Novel insights into the pig gut microbiome using metagenome-assembled genomes

Devin B. Holman[a#], Arun Kommadath[a], Jeffrey P. Tingley[bc], D. Wade Abbott[bc]

[a]Lacombe Research and Development Centre, Agriculture and Agri-Food Canada, 6000 C&E Trail, Lacombe, AB, T4L1W1, Canada

[b]Lethbridge Research and Development Centre, Agriculture and Agri-Food Canada, 5403 1st Avenue South, Lethbridge, AB, T1J 4B1, Canada

[c]Department of Biochemistry, University of Lethbridge, Lethbridge, AB T1K 6T5, Canada

#Address correspondence to Devin B. Holman, devin.holman@agr.gc.ca

Running title: Swine gut-derived metagenome-assembled genomes

Keywords: metagenome-assembled genomes, antimicrobial resistance, CAZymes, swine, gut microbiome

Abstract

Pigs are among the most numerous and intensively farmed food-producing animals in the world. The gut microbiome plays an important role in the health and performance of swine and changes rapidly after weaning. Here, fecal samples were collected from pigs at 7 different times points from 7 to 140 days of age. These swine fecal metagenomes were used to assemble 1,150 dereplicated metagenome-assembled genomes (MAGs) that were at least 90% complete and had less than 5% contamination. These MAGs represented 472 archaeal and bacterial species, and the most widely distributed MAGs were the uncultured species *Collinsella* sp002391315, *Sodaliphilus* sp004557565, and *Prevotella* sp000434975. Weaning was associated with a decrease in the relative abundance of 69 MAGs (e.g. *Escherichia coli*) and an increase in the relative abundance of 140 MAGs (e.g. *Clostridium* sp000435835, *Oliverpabstia intestinalis*). Genes encoding for the production of the short-chain fatty acids acetate, butyrate, and propionate were identified in 68.5%, 18.8%, and 8.3% of the MAGs, respectively. Carbohydrate-active enzymes associated with the degradation of arabinose oligosaccharides and mixed-linkage glucans were predicted to be most prevalent among the MAGs. Antimicrobial resistance genes were detected in 327 MAGs, including 59 MAGs with tetracycline resistance genes commonly associated with pigs such as *tet*(44), *tet*(Q), and *tet*(W). Overall, 82% of the MAGs were assigned to species that lack cultured representatives indicating that a large portion of the swine gut microbiome is still poorly characterized. The results here also demonstrate the value of MAGs in adding genomic context to gut microbiomes.

Importance

Many of the bacterial strains found in the mammalian gut are difficult to culture and isolate due to their various growth and nutrient requirements that are frequently unknown. Here, we

assembled strain-level genomes from short metagenomic sequences, so-called metagenome-assembled genomes (MAGs), that were derived from fecal samples collected from pigs at multiple time points. The majority of these MAGs represented bacterial species that have yet to be cultured or described thus underlining the need for cultivation studies that isolate and describe novel bacterial species. The genomic context of a number of antimicrobial resistance genes commonly detected in swine was also determined. In addition, our study connected taxonomy with potential metabolic functions such as carbohydrate degradation and short-chain fatty acid production.

Introduction

Pork production continues to increase globally (1) despite serious challenges, such as antimicrobial use and resistance (2) and infectious disease (3), that threaten its profitability and sustainability. Microbiome research has the potential to contribute solutions to some of these issues, given a better understanding of the swine gut microbiome. The pig gut microbiome, as in most mammals, provides the host with numerous benefits including protection against pathogen colonization, aiding immune system development and maturation, and production of certain vitamins and metabolites (4). The number of unique genes within the gut microbiome also greatly exceeds those encoded within the pig genome thereby providing the host with an expanded repertoire of genes that can degrade dietary substrates (5).

Cultivation of many of the microbes found in the swine gastrointestinal tract remains difficult due to their unique, and often unknown, growth requirements. Consequently this has traditionally limited study of the mammalian gut microbiome to those microbes that can be easily grown and characterized in the laboratory (6). However, this often represents only a small fraction of the microbial diversity in the gut. In recent years, metagenome-assembled genomes (MAGs) recovered from shotgun metagenomic sequences have greatly expanded the number of

71 microbial genomes in reference databases (7-10). Although the quality of these MAGs varies,
72 they enable researchers to connect functional potential with microbial species and strains lacking
73 cultured representatives.

74 Previously, we assessed the effects of varying weaning ages on the development of the
75 swine gut microbiome using shotgun metagenomic sequences generated from fecal samples
76 collected from pigs throughout the swine production cycle (11). Here, we assembled those
77 sequences and binned the assembled contigs into MAGs, retaining only those MAGs that were at
78 least 90% complete and had less than 5% contamination. Our main objective was to characterize
79 the functional potential represented by those MAGs, including carbohydrate-active enzymes
80 (CAZymes) and antimicrobial resistance genes (ARGs) and to associate those functions with
81 individual taxa. In addition, we aimed to determine if MAGs assembled in this study are
82 representative of the pig gut microbiome in general by using metagenomic sequences from other
83 publicly-available swine studies.

84 **Results**

85 **Metagenome-assembled genomes**

86 From 738 Gb of shotgun metagenomic sequencing data, 87,472 MAGs with greater than
87 90% completeness and less than 5% contamination were recovered. After dereplication at 99%
88 ANI (average nucleotide identity), the remaining 1,150 non-redundant MAGs represented
89 potentially unique strains that were assigned to 360 and 472 archaeal and bacterial genera and
90 species, respectively (Supplementary Table S1). The MAGs ranged in size from 0.74 to 6.14 Mb
91 with an average size of 2.28 Mb (SEM ± 0.02). The 358 MAGs that were not assigned to a
92 species in the Genome Database Taxonomy (GTDB) at a 95% ANI threshold may represent
93 potentially novel species, 32 of which also had no genus designation. When 95% ANI was used

for secondary clustering, 758 dereplicated MAGs (putative species) were produced from the original 87,472 MAGs. The vast majority of the MAGs classified using GTDB were bacteria and only 10 MAGs were assigned to the archaea, all of which were identified as methanogens. In total, 19 unique phyla were represented among the MAGs (Fig. 1). The most common species designation was *Collinsella* sp002391315 (22 MAGs), followed by *Sodaliphilus* sp004557565 (19 MAGs), *Prevotella* sp000434975 (17 MAGs), and UBA3388 sp004551865 (13 MAGs). Overall, 938 MAGs were assigned to archaeal or bacterial species that lack cultured representatives.

Functional analysis of MAGs

Functional profiling of the dereplicated MAGs using CAZymes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was carried out using the Distilled and Refined Annotation of Metabolism (DRAM) package. There were 6,656 unique KOs (KEGG Orthology) and 155,297 CAZymes within 281 unique CAZyme families identified among the MAGs. The average number of CAZymes per MAG was 135.0 ± 2.9 within 33.3 ± 0.5 CAZyme families (Supplementary Table S2). The glycoside hydrolases (GHs; n = 122) were most prevalent among the unique CAZyme families followed by carbohydrate-binding modules (CBMs; n = 58), glycosyltransferases (GTs; n = 55), polysaccharide lyases (PLs; n = 24), carbohydrate esterases (CEs; n = 14), and auxiliary activities (AAs; n = 8). The GH families GH2, GH3, GH13, GH23, GH25, GH31, GH32, GH36, GH73, and GH77 were most widely distributed. Many of the MAGs encoding the greatest number of CAZymes and CAZyme families belonged to the *Bacteroidaceae* family including *Bacteroides fragilis*, *Bacteroides heparinolyticus*, *Bacteroides stercoris, Bacteroides thetaiotaomicron*, *Bacteroides uniformis*, *Bacteroides xylanisolvens*, and *Phocaeicola vulgatus*. CAZymes involved in the degradation of arabinose oligosaccharides and

mixed-linkage glucans were found in at least 50% of the MAGs and CAZymes predicted to be involved in the digestion of amorphous cellulose, arabinan, beta-mannan, fucose oligosaccharides, pectin, rhamnose oligosaccharides, starch, xylan, and xyloglucan were encoded by 25% or more of the MAGs (Supplementary Table S3). Mucin-degrading CAZymes were identified in only 26 MAGs, with four of these classified as *Pauljensenia hyovaginalis* and another three as *Tractidigestivibacter* sp004557505.

The production of SCFAs from carbohydrates is an important function of the gut microbiome from the perspective of the host. The most significant of these are acetate, butyrate, and propionate. Here, 68.5% of the MAGs encoded acetate-producing enzymes (acetate kinase [K00925] and phosphate acetyltransferase enzymes [K00625 or K13788]) and 8.3% had the propionate CoA-transferase gene (K01026) (Supplementary Table S3). Genes for butyrate production via the butyryl-CoA:acetate CoA-transferase (K01034, K01035) or butyrate kinase (K00634, K00929) pathways were identified in 18.8% of MAGs and included known butyrate producers, such as *Anaerostipes hadrus*, *Butyricimonas virosa*, *Butyrivibrio crossotus*, *Cloacibacillus porcorum*, *Coprococcus catus*, *Gemmiger formicilis*, *Faecalibacterium prausnitzii*, *Flavonifractor plautii*, and *Megasphaera elsdenii*.

Succinate is a propionate precursor that is produced by certain bacteria species. Genes encoding a fumarate reductase/succinate hydrogenase (K00239, K00240, K00241 and K00244, K00245, K00246) were identified in 28.9% of the MAGs. Known succinate producers among the MAGs with these genes included *Akkermansia muciniphila*, *Anaerobiospirillum succiniciproducens*, *Mitsuokella jalaludinii*, *Parabacteroides distasonis*, and *P. vulgatus*. The potential for either D-lactate or L-lactate production via lactate dehydrogenase (K00016, K03778) was detected in 55.3% of the MAGs. Genes for the production of both enantiomers of lactate

were carried by 172 MAGs and over half (n = 92) of these were members of the *Treponema* genus or *Lachnospiraceae* or *Lactobacillaceae* families.

The 10 archaeal MAGs all carried the genes encoding the methyl-coenzyme M reductase complex (K00399, K00400, K00401, K00402, K03421, K03422) involved in methanogenesis (Supplementary Table S3). However, only the *Methanobacteriaceae* MAGs had the genes for the formylmethanofuran dehydrogenase complex (K00200, K00201, K00202, K00203, K00204, K00205, K11260, K11261) that is necessary for the reduction of carbon dioxide to methane. Hydrogen sulfide production in swine manure slurry has been linked to *Desulfovibrio* spp. (12) and here genes encoding the dissimilatory sulfite reductase and involved in the metabolism of sulfate were only identified in the eight *Desulfovibrionaceae* MAGs. This included *Desulfovibrio piger* and *Desulfovibrio* sp900556755 as well as one MAG classified as *Bilophila wadsworthia*. These *Desulfovibrionaceae* MAGs also carried the gene for thiosulfate reductase (K08352) which produces sulfide and sulfite through the reduction of thiosulfate.

Antimicrobial resistance genes

The 1,150 dereplicated MAGs were screened for ARGs using the comprehensive antibiotic resistance database (CARD). A total of 327 MAGs carried at least one ARG (Supplementary Table S4), and together they accounted for 115 unique ARGs, excluding those due to point mutations. The six *Escherichia coli* MAGs contained the greatest number of ARGs (52 to 60) by a large margin. However, this is expected given that the vast majority of these ARGs are widespread within this species. ARGs conferring resistance to tetracycline (*tet* genes) are frequently among the most abundant in the gastrointestinal tract of conventionally-raised pigs and here, 59 MAGs carried at least one *tet* gene. Among the MAGs with at least one *tet* gene and an overall relative abundance of at least 0.1%, were those identified as *B. fragilis* [*tet*(Q)], *B.*

163 *stercoris* [*tet*(Q)], CAG-873 sp001701165 [*tet*(Q)], *Campylobacter coli* [2 MAGs; *tet*(W/N/W),
164 *tet*(O)], *Limosilactobacillus reuteri* [*tet*(B)], *P. vulgatus* [*tet*(Q)], *Prevotella* sp000434975
165 [*tet*(Q)], *Prevotella* sp000436915 [*tet* (37)], and *Streptococcus pasteurianus* [*tet*(M)].

166 Resistance to macrolide-lincosamide-streptogramin B (MLS_B) antimicrobials is also often
167 detected in the swine gut microbiome and 48 MAGs were carried at least one MLS_B resistance
168 gene. Relatively abundant (≥ 0.1%) MAGs carrying one or more MLS_B resistance genes included
169 *B. fragilis* [*mef*(En2)], *Catenibacterium mitsuokai* [*erm*(G)], *Clostridium* sp000435835 [*erm*(Q)],
170 *Fusobacterium mortiferum* [*lnu*(C)], *Lactobacillus johnsonii* [2 MAGs; *erm*(B), *erm*(G)], *L.*
171 *reuteri [erm*(B)*]*, *Parabacteroides merdae* [*mef*(En2)], *P. vulgatus* [*mef*(En2)], *Treponema*
172 *succinifaciens* [*erm*(F)], *Schaedlerella* sp004556565 [*lnu(C)*], SFDP01 sp004558185 [*erm*(B)],
173 *and S. pasteurianus* [*lnu*(A), *lnu*(C), *erm*(B)].

174 The *vanC* cluster genes, (*vanC*, $vanR_C$, $vanS_C$, $vanT_C$, $vanXY_C$), which confer resistance
175 to vancomycin were found in one MAG classified as *Enterococcus gallinarum*. In this species
176 low-level vancomycin resistance is intrinsic due to this gene cluster (13). The beta-lactamase
177 resistance gene *cfxA2* was identified in 13 MAGs, including six given the taxonomic designation
178 *Sodaliphilus* sp004557565. Many of the other beta-lactamase genes detected were associated
179 with only one bacterial species: bla_{OXA-61} (three *C. coli* MAGs), bla_{TEM-1} (one *E. coli* MAG),
180 *cblA-1* (three *Bacteroides uniformis* MAGs), and *cepA* (one *B. fragilis* MAG). Aminoglycoside
181 resistance genes among the relatively abundant MAGs (> 0.1%) included *aph(3')-IIIa* in three
182 MAGs (*C. mitsuokai*, *F. mortiferum*, SFDP01 sp004558185), *aac(6')-Im* (*Blautia* spp.), *aad(6)*
183 (SFDP01 sp004558185), *aadA* (*E. coli*), *ant(6)-Ib* (*L. johnsonii*), *aph*(2")-IIa (*Blautia* spp.), and
184 *aph(6)-Id* (*E. coli*).

The location of ARGs within the MAGs was also determined to identify those ARGs co-located on the same contig as other ARGs and/or integrase/transposase sequences (Supplementary Fig. S2). The *aac(6')-Im* and *aph(2'')-IIa* genes were adjacent to each other in three MAGs classified as *Blautia* sp018919065, *Ruminococcus gnavus*, and CAG-238 sp. In one *T. succinifaciens* MAG, *erm*(F) and *tet*(X) were also found on the same contig as were *tet*(M) and *tet*(W/N/W) and *tet*(44) and *ant(6)-Ib* in a CAG-877 sp. MAG. In several MAGs, *lnu*(C) was co-located on the same contig as a putative transposase or integrase gene. Other ARGs potentially associated with transposases included *tet*(44) in two MAGs assigned to *Onthovivens* sp016302065 and CAG-1000 sp004552445, *tet*(M) in *Erysipelatoclostridium ramosum* and *S. pasteurianus*, and *tet*(Q) in *Onthomorpha* sp004551865 and *Prevotella* sp900548195.

Pre- vs. post-weaning changes

As these MAGs were assembled from fecal samples taken before and after weaning it was possible to identify MAGs that were differentially abundant in the fecal microbiome of pigs immediately before weaning and 7 days post-weaning. There were 69 MAGs that were more relatively abundant in samples take just prior to weaning, the most differentially abundant of which were those classified as *Limousia pullorum*, *B. fragilis*, *E. coli*, *P. hyovaginalis* (*Schaalia hyovaginalis* in NCBI), and *P. vulgatus* (Supplementary Table S5). There were also six MAGs with a relative abundance greater than 0.1% in the fecal microbiomes of nursing piglets that were not detected in samples from these same pigs 7 d later. These MAGs were classified as *B. thetaiotaomicron*, *Bulleidia* sp., *Enterococcus faecalis*, *Mediterraneibacter torques*, *Parvimonas* sp., and *P. hyovaginalis*. Among the 140 MAGs that were most enriched in the post-weaning samples were MAGs assigned to *Copromorpha* sp., *Clostridium* sp000435835, *Fusicatenibacter saccharivorans*, *Intestinibacter* sp., *Oliverpabstia intestinalis*, *Phascolarctobacterium*

208 sp004558595, *Prevotella* sp002251295, *Prevotella* sp004556065, and *Ruminococcus*
209 sp003011855.

210 **Presence of MAGs in publicly available datasets**

211 To determine how widely distributed the species/strains represented by the MAGs in the
212 present study are among pigs from other studies in different geolocations, the presence and
213 relative abundance of these MAGs within publicly available swine gut metagenomic datasets
214 was assessed. These metagenomic sequences were from 626 fecal and cecal content samples
215 within nine studies representing 13 different counties (Supplementary Table S6). On average,
216 45.5% ± 0.4% SEM of these metagenomic sequences mapped to one of the MAGs from the
217 present study (Supplementary Table S7). Two MAGs classified as *Lactobacillus amylovorus*
218 were the most relatively abundant overall. Other relatively abundant MAGs (>0.25%) included
219 those identified as *B. fragilis*, *C. mitsuokai*, *L. reuteri*, *Phascolarctobacterium succinatutens*,
220 *Prevotella pectinovora*, *Prevotella* sp002251295, *Prevotella* sp002300055, *Streptococcus*
221 *alactolyticus*, and VUNA01 sp002299625. Metagenomic sequences from 96 MAGs were
222 detected in 90% of these publicly available samples. Thirty-three of these MAGs were classified
223 within the *Oscillospiraceae* family including 12 as co-abundance gene groups (CAGs), and 8
224 each as *Dysosmobacter* spp. and *Faecousia* spp. An additional 19 MAGs were assigned to
225 *Sodaliphilus* sp004557565 and 8 as *Cryptobacteroides* spp.

226 The samples from these studies were all collected from post-weaned pigs and therefore
227 on a non-metric multidimensional scaling (NMDS) plot of the Bray-Curtis dissimilarities the pre-
228 weaned pig samples from the present study appear separate from the other samples
229 (Supplementary Fig. S1). Only eight MAGs were not detected in at least one sample among all
230 of the publicly available metagenomic samples and three of these MAGs (*Clostridium* sp.,

231 *Erysipelotrichaceae* sp., and *Negativicoccaceae* sp.,) were not identified in any of the post-
232 weaned samples in the present study either. Overall, including the samples from the current
233 study, there were 71 MAGs that were found in 85% of all samples (Fig. 2).

234 **Discussion**

235 Here we assembled and analyzed 1,150 high-quality MAGs, including 358 that could not
236 be assigned to a species and thus some of these may represent novel species. Clearly, there still
237 exists a large fraction of the swine gut microbiome that has yet to be cultivated as demonstrated
238 by the 938 MAGs that were not associated (>95% ANI) with a phenotypically characterized
239 archaeal or bacterial species. Previously, we reported on changes in the pig gut microbiome in
240 response to different weaning ages using the unassembled short reads from this study (11). We
241 were able to recover MAGs from many of the relatively abundant species in our earlier work
242 including *Anaeromassilibacillus senegalensis* (An172 in GTDB), *B. fragilis*, *E. coli*, *L.*
243 *johnsonii, L. reuteri*, *P. succinatutens*, *P. pectinovora*, and *Subdoligranulum variabile*
244 (*Gemmiger variabilis* in the GTDB). The exceptions were *Prevotella copri* and *Clostridioides*
245 *difficile*; however, this may have been due to differences in how taxonomy was assigned to the
246 unassembled reads vs. the MAGs as several closely related species were identified here.

247 After weaning, pigs are typically fed a diet that is rich in cereal grains such as corn,
248 barley, and/or wheat which, in addition to high levels of starch, contain other polysaccharides
249 such as cellulose, hemicellulose (e.g., mannan, mixed-linkage glucan, xylan, and xyloglucan),
250 and pectin. These polysaccharides escape digestion by the host and are therefore available as
251 substrates for the gut microbiome (14). As such, the pig gut microbiome carries a large repertoire
252 of genes encoding enzymes called CAZymes that can breakdown and metabolize these
253 polysaccharides. The CAZymes are grouped into families based on sequence similarity, although

CAZymes within the same family may have different substrate specificities (15). The AAs, CEs, GHs, and PLs are the CAZyme families involved in the degradation of glycans and typically multiple CAZymes are required for the digestion of specific glycans.

Many of the MAGs encoding the greatest number of CAZymes and with the potential capacity to degrade multiple types of glycans were classified within the Bacteroidales order including *Alistipes senegalensis*, *Alistipes shahii*, *B. thetaiotaomicron*, *B. uniformis*, *B. xylanisolvens*, *Parabacteroides* spp., *P. vulgatus*, and *Prevotella* spp. Bacteria within this order are well documented as having a diverse and rich set of CAZymes which may be organized into groups of genes termed polysaccharide utilization loci (PUL) (16, 17). These CAZymes and PULs likely confer an advantage to these bacteria within a highly competitive ecosystem like the mammalian gastrointestinal tract. The other taxonomic group of MAGs encoding a large number of CAZymes was the *Lachnospiraceae* family that included *Acetatifactor* sp., *Blautia* sp001304935, COE1 sp., *Eisenbergiella massiliensis*, *Hungatella* sp005845265, and *Roseburia* sp. Some members of this family have also been reported to have gene clusters of CAZymes, regulators, and transporters that are similar to PULs (18).

The potential for the degradation of arabinan, amorphous cellulose, arabinose oligosaccharides, mixed-linkage glucans, xyloglucan, and xylan was relatively widespread among the MAGs. In pigs, diets supplemented with xylan, mixed-linkage glucans, and resistant starch have been shown to increase the relative abundance of *Blautia* spp., *Prevotella* spp. and *Lachnospiraceae* spp. (19-21). Monosaccharides produced through the action of CAZymes can then be used by the CAZyme-producer or other bacteria in the gut to generate various metabolites. In particular, the potential for short-chain fatty acid (SCFA) production through the fermentation of monosaccharides is frequently a focus of many mammalian gut microbiome

277 studies as even in monogastric animals like pigs, up to 25% of daily energy requirements are met
278 by SCFAs (22). Butyrate is often the SCFA of most interest as it is the primary energy source for
279 mammalian colonic epithelial cells and can regulate apoptosis, enhance barrier function, and
280 reduce inflammation in these cells (23, 24).

281 Here, 216 MAGs carried genes for butyrate production through either the butyryl-
282 CoA:acetate CoA-transferase (*but*) or butyrate kinase (*buk*) pathways. Although several known
283 butyrate producers are included among these MAGs such as *B. virosa*, *F. prausnitzii* and *M.*
284 *elsdenii*, certain MAGs were assigned to bacterial species (e.g., *E. coli*, *E. faecalis*) that do not
285 typically produce butyrate. Instead, these genes are likely involved in other metabolic functions
286 or two component systems in these species. Typically, the *but* gene is more prevalent than the
287 *buk* gene among gut bacteria (25); however, here the number of MAGs carrying either of these
288 genes was nearly the same. There were also 17 MAGs with both *but* and *buk* genes, including *C.*
289 *porcorum*, *F. plautii*, and *Intestinimonas massiliensis*.

290 Acetate and propionate, the two other physiologically important SCFAs in the
291 mammalian gut, also have anti-inflammatory effects in addition to providing an energy source
292 for the host (26). In the swine lower gastrointestinal tract the concentration ratio of
293 acetate:propionate:butyrate is approximately 65:25:10 (27, 28) and here the number of MAGs (n
294 = 788) encoding the acetate kinase and phosphate acetyltransferase genes involved in acetate
295 production outnumbered those carrying genes for producing butyrate (n = 216) and propionate (n
296 = 95). Bacterial species represented by these MAGs are therefore attractive targets for
297 microbiome manipulation studies and culturing work to obtain isolates of these species for
298 further characterization.

There were also significant shifts in the relative abundance of a large number of MAGs 7 days post-weaning. As discussed, the diet of the pigs is abruptly changed at weaning from one that is liquid and milk-based to one that is solid and based on cereal grains. This often results in a decrease in the relative abundance of *Bacteroides* and *Escherichia* spp. and an increase in the relative abundance of *Blautia*, *Prevotella*, and *Roseburia* spp. (29-31). Many of the differentially abundant MAGs pre- and post-weaning were assigned to these genera; however, there were also several MAGS classified as bacterial species or genera that are not known to be associated with weaning. These included MAGs enriched in post-weaning pigs that were assigned to uncultured genera or species and were also identified as potential butyrate producers such as CAG-83 sp., *Aphodosoma* sp900769035, *Copromorpha* sp., *Egerieousia* sp004561775, and UMGS1668 sp004556975. Some of these placeholder names represent bacterial taxa that have been previously reported in swine gut metagenomes and await further characterization (32, 33). One MAG classified as *E. faecalis* was relatively abundant in the nursing pig samples (0.25 ± 0.07%) but was not detected in any of the post-weaning fecal samples. *E. faecalis* was previously identified among the unassembled reads post-weaning so this MAG may represent a strain of *E. faecalis* that is unique to nursing piglets.

Binning ARGs into MAGs generated from short reads is extremely challenging as they are often flanked by repeat sequences and located on mobile genetic elements such as plasmids which have different properties (e.g., G+C content) than the chromosomal DNA of their host (34). Therefore, one can assume that ARGs identified in the MAGs here are located on the bacterial chromosome. This also explains why the number (115 ARGs) and diversity of ARGs detected in the present study was much lower than in a previous study (250 ARGs) using the same short reads that were used to assemble the MAGs here (11) as well as in the metagenome

co-assembly (897 ARGs; data not shown). Despite the limitations associated with ARG binning we were able to provide genomic context for 115 ARGs including several that are relatively abundant in the swine gut such as *erm*(B), *tet*(44), *tet*(Q), and *tet*(W) (35-38).

A number of the *tet* (tetracyclines) and *erm* (MLS_B) genes were linked to bacterial species or genera that are considered to be commensal members of the pig gut microbiome such as *Bacteroides* spp., *Clostridium* spp., *L. johnsonii*, *L. reuteri*, *Prevotella* spp., *Roseburia* spp., *Ruminococcus bromii*, and *Succinivibrio* spp. (39, 40). This may explain the extensive background level of resistance to tetracyclines and MLS_B antimicrobials in swine gut bacteria even in the absence of exposure to these antimicrobials, as observed here and reported in many previous studies (37, 41-43). Until relatively recently in North America, antimicrobials were often administered to all pigs in a herd for non-therapeutic purposes, namely for growth promotion (44). The gut microbiome is vertically transferred from sow to piglet and so it highly plausible that this microbiome would have been exposed to antimicrobials at some point in the past even if the pigs used in this study were not.

Several MAGs also carried ARGs conferring resistance to two or more antimicrobials. Most notable among these were a *C. coli* MAG encoding bla_{OXA-61}, *tet*(O), and *tet*(W/N/W) and a S. *pasteurianus* MAG with *erm*(B), *lnu*(A), *lnu*(C), and *tet*(M). Both of these MAGs were also enriched in fecal samples of pre-weaned piglets. *C. coli* can be a cause of foodborne illness in humans (45) and carried by healthy pigs while *S. pasteurianus* is an opportunistic pathogen in humans and has been associated with meningitis in piglets (46). In addition, certain MAGs contained more than one ARG on the same contig, suggesting that the ARGs are linked. ARGs linked together in this manner are more likely to be co-selected and maintained within the bacterium. The aminoglycoside resistance genes *aac(6')-Im* and *aph(2'')-Iia* (also known as

345 *aph(2'')-Ib*) were adjacent to each other in three MAGs within the Clostridia class. These ARGs
346 have previously been reported together in *Enterococcus faecium* and *E. coli* strains (47). A
347 contig with *erm*(F) and *tet*(X) was also binned into a MAG classified as *T. succinifaciens.* These
348 two ARGs confer resistance to macrolides and tetracyclines, respectively, and were originally
349 described on a transposon in *B. fragilis*, although the *tet*(X) gene was reported to be inactive in
350 this species and under anaerobic conditions (48). The *tet*(44) and *ant(6)-Ib* ARGs found here
351 together on the same contig in a *Clostridium* sp. MAG have also been co-located on a transposon
352 in *C. difficile* (49) and a pathogenicity island in *Campylobacter fetus* (50).

353 There were also a number of ARGs co-located with putative transposase or integrase
354 genes. Transposases and integrases are enzymes that can transfer DNA segments, including
355 ARGs, within and between bacterial genomes (51). Here, the lincosamide resistance genes
356 *lnu*(C) and *lnu*(P) were co-located with putative transposase genes in eight different MAGs. Both
357 *lnu*(C) and *lnu*(P) have been previously identified in *Streptococcus agalactiae* (52) and
358 *Clostridium perfringens* (53), respectively, where they were located on the same genomic region
359 as transposase genes. The *tet*(44), *tet*(M), *tet*(Q), and *tet*(W/N/W) genes were also detected on
360 the same contig as putative transposase genes in certain MAGs. If these ARGs are indeed able to
361 move between bacterial genomes it may also explain their ubiquity in swine gut metagenomes. It
362 is possible that some of the contigs with ARGs may have been binned incorrectly given the
363 difficulties in assembly and binning of ARGs discussed above. However, many of the ARGs
364 were found in MAGs that were closely related to the known species range for the ARG. The use
365 of long-read sequencing would likely increase the number of ARGs binned as well as improve
366 the resolution of their genomic context.

We also evaluated the presence of the 1,150 MAGs from the present study within 626 swine gut metagenomes that were publicly available. Sequences aligning to 96 MAGs were identified in 90% or greater of all these samples and included 8 MAGs that were classified as *Dysosmobacter* spp. and 19 as *Sodaliphilus* sp004557565. *Dysosmobacter* is a new genus most closely related to *Oscillibacter* (54), thus explaining the absence of previous reports of this genus in the swine gut microbiome. The type species of this genus, *Dysosmobacter welbionis*, has recently been shown to partially protect against some of the negative effects of a high-fat diet when administered to mice (55). Similar to *Dysosmobacter*, *Sodaliphilus* is a newly described genus whose type species, *Sodaliphilus pleomorphus*, was first isolated from pig feces. Swine-derived MAGs classified as *Sodaliphilus* sp004557565 have also been recently reported (33). These results suggest that members of these genera are widespread among pigs and may represent previously unreported bacterial taxa.

Conclusions

We recovered 1,150 high-quality MAGs from fecal metagenomes of pre- and post-weaned pigs. The MAGs described here demonstrate the vast potential of the pig gut microbiome to degrade and metabolize various glycans and of certain members to provide beneficial SCFAs to the host. In addition, the significant number of ARGs found associated with MAGs assigned to bacterial species that are typically commensals in the gut, may explain why resistance to macrolides and tetracyclines persists in the absence of antimicrobial selective pressure. The large majority of the MAGs were assigned to poorly characterized taxa and thus, there still exists a large fraction of the swine gut microbiome that has yet to be cultured. This included many bacterial species that appear to be widely disseminated among pigs from different geolocations.

389 Future efforts focused on expanding the number of known bacterial species would greatly
390 improve on efforts to manipulate the gut microbiome to improve production and health.

391 **Materials and Methods**

392 **Experimental design**

393 The study design and fecal sampling were previously described in Holman et al. (11). Briefly,
394 piglets (n = 15) were assigned to be weaned at one of three ages; 14, 21, or 28 days of age. Fecal
395 swabs were collected from the piglets at d 7, 14, 21, 28, 35, 70, and 140 days of age (n = 179).
396 DNA was extracted using the QIAamp BiOstic Bacteremia DNA Kit (Qiagen, Mississauga, ON,
397 Canada) and shotgun metagenomic sequencing carried out on an Illumina NovaSeq 6000
398 instrument (Illumina Inc., San Diego, CA, USA) with a SP flowcell (2 x 250 bp) as per Holman
399 et al. 2021 (11).

400 **Bioinformatics**

401 Metagenomic sequences were trimmed (quality score < 15 over a sliding window of 4 bp;
402 minimum length of 50 bp) and sequencing adapters removed using Trimmomatic v. 0.38 (56).
403 Host sequences were removed by alignment to the *Sus scrofa* genome (Sscrofa11.1) (57) using
404 Bowtie2 v. 2.4.2-1 (58). MEGAHIT v. 1.29.0 (59) was used to co-assemble and individually
405 assemble metagenomes. Prior to co-assembly, all metagenomic samples were normalized using
406 BBNorm in BBTools v. 38.79 (https://sourceforge.net/projects/bbmap/). For the co-assembled
407 metagenome, the metagenomic sequences from each sample were mapped to the co-assembly
408 using Bowtie2 and for individual assemblies each sample was aligned to its own metagenomic
409 assembly. These contigs in each sample with a minimum length of 2,000 bp were then binned
410 using MetaBAT 2 (60) . These bins or MAGs were assessed for quality and completeness using
411 CheckM v. 1.1.2 (61) and those MAGs that were > 90% complete and had < 5% contamination

were retained. This resulted in 2,327 MAGs from the individually assembled metagenomes and 85,145 MAGs from the co-assembled metagenomes. These MAGs were then dereplicated using dRep v. 3.2.2 (62) with primary clustering at 90% and secondary clustering at 99% ANI. These 1,150 MAGs were then used for all subsequent analyses.

Taxonomy was assigned to each MAG using GTDB-tk 2.0.0 (63) and the GTDB release 207. CoverM v. 0.6.1 (https://github.com/wwood/CoverM) (parameters: --min-read-aligned-percent 75% --min-read-percent-identity 95% --min-covered-fraction 0) was used to determine the relative abundance (coverage) of each MAG within in each metagenomic sample. A phylogenetic tree of the MAGs was constructed from the alignment of 399 marker genes in PhyloPhlAn v. 3.0.60 (64) (parameters: min_num_markers=100; f = supermatrix_aa.cfg) and visualized using iTol v6. (65). DRAM v. 1.2.4 (66) together with the KEGG (release 100, October 1, 2021) and dbCAN2 (67) databases was used to annotate the MAGs. The MAGs were also screened for ARGs using the CARD-RGI v. 5.2.0 (68). Proksee v. 1.0.0a1 (https://proksee.ca) was used to visualize the location of the ARGs within each MAG as well as potential integrases and transposases as annotated by Prokka v. 1.14.6 (69) . MaAsLin2 v. 1.8.0 (70) was used to identify MAGs that were differentially abundant immediately before weaning and 7 days post-weaning. Only those MAGs with a relative abundance greater than 0.05% in these samples were included in this analysis.

Publicly available metagenomic sequences from other swine gut microbiome studies published since 2016 were downloaded and aligned to the MAGs in the present study with CoverM to assess their presence in pigs from other studies in different geographic locations. The unassembled reads as well as the MAGs from the present study are available under BioProject PRJNA629856.

435 **Acknowledgements**

436 This research was supported by funding from Alberta Agriculture and Forestry grant 2018R009R

437 and Agriculture and Agri-Food Canada.

438 **References**

439 1. FAO. 2021. Food outlook: biannual report on global food markets, November 2021.

440 http://www.fao.org/3/cb7491en/cb7491en.pdf.

441 2. Van Boeckel TP, Pires J, Silvester R, Zhao C, Song J, Criscuolo NG, Gilbert M,

442 Bonhoeffer S, Laxminarayan R. 2019. Global trends in antimicrobial resistance in

443 animals in low- and middle-income countries. Science 365:eaaw1944.

444 https://doi.org/10.1126/science.aaw1944.

445 3. VanderWaal K, Deen J. 2018. Global trends in infectious diseases of swine. Proc Natl

446 Acad Sci U S A 115:11495-11500. https://doi.org/10.1073/pnas.1806068115.

447 4. Patil Y, Gooneratne R, Ju XH. 2020. Interactions between host and gut microbiota in

448 domestic pigs: a review. Gut Microbes 11:310-334.

449 https://doi.org/10.1080/19490976.2019.1690363.

450 5. Chen C, Zhou Y, Fu H, Xiong X, Fang S, Jiang H, Wu J, Yang H, Gao J, Huang L. 2021.

451 Expanded catalog of microbial genes and metagenome-assembled genomes from the pig

452 gut microbiome. Nat Commun 12:1106. https://doi.org/10.1038/s41467-021-21295-0.

453 6. Renwick S, Ganobis CM, Elder RA, Gianetto-Hill C, Higgins G, Robinson AV,

454 Vancuren SJ, Wilde J, Allen-Vercoe E. 2021. Culturing Human Gut Microbiomes in the

455 Laboratory. Annu Rev Microbiol 75:49-69. https://doi.org/10.1146/annurev-micro-

456 031021-084116.

457 7. Stewart RD, Auffret MD, Warr A, Walker AW, Roehe R, Watson M. 2019. Compendium
458 of 4,941 rumen metagenome-assembled genomes for rumen microbiome biology and
459 enzyme discovery. Nat Biotechnol 37:953-961. https://doi.org/10.1038/s41587-019-
460 0202-3.
461 8. Xie F, Jin W, Si H, Yuan Y, Tao Y, Liu J, Wang X, Yang C, Li Q, Yan X, Lin L, Jiang
462 Q, Zhang L, Guo C, Greening C, Heller R, Guan LL, Pope PB, Tan Z, Zhu W, Wang M,
463 Qiu Q, Li Z, Mao S. 2021. An integrated gene catalog and over 10,000 metagenome-
464 assembled genomes from the gastrointestinal microbiome of ruminants. Microbiome
465 9:137. https://doi.org/10.1186/s40168-021-01078-x.
466 9. Glendinning L, Stewart RD, Pallen MJ, Watson KA, Watson M. 2020. Assembly of
467 hundreds of novel bacterial genomes from the chicken caecum. Genome Biol 21:34.
468 https://doi.org/10.1186/s13059-020-1947-1.
469 10. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, Hugenholtz
470 P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes
471 substantially expands the tree of life. Nat Microbiol 2:1533-1542.
472 https://doi.org/10.1038/s41564-017-0012-7.
473 11. Holman DB, Gzyl KE, Mou KT, Allen HK. 2021. Weaning Age and Its Effect on the
474 Development of the Swine Gut Microbiome and Resistome. mSystems 6:e0068221.
475 https://doi.org/10.1128/mSystems.00682-21.
476 12. Karnachuk OV, Rusanov II, Panova IA, Grigoriev MA, Zyusman VS, Latygolets EA,
477 Kadyrbaev MK, Gruzdev EV, Beletsky AV, Mardanov AV. 2021. Microbial sulfate
478 reduction by Desulfovibrio is an important source of hydrogen sulfide from a large swine
479 finishing facility. Sci Rep 11:1-11. https://doi.org/10.1038/s41598-021-90256-w.

480 13. Arias CA, Courvalin P, Reynolds PE. 2000. *vanC* cluster of vancomycin-resistant
481 *Enterococcus gallinarum* BM4174. Antimicrob Agents Chemother 44:1660-6.
482 https://doi.org/10.1128/AAC.44.6.1660-1666.2000.
483 14. Navarro D, Abelilla JJ, Stein HH. 2019. Structures and characteristics of carbohydrates in
484 diets fed to pigs: a review. J Anim Sci Biotechnol 10:39. https://doi.org/10.1186/s40104-
485 019-0345-6.
486 15. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The
487 carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490-5.
488 https://doi.org/10.1093/nar/gkt1178.
489 16. McKee LS, La Rosa SL, Westereng B, Eijsink VG, Pope PB, Larsbrink J. 2021.
490 Polysaccharide degradation by the Bacteroidetes: mechanisms and nomenclature.
491 Environ Microbiol Rep 13:559-581. https://doi.org/10.1111/1758-2229.12980.
492 17. Grondin JM, Tamura K, Dejean G, Abbott DW, Brumer H. 2017. Polysaccharide
493 Utilization Loci: Fueling Microbial Communities. J Bacteriol 199:e00860-16.
494 https://doi.org/10.1128/JB.00860-16.
495 18. Li H, Ma L, Li Z, Yin J, Tan B, Chen J, Jiang Q, Ma X. 2021. Evolution of the Gut
496 Microbiota and Its Fermentation Characteristics of Ningxiang Pigs at the Young Stage.
497 Animals (Basel) 11. https://doi.org/10.3390/ani11030638.
498 19. Umu OC, Frank JA, Fangel JU, Oostindjer M, da Silva CS, Bolhuis EJ, Bosch G, Willats
499 WG, Pope PB, Diep DB. 2015. Resistant starch diet induces change in the swine
500 microbiome and a predominance of beneficial bacterial populations. Microbiome 3:16.
501 https://doi.org/10.1186/s40168-015-0078-5.

20. Wang Z, Bai Y, Pi Y, Gerrits WJJ, de Vries S, Shang L, Tao S, Zhang S, Han D, Zhu Z, Wang J. 2021. Xylan alleviates dietary fiber deprivation-induced dysbiosis by selectively promoting *Bifidobacterium pseudocatenulatum* in pigs. Microbiome 9:227. https://doi.org/10.1186/s40168-021-01175-x.

21. Gorham JB, Kang S, Williams BA, Grant LJ, McSweeney CS, Gidley MJ, Mikkelsen D. 2017. Addition of arabinoxylan and mixed linkage glucans in porcine diets affects the large intestinal bacterial populations. Eur J Nutr 56:2193-2206. https://doi.org/10.1007/s00394-016-1263-4.

22. Iyayi EA, Adeola O. 2015. Quantification of short-chain fatty acids and energy production from hindgut fermentation in cannulated pigs fed graded levels of wheat bran. J Anim Sci 93:4781-7. https://doi.org/10.2527/jas.2015-9081.

23. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, Harmsen HJ, Faber KN, Hermoso MA. 2019. Short chain fatty acids (SCFAs)-mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. Front Immunol doi:10.3389/fimmu.2019.00277:277. https://doi.org/10.3389/fimmu.2019.00277.

24. Bach Knudsen KE, Laerke HN, Hedemann MS, Nielsen TS, Ingerslev AK, Gundelund Nielsen DS, Theil PK, Purup S, Hald S, Schioldan AG, Marco ML, Gregersen S, Hermansen K. 2018. Impact of Diet-Modulated Butyrate Production on Intestinal Barrier Function and Inflammation. Nutrients 10:1499. https://doi.org/10.3390/nu10101499.

25. Vital M, Howe AC, Tiedje JM. 2014. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. mBio 5:e00889. https://doi.org/10.1128/mBio.00889-14.

525 26. Riviere A, Selak M, Lantin D, Leroy F, De Vuyst L. 2016. Bifidobacteria and Butyrate-
526 Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human
527 Gut. Front Microbiol 7:979. https://doi.org/10.3389/fmicb.2016.00979.
528 27. Den Besten G, Van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM. 2013.
529 The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host
530 energy metabolism. J Lipid Res 54:2325-2340. https://doi.org/10.1194/jlr.R036012.
531 28. Li H, Ma L, Li Z, Yin J, Tan B, Chen J, Jiang Q, Ma X. 2021. Evolution of the Gut
532 Microbiota and Its Fermentation Characteristics of Ningxiang Pigs at the Young Stage.
533 Animals (Basel) 11:638. https://doi.org/10.3390/ani11030638.
534 29. Mach N, Berri M, Estelle J, Levenez F, Lemonnier G, Denis C, Leplat JJ, Chevaleyre C,
535 Billon Y, Dore J, Rogel-Gaillard C, Lepage P. 2015. Early-life establishment of the swine
536 gut microbiome and impact on host phenotypes. Environ Microbiol Rep 7:554-69.
537 https://doi.org/10.1111/1758-2229.12285.
538 30. De Rodas B, Youmans BP, Danzeisen JL, Tran H, Johnson TJ. 2018. Microbiome
539 profiling of commercial pigs from farrow to finish. J Anim Sci 96:1778-1794.
540 https://doi.org/10.1093/jas/sky109.
541 31. Wang X, Tsai T, Deng F, Wei X, Chai J, Knapp J, Apple J, Maxwell CV, Lee JA, Li Y,
542 Zhao J. 2019. Longitudinal investigation of the swine gut microbiome from birth to
543 market reveals stage and growth performance associated bacteria. Microbiome 7:109.
544 https://doi.org/10.1186/s40168-019-0721-7.
545 32. Gaio D, DeMaere MZ, Anantanawat K, Chapman TA, Djordjevic SP, Darling AE. 2021.
546 Post-weaning shifts in microbiome composition and metabolism revealed by over 25 000

pig gut metagenome-assembled genomes. Microb Genom 7. https://doi.org/10.1099/mgen.0.000501.

33. Crossfield M, Gilroy R, Ravi A, Baker D, La Ragione RM, Pallen MJ. 2022. Archaeal and Bacterial Metagenome-Assembled Genome Sequences Derived from Pig Feces. Microbiol Resour Announc 11:e0114221. https://doi.org/10.1128/mra.01142-21.

34. Maguire F, Jia B, Gray KL, Lau WYV, Beiko RG, Brinkman FSL. 2020. Metagenome-assembled genome binning methods with short reads disproportionately fail for plasmids and genomic Islands. Microb Genom 6. https://doi.org/10.1099/mgen.0.000436.

35. Tunsagool P, Mhuantong W, Tangphatsornruang S, Am-In N, Chuanchuen R, Luangtongkum T, Suriyaphol G. 2021. Metagenomics of antimicrobial and heavy metal resistance in the cecal microbiome of fattening pigs raised without antibiotics. Appl Environ Microbiol 87:e02684-20. https://doi.org/10.1128/AEM.02684-20.

36. Pollock J, Muwonge A, Hutchings MR, Mainda G, Bronsvoort BM, Gally DL, Corbishley A. 2020. Resistance to change: AMR gene dynamics on a commercial pig farm with high antimicrobial usage. Sci Rep 10:1708. https://doi.org/10.1038/s41598-020-58659-3.

37. Holman DB, Chenier MR. 2013. Impact of subtherapeutic administration of tylosin and chlortetracycline on antimicrobial resistance in farrow-to-finish swine. FEMS Microbiol Ecol 85:1-13. https://doi.org/10.1111/1574-6941.12093.

38. Zhou Y, Fu H, Yang H, Wu J, Chen Z, Jiang H, Liu M, Liu Q, Huang L, Gao J, Chen C. 2022. Extensive metagenomic analysis of the porcine gut resistome to identify indicators reflecting antimicrobial resistance. Microbiome 10:39. https://doi.org/10.1186/s40168-022-01241-y.

570 39. Holman DB, Brunelle BW, Trachsel J, Allen HK. 2017. Meta-analysis To Define a Core
571 Microbiota in the Swine Gut. mSystems 2:e00004-17.
572 https://doi.org/10.1128/mSystems.00004-17.
573 40. Valeriano VD, Balolong MP, Kang DK. 2017. Probiotic roles of *Lactobacillus* sp. in
574 swine: insights from gut microbiota. J Appl Microbiol 122:554-567.
575 https://doi.org/10.1111/jam.13364.
576 41. Funk JA, Lejeune JT, Wittum TE, Rajala-Schultz PJ. 2006. The effect of subtherapeutic
577 chlortetracycline on antimicrobial resistance in the fecal flora of swine. Microb Drug
578 Resist 12:210-8. https://doi.org/10.1089/mdr.2006.12.210.
579 42. Poulin-Laprade D, Brouard JS, Gagnon N, Turcotte A, Langlois A, Matte JJ, Carrillo CD,
580 Zaheer R, McAllister T, Topp E, Talbot G. 2021. Resistance determinants and their
581 genetic context in enterobacteria from a longitudinal study of pigs reared under various
582 husbandry conditions. Appl Environ Microbiol 87:e02612-20.
583 https://doi.org/10.1128/AEM.02612-20.
584 43. Ghanbari M, Klose V, Crispie F, Cotter PD. 2019. The dynamics of the antibiotic
585 resistome in the feces of freshly weaned pigs following therapeutic administration of
586 oxytetracycline. Sci Rep 9:4062. https://doi.org/10.1038/s41598-019-40496-8.
587 44. Lekagul A, Tangcharoensathien V, Yeung S. 2019. Patterns of antibiotic use in global pig
588 production: a systematic review. Vet Anim Sci 7:100058.
589 https://doi.org/10.1016/j.vas.2019.100058.
590 45. Igwaran A, Okoh AI. 2019. Human campylobacteriosis: A public health concern of
591 global importance. Heliyon 5:e02814. https://doi.org/10.1016/j.heliyon.2019.e02814.

592 46. Wang S, Ma M, Liang Z, Zhu X, Yao H, Wang L, Wu Z. 2021. Pathogenic investigations
593 of Streptococcus pasteurianus, an underreported zoonotic pathogen, isolated from a
594 diseased piglet with meningitis. Transbound Emerg Dis doi:10.1111/tbed.14413.
595 https://doi.org/10.1111/tbed.14413.
596 47. Chow JW, Kak V, You I, Kao SJ, Petrin J, Clewell DB, Lerner SA, Miller GH, Shaw KJ.
597 2001. Aminoglycoside resistance genes *aph(2'')-Ib* and *aac(6')-Im* detected together in
598 strains of both Escherichia coli and Enterococcus faecium. Antimicrob Agents
599 Chemother 45:2691-4. https://doi.org/10.1128/AAC.45.10.2691-2694.2001.
600 48. Speer BS, Bedzyk L, Salyers AA. 1991. Evidence that a novel tetracycline resistance
601 gene found on two *Bacteroides* transposons encodes an NADP-requiring oxidoreductase.
602 J Bacteriol 173:176-83. https://doi.org/10.1128/jb.173.1.176-183.1991.
603 49. Corver J, Bakker D, Brouwer MS, Harmanus C, Hensgens MP, Roberts AP, Lipman LJ,
604 Kuijper EJ, van Leeuwen HC. 2012. Analysis of a *Clostridium difficile* PCR ribotype 078
605 100 kilobase island reveals the presence of a novel transposon, Tn6164. BMC Microbiol
606 12:130. https://doi.org/10.1186/1471-2180-12-130.
607 50. Abril C, Brodard I, Perreten V. 2010. Two novel antibiotic resistance genes, *tet*(44) and
608 *ant(6)-Ib*, are located within a transferable pathogenicity island in *Campylobacter fetus*
609 subsp. *fetus*. Antimicrob Agents Chemother 54:3052-3055.
610 https://doi.org/10.1128/AAC.00304-10.
611 51. Aziz RK, Breitbart M, Edwards RA. 2010. Transposases are the most abundant, most
612 ubiquitous genes in nature. Nucleic Acids Res 38:4207-4217.
613 https://doi.org/10.1093/nar/gkq140.

614 52. Achard A, Villers C, Pichereau V, Leclercq R. 2005. New *lnu*(C) gene conferring
615 resistance to lincomycin by nucleotidylation in *Streptococcus agalactiae* UCN36.
616 Antimicrob Agents Chemother 49:2716-9. https://doi.org/10.1128/AAC.49.7.2716-
617 2719.2005.
618 53. Lyras D, Adams V, Ballard SA, Teng WL, Howarth PM, Crellin PK, Bannam TL,
619 Songer JG, Rood JI. 2009. tIS Cpe8, an IS*1595*-family lincomycin resistance element
620 located on a conjugative plasmid in *Clostridium perfringens*. J Bacteriol 191:6345-6351.
621 https://doi.org/10.1128/JB.00668-09.
622 54. Le Roy T, Van der Smissen P, Paquot A, Delzenne N, Muccioli GG, Collet JF, Cani PD.
623 2020. *Dysosmobacter welbionis* gen. nov., sp. nov., isolated from human faeces and
624 emended description of the genus *Oscillibacter*. Int J Syst Evol Microbiol 70:4851-4858.
625 https://doi.org/10.1099/ijsem.0.003547.
626 55. Le Roy T, de Hase EM, Van Hul M, Paquot A, Pelicaen R, Régnier M, Depommier C,
627 Druart C, Everard A, Maiter D. 2022. *Dysosmobacter welbionis* is a newly isolated
628 human commensal bacterium preventing diet-induced obesity and metabolic disorders in
629 mice. Gut 71:534-543. https://doi.org/10.1136/gutjnl-2020-323778.
630 56. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina
631 sequence data. Bioinformatics 30:2114-20.
632 https://doi.org/10.1093/bioinformatics/btu170.
633 57. Warr A, Affara N, Aken B, Beiki H, Bickhart DM, Billis K, Chow W, Eory L, Finlayson
634 HA, Flicek P, Giron CG, Griffin DK, Hall R, Hannum G, Hourlier T, Howe K, Hume
635 DA, Izuogu O, Kim K, Koren S, Liu H, Manchanda N, Martin FJ, Nonneman DJ,
636 O'Connor RE, Phillippy AM, Rohrer GA, Rosen BD, Rund LA, Sargent CA, Schook LB,

637 Schroeder SG, Schwartz AS, Skinner BM, Talbot R, Tseng E, Tuggle CK, Watson M,
638 Smith TPL, Archibald AL. 2020. An improved pig reference genome sequence to enable
639 pig genetics and genomics research. Gigascience 9:giaa051.
640 https://doi.org/10.1093/gigascience/giaa051.
641 58. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat
642 Methods 9:357-9. https://doi.org/10.1038/nmeth.1923.
643 59. Li D, Liu C-M, Luo R, Sadakane K, Lam T-WJB. 2015. MEGAHIT: an ultra-fast single-
644 node solution for large and complex metagenomics assembly via succinct de Bruijn
645 graph. 31:1674-1676.
646 60. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an
647 adaptive binning algorithm for robust and efficient genome reconstruction from
648 metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359.
649 61. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GWJGr. 2015. CheckM:
650 assessing the quality of microbial genomes recovered from isolates, single cells, and
651 metagenomes. 25:1043-1055.
652 62. Olm MR, Brown CT, Brooks B, Banfield JFJTIj. 2017. dRep: a tool for fast and accurate
653 genomic comparisons that enables improved genome recovery from metagenomes
654 through de-replication. 11:2864-2868.
655 63. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to
656 classify genomes with the Genome Taxonomy Database. Oxford University Press.
657 64. Asnicar F, Thomas AM, Beghini F, Mengoni C, Manara S, Manghi P, Zhu Q, Bolzan M,
658 Cumbo F, May UJNc. 2020. Precise phylogenetic analysis of microbial isolates and
659 genomes from metagenomes using PhyloPhlAn 3.0. 11:1-10.

660 65. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for
661 phylogenetic tree display and annotation. Nucleic Acids Res 49:W293-W296.
662 https://doi.org/10.1093/nar/gkab301.
663 66. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P,
664 Narrowe AB, Rodriguez-Ramos J, Bolduc B, Gazitua MC, Daly RA, Smith GJ, Vik DR,
665 Pope PB, Sullivan MB, Roux S, Wrighton KC. 2020. DRAM for distilling microbial
666 metabolism to automate the curation of microbiome function. Nucleic Acids Res
667 48:8883-8900. https://doi.org/10.1093/nar/gkaa621.
668 67. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018.
669 dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. Nucleic
670 Acids Res 46:W95-W101. https://doi.org/10.1093/nar/gky418.
671 68. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W,
672 Nguyen AV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran HK, Werfalli RE,
673 Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-
674 Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D,
675 Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar
676 GV, McArthur AG. 2020. CARD 2020: antibiotic resistome surveillance with the
677 comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517-D525.
678 https://doi.org/10.1093/nar/gkz935.
679 69. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics
680 30:2068-9. https://doi.org/10.1093/bioinformatics/btu153.
681 70. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, Tickle TL, Weingart
682 G, Ren B, Schwager EH, Chatterjee S, Thompson KN, Wilkinson JE, Subramanian A, Lu

683 Y, Waldron L, Paulson JN, Franzosa EA, Bravo HC, Huttenhower C. 2021. Multivariable
684 association discovery in population-scale meta-omics studies. PLoS Comput Biol
685 17:e1009442. https://doi.org/10.1371/journal.pcbi.1009442.

686

687 **Figure legends**

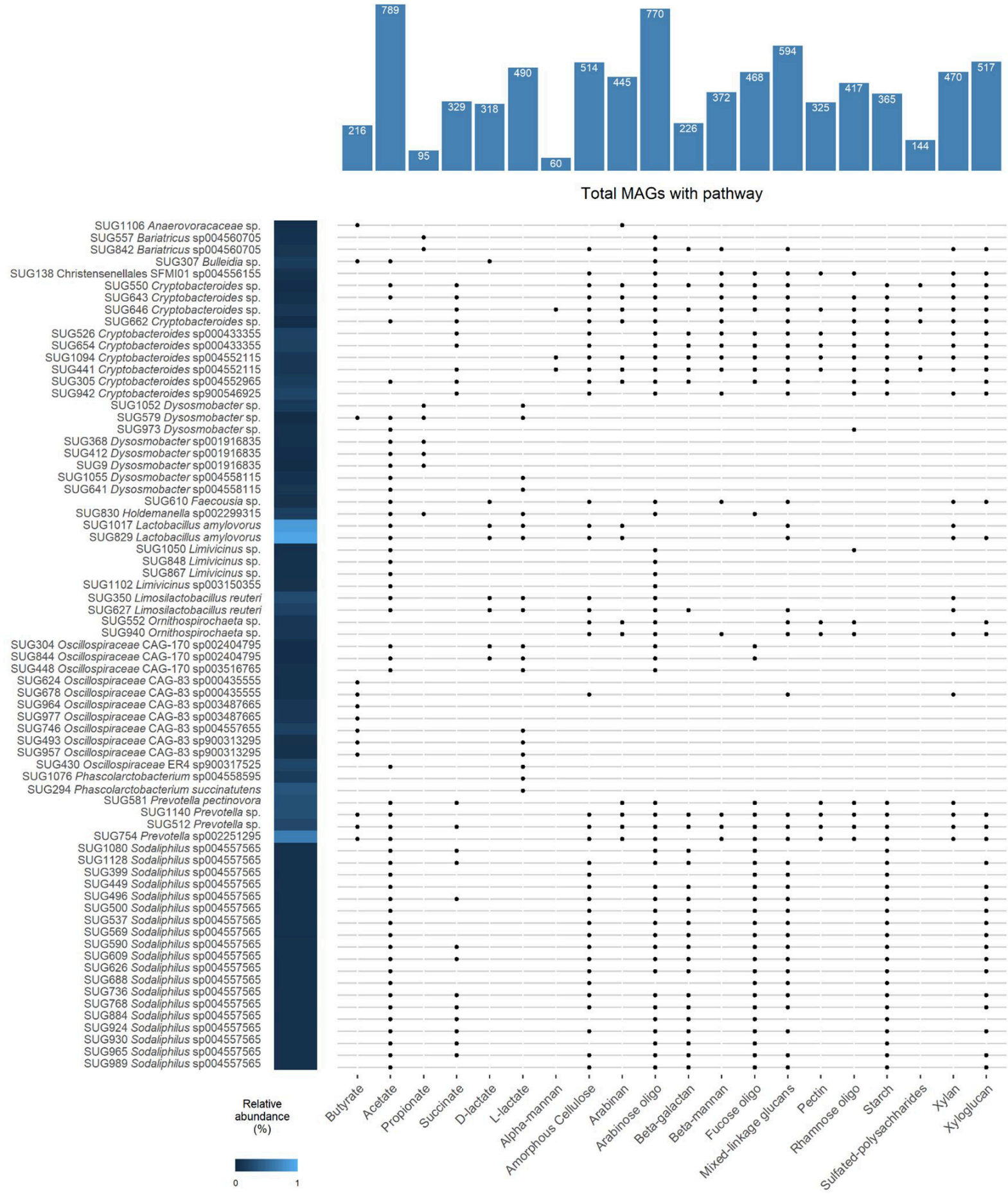
688 **Figure 1**. Maximum likelihood phylogenetic tree of 1,150 MAGs based on the alignment of 399
689 marker genes in PhyloPhlAn. MAGs colored by GTBD-tk assigned phyla are labelled in the
690 inner ring. The outer ring indicates the number of carbohydrate-active enzymes (CAZymes) per
691 MAG and the outer bars display the percent relative abundance (minimum = 0%; maximum =
692 2.73%) of each MAG in the pre- and post-weaning fecal samples.

693 **Figure 2**. Metagenome-assembled genomes (MAGs) that were identified in 85% or more of all
694 samples from this study and publicly available metagenome samples. The relative abundance
695 within all of the MAGs within these samples (n = 805) is displayed as a heat map and the
696 presence of genes encoding for pathways involved in selected short-chain fatty acid and other
697 organic acid production, as well as polysaccharide degradation (carbohydrate-active enzymes
698 [CAZymes]) is indicated by a dot. The total number of MAGs (n = 1,150) that encode these
699 pathways are displayed on the top of the plot.

700

701

702


216
789
95
329
318
490
60
514
445
770
226
372
468
594
325
417
365
144
470
517
Total MAGs with pathway
SUG1106 Anaerovoracaceae sp.
SUG557 Bariatricus sp004560705
SUG842 Bariatricus sp004560705
SUG307 Bulleidia sp.
SUG138 Christensenellales SFMI01 sp004556155
SUG550 Cryptobacteroides sp.
SUG643 Cryptobacteroides sp.
SUG646 Cryptobacteroides sp.
SUG662 Cryptobacteroides sp.
SUG526 Cryptobacteroides sp000433355
SUG654 Cryptobacteroides sp000433355
SUG1094 Cryptobacteroides sp004552115
SUG441 Cryptobacteroides sp004552115
SUG305 Cryptobacteroides sp004552965
SUG942 Cryptobacteroides sp900546925
SUG1052 Dysosmobacter sp.
SUG579 Dysosmobacter sp.
SUG973 Dysosmobacter sp.
SUG368 Dysosmobacter sp001916835
SUG412 Dysosmobacter sp001916835
SUG9 Dysosmobacter sp001916835
SUG1055 Dysosmobacter sp004558115
SUG641 Dysosmobacter sp004558115
SUG610 Faecousia sp.
SUG830 Holdemanella sp002299315
SUG1017 Lactobacillus amylovorus
SUG829 Lactobacillus amylovorus
SUG1050 Limivicinus sp.
SUG848 Limivicinus sp.
SUG867 Limivicinus sp.
SUG1102 Limivicinus sp003150355
SUG350 Limosilactobacillus reuteri
SUG627 Limosilactobacillus reuteri
SUG552 Ornithospirochaeta sp.
SUG940 Ornithospirochaeta sp.
SUG304 Oscillospiraceae CAG-170 sp002404795
SUG844 Oscillospiraceae CAG-170 sp002404795
SUG448 Oscillospiraceae CAG-170 sp003516765
SUG624 Oscillospiraceae CAG-83 sp000435555
SUG678 Oscillospiraceae CAG-83 sp000435555
SUG964 Oscillospiraceae CAG-83 sp003487665
SUG977 Oscillospiraceae CAG-83 sp003487665
SUG746 Oscillospiraceae CAG-83 sp004557655
SUG493 Oscillospiraceae CAG-83 sp900313295
SUG957 Oscillospiraceae CAG-83 sp900313295
SUG430 Oscillospiraceae ER4 sp900317525
SUG1076 Phascolarctobacterium sp004558595
SUG294 Phascolarctobacterium succinatutens
SUG581 Prevotella pectinovora
SUG1140 Prevotella sp.
SUG512 Prevotella sp.
SUG754 Prevotella sp002251295
SUG1080 Sodaliphilus sp004557565
SUG1128 Sodaliphilus sp004557565
SUG399 Sodaliphilus sp004557565
SUG449 Sodaliphilus sp004557565
SUG496 Sodaliphilus sp004557565
SUG500 Sodaliphilus sp004557565
SUG537 Sodaliphilus sp004557565
SUG569 Sodaliphilus sp004557565
SUG590 Sodaliphilus sp004557565
SUG609 Sodaliphilus sp004557565
SUG626 Sodaliphilus sp004557565
SUG688 Sodaliphilus sp004557565
SUG736 Sodaliphilus sp004557565
SUG768 Sodaliphilus sp004557565
SUG884 Sodaliphilus sp004557565
SUG924 Sodaliphilus sp004557565
SUG930 Sodaliphilus sp004557565
SUG965 Sodaliphilus sp004557565
SUG989 Sodaliphilus sp004557565
Relative abundance (%)
0
1
Butyrate
Acetate
Propionate
Succinate
D-lactate
L-lactate
Alpha-mannan
Amorphous Cellulose
Arabinan
Arabinose oligo
Beta-galactan
Beta-mannan
Fucose oligo
Mixed-linkage glucans
Pectin
Rhamnose oligo
Starch
Sulfated-polysaccharides
Xylan
Xyloglucan