the overrepresentation of dominant taxa.²¹



The PaRTI-Seq workflow incorporates the **Unison Ultralow DNA NGS Library Preparation Kit**, which is specifically designed to support metagenomic sequencing across this full range of input conditions. Using proprietary Tn5-based transposase chemistry, the kit enables efficient and reproducible library construction from as little as 10 picograms of DNA.

This input level is below the detection threshold of conventional quantification methods. The protocol reduces hands-on time to under one hour and remains compatible with both automated and manual workflows. Its combination of sensitivity and operational flexibility supports consistent, high-quality library generation from both low-biomass clinical specimens and complex microbiome samples.

Efficient and Consistent Library Preparation

Traditional library preparation protocols often involve multiple enzymatic steps, intermediate cleanups, and manual normalization. Each step can introduce variability, increase complexity, and raise the risk of sample loss. Unison streamlines the process, with a total runtime of 150 minutes and just 45 minutes of hands-on time. For samples above a certain amount of input (e.g. 1ng), a built-in library size normalization step further simplifies the workflow by equalizing library concentrations without the need for pre-amplification quantification or dilution. This design improves reproducibility, reduces error, and enables higher-throughput processing with greater consistency.

Robust Performance at Ultralow Input Levels

The Unison kit is optimized to preserve microbial community structure and taxonomic resolution even when DNA inputs are minimal. Across a range of concentrations, it has demonstrated strong performance in maintaining alpha diversity metrics, including Shannon, Simpson, Evenness, and Richness indices. These results reflect its ability to recover a comprehensive and representative microbial profile, including rare or low-abundance taxa, from samples with limited starting material.

Taken together, these features position the PaRTI-Seq workflow as a reliable, scalable, and cost-effective approach for mNGS of low-biomass samples. It enables the generation of actionable microbial insights in scenarios where conventional methods often underperform.

Case Study:

Superior Performance of Unison Ultralow DNA Library Prep in Ultra-Low Input Metagenomics

Library preparation from sub-nanogram quantities of DNA remains a critical bottleneck in both clinical and mNGS workflows.²² Clinical specimens, such as cerebrospinal fluid or plasma, frequently contain trace levels of microbial DNA, and re-sampling is often not feasible. Likewise, research samples including biopsy tissues, swabs, or host-depleted extracts routinely yield limited nucleic acid input. Under these constraints, even marginal improvements in library efficiency and fidelity can substantially influence diagnostic sensitivity and taxonomic resolution. As such, the ability to construct compositionally accurate and reproducible libraries from picogram- to low-nanogram inputs has become essential for modern low-biomass metagenomic applications.

Study Background

To assess kit performance at sub-nanogram input levels, AMILI (Singapore) conducted a benchmarking study using serial dilutions of stool-derived microbial DNA. Although not a low-biomass matrix, stool provides a compositionally rich and consistent model for testing input sensitivity under controlled conditions. This approach isolates the effect of DNA quantity on library performance while minimizing variability from extraction method or host contamination.

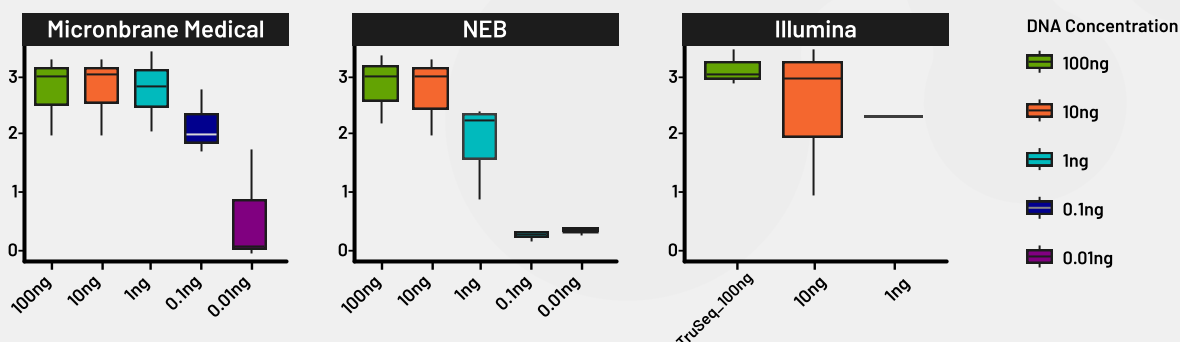
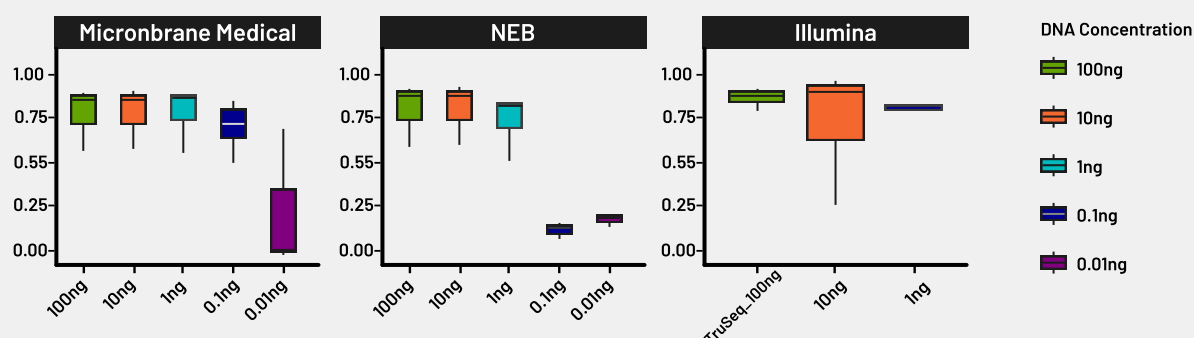
Materials and Methods

Stool microbiome DNA was extracted from a single healthy donor using the QIAamp PowerFecal Pro DNA Kit (Qiagen) and serially diluted to generate input amounts ranging from 100 ng to 10 pg. For each input level, triplicate libraries were prepared using four different library preparation kits—Unison Ultralow DNA NGS Library Preparation Kit, Illumina TruSeq DNA Nano, Illumina DNA Prep, and NEBNext Ultra II—according to their respective manufacturers' protocols. Due to input limitations and quality control outcomes, Illumina TruSeq DNA Nano was tested only at 100 ng, while Illumina DNA Prep was evaluated at 10 ng and 1 ng; lower input libraries (0.1 ng and 0.01 ng) failed to meet the quality thresholds for downstream sequencing.

All libraries were sequenced on the Illumina NovaSeq platform, generating approximately 20 million paired-end reads per sample. Raw sequencing reads underwent quality control using KneadData (v0.12.0), which included the removal of host (human) DNA and low-quality reads. Taxonomic profiling was performed using MetaPhlAn4 (v4.1.1)²³ within the bioBakery²⁴ framework. Downstream analyses, including alpha and beta diversity metrics, and taxonomic composition comparisons, were conducted in R (v4.2).

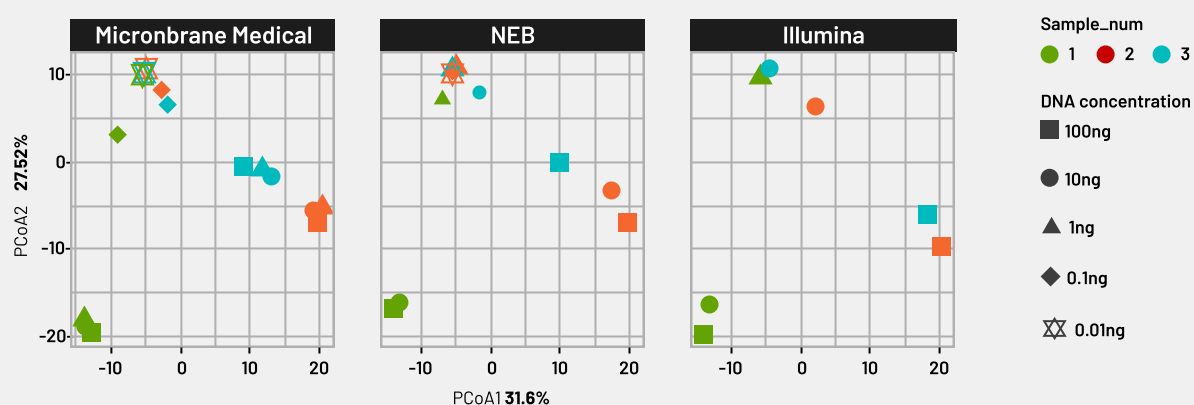
Results

Unison produced consistent alpha diversity (Shannon and Simpson indices) down to 1 ng input. In contrast, NEB required ≥ 10 ng to maintain comparable diversity, and Illumina DNA Prep showed marked diversity loss below 10 ng.

**Shannon's index (Species)****Simpson's index (Species)**

Case Study Figure 1. Unison maintains microbial diversity at inputs down to 1 ng, while NEB and Illumina show substantial declines below this threshold.

Beta diversity analysis showed tight replicate clustering for Unison down to 1 ng, indicating high reproducibility. NEB and Illumina kits showed greater inter-replicate dispersion at low input levels, suggesting amplification bias and reduced consistency.

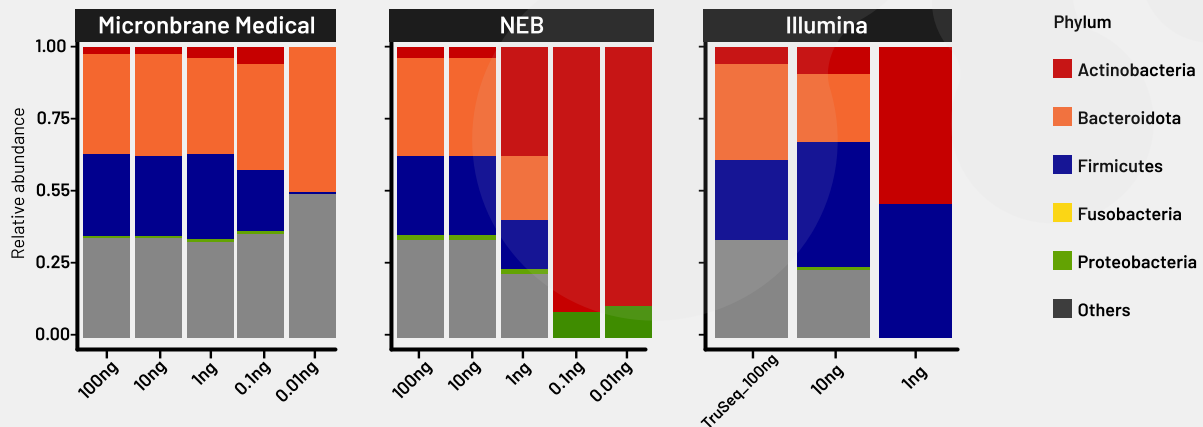
Aitchison distance (Species)

Case Study Figure 2. Unison demonstrates strong replicate consistency at 1 ng, unlike NEB and Illumina kits, which show increasing dispersion with lower input.

At the phylum level, Unison maintained stable community structure down to 0.1 ng. NEB and Illumina kits displayed compositional distortion, including overrepresentation of *Actinobacteria* likely a result of input-dependent PCR bias.



Phylum distribution



Case Study Figure 3. Unison preserves phylum-level structure at 0.1 ng input, whereas NEB and Illumina skew microbial representation.

Case Study Conclusion

Unison consistently generated high-quality libraries across a broad input range, including sub-nanogram quantities. When applied to challenging metagenomic samples, it demonstrated strong preservation of sample complexity and taxonomic structure. Key performance outcomes included:

- **Reproducibility at ultra-low input:** Reliable library generation from as little as 0.1ng of DNA
- **Preservation of diversity:** High microbial diversity retained across replicates
- **Minimized taxonomic bias:** Accurate composition maintained at low input concentrations
- **Sample integrity conservation:** Streamlined workflow reduced input loss
- **Versatility across sample types:** Compatible with degraded, low-biomass, or trace input DNA

Used as part of the PaRTI-Seq workflow, including host depletion and integrated bioinformatics, Unison supports high-confidence microbial and pathogen detection, even from limited or complex research samples.

Computational Challenges in Metagenomic Sequencing Workflows

While mNGS offers broad taxonomic coverage and the ability to detect multiple classes of microorganisms in a single assay, the data it generates is inherently complex. Each sequencing run can produce millions of reads, encompassing not only microbial genomes but also vast amounts of host DNA, environmental contaminants, and unclassified sequences.

Transforming this raw data into accurate, clinically meaningful insights requires sophisticated computational tools, high-performance infrastructure, and carefully curated reference databases.

Unlike targeted assays, mNGS lacks standardized analytical frameworks. Laboratories must differentiate true pathogens from background noise, particularly in low-biomass samples where trace contaminants can produce misleading signals. Further, challenges such as ambiguous taxonomic resolution, low-abundance organism detection, and inconsistent classification algorithms can undermine diagnostic confidence.²⁵ These barriers are exacerbated in research and clinical environments lacking dedicated bioinformaticians, computational resources, or regulatory-aligned workflows.

PaRTI-Seq Analysis: Scalable, Interpretable Solutions for Metagenomic Applications



PaRTI-Seq Analysis is a bioinformatics pipeline available in three configurations: a complimentary Research Use Only (RUO) version for exploratory applications, an AWS-based cloud deployment for scalable remote access, and PaRTIBot, a secure on-premise system designed for offline analysis in data-restricted environments.

Feature	RUO Analysis (Cloud)	Clinical Analysis (Cloud)	PaRTIBot (On-Premise Clinical Analysis)
Curated Pathogen Database	Shared database of >1,500 pathogens	Shared database of >1,500 pathogens	Shared database of >1,500 pathogens
Bioinformatics Expertise	Not required	Not required	Not required
Deployment Platform	Cloud (AWS)	Cloud (AWS)	Local offline workstation
Hardware Requirement	Regular computer	Regular computer	Requires PaRTIBot device
Clinical Thresholds	Not applied	Applied	Applied
Spike-in Normalization	Not supported	Included	Included
Quantitative Outputs	Read counts only	Genome copies/mL, coverage uniformity, abundance plots, QC flags	Genome copies/mL, coverage uniformity, abundance plots, QC flags
Environmental Contamination Flags	Not included, but can be inferred from read counts	Report includes detection of outlier species	Report includes detection of outlier species
Report Format	Spreadsheet (.csv/.xlsx)	Structured clinical PDF report	Structured clinical PDF report
Customizable Reports	Not supported	Lab branding supported	Lab branding supported
Turnaround Time	Variable	~30 minutes per sample	~30 minutes per sample
Data Privacy	Not PHI-compliant	Secure cloud handling; PHI-compliant	Secure cloud handling; PHI-compliant
Cost & Licensing	Free with code	Licensed	Hardware + software license required
Ideal Use Case	Exploratory or Translational research	Cloud-based clinical diagnostics	Local clinical diagnostics with privacy constraints

By optimizing sequencing efficiency and streamlining analysis, PaRTI-Seq reduces cost and complexity while maximizing diagnostic yield. Its flexible architecture supports mNGS adoption in both resource-limited and high-throughput environments.

More importantly, PaRTI-Seq addresses the core computational and interpretive challenges of mNGS, delivering high-confidence pathogen detection with the speed and clarity needed for timely, informed decision-making in diverse laboratory settings.

Economic Scalability Considerations in mNGS

One of the most significant barriers to widespread adoption of mNGS is the high cost per sample—driven by deep sequencing requirements, labor-intensive assays, and complex bioinformatics infrastructure. The PaRTI-Seq workflow delivers a streamlined, integrated approach that significantly reduces both fixed and variable costs.

Operational Cost Modeling and Efficiency Gains

The PaRTI-Seq Value Calculator demonstrates measurable advantages across key operational metrics. When compared to conventional mNGS methods, PaRTI-Seq consistently outperforms in terms of cost efficiency and sample throughput.

- **Up to 75% Cost Reduction**

PaRTI-Seq lowers per-sample costs by combining effective host DNA depletion, ultralow-input library preparation, and built-in normalization with reduced sequencing depth requirements.

- **Reduced Sequencing Load**

Accurate pathogen detection is achieved with only 5 million reads per sample, compared to the 20 million reads typically required by standard workflows. This substantially decreases sequencing costs and allows for higher sample multiplexing per run.

- **Increased Throughput**

For Platforms such as Illumina MiniSeq or MiSeq, the streamlined protocol enables processing of up to 1,040 samples per year, a fourfold increase over the 260 samples achievable with conventional methods—without requiring additional instrumentation or staff.

PARTI-SEQ VALUE CALCULATOR

Select your sequencing platform Illumina MimiSeq

Choose your currency USD

Calculating the Costs* per Sample

Reagents and consumables

- Sample preparation (per sample) ?
- Microbial DNA extraction (per sample) ?
- NGS library preparation and analysis (per sample) ?
- Sequencing (per run) ?

System Control ? ☒

Labor per Person

☒ annual ☐ monthly ☐ hourly

Number of Staff Dedicated to mNGS (enter a number)

- Days of sequencing runs per week ?

Maintenance per year ?

(usually 8% of purchase price)

Platform cost

- Full purchasing price of platform ?
- Salvage ?
- Depreciation period, years ?

With PaRTI-Seq

Without PaRTI-Seq

20	0
7	7
73	73
1429	1429
24.60	24.60
1	1
5	5
3960	3960
9,900.00	9,900.00
49500	49500
0	0
5	5

Cost Benefits of PaRTI-Seq

The cost per sample is NaN% less using PaRTI-Seq because of the host depletion, mNGS-grade reagents and library prep efficiency and thus requires 75% less sequencing reads with the same sensitivity.

Pathogen
identification in 5M
reads with PaRTI-Seq

537.28

Need 20M reads without
PaRTI-Seq

1,729.11

Figure 7. Calculation of cost per sample using the PaRTI-Seq Value Calculator;



Sustainable and Scalable Cost Advantages

In addition to direct financial outcomes, PaRTI-Seq contributes to broader cost savings by:

- ✓ Reducing reagent waste and reprocessing due to high success rates with low-biomass inputs
- ✓ Eliminating the need for in-house bioinformatics infrastructure through fully automated, cloud-based or on-premise analysis
- ✓ Lowering data storage and compute demands by minimizing total read volume per sample

By significantly lowering the cost per sample, increasing processing efficiency, and enabling reliable microbial detection at ultralow input levels, PaRTI-Seq supports the scalable and economically sustainable adoption of mNGS in both clinical and research laboratories. These cost benefits, validated through an ROI-based model, position PaRTI-Seq as a viable solution for institutions seeking to expand the diagnostic utility of sequencing without incurring prohibitive operational overhead.

Contamination, Complexity, Cost. Solved.

PaRTI-Seq delivers measurable improvements across the core barriers to mNGS implementation. Its zwitterionic host depletion filter removes over 99% of host nucleated cells in under two minutes, without the use of reagents or centrifugation. This substantial reduction in host background enables detection of microbial reads exceeding 16,000 RPM in clinical blood samples, compared to fewer than 150 RPM with conventional workflows.

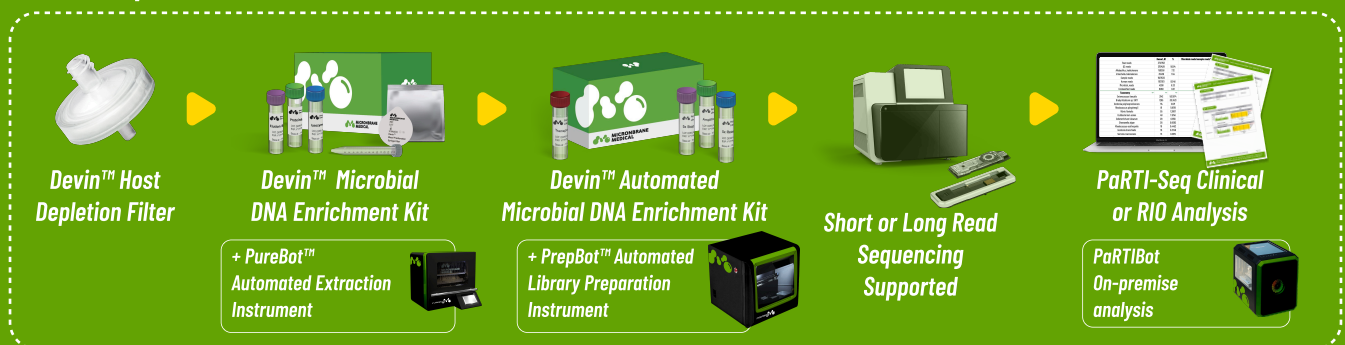
Contamination control is further strengthened through the use of batch-validated, mNGS-grade reagents that minimize false positives, particularly in low-biomass samples where trace background DNA can distort taxonomic interpretation. The ultralow-input library preparation kit supports inputs as low as 10 pg and maintains diversity metrics such as Shannon and Simpson indices across a broad biomass range.

Workflow efficiency is enhanced by reducing hands-on time to under 45 minutes and incorporating automated library normalization. PaRTI-Seq Analysis delivers taxonomically annotated pathogen reports in under 30 minutes, eliminating the need for internal bioinformatics teams. The platform supports both secure on-premise (PaRTIBot) and cloud-based deployment options.

Operationally, the system reduces sequencing and reagent costs by up to 75%, increases sample throughput by a factor of four, and enables laboratories to reach break-even at fewer than two samples per day. These efficiencies make high-resolution mNGS feasible for diagnostic and translational labs operating under time, cost, and personnel constraints.

Rather than optimizing a single step, PaRTI-Seq redefines the end-to-end mNGS workflow—improving detection performance, reducing resource demands, and expanding the practical utility of mNGS.

PaRTI-Seq Workflow



References

1. Jurasz H, Pawłowski T, Perlejewski K. Contamination issue in viral metagenomics: problems, solutions, and clinical perspectives. *Front Microbiol.* 2021;12:745076.
2. Batool M, Galloway-Peña J. Clinical metagenomics—challenges and future prospects. *Front Microbiol.* 2023;14:1123456.
3. Gu W, Deng X, Miller S. Host-Pathogen Interactions and the Application of Metagenomic Next-Generation Sequencing in Clinical Microbiology. *J Clin Microbiol.* 2019;57(8):e00326-19.
4. Santiago A, Panda S, Mengels G, Martinez-Martinez D, Be-Sarto M, Rodrigues R, et al. Impact of Host DNA and Sequencing Depth on the Taxonomic Resolution of Whole Metagenome Sequencing for Microbiome Analysis. *Front Microbiol.* 2019;10:1277.
5. Gauthier C, Gargani S, Gasc C, Le-Senechal C, Bresson E, Garel M, et al. Host DNA Depletion in Clinical Samples for Metagenomic Applications. *Front Cell Infect Microbiol.* 2022;12:917536.
6. Cormier Y, Chakradhar S, Sariya S, Hida A, Shen P, Yang J, et al. Host DNA depletion on frozen human respiratory samples enables successful metagenomic sequencing for microbiome studies. *Commun Biol.* 2024;7(1):507.
7. Islam MS, Li H, Lim B, Chau CH, Unemo M, Yang L, et al. Advances in host depletion and pathogen enrichment methods for rapid sequencing-based diagnosis of bloodstream infection. *J Mol Diagn.* 2024;26(9):741-53.
8. Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection. *Annu Rev Pathol.* 2019;14:319-38.
9. Feng Y, Chen H, Li Z, Liu Y, Nan Z, Liu Y, et al. A Host DNA Depletion Method for Metagenomic Analysis of Human-Derived Samples. *Front Microbiol.* 2021;12:772648.
10. Feehery GR, Yigit E, Oyola SO, Langhorst BW, Schmidt VT, Stewart FJ, et al. A method for selectively enriching microbial DNA from contaminating vertebrate host DNA. *PLoS One.* 2013;8(10):e76096.
11. Gu W, Crawford ED, O'Donovan BD, Wilson MR, Chow ED, Retallack H, et al. Depletion of abundant sequences by hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications. *Genome Biol.* 2016;17(1):41.
12. Longhi G, Argentini C, Fontana F, Tarracchini C, Mancabelli L, et al. Saponin treatment for eukaryotic DNA depletion alters the microbial DNA profiles by reducing the abundance of Gram-negative bacteria in metagenomics analyses. *Microbiome Res Rep.* 2023;3(1):4.
13. Liu WD, Yen TY, Liu PY, Wu UI, Bhan P, Li YC, et al. Clinical application of metagenomic next-generation sequencing in patients with hematologic malignancies suffering from sepsis. *Microorganisms.* 2021;9(11):2309.
14. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 2014;12:87.
15. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol.* 2019;27(2):105-17.
16. Burnham P, Dadhania D, Heyang M, Chen F, Westblade LF, Suthanthiran M, et al. Separating the signal from the noise in metagenomic cell-free DNA sequencing. *Microbiome.* 2022;10(1):17.
17. Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* 2016;8:24.
18. Miller S, Lurf K, O'Connell AK, Galiya S, Lyden A, Fung B, et al. Metagenomic Sequencing of Cell-Free DNA and Cellular-Fraction DNA for Diagnosis of Sepsis. *Microbiol Spectr.* 2021;9(3):e0063221.
19. Chen YC, et al. Evaluation of a metagenomic next-generation sequencing assay with a novel host depletion method for pathogen identification in septic patients. *medRxiv [Preprint].* 2023. doi:10.1101/2023.03.28.23287867.
20. Simner PJ, Miller S, Carroll KC. Understanding the promises and hurdles of metagenomic next-generation sequencing for clinical infectious diseases diagnosis. *Clin Infect Dis.* 2018;66(5):778-88.
21. Ruan Z, Lusk CP. Challenges and strategies in shotgun metagenomics. *Front Microbiol.* 2022;13:1042885.
22. Marquet M, Zöllkau J, Pastuschek J, Viehweger A, Schleußner E, Makarewicz O, et al. Sensitivity of shotgun metagenomics to host DNA: Abundance matters. *Microb Genom.* 2020;6(2):e000377.
23. Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat Biotechnol.* 2023;41:1633-44.
24. McIver LJ, Abu-Ali G, Franzosa EA, Schwager R, Morgan XC, Waldron L, et al. bioBakery: a meta'omic analysis environment. *Bioinformatics.* 2018;34(7):1235-7.
25. Salzberg SL, Breitwieser FP, Kumar A, Hao H, Burger P, Rodriguez FJ, et al. Next-generation sequencing in pathogen identification: Implications for clinical microbiology. *Clin Microbiol Rev.* 2016;29(4):687-703.