Assessing the comparability of different DNA extraction and amplification methods in gut microbial community profiling

Elizabeth K. Mallott[1,*], Ripan S. Malhi[2,3] and Katherine R. Amato[1]

Abstract

Automated, high-throughput technologies are becoming increasingly common in microbiome studies to decrease costs and increase efficiency. However, in microbiome studies, small differences in methodology – including storage conditions, wet lab methods, sequencing platforms and data analysis – can influence the reproducibility and comparability of data across studies. There has been limited testing of the effects of high-throughput methods, including microfluidic PCR technologies. In this paper, we compare two extraction methods (the QIAamp DNA Stool Mini Kit and the MoBio PowerSoil DNA Isolation kit), two taq polymerase enzymes (MyTaq HS Red Mix and Accustart II PCR ToughMix), two primer sets (V3–V4 and V4–V5) and two amplification methods (a common two-step PCR protocol and amplicon library preparation on the Fluidigm Access Array system that allows automated multiplexing of primers). Gut microbial community profiles were significantly affected by all variables. While there were no significant differences in alpha diversity measured between the two extraction methods, there was an effect of extraction method on community composition measured by unweighted UniFrac distances. Both amplification method and primers had a significant effect on both alpha diversity and community composition. The relative abundance of *Actinobacteria* was significantly lower when using the MoBio kit or Fluidigm amplification method, and the relative abundance of *Firmicutes* was lower when using the Qiagen kit. Microbial community profiles based on Fluidigm-generated amplicon libraries were not comparable to those generated with more commonly used methods. Researchers should carefully consider the limitations and biases that different extraction and amplification methods can introduce into their results. Additionally, more thorough benchmarking of automated and multiplexing methods is necessary to determine the magnitude of the potential trade-off between the quality and the quantity of data.

INTRODUCTION

The importance of the microbiome to many biological systems is becoming increasingly clear. However, large-scale studies of the microbiome are still limited by the available technology as processing large numbers of samples can be costly. Thus, there is great interest in automated, high-throughput technologies that increase efficiency and decrease costs. One popular tool to increase the speed of sample preparation that has been adopted by many labs is robotic DNA extraction, PCR setup and sample pooling. More recently, microfluidic PCR technology that allows the use of multiple primer pairs simultaneously, such as the Fluidigm Access Array system, has been introduced to maximize the types of data produced at once [1–9].

The Fluidigm system is attractive for microbiome research given that it allows for amplification of the region of interest and barcoding of samples within the same reaction. Additionally, the microfluidic nature of the system limits contamination, reduces primer bias by allowing for the use of multiple primers in separate reactions in a single run, and allows for the generation of multiple data types in one run (e.g. 16S, 18S, *COI*, *trnL*) [1, 5]. However, despite these clear benefits, drawbacks to the Fluidigm system exist. Most importantly, it confers limited control over reaction conditions, which is potentially problematic when working with low-quantity or low-quality DNA samples, or when using multiple primers with different optimal melting temperatures.

Systematic comparisons of microbiome methods indicate that small differences in methodology can influence the
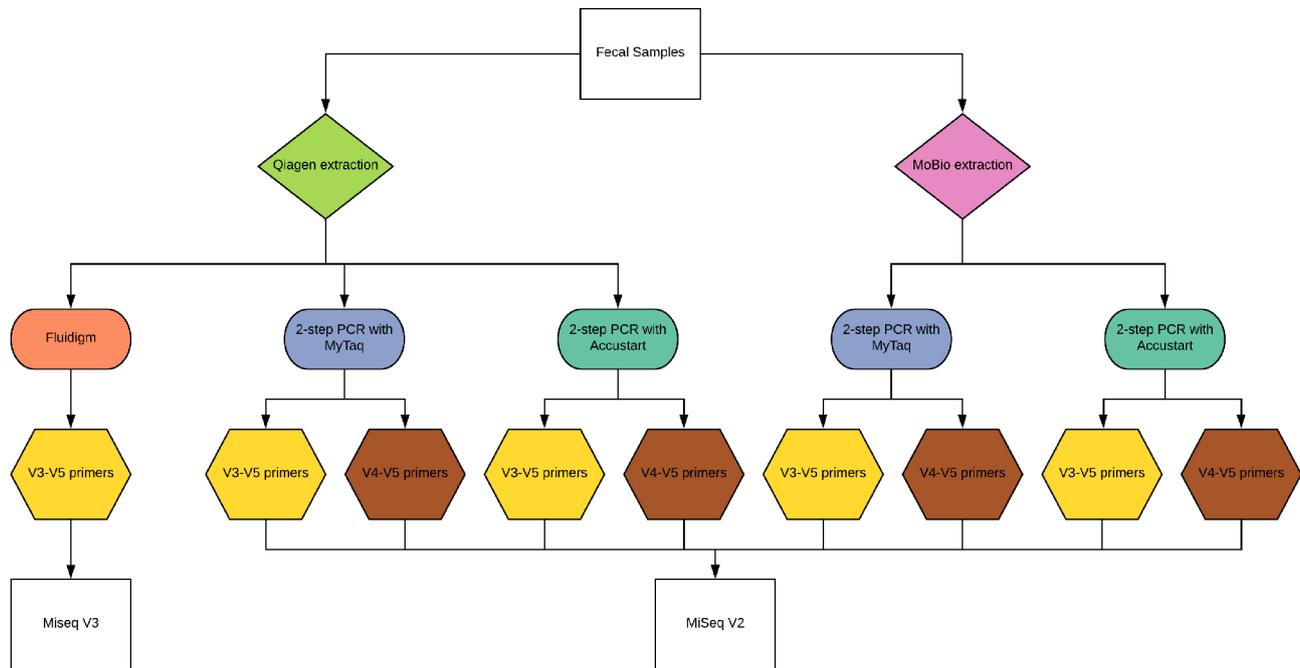
Fig. 1. The experimental design used in this study.

reproducibility and comparability of data across studies. Storage conditions, DNA extraction and amplification methods, choice of PCR primers, sequencing technologies, and data cleaning and analysis pipelines all impact the estimated diversity and composition of a given bacterial community [10–13]. There have been no comparisons between microbiome data produced using Fluidigm-generated amplicon libraries and more commonly used protocols. However, given that the Fluidigm system limits several key aspects of 16S rRNA gene amplicon data generation, namely control over reaction conditions, there are likely to be differences. Understanding the magnitude of these differences is a critical next step in the continued development and adoption of this potentially transformative technology.

In this paper we compare the effects of Fluidigm amplification of the 16S rRNA gene to the effects of DNA extraction kit, polymerase enzyme and primer set selection on estimates of microbial community structure from faecal samples collected from wild white-faced capuchin monkeys (*Cebus capucinus*). We compare samples processed using one of two common commercial DNA extraction kits (QIAamp DNA Stool Mini Kit and MoBio PowerSoil DNA Isolation kit), one of two taq polymerases (MyTaq HS Red Mix and Accustart II PCR ToughMix), one of two primer sets (V3–V5 and V4–V5), and either a two-step PCR protocol or the Fluidigm Access Array (Fig. 1). While all of the differences in sample processing were expected to have some effect on the resulting data, our goal was to determine whether Fluidigm introduces a greater source of bias than other common sources of bias.

METHODS

Sample collection

Faecal samples (*n*=16) were collected from wild white-faced capuchin monkeys at La Suerte Biological Field Station (LSBFS) in north-eastern Costa Rica (10.445° N 83.784° W) between January 2013 and January 2014. Faecal samples were collected in sterile 15 ml tubes, fixed in 90% ethanol and stored at −20 °C prior to transport to University of Illinois at Urbana-Champaign (UIUC). All data collection methods were approved by the University of Illinois IACUC, and LSBFS, MINAET, SINAC and CONAGEBIO in Costa Rica. Appropriate import permits were obtained from the CDC.

DNA extraction

A QIAamp DNA Stool Mini Kit (Qiagen) was used to extract DNA from faecal samples according to the manufacturer's protocol for 'Isolation of DNA from Stool for Human DNA Analysis'. The following modifications were used [14]: samples were mixed with Buffer ASL and then incubated for 1 h at room temperature. Next, Buffer AL was added to the samples, which were then incubated for an additional 20 min at 70 °C prior to elution. During the final elution step, Buffer AE was pre-warmed to 70 °C, added to the samples, and incubated for 20 min at room temperature prior to centrifugation.

A PowerSoil DNA Isolation Kit (MoBio) was also used to extract DNA from the same faecal samples using the manufacturer's protocol. The following modification was used: samples were incubated in Solution C1 at 65 °C for 10 min prior to vortexing horizontally for 10 min. Although Qiagen

Table 1. 16S V4–V5 and 16S V3–V5 primers used in this study. CS1 (5′-ACACTGACGACATGGTTCTACA) and CS2 (5′-TACGGTAGCAGAGACTTGGTCT) linkers were added to the forward and reverse primer sequences, respectively

Primer	Sequence	Reference
16S V4–V5 forward (515fa)	5′-GTGYCAGCMGCCGCGGTAA	[39]
16S V4–V5 reverse (926r)	5′-CCGYCAATTYMTTTRAGTTT	[39]
16S V3–V5 forward (357f)	5′-CCTACGGGAGGCAGCAG	[22]
16S V3–V5 reverse (926r)	5′-CCGTCAATTCMTTTRAGT	[22]

and MoBio extractions were performed at different times, samples were stored at −80 °C between extractions and were homogenized prior to the first extraction. Thus, within-sample variation due to different time points is minimal. The modifications to both extraction protocols were made to minimize the impact of PCR inhibitors, such as tannins and other plant secondary compounds, that are common in wild non-human primate faecal samples.

## PCR amplification and sequencing

All samples extracted with both the Qiagen and the MoBio kits were amplified using a two-step PCR amplification protocol. The same protocol was utilized for all samples, once using MyTaq HS Red Mix (Bioline) and once using AccuStart II PCR ToughMix (QuantaBio). The V4 region of the 16S rRNA gene was amplified using two different primer sets (Table 1) with Fluidigm CS1 and CS2 linker sequences added. The first PCR was carried out in a total volume of 25 µl, consisting of 2 µl of DNA sample (taken directly from the DNA extraction), 12.5 µl of 2× MyTaq or AccuStart, 1.25 µl of 10 µM forward primer, 1.25 µl of 10 µM reverse primer and 8 µl of molecular grade H_2O. The following PCR programme was used: 3 min at 95 °C; 28 cycles of 30 s at 95 °C, 45 s at 55 °C and 45 s at 72 °C; 1 min at 72 °C; and hold at 4 °C indefinitely. PCR products were verified using a 1% agarose gel. A second amplification was performed using the Fluidigm AccessArray primers containing sample-specific barcodes and Illumina sequencing adapters. The second PCR was carried out in a total volume of 20 µl, containing 1 µl of DNA (taken directly from the product of the first PCR), 10 µl of 2× MyTaq HS Red Mix or AccuStart II PCR ToughMix, 4 µl of 0.4 µM Fluidigm AccessArray Barcoded primers for Illumina, and 5 µl of molecular-grade H_2O. The following PCR programme was used: 5 min at 95 °C; eight cycles of 30 s at 95 °C, 30 s at 60 °C and 45 s at 72 °C; and hold at 4 °C indefinitely. PCR products were purified and normalized using a SequalPrep Normalization Plate and were then sequenced on the Illumina MiSeq V2 platform using a 2×250 nt sequencing kit at the University of Illinois Chicago DNA Services Facility.

PCR amplification using the Fluidigm Access Array system was performed for samples extracted with the Qiagen kit at the Roy J. Carver Biotechnology Center at the UIUC. After DNA extracts were quantified on a Qubit, they were diluted to a concentration of 2 ng µl⁻¹. For each sample the following reagents were combined to create a mastermix for amplification: 0.5 µl 10× FastStart Reaction Buffer without $MgCl_2$ (Roche), 0.9 µl 25 mM $MgCl_2$, 0.25 µl DMSO, 0.1 µl 10 mM PCR-grade Nucleotide Mix, 0.05 µl 5 U µl⁻¹ FastStart High Fidelity Enzyme Blend (Roche), 0.25 µl 20× Access Array Loading Reagent (Fluidigm Corp.) and 0.95 µl molecular-grade water. Mastermix was aliquoted to 48 wells of a PCR plate. To each well, 1 µl DNA sample and 1 µl Fluidigm Illumina linkers with unique barcode were added. A 20× primer solution was prepared by adding 2 µl of forward and reverse primers for the V3–V5 region of the 16S rRNA gene (Table 1), 5 µl of 20× Access Array Loading Reagent and water to a final volume of 100 µl, and the solution was aliquoted to a separate 48-well PCR plate. A 4 µl aliquot of sample was loaded in the sample inlets and 4 µl of primer solution was loaded in primer inlets of a previously primed Fluidigm 48.48 Access Array integrated fluidic circuit (IFC). The IFC was placed in an AX controller (Fluidigm Corp.) for microfluidic loading of all primer/sample combinations. Following the loading stage, the IFC plate was loaded on the Fluidigm Biomark HD PCR machine and samples were amplified using the following Access Array cycling programme without imaging: 2 min at 50 °C, followed by 20 min at 70 °C, followed by 10 min at 95 °C, followed by 10 cycles of 15 s at 95 °C, 30 s at 60 °C and 1 min at 72 °C, followed by two cycles of 15 s at 95 °C, 30 s at 80 °C, 30 s at 60 °C and 1 min at 72 °C, followed by eight cycles of 15 s at 95 °C, 30 s at 60 °C and 1 min at 72 °C, followed by two cycles of 15 s at 95 °C, 30 s at 80 °C, 30 s at 60 °C and 1 min at 72 °C, followed by eight cycles of 15 s at 95 °C, 30 s at 60 °C and 1 min at 72 °C, followed by five cycles of 15 s at 95 °C, 30 s at 80 °C, 30 s at 60 °C and 1 min at 72 °C. Following amplification, PCR products were harvested using 2 µl of Fluidigm Harvest Buffer loaded into the sample inlets. PCR products were then transferred to a new 96-well plate and quantified on a Qubit fluorimeter. Samples were pooled in equal amounts according to product concentration. The pooled products were size selected on a 2% agarose E-gel (Life Technologies) and extracted from the isolated gel slice with a Qiagen gel extraction kit (Qiagen). Cleaned, size-selected products were run on an Agilent Bioanalyzer to confirm the appropriate profile and determination of average size. Amplicons were sequenced on the Illumina MiSeq V3 platform using a 2×300 nt sequencing kit at the Roy J. Carver Biotechnology Center at the UIUC. Amplification and sequencing of 16S rRNA genes was performed simultaneously with amplification and sequencing of other genes of interest (*COI*, ITS4, *trnL*, *rbcL*, 12S rRNA).

Sequence processing and analysis

Both forward and reverse reads were analysed separately, allowing us to better understand the effect of primer set and to account for the minimal overlap of the longer primer set. Reads were trimmed and quality-filtered in QIIME [15], resulting in 4841508 forward and 4099102 reverse sequences. There were
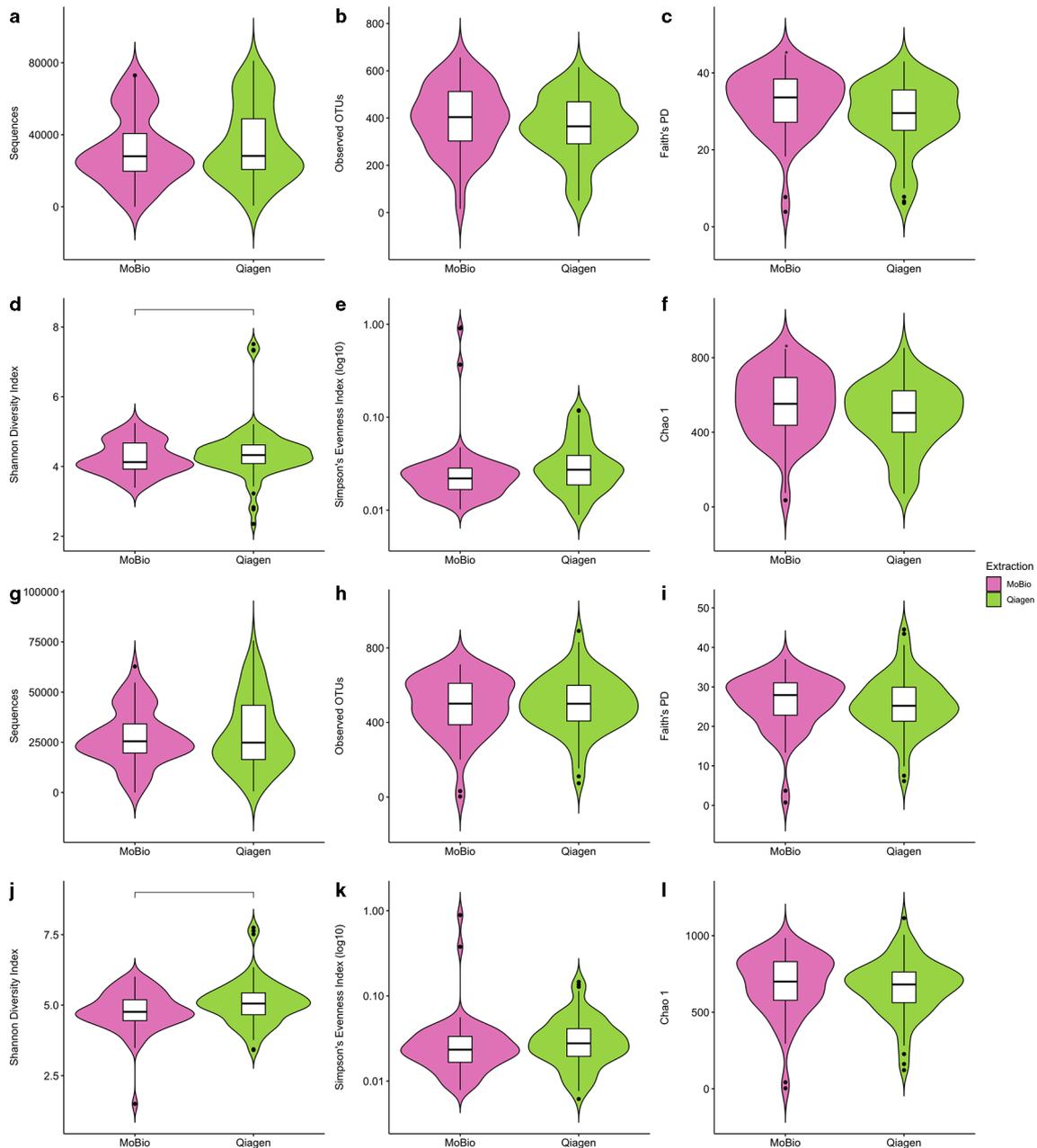
Fig. 2. The effect of extraction method (MoBio vs. Qiagen kits) on the number of sequences per sample, number of observed OTUs per sample and alpha diversity metrics (Faith's phylogenetic diversity, Shannon diversity Index, Simpson's evenness index and Chao 1) for both forward (a–f) and reverse (g–l) reads. Shannon diversity was significantly higher in Qiagen-extracted samples, but other measures of alpha diversity were similar across extraction methods.

an average of 33857 forward sequences (range=31–81071) and 28665 reverse sequences (range=13–75649) per sample. Operational taxonomic units (OTUs) were picked using an open reference method with the sortmerna_sumaclust algorithm and taxonomy was assigned using the Greengenes 13_8 database in QIIME. Samples were rarefied to an even sampling depth of 4800 sequences prior to downstream analysis. Both alpha diversity and beta diversity metrics were calculated in QIIME. Alpha diversity metrics included

Faith's Phylogenetic Distance, Chao 1 and Shannon diversity index. Beta diversity metrics included both unweighted and weighted UniFrac distances.

Statistical analysis

Linear mixed effects models were used to examine the effect of amplification method, primer set and extraction kit on the number of sequences, OTUs, alpha diversity metrics and
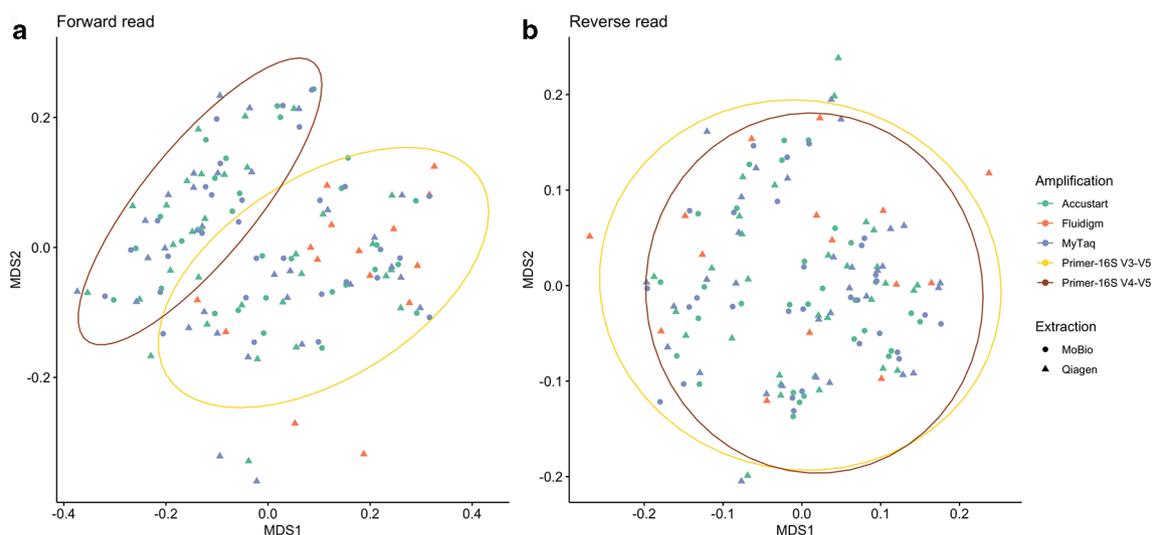
Fig. 3. Nonlinear multidimensional scaling plot visualizing the effects of amplification method, primer set and extraction method on weighted UniFrac distances for forward (a) and reverse (b) reads. Forward reads show a clear separation between primer sets (V3–V5 vs. V4–V5) and amplification methods (Fluidigm vs. Accustart and MyTaq), but not extraction methods.

relative abundances of major phyla. Pairwise comparisons between extraction methods and amplification method treatments were carried out using Tukey contrasts. Permutational analysis of variance (PERMANOVA) was used to assess the effect of both amplification and extraction method on unweighted and weighted UniFrac distances. The original sample was included in all models to control for repeated measures. LEfSe was used to determine which taxa discriminated each amplification and extraction method [16], using an linear discriminant analysis (LDA) effect size threshold of 3.

RESULTS

Effect of extraction method on alpha diversity

There were no significant differences in the number of sequences or observed OTUs between extraction methods (Fig. 2). Similarly, there was no significant effect of extraction method on most measures of alpha diversity when analysing

forward or reverse reads (Fig. 2). However, Shannon diversity was higher in Qiagen samples (forward reads: $F_{1,123}=7.070$, $P=0.009$; reverse reads: $F_{1,123}=5.925$, $P=0.016$).

Effect of extraction method on beta diversity

Extraction method had a significant effect on gut microbial community composition (Figs 3 and 4). PERMANOVA results indicated that extraction method had a significant influence on unweighted and weighted UniFrac distances when analysing both the read 1 and read 2 sequences (Table 2).

Effect of extraction method on individual taxa

Extraction method had a significant effect on the relative abundance of both *Actinobacteria* (forward reads: $F_{1,123}=5.925$, $P=0.016$; reverse reads: $F_{1,123}=8.917$, $P=0.003$) and *Firmicutes* (forward reads: $F_{1,123}=8.423$, $P=0.004$; reverse reads: $F_{1,123}=5.757$, $P=0.018$) (Fig. 5). The Qiagen extraction method had higher relative abundances of *Actinobacteria* and lower
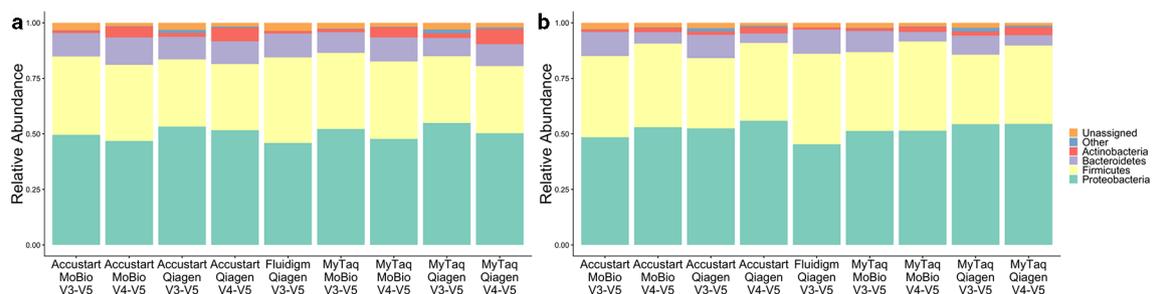
Fig. 4. Differences in the relative abundance of major phyla across different amplification methods (Accustart, Fluidigm, and MyTaq), primer sets (V3–V5 and V4–V5) and extraction methods (Qiagen and MoBio kits) for forward (a) and reverse (b) reads. All factors influenced estimates of gut microbial community composition.

Table 2. PERMANOVA results for full models examining the effect of amplification method, extraction method, primer set and original faecal sample identity on both unweighted and weighted UniFrac distances calculated separately from forward sequences and reverse sequences

		Effect of:			
		Amplification	**Extraction**	**Primer set**	**Sample**
Forward read	**Unweighted UniFrac**	F=5.688, R^2=0.067, P<0.001	F=2.550, R^2=0.015, P<0.001	F=19.809, R^2=0.117, P<0.001	F=2.959, R^2=0.017, P<0.001
	Weighted UniFrac	F=4.134, R^2=0.045, P<0.001	F=2.313, R^2=0.013, P=0.039	F=33.110, R^2=0.181, P<0.001	F=5.895, R^2=0.032, P<0.001
Reverse read	**Unweighted UniFrac**	F=7.070, R^2=0.088, P<0.001	F=3.483, R^2=0.022, P<0.001	F=7.022, R^2=0.044, P<0.001	F=3.260, R^2=0.020, P<0.001
	Weighted UniFrac	F=2.295, R^2=0.030, P=0.011	F=2.043, R^2=0.013, P=0.059	F=5.775, R^2=0.038, P<0.001	F=7.653, R^2=0.050, P<0.001

relative abundances of *Firmicutes*. Extraction method did not have a significant effect on the relative abundance of *Bacteroidetes* or *Proteobacteria*. In the LEfSe analysis of differences in extraction method, *Alcaligenaceae*, *Burkholderiaceae*, *Pseudomonadaceae*, *Xanthomonadaceae* and an unclassified family within OD1 were differentially abundant in Qiagen samples when analysing the forward read, while only *Alcaligenaceae*, *Burkholderiaceae*, *Pseudomonadaceae* and *Xanthomonadaceae* were differentially abundant in Qiagen samples when analysing the reverse read (Fig. 6). When examining forward reads, MoBio-extracted samples were discriminated by *Lachnospiraceae*, *Micrococcaceae* and an unclassified family within *Proteobacteria*, while an unclassified family within *Proteobacteria* discriminated MoBio-extracted samples when examining reverse reads (Fig. 6).

Effect of amplification method on alpha diversity

Amplification method did not have a significant effect on the number of sequences per sample but did significantly affect the number of observed OTUs per sample when analysing forward reads ($F_{2,123}$=31.637, P<0.001), but not reverse reads (Fig. 7). Pairwise comparisons of the forward reads showed significantly lower numbers of observed OTUs in the Fluidigm treatment compared with Accustart (forward: z=−4.129, P_{adj}<0.001) and MyTaq samples (z=−5.437, P_{adj}<0.001) (Fig. 7).

Amplification method also had a significant effect on three alpha diversity metrics when examining the forward sequences (Faith's PD: $F_{2,123}$=37.579, P<0.001; Chao 1: $F_{2,123}$=43.902, P<0.001; Shannon Index: $F_{2,123}$=5.535, P=0.005) (Fig. 7). Chao 1 scores differed significantly between amplification methods when analysing reverse sequences ($F_{2,123}$=3.481, P=0.034) (Fig. 8). Pairwise comparisons of read 1 data only showed that Fluidigm samples had significantly lower Faith's PD compared with both MyTaq and Accustart treatments (z=−6.411, P_{adj}<0.001 and z=−5.664, P_{adj}<0.001). Pairwise comparisons of forward reads show that Fluidigm samples had significantly lower Chao 1 scores compared with Accustart samples (z=−5.073, P_{adj}<0.001) and MyTaq samples had higher Chao 1 scores than Accustart samples (z=3.572, P_{adj}=0.0258) and Fluidigm samples (z=6.556, P_{adj}<0.001). Pairwise comparisons of forward reads showed significantly lower

Shannon Index scores in Fluidigm samples compared with both MyTaq and Accustart samples (z=−3.111, P_{adj}=0.005 and z=−3.653, P_{adj}<0.001), but found no significant differences between Shannon Indexes of MyTaq and Accustart samples. Simpson's evenness index did not differ between amplification methods when analysing either forward or reverse reads.

Effect of amplification method on beta diversity

Amplification method also affected estimates of gut microbial community composition (Figs 3 and 4). PERMANOVA results indicated that amplification method had a significant influence on both unweighted and weighted UniFrac distances when examining either read (Table 2). Pairwise PERMANOVAs showed that estimates of community composition differed significantly between Fluidigm and Accustart (forward reads: unweighted UniFrac: F=8.376, R^2=0.100, P_{adj}<0.001; weighted Unifrac: F=5.918, R^2=0.073, P_{adj}=0.001; reverse reads: unweighted UniFrac: F=10.649, R^2=0.124, P_{adj}<0.001; weighted Unifrac: F=3.564, R^2=0.045, P_{adj}=0.011) and Fluidigm and MyTaq (forward reads: unweighted UniFrac: F=8.025, R^2=0.097, P_{adj}<0.001; weighted Unifrac: F=5.342, R^2=0.066, P_{adj}=0.001; reverse reads: unweighted UniFrac: F=10.445, R^2=0.122, P_{adj}<0.001; weighted Unifrac: F=3.260, R^2=0.042, P_{adj}=0.012) samples, but did not differ between Accustart and MyTaq samples (forward reads: unweighted UniFrac: F=0.718, R^2=0.006, P_{adj}=0.918; weighted Unifrac: F=0.307, R^2=0.003, P_{adj}=0.968; reverse reads: unweighted UniFrac: F=1.407, R^2=0.011, P_{adj}=0.052; weighted Unifrac: F=0.423, R^2=0.003, P_{adj}=0.880).

Effect of amplification method on individual taxa

Amplification method had a significant effect on the relative abundance of *Actinobacteria* (forward reads: $F_{2,123}$=5.659, P=0.005; reverse reads: $F_{2,123}$=4.293, P=0.016), *Bacteroidetes* (forward reads: $F_{2,123}$=3.066, P=0.050; reverse reads: $F_{2,123}$=6.241, P=0.003) and *Firmicutes* (forward reads: $F_{2,123}$=3.497, P=0.033; reverse reads: n.s.) (Fig. 5). Pairwise comparisons of forward reads showed that MyTaq samples had significantly lower relative abundances of *Bacteroidetes* compared with both Accustart (z=−2.374, P=0.0438) and Fluidigm samples (z=−2.516, P=0.030) and that Fluidigm
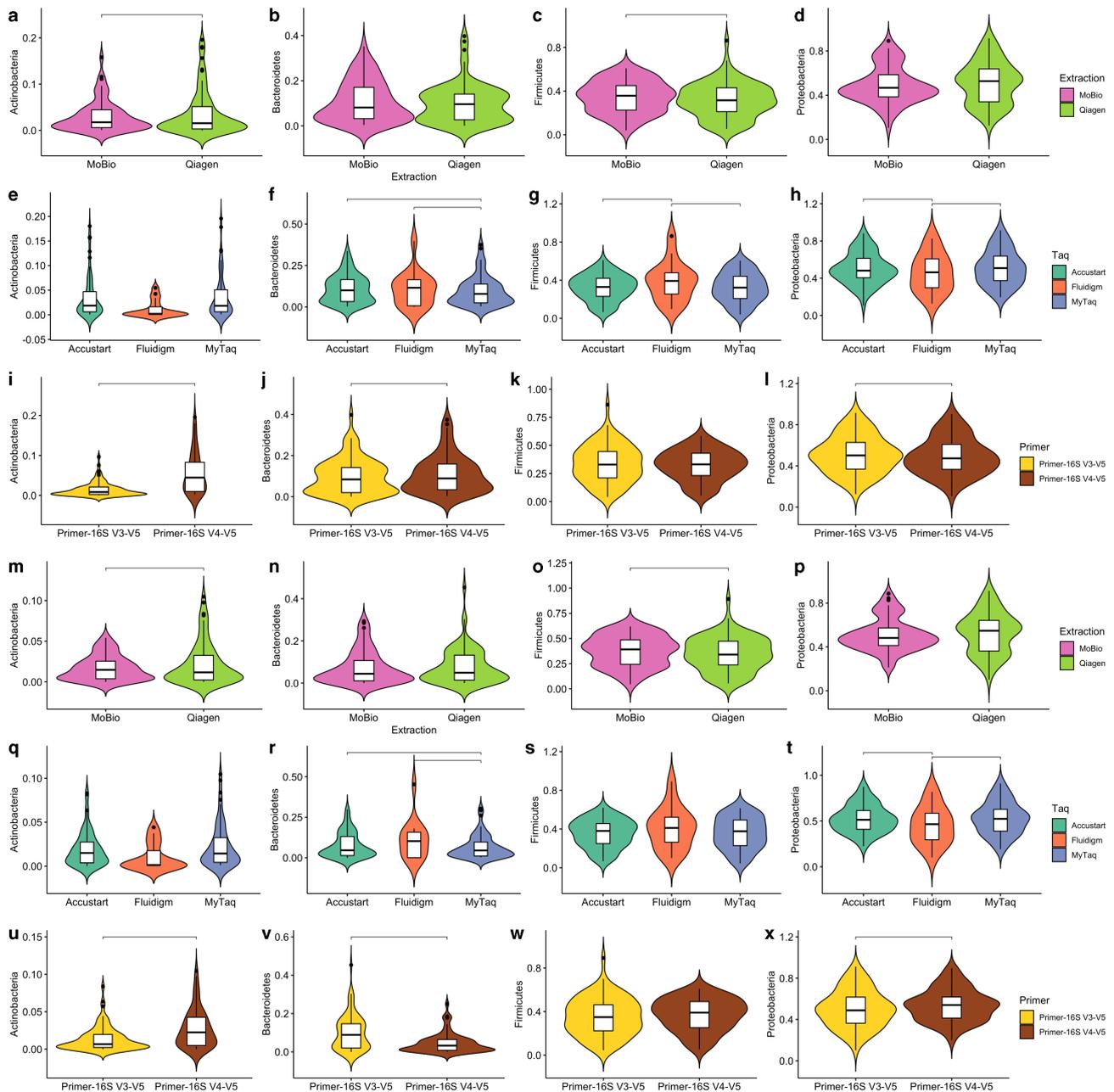
Fig. 5. Relative abundance of major phyla when using each amplification method (Accustart, Fluidigm, MyTaq), primer set (V3–V5) and extraction method (Qiagen, MoBio) for forward (a–l) and reverse (m–x) reads. Relative abundances of *Actinobacteria* and *Firmicutes* differed significantly between extraction methods; *Bacteroidetes*, *Firmicutes* and *Proteobacteria* differed significantly between amplification methods; and *Actinobacteria*, *Bacteroidetes* and *Proteobacteria* differed significantly between primer sets.

samples had significantly higher relative abundances of *Firmicutes* than either MyTaq (forward reads: z=3.069, P=0.006) or Accustart samples (forward reads: z=3.122, P=0.005). Amplification method had a significant effect on the relative abundance of *Proteobacteria*, but only when analysing the reverse reads (reverse reads: $F_{2,123}$=4.246, P=0.017). Pairwise comparisons showed that Fluidigm samples had lower relative abundances than both Accustart (forward reads: z=−2.648, P=0.020; reverse reads: z=−2.595, P=0.024) and MyTaq

samples (forward reads: z=−3.074, P=0.006; reverse reads: z=−2.805, P=0.013).

The results of the LEfSe analysis showed multiple families that were overrepresented in all treatments (Fig. 8). The specific families that were overrepresented differed based on whether forward or reverse reads were analysed. When analysing forward reads, *Ruminococcaeae* and *Coriobacteriaceae*, as well as unclassified families within *Bacteroidetes*, *Tenericutes*
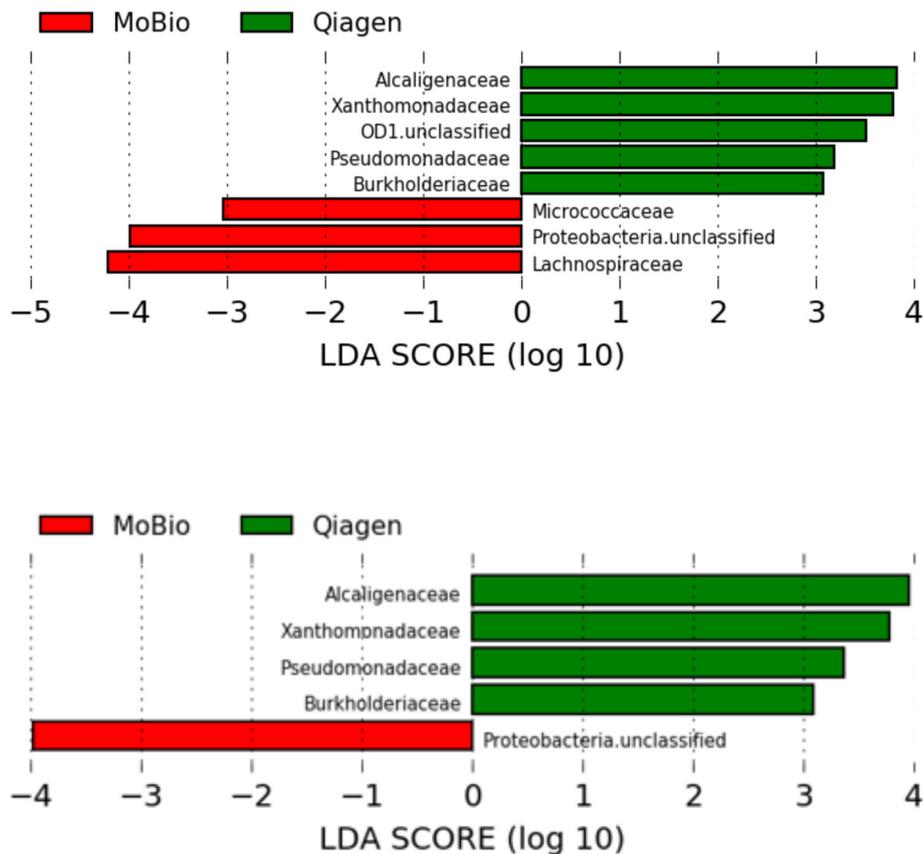
Fig. 6. LEfSe results indicating which taxa discriminate each extraction method when analysing forward (top) and reverse (bottom) reads.

and *Cyanobacteria* were overrepresented in MyTaq samples; *Acetobacteraceae, Bacteroidaceae, Burkholderiaceae, Rhizobiaceae, Sphingomonadaceae* and an unclassified family within *Firmicutes* were overrepresented in Fluidigm samples; and *Bifidobacteriaceae, Campylobacteraceae, Chitinophagaceae, Coxiellaceae, Helicobacteraceae, Neisseriaceae, Succinivibrionaceae* and an unclassified family within *Proteobacteria* were overrepresented in Accustart samples. When analysing reverse reads, *Bifidobacteriaceae, Coriobacteriaceae, Ruminococcaceae* and an unclassified family within *Tenericutes* were overrepresented in MyTaq samples; *Burkholderiaceae, Rhizobiaceae* and *Sphingomonadaceae* were overrepresented in Fluidigm samples; and *Bacteroidaceae, Campylobacteraceae, Chitinophagaceae, Helicobacteraceae, Neisseriaceae, Succinivibrionaceae* and an unclassified family within *Proteobacteria* were overrepresented in Accustart samples.

Effect of primer set on alpha diversity

Alpha diversity was significantly affected by which primer set was used (Fig. 9). While the total number of sequences did not differ between primer sets, the V3–V5 primer set had significantly fewer observed OTUs (forward reads: $F_{1,123}$=40.915, P<0.001; reverse reads: $F_{1,123}$=6.047, P=0.015). Samples amplified with the V3–V5 primer also had significantly lower Faith's

PD (forward reads: $F_{1,123}$=18.591, P<0.001; reverse reads: $F_{1,123}$=10.683, P=0.001) and Chao 1 scores (forward reads: $F_{1,123}$=44.709, P<0.001; reverse reads: n.s.). Shannon diversity was higher in V3–V5 samples, but only when analysing the reverse reads ($F_{1,123}$=25.767, P<0.001). Simpson's evenness index did not differ between primer sets.

Effect of primer set on beta diversity

Estimates of gut microbial community composition were significantly different between primer sets (Figs 3 and 4). PERMANOVAs showed a significant influence of primer set on both unweighted and weighted UniFrac distances (Table 2). The magnitude of this effect was much higher for forward sequences.

Effect of primer set on individual taxa

Samples amplified with the V4–V5 primer set had higher relative abundances of *Actinobacteria* (forward reads: $F_{1,123}$=65.730, P<0.001; reverse reads: $F_{1,123}$=21.969, P<0.001) (Fig. 5). *Bacteroidetes* and *Proteobacteria* also differed significantly between primer sets (Fig. 5). Analysing the forward reads indicated that the relative abundances of *Bacteroidetes* were higher and *Proteobacteria* were lower when using the
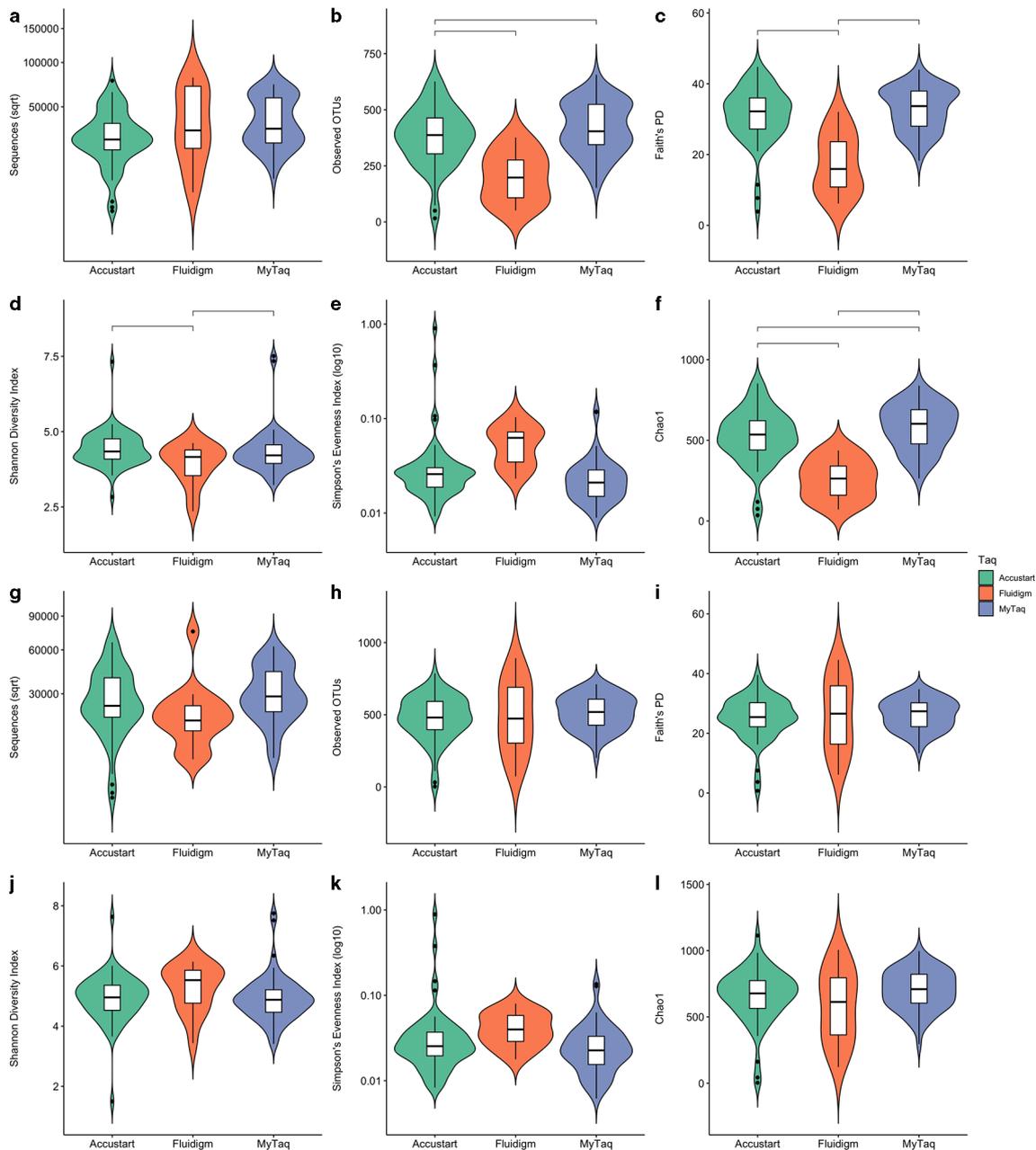
Fig. 7. Effect of amplification method (Accustart, Fluidigm, MyTaq) on the number of sequences per sample, number of OTUs per sample and alpha diversity metrics (Faith's phylogenetic diversity, Shannon diversity Index, Simpson's evenness index, and Chao 1) for both forward (a–f) and reverse (g–l) reads. Fluidigm-amplified samples had significantly lower alpha diversity than Accustart- or MyTaq-amplified samples when analysing forward reads.

V4–V5 primer set (*Bacteroidetes*: $F_{1,123}$=5.835, P=0.017; *Proteobacteria*: $F_{1,123}$=4.520, P=0.036). However, analysing the reverse reads gave the opposite result – relative abundances of *Bacteroidetes* were lower and *Proteobacteria* were higher when using the V4–V5 primers (*Bacteroidetes*: $F_{1,123}$=45.186, P<0.001; *Proteobacteria*: $F_{1,123}$=1.091, P=0.298). The relative abundances of *Firmicutes* did not differ significantly between primer sets.

LEfSe results indicated several phylogenetically diverse taxa were differentially abundant when using different primer sets (Fig. 10). When analysing forward reads, *Bifidobacteriaceae*, *Lachnospiraceae*, *Micrococcaceae*, *Neisseriaceae*, *Ruminococcaceae* and an unclassified family within *Proteobacteria* discriminated the V4–V5 primer set. When analysing forward reads, *Coxiellaceae*, *Enterococcaceae*, and unclassified families within *Firmicutes* and *Tenericutes* discriminated
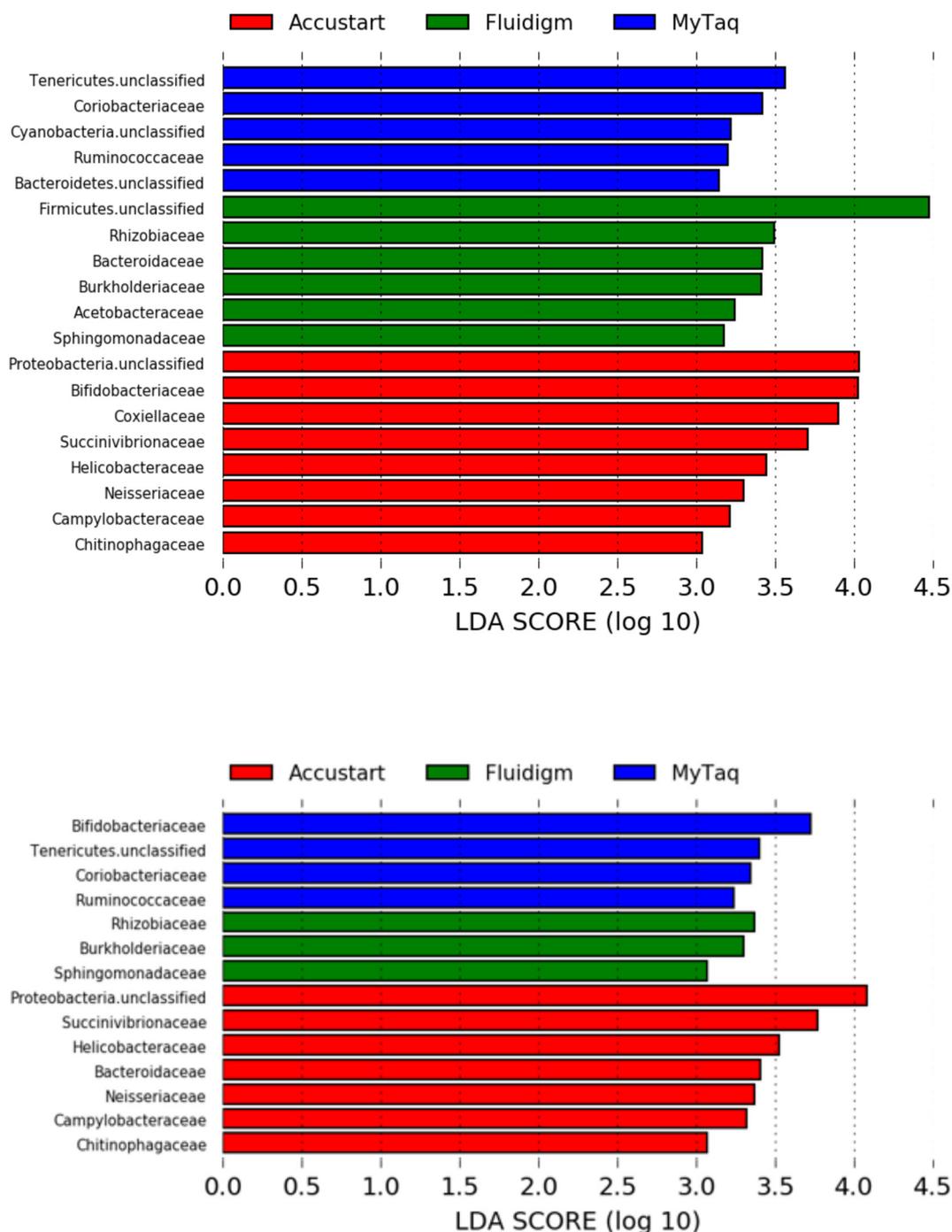
Fig. 8. LEfSe results indicating which taxa discriminate each amplification method when analysing forward (top) and reverse (bottom) reads.

the V3–V5 primer set. Additional taxa distinguished primer sets when analysing reverse reads – the V4–V5 samples were discriminated by *Beutenbergiaceae, Bifidobacteriaceae, Chthoniobacteraceae, Eubacteriaceae, Neisseriaceae, Paraprevotellaceae, Ruminococcaceae, Scytonemataceae,* and unclassified families within *Proteobacteria, Firmicutes, Bacteroidetes* and *Cyanobacteria*; while *Coxiellaceae,*

Prevotellaceae and an unclassified family within *Tenericutes* distinguished the V3–V5 primer set.

DISCUSSION

Extraction method, choice of taq polymerase and primer set all had significant effects on gut microbial community
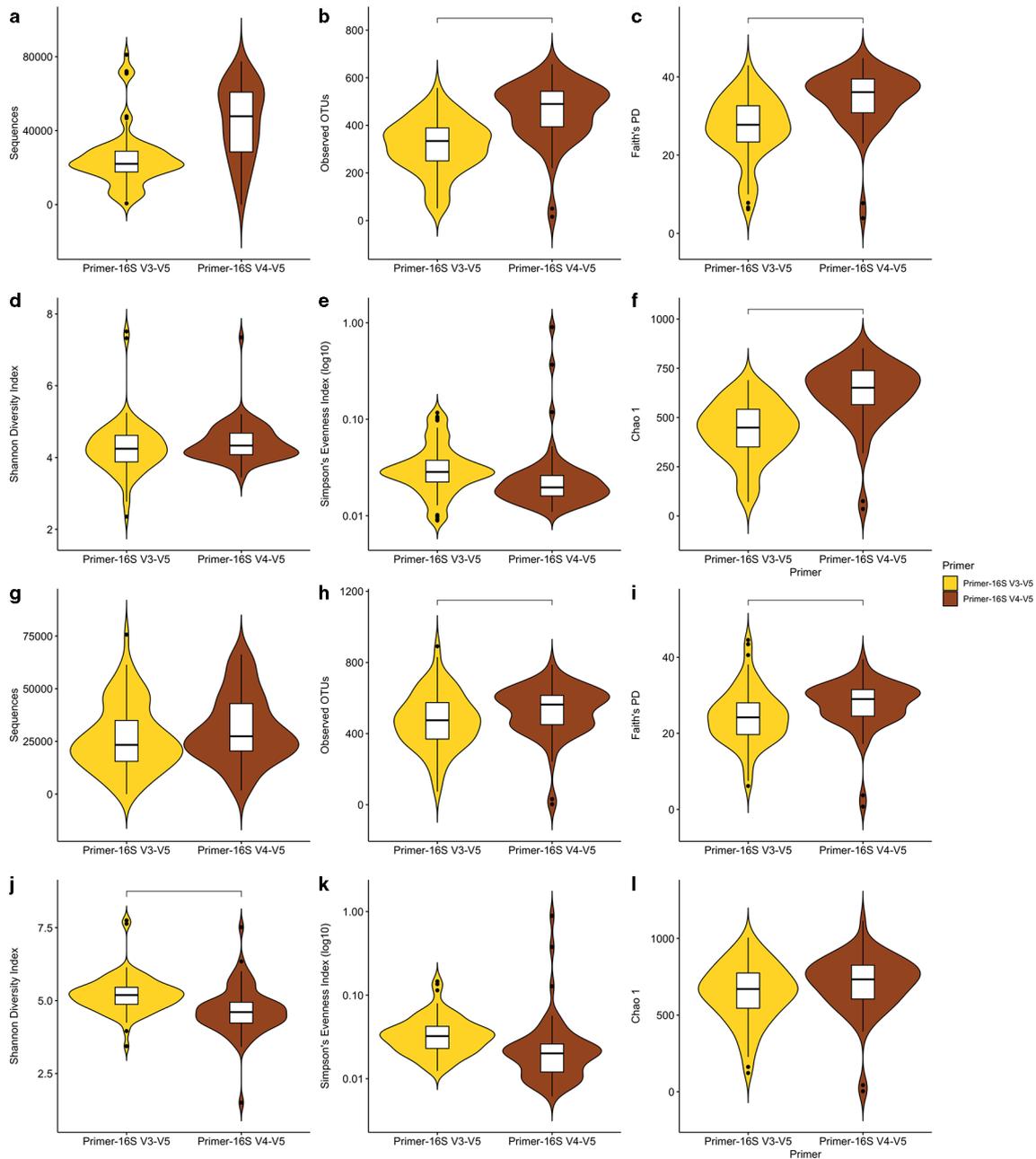
Fig. 9. Effect of primer set (V3–V5 and V4–V5) on the number of sequences per sample, number of OTUs per sample and alpha diversity metrics (Faith's phylogenetic diversity, Shannon diversity Index, Simpson's evenness index and Chao 1) for both forward (a–f) and reverse (g–l) reads. The V4–V5 primer set had higher diversity when analysing the forward reads, but lower diversity when analysing the reverse reads.

characterization. Differences between samples amplified on the Fluidigm Access Array system and samples amplified using a more common two-step PCR protocol, as well as differences between primer sets were substantial. In fact, primer set or amplification method explained the largest proportion of the variance in all models, while differences between samples accounted for more of the variation than extraction method in most models (Table 1). These findings suggest that the use

of current Fluidigm technology and choice of primers for microbiome analyses must be carefully considered.

Effects of extraction kit

Several studies have shown that extraction method can have a significant influence on DNA yield, DNA purity, amplification and sequencing success, and taxonomic community profiles [17–21]. However, others indicate that the proportion of
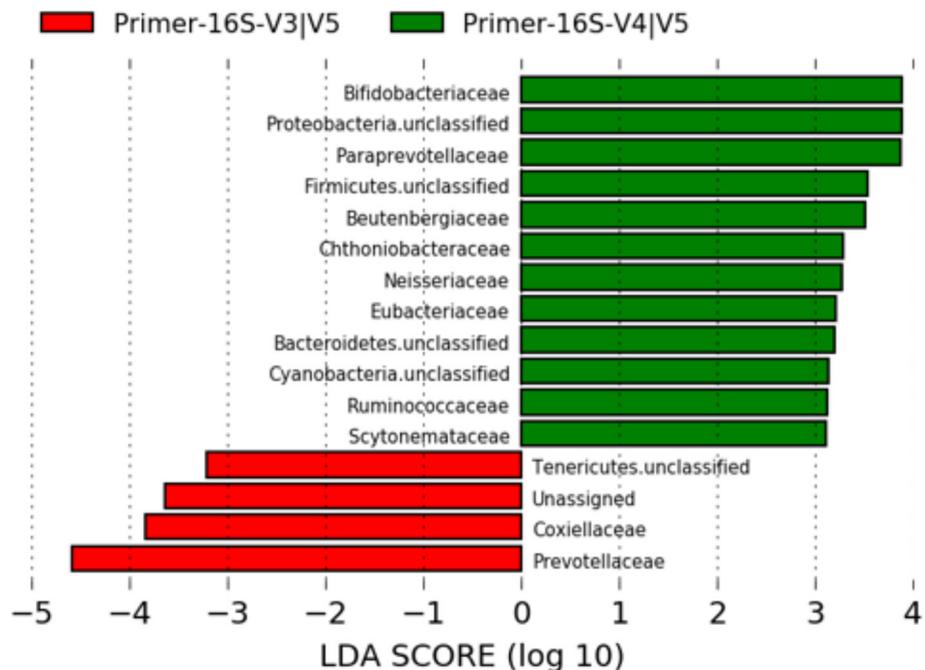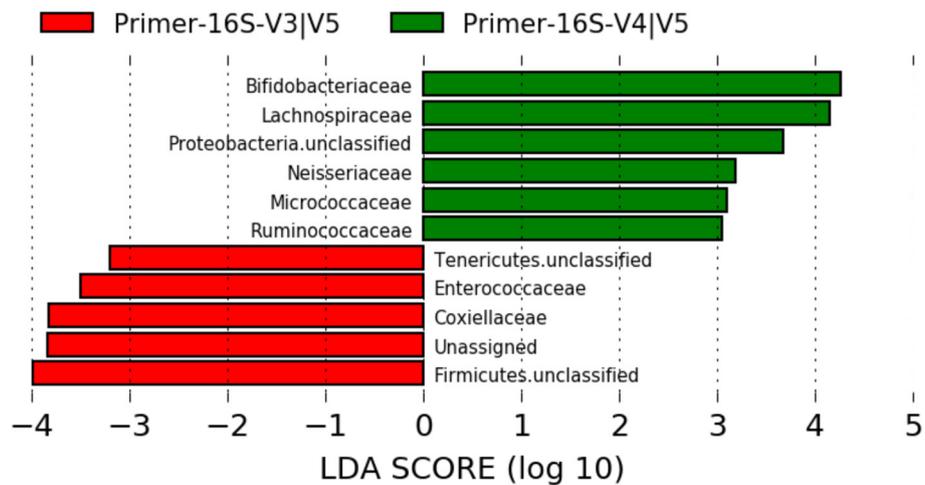
Fig. 10. LEfSe results indicating taxa that discriminate each primer set when analysing forward (top) and reverse (bottom) reads.

variation explained by extraction method is much lower than the proportion explained by other experimental factors, such as individual, population or body site [22–24]. Our findings replicate those reported by Wu *et al.* [22] and suggest that there are small, but potentially important, differences between extraction methods. Specifically, extraction method appears to limit the ability to detect the presence or absence of rarer taxa and alters the relative abundance of some microbial taxa, but it does not obscure differences between samples [18, 23, 25].

A potential major driver of the impact of extraction kit is bead-beating. Previous studies report that bead-beating is a better method for recovering DNA from Gram-positive bacteria [18, 19, 26], but see [24]. Our results support this evaluation. Bacterial taxa that discriminated the Qiagen kit tended to be Gram-negative (*Alcaligenaceae*, *Burkholderiaceae*, *Pseudomonadaceae* and *Xanthomonadaceae*), while families that distinguished samples extracted with the MoBio kit tended to be Gram-positive (*Lachnospiraceae* and *Micrococcaceae*). However, some Gram-negative taxa discriminated

samples extracted with the MoBio kit (the unclassified family of *Proteobacteria*). These patterns are less clearly explained by bead-beating but have also been reported in previous studies [21].

Effects of taq polymerase

With respect to taq polymerase, our results also generally agreed with the existing literature. A few studies have shown that the choice of taq polymerase can alter estimates of bacterial community structure and diversity [27, 28] and that using a proofreading taq polymerase can increase the rates of PCR artefacts and chimeras at high cycle numbers [29]. However, other studies have found that the effects of taq polymerase on 16S community profiling are slight, particularly when compared with the effects of other methodological differences [20, 30]. The differences between Accustart and MyTaq in our experiment were minimal. Differences between MyTaq and the Fluidigm polymerase or Accustart and the Fluidigm polymerase accounted for the majority of pairwise differences in alpha diversity, beta diversity and the abundance of specific taxa.

Samples amplified on the Fluidigm Access Array had significantly lower alpha diversity using most metrics and had a significant effect on the estimated taxonomic composition of the microbial community. Due to the constraints of the Fluidigm system, it is difficult to determine whether the observed patterns are a result of differences in PCR cycle number (33 for Fluidigm vs. 36 for the two-step PCR method), lower template concentration for Fluidigm samples, differences in reaction reagents and conditions (including taq polymerase), slightly lower sequencing depth for Fluidigm samples, the fact that sequencing of 16S rRNA gene amplicons prepared with Fluidigm was performed simultaneously with amplicons of other marker genes of interest prepared on the same Fluidigm run, differences in sequencing chemistry (Illumina V2 vs. V3), or a combination of all of these factors, as all of these may influence estimates of community composition [27, 30, 31]. It is likely that the small effects of all of these factors combined result in a large influence on estimates of community composition and are also compounded by suboptimal PCR conditions for these primers. However, regardless of the mechanism, microbial community profiles based on Fluidigm-generated amplicon libraries do not seem to be comparable to those generated with more commonly used methods.

Effects of primer set

There is a general consensus that different primer sets and target regions of 16S vary in their taxonomic coverage, taxonomic specificity and compatibility with various sequencing platforms [10–13, 32]. We saw a substantial effect of primer set on most alpha diversity metrics, both weighted and unweighted UniFrac distances, and the relative abundance of most major phyla. The observed differences are intensified when examining only the forward reads, unsurprisingly, as the reverse primer sequences in both primer sets were similar. These results are similar to the findings of other studies comparing multiple primer sets that find both differences in estimates of taxonomic composition and alpha diversity [33]. Additionally, the length of the V3–V5 primer set combined with our choice of sequencing platform limited our ability to analyse paired sequence reads.

Effects of sequencing chemistry

While we were not able to formally test for the impact of the differences in sequencing chemistry between the Fluidigm treatment and the other treatment groups, we feel it is unlikely that sequencing chemistry is causing the observed differences. While most studies of the effects of sequencing technology have focused on cross-platform differences [30, 34–37], comparisons of the V2 and V3 MiSeq chemistries result in similar estimates of taxonomic composition of microbial communities [38]. The V3 MiSeq chemistry used for the Fluidigm samples does have a longer read length, higher cluster density and increased sequencing depth, which should increase alpha diversity. In contrast, samples amplified on the Fluidigm platform in the present study have decreased alpha diversity.

Overall conclusions and recommendations

In summary, while the results produced by many of the methods we tested are generally comparable, researchers should still be aware of the limitations and biases that methodological differences can introduce into their results. Choice of primer set can have a large influence on estimates of microbial community composition, and researchers should carefully weigh the taxonomic biases of a primer set against the expected composition of the microbial community of interest. The difference between the Fluidigm Access Array amplification method and two-step PCR protocols was substantial and suggests that researchers should cautiously consider whether the benefits of using the Fluidigm Access Array for amplicon library preparation outweigh the downsides. Further benchmarking, including testing the V4–V5 primer set on the Fluidigm platform, to identify which specific factors are contributing to the differences between the amplification methods will be important. In addition, the development of primers that are more efficient with the allowable Fluidigm reaction conditions and give comparable results to the V4–V5 primer set or technological development that enables variation in reaction temperatures between individual wells within a single microfluidic PCR run may be necessary to increase the comparability of results. Microfluidic PCR has the potential to greatly increase the types of data that can be generated without increased time or costs, but further optimization is necessary prior to integrating microfluidic PCR into microbiome protocols on a large scale.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Author contributions

E.K.M. (ORCiD: 0000-0001-5446-8563) contributed to conceptualization, methodology, formal analysis, investigation, writing of the original draft and funding. R.S.M. (ORCiD: 0000-0002-1484-0292) contributed to the methodology, resources, supervision, review and editing of the manuscript, and funding. K.R.A. (ORCiD: 0000-0003-2722-9414) contributed to the conceptualization, methodology, resources, supervision, review and editing of the manuscript, project administration and funding.

Ethical statement

Data collection methods were approved by the University of Illinois IACUC, and La Suerte Biological Field Station, MINAET, SINAC and CONAGEBIO in Costa Rica. Appropriate import permits were obtained from the CDC.

References

1. Brown SP, Ferrer A, Dalling JW, Heath KD. Don't put all your eggs in one basket: a cost-effective and powerful method to optimize primer choice for rRNA environmental community analyses using the Fluidigm Access Array. *Mol Ecol Resour* 2016;16:946–956.

2. Menke S, Wasimuddin S, Meier M, Melzheimer JÃrg, Mfune JKE *et al.* Oligotyping reveals differences between gut microbiomes of free-ranging sympatric Namibian carnivores (Acinonyx jubatus, canis mesomelas) on a bacterial species-like level. *Front Microbiol* 2014;5:1–12.

3. Menke S, Meier M, Mfune JKE, Melzheimer J, Wachter B *et al.* Effects of host traits and land-use changes on the gut microbiota of the Namibian black-backed jackal (canis mesomelas). *FEMS Microbiol Ecol* 2017;93:1–16.

4. Menke S, Meier M, Sommer S. Shifts in the gut microbiome observed in wildlife faecal samples exposed to natural weather conditions: lessons from time-series analyses using next-generation sequencing for application in field studies. *Methods Ecol Evol* 2015;6:1080–1087.

5. Moonsamy PV, Williams T, Bonella P, Holcomb CL, Höglund BN *et al.* High throughput HLA genotyping using 454 sequencing and the Fluidigm access Array™ system for simplified amplicon library preparation. *Tissue Antigens* 2013;81:141–149.

6. Mallott EK, Amato KR, Garber PA, Malhi RS. Influence of fruit and invertebrate consumption on the gut microbiota of wild white-faced capuchins (*Cebus capucinus*). *Am J Phys Anthropol* 2018;165:576–588.

7. Wasimuddin BSD, Tschapka M, Page R, Rasche A *et al.* Astrovirus infections induce age-dependent dysbiosis in gut microbiomes of bats. *Isme J* 2018:Epub ahead of print.

8. Venable EB, Bland SD, Holscher HD, Swanson KS. Effects of air travel stress on the canine microbiome: a pilot study. *Int J Vet Heal Sci Res* 2016;4:132–139.

9. Frossard A, Donhauser J, Mestrot A, Gygax S, Bååth E *et al.* Long- and short-term effects of mercury pollution on the soil microbiome. *Soil Biology and Biochemistry* 2018;120:191–199.

10. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C *et al.* Best practices for analysing microbiomes. *Nat Rev Microbiol* 2018;16:410–422.

11. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC *et al.* Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* 2012;13:47–58.

12. Hamady M, Knight R. Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Res* 2009;19:1141–1152.

13. Pollock J. The madness of microbiome : attempting to find consensus. *Appl Environ Microbiol* 2018;84:e02627–17.

14. Pickett SB, Bergey CM, Di Fiore A, Fiore ADI. A metagenomic study of primate insect diet diversity. *Am J Primatol* 2012;74:622–631.

15. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;7:335–336.

16. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011;12:R60.

17. Peng X, Yu K-Q, Deng G-H, Jiang Y-X, Wang Y *et al.* Comparison of direct boiling method with commercial kits for extracting fecal microbiome DNA by Illumina sequencing of 16S rRNA tags. *J Microbiol Methods* 2013;95:455–462.

18. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilić-Stojanović M *et al.* Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods* 2010;81:127–134.

19. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* 2012;7:e33865.

20. Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM *et al.* The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol* 2015;15:1–14.

21. Maukonen J, Simões C, Saarela M. The currently used commercial DNA-extraction methods give different results of clostridial and actinobacterial populations derived from human fecal samples. *FEMS Microbiol Ecol* 2012;79:697–708.

22. Wu GD, Lewis JD, Hoffmann C, Chen Y-Y, Knight R *et al.* Sampling and pyrosequencing methods for characterizing bacterial communities in the human gut using 16S sequence tags. *BMC Microbiol* 2010;10:206.

23. Mackenzie BW, Waite DW, Taylor MW. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Front Microbiol* 2015;6:1–11.

24. Wesolowska-Andersen A, Bahl MI, Carvalho V, Kristiansen K, Sicheritz-Pontén T *et al.* Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* 2014;2:19–11.

25. Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM *et al.* The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLoS One* 2014;9:e88982–88989.

26. de Boer R, Peters R, Gierveld S, Schuurman T, Kooistra-Smid M *et al.* Improved detection of microbial DNA after bead-beating before DNA isolation. *J Microbiol Methods* 2010;80:209–211.

27. Wu J-Y, Jiang X-T, Jiang Y-X, Lu S-Y, Zou F *et al.* Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiol* 2010;10:255.

28. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A *et al.* Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol* 2016;34:942–949.

29. Ahn JH, Kim BY, Song J, Weon HY. Effects of PCR cycle number and DNA polymerase type on the 16S rRNA gene pyrosequencing analysis of bacterial communities. *J Microbiol* 2012;50:1071–1074.

30. D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R *et al*. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* 2016;17.

31. Pinto AJ, Raskin L. Pcr biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS One* 2012;7:e43093.

32. Kennedy K, Hall MW, Lynch MDJ, Moreno-Hagelsieb G, Neufeld JD. Evaluating bias of illumina-based bacterial 16S rRNA gene profiles. *Appl Environ Microbiol* 2014;80:5717–5722.

33. Fouhy F, Clooney AG, Stanton C, Claesson MJ, Cotter PD. 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol* 2016;16:1–13.

34. Salipante SJ, Kawashima T, Rosenthal C, Hoogestraat DR, Cummings LA *et al*. Performance comparison of illumina and ion torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Appl Environ Microbiol* 2014;80:7583–7591.

35. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT *et al*. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res* 2015;43:e37.

36. Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J. Analysis, optimization and verification of illumina-generated 16S rRNA gene amplicon surveys. *PLoS One* 2014;9:e94249.

37. Whon TW, Chung W-H, Lim MY, Song E-J, Kim PS *et al*. The effects of sequencing platforms on phylogenetic resolution in 16 S rRNA gene profiling of human feces. *Sci Data* 2018;5:1–15.

38. Fadrosh DW, Ma B, Gajer P, Sengamalay N, Ott S *et al*. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* 2014;2:6–7.

39. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G *et al*. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* 2016;1:e0009–0015.