MARS 16S processing pipeline methods

DNA extraction, PCR amplification, and sequencing of taxonomic marker

DNA was extracted from 0.25g of fecal sample using the MagAttract PowerSoil Pro DNA kit (Qiagen, Inc) according to the manufacturer's protocol for the Kingfisher sample handling robot. DNA extracts were quantified using the Quant-iT PicoGreen kit (invitrogen, ThermoFisher Scientific). Partial bacterial 16S rRNA (V4) and fungal IT S2 genes were amplified using 30ng extracted DNA as template. The V4 region was amplified using 515F and 806R with Illumina adapters and dual indices (8 basepair (Kozich 2013)). The ITS2 region was amplified with fITS7 (Ihrmark et al. 2012) and IT S4 (White 1990) using the same dual indexing design as the V4. Samples were amplified in triplicate 15ul reactions using Go-Taq DNA polymerase (Promega) with the addition of 3.3µg BSA (New England BioLabs). To overcome inhibition from host DNA, 0.1pmol primer without the indexes or adapters was added to the mastermix. The v4 16S PCR reaction was incubated at 95℃ for 2 minutes, then 30 cycles of 30 s at 95.0°C, 30 s at 50.0°C and 60 s at 72.0°C, followed by final extension as 72.0°C for 10 minutes. The ITS2 PCR reaction was incubated at 95℃ for 2 minutes, then 5 cycles of 30 s at 95.0°C, 60 s at 48.0°C and 60 s at 72.0°C, then 25 cycles of 30 s at 95.0°C, 60 s at 55.0°C and 60 s at 72.0°C followed by final extension as 72.0°C for 10 minutes.

PCR products were pooled for quantification and visualization using the QIAxcel DNA Fast Analysis (Qiagen). PCR products were normalized based on the concentration of DNA from 250-400 bp then pooled using the epMotion 3075 liquid handling robot. The pooled PCR products were cleaned using Omega Bio-Tek Mag-Bind Beads according to the manufacturer's protocol using 0.8x beads to PCR product. T he cleaned pool was sequenced on the MiSeq using v2 2x250 base pair kit (Illumina, Inc).

Sequence data processing and statistical analyses

Sequences were processed in Mothur v. 1.36.1 following the MiSeq SOP (Kozich et al., 2013). After demultiplexing and quality checking steps the sequences were cluster at 97% similarity. Alpha and beta diversity statistics were calculated by subsampling to 10,000 reads per sample. NMS and Permanova were run using the vegan package(Oksanen 2015) in R 3.2.0. A subsampled species matrix was used for indicator species analysis (DeCaceres 2009) and generating a heatmap.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq IlluminaSequencing Platform. Appl. Environ. Microbiol. 79, 5112–5120.

Lange, V., Böhme, I., Hofmann, J., Lang, K., Sauter, J., Schöne, B., Paul, P., Albrecht, V., Andreas, J.M., Baier, D.M., et al. (2014). Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. BMC Genomics 15, 63.

Jari Oksanen, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens and Helene Wagner (2015). vegan: Community Ecology Package. R package version 2.3-0.http://CRAN.R-project.org/package=vegan

White T J, Bruns T , Lee S, and T aylor J (1990). Amplification and direct sequencing of fungal ribosomal RNA for phylogenetics. PCR Protoc. Guide Methods Appl. 315–322.