

### **Supplementary methods: sample handling and sequencing**

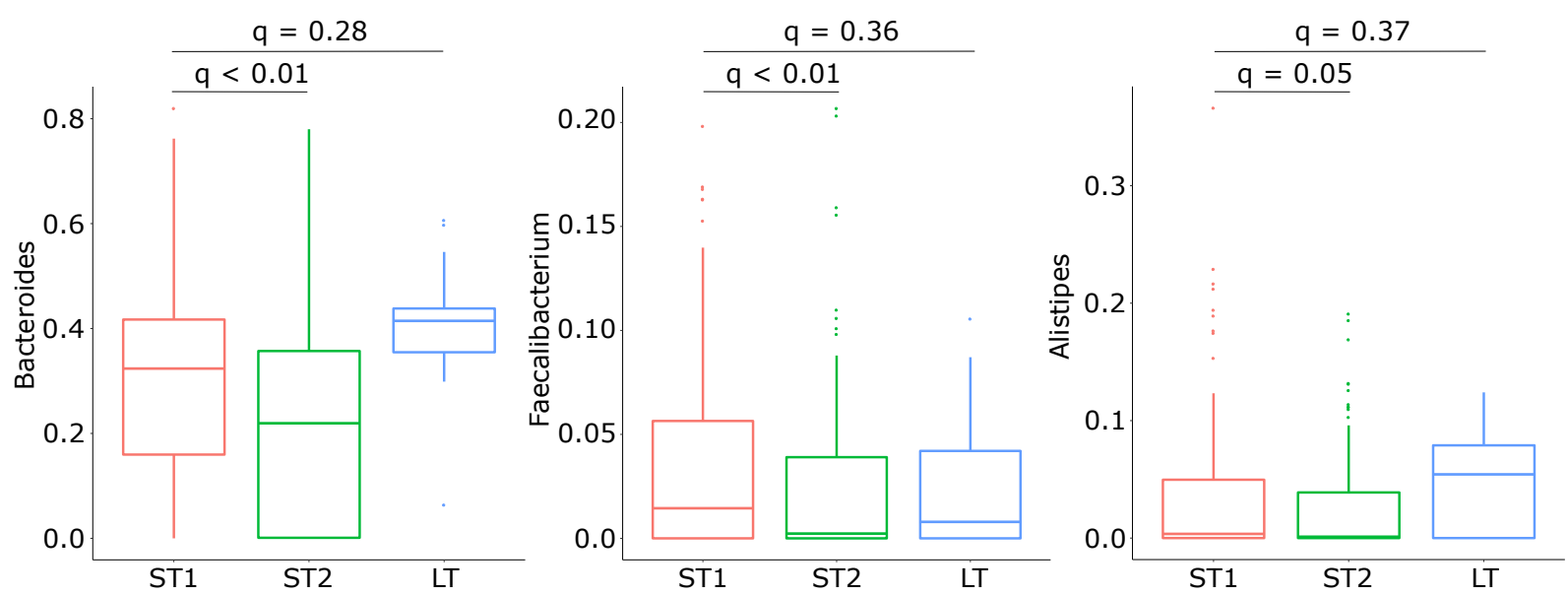
Samples were collected in 95% ethanol-filled sterile tubes and stored at -80°C. DNA was extracted using the DNeasy PowerSoil DNA isolation kit (QIAGEN, Hilden, Germany). The V4 hypervariable region of the 16S rRNA gene was amplified on an Illumina MiSeq platform (2 x 300 paired-end mode)<sup>1</sup>. Adaptor trimming was done using SHI7<sup>2</sup>, and the resulting demultiplexed fastq files were used as input to DADA2<sup>3</sup> to infer exact amplicon sequence variants (ASVs) (*dada2* package v1.18.0 in R). For filtering, we used DADA2 default parameters (PHRED score threshold of 2, maximum number of expected errors of 2 for both forward and reverse reads) and truncation lengths of 220 (forward) and 150 (backward). De-replication, de-noising, merging, and chimera removal were done using DADA2 default parameters. Taxonomic assignment was done by the naive Bayesian classifier implemented in DADA2 and the SILVA non-redundant v138.1 training set<sup>4</sup>.

### **References**

1. Gohl DM, Vangay P, Garbe J, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat. Biotechnol.* 2016;34(9):942–949.
2. Al-Ghalith GA, Hillmann B, Ang K, Shields-Cutler R, Knights D. SHI7 Is a Self-Learning Pipeline for Multipurpose Short-Read DNA Quality Control. *mSystems.* 2018;3(3.):
3. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods.* 2016;13(7):581–583.
4. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 2013;41(Database issue):D590–6.

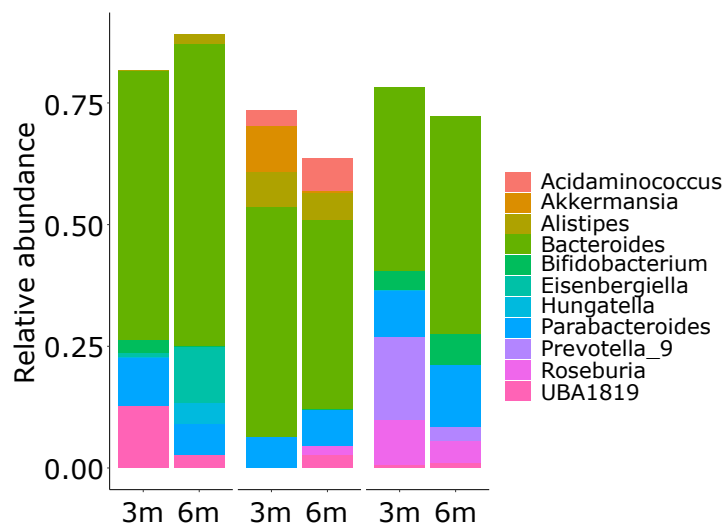
**Supplementary Table S1: Sample data**

	<b>N (samples)</b>	<b>N (subjects)</b>	<b>Average number of samples per subject</b>
<b>All samples</b>	410	52	7.9
<b>Short-term</b>	386	52	7.4
<b>Long-term</b>	24	18	1.3
<b>Samples analyzed</b>	339	51	6.6
<b>Short-term</b>	318	51	6.2
<b>Long-term</b>	21	16	1.3



**Supplementary Figure S1: Changes in the relative abundance of *Bacteroides*, *Faecalibacterium*, and *Alistipes***

Pooled relative abundances are shown in the three intervals defined in Fig. 1a.  $q$  values are from Fig. 2a-b using mixed effect models.



**Supplementary Figure S2: Microbiota**

**composition in 3-month vs. 6-month samples**

Genus-level relative abundances in the 3-month vs. 6-month sample from the same patient are visualized. Only 3 patients had both 3-month and 6-month eligible samples. Each bar represents a genus. The 5 most abundant genera in each sample were selected and the combined set of genera generated from all included samples was used to plot the stacked bars.