The Microbe 7 (2025) 100398

Contents lists available at ScienceDirect

The Microbe

journal homepage: www.sciencedirect.com/journal/the-microbe

Comparative evaluation of DNA extraction protocols for neonatal gut microbiota profiling in a resource-limited setting

Chinenye Akpulu [a,b,*], Aditya Kumar Lankapalli [a], Rida Toufiq [a], Kate Cook [a], Edward AR Portal [a,c], Rashida Yakubu Khalid [d], Aisha Mukkadas [d], Safiya Gambo [e], Aliyu Aminu [f], Kenneth Iregbu [g], Kathryn Thomson [a], Timothy Walsh [a], Kirsty Sands [a,c]

[a] University of Oxford, Ineos Oxford Institute, Oxford, United Kingdom
[b] University of Oxford, Interdisciplinary Bioscience DTP, Oxford, United Kingdom
[c] Cardiff University, Microbiology, Cardiff, United Kingdom
[d] Murtala Muhammad Specialist Hospital, Microbiology, Kano, Nigeria
[e] Murtala Muhammad Specialist Hospital, Paediatrics Unit, Kano, Nigeria
[f] Aminu Kano Teaching Hospital, Microbiology, Kano, Nigeria
[g] National Hospital Abuja, Clinical Microbiology and Parasitology, Abuja, Nigeria

ARTICLE INFO

Keywords:
Neonatal gut microbiota
DNA extraction
Metagenomics
Low-resource settings
Long-read sequencing
Sample preservation
Stool sample processing

ABSTRACT

Accurate and reproducible gut microbiome profiling depends heavily on the DNA extraction method, particularly in low-biomass samples such as neonatal stool. In this study, we evaluated the performance of three commercially available DNA extraction kits; QIAamp Fast DNA Stool Mini, DNeasy PowerSoil Pro, and ZymoBIOMICS DNA Miniprep on neonatal stool samples collected in a resource-limited hospital in Kano, Nigeria. Samples were stored under various conditions (temperature and preservatives), and DNA was extracted and sequenced using Oxford Nanopore Technologies. DNA yield differed substantially across extraction kits and storage conditions. The bead-beating-based kits: PowerSoil and ZymoBIOMICS, consistently outperformed the QIAamp Fast DNA Stool Mini kit, which produced negligible yields across all conditions. Both bead-beating kits achieved the highest DNA concentrations when samples were processed fresh and without preservatives, while yields declined sharply after just one day of storage. Although overall DNA yields were similar between PowerSoil and ZymoBIOMICS at all time points, PowerSoil extracts produced longer sequencing reads and higher-quality assemblies. Specifically, PowerSoil-derived libraries generated higher read-level N50 values than those from ZymoBIOMICS at both Day 0 and Week 6. While these differences were not statistically significant, a moderate effect size at Day 0 (rank-biserial correlation = 0.60) suggests a potential advantage in genome assembly continuity. Additionally, PowerSoil had a shorter processing time, enhancing its suitability for long-read metagenomics workflows in resource-limited settings. We conclude that same-day processing using bead-beating-based extraction kits improves yield and may reduce bias in neonatal gut microbiome studies. These findings are especially relevant for low-resource settings, where equipment limitations and delayed sample processing can impact data quality and study scalability.

1. Introduction

The human gut microbiome plays a vital role in maintaining host health, influencing immune development, nutrient absorption, protection against pathogens, and contributing to biosynthesis of essential vitamins and amino acids (Olin et al., 2018; Bäumler and Sperandio, 2016; Alsharairi, 2023). It also contributes to maintaining intestinal barrier integrity and modulating host metabolism and inflammation (Trompette et al., 2014). Disruption of this microbial community known as dysbiosis, has been associated with various health conditions, including metabolic, gastrointestinal, and immunological disorders (Gomaa, 2020; Sommer and Bäckhed, 2013; Walters et al., 2014; Leong et al., 2020).

In neonates, gut microbiome colonisation begins at birth and undergoes rapid changes in early life. Factors such as delivery mode, feeding type, antibiotic exposure, environment, and geographic location

* Corresponding author at: University of Oxford, Ineos Oxford Institute, Oxford, United Kingdom.
E-mail address: Chinenye.akpulu@linacre.ox.ac.uk (C. Akpulu).

https://doi.org/10.1016/j.microb.2025.100398
Received 22 January 2025; Received in revised form 17 May 2025; Accepted 29 May 2025
Available online 29 May 2025

can influence the initial microbial community structure (Dominguez-Bello et al., 2010; Dogra et al., 2015; Reyman et al., 2022; Rinninella et al., 2019; Thänert et al., 2022; Baumann-Dudenhoeffer et al., 2018; Pehrsson et al., 2016). This early window is especially dynamic and unstable, typically reaching a more stable, adult-like composition by two to three years of age (Xiao and Zhao, 2023). The initial microbial exposure and colonisation patterns are thought to have lasting impacts on immune programming and disease susceptibility, making it a critical period for microbiome-targeted interventions.

Despite its importance, studying the neonatal gut microbiome presents significant technical challenges. Neonatal stool, especially meconium, the first stool passed by newborns often contains low bacterial biomass, small sample volumes, and a high proportion of host (human) DNA (Rackaityte et al., 2020; Dos Santos et al., 2021). These properties make neonatal samples particularly susceptible to contamination during DNA extraction and library preparation, introducing substantial variability and increasing the risk of inaccurate microbial profiling. Furthermore, the relatively low microbial load in early stool samples may lead to random sequencing noise and misinterpretation of results, especially in metagenomic analyses where resolution and accuracy depend heavily on DNA input quality and yield (Weyrich et al., 2019; Selway et al., 2020; Jurasz et al., 2021).

The advent of long-read sequencing technologies, such as those provided by Oxford Nanopore Technologies (ONT), has expanded the possibilities of microbiome research by allowing for improved genome assembly, taxonomic characterisation, and more accurate profiling of antimicrobial resistance genes (ARGs) and mobile genetic elements (Maghini et al., 2021; Gand et al., 2023; Boostrom et al., 2022). These technologies however require high-quality, high molecular weight DNA and are particularly sensitive to fragmentation and contamination. In neonatal studies, where samples are inherently low biomass, DNA extraction becomes a critical bottleneck. Poor extraction efficiency, sample degradation, or over-fragmentation can compromise downstream sequencing quality, skew taxonomic profiles, and impair resistome detection (Fernández-Pato et al., 2024; Wagner Mackenzie et al., 2015; Wesolowska-Andersen et al., 2014).

DNA extraction protocols vary widely in their ability to lyse different microbial taxa, minimise host DNA contamination, and recover intact DNA molecules. Studies have shown that the choice of extraction method significantly affects taxonomic diversity, relative abundance estimates, and detection of low-abundance organisms or gene (Quince et al., 2017; Costea et al., 2017; Lim et al., 2020). Recognising this, the International Human Microbiome Standards (IHMS) consortium recommended a standardised protocol; Protocol Q, for adult stool DNA extraction using short-read sequencing (e.g., Illumina) (Wesolowska-Andersen et al., 2014). While increased standardisation such as Protocol Q has improved study comparability, they are often optimised for adult samples and are not well-suited for neonatal stool, which differs in consistency, composition, and microbial load. Moreover, current protocols are often complex, time-consuming, and dependent on specialised equipment and consumables that may not be available or affordable in many low- and middle-income countries (LMICs) (Maghini et al., 2021; Gand et al., 2023).

Consumable access and usage are problematic requiring resolution given the growing global interest in understanding early-life microbiome development in diverse settings. Neonatal populations in LMICs are of special interest due to high rates of antibiotic exposure, infectious disease burden, and environmental differences that may significantly shape the developing gut microbiome and resistome. Yet, despite their importance, LMIC populations remain underrepresented in microbiome studies, largely due to infrastructural and methodological limitations (Makhalanyane et al., 2023; Ayeni et al., 2024). There is a pressing need for protocols that are not only scientifically robust but also feasible, scalable, and cost-effective for use in low-resource laboratory environments.

To address these challenges, we selected three commercially available DNA extraction kits; QIAamp Fast DNA Stool Mini, DNeasy PowerSoil Pro, and ZymoBIOMICS DNA Miniprep, based on their frequent use in published gut microbiome studies, compatibility with long-read sequencing, and applicability in LMIC environments. A preliminary literature review was conducted using PubMed with search terms such as "stool DNA extraction kit metagenomics", "long-read sequencing microbiome", and "DNA extraction LMIC", with priority given to methods referenced in studies with high citation counts, minimal equipment requirements, and suitability for low biomass inputs. The selected kits represent a spectrum of trade-offs between lysis efficiency, DNA integrity, cost, and portability. For example, PowerSoil is widely used for its bead-beating effectiveness and has been applied in long-read metagenomic studies, (Maghini et al., 2021; Costea et al., 2017) while ZymoBIOMICS has shown promise in preserving high-molecular-weight DNA with minimal fragmentation (Wongsurawat et al., 2019; Sudjai et al., 2022). The QIAamp Fast DNA Stool Mini kit is frequently used in short-read and clinical contexts but is known to have limitations for hard-to-lyse organisms and low-input samples (Fernández-Pato et al., 2024). These considerations are particularly important for field studies, where factors such as cold chain limitations, access to reagents, and ease of protocol execution may outweigh laboratory preferences. In this study, we also incorporate considerations such as non-refrigerated storage, (Song et al., 2016) and the need to minimise DNA fragmentation to support high-contiguity long-read assemblies (Wongsurawat et al., 2019).

Using Oxford Nanopore Technologies, which has a comparably low start-up cost, portability, and suitability for low-resource environments, (Boostrom et al., 2022) we assessed DNA yield and downstream taxonomic profiling across multiple storage conditions, extraction protocols, and time points. Our comparison focused on identifying protocols that are not only biologically effective but also logistically practical for use in field settings where immediate extraction is not always feasible due to constraints like limited laboratory infrastructure, supply chain delays, and geographic inaccessibility. By optimising upstream sample preparation, our goal is to enable reliable, high-quality metagenomic data generation in neonatal populations from LMICs. In doing so, we aim to reduce methodological barriers to microbiome research in underrepresented settings and promote more inclusive and scalable approaches to global microbiome science.

2. Methods

2.1. Ethics approval and study setting

This study received ethical approval from the National Health Research Ethics Committee of Nigeria (NHREC/2022/3448) and the Oxford Tropical Research Ethics Committee (OxTREC/REF: 36–22). All neonates were enrolled following informed consent from a parent or guardian.

Stool samples were collected from five neonates within their first week of life at the maternity unit of Murtala Muhammed Specialist Hospital (MMSH), Kano, Nigeria; a healthcare facility with limited laboratory infrastructure. Samples were stored under varying conditions prior to microbial DNA extraction, reflecting many of the constraints associated with field work (Fig. 1).

2.2. Faecal sample collection and storage

Fresh stool samples were collected in sterile containers, kept on ice, and transported to a makeshift laboratory built from a shipping container located on the hospital premises (Supplementary Figures 1, 2). Each sample was homogenised and aliquoted into six storage conditions: (1) No preservative (processed same day), (2) DNA/RNA Shield (Zymo Research) stored for 24 h at room temperature (RT), (3) DNA/RNA Shield (Zymo Research) stored for 24 h at 4°C, (4) DNA/RNA Shield (Zymo Research) stored for 24 h at –80°C, (5) DNA/RNA Shield stored

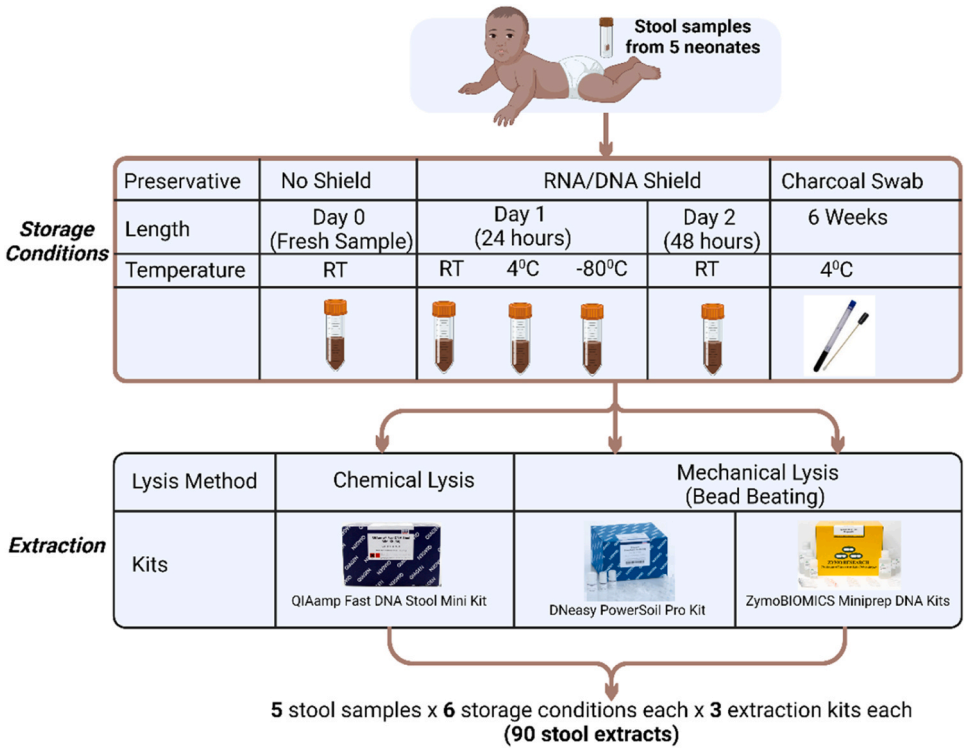

Fig. 1. Overview of experimental design for evaluating DNA extraction protocols from neonatal stool samples under varying storage conditions. Stool samples were collected from five neonates and stored under six different conditions, including fresh (no preservative), RNA/DNA Shield at varying temperatures and durations (RT, 4°C, –80°C), and charcoal swab at 4°C for 6 weeks. DNA was extracted using three commercial kits employing chemical lysis (QIAamp Fast DNA Stool Mini Kit) or mechanical lysis (DNeasy PowerSoil Pro Kit and ZymoBIOMICS Miniprep DNA Kit).

for 48 h at room temperature (RT), (6) Charcoal swabs stored for 6 weeks at 4°C.

2.3. DNA extraction kit selection and summary

Three commercial microbial DNA extraction kits were compared based on DNA yield, and microbial diversity and composition. The commercial extraction kits: QIAamp Fast DNA Stool Mini Kit (Qiagen), DNeasy PowerSoil Pro Kit (Qiagen, previously MoBio) and ZymoBIOMICS DNA Miniprep Kit (Zymo Research) designated as **protocol SM, protocol PS** and **protocol ZB** respectively, were chosen based on frequency of use in existing gut microbiome studies suitability for onsite extraction without need for third-party consumables or cold chain transport, and cost (Table 1).

Each of the five stool samples was subjected to all six storage conditions and extracted using each of the three kits, yielding a total of 90 DNA extracts. To reduce inter-operator variability, all extractions were performed by a single researcher. Samples preserved in DNA/RNA Shield were processed in suspension (200 µL), while 220–250 mg was used for solid stool extractions.

Table 1
Summary information for each DNA extraction protocol used.

DNA Extraction Methods	SM	PS	ZB
Mechanical lysis	Vortex	Bead-beating	Bead-beating
Chemical lysis	Yes	Yes	Yes
Thermal lysis	95C for 10 mins	65C for 10 mins (optional)	No
DNA purification method	Spin column-based	Spin column-based	Spin column-based
Cost of kit (50 Preparations)	~£ 290	~£ 374	~£ 417
Processing time per 12 samples	~1.5–2 h	~1.5–2 h	~2 – 2.5 h
Equipment needed	Micro-centrifuge Heat block	Micro-centrifuge Heat block (optional)	Micro-centrifuge

The QIAamp Fast DNA Stool Mini Kit (SM), DNeasy PowerSoil Pro Kit (PS), and ZymoBIOMICS DNA Miniprep Kit (ZB) differ in mechanical and thermal lysis approaches, processing time, and cost. All kits use spin column-based purification and chemical lysis. PS and ZB incorporate bead-beating, and ZB has the longest processing time

2.4. DNA extraction protocol comparison

Protocol SM (QIAamp Fast DNA Stool Mini Kit) - Stool was homogenised in InhibitEX buffer, followed by Proteinase K and Buffer AL incubation at 70°C for 10 min. Ethanol was added and lysates loaded onto spin columns, washed with AW1 and AW2, and eluted in 200 µL Buffer ATE.

Protocol PS (DNeasy PowerSoil Pro Kit) - Stool was added to bead tubes with Solution CD1, bead-beaten using a Vortex-Genie 2 for 10 min, incubated at 65°C for 10 min, and centrifuged. The supernatant was treated with Solutions CD2 and CD3, and DNA was purified using MB spin columns and washed with EA and C5 buffers. DNA was eluted in 100 µL of Solution C6.

Protocol ZB (ZymoBIOMICS DNA Miniprep Kit) - Stool was combined with lysis buffer in bead tubes and bead-beaten. After centrifugation, lysate was filtered through a Zymo-Spin III-F filter. The flow-through was bound to silica columns, washed, and eluted in 100 µL water, followed by further purification using the Zymo-Spin HRC filter.

2.5. DNA quantification

DNA concentration was measured using a Qubit 4 fluorometer (Life Technologies, USA) with dsDNA High Sensitivity and Broad Range Assay

Kits. Yield was calculated by multiplying DNA concentration by elution volume. Extracts were then shipped to the UK for metagenomics sequencing due to current capacity restraints in the makeshift laboratory to perform the sequencing *in situ*.

2.6. Library preparation and sequencing

SM kit produced negligible yields across all conditions and therefore did not provide enough material for further analysis. On the contrary, DNA extracts from the two adequate-performing kits (PS and ZB) from Day 0 and Week 6 samples was selected for metagenomic sequencing. Solid-phase reversible immobilisation (SPRI) bead purification was used to concentrate DNA prior to library prep. DNA was eluted in 30–70 μL depending on the initial yield, aiming for a final concentration of 15–20 ng/μL. Sequencing libraries were prepared per manufacturer's instructions. Libraries were prepared using the ONT SQK-RBK114.96 Rapid Barcoding Kit and sequenced on a GridION platform with R10.4.1 flow cells (Oxford Nanopore Technologies, UK). The Rapid Barcoding Kit was selected instead of the Ligation sequencing kit due to a significantly lower cost of sequencing per sample associated with reduced need for third-party consumables, as well as ease of use. The Rapid Barcoding Kit also allows for a lower input concentration per DNA sample for sequencing, which is more appropriate for low-biomass samples such as meconium.

2.7. Bioinformatics and taxonomic profiling

Base calling, demultiplexing, and adapter trimming were performed using Guppy v6.4.6. Quality control was conducted with SeqKit v2.6.1 (Shen et al., 2016) and low-quality reads (Q <10, length <100 bp) were removed using NanoQ v0.10.0 (Steinig and Coin, 2022).

Filtered reads were taxonomically classified with Kraken2 (v2.1.3) (Wood et al., 2019) using the PlusPFP database (January 2024 release 1/12/2024) (Kraken 2, 2025). Abundances were re-estimated with Bracken (Lu et al., 2022) and exported as BIOM-format file using Kraken-BIOM (v1.0.1) (Kraken-BIOM,).

2.8. Statistical and diversity analysis

All analyses were performed in R (v4.4.1). Taxonomic BIOM file was imported into phyloseq (v1.42.0) and visualised with ggplot2 (v3.5.1). Samples were rarefied to equal depth using the rarefy_even_depth function and a linear model was used to compare diversity values pre- and post-rarefaction. Diversity metrics (Observed, Shannon, Simpson, Inverse Simpson) were computed using the alpha function of microbiome (v1.20.0). Beta diversity was calculated using Bray–Curtis dissimilarity and visualised via Principal Coordinates Analysis (PCoA). Group differences (by kit and time point) were tested using multivariate analysis of variance adonis2 (vegan v2.6.6.1) on Bray-Curtis distance matrix and group variance (dispersion) assessed with betadisper that handles distances to group centroids followed by permutest for significance. Raw sequencing reads have been deposited in NCBI under BioProject accession PRJNA1260441.

Differences in DNA yield across extraction kits were assessed using the Friedman test with Wilcoxon signed-rank post-hoc tests and Bonferroni correction. The impact of storage conditions on yield was evaluated separately for PS and ZB using Kruskal–Wallis tests, followed by Bonferroni-adjusted Mann–Whitney U tests (Choo et al., 2015). Wilcoxon signed-rank tests were also used to compare N50 values and read quality metrics between kits. Non-parametric tests were selected due to small sample size and non-normal data distribution, which are common in microbiome studies (Zhang et al., 2024). To further comprehend the variation in species abundance and prevalence between Kit and Time variables, differential abundance analysis (DAA) was performed using statistical models of DESeq2 (v1.46.0), ALDEx2 (v1.38.0) and MaAsLin3 (v0.99.16). Mixed models were to analyse the impact of two independent factors Kit and Time on compositionality of the microbiome data reducing the effects from repeated measures of individuals. A model design of Abundance ~ Kit * Day + (1|Sample) was implemented to detect changes in composition of species considering Kit and Day as fixed variables along with an interaction term between Kit and Day, while accounting for random effects from individuals. The model parameters for the three approaches were setup on abundance data with TSS normalisation, log transformation, default small_random_effects, and significant taxa were recorded if fdr adjusted p-value was less than 0.05.

3. Results

3.1. DNA extraction yield across protocols

DNA yield varied substantially across the three extraction protocols: SM, PS, and ZB. Under same-day extraction without preservation (Day 0), PS and ZB produced significantly higher yields (median ~2000 ng), while SM yielded negligible amounts (Fig. 2). This trend persisted across all time points with ZB consistently outperforming both kits; PS showed moderate performance, and SM frequently fell below detection limits.

To statistically evaluate these differences while accounting for repeated measures, a Friedman test revealed a highly significant variation in DNA yield among the three kits ($\chi^2 = 41.18$, $p = 1.14 \times 10^{-9}$). Post-hoc Wilcoxon signed-rank tests with Bonferroni correction confirmed that both PS and ZB yielded significantly more DNA than SM (adjusted $p < 0.001$), with no significant difference between PS and ZB (adjusted $p = 0.69$).

Further paired comparisons across storage conditions showed no significant differences in DNA yield between PS and ZB at any time point; Day 0 ($p = 0.465$), Day 1 ($p = 0.524$), Day 2 ($p = 0.625$), or Week 6 ($p = 0.125$). These results indicate that both PS and ZB perform comparably and reliably across varied storage conditions, supporting their use in neonatal stool microbiome studies, particularly in field or low-resource settings.

3.2. Impact of storage conditions on DNA yield

Storage conditions had a pronounced impact on DNA yield across all extraction protocols. Relative to fresh samples, DNA/RNA Shield-stored samples (24–72 h) at –80°C, 4°C, and room temperature showed a reduction in DNA yield across all kits (Fig. 2). Nevertheless, protocol ZB maintained relatively high recovery even after 1–2 days of storage, particularly under 4 °C and room temperature conditions, whereas PS displayed a moderate decline. SM consistently failed to recover DNA under any storage condition. After six weeks of storage at 4 °C using charcoal swabs, both ZB and PS protocols were still capable of extracting measurable amounts of DNA, although yields were lower, and variability increased. SM protocol was ineffective under all delayed extraction conditions. To assess the significance of these impact, Kruskal–Wallis tests were conducted separately for the PS and ZB protocols. Both showed a statistically significant effect of storage time on the yield (PS: $\chi^2 = 14.61$, $p = 0.0022$; ZB: $\chi^2 = 13.45$, $p = 0.0038$). Post-hoc pairwise comparisons using Bonferroni-corrected Mann–Whitney U tests revealed that for both kits, DNA yield at Day 0 was significantly higher than at Day 1 (*adjusted* $p = 0.024$), indicating that substantial degradation or loss of extractable DNA occurs within the first 24 h of storage. For PS, DNA yield at Day 0 was also significantly higher than at Week 6 (*adjusted* $p = 0.036$), while no other comparisons reached statistical significance for either kit. These findings underscore the importance of same-day processing for maximum DNA recovery, particularly in field or low-resource settings where sample storage and preservation may be suboptimal.

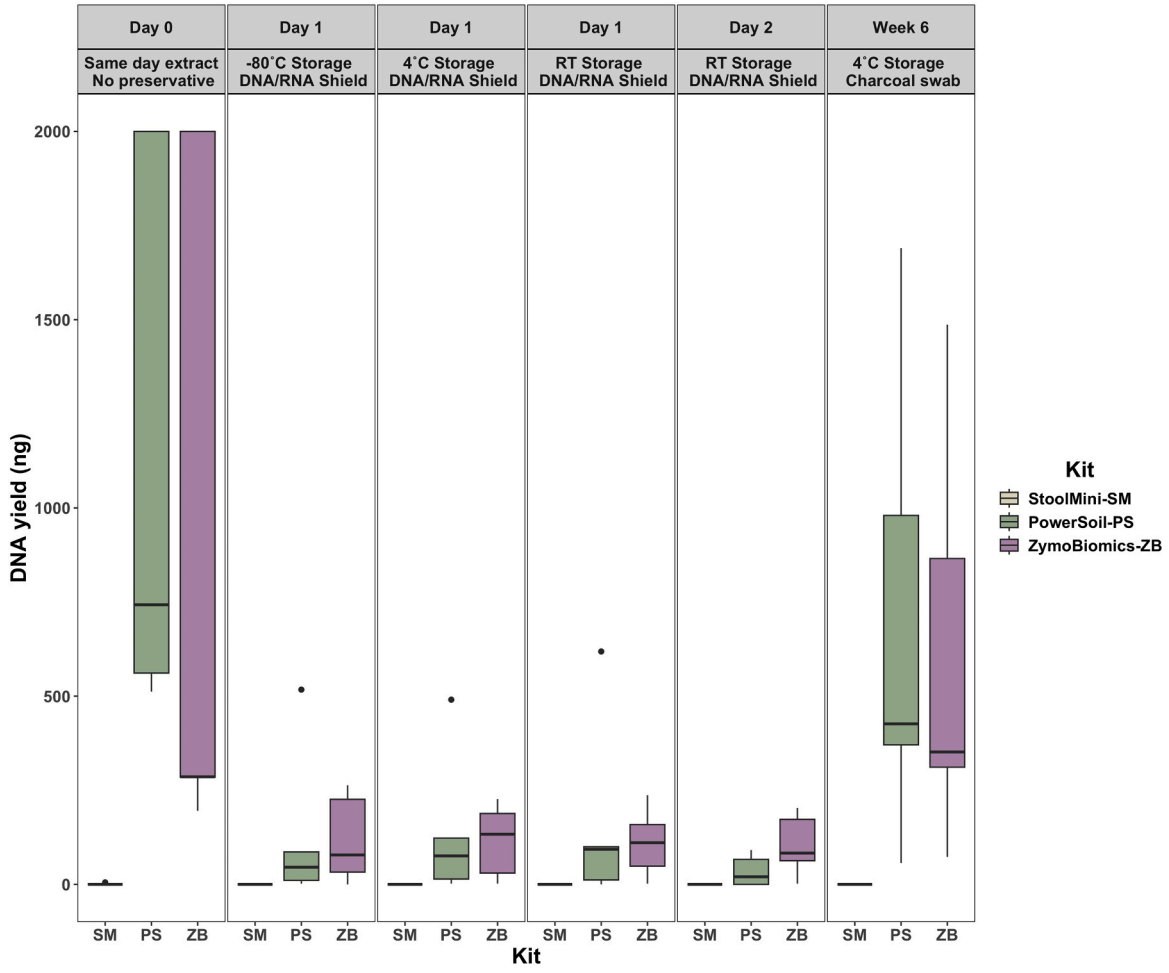

Fig. 2. Comparison of DNA yield across different storage conditions and extraction kits. DNA yield (ng) measured from neonatal stool samples stored under six conditions and extracted using the three protocols: SM, PS and Z). Each bar represents results from five neonates. Fresh, unpreserved samples (Day 0) produced the highest DNA yield, particularly with PS and ZB kits. Across preserved and delayed extraction conditions, ZB consistently outperformed SM, while PS showed moderate performance. SM kit yielded the lowest DNA yields across all conditions and therefore did not provide enough material for further analysis.

3.3. Comparing protocols for read length and distribution

Shotgun metagenomic sequencing was performed on samples extracted using PS and ZB protocols at Day 0 and Week 6. SM protocol and other time points were excluded from sequencing due to insufficient DNA yield. Sequencing produced 2,736,799 and 3,420,942 raw reads for PS and ZB samples, respectively. After quality filtering, a mean of 300,616 reads per sample was retained (SD ± 195,894), with higher retained reads observed for ZB (334,794) than PS (266,753). Read length distributions revealed that ZB yielded a higher proportion of short reads (<1000 bp), whereas PS produced more long reads (>10,000 bp), with approximately 140,000 reads exceeding this threshold (Supplementary Figure 3).

To assess sequencing read continuity, read N50 lengths were compared across the two extraction kits and storage times: D0 and W6 (Fig. 3). PS generally showed higher N50 values (4936 bp; range: 1910–6903) than the ZB protocol (3924 bp; range: 2381–5248). At Day 0, PS showed higher N50 values (Wilcoxon signed-rank test, $p = 0.313$), with a moderate effect size (rank-biserial correlation = 0.60), suggesting a trend toward better assembly contiguity. By Week 6, the difference was further reduced ($p = 0.625$; effect size = 0.33), indicating that storage may lessen the performance gap between kits.

Comparing read quality metrics between the two kits at Day 0 and Week 6 to assess their impact on sequencing performance. Average read quality as well as the proportion of bases exceeding Q20 and Q30 thresholds, were evaluated. At both time points, no statistically significant differences were observed between kits for any of the three quality metrics. Average read quality comparisons yielded $p = 0.5625$ at Day 0 and $p = 0.4375$ at Week 6. Similarly, the proportion of Q20 bases showed $p = 0.5625$ (Day 0) and $p = 0.6250$ (Week 6), while Q30 base proportions resulted in $p = 0.8125$ (Day 0) and $p = 0.6250$ (Week 6).

3.4. Bacterial composition across extraction kits and storage times

Taxonomic profiling identified over 400 bacterial species, with the 20 most abundant species shown in Fig. 4. Both PS and ZB protocols yielded broadly similar community profiles when directly comparing the same time point. Across all samples, ten species were consistently identified in ≥ 70 % of samples: *Escherichia coli*, *Streptococcus salivarius*, *Klebsiella pneumoniae, Enterococcus faecium, Bifidobacterium breve, Leclercia spp., Enterococcus faecalis, Bacteroides fragilis, Phocaeicola vulgatus, and Staphylococcus epidermidis.* Among these, *E. coli* was the most abundant, comprising > 25 % relative abundance in three samples.

Some taxa, such as *K. pneumoniae* and *Streptococcus spp.,* showed increased abundance after six weeks of storage, while others like *B. breve* and *S. hominis* declined. These shifts may reflect community restructuring due to microbial competition or selective pressures imposed by long-term storage, particularly with charcoal swabs. Overall, taxonomic profiles were more consistent between kits than between storage time points, suggesting that storage conditions had a larger effect on bacterial

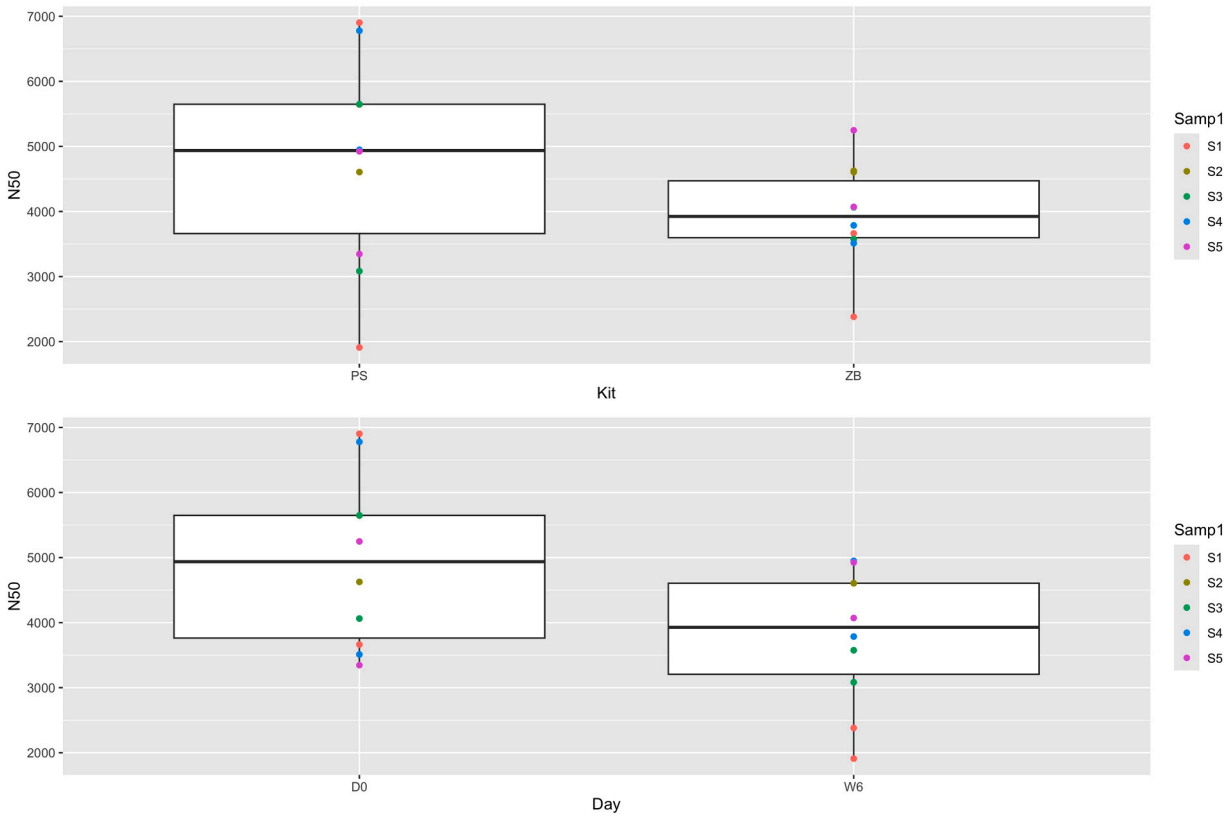

Fig. 3. Comparison of read-level N50 across DNA extraction kits and storage times. **(Top)** Boxplot comparing read level N50 values obtained using the PS and ZB protocols,combining data from both Day 0 and Week 6. Samples processed with the PS kit generally yielded higher N50 values, indicating more contiguous assemblies. **(Bottom)** Boxplot comparing N50 values between samples processed on Day 0 (fresh) and after 6 weeks of storage (Week 6). Reads from fresh samples had higher and more variable N50 values compared to those stored long-term, suggesting that delayed processing may reduce genome assembly quality. Coloured dots represent individual samples (S1–S5).

composition compared to extraction protocol. DAA analyses from the three approaches found varying number of significantly abundant species but most of these species are not shared owing to the differences in the model parametrisation, normalisation procedures and zero count handling. Most of the abundant taxa appear significant in response to storage time compared to Kit. Bacterial genus such as *Acinetobacter, Atlantibacter, Bifidobacterium, Citrobacter, Clostridioides, Enterobacter, Leclercia, Klebsiella, Pseudomonas, Staphylococcus, Streptococcus, Salmonella* were found to be significantly different among the three approaches (Fig. 6).

3.5. Microbial diversity metrics

Microbial diversity was assessed to evaluate how DNA extraction protocol and sample storage conditions influenced gut microbiome composition. Alpha diversity was assessed using Observed species, Shannon, Simpson, and Inverse Simpson indices (Fig. 5A). Despite differences in DNA yield and read characteristics, diversity measures were largely comparable between PS and ZB protocols. This suggests that both methods preserved within-sample microbial complexity equally well, even though their efficiency in DNA recovery and read length varied. Fig. 6

When diversity was compared across storage conditions, Day 0 samples generally exhibited slightly higher richness (Observed species) and diversity (Shannon index, $p = 0.0479$) than those stored for six weeks in charcoal swabs. In particular, a reduction in rare taxa was inferred from lower values in richness-based metric (Observed, $p = 0.2382$) in the Week 6 samples. This may reflect loss or degradation of low-abundance organisms over prolonged storage, or overgrowth of hardier taxa that reduced overall evenness (Simpson index, $p = 0.0124$; Inverse Simpson index, $p = 0.0613$). These trends were consistent across both extraction kits, indicating that storage duration and/or storage medium, rather than extraction method, had a greater effect on within-sample diversity.

Beta diversity was evaluated using Bray–Curtis dissimilarity and reduced to principal coordinates via PCoA (Fig. 5B–C). This analysis revealed that bacterial community composition clustered more strongly by storage time point (adonis2 $p = 0.0002$) compared to extraction protocol (adonis2 $p = 0.8250$), further suggesting that time-related shifts in community structure were more pronounced than kit-related biases (average distance to PS centroid = 0.5434, average distance to ZB centroid = 0.5646). Specifically, Week 6 samples stored in charcoal swabs showed tighter clustering, suggesting reduced inter-individual variability or a converging microbial profile after extended storage (average distance to Week 6 centroid = 0.4790). In contrast, Day 0 samples were more dispersed, consistent with greater sample-specific variation in freshly collected stools (average distance to Day 0 centroid = 0.5717).

4. Discussion

An ideal DNA extraction protocol for long-read metagenomic sequencing should lyse a broad range of microbial cells, avoid contamination, introduce minimal bias in community structure, and yield high concentrations of intact DNA suitable for downstream applications. In this study, we evaluated three commercial extraction kits—QIAamp Fast DNA Stool Mini Kit (SM), DNeasy PowerSoil Pro Kit (PS), and ZymoBIOMICS DNA Miniprep Kit (ZB) for their performance in extracting DNA from neonatal stool samples under real-world conditions in a resource-limited setting.

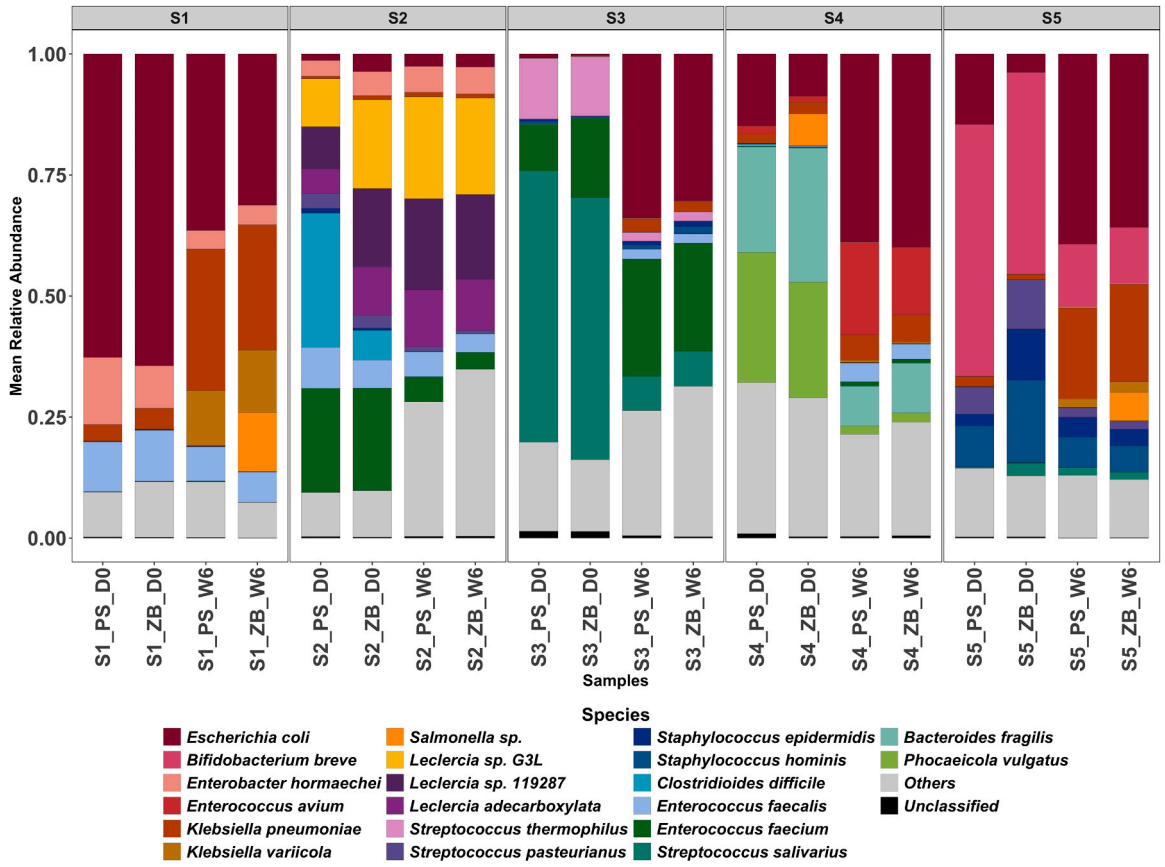

Fig. 4. Species-level taxonomic composition of neonatal stool samples extracted using protocols PS and ZB across two time points. Bar plots show the mean relative abundance of bacterial species detected in five neonates (S1–S5) using protocol PS and ZB at Day 0 (D0) and after 6 weeks of storage at 4 °C (W6).

Our results demonstrate that both the choice of DNA extraction kit and sample storage conditions have a significant impact on DNA yield from neonatal stool. Among the three kits tested, the bead-beating-based kits; PS and ZB consistently outperformed chemically based SM kit, which yielded negligible DNA even under optimal conditions. This finding is consistent with previous studies highlighting the importance of bead-beating for efficient lysis and accurate microbial profiling from stool samples (Fernández-Pato et al., 2024; Zhang et al., 2021; Yang et al., 2020). Results herein therefore emphasise that access to and investment in convenient and efficient bead-beating related equipment and consumables must be a priority to enable metagenomics research in resource-limited settings.

DNA yield was highest when samples were extracted fresh, on the same day of collection, without preservatives. Both PS and ZB kits produced median yields near 2000 ng under these conditions, while SM yielded close to zero. Across all subsequent storage durations, 1 day, 2 days, and 6 weeks DNA yield declined substantially, particularly within the first 24 h. This trend was statistically significant for both PS and ZB, indicating that early degradation or reduced recovery efficiency occurs quickly after collection. DNA yield remained stable between Day 1 and later time points, suggesting that most yield loss happens rapidly and plateaus thereafter. Notably, while PS and ZB yields declined over time, there was no significant difference between them at any storage point, reinforcing their comparable performance. These findings underscore the critical importance of immediate DNA extraction or effective sample stabilisation, particularly in low-resource settings where cold chain infrastructure is limited. Delays in processing, even by one day, can markedly reduce DNA recovery, potentially compromising downstream microbiome analysis and reducing study power.

Importantly, prior work has shown that while microbial richness and alpha diversity remained relatively stable across storage durations, the most pronounced shifts in community composition occurred within the first 24 h of storage at 4 °C (Holzhausen et al., 2021). This early compositional change, measured by Bray–Curtis dissimilarity, underscores that even short delays can alter microbial structure, further emphasising the need for immediate stabilisation or extraction in microbiome studies involving low-biomass samples. Selecting robust extraction methods and minimising storage time are therefore essential considerations for field-based microbiome studies, especially in neonatal and low-biomass contexts (Goodrich et al., 2014).

Beyond yield, sequence read quality and microbial community structure recovered by PS and ZB was largely comparable. For instance, no statistically significant differences were observed in average read quality or in the proportion of high-quality reads (Q20 +, Q30 +) across kits or time points ($p > 0.4$ for all comparisons). This finding aligns with previous studies reporting that while DNA extraction protocols can influence downstream metagenomic assembly and taxonomic profiling, they have limited impact on base-calling quality when high-integrity DNA is available for sequencing (Maghini et al., 2021). The consistency in read quality across kits is encouraging for resource-limited settings, where kit availability, reagent access, and cold chain limitations can complicate sample processing. In such contexts, the ability to maintain sequencing accuracy regardless of extraction method increases flexibility in field protocols, particularly when long-read sequencing platforms like Oxford Nanopore are used. However, as demonstrated in our study, extraction methods still play a critical role in DNA yield and assembly contiguity, both of which are essential for genome-resolved metagenomics, highlighting the need for careful selection of protocols based on both practical constraints and downstream goals.

Furthermore, taxonomic profiles at both Day 0 and Week 6 were similar between the two protocols, with shared detection of dominant species such as *E. coli*, *K. pneumoniae*, and *B. breve*. While slight

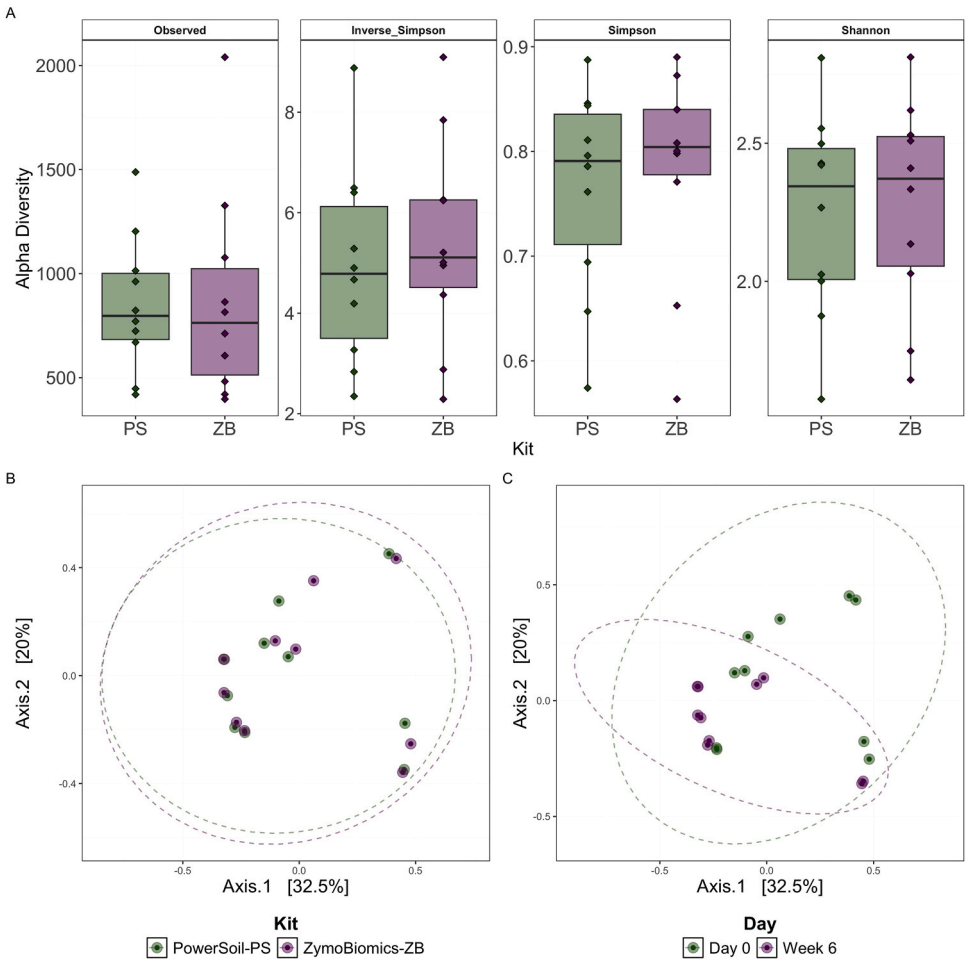

Fig. 5. Alpha and beta diversity of neonatal stool microbiomes by extraction kit and storage time. (A) Boxplots of alpha diversity indices Observed ($p = 0.3863$), Inverse Simpson ($p = 0.4889$), Simpson ($p = 0.6430$), and Shannon ($p = 0.9487$) comparing DNA extracted using protocols PS and ZB. (B) Principal Coordinates Analysis (PCoA) based on Bray-Curtis dissimilarity showing minimal clustering by extraction kit (PS vs. ZB, $p = 0.548$). (C) PCoA plot showing separation of samples by storage duration (Day 0 vs. Week 6, p-value = 0.008).

differences in relative abundance were observed, particularly in long-term stored samples, both kits preserved community diversity to a similar extent. Importantly, both protocols also showed comparable performance in alpha diversity metrics (Shannon, Simpson, Inverse Simpson), while beta diversity analysis revealed that storage duration and time to extraction from collection, not extraction protocol, had the greater influence on community composition. These findings align with those of Fernández-Pato *et al.* (2024) (Fernández-Pato et al., 2024), who reported that while DNA extraction methods can introduce compositional biases, the inclusion of bead-beating steps improves microbial representation, especially for gram-positive taxa. However, they also emphasised that standardised protocols are critical, as methodological variability can confound phenotype-microbiome associations more than biological factors alone.

The observed decline in DNA yield and diversity with delayed processing has practical implications for field-based microbiome studies. In such settings, immediate DNA extraction is often unfeasible due to unpredictable neonatal bowel habits, limited laboratory infrastructure, and delays in sample transport. Our results demonstrate that DNA can still be reliably extracted after up to six weeks of storage at 4 °C in charcoal swabs, particularly when using the PowerSoil and ZymoBIOMICS kits.

These findings align with those of Nel Van Zyl et al. (2020), who reported minimal bias in microbial profiles when stool samples were stored for 48 h at room temperature or in cool boxes, compared to immediate freezing at –80 °C. This supports the feasibility of using alternative, non-freezing storage methods in resource-limited environments.

Notably, samples preserved in DNA/RNA Shield yielded unexpectedly low DNA concentrations in our study, suggesting potential incompatibility between some commercial preservatives and specific extraction kits. While DNA/RNA Shield is marketed for long-term nucleic acid stabilisation, our findings, and those of others, (Choo et al., 2015; Tap et al., 2019) indicate that its performance is highly dependent on both the extraction kit and storage conditions. This underscores the need for clearer guidance regarding preservative–kit compatibility in microbiome workflows. In our setting, charcoal swabs offered a more reliable alternative to preservative-based storage for medium-term field applications, albeit assessing estimates of relative abundance in addition to microbial composition may be limited.

A key operational consideration in LMICs is the feasibility of implementing microbiome workflows at scale. Although PS-extracted samples tended to yield longer reads and slightly higher median DNA concentrations, the variability across samples and limited sample size precluded statistical significance. These results suggest that both kits perform comparably in terms of DNA yield under the tested conditions. Nevertheless, while both PS and ZB kits performed well in terms of DNA yield and diversity, PS offered several practical advantages. It required a shorter processing time (~1.5 h for 12 samples), used fewer components (reducing the risk of user error), and achieved longer read lengths, critical for high-quality metagenomic assembly and increased resolution of repetitive regions. The ZB protocol, although effective, required more

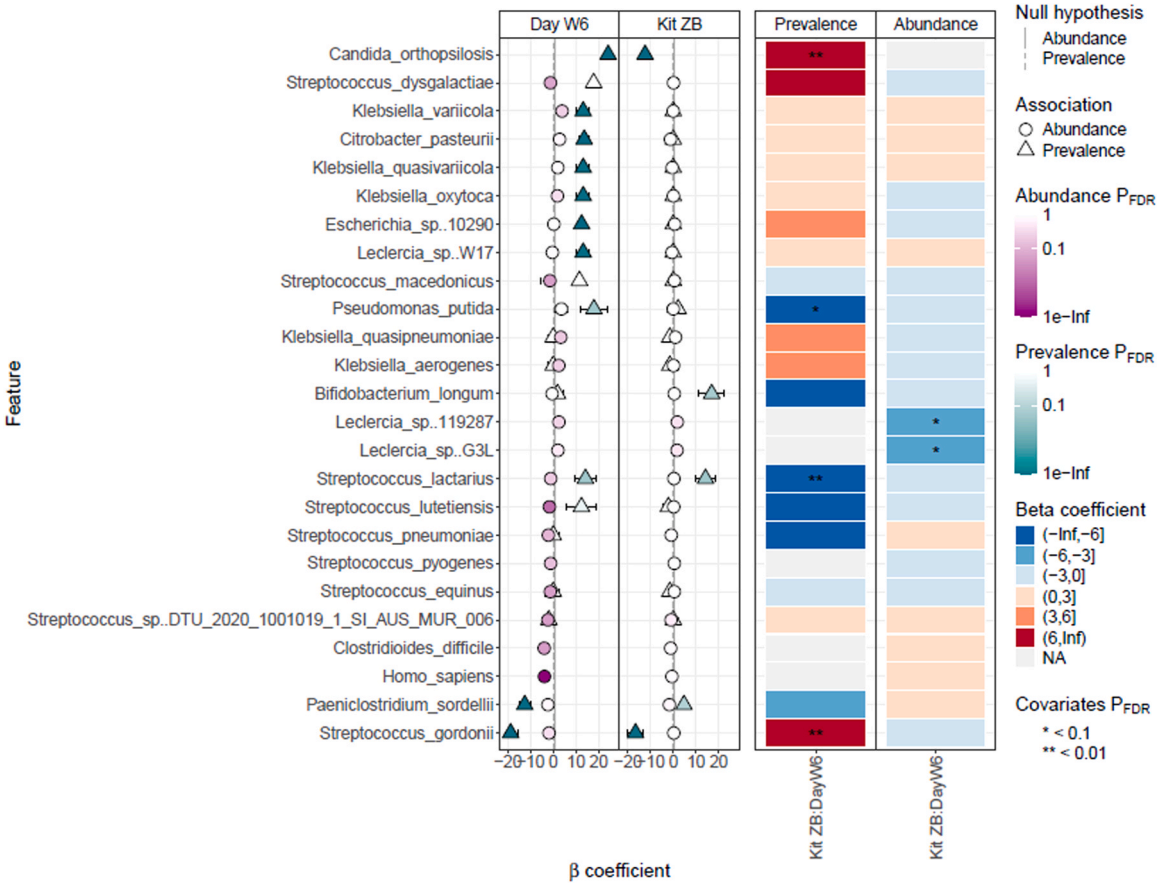

Fig. 6. Differentially abundant species identified by MaAsLin3 in response to DNA extraction kit and storage time. The left panel shows coefficient plots of significant associations with either kit (PS vs ZB) or storage time (Day 0 vs Week 6). Point shapes represent the type of association (abundance or prevalence), and colours indicate the false discovery rate (FDR)-adjusted p-values (q-values). The right panel displays a heatmap of interaction effects between kit and storage time on both abundance and prevalence. Colours represent the magnitude and direction of the β coefficients, scaled from minimum to maximum. The statistical model used was: Abundance ~ Kit * Day + (1|Sample), with PS as the reference level for Kit, D0 for Day, and S1 for Sample. Asterisks indicate significance levels ($q < 0.1$, $q < 0.01$)

handling steps and a longer bead-beating time (~40 min), which could increase turnaround time and introduce logistical bottlenecks in high-throughput settings. Additionally, kit design such as the use of similar-looking columns in ZB may increase the risk of workflow errors in constrained laboratory environments.

This study had several limitations. First, our small sample size limited statistical power and precluded robust testing of subtle differences in taxonomic profiles. Negative and positive controls (Zymo) were not sequenced in this study as the focus was on applicability of DNA extraction in a resource-limited setting but will be in our prospective cohort observational study designed following this feasibility study. Likewise, metagenomics sequencing was performed in the UK due to insufficient laboratory infrastructure, however, with continued developments in ONT approaches we envisage future research to adopt an end-to-end sample collection, extraction and sequence generation locally. Additionally, the lack of metadata (e.g., mode of delivery, feeding type) prevented analysis of host-associated factors that may influence neonatal microbiota composition. While not a focus of this study, future investigations should incorporate such metadata to better contextualise microbiome variation and validate findings across larger cohorts.

4.1. *Conclusion and LMIC implications*

In conclusion, our results show that mechanical lysis-based extraction protocols (PS and ZB) are better suited for neonatal stool samples, producing higher DNA yields than chemical lysis-based method. While PS and ZB kits are compatible with long-read sequencing producing similar taxonomic composition, PowerSoil (PS) provides a practical edge for LMIC implementation due to faster processing, fewer handling steps, and superior read length generation. Moreover, our findings reinforce the importance of timely sample processing, preferably within 24 h or using reliable preservation strategies such as charcoal swabs to minimise community drift.

To advance microbiome research in resource-limited settings, it is critical to not only select scientifically robust methods but also prioritise operational feasibility. DNA extraction protocols should be chosen not just for performance, but for their adaptability to LMIC contexts, where electricity, cold chain storage, and access to consumables may be inconsistent. Manufacturers and funders alike have a role in supporting sustainable, scalable and accessible solutions tailored to these environments. By grounding microbiome research in context-sensitive methods, we can improve data quality, promote reproducibility, and ensure broader global participation in microbiome science.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This research was funded by the Ineos Oxford Institute, University of Oxford. Chinenye Akpulu'sDPhil is supported by the Interdisciplinary Bioscience Doctoral Training partnership and Ineos Oxford Institute both at the University of Oxford.

CRediT authorship contribution statement

Aisha Mukkadas: Writing – review & editing, Project administration, Methodology, Investigation. **Khalid Rashida:** Writing – review & editing, Project administration, Methodology, Investigation. **Portal Edward:** Writing – review & editing, Validation, Methodology. **Kirsty Sands:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization. **Lankapalli Aditya:** Writing – review & editing, Visualization, Validation, Software, Resources, Formal analysis, Data curation. **Timothy Walsh:** Writing – review & editing, Validation, Supervision, Investigation, Funding acquisition. **Chinenye Akpulu:** Writing – review & editing, Writing – original draft, Validation, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Kate Cook:** Writing – review & editing, Validation, Methodology. **Rida Toufiq:** Writing – review & editing, Validation, Methodology. **Aliyu Aminu:** Writing – review & editing, Visualization, Investigation. **Safiya Gambo:** Writing – review & editing, Validation, Investigation. **Kathryn Thomson:** Writing – review & editing, Validation, Supervision, Investigation. **Kenneth Iregbu:** Writing – review & editing, Validation, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors thank the department of molecular microbiology laboratory at the Aminu Kano Teaching Hospital for providing ultra-low temperature freezer storage space.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.microb.2025.100398.

Data availability

I have shared my data within the manuscript and supplementary materials. Raw sequencing reads have been deposited in NCBI under BioProject accession PRJNA1260441

References

Alsharairi, N.A., 2023. Therapeutic potential of gut microbiota and its metabolite short-chain fatty acids in neonatal necrotizing enterocolitis. Life 13, 561. https://doi.org/10.3390/life13020561.

Ayeni, K.I., Berry, D., Ezekiel, C.N., Warth, B., 2024. Enhancing microbiome research in sub-Saharan Africa. Trends Microbiol. 32, 111–115. https://doi.org/10.1016/j.tim.2023.11.003.

Baumann-Dudenhoeffer, A.M., D'Souza, A.W., Tarr, P.I., Warner, B.B., Dantas, G., 2018. Infant diet and maternal gestational weight gain predict early metabolic maturation of gut microbiomes. Nat. Med. 24, 1822–1829. https://doi.org/10.1038/s41591-018-0216-2.

Bäumler, A.J., Sperandio, V., 2016. Interactions between the microbiota and pathogenic bacteria in the gut. Nature 535, 85–93. https://doi.org/10.1038/nature18849.

Boostrom, I., Portal, E.A.R., Spiller, O.B., Walsh, T.R., Sands, K., 2022. Comparing long-read assemblers to explore the potential of a sustainable low-cost, low-infrastructure approach to sequence antimicrobial resistant bacteria with oxford nanopore sequencing. Front. Microbiol. 13. https://doi.org/10.3389/fmicb.2022.796465.

Choo, J.M., Leong, L.E., Rogers, G.B., 2015. Sample storage conditions significantly influence faecal microbiome profiles. Sci. Rep. 5, 16350. https://doi.org/10.1038/srep16350.

Costea, P.I., Zeller, G., Sunagawa, S., Pelletier, E., Alberti, A., Levenez, F., Tramontano, M., Driessen, M., Hercog, R., Jung, F.-E., Kultima, J.R., Hayward, M.R., Coelho, L.P., Allen-Vercoe, E., Bertrand, L., Blaut, M., Brown, J.R.M., Carton, T., Cools-Portier, S., Daigneault, M., Derrien, M., Druesne, A., de Vos, W.M., Finlay, B. B., Flint, H.J., Guarner, F., Hattori, M., Heilig, H., Luna, R.A., van Hylckama Vlieg, J., Junick, J., Klymiuk, I., Langella, P., Le Chatelier, E., Mai, V., Manichanh, C., Martin, J.C., Mery, C., Morita, H., O'Toole, P.W., Orvain, C., Patil, K. R., Penders, J., Persson, S., Pons, N., Popova, M., Salonen, A., Saulnier, D., Scott, K. P., Singh, B., Slezak, K., Veiga, P., Versalovic, J., Zhao, L., Zoetendal, E.G., Ehrlich, S. D., Dore, J., Bork, P., 2017. Towards standards for human fecal sample processing in metagenomic studies. Nat. Biotechnol. 35, 1069–1076. https://doi.org/10.1038/nbt.3960.

Dogra, S., Sakwinska, O., Soh, S.-E., Ngom-Bru, C., Brück, W.M., Berger, B., Brüssow, H., Lee, Y.S., Yap, F., Chong, Y.-S., Godfrey, K.M., Holbrook, J.D., 2015. Dynamics of infant gut microbiota are influenced by delivery mode and gestational duration and are associated with subsequent adiposity. MBio 6. https://doi.org/10.1128/mBio.02419-14.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., Knight, R., 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proc. Natl. Acad. Sci. USA 107, 11971–11975. https://doi.org/10.1073/pnas.1002601107.

Dos Santos, S.J., Pakzad, Z., Elwood, C.N., Albert, A.Y.K., Gantt, S., Manges, A.R., Dumonceaux, T.J., Maan, E.J., Hill, J.E., Money, D.M., 2021. Early neonatal meconium does not have a demonstrable microbiota determined through use of robust negative controls with *cpn* 60-based microbiome profiling. Microbiol. Spectr. 9. https://doi.org/10.1128/Spectrum.00067-21.

Fernández-Pato, A., Sinha, T., Gacesa, R., Andreu-Sánchez, S., Gois, M.F.B., Gelderloos-Arends, J., Jansen, D.B.H., Kruk, M., Jaeger, M., Joosten, L.A.B., Netea, M.G., Weersma, R.K., Wijmenga, C., Harmsen, H.J.M., Fu, J., Zhernakova, A., Kurilshikov, A., 2024. Choice of DNA extraction method affects stool microbiome recovery and subsequent phenotypic association analyses. Sci. Rep. 14, 3911. https://doi.org/10.1038/s41598-024-54353-w.

Gand, M., Bloemen, B., Vanneste, K., Roosens, N.H.C., De Keersmaecker, S.C.J., 2023. Comparison of 6 DNA extraction methods for isolation of high yield of high molecular weight DNA suitable for shotgun metagenomics Nanopore sequencing to detect bacteria. BMC Genom. 24, 438. https://doi.org/10.1186/s12864-023-09537-5.

Gomaa, E.Z., 2020. Human gut microbiota/microbiome in health and diseases: a review. Antonie Van. Leeuwenhoek 113, 2019–2040. https://doi.org/10.1007/s10482-020-01474-7.

Goodrich, J.K., Di Rienzi, S.C., Poole, A.C., Koren, O., Walters, W.A., Caporaso, J.G., Knight, R., Ley, R.E., 2014. Conducting a microbiome study. Cell 158, 250–262. https://doi.org/10.1016/j.cell.2014.06.037.

Holzhausen, E.A., Nikodemova, M., Deblois, C.L., Barnet, J.H., Peppard, P.E., Suen, G., Malecki, K.M., 2021. Assessing the impact of storage time on the stability of stool microbiota richness, diversity, and composition. Gut Pathog. 13, 75. https://doi.org/10.1186/s13099-021-00470-0.

Jurasz, H., Pawłowski, T., Perlejewski, K., 2021. Contamination issue in viral metagenomics: problems, solutions, and clinical perspectives. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.745076.

Kraken 2, KrakenUniq and Bracken indexes, 2025.

Kraken-BIOM, (n.d.).

Leong, K.S.W., McLay, J., Derraik, J.G.B., Gibb, S., Shackleton, N., Taylor, R.W., Glover, M., Audas, R., Taylor, B., Milne, B.J., Cutfield, W.S., 2020. Associations of prenatal and childhood antibiotic exposure with obesity at age 4 years. JAMA Netw. Open 3, e1919681. https://doi.org/10.1001/jamanetworkopen.2019.19681.

Lim, M.Y., Hong, S., Kim, B.-M., Ahn, Y., Kim, H.-J., Nam, Y.-D., 2020. Changes in microbiome and metabolomic profiles of fecal samples stored with stabilizing solution at room temperature: a pilot study. Sci. Rep. 10, 1789. https://doi.org/10.1038/s41598-020-58719-8.

Lu, J., Rincon, N., Wood, D.E., Breitwieser, F.P., Pockrandt, C., Langmead, B., Salzberg, S.L., Steinegger, M., 2022. Metagenome analysis using the Kraken software suite. Nat. Protoc. 17, 2815–2839. https://doi.org/10.1038/s41596-022-00738-y.

Maghini, D.G., Moss, E.L., Vance, S.E., Bhatt, A.S., 2021. Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome. Nat. Protoc. 16, 458–471. https://doi.org/10.1038/s41596-020-00424-x.

Makhalanyane, T.P., Bezuidt, O.K.I., Pierneef, R.E., Mizrachi, E., Zeze, A., Fossou, R.K., Kouadjo, C.G., Duodu, S., Chikere, C.B., Babalola, O.O., Klein, A., Keyster, M., du Plessis, M., Yorou, N.S., Hijri, M., Rossouw, T., Kamutando, C.N., Venter, S., Moleleki, L.N., Murrell, C., 2023. African microbiomes matter. Nat. Rev. Microbiol. 21, 479–481. https://doi.org/10.1038/s41579-023-00925-y.

Nel Van Zyl, K., Whitelaw, A.C., Newton-Foot, M., 2020. The effect of storage conditions on microbial communities in stool. PLOS One 15, e0227486. https://doi.org/10.1371/journal.pone.0227486.

Olin, A., Henckel, E., Chen, Y., Lakshmikanth, T., Pou, C., Mikes, J., Gustafsson, A., Bernhardsson, A.K., Zhang, C., Bohlin, K., Brodin, P., 2018. Stereotypic immune system development in newborn children. Cell 174, 1277–1292.e14. https://doi.org/10.1016/j.cell.2018.06.045.

Pehrsson, E.C., Tsukayama, P., Patel, S., Mejía-Bautista, M., Sosa-Soto, G., Navarrete, K. M., Calderon, M., Cabrera, L., Hoyos-Arango, W., Bertoli, M.T., Berg, D.E., Gilman, R.H., Dantas, G., 2016. Interconnected microbiomes and resistomes in low-income human habitats. Nature 533, 212–216. https://doi.org/10.1038/nature17672.

Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., Segata, N., 2017. Shotgun metagenomics, from sampling to analysis. Nat. Biotechnol. 35, 833–844. https://doi.org/10.1038/nbt.3935.

Rackaityte, E., Halkias, J., Fukui, E.M., Mendoza, V.F., Hayzelden, C., Crawford, E.D., Fujimura, K.E., Burt, T.D., Lynch, S.V., 2020. Viable bacterial colonization is highly limited in the human intestine in utero. Nat. Med. 26, 599–607. https://doi.org/10.1038/s41591-020-0761-3.

Reyman, M., van Houten, M.A., Watson, R.L., Chu, M.L.J.N., Arp, K., de Waal, W.J., Schiering, I., Plötz, F.B., Willems, R.J.L., van Schaik, W., Sanders, E.A.M.,

Bogaert, D., 2022. Effects of early-life antibiotics on the developing infant gut microbiome and resistome: a randomized trial. Nat. Commun. 13, 1–12. https://doi.org/10.1038/s41467-022-28525-z.

Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G., Gasbarrini, A., Mele, M., 2019. What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. Microorganisms 7, 14. https://doi.org/10.3390/microorganisms7010014.

Selway, C.A., Eisenhofer, R., Weyrich, L.S., 2020. Microbiome applications for pathology: challenges of low microbial biomass samples during diagnostic testing. J. Pathol. Clin. Res. 6, 97–106. https://doi.org/10.1002/cjp2.151.

Shen, W., Le, S., Li, Y., Hu, F., 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLOS One 11, e0163962. https://doi.org/10.1371/journal.pone.0163962.

Sommer, F., Bäckhed, F., 2013. The gut microbiota — masters of host development and physiology. Nat. Rev. Microbiol. 11, 227–238. https://doi.org/10.1038/nrmicro2974.

Song, S.J., Amir, A., Metcalf, J.L., Amato, K.R., Xu, Z.Z., Humphrey, G., Knight, R., 2016. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. MSystems 1. https://doi.org/10.1128/mSystems.00021-16.

Steinig, E., Coin, L., 2022. Nanoq: ultra-fast quality control for nanopore reads. J. Open Source Softw. 7, 2991. https://doi.org/10.21105/joss.02991.

Sudjai, A., Suriye, K., Kruasuwan, W., Jenjaroenpun, P., Foongladda, S., Wongsurawat, T., 2022. Long-read 16S rRNA amplicon and metagenomic data of swine feed-additive probiotics product. Microbiol. Resour. Announc. 11. https://doi.org/10.1128/mra.00397-22.

Tap, J., Cools-Portier, S., Pavan, S., Druesne, A., Öhman, L., Törnblom, H., Simren, M., Derrien, M., 2019. Effects of the long-term storage of human fecal microbiota samples collected in RNAlater. Sci. Rep. 9, 601. https://doi.org/10.1038/s41598-018-36953-5.

Thänert, R., Sawhney, S.S., Schwartz, D.J., Dantas, G., 2022. The resistance within: antibiotic disruption of the gut microbiome and resistome dynamics in infancy. Cell Host Microbe 30, 675–683. https://doi.org/10.1016/j.chom.2022.03.013.

Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L., Marsland, B.J., 2014. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. Nat. Med. 20, 159–166. https://doi.org/10.1038/nm.3444.

Wagner Mackenzie, B., Waite, D.W., Taylor, M.W., 2015. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. Front. Microbiol. 6. https://doi.org/10.3389/fmicb.2015.00130.

Walters, W.A., Xu, Z., Knight, R., 2014. Meta-analyses of human gut microbes associated with obesity and IBD. FEBS Lett. 588, 4223–4233. https://doi.org/10.1016/j.febslet.2014.09.039.

Wesolowska-Andersen, A., Bahl, M.I., Carvalho, V., Kristiansen, K., Sicheritz-Pontén, T., Gupta, R., Licht, T.R., 2014. Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. Microbiome 2, 19. https://doi.org/10.1186/2049-2618-2-19.

Weyrich, L.S., Farrer, A.G., Eisenhofer, R., Arriola, L.A., Young, J., Selway, C.A., Handsley-Davis, M., Adler, C.J., Breen, J., Cooper, A., 2019. Laboratory contamination over time during low-biomass sample analysis. Mol. Ecol. Resour. 19, 982–996. https://doi.org/10.1111/1755-0998.13011.

Wongsurawat, T., Nakagawa, M., Atiq, O., Coleman, H.N., Jenjaroenpun, P., Allred, J.I., Trammel, A., Puengrang, P., Ussery, D.W., Nookaew, I., 2019. An assessment of Oxford nanopore sequencing for human gut metagenome profiling: a pilot study of head and neck cancer patients. J. Microbiol Methods 166, 105739. https://doi.org/10.1016/j.mimet.2019.105739.

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 257. https://doi.org/10.1186/s13059-019-1891-0.

Xiao, L., Zhao, F., 2023. Microbial transmission, colonisation and succession: from pregnancy to infancy. Gut 72, 772–786. https://doi.org/10.1136/gutjnl-2022-328970.

Yang, F., Sun, J., Luo, H., Ren, H., Zhou, H., Lin, Y., Han, M., Chen, B., Liao, H., Brix, S., Li, J., Yang, H., Kristiansen, K., Zhong, H., 2020. Assessment of fecal DNA extraction protocols for metagenomic studies. Gigascience 9. https://doi.org/10.1093/gigascience/giaa071.

Zhang, B., Brock, M., Arana, C., Dende, C., van Oers, N.S., Hooper, L.V., Raj, P., 2021. Impact of bead-beating intensity on the genus- and species-level characterization of the gut microbiome using amplicon and complete 16S rRNA gene sequencing. Front. Cell Infect. Microbiol. 11. https://doi.org/10.3389/fcimb.2021.678522.

Zhang, Z., Xiao, J., Wang, H., Yang, C., Huang, Y., Yue, Z., Chen, Y., Han, L., Yin, K., Lyu, A., Fang, X., Zhang, L., 2024. Exploring high-quality microbial genomes by assembling short-reads with long-range connectivity. Nat. Commun. 15, 4631. https://doi.org/10.1038/s41467-024-49060-z.