Supplementary File 2

Metagenomic DNA extraction, amplicon preparation and sequencing

Metagenomic DNA was extracted from proximal colon contents using a QIAamp DNA stool kit (Qiagen, Valenica, CA) with some modifications. A six-minute incubation at 95°C was used to replace the 70°C lysis recommended in the standard protocol. DNA integrity was verified using a Bioanalyzer 2000 (Agilent, Palo Alto, CA). DNA concentration was quantified using a QuantiFluor fluorometer (Promega, Madison, WI). A 570-bp region of the 16S rRNA gene (*E. coli* position 357 to 926) containing hypervariable regions V3- V5 of the 16S rDNA, selected because of their high variability (Andersson et al., 2008), was amplified from 40 ng of metagenomic DNA with eight-bp sample-specific bar coded primers using 2.5 units of AccuPrime Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA) in a 50-µl reaction buffer containing 200 nM primers, 200 nM dNTP, 60 mM Tris-SO_4, 18 mM $(NH4)_2SO_4$, 2.0 mM MgSO4, 1% glycerol, and 100 ng/ul bovine serum albumin (New England BioLabs, Ipswich, MA). PCR was performed using the following cycling profile: Initial denaturing at 95°C for two min followed by 25 cycles of 95°C 30 s, 50°C 30 sec, and 72°C 120 sec. Amplicons were generated from each metagenomic DNA sample separately, purified using a Agencourt AMPure XP kit (Beckman Coulter Genomics, Danvers, MA), and quantified using a QuantiFluor fluorometer. Amplicons from individual samples were pooled in equal mass (molar) ratios. The amplicon pool at the desired size (~672 bp including primers and adaptors) was excised from 1.0% agarose gel and purified using a QIAquick Gel Extraction Kit (Qiagen). The purified amplicon library was further verified and quantified using a BioAnalyzer 2000 (Agilent) and subject to Roche/454 pyrosequencing.

The porcine proximal colon microbiota was characterized by two sequencing approaches using the Roche/454 GS FLX Titanium chemistry, the 16S rRNA gene (hypervariable V3-V5 regions) and the whole genome shotgun (WGS). For the first approach, unidirectional sequencing of amplicon libraries was performed according to the manufacturer's instructions with a modification (App No 001-2009, Roche Applied Science, Indianapolis, IN). This modification, using a specific fusion primer design, accommodates amplification using the GS FLX Titanium emPCR Kits (Lib-L). Five hundred ng of DNA were used to generate libraries using the GS FLX Titanium Rapid Library Preparation method for WGS sequencing. Therefore, emulsion PCR (emPCR) was carried out using a Lib-L kit for both approaches. Pyrosequencing was conducted using a GS FLX Titanium System (Roche) following the manufacturer's protocol.

The analysis pipeline for WGS sequence data: