Engineered fibre enables targeted activation of butyrate-producing microbiota in the distal gut

Leszek Michalak[1], John Christian Gaby[1], Leidy Lagos[2], Sabina Leanti La Rosa[1], Nicolas Terrapon[3,4], Vincent Lombard[3,4], Bernard Henrissat[3,4,5], Magnus Ø. Arntzen[1], Live Heldal Hagen[1], Johannes Dröge[6], Margareth Øverland[2], Phillip B. Pope[1,2]*, Bjørge Westereng[1]*

[1] Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, 1432 Ås, Norway

[2] Faculty of Biosciences, Norwegian University of Life Sciences, 1432 Ås, Norway

[3] Centre National de la Recherche Scientifique, Aix-Marseille Université, UMR7257, France

[4] Institut National de la Recherche Agronomique, USC1048 Architecture et Fonction des Macromolécules Biologiques, Marseille, France

[5] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[6] Department for Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden

* Contributed equally

Corresponding authors:

Bjorge Westereng bjorwe@nmbu.no

Phillip B. Pope phil.pope@nmbu.no

John Christian Gaby john.christian.gaby@nmbu.no

ABSTRACT

Beneficial modulation of the gut microbiome has high-impact implications not only in humans, but also in livestock that sustain our current societal needs. In this context, we have engineered an acetylated galactoglucomannan (AcGGM) fibre from spruce trees to match unique enzymatic capabilities of *Roseburia* and *Faecalibacterium* species, both renowned butyrate-producing gut commensals. The accuracy of AcGGM was tested in an applied pig feeding trial, which resolved 355 metagenome-assembled genomes together with quantitative metaproteomes. In AcGGM-fed pigs, both target populations differentially expressed AcGGM-specific polysaccharide utilization loci, including novel, mannan-specific esterases that are critical to its deconstruction. We additionally observed a "butterfly effect", whereby numerous metabolic changes and interdependent cross-feeding pathways were detected in neighboring non-mannolytic populations that produce short-chain fatty acids. Our findings show that intricate structural features and acetylation patterns of dietary fibre can be customized to specific bacterial populations, with the possibility to create greater modulatory effects at large.

MAIN TEXT

Prebiotic strategies use dietary fibre to manipulate gut microbiota and promote specific populations to improve gut function in humans and production animals. Prebiotics by definition are not broadly metabolized, but rather elicit a targeted metabolic response in specific indigenous microbiota that confers health and nutrition benefits to their host. This in itself presents a challenge; as many newly identified target organisms, such as beneficial butyrate-producing *Roseburia* and *Faecalibacterium* spp.[1,2], have broad metabolic capabilities that are shared with the vast majority of fibre-fermenting microbiota in the gut ecosystem. Nevertheless, recent studies have revealed intimate connections between the enzymatic and mechanistic features of microorganisms and the glycan structures of the fibres they consume, which creates new conceptual prebiotic targets. This is

exemplified by discoveries of sophisticated polysaccharide-degrading apparatuses that enable certain microbiota to consume fibre in a "selfish" manner, whereby complex glycan-structures (such as β-mannans) are cleaved into large oligosaccharides at the cell surface, which are subsequently transported into the cell and depolymerized into monomeric sugars[3-5]. Such a mechanism restricts the release of sugars into the ecosystem for neighboring scavenging populations, thus giving a selective metabolic advantage to the selfish-degrader in the presence of these highly complex glycans.

Beta-mannans are present in human and livestock diets, and depending on their plant origins, can be heavily decorated with varying amounts of acetylation that protect the fibre from enzymatic degradation[6]. We recently demonstrated that the human gut commensal *Roseburia intestinalis* encodes a mannan-specific polysaccharide utilization locus (PUL), and "selfishly" converts highly complex mannan substrates to butyrate[4]. Within this mannan PUL, a carbohydrate esterase (*Ri*CE2) removes 3-O-, and 6-O- acetylations on mannan, whereas a novel *Ri*CEX removes oriented 2-*O*-hydroxyl acetylations[6], which are distinctive features found in limited mannan moieties and inaccessible to most esterases present in the gut microbiome. Closer genome examinations have revealed that putative CE2/CEX-containing mannan PULs are infrequently distributed amongst common fibre-degrading gut microbiota, yet they are prominent within many butyrate-producers including *Roseburia* spp., *Faecalibacterium prausnitzii*, *Ruminococcus gnavus*, *Coprococcus eutactus* and *Butyrivibrio fibrisolvens*[4,7]. It is well known that the metabolic attributes of these populations are highly desirable in the gastrointestinal tract, and that their depletion is implicated in colorectal cancer, Crohn's disease, inflammatory bowel syndrome, ulcerative colitis, forms of dermatitis and several other diseases[8,9]. These collective findings thus raised the question: could a custom fibre that was engineered to match these specialized enzymatic capabilities be harnessed to selectively engage butyrate-producers in a complex microbiome ecosystem?

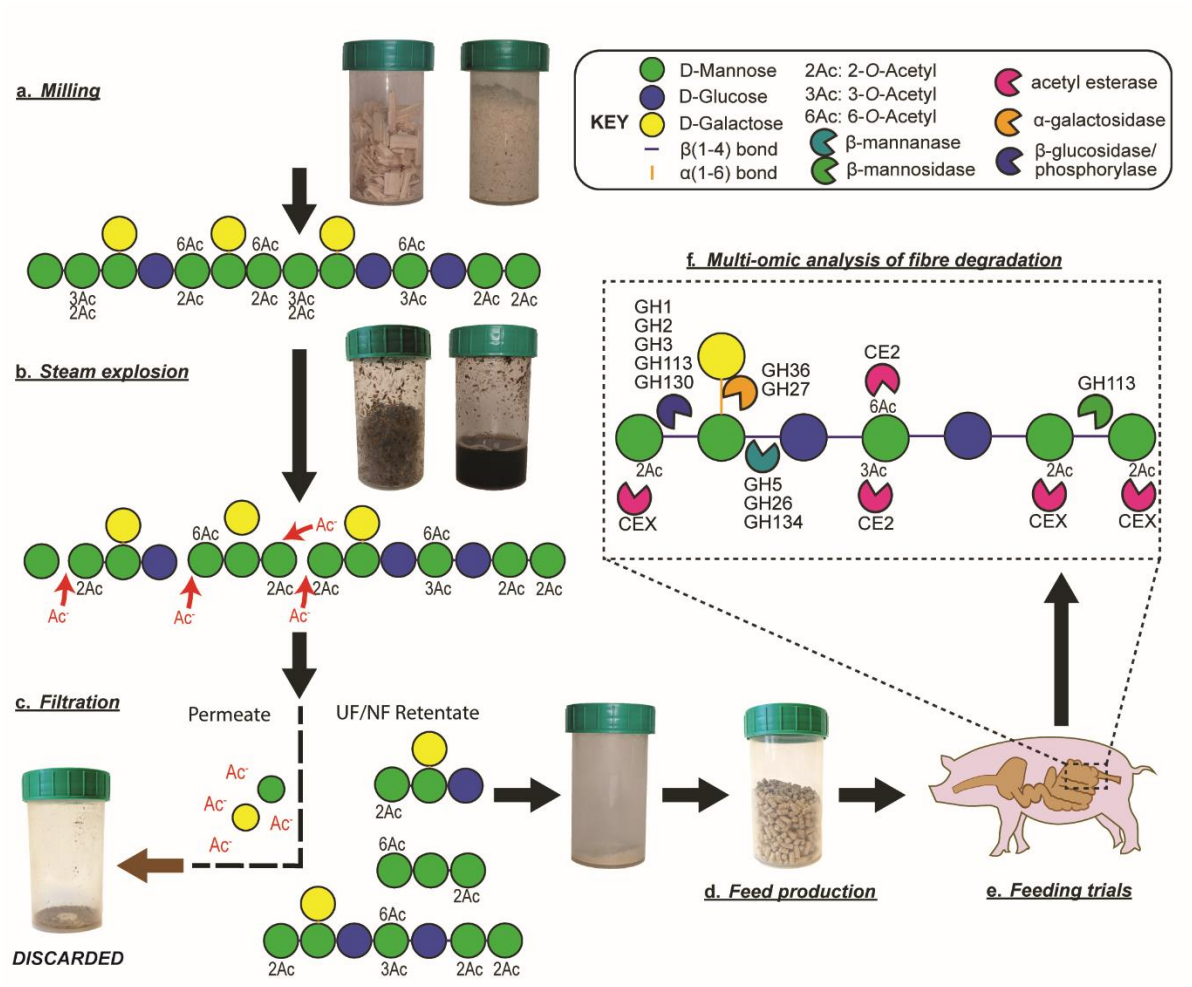
2-*O*-acetylated mannans are not commonly found in western dietary fibre sources, however 2-*O*-acetylations are present in acetylated galactoglucomannan (AcGGM), which is the main

hemicellulose in the secondary cell wall of Norway spruce (*Picea abies*)[10]. Here, we have utilized controlled steam explosion (SE), followed by ultrafiltration (UF) and fractionation to extract from spruce wood an unadulterated complex AcGGM fibre with a high degree of 2-*O*-, 3-*O*- and 6-*O*-acetylations[11], which is amenable to inclusion in animal feed production. To test (1) if our AcGGM fibre could specifically target *Roseburia* and *Faecalibacterium* species within a complex microbiome, and (2) if health benefits could be conferred to the host, we produced diets containing 0%, 1%, 2% and 4% AcGGM and fed semi-*ad libitum* to four cohorts of twelve weaning piglets. Host and gut microbiome effects were monitored temporally over a 28 day period, and metagenomics used to phylogenetically and functionally resolve the genomes of indigenous microbiota. Finally, quantitative metaproteomic analysis was used to monitor the metabolic and enzymatic response of the different microbiota to the varying AcGGM exposure. This approach deciphered how specific beneficial microbiota can be targeted and metabolically stimulated in complex gut microbiome ecosystems, with broader implications towards the evolving strategy of gut microbiome manipulations.

RESULTS AND DISCUSSION

Steam explosion allows scalable production of highly complex dietary mannan fibres from wood. Spruce galactoglucomannan consists of a backbone of β-(1,4)-linked mannose and glucose residues, decorated with α-(1,6) linked galactose branching's, and a large degree of esterification of the mannose residues by 2-*O*- and 3-*O*-, and 6-*O*- acetylation's[10] (**Fig 1a**). A crucial part of this study was the development of an efficient, large-scale extraction process entailing SE as well as ultra- and nano-filtration, which ultimately provided high quantities at high purity whilst not damaging the complexity of the AcGGM fibre (**Fig. 1b-c**). A total of 700kg of dry Norway spruce chips was processed using SE at conditions corresponding to a combined severity factor (R'_0) of 1.70. We produced 50 kg of oligo/polysaccharides for feed production (**Fig. 1c-e**), with a monosaccharide (Man: Glc: Gal) ratio of 4:1:0.6, which was in the form of β-mannooligosaccharides with DP of 2 to 10 and manno-

97 polysaccharides (DP of ≥11), and had degree of acetylation (DA = 0.36). This complexity matched the
98 enzymatic capabilities of mannan PULs encoded in human gut *Roseburia* and *Faecalibacterium* spp.[4,7]
99 and was predicted to match representatives of the same populations that are indigenous to porcine
100 gut ecosystems[12,13] (**Fig. 1f**).



101

102 **Fig. 1. Schematic representation and graphic illustration of the production pipeline for Norway spruce AcGGM**. **a**, Wood
103 chips were milled to increase the surface area exposed for hydrothermal extraction and washing the released
104 mannooligosaccharides. The naturally occurring form of mannan are long, highly branched and highly acetylated
105 polysaccharides. **b**, During hydrothermal pretreatment, acetic acid promotes hydrolysis of glycosidic bonds, releasing
106 monosaccharides, oligosaccharides, acetic acid and other breakdown products. **c**, Ultra filtration retained the longer,
107 complex oligosaccharides and discarded the monosaccharides, acetic acid and other steam explosion byproducts. **d**, The
108 purified mannan was incorporated into feed pellets at varying inclusion levels, produced by a conventional feed pelleting
109 process. **e**, Growth performance experiment and a feeding trial were conducted in a randomized block design, including four
110 inclusion levels of AcGGM. **f**, Multi-omic approaches were used to analyze the porcine gut microbiome in response to AcGGM
111 and determine if indigenous mannan PULs matched the glycan structure of the AcGGM fibre. Glycosidic bonds between the
112 β-(1,4)-mannose and glucose units in the backbone of AcGGM require hydrolysis by glycoside hydrolases (GH) from families
113 GH5 and GH26. GH36 and GH27 α-galactosidases are required to remove the α-(1,6)-galactose decorations. Single mannose
114 and glucose units are removed from the non-reducing end side of the oligosaccharides by enzymes from families GH1, GH2,
115 GH3, GH113 and GH130, while mannose in the reducing end side can be removed by GH113 family mannosidases. 3-*O*-, and
116 6-*O*- acetylations on mannan are removed by esterases from family CE2. A unique feature of particular beta-mannans is the
117 axially oriented 2-*O*- acetylation on mannose, which is the prevalent form of acetylation present on AcGGM used in this
118 study. 2-*O*-acetylations are removed by esterases homologous to the *Ri*CEX mannan-specific esterase from *Roseburia*
119 *intestinalis*, which was recently characterized by our group[4].

Spruce AcGGM altered the gut microbiome of weaned piglets in a dose-response manner. While pure culture[4], *in vitro*[7] and germ-free mice studies containing "mini-microbiota"[4] have shown that varieties of AcGGM can be metabolized by butyrate-producers, we wanted to test the accuracy of AcGGM to elicit a specific response in indigenous representatives of our target populations within a highly complex and competitive environment. To assess the potential benefit of our fibre to the host, we chose to introduce AcGGM to weaning piglets immediately after they transitioned from sow's milk to solid food. Weaning elicits a rapid, diet driven shift in the gut microbiome, which puts the animals at high risk of infection by intestinal pathogens such as enterotoxigenic *Escherichia coli* and *Salmonella enterica*[14]. These issues have been exacerbated by the banning of antibiotic use (Regulation No. 1831/2003) as growth promoters and prophylactics[15], meaning there is added urgency to develop a prebiotic compound to improve health and welfare of the animals during this crucial stage of development[16]. In this study, four separate cohorts of twelve weaned piglets were given a pelleted feed semi-*ad libitum*, which contained either 0% (control), 1%, 2% or 4% AcGGM to additionally determine the level necessary to elicit an effect on both the host and its microbiome. Fecal samples as well as animal health and growth performance metrics were taken before AcGGM administration (when piglets were assigned to pens), and subsequently at days 7, 14, 21 and 27 during the feeding trial. At day 28, the piglets were sacrificed and host gut tissue and digesta samples taken from the entire regions of the digestive tract (duodenum, jejunum, ileum, cecum and colon) for down-stream analysis. Despite widespread changes in the microbiome resulting from AcGGM inclusion, surprisingly no significant effects were observed on the host's physiology, with the average weight, feed conversion ratio, blood cell composition, T cell population, and colon morphology not differing between the control and AcGGM treatments (Supplementary Fig. 1, Supplementary Table 1-2).

Spatial and temporal microbiome changes were monitored using 16S rRNA gene analysis over the month-long trial and showed archetypical patterns with structural features of the gut microbiome varying depending on the specific gut region (Supplementary Fig. 2a). Inclusion of AcGGM into the

146 piglets feed caused a pronounced shift in the microbiome structural composition from the 21st day of
147 the trial onwards (Supplementary Fig. 2b, Supplementary Fig. 3). As expected, the AcGGM-effect was
148 more pronounced in the fibre-fermenting distal regions (cecum, colon) of the gut, where the relative
149 abundance of hundreds of phylotypes was observed to change (adjusted $p<0.05$) in response to
150 varying inclusion levels (**Fig. 2a**, Supplementary Table 4). Our target butyrate-producing populations
151 produced mixed results, whereby the relative abundance of *Faecalibacterium* affiliated phylotypes
152 increased in response to increasing levels of AcGGM (**Fig. 2b**, Supplementary Fig. 3), whereas
153 *Roseburia* affiliated phylotypes seemingly decreased (**Fig. 2b**, Supplementary Fig. 3). Reputable fiber-
154 fermenting populations affiliated to *Prevotella*, also showed varying responses, with individual
155 phylotypes increasing from 4% to 12% between the control and 4% AcGGM inclusion in both colon
156 and cecum (**Fig. 2b**, Supplementary Fig. 3). Interestingly, phylotypes affiliated to non-fibre degrading
157 taxa, such as *Catenibacterium*[17] and *Dialister*[18] demonstrated some of the highest dose-dependent
158 increases in relative abundance in response to AcGGM **(Fig. 2b**), indicating that other underlying
159 factors are likely dictating microbiome structure, besides fibre degradation.

160

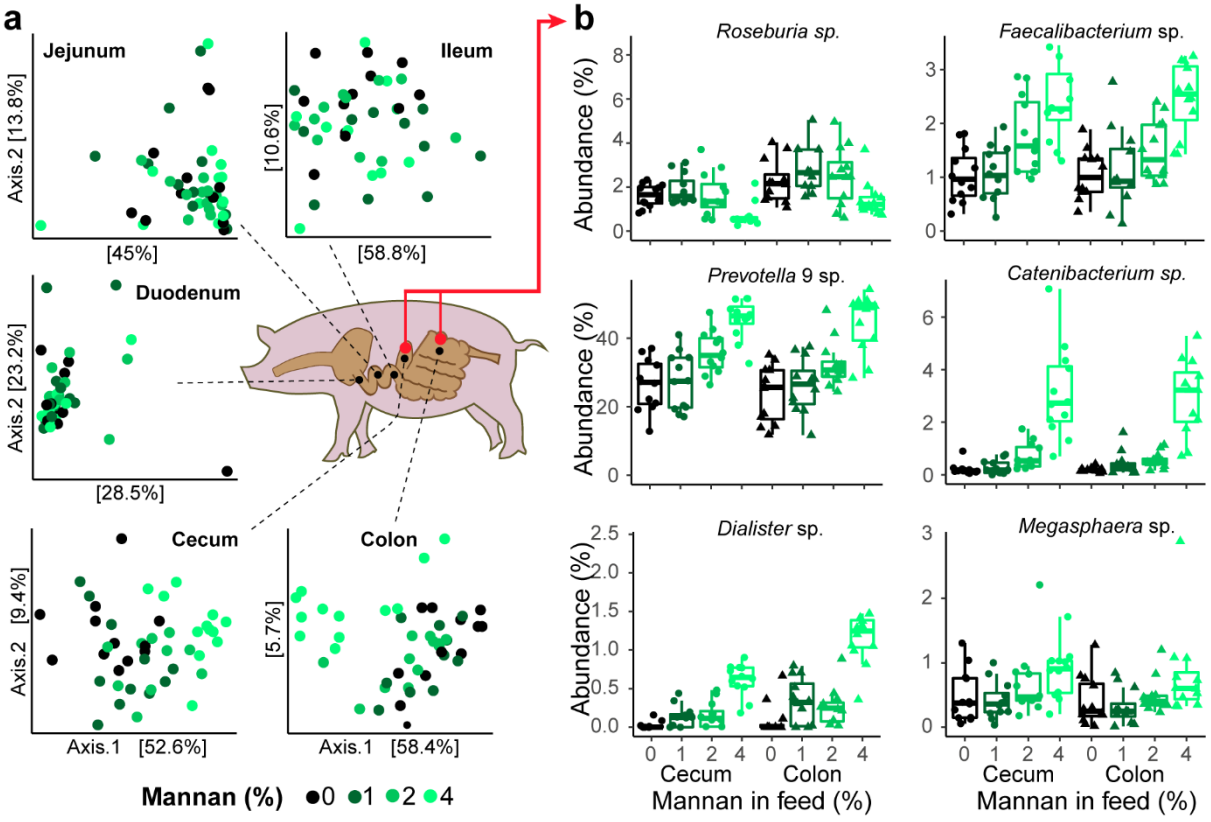


Fig. 2. Effect of AcGGM-containing diets on the pig gut microbiome. 16S rRNA gene amplicon analysis was used to monitor the effect the AcGGM fibre had on the gut microbiome structure of weaned piglets **a** Ordination plots of Bray-Curtis distances between microbial communities from pigs feed either the control or AcGGM diets (at varying inclusion levels: 1%, 2% or 4%). Samples were collected at day 27 of the feeding trials from various sections of the small and large intestine. The AcGGM-effect was more pronounced in the fibre-fermenting distal regions (cecum, colon) of the gut. **b** Relative abundance of selected bacterial genera in either cecum or colon samples collected from pigs fed AcGGM diets with varying inclusion levels. The relative abundance of the genus *Faecalibacterium* was enhanced by the inclusion of AcGGM, as were phylotypes affiliated to *Prevotella* group 9, *Catenibacterium* and *Dialister*. Despite being suspected as an active mannan degrader, *Roseburia*-affiliated phylotypes decreased in abundance with increasing % AcGGM.

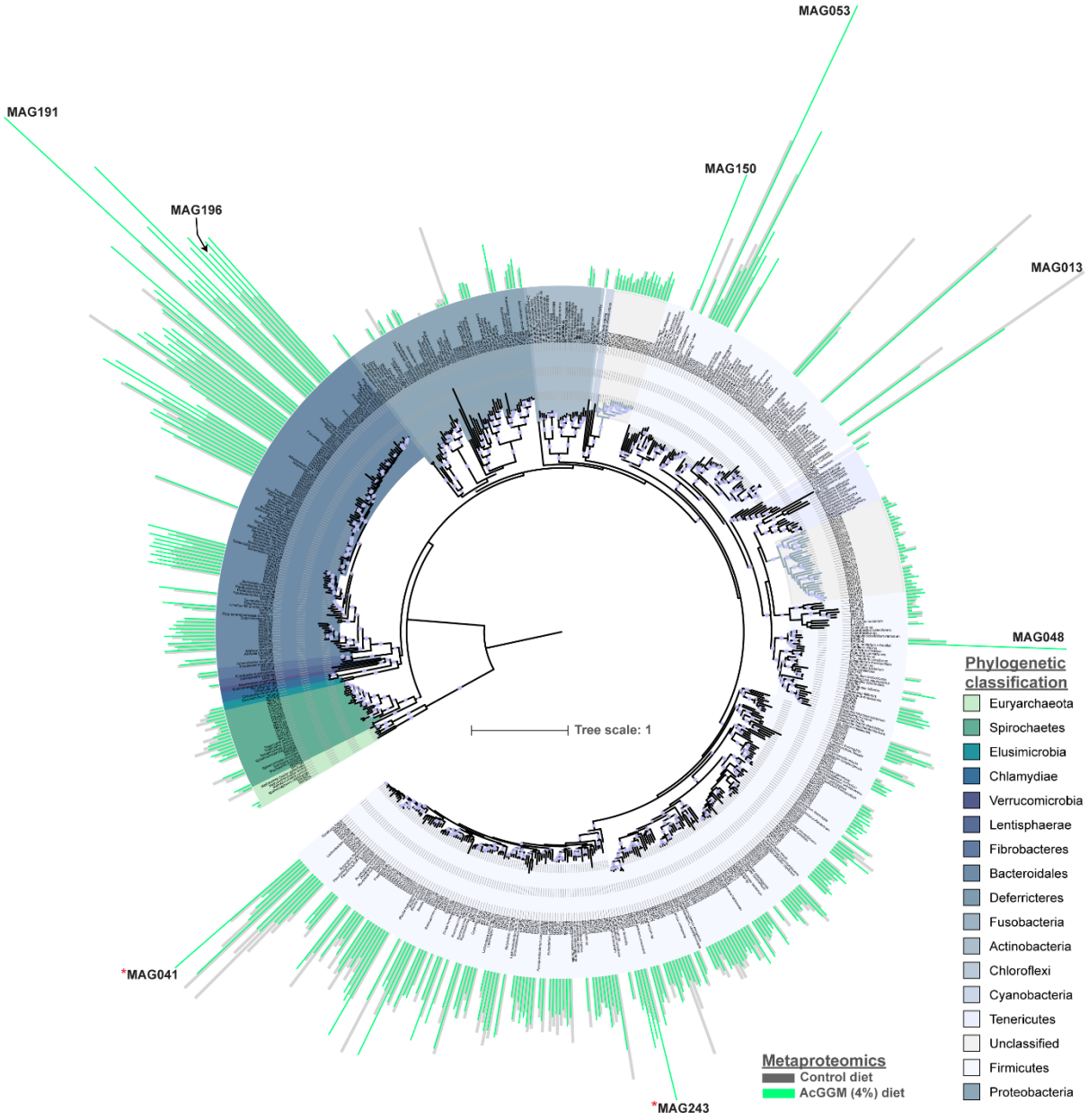
Targeted mannan PULs and butyrate-producing pathways are actively detected in *Faecalibacterium*- and *Roseburia*-affiliated populations within the colon of AcGGM-fed pigs.

To determine the effect AcGGM had on microbiome function, we analyzed 211.36 Gbps of Illumina HiSeq sequencing data obtained from the colon samples of each pig fed the control and 4% AcGGM diets (average: 8.8 Gbps per sample) (Supplementary Table 3). The metagenomic data was assembled into 355 metagenome-assembled genomes (MAGs), of which 145 had >90% completeness and were considered high quality according to the Genomics Consortium Standards[19] (Supplementary Fig. 4a). Phylogenetic relationship of the MAGs were inferred from a concatenated ribosomal protein tree

(Newick format available in Supplementary Dataset 1) that were constructed using MAGs from this study and 293 closely related reference genomes. Because our primary goal was to elucidate if our target butyrate-producing populations were activated in response to AcGGM, we conducted metaproteomic analysis on randomly selected colon samples from four control and four 4% AcGGM fed pigs, and mapped 8515 detected protein groups back against our MAG database to identify functionally active populations **(Fig. 3**, Supplementary Table 5). Community-wide analysis of the MAG genetic content (Supplementary Fig. 4b) from each sample, and distribution of their detected proteins (**Fig. 3**, Supplementary Fig. 4c), further supported our 16S rRNA gene analysis, reiterating that the microbiomes from piglets fed the control and 4% AcGGM diets were distinct.

Our MAG-centric multi-omic approach gave clear indications as to what effect the AcGGM fibre had on putative butyrate-producing *Roseburia* and *Faecalibacterium* populations in the distal gut of pigs. Ten MAGs clustered with representative *Roseburia* spp. genomes (**Fig. 3**), which reflected the multiple *Roseburia*-affiliated phylotypes that were predicted with our 16S rRNA gene analysis. (Supplementary Fig. 3). For the most part, a lower number of *Roseburia*-affiliated proteins were detected in AcGGM-fed pigs (**Fig. 3**), reintegrating our initial observations that AcGGM negatively affected *Roseburia* populations (**Fig. 2b**). However, within one specific *Roseburia*-affiliated (MAG041), we detected a higher number of total affiliated protein groups in the 4% AcGGM pig samples compared to the control (avg=94 v 53, Supplementary Table 5). Closer examination of MAG041 revealed a putative CE2/CEX-containing mannan-degrading PUL that was absent in the other *Roseburia*-affiliated MAGs and was differentially expressed in the AcGGM diet **(Fig. 4**). Importantly, the MAG041 mannan PUL encoded gene synteny to the *R. intestinalis* strain L1-82 PUL we recently biochemically characterized in detail[4] **(Fig. 4**). The predicted multi-modular mannanase (CBM27-GH26-CBM23) in the MAG041 mannan PUL is homologous to the GH26 in *R. intestinalis* L1-82 (48% identity over 87% of the sequence), and can be presumed to fulfill the same function – "selfishly" breaking down AcGGM fibres at the cell surf prior intracellular transport. Besides the detection of GH26 and esterases in AcGGM-fed pigs, other mannan-specific enzymes also responded to the dietary

206 shift within the MAG041 mannan PUL, including a phosphoglucomutase, a multiple-sugar binding
207 protein, GH130.1 4-*O*-β-D-mannosyl-D-glucose phosphorylase and a GH130.2 β-1,4-manno-
208 oligosaccharide phosphorylase (**Fig. 4**).



209

210 **Fig. 3. Phylogeny and metaproteomic detection of 355 MAGs sampled from the distal gut of weaned piglets.** Maximum
211 likelihood tree of 22 concatenated ribosomal proteins were created with 239 closely related reference genomes and
212 genomes reconstructed in this study. Branches are shaded with colour to highlight phylum-level affiliations (see legend).
213 Coloured bars on the outside of the tree depict the total number of proteins detected for each MAG in samples collected
214 from pigs fed either the control- (grey) or 4% AcGGM-diet (green). Purple circles on the inside of the tree represents nodes
215 with bootstrap support ≥ 70%, relative to size. MAG041 and MAG243 were found to encode CE2/CEX-containing mannan
216 PULs are indicated by *. All MAGs depicted in **Fig. 5** are listed. The full tree in Newick format is provided in Supplementary
217 Dataset 1.

218

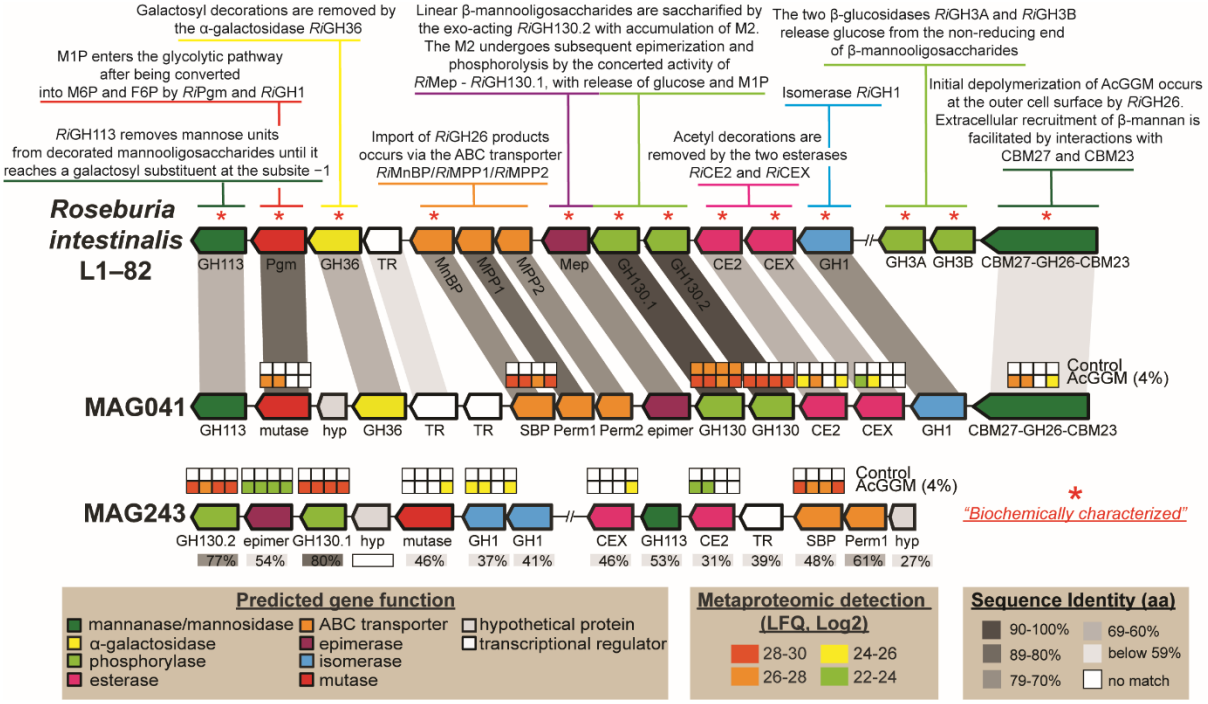


Fig. 4. Metaproteomic detection of CE2/CEX-containing mannan-PULs encoded in *Roseburia*- (MAG041) and *Faecalibacterium*- (MAG243) affiliated MAGs in pigs fed with the control or 4% AcGGM diet. Predicted gene organization and annotated gene function is colour-coded and largely derived from previous biochemical and structural characterization of the mannan degradation cluster (characterized genes indicated with *) in *R. intestinalis* L1-82[4]. Gene synteny and identity % between mannan PULs found in in *R. intestinalis* L1-82, MAG041 and MAG243 are indicated in grey boxes. Heat maps above detected enzymes show the LFQ detection levels for the four replicates sampled in control and 4% AcGGM-fed pigs. LFQ values of proteins from both clusters are in Supplementary Table 6. The predicted multi-modular mannanase (CBM27-GH26-CBM23) from MAG041 was the only extracellular protein in the locus, and the only extracellular mannanase expressed in response to AcGGM inclusion.

In contrast to *Roseburia*-affiliated MAGs, only one MAG clustered with *F. prausnitzii* (MAG243, **Fig. 3**), inferring that the multiple phylotypes that were predicted with our 16S rRNA gene data (Supplementary Fig. 3) encode high genome similarity and likely co-assembled into a representative population-level MAG. Our metaproteomic analysis predicted that MAG243 was more active in the distal gut of pigs fed 4% AcGGM (avg=32 vs 77 detected protein groups, **Fig. 3**, Supplementary Table 5), as was its CE2/CEX-containing mannan PUL, which was broadly detectable in the presence of AcGGM but absent in the control samples (**Fig. 4**, Supplementary Table 6). While, the MAG243 mannan PUL contained two GH130 manno-oligophosphorylases, a mannose 6-phosphate isomerase, phosphoglucomutase and two carbohydrate esterase (CEX and CE2), it lacked a GH26 mannanase representative, which suggests that *F. prausnitzii* is likely preferentially targeting the shorter

acetylated manno-oligosaccharides that form part of the AcGGM creation (**Fig. 1**). In addition to the mannan-PULs of MAG041 and MAG243 being activated in AcGGM-fed pigs, their butyrate-producing pathway were also detected at high levels, based on label-free quantification (LFQ) scores of detected proteins **(Fig. 5**, Supplementary Table 6), suggesting that both populations can convert mannan to butyrate (Supplementary Table 6).

Removal of 2-*O*-, 3-*O*- and 6-*O*-mannose acetylations are the key enzymatic activities required for accessing AcGGM.

A crucial step in the utilization of mannans as an energy source is the deacetylation of 2-*O*-, 3-*O*- and 6-*O*-mannose residues, which subsequently grants access to the sugar-containing backbone of the fibre. In *R. intestinalis* L1-82, AcGGM deacetylation occurs via the synergistic actions of two carbohydrate esterases (*Ri*CE2 and *Ri*CEX) that exert complementary specificities[6]. MAG041 and MAG243 both encoded CE2 homologues within their mannan PULs, sharing 63% and 31% identity (respectively) to *Ri*CE2, which has demonstrated activity on 3*O*-, (4*O*-) and 6*O*-acetylations, and is mannan specific[6]. For CEX, MAG041 and MAG243 homologs shared 65% and 46% identity (respectively) with *Ri*CEX, including the active site residues and the aromatic stacking tryptophan (Trp326), which in *Ri*CEX are associated with 2-*O*- acetylation specificity[6]. Broader screens of our MAG data revealed other CE2/CEX-containing PULs within Firmicute-affiliated MAGs from the pig colon microbiome (Supplementary Fig. 5), however aside from MAG041 and MAG243, they originated from populations that were not as metabolically detectable via metaproteomics in any of the control or AcGGM diets (**Fig. 3**). Finally, the differential proteomic detection of MAG041 and MAG243 CEs in pigs fed AcGGM diets (**Fig. 4**), strengthened our hypothesis that both these populations can accommodate the unique features of the AcGGM fibre and are actively engaging in its utilization *in vivo*.

Despite the apparent specificity of AcGGM, we observed a pronounced "butterfly effect"

Besides the activation of specific butyrate-producers that encoded CE2/CEX-mannan PULs, the AcGGM diet also altered protein expression in multiple populations within the distal gut of weaned piglets. For the most part, proteins originating from the fibre-degrading *Prevotella* were more detectable in AcGGM-fed pigs (**Fig. 3**), with MAG191 in particular accountable for the highest levels of detectable proteins in our datasets (Supplementary Table 5). Pathway annotation of abundantly detected *Prevotella* populations (such as MAG191, MAG196, MAG285, see **Fig. 5**, Supplementary Table 6) indicated active metabolism of dietary fibres such xylans, starch, cellobiose, α-galactans and mannose sugars as well as acetate, succinate and/or propionate production, which were all detected with higher LFQ scores in AcGGM-fed pigs (**Fig. 5**, Supplementary Table 6). Several mannan-targeting PULs were identified in *Prevotella*-affiliated MAGs that were configured in an archetypical "Bacteroidetes-format", which combines outer-membrane transport and carbohydrate-binding lipoproteins (SusC/D-like) as well as CAZymes[20] (Supplementary Fig. 5a). In particular, a PUL recovered from MAG196 encoded predicted SusC/D-like lipoproteins, mannanases (GH26, GH5_7), mannosyl-phosphorylases (GH130) and an esterase, although neither the mannanases nor the esterase were detected in the metaproteomes recovered from the AcGGM-fed pigs (**Fig. 5**, Supplementary Fig. 5, Supplementary Table 6). In addition, we speculate that MAG196 and MAG191 are perhaps capable of metabolizing elements of the AcGGM fibre such as the α-galactose side-chain or deacetylated manno-oligosaccharides, which was inferred via detected GH36 and GH130 representatives (**Fig. 5**, Supplementary Fig. 5, Supplementary Table 6).

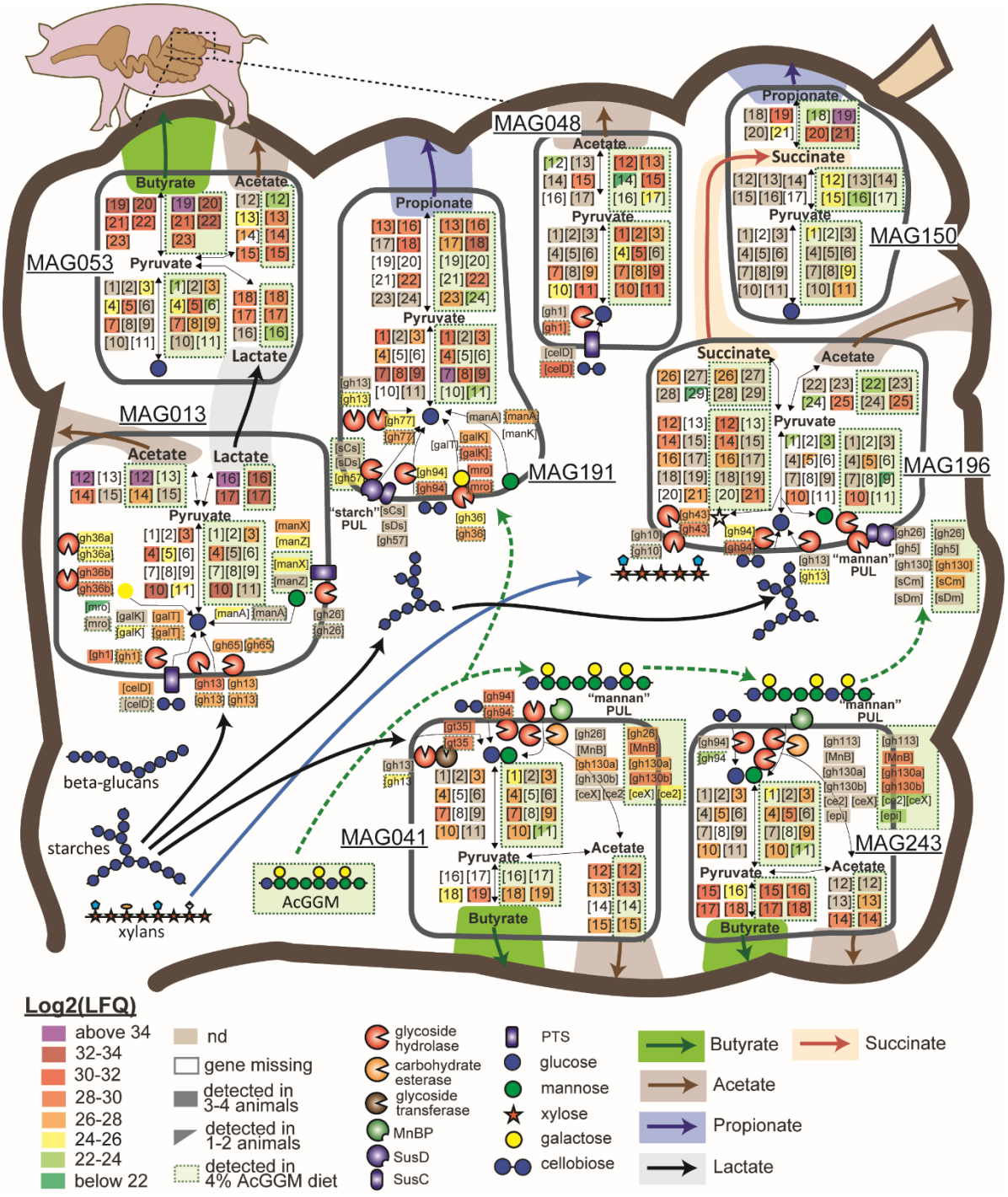


Fig. 5. Selected metabolic features of the porcine colon microbiome in response to AcGGM dietary intervention, as inferred from genome and proteome comparisons. The different metabolic pathways (fibre deconstruction, glycolysis, pentose-phosphate pathway and SCFA production) are displayed for each population MAG. Graphical representation of pathways, enzymes, CAZymes, and cellular features are based on functional annotations that are depicted as numbered or abbreviated gene boxes, which are additionally listed in Supplementary Table 6. Metaproteomic analysis (detected genes and enzyme complexes) is highlighted according to the different LFQ values of the detected proteins sampled at a single time-point (28 days) either from animals fed the control or 4% AcGGM diet (green perforated boxes). The main dietary fibres (starches, xylans, glucans and mannans) and SCFAs (butyrate, acetate, propionate, lactate and succinate) are represented by large colored arrows. Gene names and abbreviations are also provided in Supplementary Table 6.

In spite of the specificity of the AcGGM fibre to match selected mechanistic features of our target populations, our evidence suggests the effect of the AcGGM dietary intervention reverberated further down the microbial trophic networks that support conversion of dietary fibre into keystone short-chain fatty acids (SCFAs) that are of nutritional value to the host animal. Mirroring our 16S rRNA gene analysis, MAGs affiliated to the genera *Dialister* (MAG150), *Catenibacterium* (MAG048), *Lactobacillus* (MAG013) and *Megasphaera* (MAG053) demonstrated the largest transformation in response to the AcGGM diet, although none were found to encode CE2/CEX-containing mannan PULs (**Fig. 2-3**, **Fig. 5**, Supplementary Table 6). However, the MAG048 proteome increased in detection by ~4 fold in AcGGM-fed pigs, which included a putative sugar phosphotransferase system (PTS) and GH1 phospho-β-glucosidases (EC 3.2.1.86) that are predicted to catalyze the phosphorylation of di-oligosaccharides (such as cellobiose and mannobiose) and hydrolyze the PTS-transported sugars into D-glucose and D-glucose 6-phosphate. Concomitantly, glycolysis and acetate-producing pathways from MAG048 were also highly detected in AcGGM-fed pigs (**Fig. 5**, Supplementary Table 6), suggesting this population is advantageously consuming oligosaccharides that have either been generated via the actions of other fibre-degrading populations or have become available via new ecological niches that have been created via the AcGGM-derived structural shifts in the microbiome. Non-fibre degrading populations also reacted to the AcGGM diet, with both MAG053 and MAG150 predicted via our multi-omics approach to metabolize SCFAs such as lactate and succinate that were generated “in-house” by *Lactobacillus*- and *Prevotella*-affiliated populations, and produce butyrate and propionate, respectively (**Fig. 5**, Supplementary Table 6).

CONCLUSION

Mannan from woody biomass has great potential in functional diets for both human and livestock alike. It is cheap, renewable, does not compete with food sources, and common extraction methods such as SE and UF can be readily adapted to industrial scale production. The AcGGM fibre

320 described here could be produced from forestry by-products such as sawdust or lumber waste, while
321 the solid biomass discarded in this production process contains valuable, steam-exploded cellulose,
322 suitable for inclusion in other processes such as the production of biofuels and platform chemicals via
323 enzymatic treatment and fermentation.

324 Here, we characterize the impact of the AcGGM structural configuration on microbial uptake and
325 metabolism within the distal regions of the digestive tract, with the key driver of AcGGM selectivity
326 being the presence of acetylations of mannan, as well as carbohydrate composition and size[21].
327 Preserving the complexity of AcGGM resulted in a highly specific, dose-dependent shift in the
328 composition of the colon microbiome from weaned piglets, with no diverse effect on host growth
329 performance or health status. Our integrated multi-omics analysis, showed that the AcGGM fibre
330 activated a metabolic response in specific *Roseburia* and *Faecalibacterium* populations *in vivo*, as it
331 did in the previous *in vitro* experiments[4,7], with both populations expressing proteins from highly
332 sophisticated CE2/CEX-containing mannan PULs that are homologous to a biochemically characterized
333 representative in *R. intestinalis* L1-82. In conclusion, our data provide a foundation for modulatory
334 strategies to design and match custom dietary fibres to unique enzymatic features of their target
335 organisms, although they bring awareness to the fact that the greater network of interconnected
336 metabolic exchanges and trophic structures inherent to the gut microbiome are highly susceptible to
337 minor dietary interventions.

338

DATA AVAILABILITY

All sequencing reads have been deposited at the NMBU sequence read archive under BioProject PRJNA574295, with specific numbers listed in Supplementary Table 3. All annotated MAGs are publicly available via doi: 10.6084/m9.figshare.9816581. The proteomics data has been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[22] with the dataset identifier PXD015757.

ACKNOWLEDGEMENTS

We are grateful for support from The Research Council of Norway (Bionær program 244259, FRIPRO program, PBP: 250479), as well as the European Research Commission Starting Grant Fellowship (awarded to PBP; 336355 - MicroDE). The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities.

AUTHOR CONTRIBUTIONS

L.M., J.C.G, P.B.P. and B.W. conceived the study, performed the primary analysis of the data and wrote the paper (with input from all authors). L.L. and M.Ø. designed, performed and analyzed the animal experiments. S.L.L.R., L.H.H., M.Ø.A. and J.D. generated the data and contributed to the data analyses. N.T., V.L. and B.H. annotated and curated the MAGs and identified carbohydrate-active enzymes.

COMPETING INTERESTS

The authors declare there are no competing financial interests in relation to the work described.

MATERIALS AND METHODS

364 **Animals, diets and experimental design.**

365 Animal care protocols and experimental procedures were approved by the Norwegian Animal
366 Research Authority, approval no. 17/9496, FOTS ID 11314 and treated according to institutional
367 guidelines. A total of 48 cross bred piglets (Landrace x Yorkshire), 24 male and 24 female, with an
368 average initial body weight (BW) of 9,8 ± 0,5 kg, weaned at 28 days of age were sorted by litter, sex
369 and weight and randomly divided into 12 groups of four animals each (one diet per pen), but pigs were
370 housed individually during mealtime. The animals were housed in an environmentally controlled
371 facility with plastic flooring and a mechanical ventilation system. The temperature of the room was
372 maintained at 22°C.

373 Piglets were fed cereal-based diets containing increasing levels of AcGGM in the diets (1, 2
374 and 4%). Diets were pelleted with a 3 mm diameter feed formulated to meet the requirements for
375 indispensable amino acids and all other nutrients (NRC, 2012). The composition of diets is listed in
376 Supplementary Table 7. Pigs were fed semi-ad libitum twice a day at a feeding level equal to about 5%
377 of body weight. To evaluate growth performance, the BW of each pig was recorded at the beginning
378 and once a week. Feed consumption were recorded on an individual pig basis during the experiment
379 to calculate individual weight gain and feed intake. After each meal, feed leftovers were registered,
380 dried and subtracted from the total feed intake.

381

382 **Production of AcGGM.**

383 AcGGM oligosaccharides for the feeding trial were produced from Norway spruce chips milled
384 with a hammer mill to <2 mm size. Wood chips were then steam-exploded on a small pilot scale steam
385 explosion rig (100L reactor vessel) at the Norwegian University of Life Sciences (NMBU). The steam
386 explosion was conducted in batches of approximately 6kg dry matter, 14.5 bar pressure (equivalent
387 to 200° C), with 10 minutes residence time. The pH in the collected biomass slurry after the steam

388 explosion was ~3.7, which corresponds to a combined severity factor R'_0=1.70 for the process. The
389 severity was calculated by $R'_0 = (10^{-pH}) \times (t \times e(^{Texp-100)/14.75})^{23}$. Steam exploded wood was collected in
390 50 L plastic buckets that were topped up with hot (~70° C) water. The slurry was transferred to a
391 60Lcider press (Speidel, Germany) and the liquid fraction was pressed out. Milled wood was collected,
392 soaked in hot water again, and pressed for the second time. The liquid fraction was collected and
393 recirculated through a bag filter 50µm pore WE50P2VWR (Allied filter systems, England) partly filled
394 with the wood particles as a filter aid. Once free of floating wood particles, the liquid fraction of
395 hemicellulose was filtered through a 5-kDa spiral wound Polysulphone/polyethersulphone
396 ultrafiltration membrane, GR99PE polyester (Alfa Laval, Denmark) that was deliberately fouled to
397 prevent larger oligosaccharides from running through the permeate, using a GEA pilot-scale filtration
398 system Model L (GEA, Denmark). The fraction retained by the membrane was concentrated by
399 nanofiltration using a TriSep XN 45, which had a higher efficiency for permeating water. The filtrate
400 was further concentrated by vacuum evaporation (set to 65 °C) and the concentrate was freeze-dried
401 and homogenized with a grain mill. The final product consisted of 0.9% rhamnose, 2.7% arabinose,
402 13.7% xylose, 58.9% mannose, 14.9% glucose and 9.4% galactose (determined by gas chromatography
403 as alditol acetates after sulfuric acid hydrolysis as described previously[24]). AcGGM contained 0.73 %
404 ash and 2.4% protein (quantified from total nitrogen by the Kjeldahl method). The Man:Glc:Gal ratio
405 in the mannan was 4:1:0.6, and the DA=0.36 (determined by acetate release from NaOH treated
406 AcGGM by the same method as described in the SCFA section below). The dry matter content was
407 determined by drying 0.2g of sample at 105°C for 20 hours. The remaining sample was burned at 600°C
408 for 24 hours in an oven (Carbolite, Sheffield, England) to determine ash content. All measurements
409 were performed in triplicates.

410

411 **Fecal scoring.**

412 During the experiment, fecal consistency was assessed using a scoring system developed by Pedersen
413 and Toft[25] to improve and help standardize current protocols for clinical characterization of fecal
414 consistency. The scoring was based on the following 4 consistency categories: score 1 = firm and
415 shaped, score 2 = soft and shaped, score 3 = loose and score 4 = watery. Samples with score 3 or 4 are
416 considered diarrheic. Daily fecal scores for each pen were recorded throughout the trial.

417

418 **pH measurements.**

419 The pH of digesta samples from duodenum, jejunum, ileum, cecum and colon were measured
420 immediately after slaughter. Samples were placed in universal containers and pH measurements made
421 using an Inolab pH7110 pH meter (WTW, Germany).

422

423 **Blood sampling and flow cytometry.**

424 Blood samples were collected from the same six piglets per diet at 0, 7 and 27 feeding days. The blood
425 samples were taken 1-2 hours post-prandial by venipuncture in the jugular vein while pigs were kept
426 on their backs. Non-heparinised and K_3EDTA vacuum tubes (Beckman Dickson Vacutainer System)
427 were used to recollect serum and whole blood. Serum was isolated immediately by centrifugation at
428 1,500 x *g* at 4°C for 15min. Serum samples were split in PCR-tubes (200 µL) and stored at -80°C until
429 analysis. Hematological and clinical analyses were performed with an Advia® 2120 Hematology System
430 using Advia 2120 MultiSpecies System Software and clinical chemistry analyses were performed with
431 Advia 1800 Chemistry System (both from Siemens AG Healthcare Sector).

432 For flow cytometry analysis, whole blood was diluted 1:1 in RPMI 1640 and kept on ice until single
433 cells isolation. For the isolation of peripheral blood mononuclear cells (PBMCs) blood was purified by
434 centrifugation in a Ficoll gradient (Kreuzer et al. 2012). Then, isolated PBMCs were incubated with
435 Fixable Yellow Dead Cell Stain Kit (Life Technologies, Thermo Fisher Scientific Inc.) followed by primary

436 monoclonal antibodies (mAbs), brief incubation with 30% normal pig serum to block Fc-receptors, and
437 finally fluorescence-labeled secondary antibodies (Abcam plc, UK). To detect the intracellular CD3
438 epitope, surface-labeled cells were permeabilized with Intracellular Fixation and Permeabilization
439 Buffer Set (eBioscience, Affymetrix Inc.) according to the manufacturer's instructions. Labeled cells
440 were analyzed on a Gallios Flow Cytometer (Beckman Coulter, Inc.) and data were processed using
441 Kaluza 1.5 software (both Beckman Coulter, Inc.). Cell gates were designed to select for single and
442 viable mononuclear cells. Defined markers were used to identify the different immune
443 subpopulations. For monocytes, antibodies against CD45, CD3, CD14, CD163 and MHCII were used. To
444 analyze regulatory T cells (T reg) the following antibodies were used: CD45, CD3, TCR γ/δ, CD4, CD8,
445 FOXp3 and CD25, while CD45, CD8, NKp46, CD4, CD8, Ki67 and CD27 were used to identify T and NK
446 cells.

447

448 **Analysis of Serum Cytokines: MULTIPLEX**

449 Expression of GMCSF, IFNG, IL-1A, IL1B, IL-1RA, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12, IL-18 and TNFα were
450 measured in serum samples using MILLIPLEX MAP Porcine Cytokine and Chemokine Magnetic Bead
451 Panel - Immunology Multiplex Assay (Merck Millipore) following the manufacturer instructions. The
452 measurement was performed using a Bio-Plex MAGPIX Multiplex Reader (BIO-RAD).

453

454 **Small Intestine Morphology.**

455 The samples of the small intestine were collected on day 0 and 28 for determination of intestinal
456 morphology and integrity. Intestinal morphological measurements included the following indices:
457 villus height (VH), crypt depth (CD) and VH:CD. Mean values of VH, CD and their ratio were calculated.
458 Histology evaluation was performed by the Veterinary Histophalogy Center, VeHiCe, Chile.

459

460 **Microbial Sampling**

461 Fecal samples were collected from 6 piglets per experimental group (n=24) at days 0, 7, 14,
462 21, and 27 post-weaning. At the end of the trial, all piglets (n=48) were sacrificed, and samples were
463 collected from the lumen of the duodenum, jejunum, ileum, cecum, and colon. Samples were obtained
464 within the first 15 minutes after the piglets were sacrificed and the samples were flash-frozen in liquid
465 nitrogen and stored at -80°C until DNA extraction.

466

467 **DNA Extraction**

468 DNA was extracted with a MagAttract PowerMicrobiome DNA/RNA Kit (MO BIO Laboratories
469 Inc., Carlsbad, CA, USA) according to the manufacturer instructions, except for the bead beating step
470 where we used a FastPrep-96 Homogenizer (MP Biomedicals LLC., Santa Ana, CA, USA) at maximum
471 intensity for a total of 2 minutes in 4 pulses of 30s with a 5 minute cooling period between each pulse.
472 A KingFisher Flex DNA extraction robot was used for the automated steps of the protocol. The
473 extracted nucleic acids were quantified with a Qubit Fluorimeter and the Qubit dsDNA BR Assay Kit
474 (Thermo Fisher Scientific, Waltham, MA, USA) and stored at -80°C.

475

476 **16S Amplicon Sequencing and Analysis**

477 16S amplicon sequence data was obtained for all fecal and intestinal samples. The V3-V4
478 region of the 16S rRNA gene was PCR amplified using the primers Pro341F (5'-CCT ACG GGN BGC ASC
479 AG-3') and Pro805R (5'-GAC TAC NVG GGT ATC TAA TCC-3'), to which the MiSeq adaptors were
480 additionally incorporated on the 5' ends[26]. The 25 µL PCR reactions consisted of 1X iProof High-Fidelity
481 Master Mix (Biorad, Hercules, CA, USA), 0.25 µM primers, and 5 ng template DNA. PCR thermal cycling
482 began with a hot start step at 98 °C for 180 s and was followed by 25 cycles of 98 °C denaturation for
483 30 s, 55 °C annealing for 30 s, and 72 °C extension for 30 s, followed by a final, 300 s extension step at

72 °C. Amplicons were individually purified with AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA) and indexed with the Nextera XT Index Kit v2 (Illumina, San Diego, CA, USA) according to the Illumina protocol for 16S metagenomic sequencing library preparation. Next, equal volumes from each indexing reaction were pooled together, and the pool was purified with AMPure XP beads. The purified amplicon pool was then quantified with a Qubit Fluorimeter, diluted, mixed with 15% PhiX Control v3 (Illumina), and denatured according to the aforementioned Illumina protocol. The denatured library was sequenced on the Illumina MiSeq platform using the MiSeq Reagent Kit v3 (600 cycle). Data were output from the sequencer as demultiplexed FASTQ format files.

Processing of the data was done with a combination of standalone programs, QIIME[27] MOTHUR[28] and the R package Phyloseq[29]. To process the data, the paired end reads for each sample were merged with PEAR[30], specifying a minimum assembly length 400, maximum assembly length 575, minimum overlap 50, and no statistical test. Then, PRINSEQ[31] version 0.20.4 was used to filter low quality reads by requiring a minimum quality score of 10 for all bases and a minimum mean quality of 30. Primer sequences were trimmed in MOTHUR version 1.36.1, and chimeric sequences were identified and filtered out using QIIME version 1.9.1. Next, open reference $OTU_{0.97}$ clustering[32] was performed with VSEARCH[33] version 2.3.2 and the Silva database[34] release 128 as the taxonomy reference. Then, the QIIME core diversity analyses script was run. Differentially abundant phylotypes were identified in both cecum and colon for the control vs. 4% AcGGM samples using both the MetagenomeSeq fitZIG and DESeq2 negative binomial algorithms via the QIIME wrapper. The OTU table, phylogenetic tree, representative sequences, and taxonomy from QIIME were incorporated along with the sample metadata into a Phyloseq version 1.22.3 object in R for data exploration and visualization.

Whole Metagenome Sequencing and Analysis

Whole metagenome sequencing was performed at the Norwegian Sequencing Centre on 2 lanes of the Illumina HiSeq 4000 to generate 2 X 150 paired-end reads. TruSeq PCR-free libraries were prepared for 12 control and 12 AcGGM (4%) samples from the colon. All 24 samples were run in both lanes to eliminate the potential for lane-specific sequencing bias. FASTQ format files were received from the sequencing center, and prior to assembly, these were quality filtered with Trimmomatic[35] version 0.36 whereby TruSeq adaptor sequences were eliminated, sequences were required to have an average quality score above 20, leading and trailing bases with quality below 20 were removed, sequences with average quality score below 15 in a 4-base sliding window were trimmed, and the minimum read length was required to be 36 bases. Individual sample assembly was accomplished with metaSPAdes[36] version 3.11.1. MegaHIT[37] version 1.1.3 was used for co-assembly of all 24 samples together as well as co-assembly of the 12 control samples together and the 12 4% AcGGM samples together. MetaBAT[38] version 0.26.3 was used to bin the assemblies, and dRep[39] was used to dereplicate the multiple assembly and binning combinations to produce an optimal set of MAGs. MASH[40] version 2.0 used to compare the similarity of the 24 metagenomes by calculating pairwise Jaccard distances which were imported into R for NMDS ordination and visualization. Completeness and contamination was determined for each MAG using CheckM[41] version 1.0.7. Feature and functional annotation were completed with the Prokka pipeline[42] version 1.12, and the predicted protein sequences from all 355 MAGs were concatenated to create the metaproteomics reference database. Resulting annotated open reading frames (ORFs) were retrieved, further annotated for CAZymes using the CAZy annotation pipeline with libraries from July 2018 database release[43,44], and subsequently used as a reference database for the metaproteomics (with the exception of glycosyltransferases).

Metaproteomics

Proteins were extracted from each sample in quadruplicate by the following method. An aliquot (1 g) of colon digesta from pigs fed either a control diet or a diet supplemented with 4% β-mannan was dissolved 1:1 (w/v) in 50 mM TrisHCl, pH 8.4.

Lysis was performed using a bead-beating approach whereby glass beads (size ≤ 106 µm) were added to the colon digesta slurry and cells were disrupted in 3 x 60 second cycles using a FastPrep24 (MP Biomedicals, Santa Ana, CA, USA). Debris were removed by centrifugation at 16.600 x *g* for 20 minutes and proteins were precipitated overnight in 16% ice-cold TCA. The next day, proteins were dissolved in 100 µL 50 mM TrisHCl, pH 8.4 and concentration was determined using the Bradford protein assay (Bradford Laboratories, USA) using bovine serum albumin as a standard. Fifty milligrams of protein was prepared in SDS sample buffer, separated by SDS-PAGE using an Any-kD Mini-PROTEAN gel (Bio-Rad Laboratories, Hercules, CA, USA) and stained using Coomassie Brilliant Blue R250. The gel was cut into 6 slices and reduced, alkylated and digested as described previously[45]. Prior to mass spectrometry, peptides were desalted using C_{18} ZipTips (Merck Millipore, Darmstadt, Germany) according to the manufacturer's instructions.

The peptides were analyzed by nanoLC-MS/MS as described previously, using a Q-Exactive hybrid quadupole orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany)[46], and the acquired raw data was analyzed using MaxQuant[47] version 1.4.1.2. Proteins were quantified using the MaxLFQ algorithm[48]. The data was searched against a sample-specific database (602.947 protein sequences), generated from the 355 metagenome assembled genomes (MAGs), and against the pig genome (*Sus scrofa domesticus)*. In addition, common contaminants such as human keratins, trypsin and bovine serum albumin were concatenated to the database as well as reversed sequences of all protein entries for estimation of false discovery rates. Protein N-terminal acetylation, oxidation of methionine, conversion of glutamine to pyro glutamic acid, and deamination of asparagine and glutamine were used as variable modifications, while carbamidomethylation of cysteine residues was used as a fixed modification. Trypsin was used as digestion enzyme and two missed cleavages were allowed. All

557 identifications were filtered in order to achieve a protein false discovery rate (FDR) of 1% using the
558 target-decoy strategy. For a protein to be considered valid, we required the protein to be both
559 identified and quantified in both replicates, and in addition, we required at least one unique peptide
560 per protein and at least two peptides in total for every protein. The output from MaxQuant was further
561 explored in Perseus version 1.6.0.7 where filtering, data transformation, and imputation were
562 performed, and visualizations including heatmaps, hierarchical clustering, and volcano plots (for
563 identification of differentially abundant proteins between the mannan and control groups) were
564 made.

565

566 **Genome tree**

567 Phylogenetic analysis was performed using a block of 22 universal ribosomal proteins (30S ribosomal
568 protein L1, L2, L4-L6, L10, L11, L14, L15,L18 and 50S ribosomal protein S3, S5, S7-S13, S15, S17,
569 S19)[49,50]. In addition to the MAGs, we recruited 239 reference genomes for phylogenetic resolution.
570 These genomes were selected based on preliminary examination of the assembled metagenome using
571 metaQUAST[51]. The selection of reference genomes were annotated using the Prokka pipeline,
572 uniformly with annotation of the MAGs. All identified ribosomal protein sequences were aligned
573 separately with MUSCLE v3.8.31[52], and manually checked for duplications and misaligned sequences.
574 Divergent regions and poorly aligned positions were further eliminated using GBlocks[53], and the
575 refined alignment were concatenated using catfasta2phyml.pl
576 (https://github.com/nylander/catfasta2phyml) with the parameter '-c' to replace missing ribosomal
577 proteins with gaps (-). The maximum likelihood-based phylogenetic analysis of the concatenated
578 ribosomal proteins was inferred using RAxML version 8.2.12[54] (raxmlHPC-SSE3 under PROTGAMMA
579 distributed model with WAG substitution matrix) and support values determined using 100 bootstrap
580 replicates. The tree was rooted to the Euryarchaeota phylum and visualized using iTOL[55]. Clades of

581 reference genomes with only distant phylogenetic relation to the MAGs were collapsed to refine the
582 final tree in **Fig. 3**. The complete tree is available in Newick format as Supplementary Dataset 1.

583

584 **REFERENCES**

585 1 Gibson, G. R. *et al.* Expert consensus document: The International Scientific Association for
586 Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of
587 prebiotics. *Nat. Rev. Gastroenterol. Hepatol.* **14**, 491-502 (2017).
588 2 Sanders, M. E., Merenstein, D. J., Reid, G., Gibson, G. R. & Rastall, R. A. Probiotics and
589 prebiotics in intestinal health and disease: from biology to the clinic. *Nat. Rev. Gastroenterol.*
590 *Hepatol.* **doi: 10.1038/s41575-019-0173-3** (2019).
591 3 Cuskin, F. *et al.* Human gut Bacteroidetes can utilize yeast mannan through a selfish
592 mechanism. *Nature* **517**, 165-169 (2015).
593 4 La Rosa, S. L. *et al.* The human gut Firmicute *Roseburia intestinalis* is a primary degrader of
594 dietary β-mannans. *Nat. Commun.* **10**, 905 (2019).
595 5 Solden, L. M. *et al.* Interspecies cross-feeding orchestrates carbon degradation in the rumen
596 ecosystem. *Nat. Microbiol.* **3**, 1274-1284 (2018).
597 6 Michalak, L. *et al.* A pair of esterases from a commensal gut bacterium remove acetylations
598 from all positions on complex β-mannans. *BioRxiv* **https://doi.org/10.1101/788067** (2019).
599 7 La Rosa, S. L. *et al.* Wood-Derived Dietary Fibers Promote Beneficial Human Gut Microbiota.
600 *mSphere* **4**, e00554-00518 (2019).
601 8 Ferreira-Halder, C. V., Faria, A. V. S. & Andrade, S. S. Action and function of *Faecalibacterium*
602 *prausnitzii* in health and disease. *Best Pract. Res. Clin. Gastroenterol.* **31**, 643-648 (2017).
603 9 Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J. & Martinez-Medina, M. *Faecalibacterium*
604 *prausnitzii*: from microbiology to diagnostics and prognostics. *ISME J.* **11**, 841-852 (2017).

10 Xu, C. *et al.* Acetylation and characterization of spruce (Picea abies) galactoglucomannans. *Carbohydrate research* **345**, 810-816 (2010).

11 Michalak, L., Knutsen, S. H., Aarum, I. & Westereng, B. Effects of pH on steam explosion extraction of acetylated galactoglucomannan from Norway spruce. *Biotechnol. Biofuels* **11**, 311 (2018).

12 Chen, L. *et al.* The Maturing Development of Gut Microbiota in Commercial Piglets during the Weaning Transition. *Front. Microbiol.* **8**, 1688 (2017).

13 Holman, D. B., Brunelle, B. W., Trachsel, J. & Allen, H. K. Meta-analysis To Define a Core Microbiota in the Swine Gut. *MSystems* **2**, e00004.

14 Gresse, R. *et al.* Gut Microbiota Dysbiosis in Postweaning Piglets: Understanding the Keys to Health. *Trends Microbiol.* **25**, 851-873 (2017).

15 Vigre, H., Larsen, P. B., Andreasen, M., Christensen, J. & Jorsal, S. E. The effect of discontinued use of antimicrobial growth promoters on the risk of therapeutic antibiotic treatment in Danish farrow-to-finish pig farms. *Epidemiology and infection* **136**, 92-107 (2008).

16 de Lange, C. F. M., Pluske, J., Gong, J. & Nyachoti, C. M. Strategic use of feed ingredients and feed additives to stimulate gut health and development in young pigs. *Livestock Science* **134**, 124-134 (2010).

17 Kageyama, A. & Benno, Y. *Catenibacterium mitsuokai* gen. nov., sp. nov., a gram-positive anaerobic bacterium isolated from human faeces. *Int. J Syst. Bacteriol.* **50**, 1595-1599 (2000).

18 Moore, L. V. & Moore, W. E. *Oribaculum catoniae* gen. nov., sp. nov.; *Catonella morbi* gen. nov., sp. nov.; *Hallella seregens* gen. nov., sp. nov.; *Johnsonella ignava* gen. nov., sp. nov.; and *Dialister pneumosintes* gen. nov., comb. nov., nom. rev., Anaerobic gram-negative bacilli from the human gingival crevice. *Int J Syst Bacteriol.* **44**, 187-192 (1994).

19 Bowers, R. M. *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol.* **35**, 725-731 (2017).

631 20 Martens, E. C., Koropatkin, N. M., Smith, T. J. & Gordon, J. I. Complex glycan catabolism by the
632 human gut microbiota: The bacteroidetes Sus-like paradigm. *J. Biol. Chem.* **284**, 24673-24677
633 (2009).
634 21 Leth, M. L. *et al.* Differential bacterial capture and transport preferences facilitate co-growth
635 on dietary xylan in the human gut. *Nat. Micro.* **3**, 570-580 (2018).
636 22 Vizcaino, J. A. *et al.* The PRoteomics IDEntifications (PRIDE) database and associated tools:
637 status in 2013. *Nucleic Acids Res.* **41**, D1063-1069 (2013).
638 23 Chum, H. L., Johnson, D. K., Black, S. K. & Overend, R. P. Pretreatment Catalyst Effects and the
639 Combined Severity Parameter. *Appl. Biochem. Biotech.* **24**, 1-14 (1990).
640 24 Vestby, L. K., Møretrø, T., Ballance, S., Langsrud, S. & Nesse, L. L. Survival potential of wild type
641 cellulose deficient Salmonella from the feed industry. *BMC Veterinary Research* **5**, 43 (2009).
642 25 Pedersen, K. S. & Toft, N. Intra- and inter-observer agreement when using a descriptive
643 classification scale for clinical assessment of faecal consistency in growing pigs. *Prev Vet Med*
644 **98**, 288-291 (2011).
645 26 Takahashi, S., Tomita, J., Nishioka, K., Hisada, T. & Nishijima, M. Development of a prokaryotic
646 universal primer for simultaneous analysis of Bacteria and Archaea using next-generation
647 sequencing. *PloS one* **9**, e105592 (2014).
648 27 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data.
649 *Nature methods* **7**, 335-336 (2010).
650 28 Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-
651 supported software for describing and comparing microbial communities. *Appl Environ*
652 *Microbiol* **75**, 7537-7541 (2009).
653 29 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and
654 graphics of microbiome census data. *PloS one* **8**, e61217 (2013).
655 30 Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End
656 reAd mergeR. *Bioinformatics* **30**, 614-620 (2014).

31 Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863-864 (2011).

32 Rideout, J. R. *et al.* Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ* **2**, e545 (2014).

33 Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).

34 Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic acids research* **42**, D643-648 (2014).

35 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).

36 Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome research* **27**, 824-834 (2017).

37 Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674-1676 (2015).

38 Kang, D. D., Froula, J., Egan, R. & Wang, Z. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**, e1165 (2015).

39 Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *Isme j* **11**, 2864-2868 (2017).

40 Ondov, B. D. *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biology* **17**, 132 (2016).

41 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* **25**, 1043-1055 (2015).

682 42 Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069
683 (2014).
684 43 Terrapon, N., Lombard, V., Gilbert, H. J. & Henrissat, B. Automatic Prediction of Polysaccharide
685 Utilization Loci in Bacteroidetes Species. *Bioinformatics* **31**, 647-655 (2014).
686 44 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The
687 carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490-495
688 (2014).
689 45 Arntzen, M. O., Karlskas, I. L., Skaugen, M., Eijsink, V. G. & G., M. Proteomic investigation of
690 the response of *Enterococcus faecalis* V583 when cultivated in urine. *PLoS One* **10**, e0126694
691 (2015).
692 46 Arntzen, M. O., Karlskas, I. L., Skaugen, M., Eijsink, V. G. & Mathiesen, G. Proteomic
693 Investigation of the Response of Enterococcus faecalis V583 when Cultivated in Urine. *PloS*
694 *one* **10**, e0126694 (2015).
695 47 Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-
696 range mass accuracies and proteome-wide protein quantification. *Nature biotechnology* **26**,
697 1367-1372 (2008).
698 48 Cox, J. *et al.* Accurate proteome-wide label-free quantification by delayed normalization and
699 maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* **13**,
700 2513-2526 (2014).
701 49 Hug, L. A. *et al.* A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
702 50 Wu, D., Jospin, G. & Eisen, J. A. Systematic identification of gene families for use as “markers”
703 for phylogenetic and phylogeny-driven ecological studies of bacteria and archaea and their
704 major subgroups. *PloS One* **8**, e77033 (2013).
705 51 Mikheenko, A., Saveliev, V. & Gurevich, A. MetaQUAST: evaluation of metagenome
706 assemblies,. *Bioinformatics* **32**, 1088-1090 (2016).

52 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* **32**, 1792-1797 (2004).

53 Talavera, G. & Castresana, J. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst. Biol.* **56**, 564–577 (2007).

54 Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).

55 Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research* **44**, W242-W245 (2016).