A reference gene catalogue of the pig gut microbiome

Liang Xiao1†, Jordi Estellé2†, Pia Kiilerich3†, Yulias Ramayo-Caldas2, Zhongkui Xia1, Qiang Feng1†, Suisha Liang1, Anni Øyan Pedersen4, Niels Jørgen Kjeldsen6, Chuan Liu1,5, Emmanuelle Maguin6, Joël Doré6,7, Nicolas Pons7, Emmanuelle Le Chatelier7, Edi Prifti7,§, Junhua Li1,8, Huijue Jia1, Xin Liu1, Xun Xu1, Stanislav D. Ehrlich7,9, Lise Madsen1,3,10, Karsten Kristiansen1,3*, Claire Rogel-Gaillard2* and Jun Wang1,3*
Supplementary Figure 1 | Distribution of microbial gene (a) and MGS (b) annotated against the NR database. Only annotated genes/MGS (green areas in panels a & b) were included in the analysis. 50.8% of the NR genes and 69.1% of the MGS could be annotated to the Bacteria super-kingdom, demonstrating that many unknown genes are detected in the pig gut microbiota. Further taxonomic annotation was carried out for the NR genes assigned to Bacteria and Archaea, with 41.2 and 7.6% genes annotated at the phylum and genus levels, respectively. Less than 1% of the genes could be annotated at the species level (0.33%).
Supplementary Figure 2 | KEGG (a) and eggNOG (b) annotation of the NR genes assigned to bacterial species (left hand) and annotation of the NR genes clustered into MGS (right hand). The functional annotation based on all NR genes and on the genes clustered in the MGS was found to be consistent. COG: Cluster of Orthologous Groups.
Supplementary Figure 3 | KEGG annotation of gut microbiome biological functions that are 100% shared by the whole set of 287 pigs. The shared functions are consistent with those identified for the entire pig gut gene catalogue both in relation to annotation and relative abundances.
**Supplementary Figure 4 | Functional comparison of the pig, human and mouse catalogues.** The Venn diagram (a) provides the number of shared and species-specific KEGG pathways for the pig (pink), the mouse (grey), and the human (yellow) catalogues. The classification of the 2179 KEGG pathways found 100% shared by the three animal species (b) highlights the predominance of common metabolic functions related to carbohydrates and amino acids as well as environmental information processing (membrane transport), consistent with the most abundant functions found in the pig catalogue (see Supplementary Fig. 2 and Supplementary Fig. 3).
Supplementary Figure 5 | Effects of host genetics on the pig microbiome composition. The influence of host genetics was assessed by NMDS from the subset of Chinese pigs, at the levels of phylum (a), genus (b), species (c), MGS (d) and KEGG pathways (e). The MGS-based NMDS clearly distinguished three groups corresponding to the highly selected commercial breed (HybCN1, HybCN2 and Large White), the Bama and related BaRing pigs, and the Tibetan pigs. The KEGG-based NMDS still clearly separated the Tibetan pigs from the others, suggesting specific microbiota functions in this breed.
Supplementary Figure 6 | Age effect on the pig gut microbiota composition. The influence of age was assessed by NMDS at the levels of the total NR gene counts (a) and KEGG pathways (b, c), from the subsets of Danish pigs (a, c) and French pigs (b) (counterpart of Fig. 3b in the main text). The animals are distributed in a surface plot that includes lines referring to their ages, showing an effect of age, likely to be connected to the diet and environmental changes during lifetime.
**Supplementary Figure 7 | Effect of the farm system (country) on the gut microbiota composition.** NMDS based on the NR gene counts (a), the KEGG pathways (b) and the MGS counts revealed a separation between the Chinese pigs (CP), and the French (FP) and Danish (DP) pigs. This separation is likely to be partly related to different farm systems. In addition, the French pigs comprised 11 subsets divided by the breed, the environment, or both, resulting in more diversity than for the Danish pigs, as revealed by this analysis.
Supplementary Figure 8 | Country-specific relative abundances of subsets of antibiotic resistance genes (ARGs) and of KOs related to the tricarboxylic acid (TCA) cycle.  (a) Aminoglycosides; (b) Beta-lactams; (c) Sulfonamide;  (d) KOs in the TCA cycle.
Supplementary table 1. Background information on the 287 pig samples.

Supplementary table 2. Description of the assembly data from the 287 samples.

Supplementary table 3. Assembly results of the pig, human and mouse data.

Supplementary table 4. Significant differences in abundance between the gut microbiomes of castrated males and females (10 castrated males, 10 females, Svindinge farm) at the genus and species levels.

Supplementary table 5. Significant differences in abundance between the gut microbiomes of castrated males and females (10 castrated males, 10 females, Svindinge farm) at the KEGG pathway level.

Supplementary table 6. Significant differences in abundance between the gut microbiomes of males and females (11 males, 14 females, fed wet feed, Stærsminde farm) at the genus and species levels.

Supplementary table 7. Significant differences in the abundance of MGS in the gut microbiomes of males and females (11 males, 14 females, fed wet feed, Stærsminde farm).

Supplementary table 8. Significant differences in KEGG pathways in the gut microbiota of males and females (11 males, 14 females, fed wet feed, Stærsminde farm).

Supplementary table 9. List of the KEGG pathways mapped by iPATH2 that differed significantly between the gut microbiota of males and females (11 males, 14 females, fed wet feed, Stærsminde farm). The three most represented pathways are highlighted in grey. This list is derived from table S8 and corresponds to the differentially abundant functions mapping against KEGG according to iPATH2 tools.