Reproducibility and repeatability of six high-throughput 16S rDNA sequencing protocols for microbiota profiling.

Authors:

Sajan Raju[1,2], sajan.raju@helsinki.fi

Sonja Lagström[3], sonja.lagstrom@kreftregisteret.no

Pekka Ellonen[3], pekka.ellonen@helsinki.fi

Willem M. de Vos [4,5], willem.devos@wur.nl

Johan G. Eriksson[1,6,7], johan.eriksson@helsinki.fi

Elisabete Weiderpass[1,2,8,9,10], Elisabete.Weiderpass@ki.se

Trine B. Rounge[*1,2,8], trine.rounge@kreftregisteret.no

Affiliations:

[1] Folkhälsan Research Center, Helsinki, Finland,

[2] Faculty of Medicine, University of Helsinki, Helsinki, Finland.

[3] Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland

[4] RPU Immunobiology, Department of Bacteriology and Immunology, University of Helsinki, Helsinki, Finland

[5] Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands

[6] Department of General Practice and Primary Health Care, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[7] Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland

[8] Department of Research, Cancer Registry of Norway, Institute of Population-based Cancer Research, Oslo, Norway

[9] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[10] Department of Community Medicine, Faculty of Health Sciences, University of Tromsø, The Arctic University of Norway, Tromsø, Norway.

Corresponding author:

Dr. Trine B. Rounge

Folkhälsan Research Center

Biomedicum 1

P.O. Box 63 (Haartmansgatan 8)

00014 University of Helsinki, Finland

E-mail address: trine.rounge@kreftregisteret.no

Phone: +47 99604304, Fax number: +358 2941 25382

Running title: Reproducibility of saliva microbiota

Abstract

Culture-independent molecular techniques and advances in next generation sequencing (NGS) technologies make large scale epidemiological studies on microbiota feasible. A challenge using NGS is to obtain the sufficient sequence depth and accuracy to ensure high reproducibility and repeatability which is mostly attained through robust amplification. Here, we aimed to assess the reproducibility of saliva microbiota profiles produced with simplified in-house 16S amplicon assays with large number of barcodes by comparing triplicate samples. The assays included primers with Truseq (TS-tailed) or Nextera (NX-tailed) adapters and either with dual index or dual index plus a 6-bp internal index. All amplification protocol produced consistent microbial profiles for the same samples. However, in our study, reproducibility was highest for the TS-tailed method. Five replicates of a single sample, prepared with the TS-tailed 1-step protocol without internal index sequenced on the HiSeq platform provided high alpha-diversity and low standard deviation (mean Shannon and Inverse Simpson diversity was 3.19 ± 0.097 and 13.56 ± 1.634 respectively). This proposes that 16S amplicon assays using numerous barcodes suitable for large samples sizes and the TS-tailed protocol without internal index can be considered an accurate protocol that provides consistent quantification of bacterial profiles.

Introduction

Presently, there is rising interest in studying human microbiota, microbes living in or on human body, using high throughput approaches based on 16S rRNA gene sequences. This gene is as a highly abundant, evolutionary conserved and phylogenetically informative housekeeping genetic marker (Lane et al. 1985; Tringe and Hugenholtz 2008; Zheng et al. 2015). The composition and diversity of the human microbiota have been correlated to health and disease, although only few cases of causal relationships have been uncovered (Cho and Blaser 2012; Human Microbiome Project Consortium 2012; Nicholson et al. 2012; van Nood et al. 2013; Scheithauer et al. 2016). While most attention has focused on the intestinal microbiota, it is well known that the oral cavity also harbours a large microbial community that includes around 700 common bacterial species, out of which 35% are still unculturable (Dewhirst et al. 2010). Cultivation-independent molecular methods have validated these estimates, by identifying approximately 600 species or phylotypes using 16S rRNA gene sequencing techniques (Paster et al. 2001; Dewhirst et al. 2010). Oral bacteria have been linked to many oral diseases and non-oral diseases, testifying for their importance (Krishnan, Chen and Paster 2017). While metagenomic studies have provided insight in the large coding capacity of the human microbiota (Qin et al. 2010; Li et al. 2014), taxonomic studies mainly rely on amplifying and analysing hypervariable regions of 16S rRNA gene sequences.

It is well known that a precise assessment of the microbiota depends heavily on the hypervariable region selected, and primers used, whereas taxonomic resolution bias can arise with amplification of non-representative genomic regions (Zheng et al. 2015; Wen et al. 2017). Recent development of high-throughput sequencing (HTS) technology and the application of barcode indexing help to obtain thousands of sequences from a large number of samples simultaneously (Andersson et al. 2008; Hamady et al. 2008). However, reproducible identification and consistent quantification of bacterial profiles remain challenging (Ding and Schloss 2014). Studies have

shown that β-diversity metrics depicted significant correlation between oral bacterial composition for the V1–V3 and V3–V4 regions (Zheng et al. 2015). The 16S rRNA V3-V4 hypervariable region is widely used for various microbiological studies (Fadrosh et al. 2014; Belstrøm et al. 2016; Janem et al. 2017). High throughput and cost effective sequencing approaches are continuous being developed, urging researchers to use the latest technologies while abandoning the old ones. However, evaluation of new methodologies is a crucial step in conducting rigorous scientific research (Sinclair et al. 2015). This specifically applies to generating representative libraries of 16S rRNA gene amplicons that are used in shot-gun sequencing.

In this study, we aimed to simplify amplification procedure and investigate barcoding efficacy with internal indices, for sequencing 16S rRNA gene amplicons relative to sequencing quality, depth, reproducibility and repeatability. Specifically, we tested high-throughput workflows for amplicon library construction using Truseq and Nextera adapters with dual index and dual index plus 6-bp internal index (ii) and sequencing of the 16S rRNA V3–V4 hypervariable region. We assessed the reproducibility of the saliva microbiota for four saliva samples in triplicates using the Illumina MiSeq platform and the repeatability using nine control samples, including five replicates from a single individual, with the TS-tailed protocol on the Illumina Hiseq platform.

Materials and Methods

Saliva samples from four volunteers were selected for this study (Fig. 1). The study was approved by the regional Ethics Committee of the Hospital District of Helsinki and Uusimaa (169/13/03/00/10). Saliva samples in triplicates were collected in Oragene-DNA (OG-500) self-collection kits (DNA Genotek Inc, Canada). The saliva samples were mixed with stabilizing reagent within the collection tubes per manufacturer's instructions by participants, and stored at room temperature. A protocol with an intensive lysis step using a cocktail of lysozyme and mechanical disruption of microbial cells using bead-beating was employed. Fifty ml lysozyme (10 mg/ml, Sigma-Aldrich), 6 ml mutanolysin (25 KU/ml, Sigma-Aldrich), and 3 ml lysostaphin (4000 U/ml,

Sigma-Aldrich) were added to a 500 ml aliquot of cell suspension followed by incubation for 1 h at 37 ºC. Subsequently, 600 mg of 0.1- mm-diameter zirconia/silica beads (BioSpec, Bartlesville, OK) were added to the lysate and the microbial cells were mechanically disrupted using Mini-BeadBeater-96 (BioSpec, Bartlesville, OK) at 2100 rpm for 1 min (Yuan et al. 2012). After lysis, total DNA was extracted using cmg-1035 saliva kit, and Chemagic MSM1 nucleic acid extraction robot (PerkinElmer).

PCR amplification

PCR amplification and sequencing libraries was prepared according to in-house 16S rRNA gene-based PCR amplification protocols. All protocols used 16S primers (S-D-Bact-0341-b-S-17: 5' CCTACGGGNGGCWGCAG '3 and S-D-Bact-0785-a-A-21: 5' GACTACHVGGGTATCTAATCC 3') targeting the V3-V4 region as reported previously (Klindworth et al. 2013). The 16S rRNA gene-based primers were modified by adding 5' tails corresponding to the Illumina Truseq and Nextera adapter sequences to the 5'-ends. Amplification was done using primers with and without incorporated internal index (Supplementary Table S1). Two sets of index primers carrying Illumina grafting P5/P7 sequence were used: in-house index primers with Truseq adapter sequence (Supplementary Table S2) and Illumina Nextera i5/i7 adapters. All oligonucleotides (except Illumina Nextera i5/i7 adapters) were synthesized by Sigma-Aldrich (St. Louis, MO, USA).

TS-tailed 1-step amplification

Amplification was performed in 20 µl containing 1 µl of template DNA, 10 µl of 2x Phusion High-Fidelity PCR Master Mix (Thermo Scientific Inc., Waltham, MA, USA), 0,25 µM of each 16S primer carrying Truseq adapter, 0.5 µM of each Truseq index primer. The cycling conditions were as follows: initial denaturation at 98 °C for 30 seconds; 27 cycles at 98 °C for 10 sec, at 62 °C for 30 sec and at 72 °C for 15 sec; final extension at 72 °C for 10 min, followed by a hold at 10 °C.

Separate reactions were done using 16S rRNA gene-based primers with and without incorporated internal index (here after this protocol denoted as TS-tailed 1S).

TS-tailed 2-step amplification

Amplification was performed in 20 µl containing 1 µl of template DNA, 10 µl of 2x Phusion High-Fidelity PCR Master Mix (Thermo Scientific Inc., Waltham, MA, USA), 0,5 µM of each 16S primer carrying Truseq adapter. The cycling conditions were as follows: initial denaturation at 98 °C for 30 sec; 27 cycles at 98 °C for 10 sec, at 62 °C for 30 sec and at 72 °C for 15 sec; final extension at 72 °C for 10 min, followed by a hold at 10 °C. Separate reactions were done using 16S rRNA gene-based primers with and without incorporated internal index. Following PCR amplification, samples were purified using a Performa V3 96-Well Short Plate (Edge BioSystems, Gaithersburg, MD, USA) and QuickStep 2 SOPE Resin (Edge BioSystems, Gaithersburg, MD, USA) according to the manufacturer's instructions. An additional PCR step was needed to add index sequences to the PCR product. Amplification was performed in 20 µl containing 1 µl of diluted (1:100) PCR product, 10 µl of 2x Phusion High-Fidelity PCR Master Mix (Thermo Scientific Inc., Waltham, MA, USA), 0,5 µM of each Truseq index primer. The cycling conditions were as follows: initial denaturation at 98 °C for 2 min; 12 cycles at 98 °C for 20 sec, at 65 °C for 30 sec and at 72 °C for 30 sec; final extension at 72 °C for 5 min, followed by a hold at 10 °C (here after this protocol denoted as TS-tailed 2S).

NX-tailed 2-step amplification

Amplification was performed in 20 µl containing 1 µl of template DNA, 10 µl of 2x Phusion High-Fidelity PCR Master Mix (Thermo Scientific Inc., Waltham, MA, USA), 0.5 µM of each of the 16S rRNA gene-based primers carrying Nextera adapters. The cycling conditions were as follows: initial denaturation at 98 °C for 30 sec; 27 cycles at 98 °C for 10 seconds, at 62 °C for 30 sec and at 72 °C for 15 sec; final extension at 72 °C for 10 min, followed by a hold at 10 °C. Separate reactions were

done using 16S rRNA gene-based primers with and without incorporated internal index. Following PCR amplification, samples were purified using a Performa V3 96-Well Short Plate (Edge BioSystems, Gaithersburg, MD, USA) and QuickStep 2 SOPE Resin (Edge BioSystems, Gaithersburg, MD, USA) according to the manufacturer's instructions. An additional PCR step was needed to add index sequences to the PCR product. Amplification was performed according to Illumina Nextera protocol to amplify tagmented DNA with following exceptions: i) reaction volume was downscaled to 20 µl, ii) 1 µl of diluted (1:100) PCR product was used as template, iii) 1 µl of diluted (1:100) PCR product was used as template, iv) reaction mix was brought to the final volume with laboratory grade water (here after this protocol denoted as NX-tailed 2S).

Pooling, purification and quantification

Following PCR amplifications, libraries were pooled in equal volumes. Library pool was purified twice with Agencourt® AMPure® XP (Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions using equal volumes of the Agencourt® AMPure® XP and the library pool. The purified library pool was analyzed on Agilent 2100 Bioanalyzer using Agilent High Sensitivity DNA Kit (Agilent Technologies Inc., Santa Clara, CA, USA) to quantify amplification performance and yield.

Sequencing

Sequencing of PCR amplicons was performed using the Illumina MiSeq instrument (Illumina, Inc., San Diego, CA, USA). Samples were sequenced as 251x 2 bp paired-end reads and two 8-bp index reads. DNA extracted from nine blank samples, two water samples and nine control saliva samples (in which 5 samples are replicates of sample 4c) using the above mentioned protocol and amplified with TS-tailed 1S protocol without internal index, and sequencing performed (271 x 2 bp) using the Illumina HiSeq instrument.

Phylogenetic Analysis.

Sequencing quality and length filtering was carried out using Nesoni clip Version 0.130 (https://github.com/Victorian-Bioinformatics-Consortium/nesoni). Resulting sequences were processed using mothur (Version v.1.35.1) (Schloss et al. 2009) and sequences were aligned to ribosomal reference database arb-SILVA Version V119 (Quast et al. 2012). We used both SILVA database and the Human Oral Microbiome Database (HOMD) database for the alignment and classification of sequences but present here only the results from the SILVA database and taxonomy as it provides comprehensive, quality checked and regularly updated databases of aligned small (16S / 18S, SSU) and large subunits (Quast et al. 2012). To obtain high quality data for analysis, sequence reads containing ambiguous bases, homopolymers > 8 bp, more than one mismatch in the primer sequence, or less than default quality score in mothur were removed. Assembled sequences with > 460 bp length and contigs that only appeared once in the total set were assumed a result of sequencing error and removed from the analysis. Chimeric sequences were also removed from the data set using the UCHIME algorithm within the mothur pipeline (Edgar et al. 2011). The high-quality sequence reads were aligned to the Silva 16S rRNA database (Version V119) and clustered into operational taxonomic units (OTUs) at a cut-off value > 98% sequence similarity. OTUs were classified using the Silva bacteria taxonomy reference. OTUs were calculated at distance 0.02 and alpha diversity (Shannon and inverse Simpson index) was calculated per sample. These diversity indexes are shown to be a robust estimation of microbial diversity (Haegeman et al. 2013).

Statistic procedures. Microbial diversity indices, both Shannon and Inverse Simpson, were used to summarize the diversity of a population. Simpson's index is more weighted on dominant species whereas Shannon index assumes all species are represented in a sample and that they are randomly sampled (Lozupone and Knight 2008). Kruskal-Wallis (KW-test) test was performed on the alpha diversity indices to assess the statistical significance difference between microbial diversity and the methods used. We performed a principal component analysis (PCA) for 50 abundant OTUs from

all the samples using *prcomp* function in R and PCA plotted using *PCA3D* package (https://cran.r-project.org/package=pca3d). Intraclass correlation coefficients (ICC) to quantify the reproducibility, stability, and accuracy or neutrality of different protocol for six metrics included relative abundances of four major phyla (Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria) and two alpha diversity indices (Shannon & Inverse Simpson index). The ICCs were estimated using the SPSS (version 22) based on the mixed effects model (Sinha et al. 2016). All the graphics and plots were made in R using ggplot2 package (Wilkinson 2011).

Results

Illumina sequencing

Saliva microbiota sequence data of the 16S rRNA V3-V4 region for 4 individuals in triplicates using TS-tailed and NX-tailed amplification, with and without internal index, were collected on the Illumina MiSeq platform (Table 1**).** Two control water samples, nine saliva control samples (including 5 replicates) and blank samples using TS-tailed 1S protocol without internal index, were sequenced on the Illumina HiSeq platform. Samples sequenced using TS-tailed 1S and 2S protocol with and without internal index generated comparatively higher amounts of sequence reads. This was true also after trimming of low-quality sequences (Fig. 2 and Supplementary table S3**).** The sequences were clustered and assigned to 1086 OTUs. Sequence coverage and percentage of sequences passed quality check from each protocol and qualified for taxonomic classification are summarized in Table 1. The protocols with the internal index approach showed consistently 14-23 % lower OTU's per sequence for all protocol. About 61% of saliva microbiota sequences remained after quality check using NX-tailed protocol without internal index, while only 38% remained using NX-tailed protocol with internal index. In our study, the NX-tailed protocol produced slightly less sequences than the TS-tailed protocols, with 4669 and 5399 mean reads per sample respectively. About 60% of saliva microbiota passed the quality check in TS-tailed without internal index and

produced in our protocol more than 8000 reads per sample. Principal component analysis of 50 top abundant OTUs were plotted using PCA3D package in R showing the individual profiles clustered together (Fig. 3). There was clear separation into sample clusters by the microbiota profiles, indicating that the relative abundance of most OTUs were similar between samples from same individuals, as expected.

Alpha diversity of saliva microbiota is similar for all the protocols

The Shannon diversity and inverse Simpson indices used to calculate the alpha diversity showed similar diversity for each sample irrespective of the protocols used with exception of few outliers (Fig. 4a and 4b). The outliers are the samples with low diversity and low sequence depth, < 4000 sequences. Though Shannon diversity index showed less variation according to the sequence depth compared to inverse-Simpson index, we did not find any significant relationship (KW-test) between the diversity indices and the protocols used.

Consistent occurrence of bacterial abundance within the protocols

Taxonomic composition of saliva microbiota from four samples with different amplification protocols with and without internal index showed sample specific composition profile at two taxonomic levels. The bacterial relative abundance at phylum level was measured using the top five abundant phyla; Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria and Proteobacteria (Fig. 5). Similar patterns of phyla abundance were observed for the samples from same individuals using the different protocols. However, detailed comparison of the phyla abundance showed that the oral microbiota of individual 1 included a high abundance of Actinobacteria, Proteobacteria and Firmicutes, that of individual 2 included mainly Firmicutes and Bacteroidetes, whereas that of individuals 3 and 4 included mainly Firmicutes, Bacteroidetes and Proteobacteria. Sample 2b from individual 2 which was sequenced using the TS-tailed 1S protocol without internal index was an outlier with only 733 sequences.

The relative abundance at the bacterial genus level was measured using the top 30 abundant genera (Fig. 6). Similar patterns of genus abundance were also observed for the samples from same individuals using the different protocols. However, these compositions differed between the individuals in line with the differences at the phylum level (Fig. 5).

Reproducibility and stability of the protocols

Average Shannon diversity yielded for sample 1 was comparatively similar except for the TS-tailed 1S protocol with internal index. In sample 2, NX-tailed 2S and TS-tailed 1S without internal indices protocol yielded comparatively less Shannon diversity. Where as in sample 3 and sample 4 Shannon diversity was comparatively similar for all the protocols. Average Inverse Simpson diversity was comparatively less, using NX-tailed 2S protocol for sample 2, 3 and 4, TS-tailed 1S protocol for sample 1, TS-tailed 1S protocol in sample 1 and 2, and, TS-tailed 1S protocol with internal index for sample 1 (Supplementary Table S4). Intra-class correlation coefficients (ICC) used to enumerate the reproducibility and stability of different protocols for six metrics included relative abundances of four top abundant phyla and two alpha diversity indices showed comparatively better reproducibility and stability with TS-tailed 2S protocol with and without internal index (Fig. 7). Actinobacteria from TS-tailed 1S protocol with internal indices, and Shannon index from TS-tailed 1S, NX-tailed protocol, and NX-tailed protocol with internal index showed negative ICC.

Repeatability of the saliva microbiota with TS-tailed 1S protocol

Repeatability of the saliva microbiota using the TS-tailed 1S protocol, which give the reproducibility and stability, were tested with nine control samples in HiSeq Illumina platform. We also amplified and sequenced negative controls; nine blank samples and two water samples to check the effects of reagents and laboratory contamination. HiSeq platform provided 28936 mean sequences data for nine blank samples, 136554 sequences for nine control samples and 790 mean sequences for water samples. Quality check was performed, same as for the MiSeq data

283 (Supplementary Figure S5), and it showed the low diversity for blank samples sequenced and high
284 diversity for the control samples sequenced (Fig. 8). None of the sequences from the water samples
285 passed quality check. Mean Shannon diversity was 0.374 and 3.15 and standard deviation (SD) of
286 0.122 and 0.097, for blank and control samples respectively. Whereas mean inverse Simpson
287 diversity was 1.177 and 13.460 and SD of 0.097 and 1.634, for blank and control samples
288 respectively. Two abundant OTUs from the blank samples were explicitly assigned to two genera of
289 the Proteobacteria phylum, *Pseudomonas* and *Achromobacter*. Bacterial relative abundance of
290 control samples at phyla level shows high abundant of phyla Firmicutes, Bacteroidetes and
291 Proteobacteria (**Fig. 5**). Relative abundance of bacteria at genus level showed that the control
292 samples were enriched in *Veillonella*, *Prevotella*, *Rothia*, *Neisseria* and *Fusobacterium* spp. (Fig.
293 6).

294 ## Discussion

295 Several studies have successfully used the Illumina technology approaches for 16S rRNA gene
296 amplicon sequencing on diverse sample types (Claesson et al. 2010; Gloor et al. 2010; Bartram et
297 al. 2011; Zhou et al. 2011; Caporaso et al. 2012; Degnan and Ochman 2012; Kozich et al. 2013;
298 Fadrosh et al. 2014; Sinclair et al. 2015). However, protocols differ in extraction methods, primers,
299 chemistry and sequencing length between studies and a gold standard has not been established. In
300 this study, we compared the reproducibility and repeatability of six Illumina technology based
301 amplification protocols on saliva samples with primers that were modified in-house (Yuan et al.
302 2012; Klindworth et al. 2013). We aimed to simplify amplification procedure, investigate barcoding
303 efficacy and expand the number of available barcodes to make 16S assays feasible to run on the
304 HiSeq platform.

305 Cells may vary in their susceptibility to lysing methods. Various studies have shown that
306 mechanical lysis gives highest bacterial diversity in 16S rRNA gene based studies, notably when

communities carry hard to lyse Gram-positive bacteria, such as in faecal samples (Salonen et al. 2010; Yuan et al. 2012; Santiago et al. 2014; Robinson, Brotman and Ravel 2016). However, oral samples extracted using either mechanical or enzymatic lysis steps showed an overall similar microbiota profiles (Lazarevic et al. 2013). A recent study also showed that saliva sample collection, storage and genomic DNA preparation with enzymatic-mechanical lysis does not significantly influence the salivary microbiome profiles (Lim et al. 2017). All samples in this study were lysed with an identical protocol including both enzymatic and mechanical disruption of microbial cells using bead-beating to reduce the bias may arise due to the lysis step.

The four saliva samples in triplicates analysed in MiSeq using the different protocols provided comparatively high sequencing coverage for the TS-tailed protocols (>10 k) and less for all other protocols (< 10 k). With current read length of 251 x 2 bp, the V3-V4 region of the rRNA gene is a possible target for sequencing (Mizrahi-Man, Davenport and Gilad 2013), although satisfactory quality of the overlap of the forward and reverse paired-end reads may be challenging. However, extending the sequencing cycle up to 271 bp on the Illumina HiSeq platform provided sufficient overlap, and assembling these reads increases the reliability and quality in the overlapping region. In MiSeq, overall, 39% - 68% of the reads were discarded due to the low-quality score, unassembled pairs, assembled pairs with mismatched barcodes, minimum overlap length and archaeal or eukaryotic sequences. Quality trimming of the NX-tailed protocol sequence data discarded a lower number (6% -8%) of data though it yielded fewer sequences than the TS-tailed protocols. Saliva samples amplicons processed with internal index pairs had lower OTU classification per sequence. Several studies have shown that high incidence of mismatching barcodes is a main loss factor in the microbiota sequencing studies (Degnan and Ochman 2012; Sinclair et al. 2015). Protocols without internal index pairs gave comparatively high percent (>59 %) qualified for the OTU classification per sequence. This suggests that the fragment length is at the borderline of what will yield high quality sequence for the overlap between the read pairs and

adding only a few extra base pairs to the fragment will reduce output quality. Whereas protocols with and without internal index pairs gave different sequence depth and quality data, all the protocols worked sufficiently to provide similar bacterial profiles for each samples. Studies have reported that, with the dual-index approach a large number of samples can be sequenced using a number of primers equal to only twice the square root of the number of samples (Kozich et al. 2013). However, the advantage of dual index with internal index is to reduce the PCR amplification artefacts in high multiplex amplicon sequencing (Peng et al. 2015) and to reduce the cost of sequencing when the study includes a large sample size.

Our results show that low amounts of sequences usually correlate with low diversity. Our sample size was not large enough to conclude that low amounts of sequences was due to the quality or quantity of DNA, technical issues in the lab or difference in robustness of the methods. However, differences in yields using the same DNA, for example seen in sample 1a and 2b, suggest that protocol robustness may cause differences in sequencing yield (Supplementary Figure S6). Laboratory protocol, sequencing platform or error rate and bioinformatics variables can be reason for the majority of variability detected in microbiota studies (Salter et al. 2014; Sinha et al. 2015) but in our study all the protocols delivered overall similar profile of the microbes in the given saliva samples in triplicates. Three OTUs were explicitly assigned only to blank samples in HiSeq run. Negative control samples often yield contaminating bacterial species which may be due to contamination of bacterial DNA in the kits used (Salter et al. 2014). This study also reported that the presence of contaminating sequences is dependent on the amount of biomass in the samples, but we could not assess this in our samples.

Technical challenges have been reported in 16S rRNA amplicon sequencing, such as biases in estimation of population abundance in microbial communities due to the PCR primer selection, PCR template concentration and amplification conditions, pooling of multiple barcodes and sequencing (Wen et al. 2017). Hence, it is important to carefully interpret the experimental results

357 from the technical replicates to validate the reproducibility of the methods. Average alpha-diversity
358 indices for each samples in different protocols yielded comparatively similar profiles with one or
359 two exception, which may due to the low sequence depth. We used the mixed-effect model–based
360 ICC to quantify the reproducibility and stability of the Illumina MiSeq sequencing of saliva
361 microbiome. ICC measures the variability among the multiple measurements for the same sample
362 and assumes that the errors from different measurements have exactly the same statistical
363 distributions and are indistinguishable from each other (Sinha et al. 2016). In our study, based on
364 ICC, sequencing protocols using TS-tailed 2S protocol with and without internal index performed
365 better than NX-tailed protocol and TS-tailed 1S protocols. The negative ICC values observed for all
366 the NX-tailed protocol and TS-tailed 1S protocols may be due to high variation within a subject.

367 Saliva samples sequenced on HiSeq platform yielded high sequence depth ie; 48k – 398k
368 sequences. Variation in technical replicates and low reproducibility, can be overcome by increasing
369 the sequencing depth (Wen et al. 2017), obtainable by the HiSeq platform. Repeatability of the TS-
370 tailed 1S method without internal index for nine control samples sequenced in HiSeq platform was
371 given comparatively high alpha diversity and low variation (SD) among the samples. Alpha
372 diversity was similar for the sample 4 sequencing repeated in MiSeq and HiSeq platform which
373 support the repeatability of method TS-tailed without dual index as good protocol for microbiome
374 studies.

375 In conclusion, NX-tailed 2S protocol and TS-tailed both 1S and 2S protocols were able to reproduce
376 bacterial profiles for the samples sequenced, however, in our hands the reproducibility was
377 comparatively higher for the TS-tailed 2S protocols without internal index on the MiSeq platform.
378 Repeatability of the TS-tailed 1S protocol without internal dual index for nine control samples
379 provided high alpha diversity and less variation among the samples. Considering the cost and time
380 efficiency of using this simplified protocol with numerous barcodes suitable for the HiSeq platform,
381 we suggest that the TS-tailed 1S method can be considered the most effective protocol for

consistent quantification of bacterial profiles in saliva. Reproducibility and repeatability should be taken into consideration in design of a large scale epidemiological study using saliva microbiota.

Acknowledgements

We thank the individuals who participated in this study, and the FIMM biobank and FIMM tech centre. We also thank Timo Miettinen from FIMM tech centre for helping with the internal index setup. We also thank our group members for assisting with the field work of the study Nina Jokinen, Jannina Viljakainen, Stephanie von Kraemer, and the scientific advisors' Dr Eva Roos and Professor Anna Elina Lehesjoki.

Ethics approval and consent to participate

The study was approved by the regional Ethics Committee of the Hospital District of Helsinki and Uusimaa (169/13/03/00/10).

Availability of data and materials

The datasets generated during the current study are available in the NCBI-SRA repository, with the accession number SRP117317.

Competing interests

The authors have no potential conflicts of interest to declare.

Funding

This work was supported by Folkhälsan Research Foundation; Academy of Finland [grant number 250704]; Life and Health Medical Fund [grant number 1-23-28]; The Swedish Cultural Foundation in Finland [grant number 15/0897]; Signe and Ane Gyllenberg Foundation [grant number 37-1977-43]; and Yrjö Jahnsson Foundation [grant number 11486].

References

404 Andersson AF, Lindberg M, Jakobsson H *et al.* Comparative analysis of human gut microbiota by
405 barcoded pyrosequencing. *PLoS One* 2008;**3**:e2836.

406 Bartram AK, Lynch MDJ, Stearns JC *et al.* Generation of multimillion-sequence 16S rRNA gene
407 libraries from complex microbial communities by assembling paired-end Illumina reads. *Appl*
408 *Environ Microbiol* 2011;**77**:3846–52.

409 Belstrøm D, Holmstrup P, Bardow A *et al.* Temporal stability of the salivary microbiota in oral
410 health. *PLoS One* 2016;**11**:1–9.

411 Caporaso JG, Lauber CL, Walters WA *et al.* Ultra-high-throughput microbial community analysis
412 on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012;**6**:1621–4.

413 Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet*
414 2012;**13**:260–70.

415 Claesson MJ, Wang Q, O'Sullivan O *et al.* Comparison of two next-generation sequencing
416 technologies for resolving highly complex microbiota composition using tandem variable 16S
417 rRNA gene regions. *Nucleic Acids Res* 2010;**38**:e200.

418 Degnan PH, Ochman H. Illumina-based analysis of microbial community diversity. *ISME J*
419 2012;**6**:183–94.

420 Dewhirst FE, Chen T, Izard J *et al.* The human oral microbiome. *J Bacteriol* 2010;**192**:5002–17.

421 Ding T, Schloss PD. Dynamics and associations of microbial community types across the human
422 body. *Nature* 2014;**509**:357–60.

423 Edgar RC, Haas BJ, Clemente JC *et al.* UCHIME improves sensitivity and speed of chimera
424 detection. *Bioinformatics* 2011;**27**:2194–200.

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

427 Gloor GB, Hummelen R, Macklaim JM *et al.* Microbiome profiling by illumina sequencing of
428 combinatorial sequence-tagged PCR products. *PLoS One* 2010;**5**, DOI:
429 10.1371/journal.pone.0015406.

430 Haegeman B, Hamelin J, Moriarty J *et al.* Robust estimation of microbial diversity in theory and in
431 practice. *ISME J* 2013;**7**:1092–101.

432 Hamady M, Walker JJ, Harris JK *et al.* Error-correcting barcoded primers for pyrosequencing
433 hundreds of samples in multiplex. *Nat Methods* 2008;**5**:235–7.

434 Human Microbiome Project Consortium. Structure, function and diversity of the healthy human
435 microbiome. *Nature* 2012;**486**:207–14.

436 Janem WF, Scannapieco FA, Sabharwal A *et al.* Salivary inflammatory markers and microbiome in
437 normoglycemic lean and obese children compared to obese children with type 2 diabetes. *PLoS One*
438 2017;**12**:e0172647.

439 Klindworth A, Pruesse E, Schweer T *et al.* Evaluation of general 16S ribosomal RNA gene PCR
440 primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res*
441 2013;**41**:e1.

442 Kozich JJ, Westcott SL, Baxter NT *et al.* Development of a dual-index sequencing strategy and
443 curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform.
444 *Appl Environ Microbiol* 2013;**79**:5112–20.

445 Krishnan K, Chen T, Paster BJ. A practical guide to the oral microbiome and its relation to health
446 and disease. *Oral Dis* 2017;**23**:276–86.

447 Lane DJ, Pace B, Olsen GJ *et al.* Rapid determination of 16S ribosomal RNA sequences for
448 phylogenetic analyses. *Proc Natl Acad Sci U S A* 1985;**82**:6955–9.

449 Lazarevic V, Gaïa N, Girard M *et al.* Comparison of DNA Extraction Methods in Analysis of
450 Salivary Bacterial Communities. *PLoS One* 2013;**8**, DOI: 10.1371/journal.pone.0067699.

451 Li J, Jia H, Cai X *et al.* An integrated catalog of reference genes in the human gut microbiome. *Nat*
452 *Biotechnol* 2014;**32**:834–41.

453 Lim Y, Totsika M, Morrison M *et al.* The saliva microbiome profiles are minimally affected by
454 collection method or DNA extraction protocols. *Sci Rep* 2017;**7**:8523.

455 Lozupone CA, Knight R. Species divergence and the measurement of microbial diversity. *FEMS*
456 *Microbiol Rev* 2008;**32**:557–78.

457 Mizrahi-Man O, Davenport ER, Gilad Y. Taxonomic Classification of Bacterial 16S rRNA Genes
458 Using Short Sequencing Reads: Evaluation of Effective Study Designs. *PLoS One* 2013;**8**:e53608.

459 Nicholson JK, Holmes E, Kinross J *et al.* Host-Gut Microbiota Metabolic Interactions. *Science*
460 2012;**336**:1262–7.

461 van Nood E, Vrieze A, Nieuwdorp M *et al.* Duodenal Infusion of Donor Feces for Recurrent
462 *Clostridium difficile*. *N Engl J Med* 2013;**368**:407–15.

463 Paster BJ, Boches SK, Galvin JL *et al.* Bacterial diversity in human subgingival plaque. *J Bacteriol*
464 2001;**183**:3770–83.

465 Peng Q, Vijaya Satya R, Lewis M *et al.* Reducing amplification artifacts in high multiplex amplicon
466 sequencing by using molecular barcodes. *BMC Genomics* 2015;**16**:589.

467 Qin J, Li R, Raes J *et al.* A human gut microbial gene catalog established by metagenomic
468 sequencing. *Nature* 2010;**464**:59–65.

469 Quast C, Pruesse E, Yilmaz P *et al.* The SILVA ribosomal RNA gene database project: improved
470 data processing and web-based tools. *Nucleic Acids Res* 2012;**41**:D590–6.

471 Robinson CK, Brotman RM, Ravel J. Intricacies of assessing the human microbiome in
472 epidemiological studies. *Ann Epidemiol* 2016;**26**:311–21.

473 Salonen A, Nikkila J, Jalanka-Tuovinen J *et al.* Comparative analysis of fecal DNA extraction
474 methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using
475 mechanical cell lysis. *J Microbiol Methods* 2010;**81**:127–34.

476 Salter SJ, Cox MJ, Turek EM *et al.* Reagent and laboratory contamination can critically impact
477 sequence-based microbiome analyses. *BMC Biol* 2014;**12**:87.

478 Santiago A, Panda S, Mengels G *et al.* Processing faecal samples: a step forward for standards in
479 microbial community analysis. *BMC Microbiol* 2014;**14**:112.

480 Scheithauer TPM, Dallinga-Thie GM, de Vos WM *et al.* Causality of small and large intestinal
481 microbiota in weight regulation and insulin resistance. *Mol Metab* 2016;**5**:1–12.

482 Schloss PD, Westcott SL, Ryabin T *et al.* Introducing mothur: Open-source, platform-independent,
483 community-supported software for describing and comparing microbial communities. *Appl Environ*
484 *Microbiol* 2009;**75**:7537–41.

485 Sinclair L, Osman OA, Bertilsson S *et al.* Microbial community composition and diversity via 16S
486 rRNA gene amplicons: Evaluating the illumina platform. *PLoS One* 2015;**10**:1–18.

487 Sinha R, Abnet CC, White O *et al.* The microbiome quality control project: baseline study design
488 and future directions. *Genome Biol* 2015;**16**:276.

489 Sinha R, Chen J, Amir A *et al.* Collecting fecal samples for microbiome analyses in epidemiology
490 studies. *Cancer Epidemiol Biomarkers Prev* 2016;**25**:407–16.

491 Tringe SG, Hugenholtz P. A renaissance for the pioneering 16S rRNA gene. *Curr Opin Microbiol*
492 2008;**11**:442–6.

493 Wen C, Wu L, Qin Y *et al.* Evaluation of the reproducibility of amplicon sequencing with Illumina
494 MiSeq platform. *PLoS One* 2017;**12**:e0176716.

495 Wilkinson L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics*
496 2011;**67**:678–9.

497 Yuan S, Cohen DB, Ravel J *et al.* Evaluation of Methods for the Extraction and Purification of
498 DNA from the Human Microbiome. Gilbert JA (ed.). *PLoS One* 2012;**7**:e33865.

499 Zheng W, Tsompana M, Ruscitto A *et al.* An accurate and efficient experimental approach for
500 characterization of the complex oral microbiota. *Microbiome* 2015;**3**:48.

501 Zhou H-W, Li D-F, Tam NF-Y *et al.* BIPES, a cost-effective high-throughput method for assessing
502 microbial diversity. *ISME J* 2011;**5**:741–9.

503

504

505

506

507

508

509

510

511

512

513

514

515 **Table 1**

516 Sequencing statistics, quality check passed sequences and sequences classified for samples

517 (combined for methods used) sequenced in MiSeq and HiSeq platform.

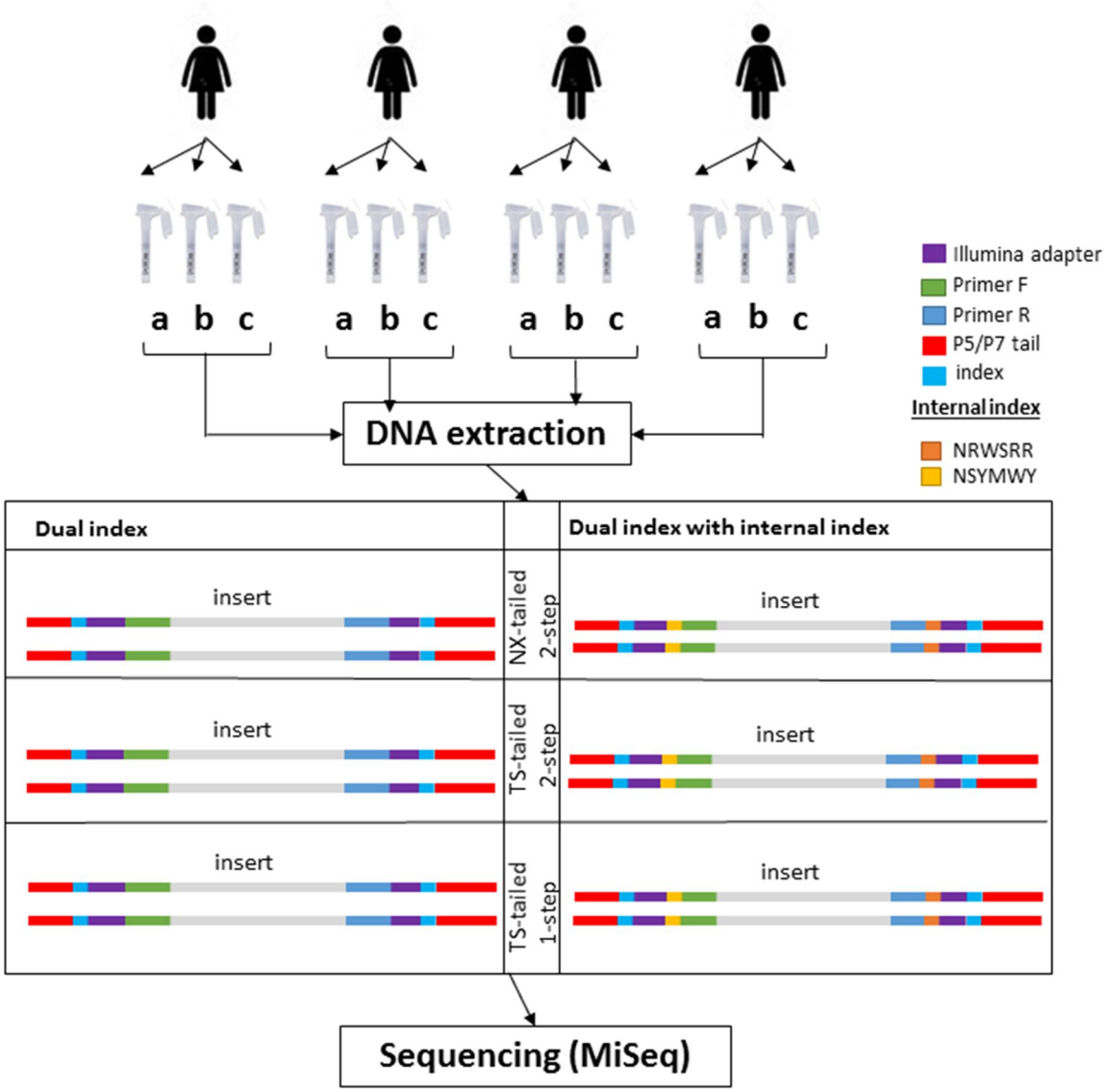
#samples	Protocol	Total sequences	Reads/sample (Min)	(Max)	(Mean)	QC passed sequences	#Sequences classified	% of sequences
12	NX-tailed 2S	56032	2273	6480	4669	52070	34332	61.27
12	NX-tailed 2S ii*	64791	2452	10192	5399	55407	24836	38.33
12	TS-tailed 2S	115736	6940	15161	9644	100788	69783	60.29
12	TS-tailed 2S ii	142880	8048	16271	11906	121134	66014	46.20
12	TS-tailed 1S	96252	733	10868	8021	79339	56761	58.97
12	TS-tailed 1S ii	123165	2187	14519	10263	89631	48479	39.36
9	TS-tailed 1S **	1228989	48100	398420	136554	974702	711088	57.86

518 * Internal indices, ** Control samples sequenced in HiSeq platform

519

Fig. 1

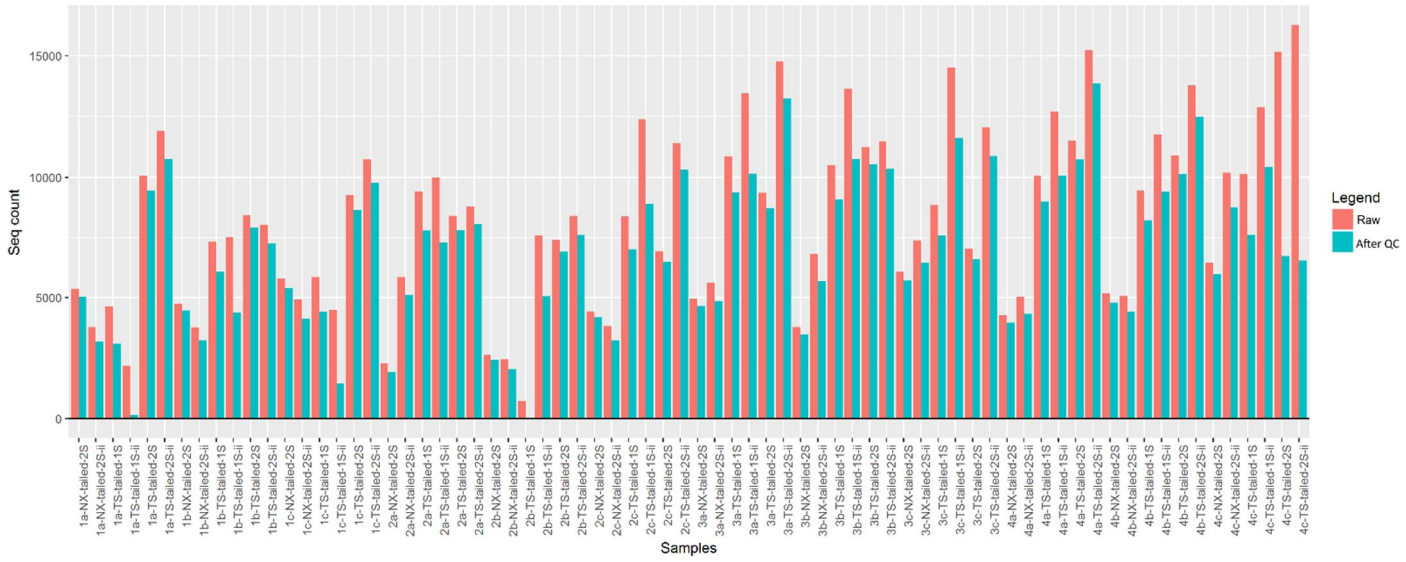
Schematic presentation of the study design showing the number of participants and different methods implemented.



524 **Fig. 2**

525 Distribution of sequences before (Raw data count in blue colour) and after (contigs count in red

526 colour) quality check for the saliva microbiota sequenced in MiSeq platform

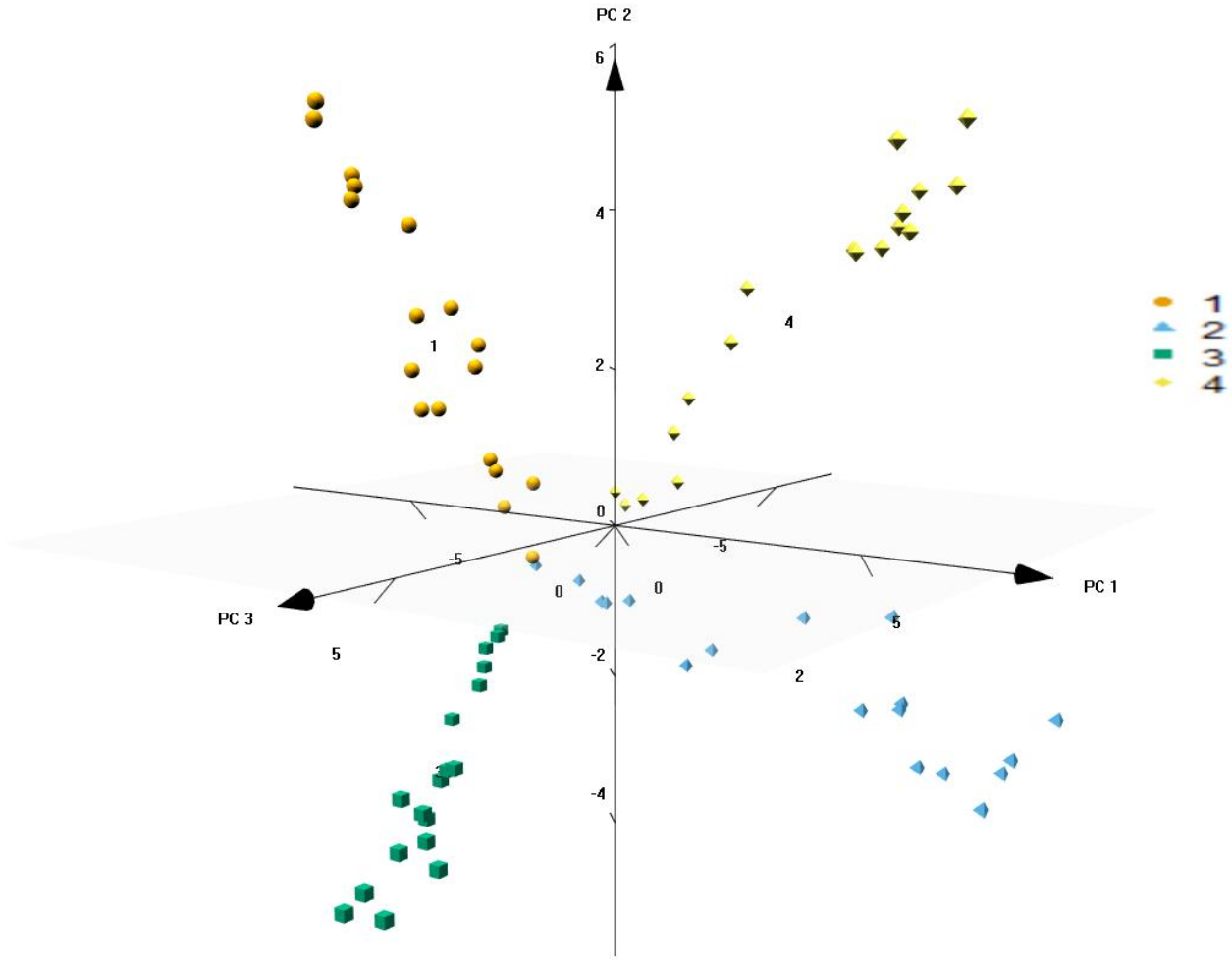


527

528 **Fig. 3**

529 Principal component analysis (PCA) plot of PC1 vs PC2 vs PC3 for the abundant 50 OTUs from all

530 the samples.



531

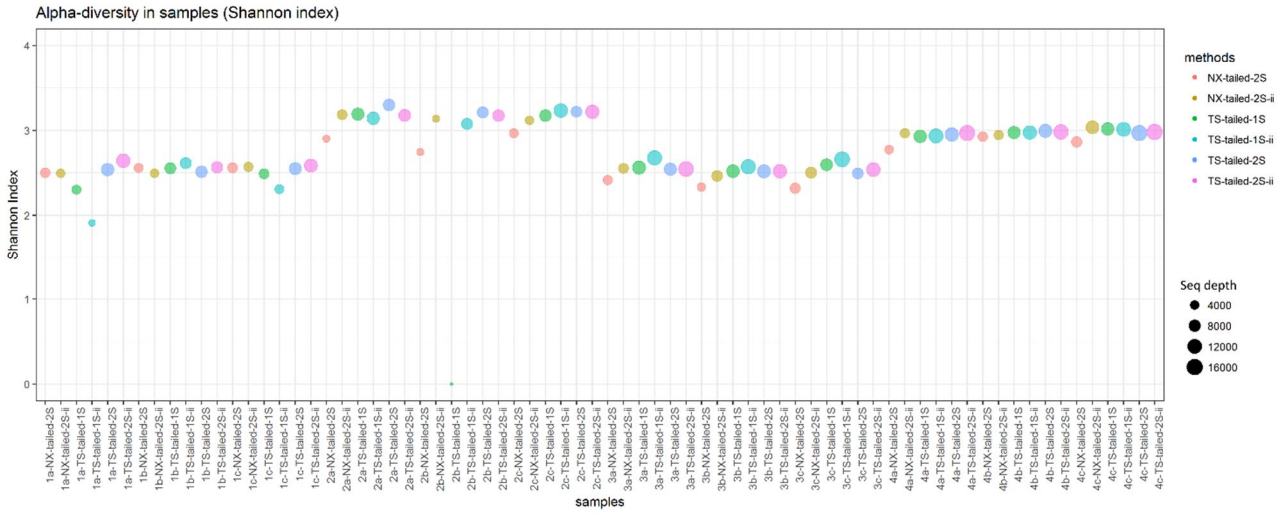
532

Fig. 4 a & b

Alpha diversity measured using Shannon (4a) and Inverse Simpson (4b) index in each replicates.

Bubble size depicts the sequence depth and bubble colour is Illumina method used.

4a



4b

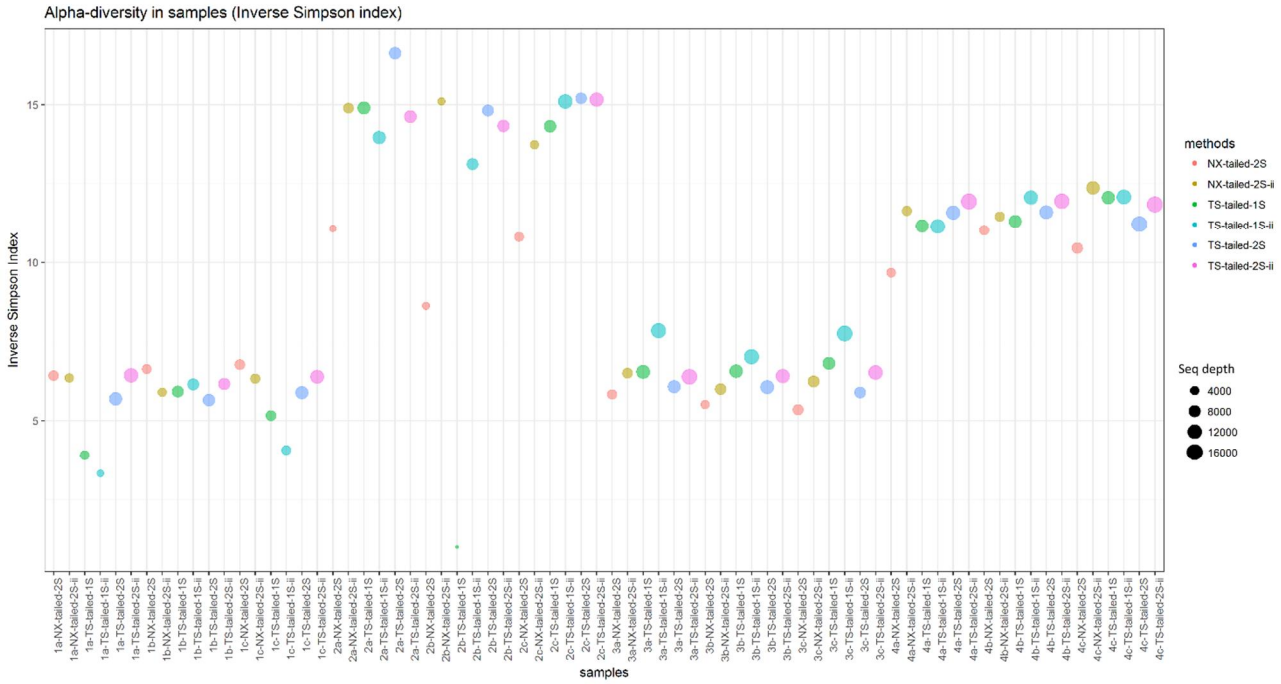


Fig. 5

Composition of abundant phylum in each sample separated by individuals/participants. Phylum composition of blank and control samples were also included in the figure as separate samples. Samples with low sequence depth are marked in red coloured box.

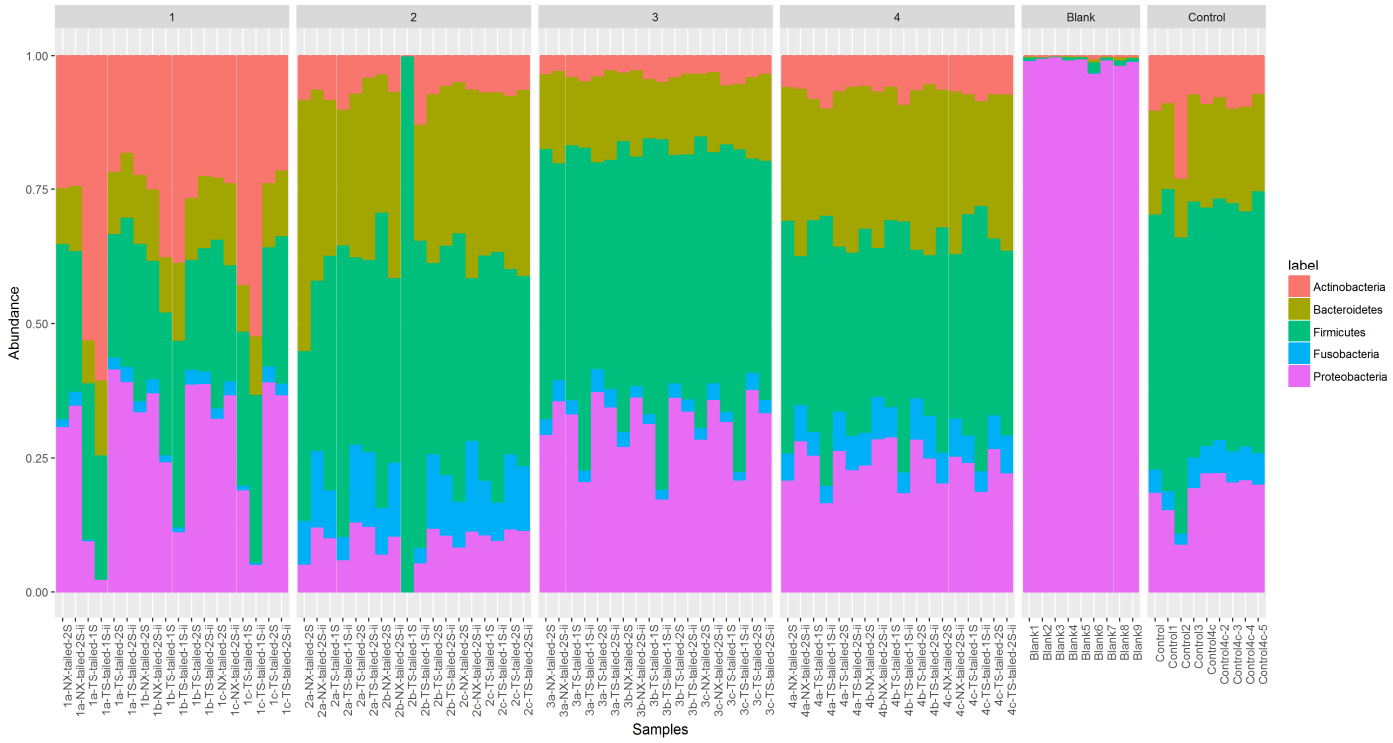
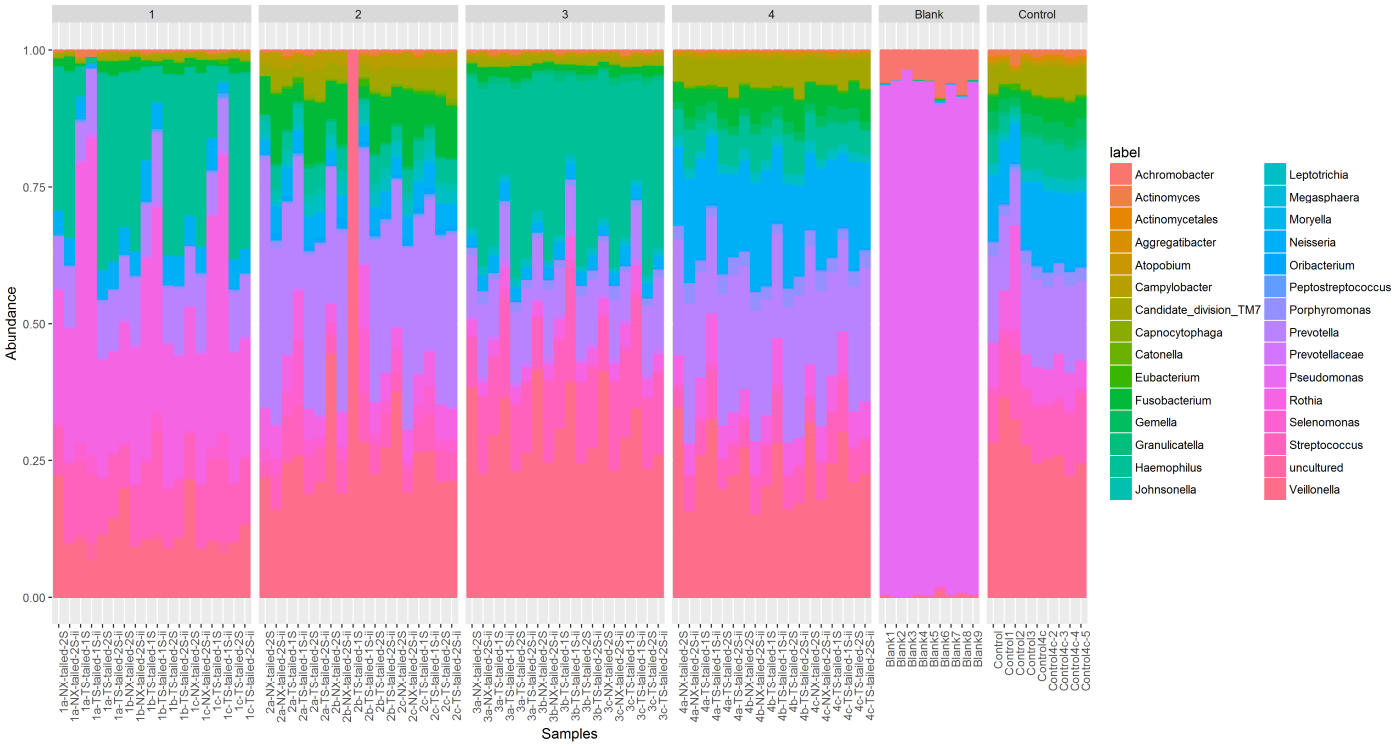


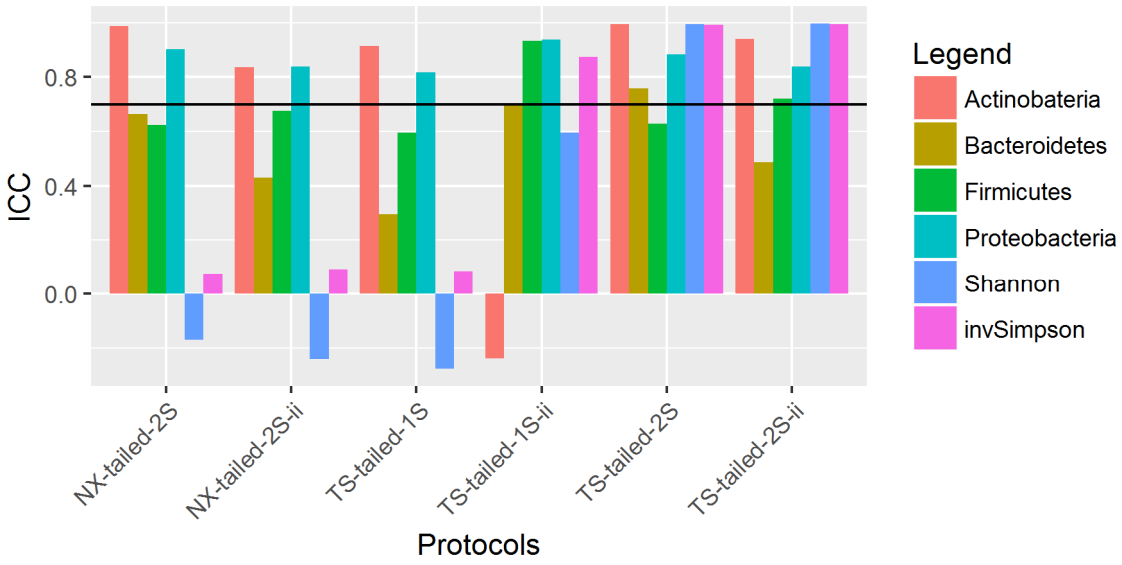
Fig. 6

Composition of abundant genus in each sample separated by individuals/participants. Genus composition of blank and control samples were also included in the figure as separate samples. Samples with low sequence depth are marked in red coloured box.



551 **Fig. 7**

552 Intra-class correlation coefficient plotted for six metrics included relative abundances of four major
553 phyla (Actinobacteria, Bacteroidetes, Firmicutes and Proteobacteria) and two alpha diversity indices
554 (Shannon & Inverse Simpson index).



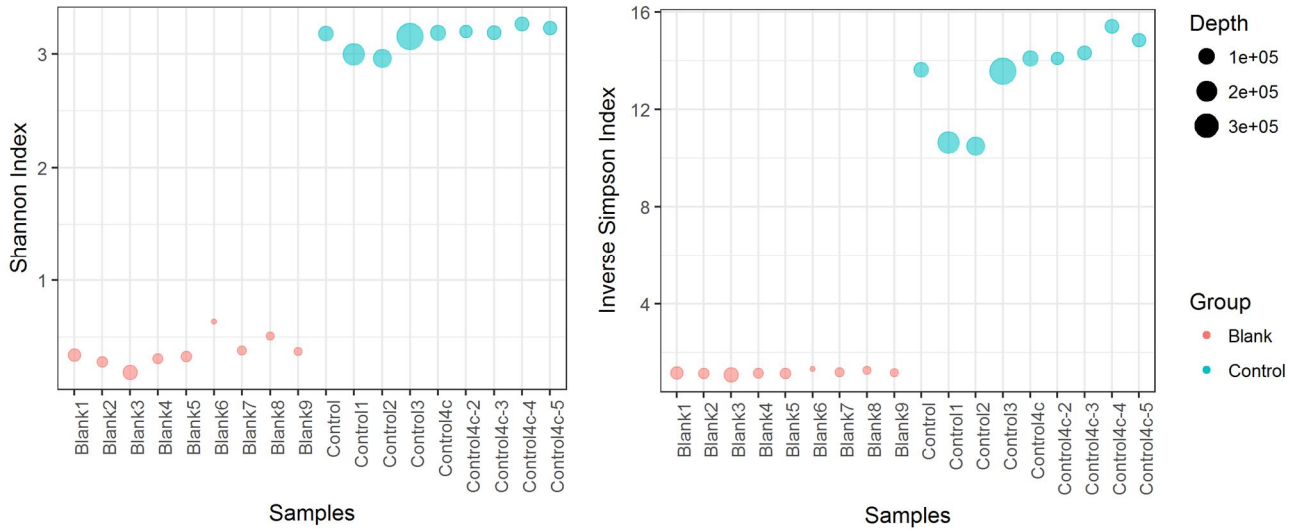
555

556

557 **Fig. 8 a & b**

558 Alpha diversity measured using Shannon (7a) and Inverse Simpson (7b) index in blank and control

559 samples data from HiSeq platform.



560

561