

Manuscript Number:	GIGA-D-17-00073	
Full Title:	The gut microbiome in Crohn's disease and modulation by exclusive enteral nutrition	
Article Type:	Research	
Funding Information:	National Natural Science Foundation of China (Nos 81470795)	Ms Qing He
	the Shenzhen Municipal Government of China (JCYJ20160229172757249)	Dr Huijue Jia
	the Shenzhen Municipal Government of China (DRC-SZ[2015]162)	Dr Huijue Jia
	the Shenzhen Municipal Government of China (CXB201108250098A)	Dr Huijue Jia
	the Shenzhen Municipal Government of China (JCYJ20140418095735538)	Mr Liang Xiao
	the Shenzhen Municipal Government of China (CXZZ20150330171521403)	Mr Liang Xiao
	the Shenzhen Municipal Government of China (JSGG20140702161403250)	Not applicable
	the Shenzhen Municipal Government of China (JSGG20160229172752028)	Not applicable
Abstract:	<p>Abstract</p> <p>Background: Crohn's disease (CD), an autoimmune disease, has become a health challenge worldwide. The gut microbiota closely interacts with the host immune system, but its functional impact in CD is unclear. Except for studies on a small number of CD patients, analyses of the gut microbiota in CD have used 16S rDNA amplicon sequencing. Here we used metagenomic shotgun sequencing to characterize compositional and functional features of the CD microbiota, comprising also unannotated bacteria, and investigated its modulation by exclusive enteral nutrition (EEN).</p> <p>Results: Based on signature taxa, CD microbiotas clustered into two distinct metacommunities, indicating individual variability in CD microbiome structure. Metacommunity-specific functional shifts in CD showed enrichment in producers of the pro-inflammatory hexa-acylated lipopolysaccharide variant and a reduction in the potential to synthesize short chain fatty acids. Disruption of ecological networks was evident in CD, coupled with reduction in growth rates of many bacterial species. Short-term EEN elicited limited impact on the overall composition of the CD microbiota, although functional changes occurred following treatment.</p> <p>Conclusions: The microbiota is massively perturbed in CD and it has functional implications in CD pathogenesis.</p>	
Corresponding Author:	Huijue Jia, postdoctor CHINA	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:		
Corresponding Author's Secondary Institution:		
First Author:	Qing He	

First Author Secondary Information:	
Order of Authors:	Qing He
	Yuan Gao
	Zhuye Jie
	Xinlei Yu, PhD
	Janne Marie Laursen, M.D.
	Liang Xiao, PhD
	Ying Li
	Lingling Li
	Faming Zhang, PhD
	Qiang Feng, PhD
	Xiaoping Li
	Jinghong Yu
	Chuan Liu
	Ping Lan, PhD
	Ting Yan
	Xin Liu, PhD
	Xun Xu, PhD
	Huanming Yang, postdoctor
	Jian wang, PhD
	Lise Madsen, PhD
Susanne Brix Pedersen, PhD	
Jianping Wang, PhD	
Karsten Kristiansen, postdoctor	
Huijue Jia, postdoctor	
Order of Authors Secondary Information:	
Opposed Reviewers:	Dusko S Ehrlich King's College London / Institute National de la Recherche Agronomique (INRA) difference of opinion
	Dan R Littman New York University School of Medicine conflict of interest
	Josh U. Scher New York University School of Medicine and Hospital for Joint Diseases conflict of interest
	Jun Yu Chinese University of Hong Kong conflict of interest
Additional Information:	
Question	Response

<p>Are you submitting this manuscript to a special series or article collection?</p>	<p>Yes</p>
<p>Please select an option from the menu: as follow-up to "Are you submitting this manuscript to a special series or article collection?"</p>	<p>Functional Metagenomics</p>
<p>Experimental design and statistics</p> <p>Full details of the experimental design and statistical methods used should be given in the Methods section, as detailed in our Minimum Standards Reporting Checklist. Information essential to interpreting the data presented should be made available in the figure legends.</p> <p>Have you included all the information requested in your manuscript?</p>	<p>Yes</p>
<p>Resources</p> <p>A description of all resources used, including antibodies, cell lines, animals and software tools, with enough information to allow them to be uniquely identified, should be included in the Methods section. Authors are strongly encouraged to cite Research Resource Identifiers (RRIDs) for antibodies, model organisms and tools, where possible.</p> <p>Have you included the information requested as detailed in our Minimum Standards Reporting Checklist?</p>	<p>Yes</p>
<p>Availability of data and materials</p> <p>All datasets and code on which the conclusions of the paper rely must be either included in your submission or deposited in publicly available repositories (where available and ethically appropriate), referencing such data using a unique identifier in the references and in the "Availability of Data and Materials" section of your manuscript.</p> <p>Have you have met the above requirement as detailed in our Minimum Standards Reporting Checklist?</p>	<p>Yes</p>

Two distinct metacommunities characterize the gut microbiota in

Crohn's disease patients

Qing He^{1,2,3†}, Yuan Gao^{4,5†}, Zhuye Jie^{4,5†}, Xinlei Yu^{4,5}, Janne Marie Laursen⁶, Liang Xiao^{4,5}, Ying Li¹, Lingling Li², Faming Zhang⁷, Qiang Feng^{4,8}, Xiaoping Li^{4,5}, Jinghong Yu^{4,5}, Chuan Liu^{4,5}, Ping Lan^{1,3}, Ting Yan², Xin Liu^{4,5}, Xun Xu^{4,5}, Huanming Yang^{4,9}, Jian Wang^{4,9}, Lise Madsen^{4,10,11}, Susanne Brix⁶, Jianping Wang^{1,3*}, Karsten Kristiansen^{4,10*}, Huijue Jia^{4,5,12*}

¹Department of Gastroenterology, The Sixth Affiliated Hospital of The Sun Yat-sen University, Guangzhou 510610, China

²Department of nutrition, The Sixth Affiliated Hospital of Sun Yat-sen University, Guangzhou 510610, China

³Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, the Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou 510610, China

⁴BGI-Shenzhen, Shenzhen 518083, China

⁵China National Genebank-Shenzhen, BGI-Shenzhen, Shenzhen 518083, China

⁶Department of Biotechnology and Biomedicine, Technical University of Denmark (DTU), Kongens Lyngby, Denmark.

⁷Digestive Endoscopy and Medical Center for Digestive Diseases, the Second Affiliated Hospital

1 19 of Nanjing Medical University, Nanjing 210011, Jiangsu Province, China

2
3
4 20 ⁸Shenzhen Engineering Laboratory of Detection and Intervention of Human Intestinal Microbiome,

5
6
7 21 BGI-Shenzhen, Shenzhen 518083, China

8
9
10
11 22 ⁹James D. Watson Institute of Genome Sciences, Hangzhou 310058, China

12
13
14 23 ¹⁰Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of

15
16
17 24 Copenhagen, Universitetsparken 13, 2100 Copenhagen, Denmark.

18
19
20
21 25 ¹¹National Institute of Nutrition and Seafood Research, Bergen, Norway.

22
23
24 26 ¹²Shenzhen Key Laboratory of Human Commensal Microorganisms and Health Research,

25
26
27 27 BGI-Shenzhen, Shenzhen 518083, China

28
29
30
31 28

32
33
34
35 29 † Contributed equally

36
37
38
39 30 * To whom correspondence should be addressed: K.K. (kk@bio.ku.dk) or H.J.

40
41
42 31 (jiahuijue@genomics.cn)

52 **Keywords:** Crohn's disease, Gut microbe, Metagenomics, Exclusive enteral nutrition

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 74 by surgery if symptoms cannot be improved pharmaceutically ⁹. However, side effects and
2
3 75 complications such as infection and malnutrition accompany these treatments ¹⁰, which
4
5
6 76 imperil the patient's life. Although not widely used, exclusive enteral nutrition (EEN) is a
7
8
9 77 low-risk, non-invasive therapy for CD that involves exclusive ingestion of 100% liquid
10
11
12 78 formula made up of either elemental or polymeric nutrients ¹¹. In pediatric CD up to 85%
13
14
15 79 remission has been achieved by EEN ¹¹. Nevertheless, in adult CD, EEN has not delivered
16
17
18 80 desirable effectiveness, which to some extent may be attributed to non-adherence and
19
20
21 81 interpersonal variations in clinical conditions ¹¹. The mechanism underlying the alleviation of
22
23
24 82 CD by EEN also remains unclear, though nutritional improvement and microbial involvement
25
26
27 83 possibly play a role ¹².

28
29 84 Through metagenomic sequencing and data analysis, we herein provide novel insights into the
30
31
32 85 CD microbiota at both compositional and inferred functional levels. We identified two
33
34
35 86 metacommunity stages within CD patients that differed by abundance of gram-negative
36
37
38 87 pro-inflammatory bacteria and presence of genes involved in production of anti-inflammatory
39
40
41 88 short-chain fatty acids. In addition, we investigated the effect of short-term EEN on the CD
42
43
44 89 microbiota. Our study highlights the presence of two microbiota severity-states related to gut
45
46
47 90 microbiota dysbiosis in CD and indicates possible functional links between the microbiota
48
49
50 91 and the underlying immunological dysbalance in CD.

1 92 **Data Description**

2
3
4 93 49 CD patients and 54 healthy controls (CTs) were enrolled in this study. 14 CD patients
5
6
7 94 underwent EEN treatment (for the clinical profiles of CD patients, see Supplementary Table
8
9
10 95 1). Fecal samples were collected from all participants at baseline and from the EEN-treated
11
12
13 96 patients after two-week EEN treatment, totaling 117 fecal samples. After DNA extraction,
14
15
16 97 DNA library of an insert size of 350bp was constructed and then sequenced on an Illumina
17
18 98 HiSeq 2000 analyzer at BGI (Shenzhen, China) using 100bp paired-end (PE) sequencing. In
19
20
21 99 total, we generated ~700Gb raw data, and 672Gb of them remained after filtering out
22
23
24 100 low-quality or host reads. The dataset is available from the EBI Database. On average ~55.65
25
26 101 million high-quality reads per sample were generated for further analyses. The proportion of
27
28
29 102 high-quality reads among all raw reads from each sample was 95.98% on average. Using both
30
31
32 103 de novo assembly and alignment against the integrated gene catalog (IGC) geneset, 2036584
33
34
35 104 genes with occurrence rate over 5% were obtained.

36
37
38 105 **Analyses**

39
40
41
42 106 **Clustering of CD microbiota into distinct metacommunities**

43
44
45
46 107 When the gut microbiotas of CD patients were compared to their non-CD counterparts, both
47
48
49 108 microbial gene counts (**Supplementary Fig. 1a**) and diversity (**Supplementary Fig. 1b**) were
50
51
52 109 considerably lower in CD patients than in CTs. For high-confidence taxonomic identification,
53
54
55 110 co-abundant genes were binned into metagenomics species (MGS) ¹³ (harboring more than
56
57 111 700 genes) which were thereafter used for taxonomic annotation. A total of 452 MGS_s were
58
59
60
61
62
63
64
65

1 112 identified, with 151 of them being assigned to existing taxonomic entities (**Supplementary**
2
3 113 **Table 2**).

4
5
6
7 114 To capture the principal differences between non-CD and CD microbiome structures, we
8
9
10 115 adopted a combinatory approach which started with sample clustering based on the dirichlet
11
12 116 multinomial mixtures (DMM) model ¹⁴, followed by the identification of discriminative
13
14
15 117 microbes using an adapted version of the linear discriminant analysis (LDA) effect size
16
17
18 118 (LEfSe) method ¹⁵. Based on Laplace approximation ¹⁴, we identified 3 clusters to exhibit
19
20
21 119 minimal negative log posterior (**Supplementary Fig. 1c**). Based on this we clustered the
22
23
24 120 microbiome samples of CD and CTs into 3 metacommunities (A, B and C), which displayed
25
26
27 121 intra-community homogeneity and inter-community dissimilarity (**Fig. 1a**). The membership
28
29
30 122 of a metacommunity was associated with disease status (Fisher's exact test with BH
31
32 123 adjustment, $q < 0.01$, **Supplementary Table 3**). Metacommunity A was dominated by CT
33
34
35 124 samples and metacommunity C exclusively by CD samples, whereas metacommunity B
36
37
38 125 contained both CT and CD samples (**Fig. 1a**). Based on a less stringent LEfSe method, 85
39
40
41 126 MGS were identified as discriminative microbes for the metacommunities or sub-groups (CT
42
43 127 and CD groups within metacommunity B) (**Fig. 1a** and **Supplementary Table 4**). The
44
45
46 128 majority of metacommunity A-enriched MGSs were reduced in metacommunity B and further
47
48
49 129 depleted in C, including short-chain fatty acid (SCFA)-producing bacteria such as
50
51
52 130 *Bifidobacterium* species, *Faecalibacterium prausnitzii*, *Alistipes shahii* and *Roseburia* species
53
54 131 (**Fig. 1a** and **Supplementary Table 4**). Among others, SCFA-producing bacteria *Bacteroides*
55
56
57 132 *cellulosilyticus*, *Bacteroides xylanisolvens*, and *Clostridium nexile*, a member of the
58
59
60
61
62
63
64
65

1 133 immunomodulatory Clostridium cluster XIVa¹⁶, were enriched in metacommunity B (**Fig. 1a**
2
3 134 and **Supplementary Table 4**). Another Clostridium cluster XIVa clade member, *Clostridium*
4
5
6 135 *symposium*, and a number of opportunistic pathogens such as *E. coli*, *Klebsiella pneumoniae*,
7
8
9 136 *Streptococcus salivarius*, and *Clostridium bolteae* were overrepresented in metacommunity C
10
11 137 (**Fig. 1a** and **Supplementary Table 4**), suggesting that subjects in this group had impaired
12
13
14 138 ability to suppress colonization by pathogenic species in their gut. We also evaluated whether
15
16
17 139 metacommunities differed in the degree of dysbiosis associated with CD through computing
18
19
20 140 the Microbial Dysbiosis index (MD-index)⁵. CD microbiotas from metacommunity C had
21
22
23 141 significantly higher values of the MD-index than those from metacommunity B ($p = 7.63e-05$,
24
25 142 **Fig. 1a** and **Supplementary Table 1**), suggesting a more severe degree of dysbiosis in this
26
27
28 143 CD subgroup. Combined, these compositionally distinct metacommunities recapitulate
29
30
31 144 disparate configurations of the microbiota under normal and CD conditions.
32
33
34
35 145 The separation of microbiomes into metacommunities was confirmed by principal coordinate
36
37 146 analysis (PCoA), which clustered samples by both metacommunity identity and disease status
38
39
40 147 (**Fig. 1b**). We determined whether the variations in microbiome composition were
41
42
43 148 accompanied with clinical phenotypes. In CD patients, 23 clinical variables together with age
44
45
46 149 correlated with microbiome variation, with uric acid (UA) and [blood leukocyte numbers](#)
47
48
49 150 being the top two covariates (effect size > 0.2) (**Supplementary Fig. 2b**). When categorized
50
51 151 into groups, various plasma biomarkers, including inflammatory markers were the strongest
52
53
54 152 classes of covariates (effect size > 0.2) (**Supplementary Fig. 2c**). However, despite the
55
56
57 153 existence of microbiome variations and their correlation with clinical states, no significant
58
59
60
61
62
63
64
65

1 154 differences were detected for these clinical variables between metacommunity B and C CD
2
3 155 patients (**Supplementary Fig. 2d**).

4
5
6
7 156

8
9
10
11 157 **CD- and metacommunity-associated functional traits**

12
13
14 158 We next analyzed the functional changes associated with disease status and differences in
15
16
17 159 microbiome structure. We made pair-wise comparisons after performing functional annotation
18
19
20 160 using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. A large number of
21
22
23 161 CD- and metacommunity-related functional shifts were identified at the level of pathways and
24
25
26 162 modules (**Fig. 2a, Supplementary Table 5 and Supplementary Table 6**). We observed
27
28
29 163 consistent changes in CD microbiotas in all within- or between- metacommunity comparisons
30
31 164 (in B-CD vs A-CT, C-CD vs A-CT, B-CD vs B-CT, and C-CD vs B-CT) (**Fig. 2a**). The
32
33
34 165 composition of the microbiota of CD patients indicated consistent changes in the potential for
35
36
37 166 carbohydrate utilization compared to the CT counterparts, with a decreased abundance of
38
39
40 167 pathways involved in starch and sucrose metabolism, and enrichment of pathways involved in
41
42
43 168 simple carbon metabolism such as fructose, mannose, and galactose in the microbiota of CD
44
45
46 169 patients (**Fig. 2a**). In addition, we observed an enrichment of genes in pathways involved in
47
48
49 170 glyoxylate, dicarboxylate, propanoate and butanoate metabolism as well as in pathways
50
51
52 171 involved in transport of simple sugars (phosphotransferase system) (**Fig. 2a**). Interestingly,
53
54
55 172 the reporter scores of numerous amino acid metabolic pathways exhibited marked decreases
56
57
58 173 or increases in CD patients compared to CTs, suggesting possible significant changes in the
59
60
61 174 amino acid metabolic profiles (**Fig. 2a**). Of note, the potential for methane metabolism was

175 also diminished in CD patients (**Fig. 2a**). By contrast, microbes in CD patients exhibited
176 enhanced potential for xenobiotic degradation (e.g. of toluene, fluorobenzoate, styrene,
177 benzoate, dioxin, and xylene) and antioxidant defense (e.g. ascorbate, aldarate and glutathione
178 metabolism) (**Fig. 2a**). In parallel, a number of pathways associated with pathogenesis and
179 virulence, including ABC transporters, bacterial secretion system, and general LPS
180 biosynthesis exhibited an incremental enrichment from metacommunity A to C (**Fig. 2a**).
181 LPS, an inherent component of Gram-negative bacteria, is an endotoxin that can have
182 opposing effects on the immune response¹⁷. Since pathway and module analyses showed an
183 enrichment of general LPS biosynthesis in the CD microbiome (**Fig. 2a**), we [took a novel](#)
184 [approach and](#) investigated the capacity amongst all Gram-negative bacteria to produce the
185 pro-inflammatory hexa-acylated LPS as compared to the antagonizing silencing
186 penta-acylated LPS variant^{18,19}. We listed bacteria with a potential for synthesizing each LPS
187 variant (**Supplementary Table 7**) and compared the abundances of these bacteria
188 (**Supplementary Table 8**). The hexa-acylated LPS producing bacteria, *E. coli* and
189 *Morganella morganii* exhibited higher abundance in CD patients from metacommunity C
190 compared to non-CD individuals from metacommunity A (**Supplementary Table 7**).
191 Consistently, compared to metacommunity A (CT), microbes in metacommunity C (CD)
192 tended to produce LPS in a higher hexa- to penta-ratio, suggested by the increase in
193 abundance of bacteria with the hexa- over the penta-acetylated LPS variant (**Fig. 2b**), which
194 in part may account for an increased inflammatory stimulation of the CD gut.

195 The abundances of Gram-positive bacteria were reduced in metacommunity C, and in

1 196 metacommunity B as compared to CTs (**Fig. 2b**). These bacteria make up the largest reservoir
2
3 197 for production of SCFAs. SCFAs are not only colonotrophic nutrients but also
4
5
6 198 immunoregulatory molecules ²⁰ that may reduce pro-inflammatory cues within the gut
7
8
9 199 environment. We estimated the abilities of the metacommunities to produce the SCFAs acetic
10
11
12 200 acid, propionic acid and butyric acid. This was done based on the presence of the genes
13
14
15 201 encoding the last enzyme within the respective biosynthetic pathway, thereby providing an
16
17
18 202 alternative method for predicting the capacity for biosynthesis of the bioactive end products
19
20
21 203 than that used in **Fig. 2a**, which was based on presence of genes involved in overall metabolic
22
23
24 204 pathways. Bacteria with a potential to produce SCFAs are listed in **Supplementary Table 7**.
25
26 205 Evidently, CD microbiotas, particularly those in metacommunity C, showed a decreased
27
28
29 206 abundance of key genes for SCFA production, including acetic acid, propionic acid and
30
31
32 207 butyric acid, when compared to the CT microbiota in metacommunity A (**Fig. 2c**).
33
34
35 208 Concordantly, the abundance of many SCFA-producing bacteria differed between CT and CD
36
37
38 209 samples (**Supplementary Table 8**). Thus, the gut microbiota in CD patients likely produces a
39
40
41 210 suboptimal amount of SCFAs compared to the healthy state.

42
43 211

44 45 46 212 **Disruption of normal gut microbial ecosystem and bacterial growth rate in CD**

47
48
49
50
51 213 The structure of a microbiota is the result of dynamic interactions between community
52
53
54 214 members. We generated correlation-based microbial interaction networks using the SparCC
55
56
57 215 algorithm (**Fig. 3, Supplementary Fig. 3**). Since metacommunity A and C were
58
59
60 216 representative of the typical CT and CD states, respectively, we first compared the

1 217 microbiome networks of these two groups (**Fig. 3a** and **3b**). The control microbiota in
2
3 218 metacommunity A was characterized by a complex network of interactions between different
4
5
6 219 taxa, especially within or between the dominant phyla Bacteroidetes and Firmicutes (**Fig. 3a**).
7
8
9 220 However, the vast majority of these relationships was no longer significant in the CD patients
10
11 221 harboring metacommunity C (**Fig. 3b**). Among the strong interactions lost in the gut
12
13 222 microbiota of the C-CD group were positive correlations ($r>0.5$) of *Bacteroides*
14
15 223 *cellulosilyticus* with *Bacteroides thetaiotaomicron* and *Bacteroides* sp., and of *Ruminococcus*
16
17 224 *bromii* with *Eubacterium ventriosum* (**Fig. 3**). Only one new strong correlation was formed
18
19
20 225 between two unidentified taxa in the C-CD group (**Fig. 3**). Thus, the CD microbiota of
21
22 226 metacommunity C showed not only alterations in composition, but also reduced
23
24
25 227 interrelationships. In comparison, CT and CD microbiotas from metacommunity B did not
26
27
28 228 differ significantly in terms of network complexity, although numerous inter-taxon
29
30
31 229 relationships were altered (**Supplementary Fig. 3**).
32
33
34
35
36
37 230 Changes in bacterial growth rate may contribute to alterations in community structures. We
38
39
40 231 calculated the growth rate from the number of sequencing reads covering the replication
41
42 232 origin relative to reads covering the replication termination site ²¹. Compared to CTs in
43
44
45 233 metacommunity A, the growth rate of many beneficial taxa decreased in metacommunity C,
46
47
48 234 including the SCFA-producing bacteria *Alistipes finegoldii*, *Alistipes shahii*, *Eubacterium*
49
50
51 235 *rectale*, *Roseburia intestinalis*, and several *Faecalibacterium prausnitzii* strains (**Fig. 3** and
52
53 236 **Supplementary Table 9**). Interestingly, certain pathogenic or opportunistic pathogenic
54
55
56 237 bacteria exhibiting an increased abundance in the C-CD group showed high growth rates (*E.*

1 238 *coli*, *Klebsiella pneumoniae*, *Bacteroides fragilis*, and *Streptococcus salivarius*) (**Fig. 3** and
2
3
4 239 **Supplementary Fig. 4** and **Supplementary Table 9**). Thus, differences in growth rate likely
5
6 240 contribute to the alterations in the relative abundance of bacteria in CTs and CD patients,
7
8
9 241 since the observed increase or decrease in growth rates largely concurred with their changes
10
11 242 in relative abundance in CD samples (**Supplementary Fig. 4**). The reduction of growth rates
12
13
14 243 for most bacteria in the C-CD group may also be an indicator that this metacommunity
15
16
17 244 structure is unlikely to shift towards increased diversity over time without specific
18
19
20 245 intervention.

24 246 **Limited remodeling of CD microbiota composition by short-term EEN**

27 247 Fourteen patients in our cohort underwent EEN treatment after baseline sampling and
28
29
30 248 provided fecal samples after two weeks of treatment. We assessed whether short-term EEN
31
32
33 249 was sufficient to alter the microbiome structure in CD patients. For all patients but one
34
35
36 250 (GZCD029, marked by * in **Fig. 4b**), such short time intervention proved insufficient to
37
38
39 251 change their metacommunity identities (**Fig. 4a**), in accord with no significant change in
40
41 252 MD-indices ($p = 0.20$, **Fig. 4a** and **Supplementary Table 1**). However, moderate changes
42
43
44 253 occurred as illustrated by the shift in the relative position of microbiomes along the two
45
46
47 254 principal coordinates within pre-identified clusters (**Fig. 4b**).

50
51 | 255 Despite the limited remodeling of the overall microbiota composition, two-weeks EEN did
52
53 256 induce a variety of functional alterations (**Fig. 4c**, and **Supplementary Table 12** and
54
55
56 257 **Supplementary Table 13**). In a reverse manner to CD-associated shifts, functions such as
57
58
59 258 LPS biosynthesis and bacterial secretion system became less enriched, while starch and

1 259 sucrose metabolism and flagellar assembly were enhanced after EEN (**Fig. 4c**), suggesting a
2
3 260 partial functional recovery. However, certain CD-driven changes, such as functions associated
4
5
6 261 with ribosomes, one carbon folate pool, PTS, and ABC transporters, were exacerbated after
7
8
9 262 two-weeks EEN (**Fig. 4c**), indicating either side effects or temporal disease progression.
10
11 263 Nevertheless, short-term EEN did not affect the abundances of LPS- or SCFA-producing
12
13 264 bacteria (**Fig 4d, e, Supplementary Table 10**) nor their growth rates (**Supplementary Table**
14
15
16
17 265 **11**). However, network re-wiring occurred (**Supplementary Fig. 5**). Rather than interacting
18
19 266 with Firmicutes, bacteria from Bacteroidetes tended to interact with each other after EEN
20
21 267 treatment (**Supplementary Fig. 5**). By contrast, a majority of Firmicutes in patients after
22
23 268 EEN treatment presented more inter-dependences with Proteobacteria and unclassified
24
25
26
27 269 species compared to those before treatment (**Supplementary Fig. 5**). Overall, the CD
28
29 270 microbiota appeared relatively stable and refractory to two-week EEN intervention. Future
30
31
32
33 271 studies will need to determine if a longer intervention period with EEN will result in
34
35
36 272 restoration of normal functional microbiota in CD patients.
37
38
39
40

41 273
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 295 more recently established that only the hexa-acylated LPS variant is able to activate
2
3 296 pro-inflammatory cues via TLR4 in humans ¹⁹, while the penta-acylated LPS variant acts as
4
5
6 297 an antagonist ¹⁸. Our finding that the CD microbiota of metacommunity C was enriched in
7
8
9 298 microbes producing hexa-acylated LPS is consistent with previous observations of the
10
11 299 increased abundance of the *Enterobacteriaceae* family members in CD ⁴⁻⁸, which are known
12
13 300 to stimulate inflammation ²⁴. Together, these changes may severely affect the host immune
14
15
16
17 301 system, leading to an unchecked inflammatory state in CD. The reduction in the network
18
19
20 302 complexity of the CD microbiota of metacommunity C reinforced the view that a globally
21
22 303 disturbed microbial ecosystem may contribute to this disease. The loss of reciprocal and
23
24
25 304 cross-inhibitory relationships may impair the survival of beneficial microbes and create
26
27
28 305 favorable conditions for the blooming of pathogens. Likewise, it appears to limit growth of
29
30
31 306 many gut bacteria found in healthy individuals. In this regard, reconstruction of the normal
32
33
34 307 ecosystem and not only the mere introduction of a single or several commensal microbes may
35
36 308 be needed to curb CD. In the case of EEN, a longer term of treatment may be needed to
37
38
39 309 achieve this goal. Analysis of the fecal microbiota is widely used as a proxy for studying the
40
41
42 310 gut microbiota composition because of the easiness and noninvasive nature of fecal sampling,
43
44
45 311 and has through the years resulted in deepening the understanding of the relationship between
46
47
48 312 the gut microbiota and IBD ^{1,6}. However, new avenues of sampling procedures open up for
49
50
51 313 more comprehensive insights into the role played by the intestinal location of microbial
52
53 314 species (luminal or mucosal layer attachment to the small and the large intestine) that, in
54
55
56 315 combination with metagenomic sequencing, would allow for deeper insights into the
57
58
59 316 inter-individual diversity in ecological dysbalance in CD patients in future studies.

1 317 Taken together, our metagenome-scale characterization of the CD gut microbiome supports
2
3 318 the notion of a shift towards enhanced pro-inflammatory capacity, which is most pronounced
4
5
6 319 in individuals harboring the severe-state metacommunity C. The level of details in this
7
8
9 320 analysis, also encompassing yet unannotated bacteria, may pave the way for elucidating
10
11 321 microbial disturbances predictive for CD by enabling the discovery of composite microbial
12
13
14 322 CD biomarkers. In addition, it may allow for the identification of future therapeutic targets
15
16
17 323 based on microbiota signatures, thereby implementing personalized medicine to CD patients
18
19
20 324 based on the individual microbiome composition.
21

22 | 325
23 |
24 |
25 |

26 326 **Methods**

27 28 29 30 327 **Study cohort, EEN treatment and sample collection**

31
32
33
34 328 49 CD patients and 54 healthy controls were enrolled in this study at the Sixth Affiliated
35
36
37 329 Hospital of the Sun Yat-sen University, Guangdong, China. All patients met the diagnostic
38
39
40 330 criteria for CD, according to the Montreal classification system ²⁵. Patients diagnosed with
41
42
43 331 diabetes, tumor, cardiovascular, kidney, liver, and metabolic diseases were excluded from this
44
45 332 study.
46

47 333

48
49
50 334 Among these participants, 14 CD patients underwent EEN treatment. ENSURE® (Abbott
51
52
53 335 Laboratories, Abbott Park, USA), PEPTISON®, NUTRISON POWDER® (NUTRICIA,
54
55
56 336 Danone, Netherlands) and FRESUBIN® (Sino-Swed Pharmaceutical Corp. Ltd, China) were
57
58
59 337 used as the standard oral polymeric formulas, and their ingredients are detailed in
60

1 338 Supplementary Table 14. Patients chose from these formulas, with 8 patients selecting
2
3 339 ENSURE® and the others selecting a mixture of two or more formulas. Formulas were
4
5
6 340 consumed at 30 kcal/kg per day as the sole nutrient source. Patients who adhered to EEN
7
8
9 341 treatment had their lesion healed.

10
11 342

12
13
14 343 Fecal samples were collected from all participants at baseline (n=103), and from the
15
16
17 344 EEN-treated CD patients after 2 weeks of treatment (n=14), totaling 117 samples. The fecal
18
19
20 345 samples were immediately frozen and stored at -80°C until being processed. DNA extraction
21
22
23 346 was performed according to the protocols described previously ²⁶.

24
25 347

26
27
28 348 All protocols in this study were approved by the institutional review boards at Sixth Affiliated
29
30
31 349 Hospital of Sun Yat-sen University and BGI-Shenzhen, and they were conducted in
32
33
34 350 compliance with the Declaration of Helsinki. Explicit informed consent was obtained from all
35
36
37 351 subjects.

38
39 352

40 41 42 43 353 **Metagenomic sequencing and assembly**

44
45
46 354 Paired-end metagenomic sequencing was conducted on the Illumina platform (insert size, 350
47
48
49 355 bp; read length, 100 bp). Quality control was performed and adaptor and host contamination
50
51
52 356 were filtered. Sequencing reads were de novo assembled into contigs with SOAPdenovo
53
54
55 357 v2.04 ²⁷ as described previously ²⁶.

56
57
58 358

1 **359 Co-abundance gene groups identification and functional annotation**

2
3
4 360 Applying the metagenomic species (MGS) clustering method ¹³, we clustered genes according
5
6
7 361 to their co-variations in abundance across samples. A group of co-abundant genes was
8
9
10 362 identified as a MGS if it contained 700 or more genes. These MGS were subjected to
11
12
13 363 subsequent analysis. Taxonomic assignment of the mapped genes was performed according to
14
15
16 364 the Integrated Microbial Genomes (IMG, v400) database using an in-house pipeline detailed
17
18 365 previously ²⁶, with 70% overlap and 65% identity for assignment to phylum, 85% identity to
19
20
21 366 genus, and 95% identity to species. The relative abundance of a co-abundance gene group was
22
23
24 367 calculated from the relative abundance of its genes.

25
26 368 Differentially enriched KO pathways or modules were identified according to their reporter
27
28
29 369 scores ²⁸, which were calculated from the Z-scores of individual KOs.

30
31
32 370 We assessed the production capacity for the two LPS forms based on the abundances of genes
33
34
35 371 of the entire lipid A biosynthesis pathway, and separated them into penta-acylated LPS
36
37
38 372 producers (harboring all lipid A pathway genes except for LpxM), and pro-inflammatory
39
40
41 373 hexa-acylated LPS producers (all lipid A pathway genes). MGSs with no lipid A pathway
42
43
44 374 genes were assigned as Gram-positive bacteria.

45
46 375 Sequences of SCFA-producing enzymes were retrieved as previously described ²⁹. Genes in
47
48
49 376 the reference gut microbiome gene catalog³⁰ were identified as these enzymes (best match
50
51
52 377 according to BlastP, identity > 35%, score > 60, E<1e-3), and their relative abundances could
53
54
55 378 then be determined accordingly.

56
57 379

1 380 **α -Diversity and gene count**

2
3
4 381 α -Diversity (within-sample diversity) was calculated on the basis of the gene profile of each
5
6
7 382 sample according to the Shannon index as described previously ²⁶. The total gene count in
8
9
10 383 each fecal sample was determined as in ref. ³¹. Genes with at least one mapped read were
11
12
13 384 considered present.

14
15 385

16
17
18
19 386 **PERMANOVA of the influence of clinical and lifestyle factors**

20
21
22
23 387 Permutational multivariate analysis of variance (PERMANOVA) ²⁶ was performed on the
24
25
26 388 gene-abundance profiles of the samples to assess the effect of each of the factors listed in
27
28
29 389 Table 1. We used Bray-Curtis distance and 9,999 permutations in R (3.10, vegan package) ³².

30
31 390

32
33
34 391 **Details of LEfSe algorithm**

35
36
37
38 392 Differential abundance analyses were performed using the LEfSe algorithm to identify feature
39
40
41 393 microbes whose abundances differed at least in one comparison ⁵. Metacomunities and
42
43
44 394 subgroups in metacomunity B were included for comparisons. The biomarker relevance was
45
46
47 395 ranked according to bootstrapped (n=30) logarithmic linear discriminant analysis scores of at
48
49 396 least 2.

50
51 397

52
53
54
55 398 **Effect size analysis**

1 399 24 metadata covariates and their combined effect size when pooled into the broader
2
3 400 predefined categories (blood fat, coagulation, inflammation markers, and plasma biomarkers)
4
5
6 401 was estimated with the *bioenv* function in the vegan R package, which selects the
7
8
9 402 combination of covariates with strongest correlation to microbiota variation (Pearson
10
11
12 403 correlation between Gower distances of covariates and microbiome Bray-Curtis dissimilarity,
13
14 404 Supplementary Fig. 2A).
15

16
17 405

20 21 406 **Correlation network inferred by phylogenetic marker genes**

22
23 407 Eighty-five MGS, which were previously selected via the detection of microbial community
24
25
26 408 clusters through DMM modelling, were subjected to compositionality data analysis using the
27
28
29 409 SparCC algorithm³³. Taxon–taxon correlation coefficients were estimated as the average of
30
31
32 410 20 inference iterations with the strength threshold of 0.25. Correlations with the
33
34
35 411 corresponding empirical P values less than 0.01 were retained, which was calculated via a
36
37
38 412 total of 10,000 simulated data sets. This set of iterative procedures was applied separately to
39
40
41 413 data from CTs and CD patients, and to patients’ data before and after EEN to infer the
42
43
44 414 correlation values. Correlation coefficients with magnitude of 0.3 or above were selected for
45
46 415 visualization in Cytoscape (version 3.3.0).
47

48
49 416

50 51 417 **Availability and requirements**

52
53
54 418 Project name: Kruskal.EffectSize.R

55
56
57 419 Project home page: <https://github.com/andriaYG/LDA-EffectSize>
58
59

1 420 Operating system: Linux

2
3 421 Programming language: R

4
5
6 422 Other requirements: N/A

7
8
9 423 License: N/A

10
11 424 **Availability of supporting data**

12
13
14 425 The data sets supporting the results of this article are available in the GigaDB repository, on
15
16
17 426 the

18
19
20 427 **List of abbreviations**

21
22 428 CD, Crohn's disease; CT, controls; EEN, exclusive enteral nutrition; IBD, inflammatory

23
24
25 429 bowel disease; GI, gastrointestinal; AIEC, adherent-invasive *Escherichia coli*; MAP,

26
27
28 430 *Mycobacterium avium paratuberculosis*; MGS, metagenomics species; LDA, linear

29
30
31 431 discriminant analysis; LEfSe, linear discriminant analysis effect size; MD-index, microbial

32
33
34 432 dysbiosis index; PCoA, principal coordinate analysis; UA, uric acid; KEGG, Kyoto

35
36
37 433 Encyclopedia of Genes and Genomes; LPS, lipopolysaccharide; SCFA, short-chain fatty acid;

38
39 434 PAMP, pathogen-associated molecular pattern; IMG, Integrated Microbial Genomes.

40
41
42 435 **Competing interests**

43
44
45 436 The authors declare that they have no competing interests

46
47 437 **Funding**

48
49
50 438 This research was supported by the National Natural Science Foundation of China (Nos

51
52
53 439 81470795), the Shenzhen Municipal Government of China (grant No. DRC-SZ[2015]162,

54
55
56 440 JSGG20140702161403250, JSGG20160229172752028, JCYJ20160229172757249,

57
58
59 441 JCYJ20140418095735538, CXZZ20150330171521403, CXB201108250098A).

1 442

2
3 443 **Authors' contributions**

4
5
6 444 All authors read and approved the final manuscript. Q.H., Jian W., Huanming Y., X.X. and
7
8
9 445 X.L. conceived the study. Q.H. participated in the design of the study. L.X., Y.L., L.L, Faming
10
11
12 446 Z., Q.F., Xiaoping L., J.Y., C.L., J.C., and T.Y. carried out the sample collection and
13
14
15 447 preparation. Y.G. and Z.J. participated in sequence assembly, gene mapping and MGS
16
17
18 448 identification. J.M.L. and S.B. performed the analysis of LPS variants. Z.J. generated the
19
20
21 449 SCFA abundance profile. Y.G. carried out the bioinformatics analysis of metacommunities,
22
23
24 450 functions and networks. Y.G., X.Y., L.M, S.B., K.K. and H.J. wrote the manuscript. K.K., S.B.
25
26 451 and H.J. supervised project.

27
28 452 **Acknowledgements**

29
30
31 453 We gratefully acknowledge colleagues at BGI-Shenzhen for DNA extraction, library
32
33
34 454 construction, sequencing, and discussions.

35
36
37 455
38
39 456 **References**

- 40
41
42
43 457 1 Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal
44
45
46 458 immune responses during health and disease. *Nat Rev Immunol* **9**, 313-323,
47
48
49 459 doi:10.1038/nri2515 (2009).
- 50
51 460 2 Barnich, N. & Darfeuille-Michaud, A. Adherent-invasive *Escherichia coli* and
52
53
54 461 Crohn's disease. *Curr Opin Gastroenterol* **23**, 16-20,
55
56
57 462 doi:10.1097/MOG.0b013e3280105a38 (2007).

1 463 3 Hermon-Taylor, J. *et al.* Causation of Crohn's disease by Mycobacterium
2
3 464 avium subspecies paratuberculosis. *Can J Gastroenterol* **14**, 521-539 (2000).
4
5
6 465 4 Ricanek, P. *et al.* Gut bacterial profile in patients newly diagnosed with
7
8 466 treatment-naive Crohn's disease. *Clin Exp Gastroenterol* **5**, 173-186,
9
10 467 doi:10.2147/CEG.S33858 (2012).
11
12
13
14 468 5 Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's
15
16 469 disease. *Cell Host Microbe* **15**, 382-392, doi:10.1016/j.chom.2014.02.005
17
18 470 (2014).
19
20
21
22 471 6 Imhann, F. *et al.* Interplay of host genetics and gut microbiota underlying the
23
24 472 onset and clinical presentation of inflammatory bowel disease. *Gut*,
25
26 473 doi:10.1136/gutjnl-2016-312135 (2016).
27
28
29
30 474 7 Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory
31
32 475 bowel disease and treatment. *Genome Biol* **13**, R79,
33
34 476 doi:10.1186/gb-2012-13-9-r79 (2012).
35
36
37
38 477 8 Thorkildsen, L. T. *et al.* Dominant fecal microbiota in newly diagnosed
39
40 478 untreated inflammatory bowel disease patients. *Gastroenterol Res Pract* **2013**,
41
42 479 636785, doi:10.1155/2013/636785 (2013).
43
44
45
46 480 9 Torres, J., Mehandru, S., Colombel, J. F. & Peyrin-Biroulet, L. Crohn's disease.
47
48 481 *Lancet*, doi:10.1016/S0140-6736(16)31711-1 (2016).
49
50
51
52 482 10 Buchman, A. L. Side effects of corticosteroid therapy. *J Clin Gastroenterol* **33**,
53
54 483 289-294 (2001).
55
56
57
58 484 11 Wall, C. L., Day, A. S. & Geary, R. B. Use of exclusive enteral nutrition in
59
60

1 485 adults with Crohn's disease: a review. *World J Gastroenterol* **19**, 7652-7660,
2
3 486 doi:10.3748/wjg.v19.i43.7652 (2013).
4
5
6 487 12 Day, A. S. & Lopez, R. N. Exclusive enteral nutrition in children with Crohn's
7
8
9 488 disease. *World J Gastroenterol* **21**, 6809-6816, doi:10.3748/wjg.v21.i22.6809
10
11 489 (2015).
12
13
14 490 13 Nielsen, H. B. *et al.* Identification and assembly of genomes and genetic
15
16
17 491 elements in complex metagenomic samples without using reference genomes.
18
19
20 492 *Nat Biotechnol* **32**, 822-828, doi:10.1038/nbt.2939 (2014).
21
22
23 493 14 Holmes, I., Harris, K. & Quince, C. Dirichlet multinomial mixtures:
24
25
26 494 generative models for microbial metagenomics. *PLoS One* **7**, e30126,
27
28 495 doi:10.1371/journal.pone.0030126 (2012).
29
30
31 496 15 Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome*
32
33
34 497 *Biol* **12**, R60, doi:10.1186/gb-2011-12-6-r60 (2011).
35
36
37 498 16 Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous
38
39
40 499 *Clostridium* species. *Science* **331**, 337-341, doi:10.1126/science.1198469
41
42 500 (2011).
43
44
45 501 17 Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu Rev*
46
47
48 502 *Biochem* **71**, 635-700, doi:10.1146/annurev.biochem.71.110601.135414
49
50 503 (2002).
51
52
53 504 18 Park, B. S. *et al.* The structural basis of lipopolysaccharide recognition by the
54
55
56 505 TLR4–MD-2 complex. *nature* **458**, 1191-1195 (2009).
57
58
59 506 19 Brix, S., Eriksen, C., Larsen, J. M. & Bisgaard, H. Metagenomic heterogeneity

1 507 explains dual immune effects of endotoxins. *Journal of Allergy and Clinical*
2
3 508 *Immunology* **135**, 277 (2015).
4
5
6 509 20 Puertollano, E., Kolida, S. & Yaqoob, P. Biological significance of short-chain
7
8 510 fatty acid metabolism by the intestinal microbiome. *Current opinion in clinical*
9
10
11 511 *nutrition and metabolic care* **17**, 139-144,
12
13
14 512 doi:10.1097/mco.000000000000025 (2014).
15
16
17 513 21 Korem, T. *et al.* Growth dynamics of gut microbiota in health and disease
18
19 514 inferred from single metagenomic samples. *Science* **349**, 1101-1106,
20
21
22 515 doi:10.1126/science.aac4812 (2015).
23
24
25 516 22 Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the
26
27 517 differentiation of colonic regulatory T cells. *Nature* **504**, 446-450,
28
29
30 518 doi:10.1038/nature12721 (2013).
31
32
33 519 23 Singh, N. *et al.* Activation of Gpr109a, receptor for niacin and the commensal
34
35 520 metabolite butyrate, suppresses colonic inflammation and carcinogenesis.
36
37
38 521 *Immunity* **40**, 128-139, doi:10.1016/j.immuni.2013.12.007 (2014).
39
40
41 522 24 Jensen, S. R. *et al.* Distinct inflammatory and cytopathic characteristics of
42
43 523 *Escherichia coli* isolates from inflammatory bowel disease patients.
44
45
46 524 *International Journal of Medical Microbiology* **305**, 925-936 (2015).
47
48
49 525 25 Silverberg, M. S. *et al.* Toward an integrated clinical, molecular and
50
51 526 serological classification of inflammatory bowel disease: Report of a Working
52
53
54 527 Party of the 2005 Montreal World Congress of Gastroenterology. *Canadian*
55
56
57 528 *Journal of Gastroenterology and Hepatology* **19**, 5A-36A (2005).
58
59
60
61
62
63
64
65

1 529 26 Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2
2
3 530 diabetes. *Nature* **490**, 55-60 (2012).
4
5
6 531 27 Luo, R. *et al.* SOAPdenovo2: an empirically improved memory-efficient
7
8 532 short-read de novo assembler. *GigaScience* **1**, 1 (2012).
9
10
11 533 28 Patil, K. R. & Nielsen, J. Uncovering transcriptional regulation of metabolism
12
13 534 by using metabolic network topology. *Proceedings of the National Academy of*
14
15 535 *Sciences of the United States of America* **102**, 2685-2689 (2005).
16
17
18
19 536 29 Claesson, M. J. *et al.* Gut microbiota composition correlates with diet and
20
21 537 health in the elderly. *Nature* **488**, 178-184, doi:10.1038/nature11319 (2012).
22
23
24
25 538 30 Li, J. *et al.* An integrated catalog of reference genes in the human gut
26
27 539 microbiome. *Nature biotechnology* **32**, 834-841 (2014).
28
29
30
31 540 31 Le Chatelier, E. *et al.* Richness of human gut microbiome correlates with
32
33 541 metabolic markers. *Nature* **500**, 541-546 (2013).
34
35
36 542 32 Zapala, M. A. & Schork, N. J. Multivariate regression analysis of distance
37
38 543 matrices for testing associations between gene expression patterns and related
39
40 544 variables. *Proceedings of the national academy of sciences* **103**, 19430-19435
41
42 545 (2006).
43
44
45
46
47 546 33 Friedman, J. & Alm, E. J. Inferring correlation networks from genomic survey
48
49 547 data. *PLoS Comput Biol* **8**, e1002687 (2012).
50
51 548
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 549

2
3
4 550 **Figure legends**

5
6
7
8 551 **Figure 1. Clustering of gut microbiota into metacommunities associated with CD. (a)**

9
10 552 Heatmap of signature microbes for three metacommunities determined by DMM model.

11
12 553 Rows correspond to 85 discriminative MGS, with hierarchical clustering by their relative

13
14 554 abundances. Taxonomic annotations of these MGS are indicated at the right and colored by

15
16 555 phylum. Each column corresponds to one sample. The disease status (the first horizontal bar)

17
18 556 and metacommunity membership (the second horizontal bar) of samples are indicated by

19
20 557 color at the top, and MD index for each sample is represented by gray scale (the third

21
22 558 horizontal bar). (b) PCoA of the 85 MGS based on Jensen-Shannon distance (JSD). Colors

23
24 559 indicate metacommunity memberships, and shapes (triangle or round) denote disease states

25
26 560 (CT or CD).

27
28 561

29
30 562 **Figure 2. Functional alterations of the gut microbiota in CD. (a)** Heatmap and hierarchical

31
32 563 clustering of KEGG pathways that are differentially enriched between the microbiota groups

33
34 564 identified in Fig 1a. Color scale represents reporter score, and only KEGG pathways with a

35
36 565 reporter score greater than 1.9 are shown. (b) Relative abundances of Gram-negative MGS

37
38 566 (the first left panel), Gram-positive MGS (the second left panel), penta-acylated LPS

39
40 567 producing MGS (the middle panel), hexa-acylated LPS producing MGS (the second last

41
42 568 panel), and the ratio of hexa- to penta-acylated LPS producing MGS (the last panel) across

43
44 569 different groups. The value of relative abundance was log-transformed. (c) Relative

1 570 abundances of genes encoding key enzymes for the biosynthesis of different SCFAs across
2
3 571 different microbiota groups. Carbon monoxide dehydrogenase and acetyl CoA synthase
4
5
6 572 complex are crucial for acetic acid production; propionyl-CoA transferase and
7
8
9 573 propionyl-CoA/succinyl-CoA transferase are responsible for propionate acid synthesis;
10
11
12 574 butyryl CoA transferase accounts for butyric acid generation. Their relative abundances were
13
14 575 log-transformed. **(b,c)** Statistical comparison by Wilcoxon test followed by a
15
16
17 576 Benjamini-Hochberg correction for significance level; * $q < 0.2$; ** $q < 0.1$; *** $q < 0.05$;
18
19
20 577 **** $q < 0.001$.

21
22 578

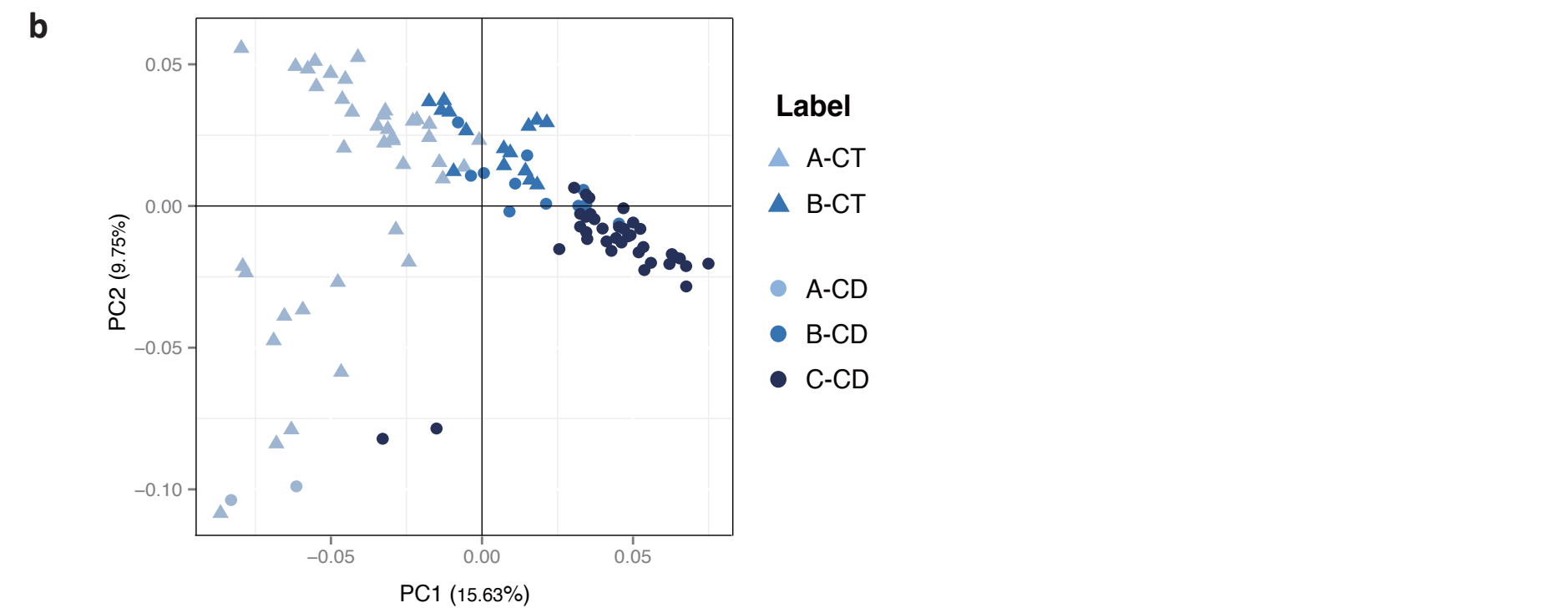
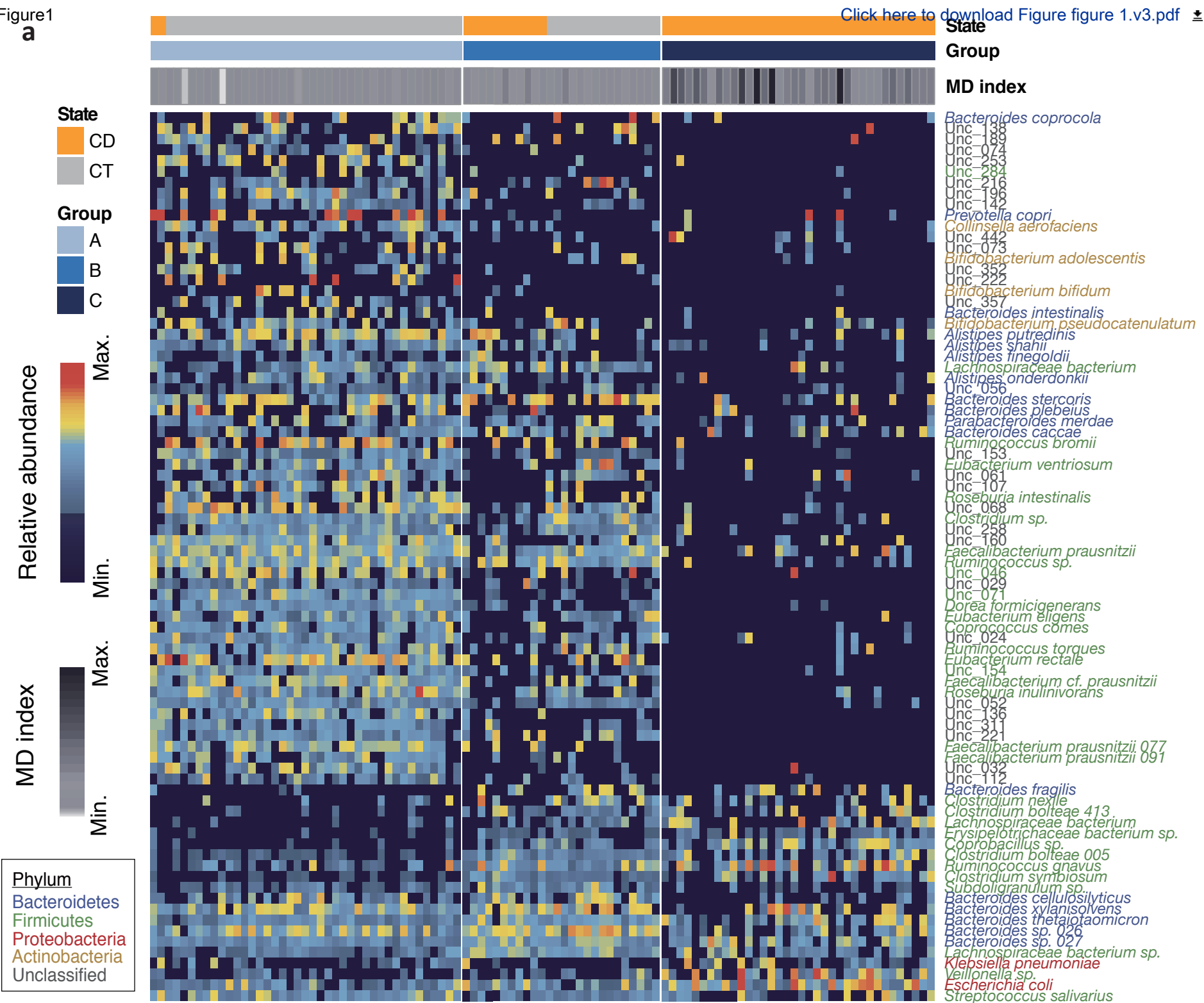
23
24
25 579 **Figure 3. Reconstruction of microbial interaction networks by CD.** Co-occurrence (blue)
26
27
28 580 relationships and co-exclusion (red) between taxa were estimated by SparCC algorithm, and
29
30
31 581 correlation networks were compared between non-CD samples from metacommunity A (**a**,
32
33
34 582 A-CT) and CD samples from metacommunity C (**b**, C-CD). Only relationships with
35
36 583 coefficients above 0.3 are visualized, and the thickness of lines denotes strength of correlation
37
38
39 584 as indicated in the legend. Node size represents mean taxon abundance in networks, and node
40
41
42 585 color represents the growth rate of each species (grey indicates no detection). Taxa of the
43
44
45 586 same bacterial phylum are encircled by dashed lines.

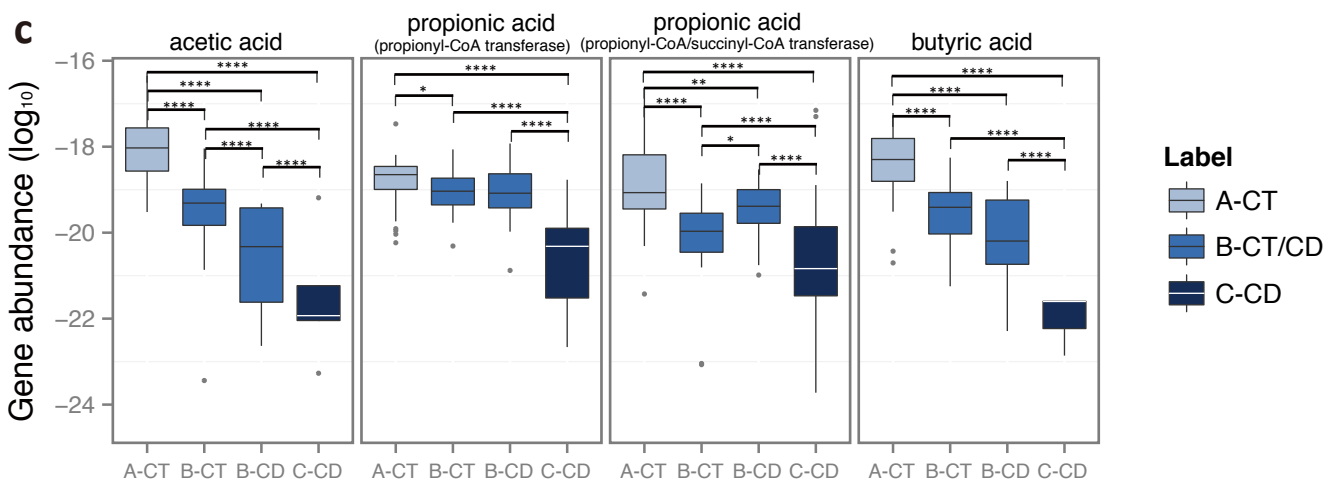
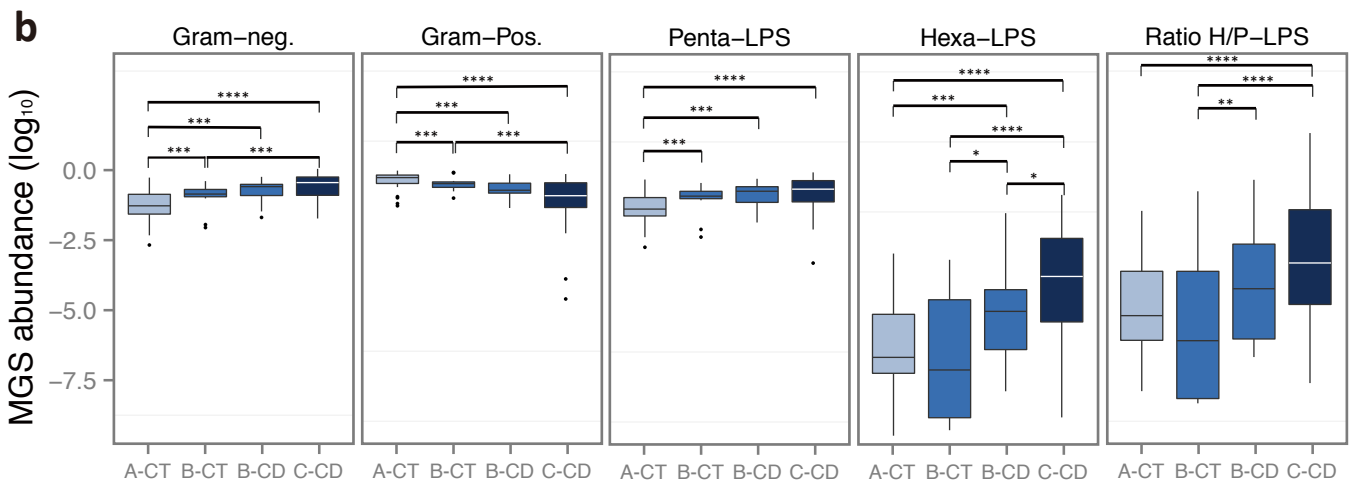
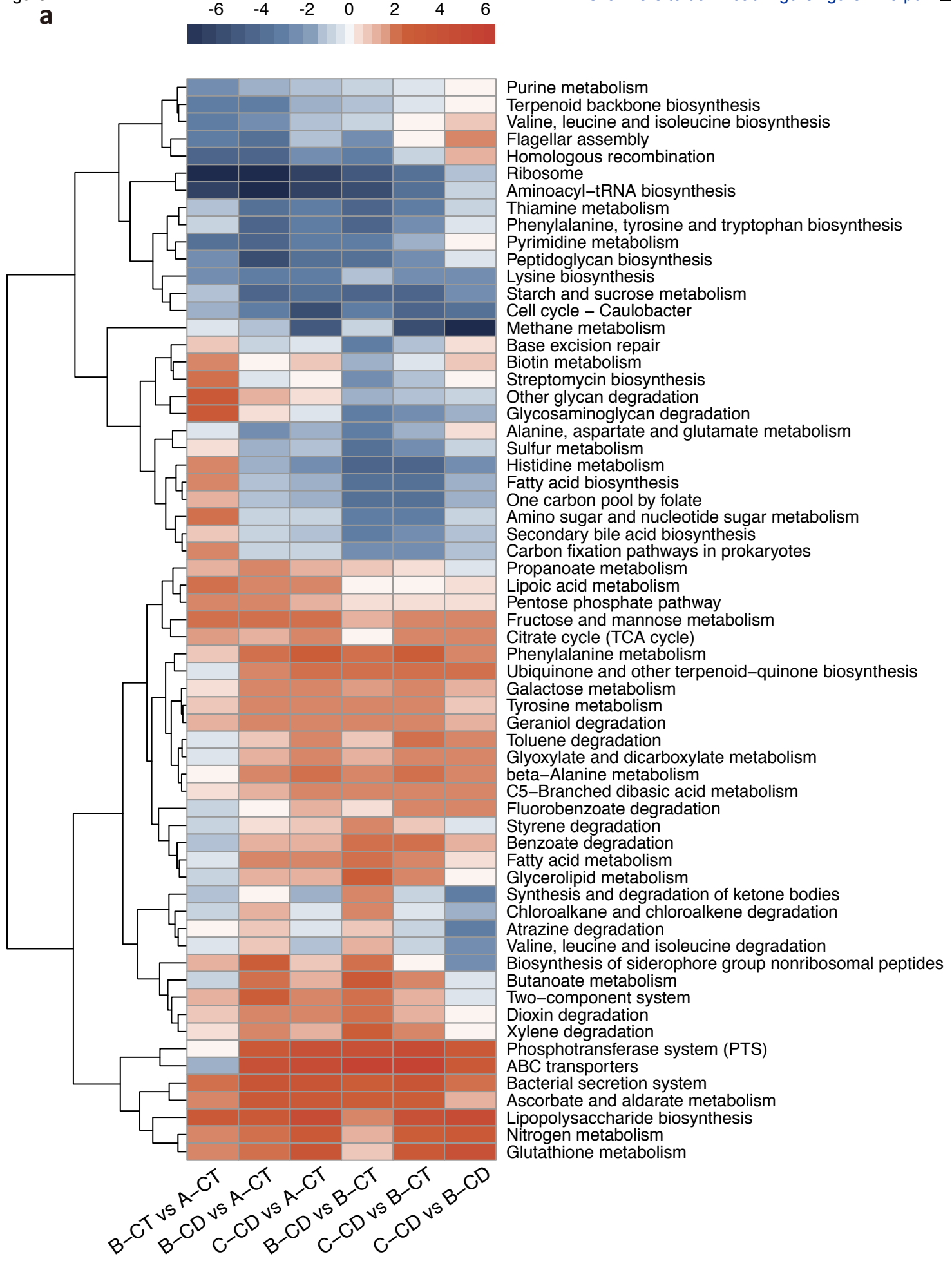
46
47 587

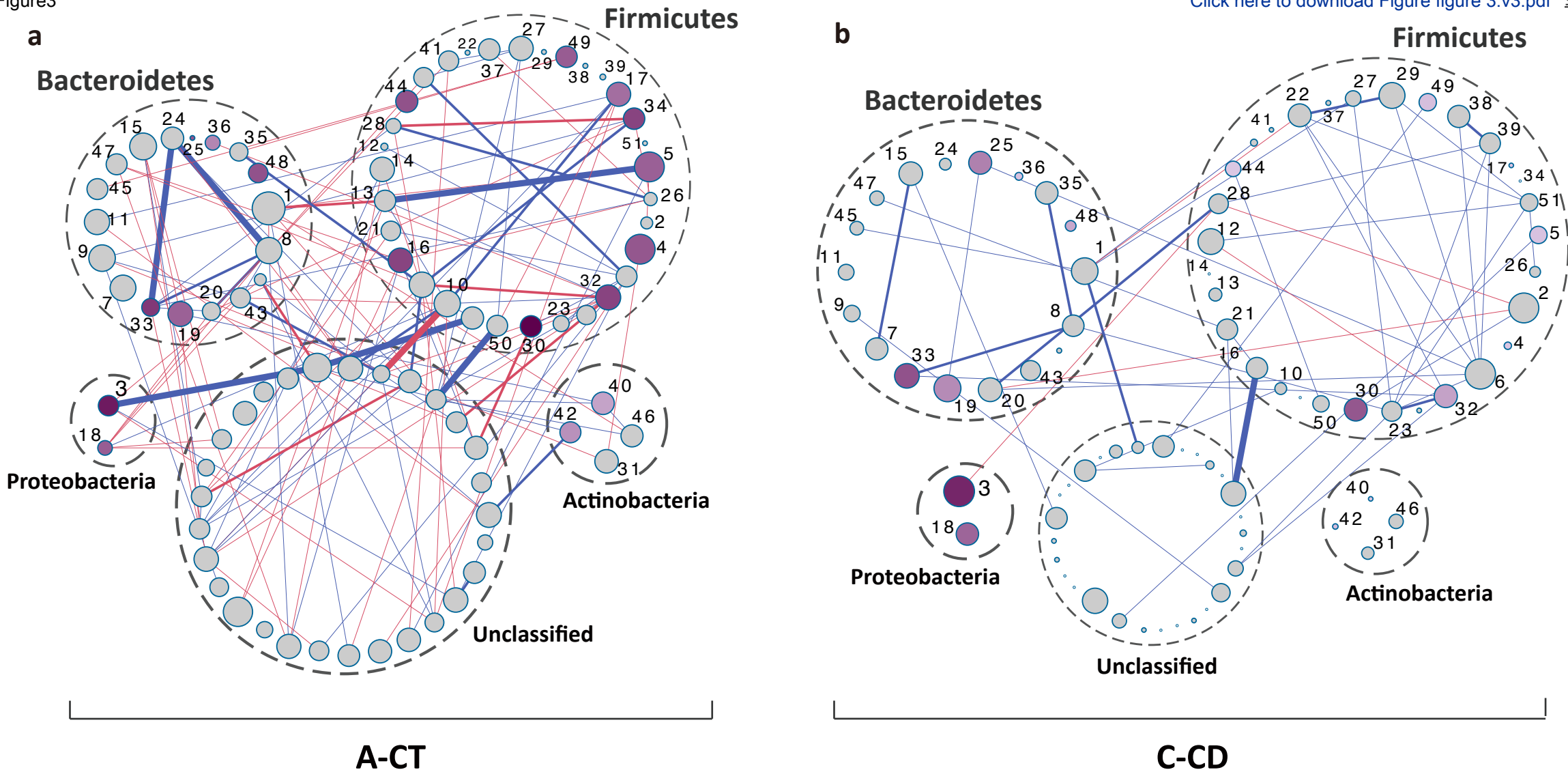
48
49
50 588 **Figure 4. Moderate modification of CD microbiota by EEN treatment.** **(a)** Gut MGS from
51
52
53 589 CD patients (n=14) before and after 14 days of EEN were clustered into metacommunities
54
55
56 590 and visualized as a heatmap representing the 85 discriminative MGS (as in Fig. 1a). Each
57
58
59 591 column corresponds to one sample. **(b)** PCoA of pre- and post-EEN CD microbiota based on

1 592 Jensen-Shannon distance (JSD). Arrows indicate the shift of position along the first two
2
3 593 principal coordinates pre- to post-EEN treatment. The sample whose metacommunity identity
4
5
6 594 changed after EEN treatment is marked with an asterisk (GZCD029). (c) Heatmap and
7
8
9 595 hierarchical clustering KEGG pathways that were enriched or decreased in post- versus
10
11 596 pre-EEN. Color scale represents reporter score, and only KEGG pathways with a reporter
12
13
14 597 score greater than 1.9 are shown. (d) Log_{10} relative abundances of Gram-negative MGS (the
15
16
17 598 first left panel), Gram-positive MGS (the second left panel), penta-acylated LPS producing
18
19
20 599 MGS (the middle panel), hexa-acylated LPS producing MGS (the second last panel), and the
21
22
23 600 ratio of hexa- to penta-acylated LPS producing MGS (the last panel) in pre- versus post-EEN.
24
25
26 601 (e) Log_{10} relative abundances of genes encoding key enzymes for the biosynthesis of different
27
28
29 602 SCFAs in pre- versus post-EEN, as calculated in Figure 2c. (d,e) Statistical comparison by
30
31 603 Wilcoxon test followed by a Benjamini-Hochberg correction for significance level showed no
32
33
34 604 changes between groups.

35
36 605
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65







A-CT

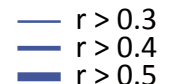
C-CD

- | | | | |
|--|---|---|--|
| 1. <i>Prevotella copri</i> | 14. <i>Faecalibacterium cf. prausnitzii</i> | 27. <i>Ruminococcus sp.</i> | 40. <i>Bifidobacterium bifidum</i> |
| 2. <i>Veillonella sp.</i> | 15. <i>Bacteroides plebeius</i> | 28. <i>Lachnospiraceae bacterium sp.</i> | 41. <i>Dorea formicigenerans</i> |
| 3. <i>Escherichia coli</i> | 16. <i>Faecalibacterium prausnitzii</i> 077 | 29. <i>Clostridium symbiosum</i> | 42. <i>Bifidobacterium adolescentis</i> |
| 4. <i>Eubacterium rectale</i> | 17. <i>Roseburia intestinalis</i> | 30. <i>Streptococcus salivarius</i> | 43. <i>Alistipes onderdonkii</i> |
| 5. <i>Ruminococcus bromii</i> | 18. <i>Klebsiella pneumoniae</i> | 31. <i>Collinsella aerofaciens</i> | 44. <i>Eubacterium eligens</i> |
| 6. <i>Ruminococcus gnavus</i> | 19. <i>Bacteroides xylanisolvens</i> | 32. <i>Faecalibacterium prausnitzii</i> 094 | 45. <i>Bacteroides intestinalis</i> |
| 7. <i>Bacteroides stercoris</i> | 20. <i>Bacteroides sp.</i> 027 | 33. <i>Bacteroides thetaiotaomicron</i> | 46. <i>Bifidobacterium pseudocatenulatum</i> |
| 8. <i>Bacteroides sp.</i> 026 | 21. <i>Lachnospiraceae bacterium</i> 065 | 34. <i>Faecalibacterium prausnitzii</i> 091 | 47. <i>Parabacteroides merdae</i> |
| 9. <i>Alistipes putredinis</i> | 22. <i>Coprobacillus sp.</i> | 35. <i>Bacteroides caccae</i> | 48. <i>Alistipes shahii</i> |
| 10. <i>Roseburia inulinivorans</i> | 23. <i>Clostridium nexile</i> | 36. <i>Alistipes finegoldii</i> | 49. <i>Ruminococcus torques</i> |
| 11. <i>Bacteroides coprocola</i> | 24. <i>Bacteroides cellulosilyticus</i> | 37. <i>Coprococcus comes</i> | 50. <i>Clostridium sp.</i> |
| 12. <i>Lachnospiraceae bacterium</i> 100 | 25. <i>Bacteroides fragilis</i> | 38. <i>Clostridium bolteae</i> 413 | 51. <i>Erysipelotrichaceae bacterium sp.</i> |
| 13. <i>Eubacterium ventriosum</i> | 26. <i>Subdoligranulum sp.</i> | 39. <i>Clostridium bolteae</i> 005 | |

Growth rate



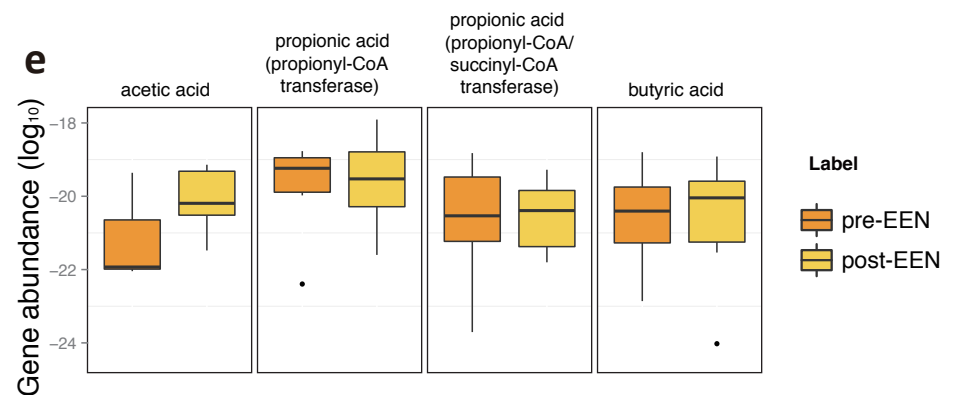
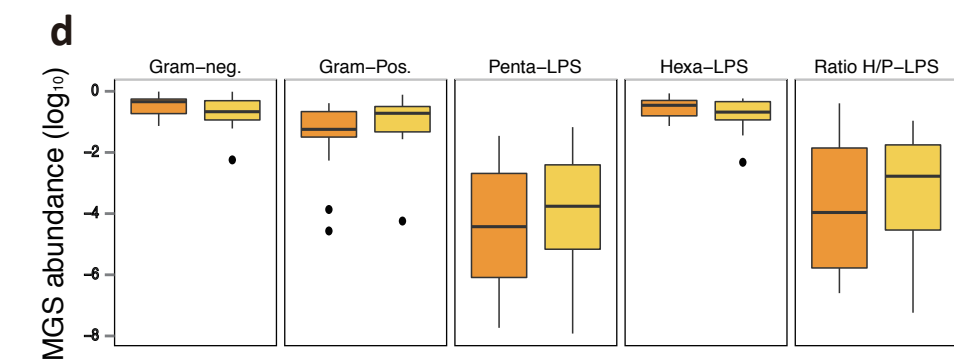
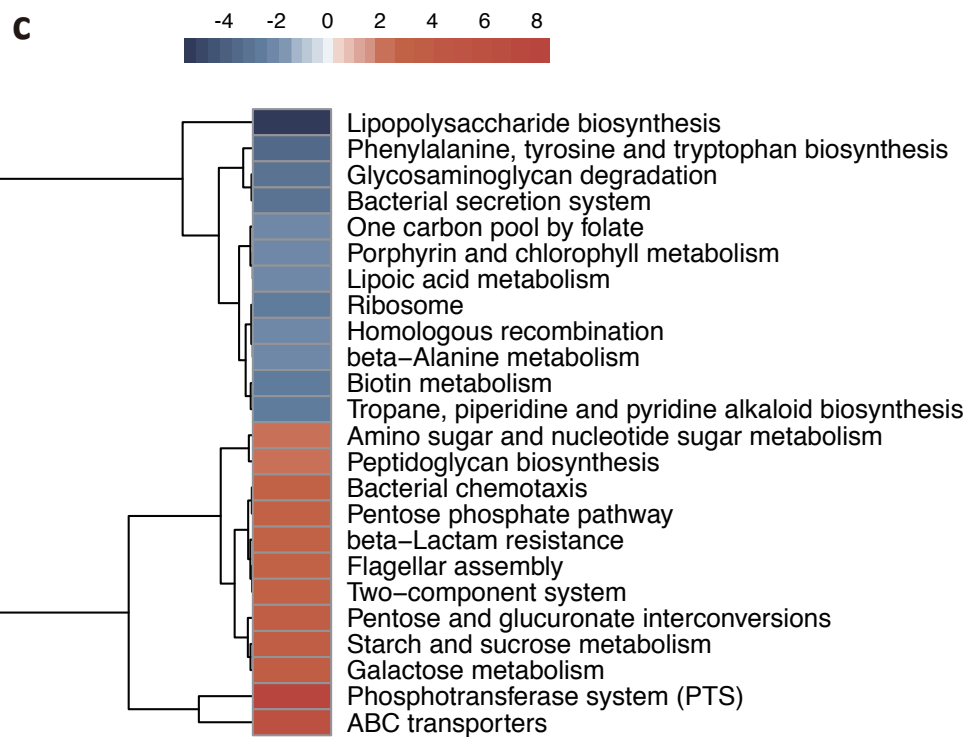
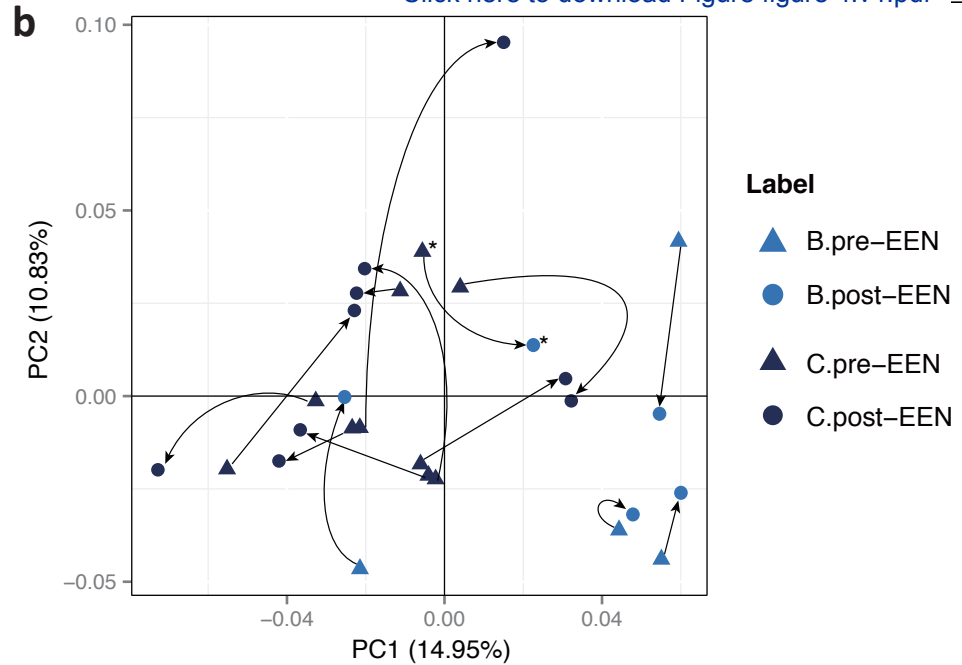
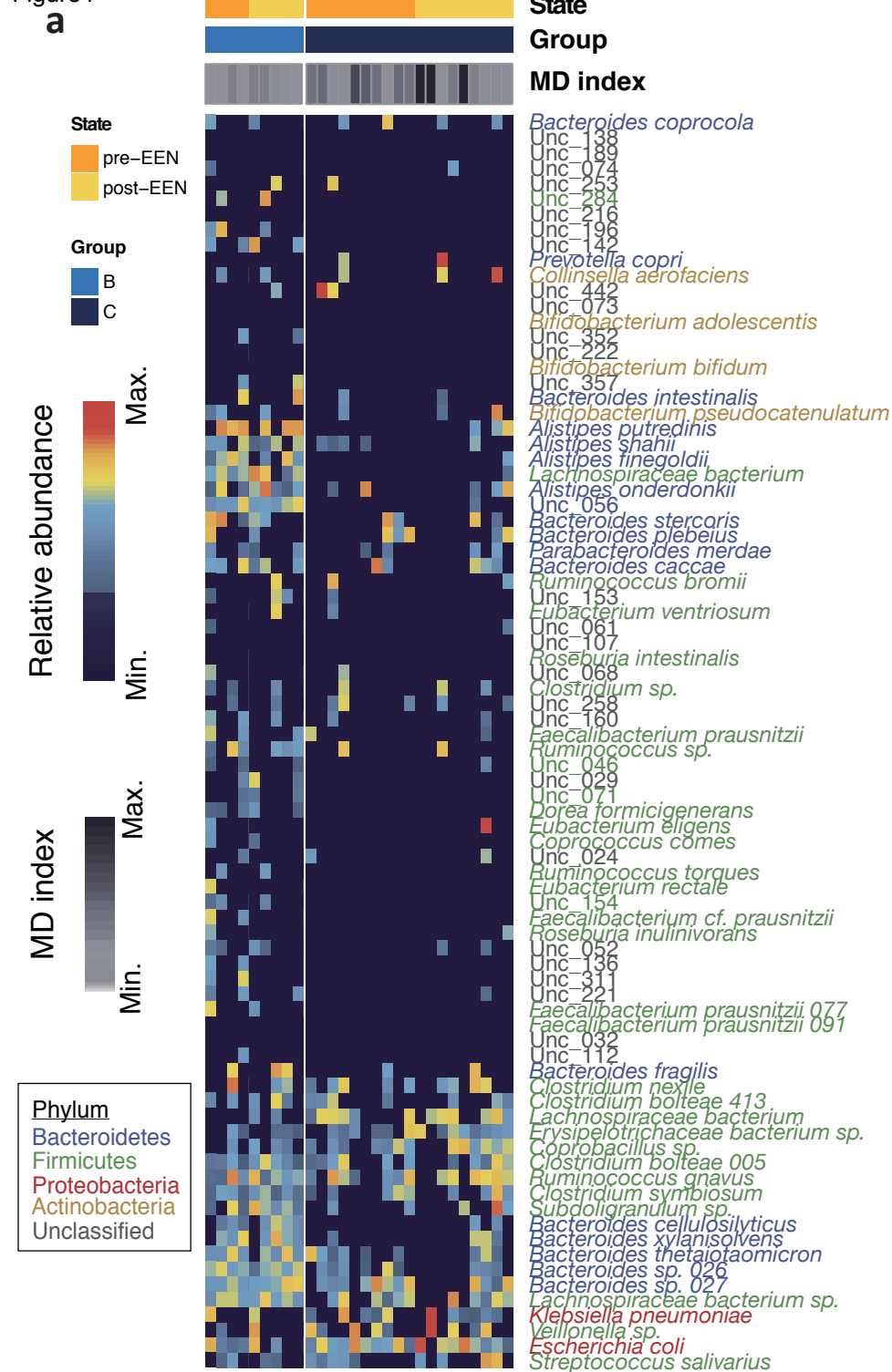
Co-occurring



Co-excluding



Figure 4

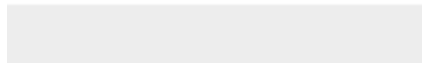




Click here to access/download

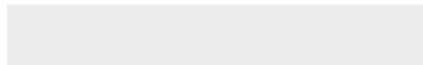
Supplementary Material

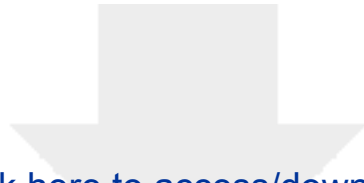
Table & Supplementary Table.CD.xlsx





Click here to access/download
Supplementary Material
SI-CD-paper-gigascience.docx





[Click here to access/download](#)

Supplementary Material

2017-04-17 GIGA-D-17-00073 rebuttal.docx

