

## Supplementary material and methods

### Faecal microbiota analyses

In order to extract metagenomic DNA, 250 mg of the frozen faecal samples were homogenised in phosphate buffered saline (PBS) and were centrifuged for 1 minute at 900 rpm. For cell lysis, a combination of chemical, mechanical and thermal disruption was used. A lysis buffer containing 1M Tris-HCl, 0.5M EDTA, 5M sterile NaCl and SDS (final concentration 4%) was filled into bead tubes of the Ambion MagMax™ Total Nucleic Acid Isolation Kit (*Thermo Fisher Scientific*) and mixed with 175 µl supernatant of faeces in PBS. Mechanical disruption consisted of a bead-beating procedure using the Fastprep™ Homogenizer (5,5 ms for 3x1 min; resting 1 min in between, *MP Biomedicals*). Samples were subsequently incubated for 15 minutes at 95°C with gentle shaking. After centrifugation for five minutes at 11000 rpm, supernatant was filled in an Eppendorf tube. Afterwards, a second round of bead beating and incubation was performed and supernatants were pooled and stored at -20°C until further analysis. 200 µl of the supernatants were introduced into a KingFisher 96-wells deep well plate (*Thermo Fisher Scientific*), together with bead mix of the Ambion MagMax™ Total Nucleic Acid Isolation Kit (*Thermo Fisher Scientific*), isopropanol, and lysis buffer. Other plates were filled with wash buffers, elution buffer (+RNAse), and 96-tips for DW magnets (*Thermo Fisher Scientific*). Afterwards, the prepared plates were introduced into the KingFisher system and the DNA isolation was performed according to the manufacturer's standard protocol (*Thermo Fisher Scientific*). After removal of the plates from the system, the plate containing

purified nucleic acids was incubated for 15 minutes at 37°C for degradation of RNA.

Subsequently, according to current international accepted standards [1] the V4 hypervariable region of the 16S rRNA gene was amplified in triplicate using the 515F/806R barcoded primer pair as described previously [2]. Pooled amplicons from the triplicate reactions were purified using AMPure XP purification (Agencourt) according to the manufacturer's instructions and eluted in 25 µl 1 × low TE (10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0). Quantification of amplicons was subsequently performed by the Quant-iT PicoGreen dsDNA reagent kit (Invitrogen) using a Victor3 Multilabel Counter (*Perkin Elmer, Waltham, USA*). Amplicons were mixed in equimolar concentrations to ensure equal representation of each sample and sequenced on an Illumina MiSeq instrument (MiSeq Reagent Kit v3, 2 × 250 cycles, 10% PhiX) to generate paired-end reads of 250 bases (~25.000 reads/sample) [3].

The pre-processing of sequencing data, using an in-house pipeline based upon DADA2 (R version 4.0.3) [4], consisted of the following steps: reads filtering, identification of sequencing errors, dereplication, and removal of chimeric sequences. In order to assign taxonomy, DECIPHER [5] was used to annotate to the genus level. Data were expressed as amplicon sequence variants (ASVs). Decontam was used with the either setting, which combines the two statistical methods prevalence and frequency for the identification of contamination in marker-gene and metagenomics data [6]. Contaminated ASVs identified by decontam were filtered out together with ASVs presented

in less than 5% of all samples and a total abundance of less than 0.001%. A total of 908 ASVs were maintained for downstream analysis. The final file was saved in the phyloseq format [7].

## References

- [1] Liu PY, Wu WK, Chen CC, Panyod S, Sheen LY, Wu MS. Evaluation of Compatibility of 16S rRNA V3V4 and V4 Amplicon Libraries for Clinical Microbiome Profiling. *bioRxiv*. 2020.
- [2] Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6:1621-4.
- [3] Galazzo G, Tedjo DI, Wintjens DSJ, Savelkoul PHM, Masclee AAM, Bodelier AGL, et al. Faecal Microbiota Dynamics and their Relation to Disease Course in Crohn's Disease. *J Crohns Colitis*. 2019;13:1273-82.
- [4] Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581-3.
- [5] Wright ES. Using DECIPHER v2.0 to Analyze Big Biological Sequence Data in R. *The R Journal*. 2016;8(1):352-9.
- [6] Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*. 2018;6:226.
- [7] McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8:e61217.