bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

1 **Comparative genomics guides elucidation of vitamin B12 biosynthesis in novel human**
2 **associated *Akkermansia***
3
4 Nina Kirmiz[1], Kadir Galindo[1], Karissa L. Cross[2,3], Estefani Luna[1], Nicholas Rhoades[1,4], Mircea
5 Podar[2,3], and Gilberto E. Flores[1*]
6
7 [1]Department of Biology, California State University, Northridge, CA 91330, United States.
8 [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States.
9 [3]Microbiology Department, University of Tennessee Knoxville, Knoxville TN
10 [4]Department of Molecular Biology and Biochemistry, University of California, Irvine, CA
11 92697, United States.
12
13 [*]Corresponding author: Gilberto E. Flores
14 Department of Biology
15 California State University, Northridge
16 18111 Nordhoff Street
17 Northridge, CA 91330-8303
18 Phone: +1 (818) 677-4276
19 Email: gilberto.flores@csun.edu
20 nina.kirmiz@csun.edu, kadir.galindo.150@my.csun.edu, crosskl@ornl.gov,
21 estefani.luna.438@my.csun.edu, rhoadesn@uci.edu, podarm@ornl.gov

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

22 **ABSTRACT**

23 **Background:** *Akkermansia muciniphila* is a mucin-degrading bacterium found in the gut of most
24 healthy humans and is considered a 'next-generation probiotic.' However, knowledge of the
25 genomic and physiological diversity of human associated *Akkermansia* is limited, as only one
26 species has been formally described.
27 **Results:** To begin to fill this knowledge gap, we reconstructed 35 high-quality metagenome
28 assembled genomes from children and combined them with 40 other publicly available genomes
29 from adults and mice for comparative genomic analysis. We identified at least four species-level
30 phylogroups (AmI-AmIV) with distinct functional potentials. Most notably, we identified the
31 presence of putative cobalamin (vitamin B12) biosynthesis genes within the AmII (n=26/28) and
32 AmIII (n=2/2) phylogroups. To test these predictions, 10 novel strains of *Akkermansia* were
33 isolated from adults and screened for essential vitamin B12 biosynthesis genes via PCR. Two
34 strains of the AmII phylogroup were positive for the presence of vitamin B12 biosynthesis genes,
35 while all AmI strains, including the type strain *A. muciniphila* Muc^{T}, were negative. To
36 demonstrate vitamin B12 biosynthesis, we measured the production of acetate, succinate, and
37 propionate in the presence and absence of vitamin supplementation in representative strains of
38 the AmI and AmII phylogroups since cobalamin is a cofactor in propionate metabolism. Results
39 show that the *Akkermansia* AmII strain produced acetate and propionate in the absence of
40 supplementation, which is indicative of *de novo* vitamin B12 biosynthesis. In contrast, acetate
41 and succinate were the main fermentation products for the AmI strains when vitamin B12 was
42 not supplied in the culture medium.
43 **Conclusions:** We identified *Akkermansia* strains as potentially important vitamin B12
44 biosynthetic bacteria in the human gut. This novel physiological trait of human associated

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

45 *Akkermansia* may impact how these bacteria interact with the human host and other members of
46 the human gut microbiome.
47
48 **KEYWORDS:** *Akkermansia*, intestinal bacteria, vitamin B12, human gut microbiome,
49 probiotics
50
51 **BACKGROUND**

52 *Akkermansia muciniphila* is a mucin degrading, gram-negative intestinal bacterium
53 widely present in the human population, typically at 1 to 4% relative abundance [1, 2]. A number
54 of studies in humans [3-5] and rodents [6-8] have found positive associations between its
55 abundance and intestinal health, suggesting that *Akkermansia* may be a beneficial member of the
56 gut microbiome and could be used as a biomarker of a healthy gut [9-11]. However, despite a
57 diversity of phylotypes being reported by previous sequence-based studies, *A. muciniphila* Muc^T
58 (ATCC BAA-835) represents the sole described species of the Verrucomicrobia phylum
59 associated with humans [2, 12, 13]. Therefore, before we can fully assess the health potential of
60 human associated *Akkermansia*, a comprehensive understanding of the genomic and
61 physiological diversity of this lineage is needed.

62 Recently, a pangenomic study that included 33 new isolates from adults and 6 from
63 laboratory mice provided insights into the population structure and evolutionary history of the
64 *Akkermansia* lineage [14]. Specifically, this study revealed an open pangenome with at least
65 three species-level phylogroups (AmI, AmII, and AmIII) that appear to be evolving
66 independently. Although genomic differences amongst phylogroups were noted, the
67 physiological consequences were not explored.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

68 To continue to expand our understanding of the genomic content and functional potential
69 of human associated *Akkermansia*, we reconstructed 35 *Akkermansia* genomes from children
70 aged 2-9 years and combined our genomes with those from Guo et al. [14]. With these genomes,
71 we identified novel diversity and several putative functional differences amongst the
72 *Akkermansia* phylogroups. Most notably, we identified the presence of genes associated with *de*
73 *novo* cobalamin (vitamin B12) biosynthesis in selected phylogroups of *Akkermansia*.
74 Furthermore, using isolates obtained from healthy adults, we tested these genomic predictions
75 and confirmed vitamin B12 biosynthesis by select strains of human-associated *Akkermansia*.
76 These results build upon our understanding of the physiological capabilities of human-associated
77 *Akkermansia* and demonstrate an important biosynthetic activity by this bacterial lineage that
78 further expands its potential beneficial role in the intestinal environment.
79

80 **RESULTS**

81 ***Comparative genomics***

82 A total of 334.9 Gbp of metagenomic sequence data was obtained from 70 children aged
83 2-9 years. Using SPAdes [15] to assemble contigs and MetaBAT [16] to bin contigs, we
84 recovered 35 high quality metagenome assembled genomes (MAGs) of human associated
85 *Akkermansia* (Table 1) from 35 of the 70 children. Completeness of the MAGs was relatively
86 high, ranging from 68.5% to 95.5% with 31 of 35 MAGs > 90% complete. Likewise,
87 contamination of the MAGs were low with all < 1%. On average, each MAG was 2.87 Mbp in
88 length and contained approximately 2,420 protein-coding genes.
89 To explore the genomic diversity of human associated *Akkermansia*, we performed a
90 pangenomic analysis using tools in anvi'o [17, 18]. These analyses included the closed genome

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

91 of the type strain [12] as well as 33 other human-associated and 6 mouse-associated *Akkermansia*
92 genomes [14]. Previously, these 40 *Akkermansia* genomes were used to define three species-
93 level phylogroups, AmI, AmII, and AmIII [14]. Merging our 35 MAGs with these 40 other
94 genomes, we were able to regenerate the three phylogroups, but also observed a fourth
95 phylogroup (AmIV) based on average nucleotide identity (ANI) calculated using PyANI [19]
96 (heatmap in Figure 1). Phylogroup AmI, which includes the type strain *A. muciniphila* Muc^T,
97 contained the largest number of genomes with 40, followed by AmII (n=26), AmIV (n=7), and
98 AmIII (n=2). Phylogroup AmIII was not observed in any of our 35 MAGs. Interestingly, both
99 AmI and AmII included isolates obtained from mice. Within each phylogroup, ANI ranged from
100 93.94% to 99.98% across > 65% of each pair of genomes (Supplemental Figure 1). All between
101 phylogroup ANI comparisons were < 92%. One genome in AmIV (CDI-148A-8) showed lower
102 similarity (on average ~94%) to other genomes within this phylogroup, possibly indicating
103 further species level diversity within human associated *Akkermansia*. Across all phylogroups, we
104 identified 6,557 gene clusters (GCs) with 1,021 found in all 75 genomes and 1,240 found in only
105 one genome (Figure 1). Functional genes within the core included the cytochrome bd [20]
106 (Amuc_1694 and Amuc_1695) and Type IV pili genes [21-23] (Amuc_1098 – Amuc_1102)
107 previously characterized from *A. muciniphila* Muc^T.

108 Next, we were interested in identifying functional gene predictions that differed amongst
109 the phylogroups. Using Clusters of Orthologous Group (COG) annotations of GCs implemented
110 in anvi'o, we observed 7 GCs putatively involved in the corrin ring stage of cobalamin (vitamin
111 B12) biosynthesis within the AmII (n=24/26) and AmIII (n=2/2) phylogroups (Supplemental
112 Data 1). To investigate these genes in greater detail, we manually inspected the annotations of all
113 75 genomes using Integrated Microbial Genome (IMG) [24] and Geneious 7.1.3

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

114 (https://www.geneious.com). With this approach, we confirmed the COG annotations and
115 identified a cluster of 8 genes that appears to code for the corrin ring biosynthesis proteins in a
116 subset of *Akkermansia* genomes (Figure 2). Included in this genomic region were genes *cbiK/X*,
117 *cbiL*, *cbiC*, *cbiD*, *cbiET*, *cbiFGH*, and *cbiA* that encode the enzymes associated with the
118 anaerobic pathway of corrin ring biosynthesis. This cluster also contains a gene annotated as a
119 hypothetical protein, which shows some similarity to a putative cobalt transporter [25]. The
120 content and arrangement of these genes was similar to the only other named species of the
121 *Akkermansia* genus, *Akkermansia glycaniphila* Pyt^T previously isolated from a python [26].
122 Additionally, all 75 genomes contained most of the genes associated with the upstream
123 (tetrapyrrole precursor biosynthesis: e.g. Amuc_0090 – 0091, Amuc_0417, Amuc_0896, and
124 Amuc_1730) and downstream (nucleotide loop assembly, e.g. Amuc_1678 – Amuc_1683) stages
125 of vitamin B12 biosynthesis [27]. Genes annotated as a TonB dependent transporter (e.g.
126 Amuc_1684) and an extracellular solute-binding family 5 protein (e.g. Amuc_1685) that may be
127 involved with vitamin B12 import were also identified adjacent to the nucleotide loop assembly
128 genes in all but one genome.

129 *A. muciniphila* Muc^T was previously classified as a Cbi salvager because it lacked the
130 genes coding for the enzymes to synthesize the corrin ring of vitamin B12, yet it needs this
131 cofactor for methionine synthesis, nucleotide synthesis, queuosine synthesis and propionate
132 metabolism [27]. Indeed, genes associated with these cellular functions were conserved across all
133 phylogroups (Supplemental Data 1). Interestingly, the vitamin B12 independent methionine
134 synthase II gene (*metE*) was present in 25 of 40 AmI genomes but not in any of the other
135 genomes including the type strain Muc^T. Together, these observations suggest that all

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

136 *Akkermansia* examined here are able to acquire and likely remodel corrinoids from the
137 environment for use, but some are also able to synthesize this important cofactor.
138

139 ***Cultivation and validation of vitamin B12 biosynthesis***

140 To determine if specific *Akkermansia* species/strains are indeed able to *de novo*
141 synthesize vitamin B12, we first isolated several strains of *Akkermansia* from healthy adults and
142 compared their near-full length 16S rRNA gene sequences with those from Guo et al. [14] in
143 ARB [28] to determine phylogroup affiliation. Across phylogroups AmI, AmII, and AmIII, 16S
144 rRNA gene sequences were all greater than 97% identical but nevertheless clustered into the
145 known phylogroups (Supplemental Figure 2). Based on this approach, we identified eight AmI
146 and two AmII isolates in our culture collection. Because our MAGs did not contain any full-
147 length 16S rRNA gene sequences, we could not positively identify AmIV members among the
148 isolates.

149 Next, using the AmII and AmIII genomes and the genome of *A. glycaniphila* Pyt^T, we
150 designed degenerate polymerase chain reaction (PCR) primers targeting four genes, *cbiL*, *cbiC*,
151 *cbiD*, and *cbiFGH,* of the corrin ring biosynthesis gene cluster, which encode a cobalt-factor II
152 C20-methyltransferase, a cobalt-precorrin-8 methylmutase, a cobalt-precorrin-5B (C(1))-
153 methyltransferase, and a cobalt-precorrin-4 methyltransferase/precorrin-3B C17-
154 methyltransferase, respectively (Supplemental Table 1). These genes were selected because they
155 are predicted to give the best indication of cobamide production as described by Shelton et al.
156 [27]. As expected, only isolates from the AmII phylogroup (CSUN-17 and CSUN-34) and *A.*
157 *glycaniphila* Pyt^T showed positive amplification, whereas all AmI isolates (including *A.*
158 *muciniphila* Muc^T) failed to amplify (Table 2). Sequencing and BLASTing of these PCR

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

159 amplicons from CSUN-17 against *A. glycaniphila* ERS 1290231 and *Desulfovibrio vulgaris* str.
160 Hildenborough confirmed the identity of these gene fragments (Supplemental Table 2) clearly
161 demonstrating the presence of select *cbi* genes in the AmII phylogroup.

162 It is known that many fermentative bacteria including *A. muciniphila* Muc^T, use vitamin
163 B12 to activate methylmalonyl-CoA synthase to convert succinate to propionate [29, 30].
164 Therefore, to demonstrate vitamin B12 biosynthesis *in vitro*, we quantified the production of
165 succinate and propionate (and acetate) in the presence and absence of vitamin B12 in mucin
166 medium (Figure 3). Our predictions were that the AmI phylogroup (represented by *A.*
167 *muciniphila* Muc^T) would produce acetate and succinate in the absence of vitamin B12, and
168 acetate and propionate when B12 was present. For AmII, we predicted that acetate and
169 propionate would be produced regardless of whether the culture medium was supplemented with
170 vitamin B12. Results show that the AmI isolate produced propionate in a vitamin B12
171 concentration-dependent manner (Figure 3B,C). Also as expected, the CSUN-17 isolate (AmII),
172 produced significant amounts of acetate and propionate in the absence and presence of vitamin
173 B12, but production was more rapid with supplementation (Figure 3D-F). These results clearly
174 indicate vitamin B12 biosynthesis by the *Akkermansia* species represented in the AmII
175 phylogroup.

176

177 **DISCUSSION**

178 *A. muciniphila* is a common gut bacterium highly regarded as beneficial member of the
179 human gut microbiome with important probiotic potential [10, 31]. Various studies have
180 described positive associations between the abundance of *Akkermansia* and intestinal health [3,
181 4]. For example, *A. muciniphila* affects glucose metabolism, intestinal immunity, and its

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

182 abundance in the gastrointestinal tract (GIT) is inversely correlated with diseases including
183 Crohn's disease, ulcerative colitis, and acute appendicitis [32-35]. Although a number of 16S
184 rRNA gene variants have been observed [12] and dozens of isolates have been obtained [14],
185 human-associated *Akkermansia* have largely been thought of as a single species and the
186 functional potential beyond mucin degradation has gone largely unexplored. Here, we
187 demonstrate that there are significant genomic and physiological differences amongst the human
188 associated *Akkermansia*. Through comparative genomic analysis, we identified four phylogroups
189 of human-associated *Akkermansia,* expanding the known genomic diversity of this lineage.
190 Although all 16S rRNA gene sequenced examined here and elsewhere [32] are >97% identical,
191 using an ANI of 95% across genomes as a species level delineation [36, 37] would suggest that
192 each phylogroup represents a different species of *Akkermansia*. When we examined gene
193 content, several phylogroup specific genes were identified that are predicted to code for
194 functional differences amongst phylogroups further supporting species delineation. Most
195 notably, we identified a complete set of genes involved in *de novo* biosynthesis of cobalamin, or
196 vitamin B12, in two of the four phylogroups. We were able to validate these predictions *in vitro*
197 using novel strains obtained from healthy adults. These findings demonstrate an ecological
198 important function [38] not previously associated with human-associated *Akkermansia,*
199 fundamentally altering our understanding of the diversity and physiology of this lineage. More
200 broadly, these results continue to demonstrate the importance of merging next-generation
201 sequencing approaches with traditional cultivation approaches to understand the basic biology of
202 microorganisms of significance.

203 A recent comparative genomic analysis examining 11,000 bacterial genomes for
204 cobamide production revealed that approximately 37% of bacteria are predicted to synthesize

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

205 cobamides, yet 86% require them for at least one cellular function [27, 39]. Additionally, Degnan
206 et al. found that most vitamin B12 dependent human gut bacteria lack the ability to synthesize
207 vitamin B12 [39]. The type strain *A. muciniphila* Muc^T was included in the analysis by Shelton
208 and colleagues [27] and was described as a Cbi salvager able to use exogenous sources of
209 vitamin B12. Indeed, based on previous *in vitro* coculture experiments, *A. muciniphila* Muc^T can
210 use at least three types of cobamides; cyanocobalamin supplied in culture medium,
211 pseudovitamin B12 produced by *Eubacterium hallii* L2-7 [29], and an unknown form produced
212 by *Anaerostipes caccae* [40]. Presumably, *Akkermansia* are able to import these various forms of
213 cobalamin and use them directly or remodel the lower ligand to suit their needs. With our
214 findings, some *Akkermansia* can now be considered producers of corrinoids, altering our
215 understanding of how they interact with other members of the human gut microbiome and
216 potentially their human host. However, questions remain regarding the type of cobalamin
217 produced by AmII members and more generally regarding the specificity and efficiency of
218 cobamide import and remodeling by all *Akkermansia*.
219 Cobalamin produced by bacteria and archaea in the large intestine is not readily available
220 to the human host for two main reasons [38]. First, the receptors responsible for cobalamin
221 absorption are found in the small intestine, which is not as densely colonized by bacteria as the
222 large intestine in times of health. Second, although bacteria produce many different types of
223 cobalamin, their contribution to the available pool of cobalamin is small because many of the
224 forms produced by bacteria are not recognized by human receptors. Thus, bacteria are thought of
225 more as competitors for dietary cobalamin than suppliers. However, if a bacterium colonized the
226 small intestine and produced an appropriate form of cobalamin then the cofactor is possibly
227 available to the human host. With regards to *Akkermansia*, we do not yet know the form of

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

228 cobalamin produced by AmII members but *Akkermansia*-like organisms have been observed
229 throughout the human gastrointestinal tract, including in the small intestine (reviewed in [32]).
230 Interestingly, phylogenetic analyses consistently group AmII and AmIII isolates [14] with clones
231 and other sequences previously observed in the small intestine [32]. Because our genomic
232 sequence data and isolates were obtained from fecal samples, we could not determine if the
233 different phylogroups colonize different segments of the gastrointestinal tract, but it is intriguing
234 to speculate.

235 Although we do not yet know if humans can directly benefit from vitamin B12 produced
236 by *Akkermansia*, there are indirect benefits resulting from the altered metabolites produced when
237 vitamin B12 is available. Specifically, the type and quantity of short chain fatty acids (SCFA)
238 produced during fermentation influences host health [41-43]. For example, propionate is known
239 to help regulate appetite by stimulating the release of peptide YY (PYY) and glucagon like
240 peptide-1 (GLP-1) by human colonic cells [44]. Less is known about the potential benefits of
241 succinate in the human gut but in the mice cecum, succinate does improve glucose homeostasis
242 via intestinal gluconeogenesis [45]. Conversely in the human small intestine, succinate has been
243 shown to trigger a type 2 immune inflammatory response initiated by epithelial tuft cells [46].
244 Thus, possessing the ability to synthesize vitamin B12 *de novo* would suggest that the AmII and
245 AmIII phylogroups have the potential to consistently produce more propionate than succinate
246 during mucin fermentation and, as a result, influence gut epithelial cell behavior. If the AmII
247 and/or AmIII phylogroups do colonize the small intestine, being able to consistently produce
248 propionate over succinate could have significant health implications.

249 In addition to propionate metabolism, *Akkermansia* are predicted to use vitamin B12 as a
250 cofactor for methionine biosynthesis using methionine synthase type I (MetH). All genomes

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

251 possessed the *metH* gene; however, select AmI genomes (n=25/40) also contain the B12
252 independent methionine synthase II gene (*metE*), suggesting that these select AmI strains can
253 generate methionine in the absence of vitamin B12. Given that the AmI phylogroup does not
254 synthesize vitamin B12, this would allow production of this essential amino acid when
255 exogenous corrinoids were unavailable. How readily available corrinoids are to *Akkermansia*
256 either from other bacterial producers or host diet is unknown but possessing both variants may be
257 an adaptive strategy for AmI strains.

258 *A. muciniphila* is being explored as a commercial probiotic and/or therapeutic agent [35].
259 Recent studies have reported large scale cultivation of *A. muciniphila* on a defined medium, safe
260 for human consumption [22] and evaluated the stability and viability of the bacterium in dark
261 chocolate [47]. Our results nevertheless indicate that there are still gaps in understanding the
262 diversity and physiology of human associated Verrucomicrobia that need to be explored.
263

264 **CONCLUSION**

265 Here we carried out a pangenomic analysis of 75 *Akkermansia* genomes and identified at
266 least four species-level phylogroups (AmI-AmIV) with differing functional potentials. However,
267 a polyphasic taxonomic characterization that includes robust phenotypic and genomic analyses is
268 needed to verify species designations. Quantification of SCFA by select strains in the presence
269 and absence of vitamin B12 supplementation demonstrated cobalamin biosynthesis by AmII
270 strains. This work alters our understanding of how *Akkermansia* interacts with its human host
271 and other members of the human gut microbiome in its unique environment. Future work will
272 focus on other genomic similarities and differences identified in our analysis, but also continue
273 to explore vitamin B12 production and acquisition using our culture collection. We are also

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

274 continuing to isolate novel strains from healthy adults attempting to obtain representatives of
275 each phylogroup observed or others that have yet to be observed.
276
277 **METHODS**
278
279 METAGENOMIC STUDIES
280
281 ***Recruitment and Sampling***
282 Samples used for metagenomic sequencing were obtained from healthy children aged 2-9
283 years as described elsewhere (Herman et al., *in review*). These participants were consented under
284 protocol #1314-223 approved by the Institutional Review Board (IRB) at California State
285 University, Northridge (CSUN). Supplemental Data 2 provides unidentifiable demographic
286 information of each child included in this study.
287
288 ***DNA extraction, Library prep, sequencing***
289 Parents collected fecal samples in the privacy of their homes using sterile, double-tipped
290 swabs by swabbing toilet paper (or diapers) after use. Samples were then frozen at -20 °C within
291 24 hrs of collection, and transported on blue ice to the lab (<30 min transit), where they were
292 stored at -80 °C. This protocol is minimally invasive and has been successfully used in many
293 similar, community-based research projects [48-50].
294 DNA was extracted from approximately ~0.1g of collected samples using the MOBIO
295 Power Soil®, DNA Isolation kit following a modified extraction protocol [51]. Extracts were
296 then quantified using a Qubit 2.0 with high sensitivity reagents, and 100ng of DNA from each

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

297 sample was sheared into 300bp fragments using a Covaris M220 [52]. The NEBNext®
298 Ultra™ DNA Library Prep Kit for Illumina® [53] was used to prep dual indexed metagenomic
299 libraries from the sheared samples. Libraries were confirmed using a BIO-RAD Experion™
300 Automated Electrophoresis System, and KAPA qPCR NGS library quantification. Two
301 sequencing runs of the multiplexed libraries were conducted on an Illumina Hiseq 2000 (2 x
302 100bp) at the University of California Irvine, Genomics High-Throughput Facility.
303
304 ***Metagenomic Sequence processing***
305 Raw fastq files from each sample were trimmed using TRIMMOMATIC [54]
306 (ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15
307 MINLEN:36). Trimmed sequences were then screened against the human genome (GRCh38)
308 using DeconSeq [55] in order to remove any potential human DNA sequences. Non-human
309 sequences were further cleaned using PRINSEQ [56] with the following parameters: -
310 min_qual_mean 20 and -ns_max_n 3. Remaining sequences without a matepair were removed
311 and paired sequences were assembled using the default parameters for metagenomes in SPAdes
312 [15]. Resulting contigs > 2 Kbp were binned used MetaBAT [16] with default parameters and the
313 taxonomy and completeness of bins were verified using the taxonomy workflow of CheckM [57]
314 against the Verrucomicrobia phylum. We determined that bins confidently identified as
315 'k_Bacteria (UID2982)' were *Akkermansia* and evaluated the quality of those bins further using
316 MiGA [58]. Assembled contigs (>2kb) from each child with a high quality *Akkermansia* bin
317 were submitted to IMG-M where they were annotated using their workflow [24]. Both IMG and
318 Geneious 7.1.3 (https://www.geneious.com) were used to manually inspect annotations of
319 interest. Vitamin B12 associated genes were detected by searching for annotations from Enzyme

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

320 Commission (EC) numbers, IMG terms, pfam, and Clusters of Orthologous Groups (COG) [24,
321 59, 60]. The annotations included those used by Shelton et al. [27] and Degnan et al. [39].
322
323 ***Pangenome analysis***
324 To explore the *Akkermansia* pangenome, we combined our 35 MAGs with 40 other
325 publically available genomes in anvi'o [17, 18]. Assembled fasta files were first converted to
326 'db' files using the 'anvi-script-FASTA-to-contigs-db' command that uses Prodigal [61] to call
327 open reading frames. Each 'db' file was then annotated against the COG database [62] using
328 'anvi-run-ncbi-cogs' with the '--use-ncbi-blast' flag. After generating the genome storage file
329 with 'anvi-gen-genomes-storage,' the 'anvi-pan-genome' command was run with the identical
330 parameters (--num-threads 12, --minbit 0.5, --mcl-inflation 10, --use-ncbi-blast) outlined in
331 Delmont and Eren [17, 63, 64]. The pangenome was visualized and aesthetics were modified
332 using the 'anvi-display-pan' command. To calculate average nucleotide identity (ANI) in anvi'o,
333 the 'anvi-compute-ani' command, which utilizes PyANI [19], was used. To identify functions
334 (i.e. COG annotations) that were differentially distributed amongst the phylogroups, we used the
335 'anvi-get-enriched-functions-per-pan-group' with phylogroups (AmI-AmIV) as the category.
336
337 CULTIVATION STUDIES
338
339 ***Recruitment and Sampling***
340 Fecal samples used in culturing of *Akkermansia* isolates were obtained from healthy
341 adults using swabs as previously described [49] under IRB protocol #1516-146. Collected

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

342 samples were refrigerated (4 °C) and transferred to culture medium (see below) within 24-hours
343 of collection.
344
345 ***Enrichment, isolation, genomic DNA extraction, and 16S rRNA gene sequencing***
346 Anaerobic mucin medium was modified slightly from Derrien et al. [13] and contained (l^{-}
347 ^{1}) 0.4 g KH_2PO_4, 0.53 g Na_2HPO_4, 0.3 g NH_4Cl, 0.3 g NaCl, 0.1 g $MgCl_2 \cdot 6H_2O$, 0.4 g $NaHCO_3$,
348 1 mg resazurin, and 10 ml trace mineral solution as described by Ferguson and Mah [65]. pH of
349 the medium was adjusted to 6.5. Medium was prepared with boiled MiliQ water under constant
350 gassing with a gas mixture consisting of N_2/CO_2 (80:20, v/v). Culture medium was later modified
351 to include 1mM L-threonine and 10 g/L tryptone (Oxoid) as described previously [30]. Broth
352 medium was prepared in serum tubes or bottles and sealed with butyl rubber stoppers and
353 aluminum crimp caps prior to autoclaving at 121°C and 15 psi for 15 minutes. Prior to
354 inoculation, medium was reduced with autoclaved 0.05% $Na_2S \cdot 9H_2O$ and supplemented with
355 0.5% - 1.0% purified hog gastric mucin (Type III, Sigma-Aldrich, St. Louis, MO). Purified
356 mucin was prepared by first autoclaving a 5% or 10% solution prepared in 0.01 M phosphate
357 buffer (stock = 88.46 g/L KH_2PO_4 and 60.97 g/L K_2HPO_4), performing dialysis using a 12-14
358 kD membrane (Spectra/Por 4, Spectrum Laboratories, Rancho Dominguez, CA), centrifuging
359 twice for 10 min at 10,000 rpm, and filter sterilizing through 0.2 µm syringe filters (Whatman
360 GE Healthcare Life Sciences, Chicago, IL) into growth medium. For solid medium, Noble Agar
361 (Difco, Detroit, MI) was added and plates were poured in an anaerobic chamber (Bactron IV,
362 Sheldon Manufacturing, Inc., Cornelius, OR) under an atmosphere of $N_2/CO_2/H_2$ (80/15/5, v/v).
363 All incubations were performed at 37 °C in the Bactron IV anaerobic chamber.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

364 Enrichments cultures targeting mucin-degrading bacteria were initiated by transferring
365 fecal swabs into 5ml of anaerobic mucin medium in serum tubes and performing ten-fold serial
366 dilutions up the 10^{-7}. Cultures were incubated for up to 5 days, monitored daily for changes in
367 turbidity, and inspected using phase-contrast microscopy (Zeiss Axioskop). Positive cultures
368 with oval cells in pairs were further diluted in broth medium and/or transferred to solid medium
369 until purity could be verified microscopically and by sequencing of the 16S rRNA gene. For
370 sequencing, genomic DNA was isolated using the MoBio Ultraclean Microbial DNA Isolation
371 Kit (Mo Bio, Carlsbad, CA) following the manufacturer's instructions. Briefly, 1.8 mL of
372 overnight bacterial culture was centrifuged at 10,000 x *g* for 30 seconds, the pellet was
373 resuspended in 300 µL of MicroBead Solution (Mo Bio, Carlsbad, CA), and the DNA was
374 subsequently isolated following the manufacturer's instructions. For amplification of the 16S
375 rRNA gene via PCR, 2 µL of extracted genomic DNA was added to 25 µL of GoTaq Green
376 Master Mix (Promega, Madison, WI) and 1 µL of 10 µM universal primers 8F (5'-
377 AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-TACGGTTACCTTGTTACGA-3') using a
378 50 µL final PCR reaction volume. PCR was conducted using an Eppendorf Mastercycler Pro S
379 96 well thermocycler using a program of an initial denaturation at 95°C for 3 min, followed by
380 30 cycles of 95°C for 45 sec, annealing at 45°C for 1 min, 72°C for 1 min, a final extension of
381 72°C for 7 min, and holding at 4°C. PCR reactions were purified using the QIAquick PCR
382 Purification Kit (Qiagen). Initial sequencing of the 16S rRNA gene was performed using either
383 the 8F or 1492R primer on an ABI Prism 3730 DNA sequencer (Laragen Sequencing and
384 Genotyping in Culver City, CA). If cultures were pure and positively BLASTed to *A.*
385 *muciniphila*, the near full length 16S rRNA gene was sequenced with additional primers (515F
386 (GTGCCAGCMGCCGCGGTAA), 806R (GGACTACHVGGGTWTCTAAT), and 8F or

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

387 1492R). Sequences associated with each isolate were then assembled in Geneious 7.1.3
388 (https://www.geneious.com) and imported into ARB [28] as discussed below. General
389 demographic information about donors is provided as Supplemental Table 3.
390
391 ***16S rRNA gene phylogeny***
392 To determine phylogroup affiliation of our isolates, 16S rRNA gene sequences of the
393 Guo et al. [14] isolates were first extracted from their genomic sequence data and imported into
394 ARB [28]. Once in ARB, gene sequences were aligned with secondary structure constraints
395 against the 16S rRNA gene sequence of *A. muciniphila* Muc^T, manually inspected, and those that
396 were < 1000 bp were discarded. Similarly, 16S rRNA gene sequences of our novel isolates were
397 imported and aligned in ARB. A custom alignment mask excluding nucleotide positions found in
398 less than half of all isolates was generated and masked alignments were imported into MEGA7
399 [66] where phylogenetic reconstruction was generated using the maximum-likelihood approach.
400 Because we knew the affiliation of the Guo et al [14] isolates, we were able to place our isolates
401 in this framework based on placement in the 16S rRNA gene tree.
402
403 ***Corrin biosynthesis PCR screen of isolates and gene sequencing***
404 To amplify conserved regions of corrin biosynthesis associated genes, degenerate primers
405 were designed (Supplemental Table 1). Select corrin biosynthesis associated homologous
406 sequences were aligned using BioEdit Sequence Alignment Editor Version 7.0.5
407 (http://www.mbio.ncsu.edu/BioEdit/page2.html) (locus tags of sequences used in the
408 alignments are shown in Supplemental Table 4). All gene sequences were obtained from JGI
409 IMG/ER. Conserved regions were found using the Accessory Application ClustalW Multiple

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

410 alignment tool in BioEdit [67]. For amplification of corrin biosynthesis genes *cbiL*, *cbiC*, *cbiD*,
411 and *cbiFGH*, 1uL of genomic DNA was added to 12.5 µL of GoTaq Green Master Mix
412 (Promega, Madison, WI) and 1 µL of each 10 µM primer using a 25 µL final PCR reaction
413 volume. PCR conditions were optimized, and a PCR screen of isolates was carried out in
414 duplicates using a PCR program of an initial denaturation at 95°C for 2 min, followed by 25-35
415 cycles of 95°C for 45 sec, annealing at 52°C to 62°C for 30 sec to 1 min, 72°C for 45 sec, a final
416 extension of 72°C for 5 min, and holding at 4°C. PCR amplicons were separated and visualized
417 using 1% agarose gel. PCR products were purified using the QIAquick PCR Purification Kit
418 (Qiagen). For the amplification of *cbiFGH*, the amplicon was excised out of the gel and gel
419 purified using the PureLink Quick Gel Extraction and PCR Purification Combo Kit (Invitrogen).
420 The amplicons were sequenced as described above. BioEdit Version 7.0.5 was used to analyze
421 the sequences. Sequences of PCR amplicons from CSUN-17 were checked by blastx using the
422 IMG and the NCBI database to examine similarity to vitamin B12 associated genes from the
423 genomes *A. glycaniphila* ERS 1290231 and *Desulfovibrio vulgaris* str. Hildenborough
424 (Supplemental Table 2).
425
426 ***Quantification of short chain fatty acids via HPLC***
427 To quantify production of short-chain fatty acids with and without vitamin
428 supplementation *A. muciniphila* Muc^T (AmI) and CSUN-17 (AmII) were grown in anaerobic
429 mucin medium supplemented with 1mM L-threonine, 10 g/L tryptone (Oxoid), 1% purified
430 mucin and vitamin supplementation depending on treatment conditions. For vitamin
431 supplementation, we first performed the experiment using the ATCC MD-VS at the
432 recommended concentration (10 ml/L). Because the concentration of vitamin B12 in the

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

433 formulation is 100 fold less than those reported by Belzer et al. [29], we subsequently performed
434 a second experiment with pure vitamin B12 (Sigma-Aldrich, St. Louis, MO) using a final
435 concentration of 100 ng/ml. For all experiments, overnight cultures were transferred in
436 appropriate medium 3 times with the final transfer used to inoculate 25ml of medium at 5% in
437 quadruplicate for each isolate and treatment. OD_{600nm} (Eppendorf BioPhotometer plus) was
438 recorded at inoculation and at 12, 16, and 20 hours. An additional 1.25 ml of culture was
439 removed at each time point, centrifuged at 15,000 x g for 10 minutes, and the cell-free
440 supernatant was filtered through a 13mm, 0.2μm SPARTAN HPLC syringe filter. Samples were
441 stored at -20°C until HPLC analysis.
442 High-performance liquid chromatography (HPLC) was performed using a Waters Breeze
443 2 system (Waters Corp., Milford, MA, USA) equipped with a refractive index detector (model
444 2414). An Aminex HPX-87H column (Bio-Rad Laboratories) was used to measure production of
445 short chain fatty acids (SCFA). Sulfuric acid (5mM) was used as the mobile phase at a flow rate
446 of 0.6 mL/min. Peak areas and retention times were compared against known standards. Samples
447 were also compared against a media-only control to determine background levels of acetate,
448 propionate, and succinate present in the starting medium before growth. Approximately 3mM
449 propionate was detected in the culture medium and subtracted from all respective measurements.
450

451 **LIST OF ABBREVIATIONS**

452 Gastrointestinal tract (GIT), short chain fatty acids (SCFA), polymerase chain reaction (PCR),
453 high-performance liquid chromatography (HPLC), metagenome assembled genomes (MAGs),
454 average nucleotide identity (ANI), gene cluster (GC), Clusters of Orthologous Groups (COG),
455 Enzyme Commission (EC), Integrated Microbial Genomes (IMG)

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

456

457 **DECLARATIONS**

458 *Ethics approval and consent to participate:* Aspects of this work that included human
459 subjects was approved by the Institutional Review Board at California State University,
460 Northridge under protocol #1314-223 (metagenomic study) and #1516-146 (cultivation study).
461 For the metagenomic study, verbal assent was obtained from each child and written consent was
462 obtained from one parent/guardian. Written consent was obtained from each subject in the
463 cultivation studies.

464 *Consent for publication:* Not applicable.

465 *Availability of data and material:* Genomic sequence data from Guo et al. [14] is
466 available at GenBank under BioProject #PRJNA331216. Our quality filtered metagenomic
467 sequence data are available at GenBank under BioProject #PRJNA525290. Additionally,
468 assembled conitgs >2kb from children who contained an *Akkermansia* bin are available in IMG
469 under GOLD Study ID Gs0133482. It is important to note that contigs available in IMG include
470 not only *Akkermansia* contigs, but also all contigs from each child. Supplemental Data 3 has a
471 list of the *Akkermansia* contigs in IMG that were included in our analysis. Near full-length 16S
472 rRNA gene sequences of our isolates are available in GenBank under accession numbers
473 MK577303 – MK577312. Corrin gene sequences of isolate CSUN-17 are available in GenBank
474 under accession numbers MK585566 – MK585569.

475 *Competing interests:* The authors declare that they have no competing interests.

476 *Funding:* Research reported in this publication was supported by the National Institute Of
477 General Medical Sciences (NIGMS) of the National Institutes of Health (NIH) under Award
478 Number SC2GM122620 to GEF. NK and KG were supported under Award Numbers

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

479 TL4GM118977, RL5GM118975, and UL1GM118976 also from NIGMS. KLC and MP were
480 supported by grant R01DE024463 from the National Institute of Dental and Craniofacial
481 Research (NIDCR) of the U.S. National Institutes of Health. The content is solely the
482 responsibility of the authors and does not necessarily represent the official views of the National
483 Institutes of Health. Oak Ridge National Laboratory (ORNL) is managed by UT-Battelle, LLC
484 for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

485 *Authors' contributions:* NK conducted wet lab work, analyzed and interpreted the data,
486 and wrote the paper. KG performed bioinformatic analysis, and analyzed and interpreted the
487 data. KLC conducted wet lab work and analyzed and interpreted the data. EL collected samples,
488 conducted wet lab work, and analyzed and interpreted the data. NR conducted wet lab work and
489 analyzed and interpreted the data. MP analyzed and interpreted the data. GEF conceived of and
490 designed the study, performed bioinformatics analysis, analyzed and interpreted the data, and
491 wrote the paper. All authors read and approved the final manuscript.

492 *Acknowledgements:* The authors would like to thank Dara Fluke, Erik Hearn, Christopher
493 Herrera, Arinnae Kurdian, Jonathan Lemus, Claudia Mendoza, and Priscilla Salcedo for help in
494 the isolations. We would also like to thank Dena Herman, Joan Maltese, Zelzah Guzman, Alison
495 Gambou, Ann-Marie Pham, Griseida Ruiz, and Jessica Saavedra for all their efforts in participant
496 recruitment for the metagenomic studies. Finally, thanks to all study participants that provided
497 fecal samples, without you none of this would be possible.

498

499 **Figure Captions**

500

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

501 **Figure 1.** Pangenome of 75 Akkermansia genomes generated using anvi'o [17, 18]. Each
502 concentric circle represents a bacterial genome with purple circles belonging to AmI, blue AmII,
503 orange AmIII, and green AmIV phylogroups. Blank areas in each circle indicate absence of a
504 particular gene cluster (GC) in that genome. A total of 6,557 GCs were observed across all
505 genomes. Genomes are ordered by average nucleotide identity (ANI) as depicted by the pink
506 heatmap in the upper right section of the figure. Host organisms are indicated just below the
507 heatmap in white (human) or black (mouse) boxes. Similarly, genome sources are indicated in
508 white [12], grey [14] and black (this work) boxes. The outer most ring is colored by presence
509 (red) or absence (grey) of functional COG annotations. The next ring indicates the number of
510 genomes that particular gene cluster was observed in. Singleton (blue) and core genes (green) are
511 indicated outside of the concentric circles. Corrin ring biosynthesis genes are indicated in the
512 AmII (blue) and AmIII (orange) genomes.
513
514 **Figure 2.** (A) Corrin ring biosynthesis gene cluster from isolate *A. muciniphila* CSUN-17
515 (phylogroup AmII). Presence of genes in the corrin ring biosynthesis gene cluster in
516 phylogroups AmI, AmII, AmIII, and AmIV. Plus sign (+) above table indicates presence of gene
517 in *A. muciniphila* CSUN phylogroup AmII isolates using PCR screen of *A. muciniphila* CSUN
518 isolates. (B) Proposed strategy of propionate production in *A. muciniphila* CSUN-17
519 (phylogroup AmII) is shown involving *de novo* vitamin B12 biosynthesis leading to activation of
520 methylmalonyl-CoA synthase and conversion of succinate to propionate.
521
522 **Figure 3.** Production of acetate, succinate, and propionate through time by two strains of human
523 associated *Akkermansia* grown on purified hog gastric mucin (1%) in the absence of vitamins (A,

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

524 D), with ATCC MD-VS supplementation (B, E; 1% v/v vitamin solution/medium, vitamin B12
525 at a final concentration of 1 μg/L medium), and with vitamin B12 as cyanocobalamin (C, F; 100
526 μg/L). All values were averaged from four replicates and error bars represent the standard
527 deviation. Background levels of organic acids present in the culture medium were subtracted
528 from calculated averages when necessary.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

529 **References:**

530 1. Collado MC, Derrien M, Isolauri E, de Vos WM, Salminen S: **Intestinal integrity and**
531 **Akkermansia muciniphila, a mucin-degrading member of the intestinal**
532 **microbiota present in infants, adults, and the elderly**. *Applied and Environmental*
533 *Microbiology* 2007, **73**(23):7767-7770.
534 2. Derrien M, Collado MC, Ben-Amor K, Salminen S, de Vos WM: **The Mucin degrader**
535 **Akkermansia muciniphila is an abundant resident of the human intestinal**
536 **tract**. *Applied and Environmental Microbiology* 2008, **74**(5):1646-1648.
537 3. Karlsson CL, Onnerfalt J, Xu J, Molin G, Ahrne S, Thorngren-Jerneck K: **The**
538 **microbiota of the gut in preschool children with normal and excessive body**
539 **weight**. *Obesity (Silver Spring, Md)* 2012, **20**(11):2257-2261.
540 4. T FST, Grzeskowiak LM, Salminen S, Laitinen K, Bressan J, Gouveia Peluzio Mdo C:
541 **Faecal levels of Bifidobacterium and Clostridium coccoides but not plasma**
542 **lipopolysaccharide are inversely related to insulin and HOMA index in women**.
543 *Clinical nutrition (Edinburgh, Scotland)* 2013, **32**(6):1017-1022.
544 5. Dao MC, Everard A, Aron-Wisnewsky J, Sokolovska N, Prifti E, Verger EO, Kayser BD,
545 Levenez F, Chilloux J, Hoyles L *et al*: **Akkermansia muciniphila and improved**
546 **metabolic health during a dietary intervention in obesity: relationship with**
547 **gut microbiome richness and ecology**. *Gut* 2016, **65**(3):426-436.
548 6. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, Guiot Y, Derrien
549 M, Muccioli GG, Delzenne NM *et al*: **Cross-talk between Akkermansia muciniphila**
550 **and intestinal epithelium controls diet-induced obesity**. *Proceedings of the*
551 *National Academy of Sciences of the United States of America* 2013, **110**(22):9066-
552 9071.
553 7. Shin NR, Lee JC, Lee HY, Kim MS, Whon TW, Lee MS, Bae JW: **An increase in the**
554 **Akkermansia spp. population induced by metformin treatment improves**
555 **glucose homeostasis in diet-induced obese mice**. *Gut* 2014, **63**(5):727-735.
556 8. Hanninen A, Toivonen R, Poysti S, Belzer C, Plovier H, Ouwerkerk JP, Emani R, Cani
557 PD, De Vos WM: **Akkermansia muciniphila induces gut microbiota remodelling**
558 **and controls islet autoimmunity in NOD mice**. *Gut* 2018, **67**(8):1445-1453.
559 9. Belzer C, de Vos WM: **Microbes inside--from diversity to function: the case of**
560 **Akkermansia**. *The ISME journal* 2012, **6**(8):1449-1458.
561 10. Derrien M, Belzer C, de Vos WM: **Akkermansia muciniphila and its role in**
562 **regulating host functions**. *Microb Pathog* 2017, **106**:171-181.
563 11. Lopez-Siles M, Enrich-Capo N, Aldeguer X, Sabat-Mir M, Duncan SH, Garcia-Gil LJ,
564 Martinez-Medina M: **Alterations in the Abundance and Co-occurrence of**
565 **Akkermansia muciniphila and Faecalibacterium prausnitzii in the Colonic**
566 **Mucosa of Inflammatory Bowel Disease Subjects**. *Frontiers in cellular and*
567 *infection microbiology* 2018, **8**:281.
568 12. van Passel MW, Kant R, Zoetendal EG, Plugge CM, Derrien M, Malfatti SA, Chain PS,
569 Woyke T, Palva A, de Vos WM *et al*: **The genome of Akkermansia muciniphila, a**
570 **dedicated intestinal mucin degrader, and its use in exploring intestinal**
571 **metagenomes**. *Plos One* 2011, **6**(3):e16876.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

572 13. Derrien M, Vaughan EE, Plugge CM, de Vos WM: **Akkermansia muciniphila gen.
573 nov., sp. nov., a human intestinal mucin-degrading bacterium**. *International
574 Journal of Systematic and Evolutionary Microbiology* 2004, **54**(Pt 5):1469-1476.
575 14. Guo X, Li S, Zhang J, Wu F, Li X, Wu D, Zhang M, Ou Z, Jie Z, Yan Q *et al*: **Genome
576 sequencing of 39 Akkermansia muciniphila isolates reveals its population
577 structure, genomic and functional diverisity, and global distribution in
578 mammalian gut microbiotas**. *BMC Genomics* 2017, **18**(1):800.
579 15. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,
580 Nikolenko SI, Pham S, Prjibelski AD *et al*: **SPAdes: a new genome assembly
581 algorithm and its applications to single-cell sequencing.** *J Comput Biol* 2012,
582 **19**(5):455-477.
583 16. Kang DD, Froula J, Egan R, Wang Z: **MetaBAT, an efficient tool for accurately
584 reconstructing single genomes from complex microbial communities**. *PeerJ*
585 2015, **3**:e1165.
586 17. Delmont TO, Eren AM: **Linking pangenomes and metagenomes: the
587 Prochlorococcus metapangenome**. *PeerJ* 2018, **6**:e4320.
588 18. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO: **Anvi'o:
589 an advanced analysis and visualization platform for 'omics data**. *PeerJ* 2015,
590 **3**:e1319.
591 19. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK: **Genomics and
592 taxonomy in diagnostics for food security: soft-rotting enterobacterial plant
593 pathogens**. *Analytical Methods* 2016, **8**(1):12-24.
594 20. Ouwerkerk JP, van der Ark KCH, Davids M, Claassens NJ, Finestra TR, de Vos WM,
595 Belzer C: **Adaptation of Akkermansia muciniphila to the Oxic-Anoxic Interface
596 of the Mucus Layer**. *Appl Environ Microbiol* 2016, **82**(23):6983-6993.
597 21. Ottman N, Huuskonen L, Reunanen J, Boeren S, Klievink J, Smidt H, Belzer C, de Vos
598 WM: **Characterization of Outer Membrane Proteome of Akkermansia
599 muciniphila Reveals Sets of Novel Proteins Exposed to the Human Intestine**.
600 *Front Microbiol* 2016, **7**:1157.
601 22. Plovier H, Everard A, Druart C, Depommier C, Van Hul M, Geurts L, Chilloux J, Ottman
602 N, Duparc T, Lichtenstein L *et al*: **A purified membrane protein from
603 Akkermansia muciniphila or the pasteurized bacterium improves metabolism
604 in obese and diabetic mice.** *Nature medicine* 2017, **23**(1):107-113.
605 23. Ottman N, Reunanen J, Meijerink M, Pietila TE, Kainulainen V, Klievink J, Huuskonen
606 L, Aalvink S, Skurnik M, Boeren S *et al*: **Pili-like proteins of Akkermansia
607 muciniphila modulate host immune responses and gut barrier function**. *PLoS
608 One* 2017, **12**(3):e0173004.
609 24. Chen IA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese
610 N, White JR, Seshadri R *et al*: **IMG/M v.5.0: an integrated data management and
611 comparative analysis system for microbial genomes and microbiomes**. *Nucleic
612 Acids Res* 2019, **47**(D1):D666-D677.
613 25. Rodionov DA, Dubchak I, Arkin A, Alm E, Gelfand MS: **Reconstruction of regulatory
614 and metabolic pathways in metal-reducing delta-proteobacteria**. *Genome
615 biology* 2004, **5**(11):R90.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

616 26. Ouwerkerk JP, Aalvink S, Belzer C, de Vos WM: **Akkermansia glycaniphila sp. nov.,**
617 **an anaerobic mucin-degrading bacterium isolated from reticulated python**
618 **faeces**. *Int J Syst Evol Microbiol* 2016, **66**(11):4614-4620.
619 27. Shelton AN, Seth EC, Mok KC, Han AW, Jackson SN, Haft DR, Taga ME: **Uneven**
620 **distribution of cobamide biosynthesis and dependence in bacteria predicted**
621 **by comparative genomics**. *The ISME journal* 2018.
622 28. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T,
623 Steppi S, Jobb G *et al*: **ARB: a software environment for sequence data**. *Nucleic*
624 *Acids Res* 2004, **32**(4):1363-1371.
625 29. Belzer C, Chia LW, Aalvink S, Chamlagain B, Piironen V, Knol J, de Vos WM:
626 **Microbial Metabolic Networks at the Mucus Layer Lead to Diet-Independent**
627 **Butyrate and Vitamin B12 Production by Intestinal Symbionts**. *MBio* 2017, **8**(5).
628 30. Ottman N, Davids M, Suarez-Diez M, Boeren S, Schaap PJ, Martins Dos Santos VAP,
629 Smidt H, Belzer C, de Vos WM: **Genome-Scale Model and Omics Analysis of**
630 **Metabolic Capacities of Akkermansia muciniphila Reveal a Preferential Mucin-**
631 **Degrading Lifestyle**. *Appl Environ Microbiol* 2017, **83**(18).
632 31. Cani PD, de Vos WM: **Next-Generation Beneficial Microbes: The Case of**
633 **Akkermansia muciniphila**. *Front Microbiol* 2017, **8**:1765.
634 32. Geerlings SY, Kostopoulos I, de Vos WM, Belzer C: **Akkermansia muciniphila in**
635 **the Human Gastrointestinal Tract: When, Where, and How?** *Microorganisms*
636 2018, **6**(3).
637 33. Png CW, Linden SK, Gilshenan KS, Zoetendal EG, McSweeney CS, Sly LI, McGuckin
638 MA, Florin TH: **Mucolytic bacteria with increased prevalence in IBD mucosa**
639 **augment in vitro utilization of mucin by other bacteria**. *The American journal of*
640 *gastroenterology* 2010, **105**(11):2420-2428.
641 34. Rajilic-Stojanovic M, Shanahan F, Guarner F, de Vos WM: **Phylogenetic analysis of**
642 **dysbiosis in ulcerative colitis during remission**. *Inflammatory bowel diseases*
643 2013, **19**(3):481-488.
644 35. Naito Y, Uchiyama K, Takagi T: **A next-generation beneficial microbe:**
645 **Akkermansia muciniphila**. *Journal of clinical biochemistry and nutrition* 2018,
646 **63**(1):33-35.
647 36. Richter M, Rossello-Mora R: **Shifting the genomic gold standard for the**
648 **prokaryotic species definition**. *Proc Natl Acad Sci U S A* 2009, **106**(45):19126-
649 19131.
650 37. Kim M, Oh HS, Park SC, Chun J: **Towards a taxonomic coherence between**
651 **average nucleotide identity and 16S rRNA gene sequence similarity for species**
652 **demarcation of prokaryotes**. *Int J Syst Evol Microbiol* 2014, **64**(Pt 2):346-351.
653 38. Degnan PH, Taga ME, Goodman AL: **Vitamin B12 as a modulator of gut microbial**
654 **ecology**. *Cell Metab* 2014, **20**(5):769-778.
655 39. Degnan PH, Barry NA, Mok KC, Taga ME, Goodman AL: **Human gut microbes use**
656 **multiple transporters to distinguish vitamin B(1)(2) analogs and compete in**
657 **the gut**. *Cell host & microbe* 2014, **15**(1):47-57.
658 40. Chia LW, Hornung BVH, Aalvink S, Schaap PJ, de Vos WM, Knol J, Belzer C:
659 **Deciphering the trophic interaction between Akkermansia muciniphila and**
660 **the butyrogenic gut commensal Anaerostipes caccae using a**
661 **metatranscriptomic approach**. *Antonie Van Leeuwenhoek* 2018, **111**(6):859-873.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

662 41. Koh A, De Vadder F, Kovatcheva-Datchary P, Backhed F: **From Dietary Fiber to**
663 **Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites**. *Cell*
664 2016, **165**(6):1332-1345.
665 42. Scheppach W: **Effects of short chain fatty acids on gut morphology and function**.
666 *Gut* 1994, **35**(1 Suppl):S35-38.
667 43. Hu J, Lin S, Zheng B, Cheung PCK: **Short-chain fatty acids in control of energy**
668 **metabolism**. *Critical reviews in food science and nutrition* 2018, **58**(8):1243-1249.
669 44. Chambers ES, Viardot A, Psichas A, Morrison DJ, Murphy KG, Zac-Varghese SE,
670 MacDougall K, Preston T, Tedford C, Finlayson GS *et al*: **Effects of targeted delivery**
671 **of propionate to the human colon on appetite regulation, body weight**
672 **maintenance and adiposity in overweight adults**. *Gut* 2015, **64**(11):1744-1754.
673 45. De Vadder F, Kovatcheva-Datchary P, Zitoun C, Duchampt A, Backhed F, Mithieux G:
674 **Microbiota-Produced Succinate Improves Glucose Homeostasis via Intestinal**
675 **Gluconeogenesis**. *Cell Metab* 2016, **24**(1):151-157.
676 46. Nadjsombati MS, McGinty JW, Lyons-Cohen MR, Jaffe JB, DiPeso L, Schneider C,
677 Miller CN, Pollack JL, Nagana Gowda GA, Fontana MF *et al*: **Detection of Succinate**
678 **by Intestinal Tuft Cells Triggers a Type 2 Innate Immune Circuit**. *Immunity*
679 2018, **49**(1):33-41 e37.
680 47. Marcial-Coba MS, Saaby L, Knochel S, Nielsen DS: **Dark chocolate as a stable**
681 **carrier of microencapsulated Akkermansia muciniphila and Lactobacillus**
682 **casei**. *FEMS Microbiol Lett* 2019, **366**(2).
683 48. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R: **Bacterial**
684 **community variation in human body habitats across space and time**. *Science*
685 2009, **326**(5960):1694-1697.
686 49. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW,
687 Vazquez-Baeza Y, Gonzalez A, Knight R *et al*: **Temporal variability is a**
688 **personalized feature of the human microbiome**. *Genome biology* 2014,
689 **15**(12):531.
690 50. Song SJ, Lauber C, Costello EK, Lozupone CA, Humphrey G, Berg-Lyons D, Caporaso
691 JG, Knights D, Clemente JC, Nakielny S *et al*: **Cohabiting family members share**
692 **microbiota with one another and with their dogs**. *eLife* 2013, **2**:e00458.
693 51. Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, Knights
694 D, Gajer P, Ravel J, Fierer N *et al*: **Moving pictures of the human microbiome**.
695 *Genome biology* 2011, **12**(5):R50.
696 52. Head SR, Komori HK, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR,
697 Ordoukhanian P: **Library construction for next-generation sequencing:**
698 **overviews and challenges**. *BioTechniques* 2014, **56**(2):61-64, 66, 68, passim.
699 53. Luck AN, Anderson KG, McClung CM, VerBerkmoes NC, Foster JM, Michalski ML,
700 Slatko BE: **Tissue-specific transcriptomics and proteomics of a filarial**
701 **nematode and its Wolbachia endosymbiont**. *BMC genomics* 2015, **16**:920.
702 54. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina**
703 **sequence data**. *Bioinformatics* 2014, **30**(15):2114-2120.
704 55. Schmieder R, Edwards R: **Fast identification and removal of sequence**
705 **contamination from genomic and metagenomic datasets**. *PloS one* 2011,
706 **6**(3):e17288.

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

707 56. Schmieder R, Edwards R: **Quality control and preprocessing of metagenomic**
708 **datasets**. *Bioinformatics* 2011, **27**(6):863-864.
709 57. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW: **CheckM: assessing**
710 **the quality of microbial genomes recovered from isolates, single cells, and**
711 **metagenomes**. *Genome research* 2015, **25**(7):1043-1055.
712 58. Rodriguez RL, Gunturu S, Harvey WT, Rossello-Mora R, Tiedje JM, Cole JR,
713 Konstantinidis KT: **The Microbial Genomes Atlas (MiGA) webserver: taxonomic**
714 **and gene diversity analysis of Archaea and Bacteria at the whole genome level**.
715 *Nucleic Acids Res* 2018, **46**(W1):W282-W288.
716 59. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob
717 B, Huang J, Williams P *et al*: **IMG: the Integrated Microbial Genomes database**
718 **and comparative analysis system**. *Nucleic acids research* 2012, **40**(Database
719 issue):D115-122.
720 60. Cornish-Bowden A: **Current IUBMB recommendations on enzyme nomenclature**
721 **and kinetics**. *Perspectives in Science* 2014, **1**(1):74-87.
722 61. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: **Prodigal:**
723 **prokaryotic gene recognition and translation initiation site identification**. *BMC*
724 *Bioinformatics* 2010, **11**:119.
725 62. Galperin MY, Makarova KS, Wolf YI, Koonin EV: **Expanded microbial genome**
726 **coverage and improved protein family annotation in the COG database**. *Nucleic*
727 *Acids Res* 2015, **43**(Database issue):D261-269.
728 63. Benedict MN, Henriksen JR, Metcalf WW, Whitaker RJ, Price ND: **ITEP: an**
729 **integrated toolkit for exploration of microbial pan-genomes**. *BMC Genomics*
730 2014, **15**:8.
731 64. van Dongen S, Abreu-Goodger C: **Using MCL to extract clusters from networks**.
732 *Methods Mol Biol* 2012, **804**:281-295.
733 65. Ferguson TJ, Mah RA: **Isolation and characterization of an h(2)-oxidizing**
734 **thermophilic methanogen**. *Appl Environ Microbiol* 1983, **45**(1):265-274.
735 66. Kumar S, Stecher G, Tamura K: **MEGA7: Molecular Evolutionary Genetics**
736 **Analysis Version 7.0 for Bigger Datasets**. *Mol Biol Evol* 2016, **33**(7):1870-1874.
737 67. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of**
738 **progressive multiple sequence alignment through sequence weighting,**
739 **position-specific gap penalties and weight matrix choice**. *Nucleic acids research*
740 1994, **22**(22):4673-4680.
741

742

743 **Table 1.** Summary of 35 *Akkermansia* metagenome assembled genomes (MAGs) recovered from a diverse population of children

744 aged 2-9 years living in Los Angeles, CA, USA. The genome of *A. muciniphila* Muc^T was not included in the averages presented at

745 the bottom of the table

		Genome Properties				Assembly Properties		
Genome name	**Phylogroup**	**Completeness (%)**	**Contamination (%)**	**Predicted Proteins**	**Coding Density (%)**	**Contigs**	**Length (Mbp)**	**GC Content (%)**
A. muciniphila Muc^T	AmI	100	0.0	2,238	88.6	1	2.66	55.8
CDI-75C-7	AmI	95.5	0.0	2,433	88.4	26	2.82	55.5
CDI-92A-19	AmI	68.5	0.0	2,117	89.0	317	2.27	55.8
CDI-93C-15	AmI	91.0	0.0	2,345	88.8	231	2.59	55.8
CDI-16B-22	AmI	94.6	0.0	2,291	88.5	79	2.65	55.6
CDI-158B-12	AmI	95.5	0.0	2,302	88.2	42	2.70	55.3
CDI-50B-13	AmI	95.5	0.0	2,293	88.4	22	2.72	55.6
CDI-51A-11	AmI	95.5	0.0	2,295	88.4	22	2.72	55.6
CDI-28A-8	AmI	95.5	0.0	2,301	88.2	27	2.71	55.4
CDI-85A-12	AmI	95.5	0.0	2,225	88.4	19	2.66	55.4
CDI-30A-11	AmI	95.5	0.0	2,272	88.4	22	2.68	55.5
CDI-42C-15	AmI	95.5	0.0	2,340	88.4	25	2.74	55.2
CDI-151B-10	AmI	94.6	0.0	2,383	88.3	64	2.77	55.4
CDI-193A-6	AmI	95.5	0.0	2,416	88.3	32	2.83	55.4
CDI-143C-7	AmII	81.1	0.9	2,301	88.4	206	2.67	58.7
CDI-10B-12	AmII	94.6	0.0	2,439	88.1	32	2.98	58.3
CDI-128B-11	AmII	92.8	0.0	2,428	88.1	22	2.96	58.3
CDI-129B-12	AmII	88.3	0.0	2,375	88.1	229	2.70	58.5
CDI-77C-9	AmII	95.5	0.0	2,435	88.0	25	2.99	58.2
CDI-24B-9	AmII	94.6	0.0	2,450	88.2	31	3.02	58.1
CDI-182B-6	AmII	95.5	0.0	2,478	88.1	22	3.02	58.3
CDI-198C-9	AmII	95.5	0.0	2,512	88.3	51	3.00	58.2
CDI-69C-9	AmII	95.5	0.0	2,421	88.2	24	2.96	58.2
CDI-138A-11	AmII	95.5	0.0	2,483	88.0	25	3.01	58.2
CDI-70C-8	AmII	95.5	0.0	2,481	88.0	32	3.01	58.2
CDI-26A-8	AmII	92.8	0.9	2,610	88.1	251	2.95	58.0
CDI-34A-8	AmII	95.5	0.0	2,558	87.2	29	3.09	57.8
CDI-65B-6	AmII	87.4	0.0	2,538	87.4	55	3.04	57.8
CDI-203B-7	AmII	94.6	0.9	2,479	87.8	25	2.99	58.1
CDI-150B-9	AmIV	95.5	0.0	2,457	87.7	29	2.99	57.2
CDI-12C-16	AmIV	95.5	0.0	2,461	87.8	32	2.99	57.2
CDI-156A-7	AmIV	95.5	0.9	2,532	87.5	56	3.05	56.7
CDI-74B-7	AmIV	94.6	0.9	2,502	87.4	48	3.01	56.7

CDI-18B-8	AmIV	94.6	0.0	2,509	88.0	124	2.95	56.9
CDI-148A-8	AmIV	95.5	0.9	2,557	87.3	46	3.04	56.0
CDI-13A-11	AmIV	95.5	0.0	2,670	88.1	66	3.20	56.6
AVERAGE	**-**	**93.4**	**0.2**	**2,419.7**	**88.1**	**68.23**	**2.87**	**56.9%**

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

746 **Table 2.** Presence of select corrin ring biosynthesis associated genes in CSUN *Akkermansia*

747 isolates as determined by PCR.

Isolate	**Phylogroup**	***cbiL***	***cbiC***	***cbiD***	***cbiFGH***
CSUN-7	AmI	-	-	-	-
CSUN-12	AmI	-	-	-	-
CSUN-17	AmII	+	+	+	+
CSUN-23	AmI	-	-	-	-
CSUN-27	AmI	-	-	-	-
CSUN-28	AmI	-	-	-	-
CSUN-33	AmI	-	-	-	-
CSUN-34	AmII	+	+	+	+
CSUN-31	AmI	-	-	-	-
CSUN-36	AmI	-	-	-	-
A. muciniphila ATCC BAA-835	AmI	-	-	-	-
A. glycaniphila ERS 1290231	NA	+	+	+	+

748 (+) indicates PCR product of predicted amplicon size.
749

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

750 **Supplemental Table 1.** Corrin ring biosynthesis degenerate primers.

Gene	Gene length (bp)	Primer name	Primer 5'-3'	Expected amplicon size (bp)
cbiL	708	precorrin-2 Forward	TYTTCAGCATGTCSCGYGAC	358
cbiL	708	precorrin-2 Reverse	GCGGCTRCGGTAGGTYTT	358
cbiC	663	cbiC Forward	ATCCACACCACGGCRGAC	500
cbiC	663	cbiC Reverse	GGCGTGCAGGGTRGT	500
cbiFGH	2619	cbiG Forward	GTSAGCAGCGTYTTYG	340
cbiFGH	2619	cbiG Reverse	ATGAGSGCCTGCCKKCCGA	340
cbiD	1119	cbiD Forward	GACCCSGACTGCACSCA	379
cbiD	1119	cbiD Reverse	TAGGCTTCRTGGCTG	379

751 Y=C or T, S=C or G, B=G or T or C, R=G or A

752

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

753 **Supplemental Table 2.** Corrin ring biosynthesis sequenced amplicons from CSUN-17 homology
754 to *A. glycaniphila* and *Desulfovibrio vulgaris.*

Gene name	*A. glycaniphila* gene name[a]	Percent identity to *A. glycaniphila*[b]	*Desulfovibrio vulgaris* gene name[a]	Percent identity to *Desulfovibrio vulgaris*[b]
cbiL	precorrin-2 C20-methyltransferase /cobalt-factor II C20-methyltransferase	65%	precorrin-2 C20-methyltransferase	48%
cbiC	precorrin-8X methylmutase	80%	precorrin-8X methylmutase	49%
cbiD	cobalt-precorrin-5B (C1)-methyltransferase	70%	cobalamin biosynthesis protein CbiD	46%
cbiFGH	cobalt-precorrin 5A hydrolase/precorrin-3B C17-methyltransferase	56%	precorrin-4 C11-methyltransferase	32%

755 [a]IMG/ER gene names using BLASTx against *A. glycaniphila* ERS 1290231 and *Desulfovibrio vulgaris*
756 (Hildenborough)
757 [b]Percent identity of CSUN-17 sequence using BLASTx against *A. glycaniphila* ERS 1290231 and *Desulfovibrio*
758 *vulgaris* (Hildenborough)
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

781 **Supplemental Table 3.** Demographic information of individuals who provided fecal samples

782 from which strains of *Akkermansia* were isolated.

Isolate	Phylogroup	Age	Gender	Ethnicity	Diet
CSUN-7	AmI	63 years	Male	Caucasian	Omnivore
CSUN-12	AmI	22 years	Male	Hispanic	Omnivore
CSUN-17	AmII	32 years	Male	Caucasian	Omnivore
CSUN-23	AmI	27 years	Male	Caucasian	Omnivore
CSUN-27	AmI	28 years	Male	Caucasian	Omnivore
CSUN-28	AmI	33 years	Male	Hispanic	Omnivore
CSUN-31	AmI	23 years	Female	Hispanic	Vegan
CSUN-33	AmI	22 years	Male	Hispanic	Vegetarian
CSUN-34	AmII	22 years	Male	Hispanic	Omnivore
CSUN-36	AmI	39 years	Female	Caucasian	Omnivore

783

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

784 **Supplemental Table 4:** Locus tags of sequences used for corrin ring biosynthesis degenerate

785 primer design.

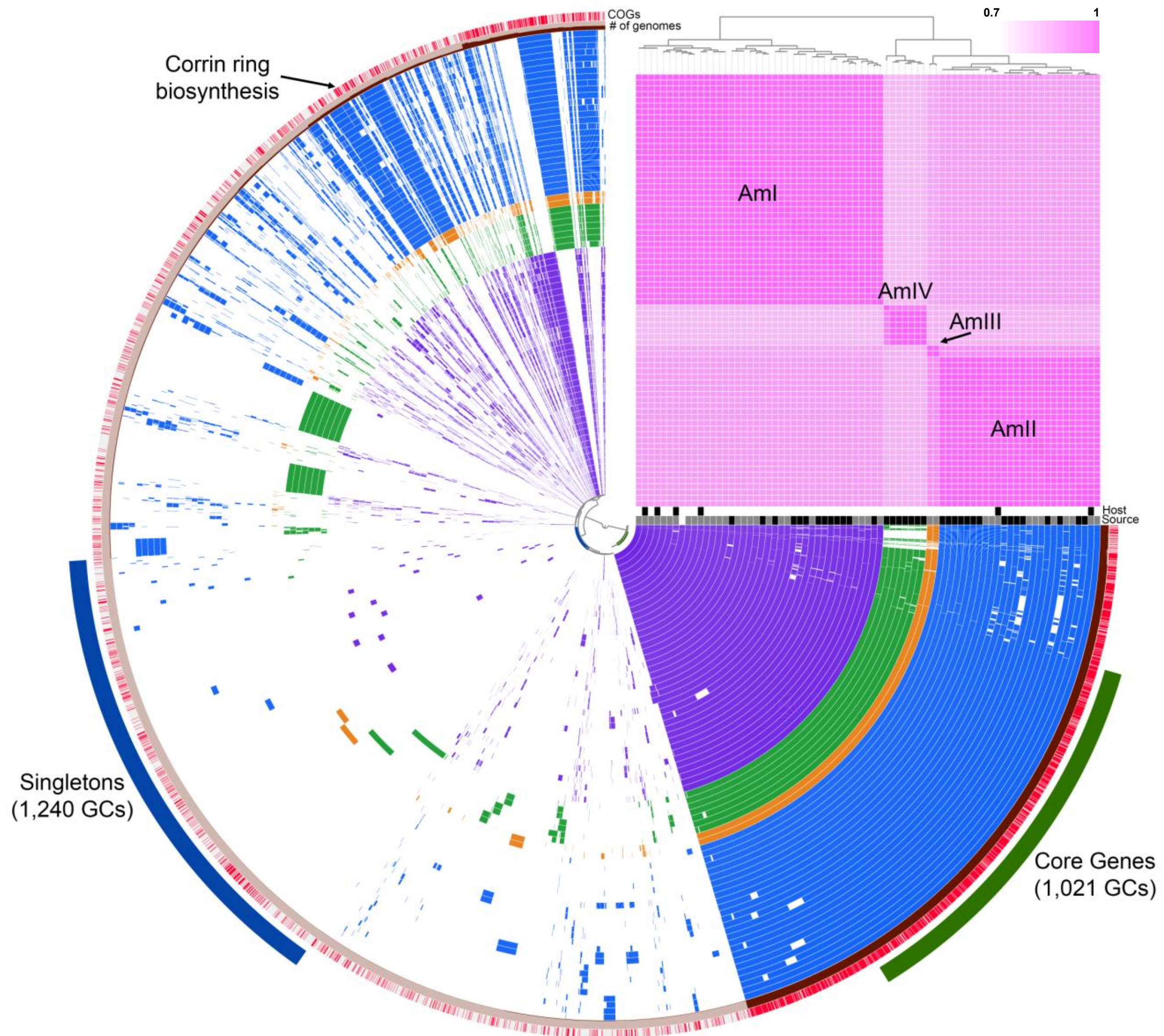
Gene	Locus tag
cbiL	Ga0175004_112520 T370DRAFT_00397 Ga0256628_10023618 Ga0257029_1018451 Ga0257030_10082613 Ga0257062_100012256 Ga0257032_118796 Ga0257040_100005132 Ga0257041_10019915 Ga0257042_100005192 Ga0257043_10000156 Ga0257044_106873 Ga0257047_1015355 Ga0257051_1033411 Ga0257052_10000756 Ga0257053_10015516 Ga0257060_10006251
cbiC	Ga0175004_112518 DESPIG_00351 Ga0256628_10023616 Ga0257029_1018453 Ga0257030_10082611 Ga0257062_100012254 Ga0257032_118794 Ga0257040_100005134 Ga0257041_10019913 Ga0257042_100005190 Ga0257043_10000158 Ga0257044_106871 Ga0257047_1015357 Ga0257051_1033413 Ga0257052_10000758 Ga0257053_10015514 Ga0257060_10006249
cbiG	Ga0175004_112515 HMPREF0179_00024 Ga0256628_10023613 Ga0257029_1018456 Ga0257030_1008268

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

	Ga0257062_100012251 Ga0257032_118791 Ga0257040_100005137 Ga0257041_10019910 Ga0257042_100005187 Ga0257043_10000161 Ga0257044_108973 Ga0257047_1015360 Ga0257051_1033416 Ga0257052_10000761 Ga0257053_10015511 Ga0257060_10006246
cbiD	Ga0175004_112517 HMPREF0178_02483 Ga0256628_10023615 Ga0257029_1018454 Ga0257030_10082610 Ga0257062_100012253 Ga0257032_118793 Ga0257040_100005135 Ga0257041_10019912 Ga0257042_100005189 Ga0257043_10000159 Ga0257044_108971 Ga0257047_1015358 Ga0257051_1033414 Ga0257052_10000759 Ga0257053_10015513 Ga0257060_10006248

786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



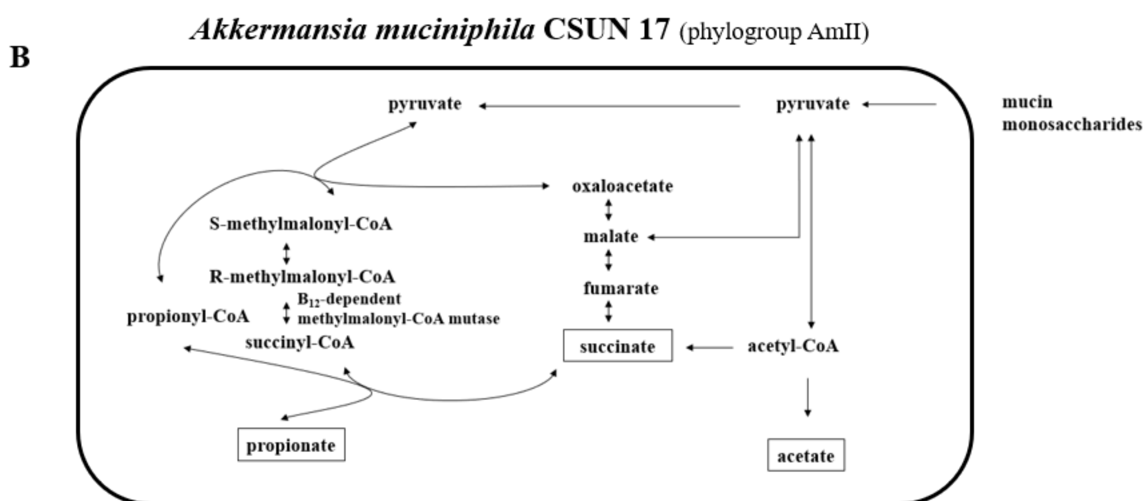
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A

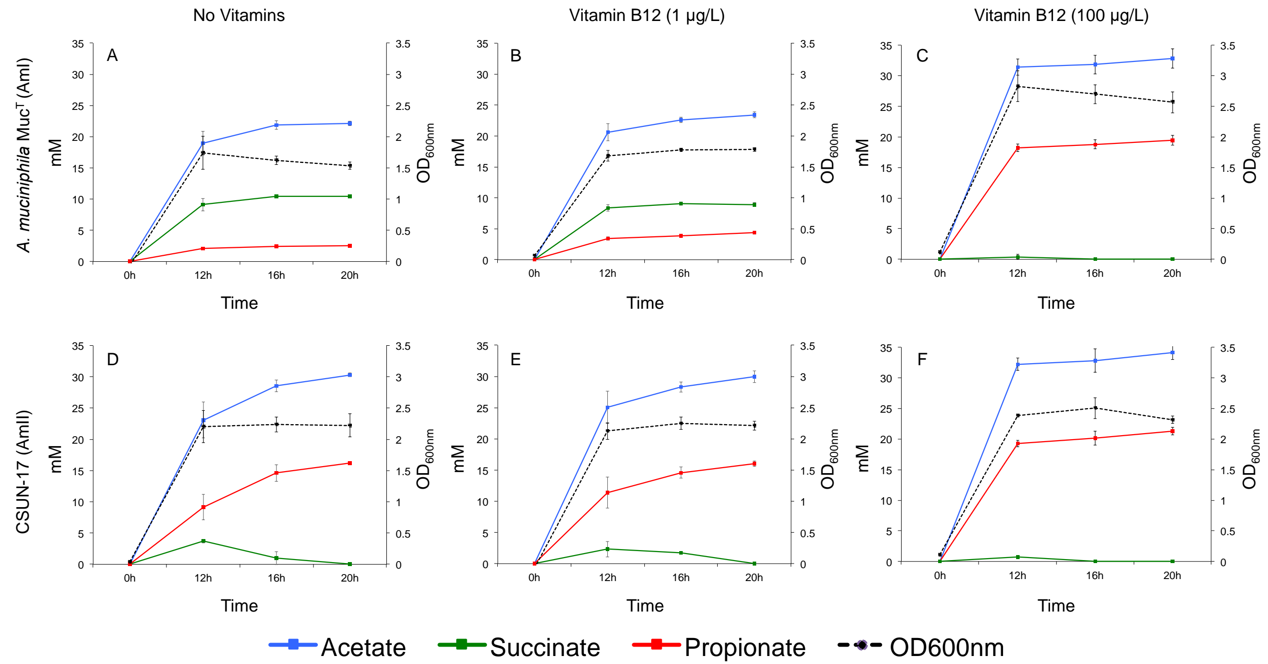
	cbiK/X	*cbiL*	*hypothetical*	*cbiC*	*cbiD*	*cbiET*	*cbiFGH*	*cbiA*	*cobA*
		+		+	+		+		
AmI	0/40	0/40	0/40	0/40	0/40	0/40	0/40	0/40	0/40
AmII	24/26	24/26	24/26	24/26	24/26	24/26	24/26	24/26	24/26
AmIII	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2	2/2
AmIV	0/7	0/7	0/7	0/7	0/7	0/7	0/7	0/7	0/7

B



818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842

bioRxiv preprint doi: https://doi.org/10.1101/587527; this version posted March 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



843