Accurate profiling of the gut microbiota using surplus clinical Faecal Immunochemical Test (FIT) samples

Merel A van den Haak*[1,2], Jakub T Zbikowski*[1,2], Aliu Moomin[1,2], Joan Wilson[3], Chris Halsey[4], Charlie Gourley[5], Farhat Din[5], Stephen T McSorley[4], Elaina SR Collie-Duguid[6], Graham Horgan[7], Alan W Walker[1], Alexandra M Johnstone[1], Anne E Kiltie[1,2]

[1]The Rowett Institute, University of Aberdeen, UK

[2]Aberdeen Cancer Centre, University of Aberdeen, UK

[3]NHS Grampian Biorepository, Aberdeen Royal Infirmary, UK

[4]School of Cancer Sciences, University of Glasgow, UK

[5]Edinburgh Cancer Research, University of Edinburgh, UK

[6]Centre for Genome-Enabled Biology and Medicine, University of Aberdeen, UK

[7]Biomathematics & Statistics Scotland, Foresterhill, Aberdeen, UK

*These authors contributed equally to this work.

Corresponding author:

Professor Anne Kiltie, anne.kiltie@abdn.ac.uk

ORCID: https://orcid.org/0000-0001-7208-2912

Abstract

Background

Numerous countries use the EXTEL HEMO-AUTO MC Quantitative Faecal Immunochemical Test (qFIT) to screen for faecal haemoglobin. We aimed to determine if bacterial 16S rRNA gene sequencing results (16S V1V2) from the leftover qFIT cassettes would be stable over time and comparable with larger volume faecal collection protocols.

Methods

Four qFIT probe samples were taken from each of the sixteen fresh healthy volunteer stool samples and 16S results were compared after 0, 4, 7 and 14 days, to provide a baseline control and mimic postage and sample processing conditions in cancer screening programmes. qFIT results were then compared to those of standard laboratory processing of larger whole-stool samples. DNA was extracted from 100 NHS surplus qFIT samples from symptomatic patients reporting rectal bleeding and quantified to assess suitability for 16S sequencing.

Results

Bacterial composition and diversity from healthy volunteer qFITs remained stable over 14 days with no differences compared to baseline (day 0) and larger stool control samples; at least 75% of the symptomatic qFITs yielded sufficient DNA for 16S sequencing.

Conclusion

qFIT samples were not significantly different to control samples and stable over 14 days, allowing them to be used for large-scale low-cost population-based intestinal microbiota studies.

Clinical Trial Registration

The study was registered on clinicaltrials.gov (NCT06100549).

Introduction(4962 words)

Bowel screening programmes, active in many countries around the world, use qFIT (quantitative Faecal Immunochemical Testing) to screen a small faecal sample for the presence of haemoglobin, followed by a colonoscopy when the test is positive, thereby improving early detection of colorectal cancer. For example, approximately 0.5 million samples are obtained in the Scottish National Bowel screening programme (SNBSP) per year. Faecal haemoglobin sampling and measurement is therefore relevant to the majority of the 44,000 patients diagnosed with CRC each year in the UK (1–3). This test, the qFIT, is also recommended for testing patients reporting symptoms considered to be concerning for colorectal cancer (CRC) to their General Practitioner (GP) (4). In numerous sites throughout Scotland, England, Ireland, Wales, Japan and Taiwan, FIT testing uses the EXTEL HEMO-AUTO MC/ HM-JACKarc system (James Rae, Alpha Laboratories Ltd, personal communication), while some other countries use the OC-Sensor (Eiken, Japan) or FOB Gold (Sysmex, UK) FIT systems.

The gut microbiota, encompassing all microorganisms living in the gastrointestinal tract, plays important roles in human health. Perturbations in microbiota composition have been linked to cardiovascular disease (5), diabetes (6), hypertension (7), inflammatory bowel disease (8) and the development and progression of colorectal cancer (9). One of the most widely used methods to study gut microbiota composition and diversity is 16S rRNA gene sequencing. The 16S rRNA gene is present in all bacteria. Conserved regions in the 16S rRNA gene allow for universal amplification of bacterial DNA, and variable regions in the 16S gene, like the V1 and V2 region, are utilised to distinguish between different bacterial taxa (10,11).

If faecal samples already collected for bowel cancer screening or symptomatic testing reliably represent the gut microbiota, then such samples could provide large-scale low-cost population-based faecal gut microbiota data. At a population level, this could be linked to ‘big data’ within trusted research environments, to provide secure access to local health data for

research (12–15). With the possibility of longitudinal sampling and data linkage, the data obtained using routine qFIT faecal sampling could be used to definitively explore associations between the gut microbiota, ageing and diseases including cancer (16) in large, condition-relevant cohorts. Such an approach would provide unparalleled statistical power to answer questions on the role of the gut microbiota in some of the most prevalent causes of poor health, including cardiovascular disease, chronic inflammatory disease and diabetes, as well as providing vital insights into the role of the gut microbiota in cancer development and treatment response. Stool sampling using FIT has been reported to be widely acceptable and feasible in patients with lower GI symptoms in primary care (17) and is familiar to clinical teams. This, coupled with existing embedded pathways and protocols for requesting, stocking, collection, transport, processing and storage of FITs, provides a significant advantage over fresh stool for both sampling at scale and repeat sampling in large patient populations.

Several studies have already investigated the possibility of using the OC-sensor (Eiken, Japan) FIT test for gut microbiota research. Microbiota testing of DNA from faecal samples collected in the OC-sensor cartridges has demonstrated conserved community membership, alpha diversity and proportional abundance of genera when raw stool and FIT samples were compared (18,19), and the ten most dominant genera from the OC-sensor samples were consistent with other studies using stool samples (20). Another study examined three bacterial species and confirmed their presence in the OC-sensor, and reported no loss in detection levels for up to 6 days (21). The OC-sensor cartridges also conserved bacterial diversities when compared to immediately frozen raw stool from healthy participants, providing further evidence that microbiota profiles in OC-sensor collected faecal samples and stool samples are similar (22). The OC-sensor FITs have a stable faecal microbial content for up to seven days at room temperature (RT), which is important as FITs are shipped at ambient temperatures before analysis (21–23). Surplus samples after use in colorectal cancer screening programmes were also largely found to provide stable gut microbiota results over time, although the gut

microbiota composition was less stable when storing the FITs at 30°C, and differences in proportional abundance of some taxa were observed when stored at 4°C (24).

However, the OC-sensor collects 10 mg faeces in 2 ml buffer, in contrast to the 2 mg faeces in 2 ml buffer for the EXTEL HEMO-AUTO MC collection picker. Although the buffer compositions of the OC-sensor and EXTEL HEMO-AUTO MC collection pickers are proprietary, their composition likely differs, as the companies report different sample stabilities for their respective FITs (25,26). While microbial profiles of OC-sensor faecal samples have been studied extensively, to our knowledge, no microbial research has been done on the EXTEL HEMO-AUTO MC qFIT collection device, which is an important knowledge gap given the 5-fold lower amount of faeces collected and its widespread use in various countries for routine screening programmes.

Our aim was to test the hypotheses that a) the small amount of faecal material contained within the EXTEL HEMO-AUTO MC qFIT cartridges is sufficient to perform 16S rRNA gene sequencing, b) results would be stable over time (0, 4, 7 and 14 days) commensurate with clinical sample handling, and c) the microbial community found in the qFIT is comparable with larger volume faecal collection protocols. We further tested whether patients' routine use of the test kits results in sufficient DNA yield for microbiota sequencing.

Methods

Study participants

Two frozen stool samples from patients diagnosed with prostate cancer, stored at -70°C (18 months for sample 1 and 20 months for sample 2), were obtained from the NHS Research Scotland Biorepository for assay optimisation tests (27).

NHS Grampian provided 100 anonymous surplus qFIT samples from symptomatic patients who reported rectal bleeding and were therefore asked to do a qFIT test by their GP. These samples were received from the Biochemistry laboratory in Aberdeen Royal Infirmary, after 3-6 days of sample storage at 4°C following haemoglobin testing.

Sixteen healthy participants were also recruited, 6M and 10F, mean volunteer age 28.9 years (range: 19 to 57), mean body mass index (BMI) 24.1 kg/m^2 (range 20.4 to 29.1). Volunteers had no history of cardiovascular disease, diabetes, bowel disease, autoimmune disorders or cancer. Participants were excluded if they had taken any type of antibiotic in the three months before producing the faecal sample. All participants provided written informed consent before entering the study. The study protocol was approved by the Rowett Institute Ethical Committee and the NHS Grampian Biorepository (TR000330), and the study was registered on clinicaltrials.gov (NCT06100549).

Sample collection for determining optimal qFIT buffer volume for 16S rRNA gene sequencing

The qFIT probes (EXTEL HEMO-AUTO MC collection picker, Minaris Medical, now Canon Medical Diagnostics Ltd.), generously gifted to us by Alpha Laboraties Ltd, are designed to collect 2 mg of stool (excess stool is removed on probe insertion) into the collection device holding 2 ml of buffer. Haemoglobin testing of the qFIT sample using the HM-JACKarc analyser (Hitachi Chemical Diagnostics Systems, Japan) only requires 6 µl of the 2 ml buffer/faeces

suspension, after which the qFIT cartridge is stored at the NHS laboratory at 4°C to allow for a repeat screening in case of failure, before being discarded. Symptomatic samples are stored for a week post analysis at 4°C (Catriona Cameron, NHS Grampian, personal communication).

Initial testing was performed on frozen faecal samples obtained for a previous study (28) from two patients diagnosed with prostate cancer (Figure 1A). Nine separate qFIT probes were inserted into each sample and placed in their respective cassettes. Cartridges were vortexed for 2 minutes to ensure proper homogenisation before the buffer containing faeces was removed. Different volumes of buffer/faecal suspension were extracted from the qFIT cassette using a sterile glass Pasteur pipette to determine the optimal buffer/faecal suspension volume for 16S rRNA gene sequencing, and to establish if the entire sample was required to obtain satisfactory results. From qFITs 1-3, 450 µl qFIT sample was directly used for DNA extraction. The residual sample was centrifuged at 3000 rpm for 15 minutes to form a pellet, supernatant was removed, and the pellet was resuspended in 450 µl of the removed supernatant and used for DNA extraction. The entire qFIT sample from qFITs 4-6 was centrifuged, supernatant was removed, and the pellet resuspended in 450 µl of the supernatant and used for DNA extraction. qFITs 7-9 were frozen at -70°C for 24 days and processed using the same process as qFITs 4-6.

Symptomatic qFIT sample processing

For the 100 anonymous samples obtained from symptomatic patients (Figure 1B), 2 ml qFIT buffer containing faecal sample was extracted from the cartridges, centrifuged and the supernatant removed. The pellet was resuspended in 450 µl of the removed qFIT sample supernatant prior to DNA extraction. After Qubit DNA quantification to establish DNA yield, the samples were stored at -20°C for 3.5 months before determining DNA integrity by running extracted DNA on 0.8% agarose gels.

Healthy volunteer faecal sample collection and processing

The sixteen healthy participants were instructed to collect a 17-18 ml stool sample into a Sterilin 30 ml stool sample collection tube (Alpha Laboratories, UK) using Fe-Col® faecal collection paper (Alpha Laboratories, UK) and to deliver the sample to the lab within 18 hours. The samples were kept cool in cool bags containing frozen freezer blocks, while ensuring they did not freeze, and sample processing commenced immediately upon delivery to the Rowett Institute.

Four different qFIT probes were inserted multiple times into each sample until both serrations of the sample collection stick were appropriately filled (Figure 1C). Probe one was immediately processed for DNA extraction upon delivery of the sample (t=0). The remaining qFIT cartridges were stored for times and temperatures commensurate with postage and sample processing in the Scottish National Bowel Screening programme, corresponding to a time course of 4 days to 2 weeks. Probe 2 was stored at room temperature and DNA extracted on t=4 days to mimic postage/shipping conditions of the qFIT to the lab. As testing laboratories store the qFIT cartridges at 4°C, after storage at RT for 4 days, probes 3 and 4 were then stored at 4°C. DNA was extracted at t=7 days and t=14 days post-sample collection respectively. The cartridges were vortexed for 2 minutes to ensure proper homogenisation before processing. Then the entire sample was removed from the qFIT cartridges using sterile Pasteur pipettes, centrifuged at 13,000rpm for 15 minutes, the supernatant removed, and the pellet resuspended in 450 µl of the removed supernatant.

The qFIT samples were compared to a faecal sample collected and processed using a method which we have demonstrated to be acceptable in a clinical setting (29), to determine comparability of the sequencing data. A hand-homogenised, 5 g aliquot of raw stool was taken from the main faecal sample, suspended in 10 ml PBS solution with 30% glycerol and homogenised using a gentleMACS™ Dissociator (Miltenyi Biotec Ltd., UK). DNA was immediately extracted from a 450 µl aliquot. This was repeated at t=4 using another 5 g aliquot of the raw stool sample that had been stored at 4°C (Figure 1).

DNA extraction

The FastDNA SPIN Kit for Soil kit (MP Biomedicals, UK) was used for all DNA extractions (30). Processing followed the manufacturer's instructions with following minor adaptations: Resuspended qFIT samples (450 µl) and homogenised faecal matter from the whole stool-derived samples (450 µl) were placed in individual, separate lysing matrix E tubes, and 800 µl sodium phosphate buffer and 122 µl MT buffer were added to each tube. 450 µl of liquid sample was used, and 1 ml of Binding Matrix Solution supernatant was discarded. DNA was quantified using a Qubit 4 Fluorometer (Invitrogen, UK) using the standard protocol. Extracted DNA was stored at -20°C.

16S rRNA gene sequencing

16S rRNA gene sequencing was carried out at the Centre for Genome-Enabled Biology (CGEBM), University of Aberdeen. The 16S rRNA gene V1-V2 region was used for bacterial community profiling of the faecal samples as described previously (29).

Briefly, Illumina compatible libraries were prepared using 2-step PCR with 2.5 µl of input gDNA or study negative control (PBS/ 30% glycerol buffer in place of sample) per sample, 200 nM each of forward and reverse PCR primers and 1x Kapa HotStart ReadyMix (Roche) in a reaction volume of 25 µl. PCR primer sequences were 16S V1 (Forward: 5' **TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG** AGM GTT YGA TYM TGG CTC AG 3') and 16S V2 (Reverse: 5' **GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G**GC TGC CTC CCG TAG GAG T 3'), and included a specific Illumina compatible sequence (bold) that was the primer binding site in the second PCR. Plates were sealed and PCR was performed in a thermal cycler with initial denaturation at 95°C for 3 minutes, followed by 15 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, and a final extension at 72°C for 5 min. Plates were held at 4°C until further processing. PCR negative (water in place of gDNA) (n=3) and positive (ATCC MSA-1003 mock community) controls were included. For each experimental sample and control, 3 PCR reactions were

prepared to improve detection of rarer taxa. Following PCR, PCR reactions were combined per sample and purified using SPRI magnetic beads (Beckman-Coulter, CA), according to the manufacturer's instructions using a 0.8x ratio of beads:sample. Then 5 µl of purified amplicons were used as template in the second PCR with 1x Kapa HiFi Hotstart Readymix (Roche, CH) and 5 µl each of Nextera XT Index 1 and 2 primers (Nextera XT index kit v2, Illumina, CA), in a reaction volume of 50 µl. Plates were sealed and PCR was performed in a thermal cycler with initial denaturation at 95°C for 3 minutes, followed by 8 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec, and a final extension at 72°C for 5 min, then h~~o~~eld at 4°C. Final libraries were quantified by fluorimetry (Qubit, Thermo Fisher, MA) and average library size determined on a Tapestation 4200 (Agilent Technologies, CA), according to the manufacturer's instructions, to calculate molarity. All samples were equimolar pooled and sequenced on an Illumina MiSeq with 300 base pair (bp) paired end reads. Base calling and demultiplexing of raw data were performed in Illumina BaseSpace and these raw data fastq were used for downstream analysis.

Illumina MiSeq sequences were analysed in mothur 1.48.0 (31) based on the mothur Miseq standard operating procedures (32). Forward and reverse reads were paired into contigs, and poor-quality sequences (<260 bp or >470 bp) were removed from the dataset, as well as all reads with any ambiguous bases and homopolymeric stretches longer than 7 bp. Unique sequences were aligned with the SILVA reference database and incorrectly aligned sequences were removed. The pre.cluster command was run to merge sequences with <3 base differences in order to mitigate the impact of sequencing errors, and split.abund was used to remove singletons (33). The Ribosomal Database Project (RDP) version 19 reference database was used for taxonomic classification (34). Reads were clustered into operational taxonomic units (OTUs) with 97% similarity. Subsampling was performed at 31,284 bp for the samples donated by the prostate cancer patients, excluding one 2 ml sample due to low yield, resulting in an average Good's coverage estimate of 99.6%. Subsampling at 34,571 bp for the 16 healthy volunteer

samples resulted in the exclusion of seven low-yield samples, with a final average Good's coverage of 99.8%.

Statistical analysis and bioinformatics

For the two prostate cancer patients, Metastats software in mothur was used to investigate significant differences in bacterial abundance among the four buffer volumes, and between the four buffer volumes and the whole stool-derived sample at the family, genus and OTU level, focussing on OTUs or phylotypes in the top 100 when more than 100 OTUs or phylotypes were present (35). For the samples from the 16 healthy volunteers, Metastats was used to investigate significant differences in bacterial proportional abundance between whole stool-derived samples t=0 and t=4, qFIT t=0 and whole stool-derived samples t=0, qFIT t=4 and whole stool-derived samples t=4, qFIT t=0 and qFIT t=4, qFIT t=0 and qFIT t=7 and qFIT t=0 and qFIT t=14. All p-values were corrected with the Benjamini Hochberg method with cut-off set to $p<0.05$ to control for the false discovery rate (36). Comparing the same groups for both the symptomatic patients and healthy volunteers, OTU's and phylotypes were also tested using LEfSe (37) with FDR-adjustment in the software package MicrobiomeAnalyst, adding complementary statistical analysis to the Metastats findings (38). Using indices of alpha and beta diversity, it was established whether the bacterial composition of the gut microbiota differs between groups and over time. Where alpha diversity measures the microbial community features in individual samples, beta diversity measures the variation of species composition between samples (13). The alpha diversity for both the symptomatic patients and healthy volunteers was assessed by calculating the observed richness, inverted Simpsons index and Shannon diversity index in mothur and analysed using ANOVA in R studio (version 2024.12.0) (39) using the same grouping factors as described above. Principal Coordinate analysis (PCoA) plots were created in MicrobiomeAnalyst using the Bray Curtis Index calculator by generating a distance matrix on the "shared" file from mothur, visualising the beta diversity, and significance was established using nonparametric analysis of molecular variance (AMOVA) in mothur. Dendograms were

generated in mothur using the tree.shared command with the Bray Curtis calculator and visualised using iTOL (40). Parsimony software in mothur was used to calculate whether clustering patterns of dendrograms were significantly different by monitoring shared branches on the tree (41). Graphs were made in GraphPad (Version 5.04).

Results

Optimising the amount of qFIT faecal homogenate for 16S rRNA gene sequencing

Two donor samples from patients with prostate cancer were used to investigate the optimal qFIT buffer/faeces suspension volume for gut microbiota analysis. No statistically significant differences were found between the gut microbiota profiles of the qFIT samples at family, genus and OTU level within the patients, indicating a similar bacterial profile for all sample volumes and no effect of freezing for 24 days. Similarly, no statistically significant differences in bacterial proportional abundance were found between the qFIT samples and the whole stool sample (Figure 2A, Supplementary Figure 1A).

There was a statistically significant increase in the number of observed OTUs present in the aliquots taken from the whole stool-derived samples versus the four qFIT buffer/faeces suspension volumes (ANOVA, $p=<0.001$)(Figure 2B), but no statistically significant differences in diversity were found using both the Inverse Simpson and Shannon indices, which take into account both species richness and the evenness of species distribution (Supplementary Figure 1B, C). This indicates that although samples taken from whole stool-derived samples contained more taxa when compared to the qFIT samples, the evenness of species distribution was comparable between both sample types.

Next, beta diversity was visualised using a Bray Curtis-based Principal coordinate analysis (PCOA) plot (Figure 2C). AMOVA testing confirmed that samples clustered by donor ($p=<0.001$), but not by sample volume ($p=0.55$), indicating similar results independent of sampling method (Figure 2C). This was further supported by dendrogram analysis (Figure 2D), which confirmed that samples clustered by donor (Parsimony test, $p<0.001$) but not by sample volume (p=NS).

As these results showed that the volume of qFIT sample used for DNA extraction did not influence the results, the entire 2 ml qFIT buffer/faeces suspension was used for all further testing.

Symptomatic qFIT samples

DNA was extracted from qFIT buffer/faecal suspension from 100 symptomatic patients provided by the NHS laboratory to establish whether samples taken in real-world settings by patients would result in similar DNA yields as the qFITs tested in the research laboratory. Overall, the 100 cartridges had a mean DNA yield of 1.19 ng/µl (SD 1.04 ng/µl) (Figure 3 shows a representative agarose gel image). Twenty-five samples had a DNA concentration below 0.432 ng/µl, which was the lowest DNA concentration in samples successfully 16S rRNA gene sequenced in this study, indicating that the DNA yield of at least 75 samples was sufficient for 16S sequencing.

Comparison of samples taken from whole stool versus qFIT faecal samples

DNA was extracted at t=0 and t=4 for the whole stool samples, and at t=0, t=4, t=7 and t=14 days for the four qFIT samples for each of the sixteen healthy volunteers. When comparing alpha diversity measures, it was found that the whole stool-derived samples contained a significantly higher OTU richness compared to the qFIT buffer/faeces suspension samples (ANOVA, $p=<0.001$) (Figure 4A). However, there were no statistically significantly differences between the groups when comparing the Shannon and inverse Simpson diversity indices (Figure 4B, 4C), indicating a similar evenness of species distribution between the whole stool-derived samples and the four qFIT buffer/faecal suspensions. When comparing overall microbiota structure, whole stool-derived samples and qFIT samples clustered primarily by volunteer (Figure 5), with Parsimony testing indicating no significant differences between the samples derived from whole stool and the qFIT samples at t=0 and t=4 respectively, or between the qFIT timepoints, indicating a comparable species composition and further strengthening the evidence for similarity between the whole stool-derived samples and the qFIT samples containing faeces suspended in buffer over time.

When comparing individual taxa within the microbiota datasets, Metastats and LEfSe analysis showed no significant difference after Benjamini-Hochberg correction for multiple comparisons in either bacterial families, genera or OTUs when comparing the whole stool-derived samples at t=0 and t=4 with the qFIT samples containing faeces suspended in buffer at t=0 and t=4 respectively (Figure 6, Supplementary Figure 2). These results therefore further indicate a similar gut microbiota profile between whole stool samples and qFIT samples. Similarly, no statistically significant differences in proportional bacterial abundances were found with Metastats or LEfSe at the family, genus and OTU level between qFIT t=0 and t=4, qFIT t=0 and t=7 and FIT t=0 and t=14, suggesting that the qFIT samples remained stable over time.

Discussion

Our data supports the use of the EXTEL HEMO-AUTO MC qFIT cartridge to accurately measure gut microbiota. This unlocks the potential of repurposing leftover qFIT bowel screening samples for microbiome research.

Initial testing on two frozen stool samples from patients with prostate cancer demonstrated that the volume of qFIT buffer/faecal suspension used did not alter 16S rRNA gene sequencing results. It is noteworthy that the re-frozen qFIT samples closely resembled the other qFIT samples, indicating that samples could generate representative microbiota profiles for up to 24 days after re-freezing. It was also found that at least 75% of samples collected by patients in the real-world setting and processed through the NHS symptomatic testing service for qFIT haemoglobin testing showed a suitable DNA yield for 16S rRNA gene sequencing. The primary aim of this study was to compare the bacterial microbiota from 16 healthy participants using EXTEL HEMO-AUTO MC qFIT collection cassettes with samples collected from whole stool. The qFIT samples were comparable to the whole stool-derived samples except for the observed richness, likely due to improved detection of rare taxa from the larger faecal input amount (5 g versus 2 mg), demonstrating the suitability of this qFIT model for gut microbiota research. Moreover, there was no significant difference in bacterial proportional abundance, alpha diversity and beta diversity with sample storage over time, demonstrating that shipping and clinical laboratory sample handling in bowel screening programmes should not interfere with subsequent 16S rRNA gene sequence data quality, provided it is carried out within 14 days of stool collection.

This work, showing that the EXTEL HEMO-AUTO MC qFIT gives a similar gut microbiota composition when compared to raw stool samples, is in agreement with previous studies where faeces was collected with OC-sensor FIT cartridges (20,24,42). This is despite a five-fold

reduction in the amount of faeces collected in the EXTEL HEMO-AUTO MC cartridge and likely differences in composition of the proprietary buffers in each FIT type (the OC-sensor guarantees sample stability for 7 days at RT or 28 days at 2-8°C (26), whereas the EXTEL HEMO-AUTO MC cartridge is stable for 14 days at RT or 120 days at 2-8°C (25)). OC-sensor FITs also demonstrated microbiota clustering by volunteer rather than sampling method (43,44) and displayed comparable alpha diversity and composition between the FIT and fresh whole stool-derived sample (19). However, in contrast, several other papers have found small differences in bacterial proportional abundance between stool and OS-sensor FIT samples (42,44–46), which we did not find when comparing the EXTEL HEMO-AUTO MC with the whole stool samples, potentially explained by technical or biological variation.

We showed sample stability of the EXTEL HEMO-AUTO MC over time. Similar results were found for the OC-sensor FIT, with stable faecal microbiota results over multiple days when stored at RT (21–23). A change in bacterial profile was, however, demonstrated over time when the OC-sensor FIT was stored at 4°C and 30°C (24,47), and previous work has indicated that storage of stool samples at RT for four days resulted in changes of *Bifidobacterium* species, and a decrease in the relative abundance of *Faecalibacterium*, *Ruminococcus*, *Lachnospiraceae* and *Anaerostipes* (46). We did not observe such changes in the qFIT samples stored at RT for 4 days or stored at 4°C (Supplementary Figure 2), suggesting that the preservative in the EXTEL HEMO-AUTO MC qFIT buffer/faecal suspension could stabilise the overall microbiota profiles. Other previous work on the OC-sensor did support sample stability of whole stool samples at 4°C, in agreement with our findings (46,48). A limitation of our study is that we only extracted DNA at t=14 (after 4 days of storage at RT and 10 days of storage at 4°C). As such, we can only confirm sample stability and comparability of EXTEL HEMO-AUTO MC qFITs for 10 days post sample delivery to the laboratory.

It has previously been reported that incorrect sample collection using qFITs can alter the amount of haemoglobin measured (49), and improper sampling might explain potentially

insufficient DNA recovery for 25% of the symptomatic qFITs. Gudra et al. explored lyophilisation of OC-sensor FIT samples with a low DNA amount, and found that this increased the DNA yield by up to a factor of 30, thus allowing for effective 16S rRNA gene sequencing (43). A similar approach may be worthy of further investigation for the EXTEL HEMO-AUTO MC method. Considerations regarding the gut microbiota also need to be made when sequencing symptomatic patients, as blood in the stool has been shown to influence the composition of the gut microbiota in OC-AutoFIT samples, potentially due to the varying absorption and utilisation of iron from haemoglobin by different bacterial species (50,51).

Importantly, our results indicate that qFIT can be used for both haemoglobin screening and microbiota analysis. Given the challenges of collecting, storing, processing and comprehensively assessing the gut microbiota using larger volume/whole stool samples in both clinical and research settings, the utility of such an acceptable, easy to use, reliable and reproducible method of sampling which already has widespread clinical use cannot be understated. Additionally, as not all qFIT buffer/faecal suspension is required for 16S rRNA gene sequencing, the remainder of the buffer could be used for other analyses.

Several studies of OC-sensor and patient outcome data have investigated a relationship between the gut microbiota and colonoscopy results (20,42), and this would be worthy of studying with the EXTEL HEMO-AUTO MC system. Significant but subtle differences were found when comparing the microbial composition between different colonoscopy outcomes in individuals with haemoglobin positive OC-sensor FIT samples, suggesting a diagnostic potential of the bacterial composition found in FIT samples. Some 41 differentially abundant species were found when comparing colorectal cancer to non-colorectal cancer proportional abundances from the OC-sensor faecal samples of different patients (20). Another study found an increased abundance of bacterial content in patients with high-grade dysplasia and CRC (21), and models used to distinguish patients with carcinomas from healthy patients using the gut microbiota were equally predictive using stool-derived DNA or OC-sensor FIT-derived DNA

(42). The EXTEL HEMO-AUTO MC could also play an important role in established and advanced CRC given the developing evidence base relating the gut microbiome to key factors such as treatment resistance and disease recurrence (52), with samples at diagnosis and during treatment informing treatment allocation (53).

To our knowledge, no FIT has been tested yet in relation to other cancers and diseases. As manipulation of the gut microbiome using probiotics and faecal microbiota transplantation is already considered to aid patients with adverse microbial signatures (54), accessible qFIT sampling offers the potential to both guide treatment and measure treatment responses. The EXTEL HEMO-AUTO MC could also be employed to study microbiome-associated diseases including inflammatory bowel disease (55), cardiovascular and metabolic diseases and Alzheimer's disease (37–40), providing timely and clinically relevant results which are easier to obtain and process than with freshly collected stool samples. As individuals are screened in the bowel screening programme routinely every two years from 50-74 years, using qFITs would also allow for longitudinal screening of the gut microbiota (13,56). For example, generating a repository of microbiota data from all biennial qFITs of a subgroup of people at low or high risk for certain diseases could be used to investigate any potential biomarkers that precede development of symptoms. More generally, longitudinal studies of healthy cohorts will facilitate study of the effects of aging on the gut microbiota (15,57,58).

In conclusion, our work provides evidence for the utility of the EXTEL HEMO-AUTO MC qFIT sampling system, currently used in Scotland, Ireland, Wales, Japan and Taiwan for symptomatic testing and bowel cancer screening and in some NHS trusts in England for symptomatic testing, in determining gut microbiota profiles of individuals. This has potential for direct impact on diagnostics and therapeutics in intestinal disease and, in the longer term, the prospect of future intervention studies to improve outcomes in a large range of chronic diseases linked to microbiota perturbations.

Additional information

Acknowledgements

We thank all volunteers that provided us with faecal samples. We thank the NHS Grampian Biorepository at Aberdeen Royal Infirmary for providing the faecal samples from cancer patients and surplus symptomatic qFIT samples, and Fiona Brandie, Katie Cameron and Julie Hawken, Biochemistry Department, Aberdeen Royal Infirmary, for their expert handling of the symptomatic qFIT samples. We thank Zeynab Heidari for her expert technical processing of the 16S rRNA gene sequencing samples. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission

Author contributions (CRediT statement)

Conceptualisation: CH, CG, AEK

Patient recruitment: MAvdH and JTZ

Resources: JW, ECD

Methodology and investigation: MAvdH, JTZ, ECD

Data analysis and curation: MAvdH, AM, AWW, GH

Writing - Original draft preparation: MAvdH

Writing - review and editing, visualisation: AEK, AMJ, CH, CG, SMcS, FD, JTZ, MAvdH, ECD, GH, AM, AWW, JW

Supervision: AEK, AMJ

Funding acquisition: AEK, SMcS, FD

All authors reviewed the final manuscript and agreed to publication.

Ethics approval and consent to participate

The study protocol was approved by the Rowett Ethics committee and the NHS Grampian Biorepository (TR000330), and the study was registered on clinicaltrials.gov (NCT06100549). All healthy participants provided informed consent before starting the study. The two samples from patients with prostate cancer and 100 symptomatic samples were anonymously obtained under NHS Grampian Biorepository consent.

Consent for publication

Not applicable

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

The raw sequencing datasets used for this study have been deposited in the NCBI Short Read Archive, under accession number PRJNA1268008.

Competing interests

AMJ holds a voluntary committee role with the Nutrition Society, Association for the study of Obesity and British Nutrition Foundation. AWW has received research funding from ZOE Ltd, and undertakes consultancy for EnteroBiotix Ltd. The other authors declare no conflict of interest.

Funding information

This work was funded by the NHS Grampian Endowment Fund. CRUK Scotland Centre funding enabled collaboration and research development (CTRQQR-2021\100006) and this project was supported by the RSE Research Workship grant 2399. MvdH is funded by a PhD studentship funded by the University of Aberdeen Development Trust and Friends of ANCHOR.

AWW and the Rowett Institute receive core funding from the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS). AEK's Chair is funded by Friends of ANCHOR. FVND's Chair is funded by Bowel Cancer UK-Royal College of Surgeons of Edinburgh.

References

1. Gerrard AD, Maeda Y, Miller J, Gunn F, Theodoratou E, Noble C, et al. Double faecal immunochemical testing in patients with symptoms suspicious of colorectal cancer. Br J Surg. 2023 Mar 30;110(4):471–80.
2. Gerrard AD, Garau R, Xu W, Maeda Y, Dunlop MG, Theodoratou E, et al. Repeat Faecal Immunochemical Testing for Colorectal Cancer Detection in Symptomatic and Screening Patients: A Systematic Review and Meta-Analysis. Cancers (Basel). 2024 Sep 19;16(18):3199.
3. Cancer Research UK [Internet]. 2015 [cited 2025 Apr 24]. Bowel cancer statistics. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer
4. Overview | Quantitative faecal immunochemical testing to guide colorectal cancer pathway referral in primary care | Guidance | NICE [Internet]. NICE; 2023 [cited 2025 Apr 24]. Available from: https://www.nice.org.uk/guidance/dg56
5. Nesci A, Carnuccio C, Ruggieri V, D'Alessandro A, Di Giorgio A, Santoro L, et al. Gut Microbiota and Cardiovascular Disease: Evidence on the Metabolic and Inflammatory Background of a Complex Relationship. Int J Mol Sci. 2023 May 22;24(10):9087.
6. Iatcu CO, Steen A, Covasa M. Gut Microbiota and Complications of Type-2 Diabetes. Nutrients. 2021 Dec 30;14(1):166.
7. Yang Z, Wang Q, Liu Y, Wang L, Ge Z, Li Z, et al. Gut microbiota and hypertension: association, mechanisms and treatment. Clinical and Experimental Hypertension. 2023 Dec 31;45(1):2195135.
8. Qiu P, Ishimoto T, Fu L, Zhang J, Zhang Z, Liu Y. The Gut Microbiota in Inflammatory Bowel Disease. Front Cell Infect Microbiol. 2022 Feb 22;12:733992.
9. Kim J, Lee HK. Potential Role of the Gut Microbiome In Colorectal Cancer Progression. Front Immunol. 2022 Jan 7;12:807648.
10. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019 Nov 6;10(1):5029.
11. Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, et al. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. mSphere. 6(1):e01202-20.
12. Safe Havens | NHS Research Scotland | NHS Research Scotland [Internet]. [cited 2025 Mar 28]. Available from: https://www.nhsresearchscotland.org.uk/research-in-scotland/data/safe-havens
13. Björk JR, Bolte LA, Maltez Thomas A, Lee KA, Rossi N, Wind TT, et al. Longitudinal gut microbiome changes in immune checkpoint blockade-treated advanced melanoma. Nat Med. 2024 Mar;30(3):785–96.
14. Rezasoltani S, Azizmohammad Looha M, Asadzadeh Aghdaei H, Jasemi S, Sechi LA, Gazouli M, et al. 16S rRNA sequencing analysis of the oral and fecal microbiota in colorectal cancer positives versus colorectal cancer negatives in Iranian population. Gut Pathogens. 2024 Feb 20;16(1):9.
15. Ragonnaud E, Biragyn A. Gut microbiota as the key controllers of "healthy" aging of elderly people. Immunity & Ageing. 2021 Jan 5;18(1):2.
16. Inamura K, Hamada T, Bullman S, Ugai T, Yachida S, Ogino S. Cancer as microenvironmental, systemic, and environmental diseases: opportunity for transdisciplinary microbiomics science. Gut. 2022 Jul 12;gutjnl-2022-327209.
17. Georgiou Delisle T, D'Souza N, Davies B, Benton S, Chen M, Ward H, et al. Faecal immunochemical test for suspected colorectal cancer symptoms: patient survey of usability and acceptability. BJGP Open. 2022 Mar;6(1):BJGPO.2021.0102.

18. Baxter NT, Koumpouras CC, Rogers MAM, Ruffin MT, Schloss PD. DNA from fecal immunochemical test can replace stool for detection of colonic lesions using a microbiota-based model. Microbiome. 2016 Nov 14;4:59.
19. Rounge TB, Meisal R, Nordby JI, Ambur OH, de Lange T, Hoff G. Evaluating gut microbiota profiles from archived fecal samples. BMC Gastroenterol. 2018 Nov 8;18:171.
20. Khannous-Lleiffe O, Willis JR, Saus E, Moreno V, Castellví-Bel S, Gabaldón T. Microbiome Profiling from Fecal Immunochemical Test Reveals Microbial Signatures with Potential for Colorectal Cancer Screening. Cancers (Basel). 2022 Dec 25;15(1):120.
21. Grobbee EJ, Lam SY, Fuhler GM, Blakaj B, Konstantinov SR, Bruno MJ, et al. First steps towards combining faecal immunochemical testing with the gut microbiome in colorectal cancer screening. United European Gastroenterol J. 2020 Apr;8(3):293–302.
22. Vogtmann E, Chen J, Amir A, Shi J, Abnet CC, Nelson H, et al. Comparison of Collection Methods for Fecal Samples in Microbiome Studies. Am J Epidemiol. 2017 Jan 15;185(2):115–23.
23. Krigul KL, Aasmets O, Lüll K, Org T, Org E. Using fecal immunochemical tubes for the analysis of the gut microbiome has the potential to improve colorectal cancer screening. Sci Rep. 2021 Oct 1;11(1):19603.
24. Zouiouich S, Mariadassou M, Rué O, Vogtmann E, Huybrechts I, Severi G, et al. Comparison of fecal sample collection methods for microbial analysis embedded within colorectal cancer screening programs. Cancer Epidemiol Biomarkers Prev. 2022 Feb;31(2):305–14.
25. Faecal immunochemical test sample Collection Picker. To collect, preserve, and test samples for analysis on the HM-JACKarc analyser. Pack of 200. [Internet]. [cited 2024 Jul 31]. Available from: https://www.alphalabs.co.uk/063631
26. FIT Screening - UK - FAQs [Internet]. [cited 2024 Jul 31]. Available from: http://www.fit-screening.co.uk/resources/faqs
27. NHS Grampian Tissue Bank [Internet]. [cited 2025 Jun 16]. Available from: https://www.biorepository.nhsgrampian.org/
28. Cancer Research UK [Internet]. 2023 [cited 2025 Apr 28]. A study looking at how bacteria in the gut affects treatment for pelvic cancer (PELICAN-23). Available from: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial/a-study-looking-at-how-bacteria-in-the-gut-affects-treatment-for-pelvic-cancer-pelican-23
29. Then CK, Paillas S, Moomin A, Misheva MD, Moir RA, Hay SM, et al. Dietary fibre supplementation enhances radiotherapy tumour control and alleviates intestinal radiation toxicity. Microbiome. 2024 May 14;12:89.
30. Mukhopadhya I, Martin JC, Shaw S, McKinley AJ, Gratz SW, Scott KP. Comparison of microbial signatures between paired faecal and rectal biopsy samples from healthy volunteers using next-generation sequencing and culturomics. Microbiome. 2022 Oct 14;10(1):171.
31. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl Environ Microbiol. 2009 Dec;75(23):7537–41.
32. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. Applied and Environmental Microbiology. 2013 Sep;79(17):5112–20.
33. Huse SM, Welch DM, Morrison HG, Sogin ML. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol. 2010 Jul;12(7):1889–98.
34. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol. 2007 Aug;73(16):5261–7.
35. White JR, Nagarajan N, Pop M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol. 2009 Apr;5(4):e1000352.

36. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57(1):289–300.
37. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biology. 2011 Jun 24;12(6):R60.
38. Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. Nat Protoc. 2020 Mar;15(3):799–821.
39. Finotello F, Mastrorilli E, Di Camillo B. Measuring the diversity of the human microbiota with targeted next-generation sequencing. Briefings in Bioinformatics. 2018 Jul 20;19(4):679–92.
40. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Research. 2024 Jul 5;52(W1):W78–82.
41. Schloss PD, Handelsman J. Introducing TreeClimber, a test to compare microbial community structures. Appl Environ Microbiol. 2006 Apr;72(4):2379–84.
42. Baxter NT, Koumpouras CC, Rogers MAM, Ruffin MT, Schloss PD. DNA from fecal immunochemical test can replace stool for detection of colonic lesions using a microbiota-based model. Microbiome. 2016 Nov 14;4(1):59.
43. Gudra D, Shoaie S, Fridmanis D, Klovins J, Wefer H, Silamikelis I, et al. A widely used sampling device in colorectal cancer screening programmes allows for large-scale microbiome studies. Gut. 2019 Sep;68(9):1723–5.
44. Brezina S, Borkovec M, Baierl A, Bastian F, Futschik A, Gasche N, et al. Using fecal immmunochemical cartridges for gut microbiome analysis within a colorectal cancer screening program. Gut Microbes. 15(1):2176119.
45. Byrd DA, Chen J, Vogtmann E, Hullings A, Song SJ, Amir A, et al. Reproducibility, stability, and accuracy of microbial profiles by fecal sample collection method in three distinct populations. PLoS One. 2019 Nov 18;14(11):e0224757.
46. Choo JM, Leong LE, Rogers GB. Sample storage conditions significantly influence faecal microbiome profiles. Sci Rep. 2015 Nov 17;5:16350.
47. Birkeland E, Ferrero G, Pardini B, Umu SU, Tarallo S, Bulfamante S, et al. Profiling small RNAs in fecal immunochemical tests: is it possible? Mol Cancer. 2023 Oct 3;22(1):161.
48. Holzhausen EA, Nikodemova M, Deblois CL, Barnet JH, Peppard PE, Suen G, et al. Assessing the impact of storage time on the stability of stool microbiota richness, diversity, and composition. Gut Pathog. 2021 Dec 20;13:75.
49. Gies A, Gruner LF, Schrotz-King P, Brenner H. Effect of Imperfect Compliance With Instructions for Fecal Sample Collection on Diagnostic Performance of 9 Fecal Immunochemical Tests. Clinical Gastroenterology and Hepatology. 2019 Aug 1;17(9):1829-1839.e4.
50. Chénard T, Malick M, Dubé J, Massé E. The influence of blood on the human gut microbiome. BMC Microbiol. 2020 Mar 3;20:44.
51. Seyoum Y, Baye ,Kaleab, and Humblot C. Iron homeostasis in host and gut bacteria – a complex interrelationship. Gut Microbes. 2021 Jan 1;13(1):1874855.
52. Teng H, Wang Y, Sui X, Fan J, Li S, Lei X, et al. Gut microbiota-mediated nucleotide synthesis attenuates the response to neoadjuvant chemoradiotherapy in rectal cancer. Cancer Cell. 2023 Jan 9;41(1):124-138.e6.
53. Cheng WY, Wu CY, Yu J. The role of gut microbiota in cancer treatment: friend or foe? Gut. 2020 Oct;69(10):1867–76.
54. Zhao W, Lei J, Ke S, Chen Y, Xiao J, Tang Z, et al. Fecal microbiota transplantation plus tislelizumab and fruquintinib in refractory microsatellite stable metastatic colorectal cancer: an open-label, single-arm, phase II trial (RENMIN-215). eClinicalMedicine [Internet]. 2023 Dec 1 [cited 2025 Apr 24];66. Available from: https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(23)00492-3/fulltext

55. Bethlehem L, Estevinho MM, Grinspan A, Magro F, Faith JJ, Colombel JF. Microbiota therapeutics for inflammatory bowel disease: the way forward. Lancet Gastroenterol Hepatol. 2024 May;9(5):476–86.
56. Olsson LM, Boulund F, Nilsson S, Khan MT, Gummesson A, Fagerberg L, et al. Dynamics of the normal gut microbiota: A longitudinal one-year population study in Sweden. Cell Host & Microbe. 2022 May 11;30(5):726-739.e3.
57. Ghosh TS, Shanahan F, O'Toole PW. The gut microbiome as a modulator of healthy ageing. Nat Rev Gastroenterol Hepatol. 2022 Sep;19(9):565–84.
58. Salazar J, Durán P, Díaz MP, Chacín M, Santeliz R, Mengual E, et al. Exploring the Relationship between the Gut Microbiota and Ageing: A Possible Age Modulator. Int J Environ Res Public Health. 2023 May 17;20(10):5845.

Figure Legends

Figure 1. Schematic showing sample processing steps for each participant group. A) Two stool samples were obtained from a previous study for assay optimisation. B) DNA was extracted from 100 anonymised qFIT samples from symptomatic patients. C) DNA was extracted from four qFIT cartridges and two whole stool-derived samples per healthy volunteer (n=16). The 16S rRNA gene sequencing results from the qFIT samples were compared to the whole stool samples. The figure was created with BioRender.com.

Figure 2. Neither the amount of qFIT buffer used for DNA extraction nor freezing influenced the 16S rRNA gene sequencing results. A) Microbiota composition, showing the 15 most proportionally abundant gut microbiota families. Metastats with Benjamini-Hochberg correction and Benjamini-Hochberg adjusted LEfSe were used to confirm the absence of statistically significant differences between the groups for the 100 most abundant OTU's and family and genus level phylotypes; B) Observed OTU richness for each of the samples. ANOVA was conducted to compare groups for statistically significant differences; C) Principal coordinate analysis using Bray-Curtis dissimilarity. AMOVA was performed to identify significant differences across groups; D) Dendrogram made with the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method based on Bray Curtis assessment of community structure. Parsimony testing was employed to determine whether group differences were statistically significant.

Figure 3. DNA from qFIT faecal samples from healthy controls and symptomatic individuals. The gel shows samples from patients 1-10. 1kb ladders are shown on far left and before the controls, and 5 µl of each sample was added to the gel. DNA concentrations for the ten symptomatic individuals left to right: 4.69, 5.12, 13.32, 10.80, 5.10, 2.47, 4.93, 10.32, 2.12 and 0.0 ng/µl per lane respectively. Controls from left to right: healthy volunteer 2 (t=4),

volunteer 11 (t=7), volunteer 8 (t=4) and volunteer 1 (t=0) with DNA concentrations of 8.8, 14.1, 9.06 and 8.13 ng/µl respectively.

Figure 4. Whole stool-derived samples contained more taxa when compared to the qFIT samples, but with a comparable evenness of species distribution between both sample types. Alpha diversity measures are shown for whole stool-derived samples and qFIT samples over time. ANOVA was carried out to evaluate potential statistical differences between the groups. A) the observed OTU richness. B) Shannon diversity index. C) Inverse Simpson diversity index.

Figure 5. Stool-derived samples and qFIT samples cluster by volunteer rather than sampling type. The dendrogram was created using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) based on the Bray Curtis-calculated dissimilarities in community structure. Individual volunteers are distinguished by colour (inner ring), and the outer ring shows the compositional profiles of the 15 most proportionally abundant gut microbiota families.

Figure 6. qFIT and whole stool-derived samples showed a similar microbiota profile at the family and genus levels, with qFIT profiles remaining stable over time. A) composition of the 15 most proportionally abundant gut microbiota families. B) composition of the 15 most proportionally abundant gut microbiota genera. Samples are grouped by individual volunteer, shown by volunteer number, to display individual variation. Metastats on the 100 most abundant OTU's and phylotypes with Benjamini-Hochberg correction and Benjamini-Hochberg adjusted LEfSe were conducted to evaluate statistically significant differences. The black vertical columns represent missing data.

Figures

Figure 1

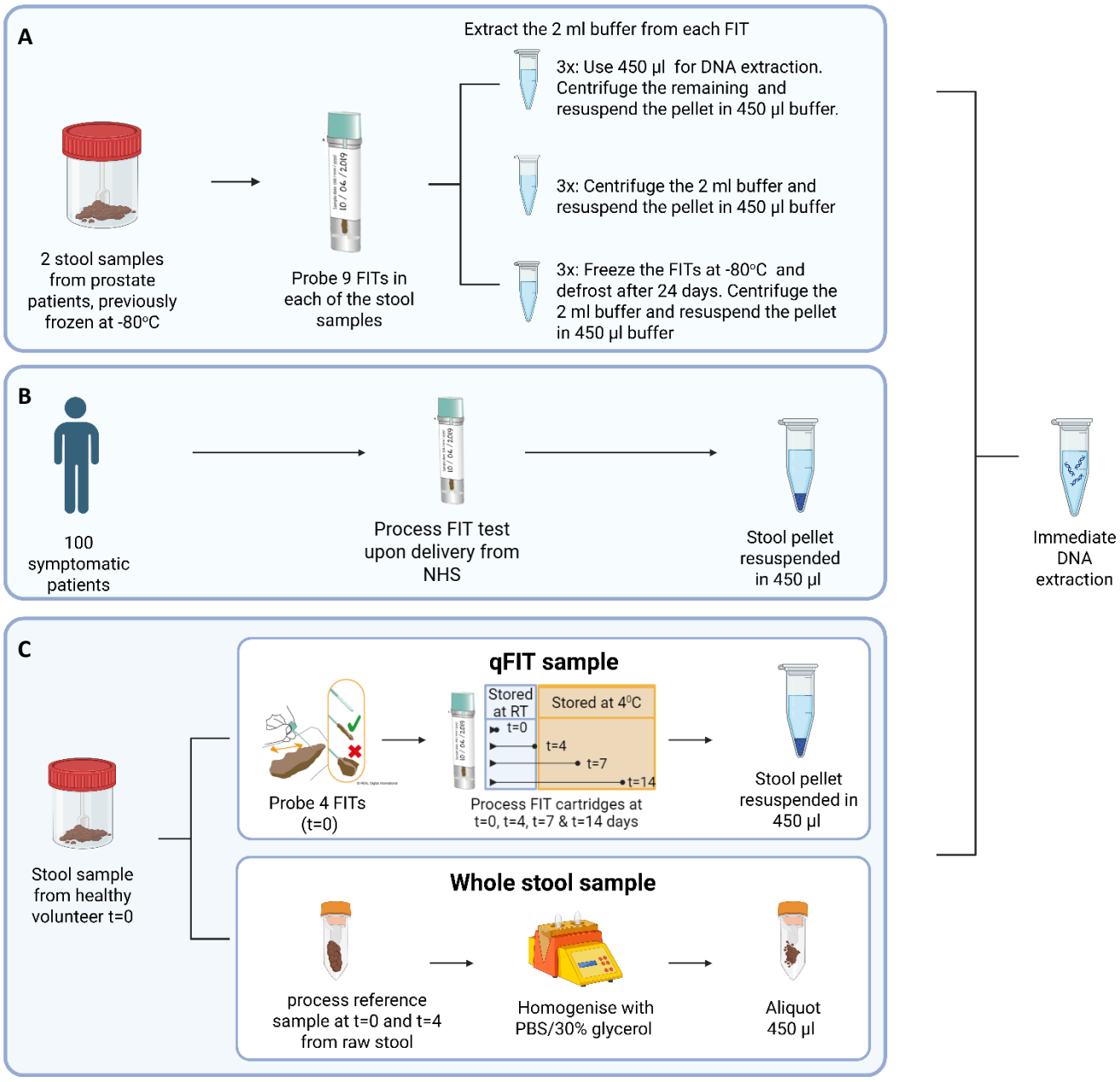


Figure 2

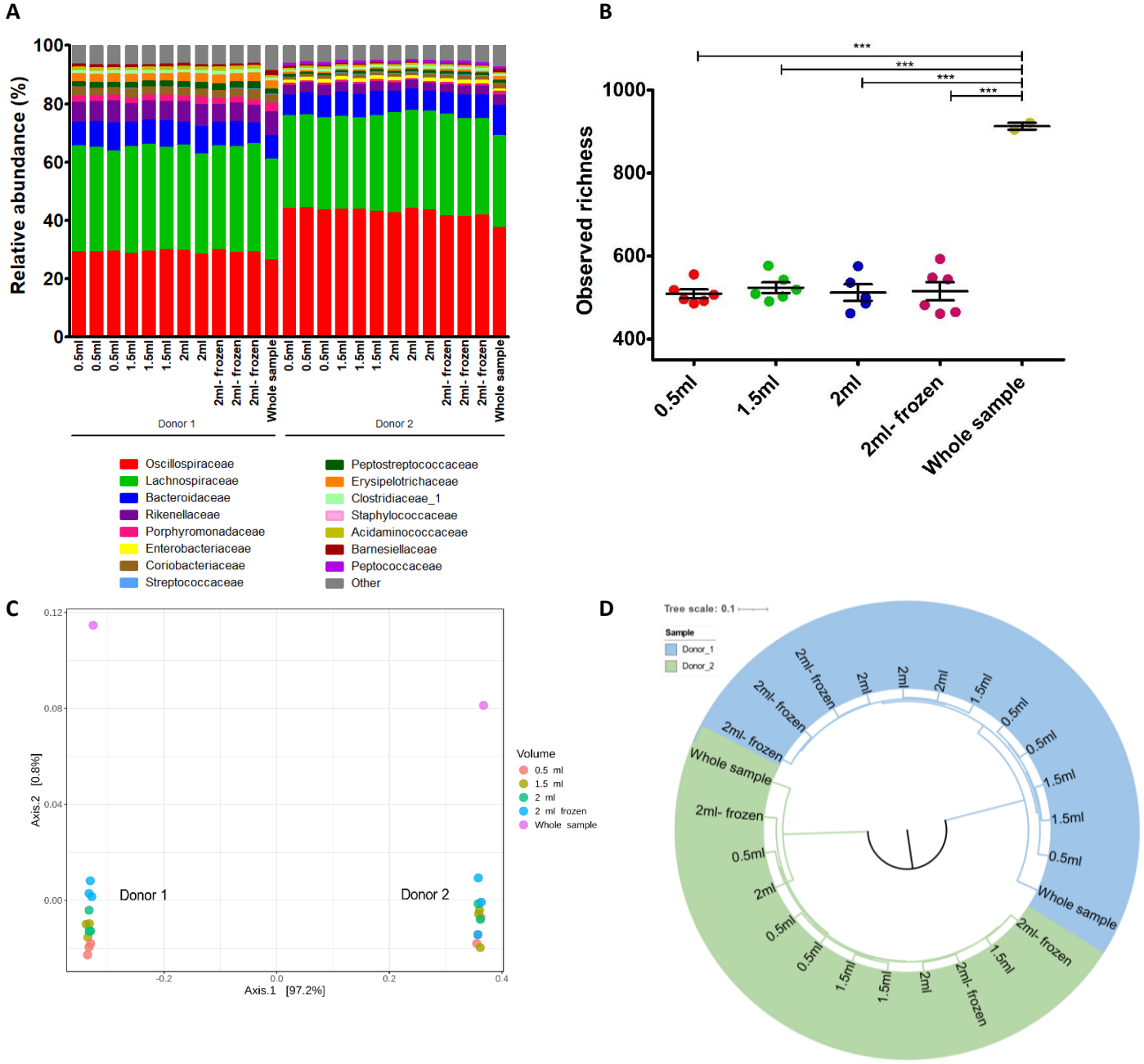

A
Relative abundance (%)
Donor 1
Donor 2
Oscillospiraceae
Lachnospiraceae
Bacteroidaceae
Rikenellaceae
Porphyromonadaceae
Enterobacteriaceae
Coriobacteriaceae
Streptococcaceae
Peptostreptococcaceae
Erysipelotrichaceae
Clostridiaceae_1
Staphylococcaceae
Acidaminococcaceae
Barnesiellaceae
Peptococcaceae
Other
B
Observed richness
0.5ml
1.5ml
2ml
2ml- frozen
Whole sample
C
Axis.2 [0.8%]
Axis.1 [97.2%]
Donor 1
Donor 2
Volume
0.5 ml
1.5 ml
2 ml
2 ml frozen
Whole sample
D
Tree scale: 0.1
Sample
Donor_1
Donor_2

Figure 3

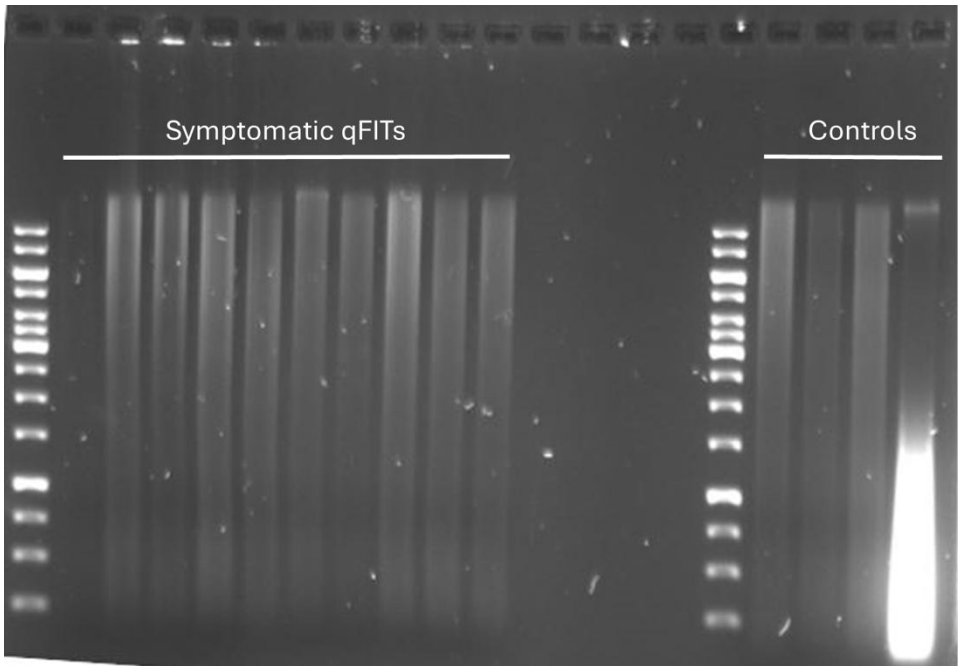


Figure 4

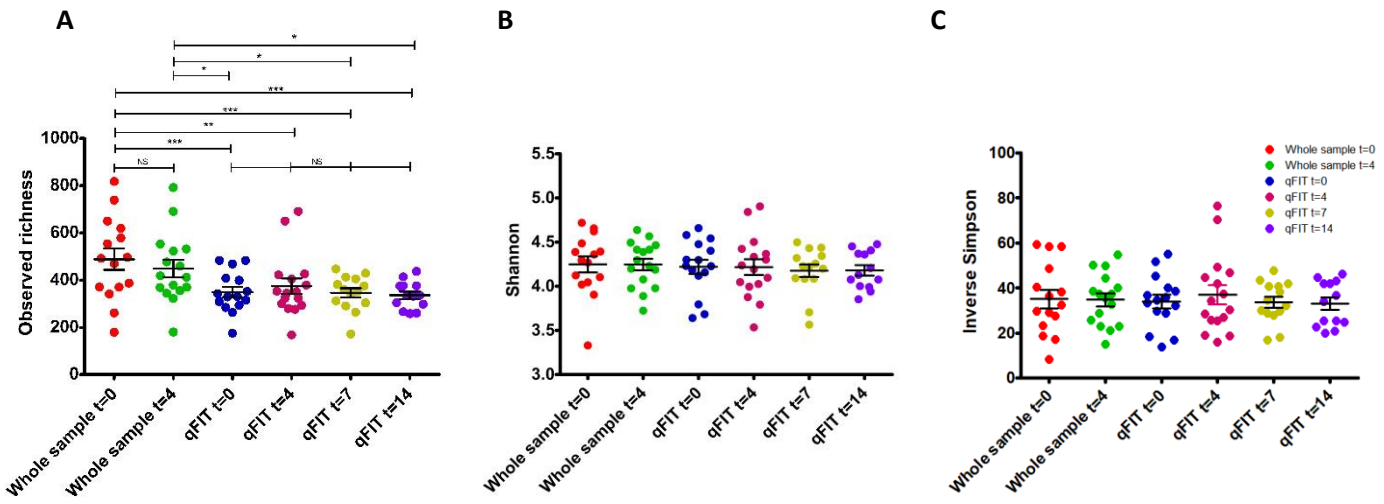

A
B
C
Observed richness
Shannon
Inverse Simpson
Whole sample t=0
Whole sample t=4
qFIT t=0
qFIT t=4
qFIT t=7
qFIT t=14
NS
NS

Figure 5

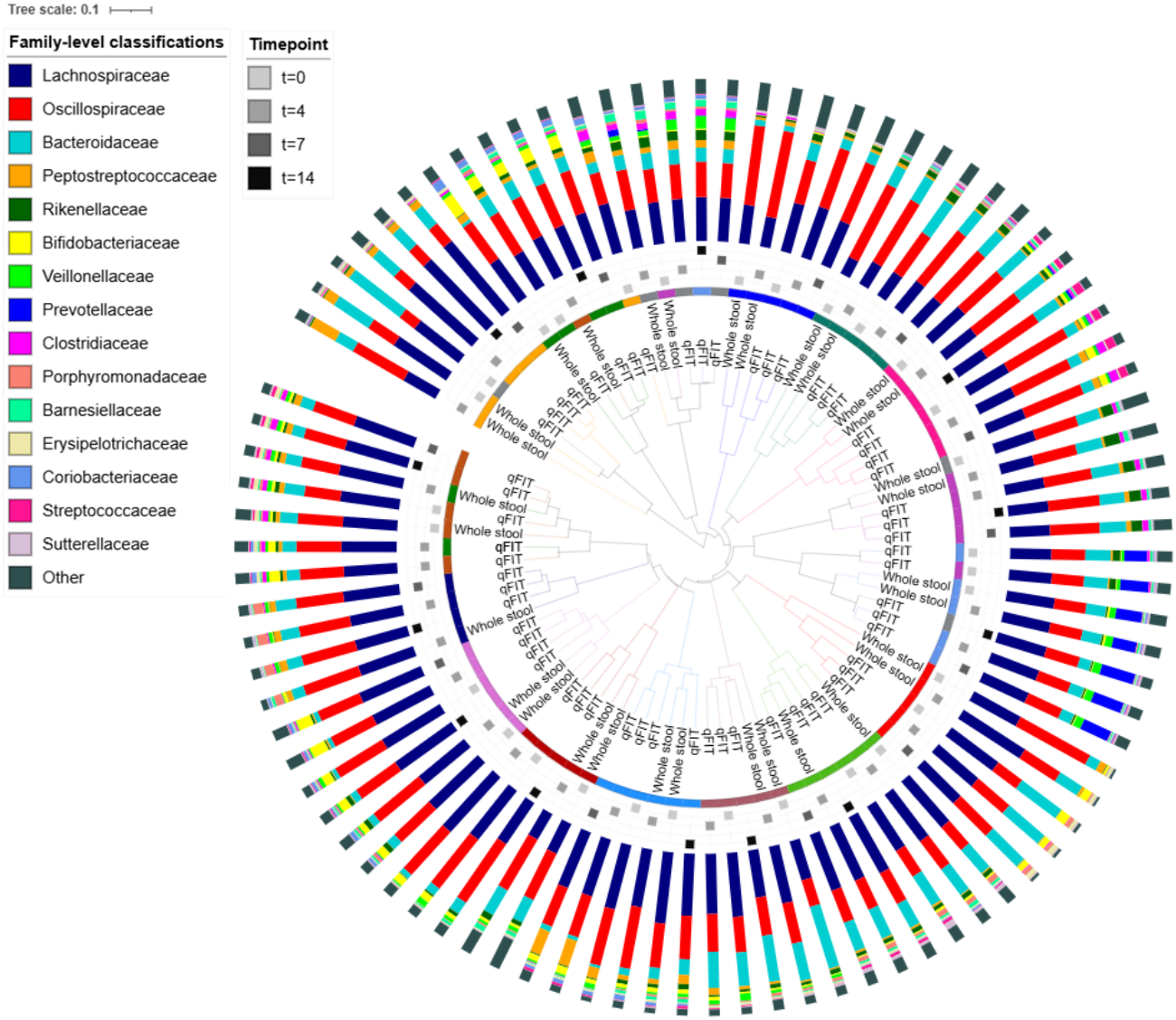
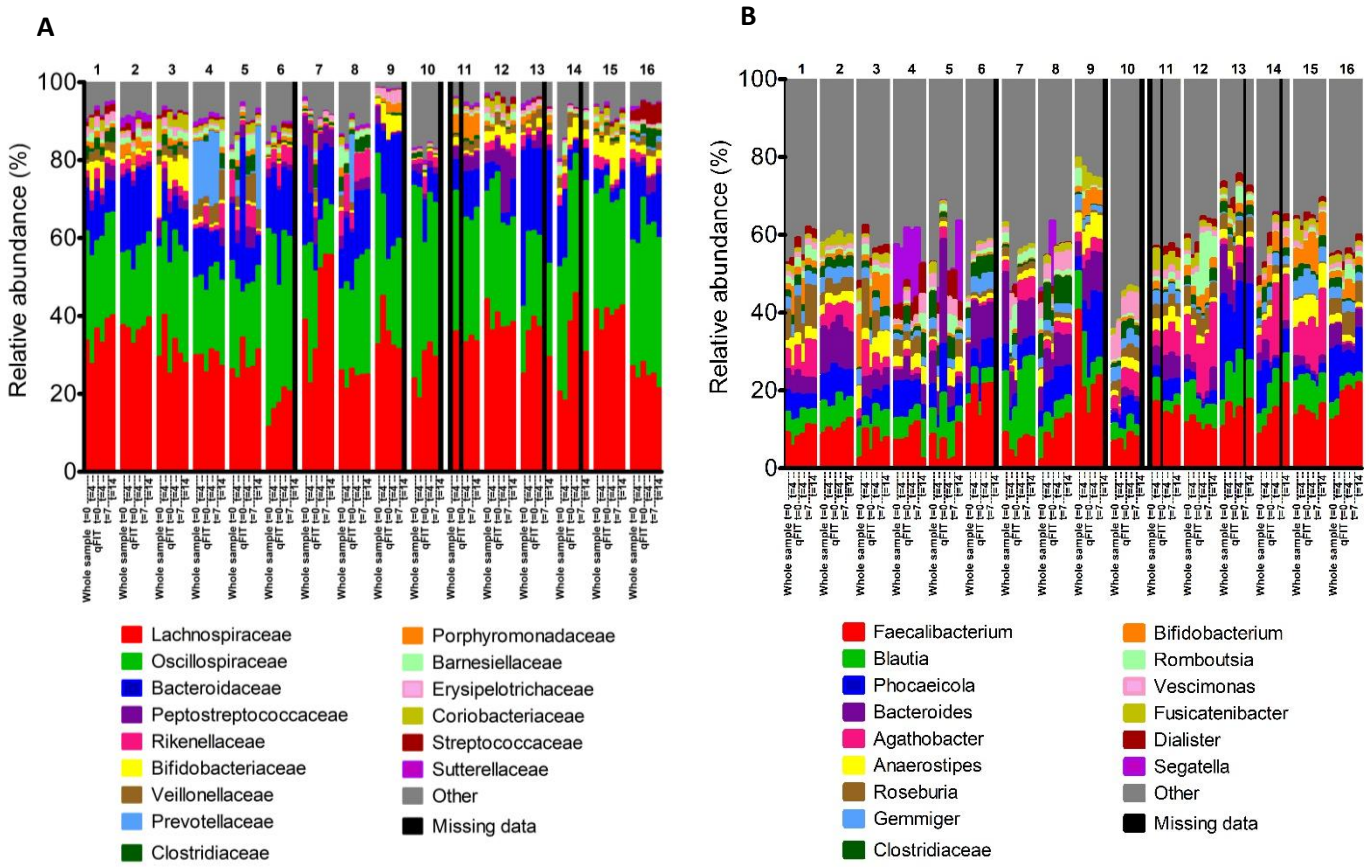

Tree scale: 0.1
Family-level classifications
Lachnospiraceae
Oscillospiraceae
Bacteroidaceae
Peptostreptococcaceae
Rikenellaceae
Bifidobacteriaceae
Veillonellaceae
Prevotellaceae
Clostridiaceae
Porphyromonadaceae
Barnesiellaceae
Erysipelotrichaceae
Coriobacteriaceae
Streptococcaceae
Sutterellaceae
Other
Timepoint
t=0
t=4
t=7
t=14
Whole stool
qFIT

Figure 6


A
B
Relative abundance (%)
Lachnospiraceae
Oscillospiraceae
Bacteroidaceae
Peptostreptococcaceae
Rikenellaceae
Bifidobacteriaceae
Veillonellaceae
Prevotellaceae
Clostridiaceae
Porphyromonadaceae
Barnesiellaceae
Erysipelotrichaceae
Coriobacteriaceae
Streptococcaceae
Sutterellaceae
Other
Missing data
Faecalibacterium
Blautia
Phocaeicola
Bacteroides
Agathobacter
Anaerostipes
Roseburia
Gemmiger
Clostridiaceae
Bifidobacterium
Romboutsia
Vescimonas
Fusicatenibacter
Dialister
Segatella
Other
Missing data