1 **Format: Article**

2 **Prevalent human gut *Enterocloster* spp. produce urolithin A from its dietary precursor**
3 **using a molybdenum-dependent dehydroxylase**

4 Reilly Pidgeon[1], Sacha Mitchell[1], Michael Shamash[2], Layan Suleiman[1], Lharbi Dridi[1], Corinne
5 Maurice[2], Bastien Castagner[1*]

6 [1] Department of Pharmacology & Therapeutics, McGill University, 3655 Prom. Sir-William-
7 Osler, Montreal, Quebec, H3G 1Y6, Canada.

8 [2] Department of Microbiology & Immunology, McGill University, 3775 University Street,
9 Montreal, Quebec, H3A 2B4, Canada

10 **Abstract**

11 Urolithin A is a polyphenol derived from the multi-step metabolism of dietary ellagitannins by
12 the human gut microbiota that can affect host health. Most, but not all, individuals harbor a
13 microbiota capable of urolithin A production; however, the enzymes that dehydroxylate its
14 dietary precursor (urolithin C) are unknown. Here, we use a combination of untargeted
15 transcriptomics, proteomics, and comparative genomics to uncover an inducible urolithin C
16 dehydroxylase (*ucd*) operon in *Lachnospiraceae*. Biochemical characterization and structural
17 modeling of proteins encoded by *ucd* demonstrated that urolithin C dehydroxylation was NADH-
18 dependent, required strict anaerobic conditions, and used urolithin C as an electron acceptor.
19 By combining functional assays in complex fecal communities with publicly available
20 metagenomic data, our study reveals that both urolithin C-metabolizing bacteria and *ucd* operon
21 genes are widely distributed in gut metagenomes and likely comprise keystone species in the
22 metabolism of urolithins by the human gut microbiota.

23 **Main**

24 The human gut microbiota is a collection of trillions of microorganisms that colonize the
25 intestinal tract and play a pivotal role in host health and disease [1]. Gut bacteria help maintain
26 homeostasis by regulating host immune cell activity, gut barrier integrity, and nutrient availability
27 [2]. One of the main mediators of microbiota-host interactions are microbial metabolites. Gut
28 bacteria possess an immense metabolic repertoire (nearly 1000-fold more protein coding
29 sequences than the human genome [3]) to perform four main classes of reactions: hydrolysis,
30 conjugation, cleavage and reduction [4-7]. These reactions are ubiquitous in the gut environment
31 have been linked to microbiota-dependent activities whereby bacteria can alter the availability
32 and efficacy of therapeutic drugs [8-10], modify host bile acids [6,11,12], and metabolize diet-derived
33 compounds [13-15].

34 Diet is a strong modulator of the composition and function of the gut microbiota [16-19]. Diet-
35 derived polyphenols are a diverse class of plant secondary metabolites found in fruits,
36 vegetables, and nuts (reviewed in [20]) that are poorly absorbed by the host and reach the large
37 intestine relatively intact [7,21]. Ellagitannins are a sub-group of polyphenols that belong to the
38 family of hydrolysable tannins and are characterized by a central glucose (open-chain or cyclic
39 forms) linked to a variety of pyrogallol-like moieties [20]. Depending on microbiota composition,
40 ellagitannins can be hydrolyzed and reduced by bacteria into bioactive metabolites (e.g., ellagic
41 acid, urolithins, nasutins) according to different metabolic phenotypes characterized by the
42 terminal metabolites observed in biological fluids [22] (Extended Data Fig. 1).

43 Urolithin A (uroA) is the most common terminal metabolite of ellagitannin metabolism [23]. Clinical
44 trials in healthy individuals have demonstrated that uroA is safe, bioavailable, and can be
45 detected in its aglycone, glucuronidated, and sulfated forms in plasma [23,24]. Once absorbed,

46 uroA can impact host health by improving mitochondrial function (mitophagy) [24-26], protecting
47 gut barrier integrity [27,28], and inhibiting inflammation [29]. Most notably, uroA modulates immune
48 function through the activation of Pink1 and aryl hydrocarbon receptor (AhR) signaling
49 pathways, thus showing promise as a postbiotic to treat age-related conditions [25,27,30,31].

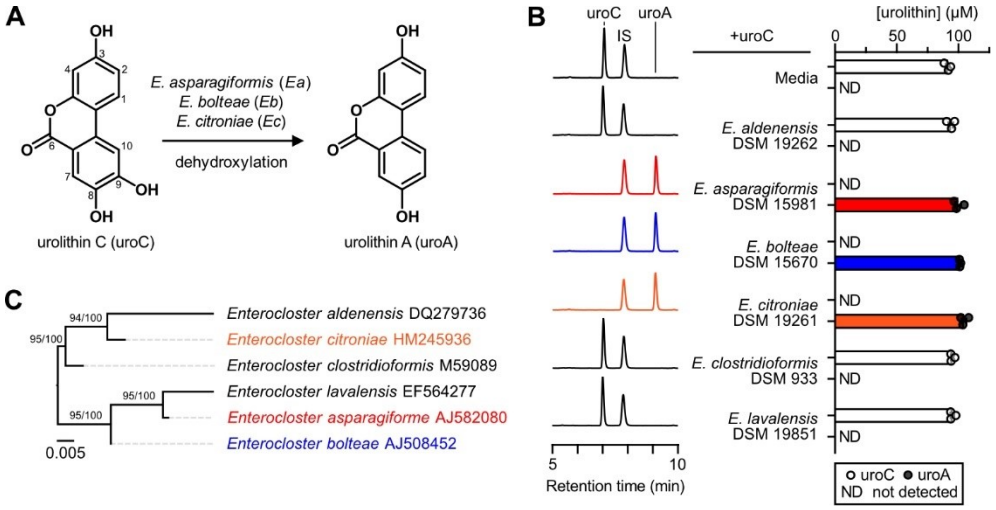
50 While urolithin metabolism is prevalent in human populations, only few gut bacteria have
51 reported urolithin metabolism activity [32-34]. Most urolithin metabolizers belong to the
52 *Eggerthellaceae* family (*Gordonibacter urolithinfaciens*, *Gordonibacter pamelaeae*, *Ellagibacter*
53 *isourolithinifaciens*) and can perform multiple metabolic steps in the urolithin metabolism
54 pathway, yielding either uroC or isourolithin A (isouroA) from ellagic acid [35]. Recently, certain
55 members of the *Enterocloster* spp. (*Lachnospiraceae* family) were reported to dehydroxylate
56 urolithin C (uroC) to uroA and isouroA to urolithin B (uroB) both *in vitro* and *in vivo* [36,37]. These
57 findings shed light on the minimal bacterial community required for complete metabolism of
58 ellagic acid to uroA; however, the genes and enzymes that catalyze these dehydroxylation
59 reactions remain unknown (Fig. 1A).

60 Here, we used a multi-omics enzyme identification framework to uncover uroC dehydroxylase
61 (*ucd*) genes and enzymes in *Enterocloster* spp. and their relative, *Lachnoclostridium pacaense*.
62 We found that the Ucd enzyme complex specifically dehydroxylates 9-hydroxy urolithins and
63 that both metabolizing species and *ucd* genes are prevalent and actively transcribed in human
64 feces during metabolism. Our study sheds light on the genetic basis underlying the complex
65 reciprocal interactions between urolithins, the gut microbiota, and the host.

66 **Results**

67 **A subset of *Enterocloster* species converts urolithin C to urolithin A *in vitro*.**

68 Members of the *Enterocloster* spp. have previously been shown to metabolize
69 (dehydroxylate) uroC *in vitro* [36] and *in vivo* [37] under anaerobic conditions (Fig. 1A, full metabolic
70 pathway in Extended Data Fig. 1). To determine the prevalence of uroC metabolism within this
71 genus, we incubated all available *Enterocloster* spp. type strains (Methods) with uroC and
72 quantified urolithin concentrations by liquid chromatography-mass spectrometry (LC-MS). Of
73 the tested bacteria, only *E. asparagiformis*, *E. bolteae*, and *E. citroniae* dehydroxylated uroC to
74 produce uroA (Fig. 1B). Interestingly, uroC metabolism was not predicted by phylogeny, as
75 uroC-metabolizing species did not cluster based on 16S rRNA genes, genomes, or proteomes
76 (Fig. 1C, Extended Data Fig. 2A,B), suggesting gain or loss of metabolic gene clusters (MGCs)
77 throughout the evolution of *Enterocloster* spp. Based on these results, we chose to perform
78 more in-depth analysis on *E. asparagiformis* and *E. bolteae* to identify the MGCs involved in
79 uroC dehydroxylation.



80

81 **Figure 1. Urolithin C metabolism by *Enterocloster* spp. is not predicted by phylogeny.**

82 **A)** Reaction scheme of uroC dehydroxylation by gut resident *Enterocloster* spp. via unknown
83 enzymes. **B)** LC-MS screen of *Enterocloster* spp. type strains for dehydroxylation activity. UroC

84 (100 μM) was added to cultures (in mABB+H media) at the start of growth and urolithins were
85 extracted after 24 h anaerobic incubation, then analyzed by LC-MS. Left: Representative
86 chromatograms (λ = 305 nm) for each experimental group (from one representative biological
87 replicate). The same scale was used for each chromatogram. Right: Quantification of urolithin
88 peak areas relative to a salicylic acid internal standard (IS) (n = 3 biological replicates). Data
89 are represented as mean ± SEM. **C)** Phylogenetic tree of tested *Enterocloster* spp. type strain
90 16S rRNA sequences constructed using the Genome-to-Genome Distance Calculator (GGDC)
91 Phylogeny Server [55]. Maximum likelihood (ML) tree inferred under the GTR+GAMMA model
92 and rooted by midpoint-rooting. The branches are scaled in terms of the expected number of
93 substitutions per site. The numbers above the branches are support values when larger than
94 60% from ML (left) and maximum parsimony (right) bootstrapping. The GenBank accession
95 numbers are provided to the right of each taxon. Source data and statistical details are provided
96 as a Source data file.

97

98 **A putative urolithin C dehydroxylase metabolic gene cluster is upregulated upon**

99 **urolithin C treatment.**

100 To understand when uroC metabolism machinery was being expressed, we first sought
101 to characterize the kinetics of uroC dehydroxylation. Therefore, a simultaneous growth and
102 metabolism experiment was designed, whereby uroC was spiked into cultures during the
103 exponential phase of growth and metabolites were measured by LC-MS (Fig. 2A). Treatment
104 with uroC did not affect growth of either bacterium compared to the DMSO control (Fig. 2B). In
105 both bacteria, quantitative conversion of uroC to uroA occurred within 4 h post-spike-in (Fig.
106 2C,D), demonstrating that metabolism in rich media is fast and robust.

107 We next sought to determine whether uroC metabolism is inducible or constitutive. To test for
108 inducibility, both bacteria were treated with DMSO or uroC during exponential growth, then
109 washed and resuspended in PBS, yielding resting cells unable to synthesize new proteins.
110 Metabolism of uroC to uroA was inducible as only resting cells originating from bacteria grown
111 with uroC were capable of uroA production (Fig. 2E). We therefore performed RNA-sequencing
112 to compare gene expression in DMSO and uroC-treated cultures of *E. asparagiformis* and *E.*
113 *bolteae*. Since uroA was detected in both bacterial cultures 2 h post spike-in (Fig. 2C), this

timepoint was selected to isolate mRNA. RNA sequencing of uroC-induced cultures revealed a distinct MGC, which we term the uroC dehydroxylase (*ucd*) cluster, that was highly and differentially expressed ($log_2FC > 8$) in both *E. asparagiformis* (Fig. 2F) and *E. bolteae* (Fig. 2G). In both bacteria, these MGCs contained adjacent genes that were expressed to similar log_2FC values: a xanthine dehydrogenase family protein subunit M, a (2Fe-2S)-binding protein, and a molybdopterin-dependent oxidoreductase (Fig. 2F,G insets). These genes, which belong to the xanthine dehydrogenase family of proteins, will hereafter be referred to as *ucdC* (for coenzyme), *ucdF* (for ferredoxin), and *ucdO* (for oxidoreductase), respectively.

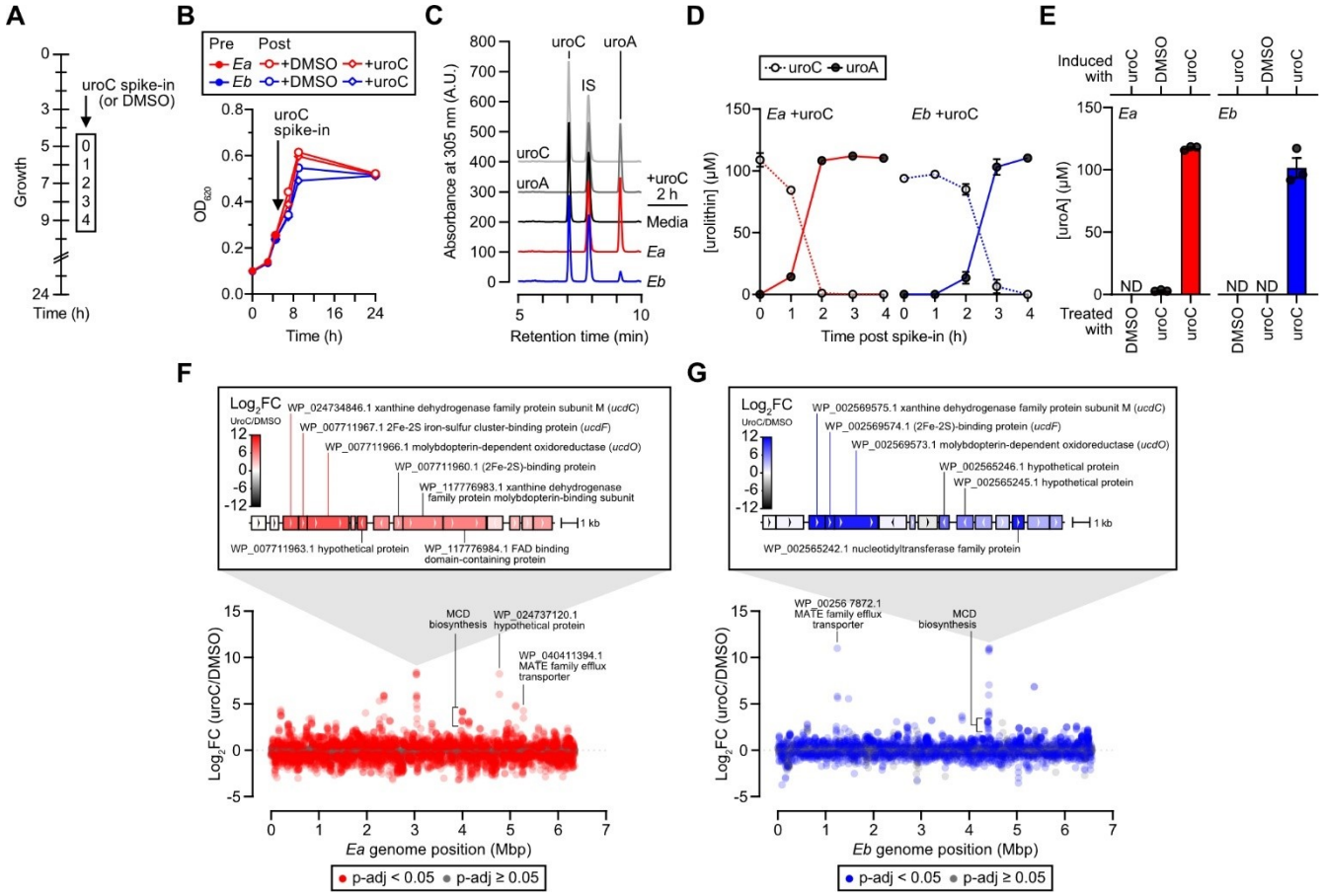


Figure 2. Urolithin C treatment upregulates a putative dehydroxylase metabolic gene cluster.

A) Experimental design of uroC (100 μM) spike-in experiments during the exponential phase of growth. For each biological replicate in this design, growth (B), metabolism (C,D), and RNA-

127 seq (F-G) results are matched. **B)** Growth curve (optical density (OD) at 620 nm) of DMSO or
128 uroC-spiked *E. asparagiformis* (*Ea*) and *E. bolteae* (*Eb*) type strain cultures according to the
129 design in (A). 200 µL of culture were sampled at each timepoint and OD_{620} was measured in a
130 96-well plate (n = 4 biological replicates). The same sampled culture was then frozen and
131 extracted for analysis by LC-MS (C,D). **C)** Representative chromatograms (λ = 305 nm) of
132 cultures sampled 2 h post-spike-in (from one representative biological replicate). The same
133 scale was used for each chromatogram. **D)** Quantification of urolithin concentrations from peak
134 areas relative to a salicylic acid internal standard (IS) over 4 h in uroC-spiked *Ea* and *Eb* type
135 strain cultures (n = 4 biological replicates). **E)** Quantification of urolithin A concentrations in
136 DMSO- or uroC-treated *Ea* and *Eb* resting cell suspensions. Resting cell suspensions were
137 prepared from *Ea* and *Eb* cells grown with either DMSO or 50 µM uroC. The cells were washed
138 and resuspended in PBS to halt the production of new enzymes, then treated with DMSO or
139 uroC (n = 3 biological replicates). **F,G)** Manhattan plots of genes altered by uroC treatment in
140 **F)** *Ea* and **G)** *Eb* based on DESeq2 analysis (n = 4 biological replicates). Data points are colored
141 according to their adjusted p-value (based on the Benjamini-Hochberg-corrected Wald
142 statistic). Grey, p-adj ≥ 0.05. Red or blue, p-adj < 0.05 for *Ea* and *Eb*, respectively. The most
143 highly and differentially expressed genes by RNA-seq are depicted in the insets, which show
144 the genomic organization around the differentially expressed genes (generated from the NCBI
145 Sequence Viewer). Genes in the insets are colored according to their log_2FC values. NCBI
146 accessions for select proteins encoded by highlighted genes are provided. Data are
147 represented as mean ± SEM in (B,D,E). FC, fold change (uroC/DMSO); Source data and
148 statistical details are provided as a Source data file.

149

150 **Presence of *ucd* metabolic gene cluster homologs in genomes predicts urolithin C**

151 **metabolism by gut bacteria.**

152 We wondered whether novel metabolizers of uroC could be discovered based on

153 nucleotide sequence homology to the *ucd* MGC. Homology searches using the *E. bolteae ucd*

154 MGC sequence confirmed that uroC-metabolizing *Enterocloster* spp. (*E. asparagiformis*, *E.*

155 *bolteae*, and *E. citroniae*) possessed homologs of the *ucdCFO* genes with a similar organization

156 (Extended Data Fig. 3). In addition, the gut bacterium *Lachnoclostridium pacaense* (or

157 *Clostridium pacaense*) [38] was identified as a hit (Extended Data Fig. 3). The type strain of this

158 bacterium (CCUG 71489T = Marseille-P3100) was closely related to *Enterocloster* spp. based

159 on 16S rRNA, whole genome, and whole proteome phylogenies (Fig. 3A, Extended Data Fig.

160 2A,B, respectively). *L. pacaense* possessed genomic sequences with high homology (86.5%

161 nucleotide identity) and identical functional annotations to the *E. bolteae ucd* MGC sequence

(Fig. 3B). When incubated with uroC, *L. pacaense* CCUG 71489T quantitatively produced uroA (Fig. 3C,D). Other urolithin and catechol-metabolizing bacteria belonging to the *Eggerthellaceae* family lack uroC dehydroxylase activity [33,35], which correlates with the absence of *ucd*-like operons in their genomes (Fig. 3B). These comparative genomics data indicate that the presence of a *ucd* MGC in genomes predicts uroC metabolism by gut bacteria.

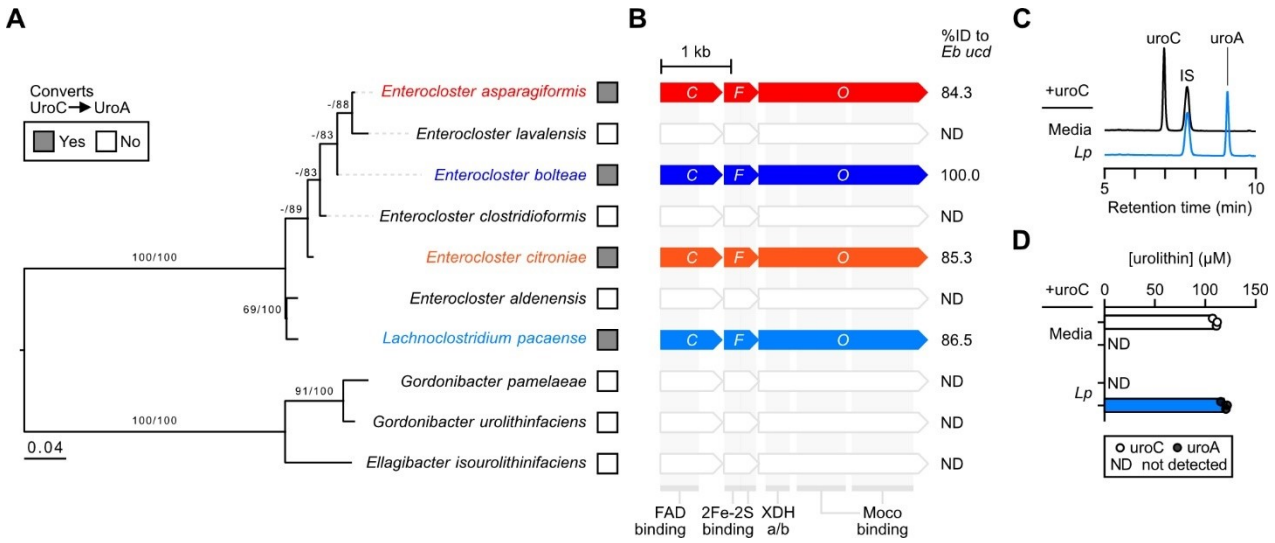


Figure 3. Urolithin C metabolism correlates with *ucd* metabolic gene cluster prevalence in gut bacteria.

A) Phylogenetic tree of *Enterocloster* spp., *Lachnoclostridium pacaense*, and catechol-metabolizing *Eggerthellaceae* type strain 16S rRNA sequences constructed using the Genome-to-Genome Distance Calculator (GGDC) Phylogeny Server [55]. Maximum likelihood (ML) tree inferred under the GTR+GAMMA model and rooted by midpoint-rooting. The branches are scaled in terms of the expected number of substitutions per site. The numbers above the branches are support values when larger than 60% from ML (left) and maximum parsimony (right) bootstrapping. Bacteria that convert uroC to uroA are labeled with grey squares. **B)** NBCI Multiple Sequence Aligner viewer hits for BLASTn searches using the *E. bolteae* DSM 15670 *ucd* MGC nucleotide sequence as a query against the NCBI refseq_genomes database. Only hits with ≥ 90 % query coverage and species-level taxonomic resolution are displayed with % identity to the query sequence. Domain annotations for each gene are denoted below according to InterPro annotations for corresponding proteins. ND, Not detected; Moco, Molybdenum cofactor. **C,D)** *In vitro* metabolism of uroC by *Lachnoclostridium pacaense* CCUG 71489T. UroC (100 μM) was added to cultures (in mABB+H media) at the start of growth and urolithins were extracted after 24 h anaerobic incubation, then analyzed by LC-MS. **C)** Representative chromatograms (λ = 305 nm) for each experimental group (from one representative biological replicate). The same scale was used for each chromatogram. **D)** Quantification of urolithin peak areas relative to a salicylic acid internal standard (IS) (n = 3 biological replicates). Data are

188 represented as mean ± SEM. Source data and statistical details are provided as a Source data
189 file.
190

191 **The *E. bolteae ucd* metabolic gene cluster is organized in an operon.**

192 We next sought to characterize the genomic organization of the *ucd* MGC in *E. bolteae*.
193 Based on the proximity, sense, and expression levels of each of the three genes by RNA-seq
194 (Fig. 2G), we hypothesized that all three genes in the MGC were organized in an operon. We
195 designed a gene-specific RT-PCR assay that would enable the detection of full-length
196 polycistronic *ucdCFO* genes using cDNA from DMSO- or uroC-treated *E. bolteae* as a template
197 (Fig. 4A). An amplicon of the expected size (~3.6 kb) was detected only in cDNA derived from
198 uroC-treated *E. bolteae*, validating the inducibility of these genes (Fig. 4B). Long-read
199 sequencing of the obtained amplicon yielded a sequence corresponding to the *E. bolteae*
200 *ucdCFO* MGC with 100% identity (Supplementary Information). Using an independent set of *E.*
201 *bolteae* cultures, we then performed RT-qPCR on DMSO- or uroC-treated *E. bolteae* cultures
202 with all three genes in the putative operon as targets. Similar to our RNA-seq results, all three
203 genes were significantly induced (Fig. 4C) relative to DMSO controls and were expressed at
204 the same levels (Fig. 4D). These results indicate that the *ucdCFO* genes are transcribed as a
205 single polycistronic mRNA and therefore form a uroC-inducible operon.

206 **The *E. bolteae ucd* operon is induced by other 9-hydroxy urolithins**

207 Next, we aimed to determine the substrate scope of the *ucd* operon. Multiple urolithins
208 possess pyrogallol, catechol, and hydroxyl structural motifs that are dehydroxylated at various
209 positions by gut bacteria (Fig. 4E, Extended Data Fig. 1). Interestingly, *E. bolteae* only
210 metabolizes the 9-position hydroxyl group of urolithins and does not require adjacent hydroxyl
211 groups for dehydroxylation (Fig. 4F,G, Extended Data Fig. 4A,B). Since dehydroxylation of

urolithins in *E. bolteae* is position specific, we hypothesized that the *ucd* operon would also be induced by these substrates. We performed RT-qPCR on DMSO-, uroM6-, uroC-, or isouroA-treated *E. bolteae* cultures using the *ucdO* gene as a target. While each urolithin significantly induced the expression of the *ucd* operon compared to DMSO (Fig. 4H), no significant difference in expression was observed between urolithin treatments (Fig. 4I). In addition, *E. bolteae* resting cell suspensions induced with uroC were capable of dehydroxylating uroM6 and isouroA, indicating that the same proteins induced by uroC can metabolize structurally similar 9-hydroxy urolithins (Extended Data Fig. 4C,D). Thus, it is likely that the same metabolic enzymes, encoded by the *E. bolteae ucd* operon, are acting on 9-hydroxy urolithins.

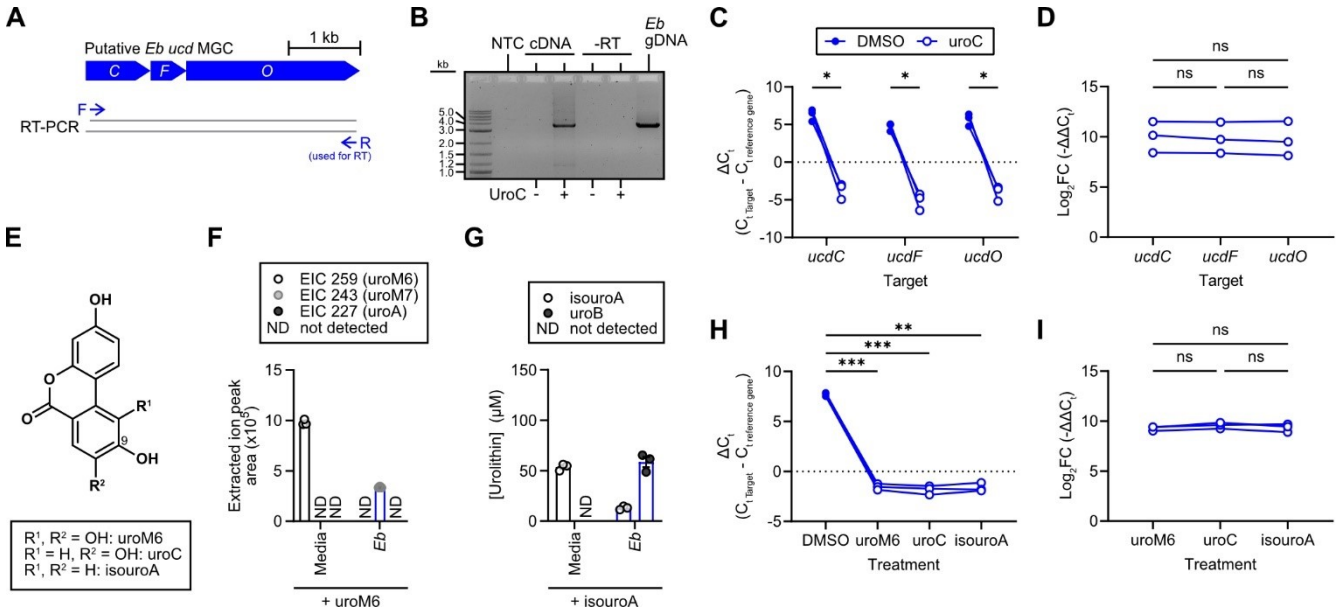


Figure 4. The *E. bolteae ucd* is organized in a three-gene operon that is induced by 9-hydroxy urolithins.

A) Primer design for RT-PCR (B) experiment targeting the *E. bolteae* (*Eb*) *ucd* MGC. The same reverse primer was used for both the reverse transcription step and the subsequent PCR reaction. **B)** 1% agarose gel image of RT-PCR amplicons using primers (A) that span the full-length *Eb ucd* MGC (from one biological replicate). Growing *Eb* cultures were treated with DMSO or uroC (100 μM) for 2 h before RNA isolation and reverse transcription. NTC, no template control. **C,D)** RT-qPCR expression of each gene in the *Eb ucd* MGC: molybdopterin-dependent oxidoreductase (*ucdO*), (2Fe-2S)-binding protein (*ucdF*), xanthine dehydrogenase family protein subunit M (*ucdC*). Growing *Eb* cultures were treated with DMSO or uroC (100

232 μM) for 2 h before RNA isolation and reverse transcription (n = 3 biological replicates). **C)**
233 Differential gene expression comparing DMSO vs. uroC is displayed as target-specific ΔCt ($C_{t\ \text{Target Gene}}$ - $C_{t\ \text{dnaK Reference Gene}}$) values. Data are presented as individual ΔCt values with lines
234
235 connecting paired biological replicates; multiple paired t tests with Benjamini, Krieger, and
236 Yekutieli FDR correction (Q set to 0.05); *, $q < 0.05$. **D)** Gene expression profiles of each target
237 gene in the *Eb ucd* MGC displayed as $\log_2$FC (equivalent to $-\Delta\Delta C_t$, where $\Delta\Delta C_t = \Delta C_{t\ \text{uroC}} - \Delta C_t$
238 $_{\text{DMSO}}$). Data are presented as individual $\log_2$FC values with lines connecting paired biological
239 replicates; repeated-measures one-way ANOVA with Tukey's multiple comparisons test; ns,
240 not significant. **E)** Chemical structures of urolithin M-6, urolithin C, and isourolithin A highlighting
241 the common 9-position hydroxyl group that gets dehydroxylated. **F)** Quantification of extracted
242 ion chromatogram (EIC) peak areas from *Eb* cultures sampled after 24 h of growth with 100 μM
243 uroM6 (n = 3 biological replicates). **G)** Quantification of urolithin peak areas from *Eb* cultures
244 sampled after 24 h of growth with 100 μM isouroA (n = 3 biological replicates). **H,I)** RT-qPCR
245 expression of the *Eb ucdO* gene. Growing *Eb* cultures were treated with DMSO, uroM6, uroC,
246 or isouroA (100 μM) for 2 h before RNA isolation and reverse transcription (n = 3 biological
247 replicates). **H)** Differential *Eb ucdO* gene expression comparing DMSO, uroM6, uroC, and
248 isouroA is displayed as target-specific ΔCt ($C_{t\ \text{MoO Gene}}$ - $C_{t\ \text{dnaK Reference Gene}}$) values. Data are
249 presented as individual ΔCt values with lines connecting paired biological replicates (from the
250 same pre-spike culture); repeated-measures one-way ANOVA with Dunnett's multiple
251 comparisons test; **, $p < 0.01$; ***, $p < 0.001$. **I)** Gene expression profile of the *Eb ucdO* gene
252 in different urolithin treatment groups displayed as $\log_2$FC (equivalent to $-\Delta\Delta C_t$, where $\Delta\Delta C_t$ =
253 $\Delta C_{t\ \text{urolithin}} - \Delta C_{t\ \text{DMSO}}$). Data are presented as individual $\log_2$FC values with lines connecting
254 paired biological replicates; repeated-measures one-way ANOVA with Tukey's multiple
255 comparisons test. ns, not significant; FC, fold change; FDR, false discovery rate. Source data
256 and statistical details are provided as a Source data file. The uncropped and unannotated gel
257 image is provided in the Supplementary Information.

258

259 **A molybdopterin cofactor biosynthetic gene cluster is upregulated upon urolithin C**

260 **treatment.**

261 Molybdoenzymes are organized into 3 main superfamilies (DMSO reductases, sulfite
262 oxidases, and xanthine oxidases) that have specific cofactor requirements (bis-MGD, Moco,
263 MCD, respectively) [39]. In addition to the three genes in the *ucd* operon, we observed a
264 significant increase ($\log_2$FC ≥ 2.6) in 9 molybdopterin cytosine dinucleotide cofactor (MCD)
265 biosynthesis genes upon uroC treatment (Extended Data Fig. 5A). These 9 genes, which
266 recapitulate the function of 10 genes in *E. coli* [39], are involved in molybdenum cofactor
267 biosynthesis (*moaAC*, *mogA*, *moeA*), molybdate ion transport (*modABCE*), cytosine addition to
268 the molybdenum cofactor (*mocA*), and MCD cofactor insertion into the active site (*xdhC*)

269 (Extended Data Fig. 5B) [40]. All 9 genes cluster in the genomes of *E. asparagiformis* and *E.*
270 *bolteae* and are organized into 2 adjacent operons (Extended Data Fig. 5C) that are induced
271 upon uroC treatment. Based on sequence homology to *E. coli* oxidoreductases, proteins
272 encoded by the *ucd* operon belong to the xanthine dehydrogenase (Xdh) family, which are
273 reported to bind an MCD cofactor in the oxidoreductase subunit (XdhB) of the active XdhAB
274 enzyme complex [40]. These findings imply that uroC dehydroxylation is MCD-dependent, which
275 differs from the bis-MGD requirement of catechol dehydroxylases in *Eggerthellaceae* [10,14].

276 **The UcdCFO complex enables anaerobic electron transport from NADH to uroC**

277 Since oxidoreductases utilise a variety of cofactors and coenzymes for catalytic activity,
278 we sought to determine the redox coenzymes and conditions necessary for uroC
279 dehydroxylation. Therefore, we performed metabolism assays using crude lysates from uroC-
280 induced *E. bolteae*. As crude lysates alone did not metabolize uroC (Fig.5A), various redox
281 coenzymes (NADPH, NADH, and FAD) were added to lysates to promote uroC
282 dehydroxylation. Only NADH-treated lysates yielded quantitative dehydroxylation of uroC to
283 uroA compared to the no cofactor control (Fig. 5A). Interestingly, the addition of free FAD
284 partially inhibited uroC dehydroxylation in NADH-treated lysates (Fig. 5A), likely by decreasing
285 the free NADH pool. NADPH, which differs from NADH by a phosphate group on the 2'-OH
286 group of the adenosine moiety, was unable to promote uroC dehydroxylation, indicating some
287 specificity in the redox cofactors necessary to sustain catalysis. Aerobic incubation of crude
288 lysates supplemented with NADH completely inhibited uroC dehydroxylation (Fig. 5B),
289 indicating that the active enzyme complex requires a strictly anaerobic environment for
290 dehydroxylation, as has been demonstrated for various metalloenzymes [41].

291 We attempted heterologous expression of the *E. bolteae ucd* operon in *E. coli* to validate
292 its function; however, all expression experiments were unsuccessful due to inclusion body

293 formation (data not shown). This lack of solubility and activity likely results from the choice of
294 heterologous host and from the complex assembly of active molybdoenzymes [40]. To confirm
295 that *ucd* operon-encoded proteins were expressed in *E. bolteae* crude lysates, we performed
296 untargeted proteomics and compared protein expression upon DMSO or uroC treatment.
297 Indeed, all 3 proteins encoded by the *ucd* operon (UcdC, UcdF, and UcdO) were the most
298 differentially expressed proteins in the uroC treatment group (Fig. 5C). In addition, proteins
299 involved in MCD biosynthesis were also strongly increased upon uroC treatment (Fig. 5C,D,
300 Extended Data Fig. 5A), pointing to the coordination between MCD biosynthesis and active
301 oxidoreductase assembly. These multi-omics datasets implicate all three genes and their
302 products in the metabolism of uroC to uroA, as demonstrated by the strong positive correlation
303 between transcript and protein differential expression (Fig. 5D).

304 To gain an understanding of the structural organization of proteins encoded by the *ucd*
305 operon, we performed modeling using AlphaFold2 [42,43]. Structures of each protein encoded by
306 the *E. bolteae ucd* operon (Fig. 5E) were superposed onto published X-ray crystal structures of
307 oxidoreductases with similar folds (using the UcdO amino acid sequence as a query) [44], yet
308 from different taxonomic domains: carbon monoxide dehydrogenase from *Afipia*
309 *carboxidovorans* [45] and xanthine dehydrogenase from *Bos taurus* [46]. The 3 proteins encoded
310 by the *ucd* operon formed subunits in an oxidoreductase complex with a similar quaternary
311 structure to the published crystal structures (Fig. 5F, Extended Data Fig. 6A,B). The predicted
312 quaternary structure of the UcdCFO enzyme complex supported a complete electron transport
313 chain whereby electrons would flow from reduced FAD to two 2Fe-2S clusters, then to the MCD
314 cofactor, and finally to uroC as the terminal electron acceptor (Fig. 5G, Extended Data Fig.
315 6C,D). Using homology modeling, we further identified the putative uroC binding site in UcdO,
316 which overlaps with the salicylic acid ligand in the *Bos taurus* xanthine dehydrogenase

(Extended Data Fig. 6E). This putative uroC binding site contains multiple tyrosine (Y375, Y538, Y624, Y632), tryptophan (W345), and phenylalanine (F458, F464) residues that could form π-π stacking interactions with uroC (Extended Data Fig. 6F), orienting it towards the molybdenum cofactor.

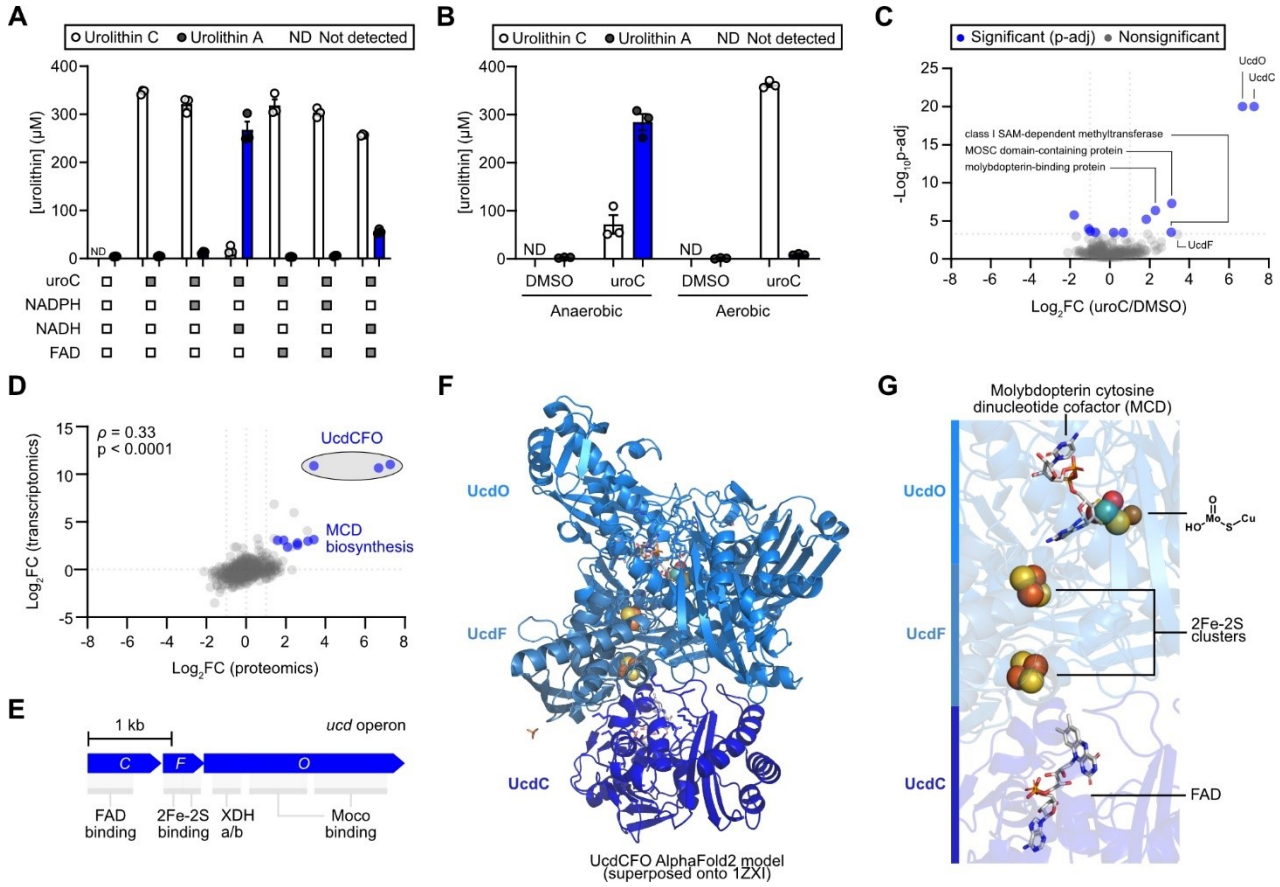


Figure 5. The UcdCFO complex enables anaerobic electron transport from NADH to uroC.

A) Quantification of urolithin peak concentrations in crude uroC-induced *Eb* lysates re-treated with DMSO or uroC (350 μM) and various coenzymes (n = 3 biological replicates). Data are represented as mean ± SEM. NADPH, nicotinamide adenine dinucleotide phosphate; NADH, nicotinamide adenine dinucleotide; FAD, flavin adenine dinucleotide. **B)** Quantification of urolithin peak concentrations in crude uroC-induced *Eb* lysates re-treated with DMSO or uroC (350 μM) and NADH in anaerobic or aerobic environments (n = 3 biological replicates). **C)** Volcano plot of untargeted proteomics analysis on DMSO or uroC-treated *Eb* (n = 3 biological replicates). Data points are colored according to their significance (Fisher's exact test with Benjamini-Hochberg correction for multiple comparisons). Grey, p-adj ≥ cutoff p-value (0.00048). Blue, p-adj < cutoff p-value (0.00048). **D)** Scatter plot showing the correlation

between gene and protein expression (log_2FC values) induced in uroC-treated *Eb* using the datasets in Fig. 2G and Fig. 5C, respectively. The non-parametric Spearman rank correlation test was used for statistical analysis. **E)** Domains of genes in the *ucd* operon based on InterPro annotations. **F)** Quaternary structure prediction of the proteins encoded by the *Eb ucd* operon. AlphaFold2 structures for each protein were superposed onto the X-ray crystal structure of PDB 1ZXI (carbon monoxide dehydrogenase from *Afipia carboxidovorans* OM5). **G)** Small molecule ligands from PDB 1ZXI in the superposed UcdCFO model form a complete electron transport chain from FAD to two 2Fe-2S clusters to a molybdopterin cytosine dinucleotide cofactor, which can then reduce uroC (terminal electron acceptor). Source data and statistical details are provided as a Source data file.

UroC-metabolizing species and *ucd* genes are prevalent and correlate with uroC metabolism in human fecal samples.

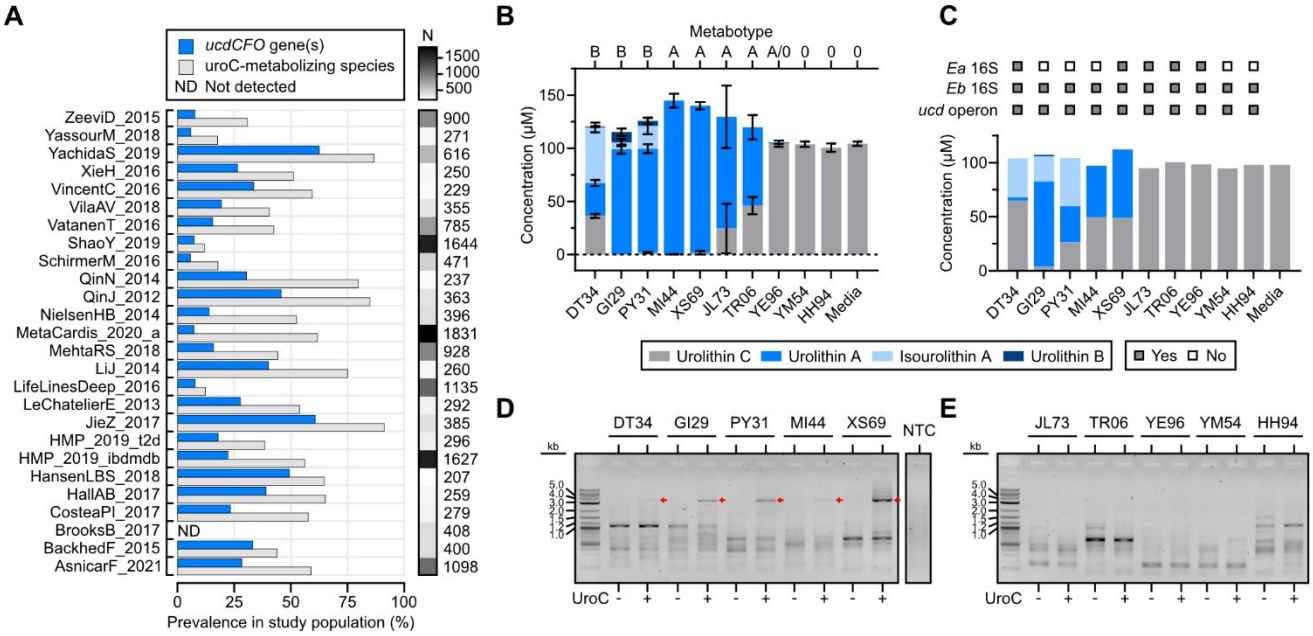
We next wondered whether uroC-metabolizing *Enterocloster* spp. and their *ucd* operons were prevalent and active in human fecal samples. We first utilized uniformly processed metagenomic data from the curatedMetagenomicData R package [47]. After filtering for fecal samples (86 studies, n = 21,030 subjects), we counted the prevalence of at least one uroC-metabolizing species and at least one *ucd* gene homolog (Methods). The prevalence of both features was variable across studies (Fig. 6A for studies with >200 participants, Extended Data Fig. 7A for all studies). Combining all studies, the prevalence of at least one uroC-metabolizing species and at least one *ucdCFO* gene homolog was 9,343/21,030 (44.9%) and 4,356/21,030 (20.7%), respectively. *E. bolteae* was the most prevalent and abundant uroC-metabolizing species detected in gut metagenomes (Extended Data Fig. 7B,C). These findings suggest that uroC-metabolizing *Enterocloster* spp. and *ucd* operon genes are prevalent in human fecal metagenomic samples and reflect the variable urolithin metabolism profiles (metabotypes) in the general population [23,48].

Next, we performed *ex vivo* metabolism assays to determine whether *Enterocloster* spp. could metabolize uroC in the context of a complex community. Fecal slurries from 10 healthy individuals were first profiled according to their uroC metabotypes (Fig. 6B) [22]. Individuals

363 clustered into metabotypes A (only uroA produced), B (uroA and isouroA/uroB), and 0 (no
364 terminal urolithin metabolites). While we observed metabotypes A and B in uroC-metabolizing
365 fecal slurries, all slurries produced some amount of uroA from uroC. Stools JL73, TR06, and
366 YE96 displayed variable metabolism patterns and did not metabolize uroC in some
367 experiments, likely reflecting differences in activity between aliquots of feces (Fig. 6B). We then
368 repeated metabolism assays using fecal slurries from all 10 healthy individuals and extracted
369 urolithins, DNA, and RNA from each culture. In this experiment, only 5/10 fecal slurries
370 metabolized uroC to uroA (Fig. 6C). We hypothesized that differences in metabotypes could be
371 explained by microbial composition. Therefore, V1-V9 16S rRNA sequencing was performed
372 on fecal slurries. Both DMSO- and uroC-treated fecal slurries within individuals had similar
373 microbial compositions but showed differences in composition between individuals (Extended
374 Data Fig. 8A,B). Surprisingly, all samples contained 16S rRNA sequences mapping to *E.*
375 *bolteae*, and many non-metabolizing fecal slurries contained *E. asparagiformis* (Fig. 6C,
376 Extended Data Fig. 8C). We then assayed the treated fecal slurries for the presence of the *ucd*
377 operon by PCR and found that 10/10 individuals (19/20 conditions) yielded a detectable
378 amplicon of the expected size (~3.6 kb) (Extended Data Fig. 8D,E). These data indicate that
379 the prevalence of uroC-metabolizing *Enterocloster* spp. 16S rRNA and *ucd* operon genes does
380 not predict metabolism in fecal samples .

381 We then surmised that the *ucd* operon would be transcribed only in fecal slurries actively
382 metabolizing uroC. Using a gene-specific reverse primer that binds to *ucdO*, the full-length *ucd*
383 operon was reverse transcribed and amplified in RNA extracted from DMSO- and uroC-treated
384 fecal slurries. An amplicon (~3.6 kb) corresponding to the *ucd* operon was only detected in
385 uroC-metabolizing fecal slurries when treated with uroC and entirely absent from non-
386 metabolizing slurries (Fig. 6D). This amplicon was absent in no reverse transcriptase controls,

387 indicating no gDNA contamination (Extended Data Fig. 8F,G). These data demonstrate that
388 *ucd* transcription correlates with uroC metabolism in complex fecal communities and that *E.*
389 *bolteae* is keystone species involved in urolithin A production.



390

391 **Figure 6. The *ucd* operon is prevalent in metagenomes and actively transcribed in**
392 **urolithin C-metabolizing human fecal samples.**

393 **A)** Prevalence of *ucd* operon (at least one gene) and of a uroC-metabolizing species (at least
394 one species) in fecal metagenomes from the CuratedMetagenomicData R package. Only
395 studies with ≥200 participants are depicted. All 86 studies are available in Extended Data Fig.
396 7. **B)** Summary of urolithin concentrations in fecal slurries (n = 10 healthy donors) incubated
397 with 100 μM uroC for 48 h. Data are representative of 3-6 experimental replicates. **C)** Summary
398 of urolithin concentrations in fecal slurries (n = 10 healthy donors) incubated with 100 μM uroC
399 for 48 h. Presence of uroC-metabolizing species and of the *ucd* operon is denoted above the
400 graph if the bacterium or operon was detected in the uroC-treated fecal slurry (see Extended
401 Data Fig. 8C,D). Data are representative of 1 replicate where DNA and RNA was also extracted
402 from fecal slurries. **D,E)** *ucd* gene-specific RT-PCR on fecal microbiota communities from 10
403 healthy donors using the primer set described in Fig. 4A. Samples are matched to the urolithin
404 metabolism data in C) and 16S rRNA sequencing in Extended Data Fig. 8A,C. 1% agarose gel
405 of amplicons derived from **D)** uroC-metabolizing and **E)** non-metabolizing fecal slurries. Bands
406 corresponding to the *E. bolteae ucd* operon (~3.6 kb) are labeled with red arrows. The no
407 template control (NTC) is the same for D,E). See Extended Data Fig. 8D,E for the no reverse
408 transcriptase control PCR reactions on the same samples. Source data are provided as a
409 Source data file. The uncropped and unannotated gel images are provided in the
410 Supplementary Information.

411 **Discussion**

412 Here, we identified genes and proteins that are essential for the metabolism of urolithins
413 by gut resident *Lachnospiraceae* through a combination of transcriptomics, comparative
414 genomics, and untargeted proteomics. Our study reveals a novel multi-subunit molybdoenzyme
415 (urolithin C dehydroxylase, Ucd) that catalyzes the dehydroxylation of 9-hydroxy urolithins like
416 uroM6, uroC, and isouroA. Importantly, prevalence analysis in published data and *ex vivo*
417 transcriptomics established *E. bolteae* as a keystone urolithin-metabolizing member of the gut
418 microbiota.

419 Catechol dehydroxylases are widespread in *Eggerthella lenta* and *Gordonibacter*
420 *pamelaeae* [10,14,49]. These molybdoenzymes, which belong to the DMSO reductase superfamily
421 (bis-MGD cofactor) [50], dehydroxylate substrates like catechol lignan (Cldh), DOPAC (Dodh),
422 dopamine (Dadh), hydrocaffeic acid (Hcdh), and caffeic acid (Cadh). The UcdCFO enzyme
423 complex differs from catechol dehydroxylases in *Eggerthellaceae* in important ways as it 1)
424 does not require a catechol structural motif for activity, 2) belongs to the xanthine oxidase
425 superfamily (MCD cofactor), and 3) is composed of 3 subunits (UcdCFO) that each lack signal
426 sequences. A recent survey of reductases in gut bacteria established that most respiratory
427 reductases contain N-terminal signal sequences and are translocated across the cytoplasmic
428 membrane, while non-respiratory reductases, which lack signal sequences, remain in the
429 cytoplasm [51]. Based on the absence of signal sequences and the cytoplasmic localization of
430 xanthine dehydrogenases, the *ucd* operon encodes for a non-respiratory reductase (see [51]
431 Supplementary Table 2).

432 In the host, UroA has reported pharmacological activities both within the gut environment
433 and systemically following absorption. In the gut, uroA can attenuate colitis by increasing the

434 expression of epithelial tight junction proteins [27] *via* the activation of AhR-Nrf2 pathways [28].
435 Additionally, uroA can enhance immunotherapy in colorectal cancer models by triggering Pink1-
436 dependent mitophagy pathways in T cells, improving anti-tumor CD8+ T cell immunity [25]. Once
437 absorbed by the host, uroA can trigger mitophagy in muscle cells, improving muscle function in
438 animal models of ageing and Duchenne muscular dystrophy [39,52]. Overall, uroA can enhance
439 gut barrier integrity, reduce inflammation, and promote mitochondrial health in the host.

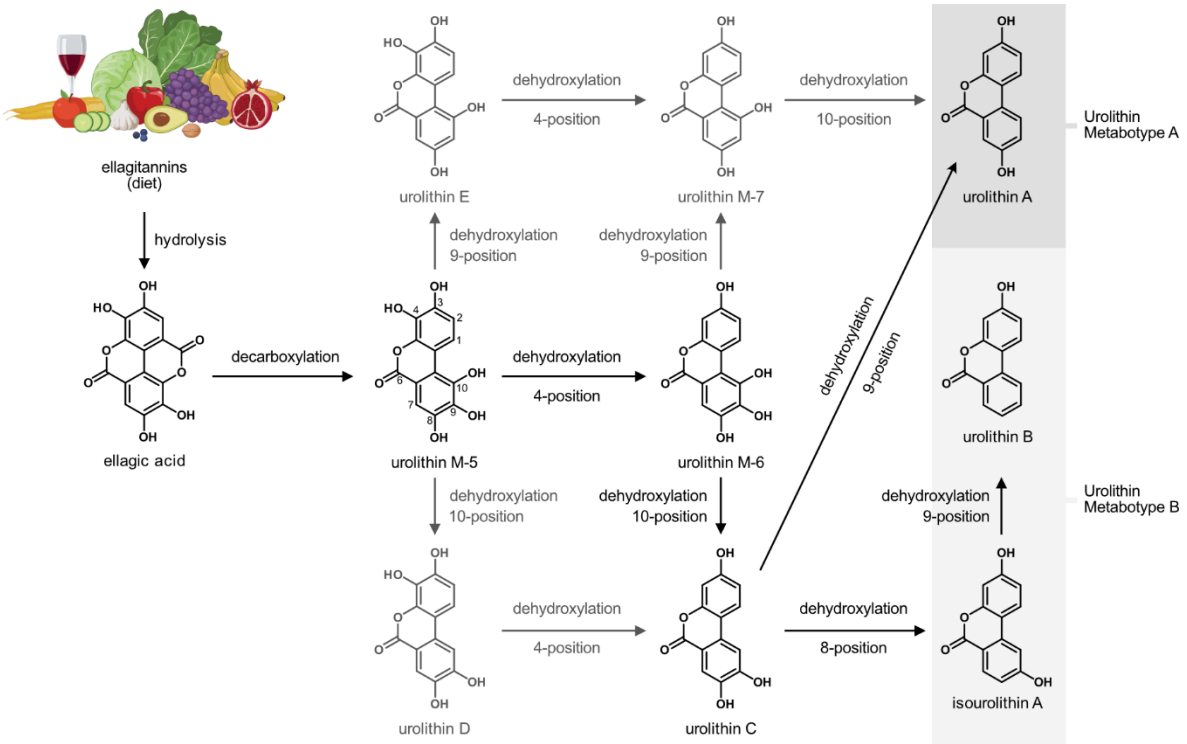
440 While uroA is the most common terminal metabolite following ellagitannin consumption
441 in humans, its production varies widely [34]. Interestingly, the ability of a fecal sample to produce
442 uroA from uroC did not correlate with the presence of widespread uroC-metabolizing
443 *Enterocloster* spp. or a *ucd* gene homologue, likely owing to poor viability or dead bacteria in
444 fecal samples. However, the active transcription of the *ucd* operon correlated perfectly with
445 metabotypes. By identifying the genetic basis for metabolism of uroC, we found a novel
446 metabolizing species that could not have been predicted based on phylogeny alone. Our data
447 suggests that *ucd*-containing *Enterocloster* spp., and the closely related *L. pacaense*, are the
448 main drivers of urolithin A production in the gut microbiota based on their prevalence in
449 metagenomes and activity in fecal samples. These findings further emphasize the importance
450 of functional assays such as transcriptomics and *ex vivo* metabolism to understand the
451 metabolism of xenobiotics by the gut microbiota.

452 Our study has several limitations. Similar to other molybdoenzyme studies [14], we were
453 unable to express a soluble, active UcdCFO enzyme complex in a heterologous host to validate
454 enzyme function. In addition, while we demonstrate that *Enterocloster* spp. actively
455 dehydroxylate uroC in human fecal samples, we cannot conclude that they are solely
456 responsible for this activity. Rare, strain-specific urolithin A production has been reported for
457 *Bifidobacterium pseudocatenulatum* INIA P815 [34], *Streptococcus thermophilus* FUA329 [53], and

458 *Enterococcus faecium* FUA027 [54], which may be a result of horizontal gene transfer since it is
459 not shared by other members of the taxa.

460 In conclusion, our studies reveal the genetic basis for urolithin A production by gut
461 bacteria and broaden our understanding of the molecular mechanisms underlying urolithin
462 metabotypes in human populations. Since diet can modulate gut microbiota function and host
463 health, elucidating the xenobiotic metabolism genes encoded by gut bacteria will be key to
464 developing dietary interventions targeting the gut microbiota.

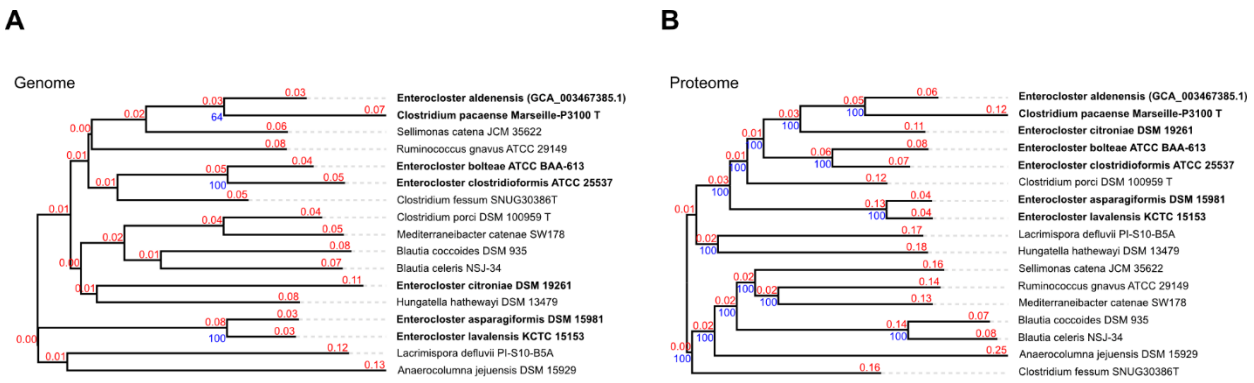
465 **Extended Data Figures**



466

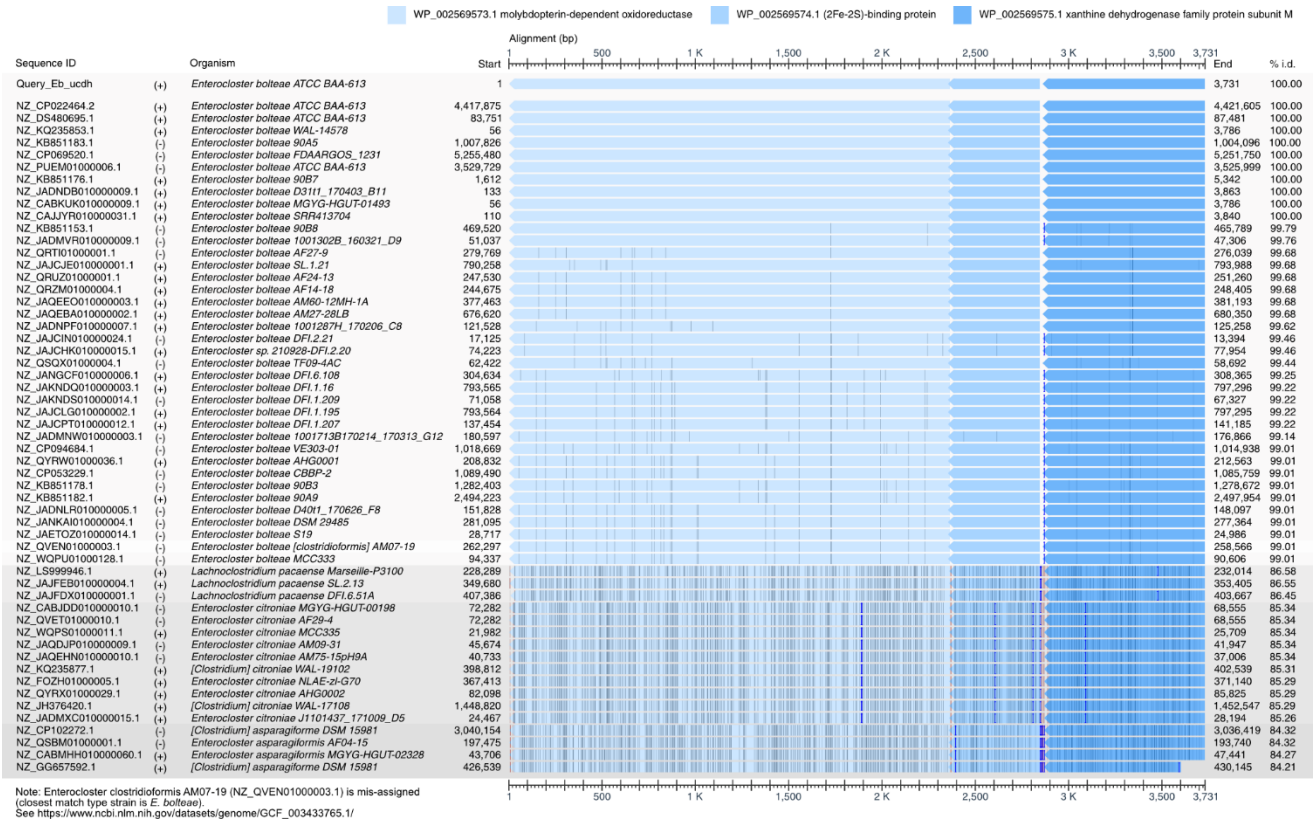
467 **Extended Data Figure 1. Ellagitannins are metabolized by gut bacteria.**

468 Reaction scheme of dietary ellagitannin metabolism by the human gut microbiota. Larger
469 ellagitannin structures are hydrolyzed during gut transit, releasing hexahydroxydiphenic acid,
470 which spontaneously lactonizes into ellagic acid. Once in the gut lumen, members of the
471 *Gordonibacter* spp. and *Ellagibacter isourolithinifaciens* can decarboxylate ellagic acid, forming
472 urolithin M-5. The resulting urolithin M-5 can be further dehydroxylated (at the 4,10- or 4,8,10-
473 positions) to uroC or isourolithin A by *Gordonibacter* spp. or *Ellagibacter isourolithinifaciens*,
474 respectively. Compounds colored in light gray are urolithin metabolites that are rarely observed
475 during *ex vivo* metabolism assays on ellagitannins. Once uroC or isouroA are produced,
476 *Enterocloster* spp. can further dehydroxylate the 9-position, yielding uroA or uroB, respectively.
477 The cartoon was generated in BioRender.



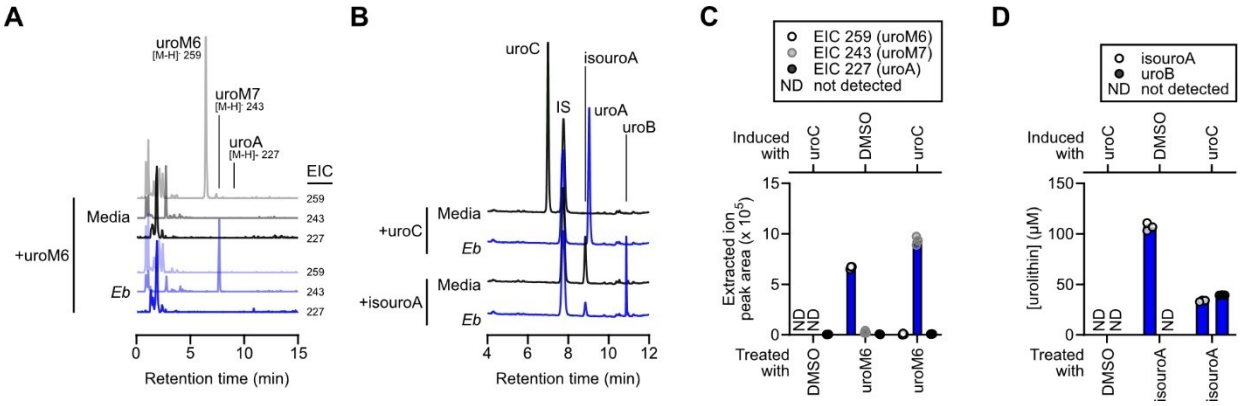
Extended Data Figure 2. *Lachnoclostridium pacaense* clusters with *Enterocloster* spp. in whole genome and proteome phylogenetic trees.

A) Whole genome phylogenetic tree of *Enterocloster* spp. The tree was inferred with FastME 2.1.6.1 from Genome BLAST Distance Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 36.8 %. The tree was rooted at the midpoint. **B)** Whole proteome phylogenetic tree of *Enterocloster* spp. The tree was inferred with FastME 2.1.6.1 from whole-proteome-based GBDP distances. The branch lengths are scaled via GBDP distance formula d5. Branch values are GBDP pseudo-bootstrap support values > 60 % from 100 replications, with an average branch support of 100.0 %. The tree was rooted at the midpoint.



Extended Data Figure 3. BLASTn searches using the *E. bolteae ucd* MGC genomic sequence identifies homologues in gut bacteria.

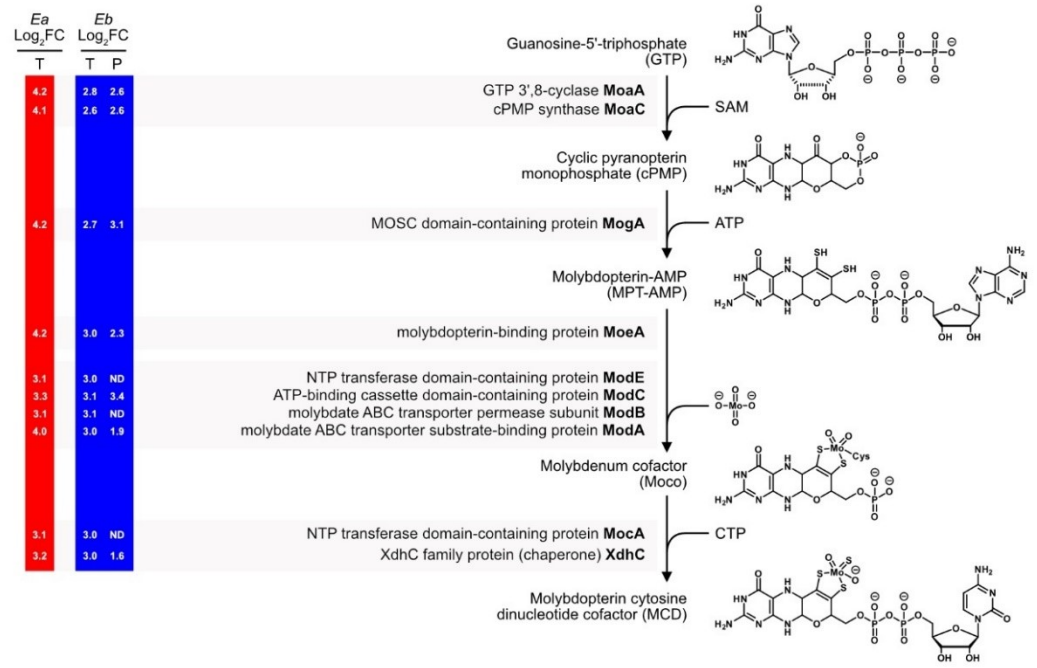
NBCI Multiple Sequence Aligner viewer hits for BLASTn searches using the *E. bolteae* DSM 15670 *ucd* MGC nucleotide sequence as a query against the NCBI refseq_genomes database (limited to Bacteria). Vertical lines in the sequence alignment represent nucleotide differences (show differences option) and insertions relative to the query sequence.



Extended Data Figure 4. *E. bolteae* dehydroxylates urolithins at the 9-position.

A) Representative extracted ion chromatograms (EIC) of *Eb* cultures sampled after 24 h of growth with 100 μM uroM6 (from one representative biological replicate). The same scale was used for each chromatogram. Analytes of interest and their respective $[M-H]^-$ m/z values are provided above the figure. **B)** Representative chromatograms (λ = 305 nm) of *Eb* cultures sampled after 24 h of growth with 100 μM uroC or isouroA (from one biological replicate). The same scale was used for each chromatogram. **C)** Quantification of extracted ion chromatogram (EIC) peak areas in DMSO- or uroC-treated *Eb* resting cell suspensions. Resting cell suspensions were prepared from *Eb* cells grown with either DMSO or 50 μM uroC. The cells were washed and resuspended in PBS to halt the production of new enzymes, then treated with 100 μM uroM6 (n = 3 biological replicates). **D)** Quantification of urolithin concentrations in DMSO- or uroC-treated *Eb* resting cell suspensions. Resting cell suspensions were prepared as described in C) (n = 3 biological replicates). For C-D, data are represented as mean ± SEM. ND, not detected.

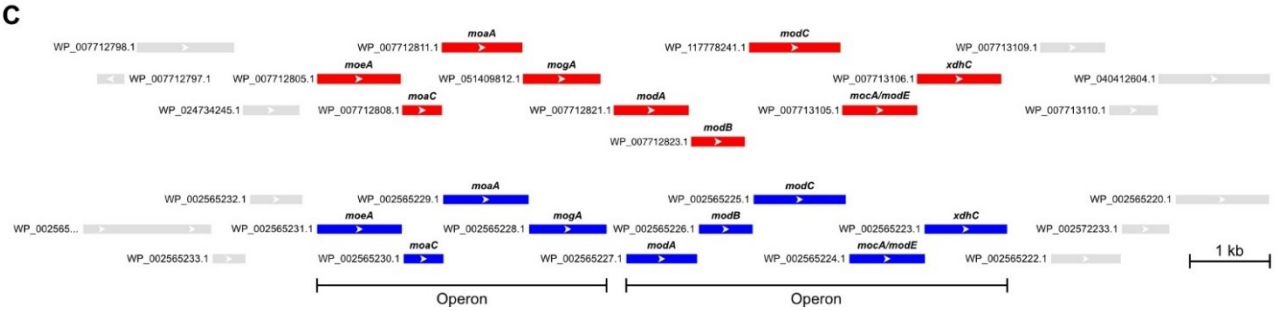
A



B

Protein annotation	Role(s)	*Ea* NCBI Accession	*Eb* NCBI Accession	#AA in *Eb*	Domain(s) [positions]
GTP 3',8-cyclase	**MoaA**	WP_007712811.1	WP_002565229.1	355	Radical SAM core [7-238]
Cyclic pyranopterin monophosphate synthase	**MoaC**	WP_007712808.1	WP_002565230.1	164	Molbdopterin cofactor biosynthesis C (MoaC) [14-149]
MOSC domain-containing protein	**MogA**	WP_051409812.1	WP_002565228.1	322	Molbdenum cofactor sulfurase C-terminal (MOSC) [28-153]
Molybdopterin-binding protein / Molybdopterin molybdenumtransferase	**MoeA**	WP_007712805.1	WP_002565231.1	352	MoaB/Mog [177-317]
NTP-transferase domain-containing protein / LysR family transcriptional regulator	**MocA/ModE**	WP_007713105.1	WP_002565224.1	313	MobA-like NTP transferase [5-160] Helix-turn-helix lysR-type [222-277]
ATP-binding cassette domain-containing protein	**ModC**	WP_117778241.1	WP_002565225.1	382	ABC transporter [2-259]
Molybdate ABC transporter permease subunit	**ModB**	WP_007712823.1	WP_002565226.1	223	ABC transmembrane type-1 [6-215]
Molybdate ABC transporter substrate-binding protein	**ModA**	WP_007712821.1	WP_002565227.1	294	Signal [1-25] Molybdate ABC transporter substrate-binding protein [26-294]
XdhC family protein	**XdhC**	WP_007713106.1	WP_002565223.1	343	XdhC- CoxI [12-70] XdhC Rossmann [189-331]

C

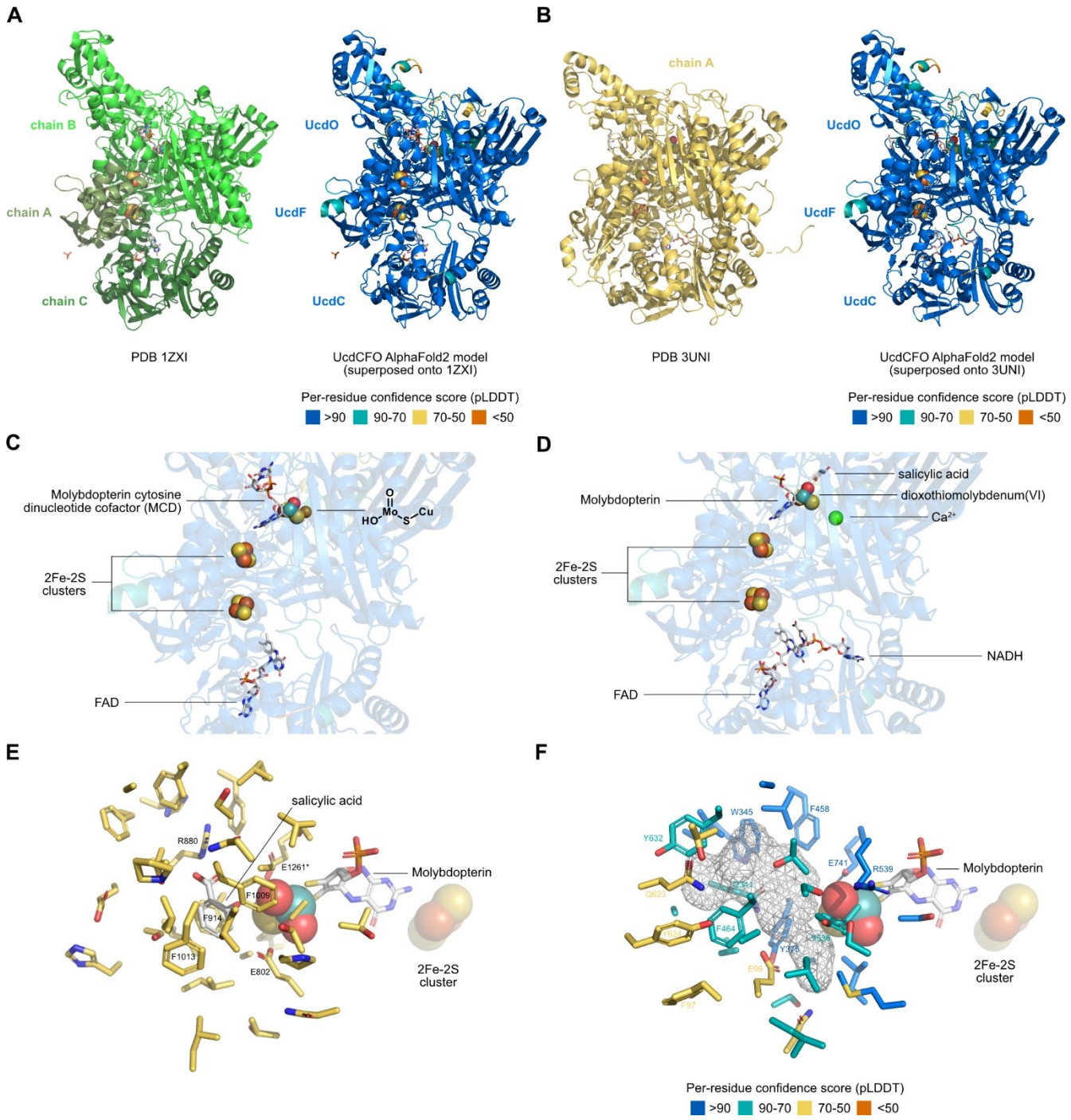


512

513 **Extended Data Figure 5. Urolithin C treatment upregulates molybdopterin cytosine**

514 **dinucleotide cofactor biosynthetic gene clusters.**

515 **A)** Molybdopterin cytosine dinucleotide (MCD) cofactor biosynthetic pathway. The $\log_2$FC
516 values of MCD cofactor biosynthetic genes upregulated by uroC in both transcriptomics (T) and
517 proteomics (P) datasets (when available) are provided to the left of the figure for both *Ea* and
518 *Eb*. **B)** Table of molybdenum cofactor biosynthetic genes found in the genomes of *Ea* and *Eb*.
519 Annotations are based on NCBI (GTF files) and UniProt gene/protein names. Roles are
520 assigned based on required proteins for molybdenum cofactors in the xanthine
521 oxidase/dehydrogenase family of enzymes. Accessions, primary sequence length, and
522 annotated domains (with positions within the primary sequence) are also provided. **C)** Genomic
523 organization of the MCD cofactor biosynthetic genes for *Ea* and *Eb* (generated from the NCBI
524 Sequence Viewer).

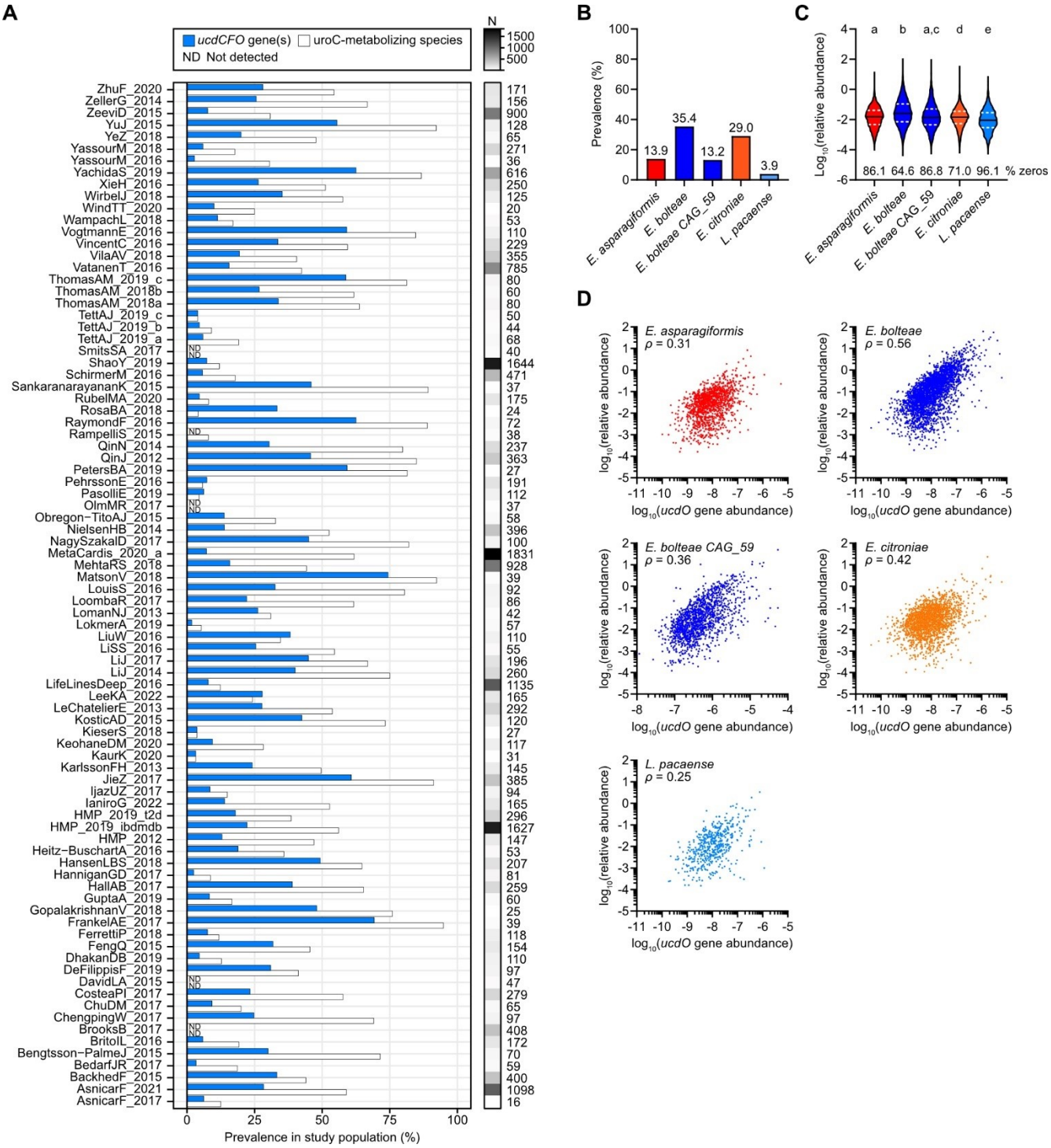


Extended Data Figure 6. The AlphaFold2 model of the *E. bolteae* UcdCFO enzyme complex has a similar quaternary structure to xanthine dehydrogenase superfamily crystal structures.

A,B) Structural superposition of AlphaFold2-predicted *Eb* Ucd proteins (right) onto the X-ray crystal structures of **A)** PDB 1ZXI (carbon monoxide dehydrogenase from *Oligotropha carboxidovorans* OM5) X-ray crystal structure (left) and **B)** PDB 3UNI (bovine milk xanthine

532 dehydrogenase with NADH bound). PDB 1ZXI is colored in shades of green (according to chain
533 ID), PDB 3UNI is colored in gold (single chain A shown), and the *Eb* Ucd enzyme complex is
534 colored according to its per-residue confidence score as indicated in the legend. The various
535 ligands (cofactors, coenzymes, ions, small molecules) of PDB 1ZXI and PDB 3UNI are included
536 in the respective *Eb* Ucd enzyme models. **C,D)** Cofactors from X-ray crystal structures **C)** PDB
537 1ZXI and **D)** PDB 3UNI modeled into the AlphaFold2 *Eb* Ucd enzyme complex, showing a
538 complete electron transport chain from a bound FAD molecule to a molybdopterin cofactor *via*
539 two 2Fe-2S clusters. **E)** Xanthine dehydrogenase active site from PDB 3UNI. Side chains within
540 8 Å of the salicylic acid ligand are colored in gold. Residues important for substrate (purine)
541 binding and catalysis are labeled with their one letter amino acid code and sequence position.
542 Position E1261* is catalytically important (acts as a general base) and is conserved in
543 XDH/xanthine oxidase enzymes [56]. **F)** *Eb* Ucd enzyme complex active site modeling. Side
544 chains within 8 Å of the salicylic acid ligand from PDB 3UNI are colored on the superposed *Eb*
545 Ucd enzyme complex according to their per-residue confidence score as indicated in the
546 legend. Residues surrounding the predicted active site are labeled with their one letter amino
547 acid code and sequence position (in the *Eb* MoO protein). The predicted urolithin binding site
548 is depicted by the surface (mesh) created by the active site residues. The surface was rendered
549 using the cavities and pockets only (culled) setting with a cavity detection cutoff of 5 solvent
550 radii in PyMOL.

551

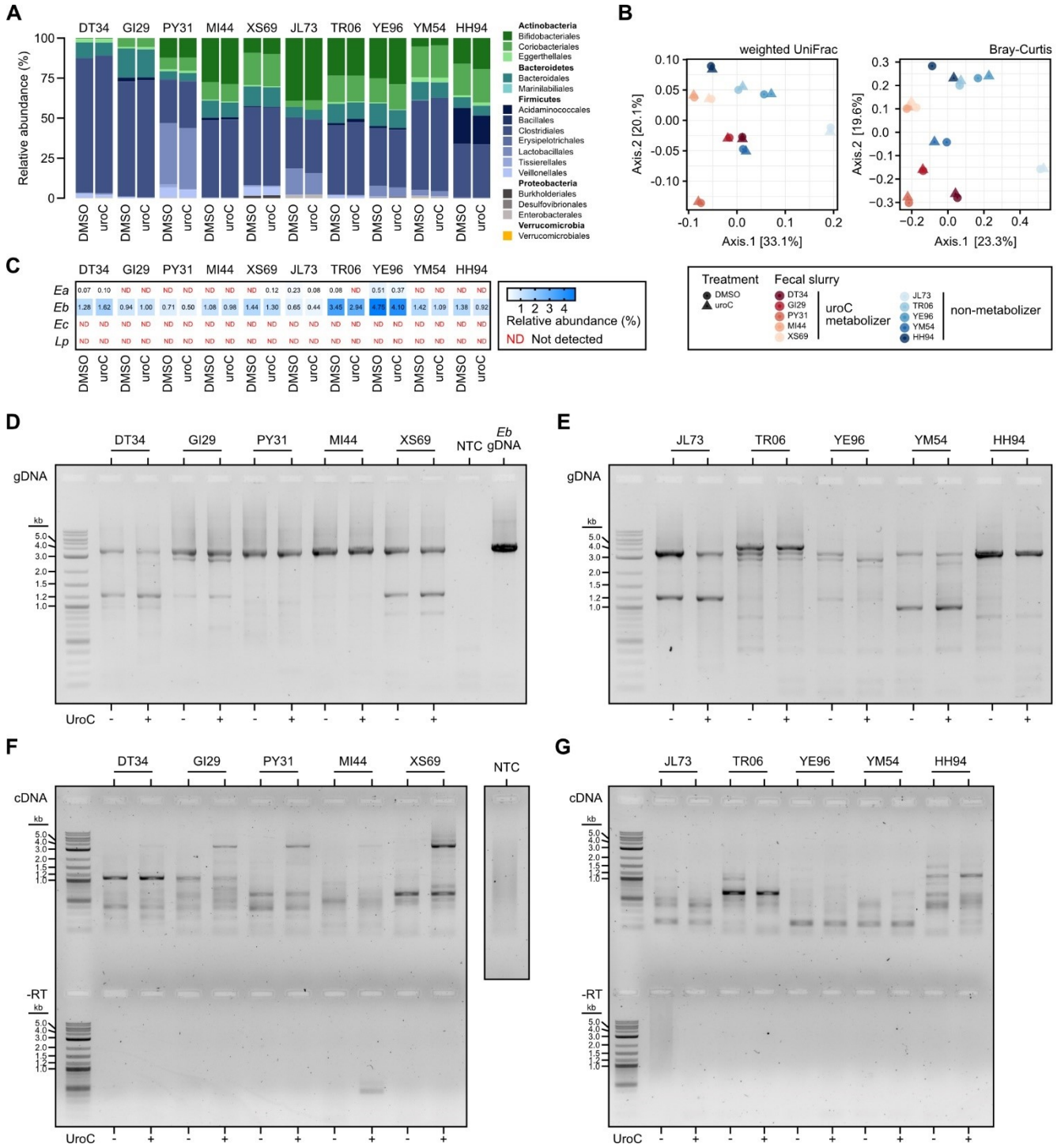


Extended Data Figure 7. Urolithin C-metabolizing species and genes are prevalent and correlate with *ucdO* gene abundance in human gut metagenomes.

A) Related to Fig. 6A. Prevalence of at least one *ucd* operon gene (blue bars) and uroC-metabolizing species (white bars) in fecal metagenomes across all 86 studies (in reverse alphabetical order). The number of participants in each study are represented to the right of the

558 prevalence plot as a heatmap. Details on the study populations (from the
559 curatedMetagenomicData R package) can be found in the Source Data file. **B)** Prevalence of
560 uroC-metabolizing species in human fecal metagenomes from the curatedMetagenomicData R
561 package. Data are reported for 86 studies (N=21,030 individuals) and are colored according to
562 the species. **C)** Violin plot of the $\log_{10}$(relative abundance) of uroC-metabolizing species in
563 human fecal metagenomes. The solid horizontal line corresponds to the median and the dashed
564 white lines correspond to the first and third quartiles. The percentage of zeroes are denoted
565 below the plotted distributions. Differences between groups were determined using the Kruskal-
566 Wallis test on untransformed relative abundance values. Significant differences between
567 groups are denoted by a different lowercase letter above each plot. **D)** Correlation between the
568 *ucdO* gene abundance in reads per kilobase per million mapped reads (RPKM) and the relative
569 abundance of each uroC-metabolizing species in fecal metagenomes. Both values are
570 illustrated on a $\log_{10}$ scale. Spearman rho (ρ) values are denoted above the scatter plots.

571


A
DT34 GI29 PY31 MI44 XS69 JL73 TR06 YE96 YM54 HH94
Relative abundance (%)
100
75
50
25
0
DMSO uroC
Actinobacteria
Bifidobacteriales
Coriobacteriales
Eggerthellales
Bacteroidetes
Bacteroidales
Marinilabiliales
Firmicutes
Acidaminococcales
Bacillales
Clostridiales
Erysipelotrichales
Lactobacillales
Tissierellales
Veillonellales
Proteobacteria
Burkholderiales
Desulfovibrionales
Enterobacterales
Verrucomicrobia
Verrucomicrobiales
B
weighted UniFrac
Bray-Curtis
Axis.2 [20.1%]
Axis.1 [33.1%]
Axis.2 [19.6%]
Axis.1 [23.3%]
C
Ea 0.07 0.10 ND ND ND ND ND ND ND 0.12 0.23 0.08 0.08 ND 0.51 0.37 ND ND ND ND
Eb 1.28 1.62 0.94 1.00 0.71 0.50 1.08 0.98 1.44 1.30 0.65 0.44 3.45 2.94 4.75 4.10 1.42 1.09 1.38 0.92
Ec ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND
Lp ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND ND
1 2 3 4
Relative abundance (%)
ND Not detected
Treatment
DMSO
uroC
Fecal slurry
DT34 GI29 PY31 MI44 XS69
uroC metabolizer
JL73 TR06 YE96 YM54 HH94
non-metabolizer
D
DT34 GI29 PY31 MI44 XS69 NTC Eb gDNA
gDNA
kb
5.0 4.0 3.0 2.0 1.5 1.2 1.0
UroC - + - + - + - + - +
E
JL73 TR06 YE96 YM54 HH94
gDNA
F
DT34 GI29 PY31 MI44 XS69 NTC
cDNA
-RT
G
JL73 TR06 YE96 YM54 HH94
cDNA
-RT

572

573 **Extended Data Figure 8. Urolithin C-metabolizing species and the *ucd* operon are**

574 **prevalent in fecal slurries, but only *ucd* transcription correlates with urolithin C**

575 **metabolism.**

576 **A)** Stacked bar plot of bacterial percent relative abundance (based on V1-V9 16S rRNA gene
577 sequencing) in DMSO- or uroC-treated fecal slurries from 10 healthy donors (from one
578 experimental replicate). Bars are colored according to the phylum (bold) and order. **B)**
579 Principal coordinate analyses of dissimilarities between 16S rRNA compositions based on the
580 (left) weighted UniFrac and (right) Bray-Curtis distance methods. **C)** Heatmap of the percent
581 relative abundance of uroC-metabolizing *Enterocloster* spp. in A). **D,E)** PCR with *ucd*-specific
582 primers on gDNA extracted from an *E. bolteae* isolate and fecal microbiota communities from
583 10 healthy donors (from one experimental replicate). 1% agarose gel of amplicons derived
584 from **D)** uroC-metabolizing and **E)** non-metabolizing fecal slurries. Samples are matched to
585 the urolithin metabolism data in Fig. 6C and 16S rRNA sequencing in A) and C). The no
586 template control (NTC) and *E. bolteae* positive control are the same for both gels. **F,G)** 1%
587 agarose gel of amplicons from Fig. 6D,E including the no reverse transcriptase (-RT) in **F)**
588 uroC-metabolizing and **G)** non-metabolizing fecal slurries. The NTC is derived from the final
589 lane of the gel in F). The uncropped and unannotated gel image is provided in the
590 Supplementary Information.

591

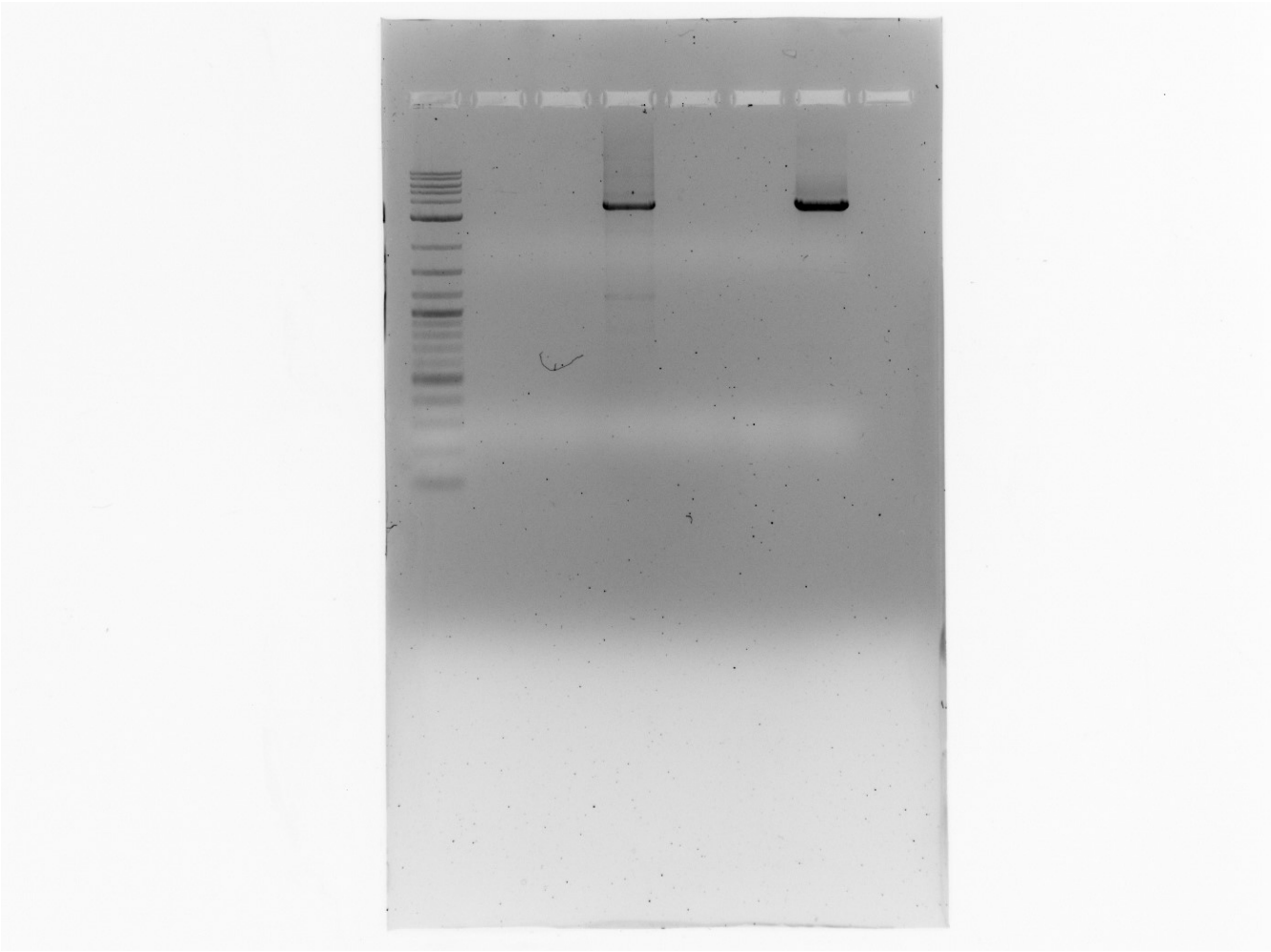
592 **Supplementary Information**

593 cDNA sequence reverse transcribed and amplified from uroC-induced *E. bolteae* RNA with *ucd*-
594 specific primers, 3605 bp (band can be seen on 1% agarose gel in Fig. 4B)

595 >Eb_ucd_RT-PCR_band coverage: 4.94e+03x

596 AAGCACAGCCAGGGTTATGGCAGGCGCAACGGATCTGATTCCGCCAATGAAAGACAAGGTTATATCACCAGAGTATATCATTG
597 ACCTTAAGAAAATTCCAGGTTTGGATTATCTGGAATACGATGACAGGGAAGGTTTAAAAATAGGGGCGCTGACAACACTGCGT
598 ACCATAGAGACATCTCCTCTGGTTAAAGAAAAGAATCCGGCAGTGGCTCATGCCGCCAAGGTAGTGGCATCCACACAAATCAG
599 GACAAAGGGCACCATGGCAGGCAATATCTGCAATGCGTCCCCATCCTGCGATACAGCCCCCAACCTACTGGCTCAGGGCGCTA
600 AGATATTGGTACAGGGTCCCAACAAGGACCGGGTCATTCAAATCGAAGACTTTTTCCTGGGTGTCAAAAAGACTTCCCTGGAG
601 CCGGGAGAAATCGTGACAGGCATTGTGATTCCGCCGCTGGCTGAGAATGAGAGGGCTGCATACATAAAACACGCGGTCAGAAA
602 GGCGATGGACCTGGCCATCATCGGCGTCGCAGTCAAGATTAAGGTGGAAGACGGCGTCTGCACAGACGCCAGGATTGCCCTGG
603 GCGCGGTTGCGGCCACGCCGGTACGGGCACCCGGGGCGGAGGAAGCCCTGATTGGCAAAGAACTTACCGATGAAGTGATTGTG
604 AAGGCATCGGAGGAAGCAATGAATTCCTGTCATCCAATCTCTGATATCCGCGCATCGGCCGAATACAGGAAGGATATGATCCG
605 CGTGTTTACAAAACGTGCCATTAAGCAGGCCATGGAATGCTATAACTAGGAGGGAGACAGGAATGAAACATAGTATAAGTTTA
606 ATCATCAACGGAGATCCGGTTGATGCTATTGTGAAGGATAACCTGACGCTTCTGGATTTTCTTCGTGACCAGTTGTTTCTGAC
607 AGGCACGAAAAAGGGGTGTGAGGAAGGGGAGTGTGGCGCCTGTACCGTGATGCTTGACGGTAAGCCGGTTAACTCCTGCTGTA
608 CCCTGGCCGTGGAGTGCGACGGCCATGAGATCATAACCGTGGAGGGAATTGCCAGGGAGGGGATGCTTCATCCCATTCAGAAG
609 CAGTTTATAGAGAAGTGGGCCATGCAGTGCGGCTACTGTACTCCGGGTATGATTATGTCCGCCAAAGCTCTTCTGGATGTGAA
610 TAAACATCCTACGGAACTGGAAATCCGGGAAGCCATTGAGGGCAATCTGTGCCGCTGCACCGGTTATGCAAAGATTGTGGAGG
611 CAATCCAGGCAGCCGCTGCGCAGATGAACTGGGAGGAGGAAGCAAAGAATGCATAAAGACTGTGACAAACATTATTTTAAAAA
612 ACCGGAATTTTACCGTCTGACCGGCGAGAACAATTATGTCAGGATTGACGCTGAAGACAAGGTAACAGGACACGGCCAGTATG
613 TGGGTGACATCATGTTTCCGGATATGCTTACCGGAAAAATGGTCAGAAGCCCTTATGCATCTGCAAAAATCCTGTCCATTGAC
614 ACCAGCGAGGCAGAGAAGCTTCCAGGTGTAAAGTGCATCCTCACTGCCAGGGACTTTGAGTGGAAGTCCCTGGTGGGAAATGG
615 AGAATTTGCAGCTGAGTTCGCGGACAAGGAAGTATTGTGTTCTGAGAAGGTGCGCCAGGTTGGTGACGATGTGGCAGCCGTGG
616 CAGCCGTGGATGAGGAAACCGCCCAGCGGGCCGCGGATCTGATTAAGGTAGAGTACCAGGTGCTTCCGGGGGTGTTTGACCCC
617 TTTGAGGCAATGGAGGAGAATGCCCCCGAGGTAAACTGGGAGGGCAAAGGCATACACAATATCGGCATGCAGTCCGTGATGAA
618 GGCGGGCACGGATATTGATGAGGAATTTGACCGCGCATCCTATGTGCAGCACAGGGATTACAAGACCCACCGTATGGTACATG
619 CGGCCATGGAGCCCCACGGTGCGGTGGCTACCTACAGGAATGGGACCTACACCATCTGGATGTCCACCCAGATGTCCTTTGTG
620 GACCAGTTCTGGTATGCCCGCTGCCTTGGCGTTGGAGAGAACCAGGTGCGGGTAATCAAGCCTCTTGTGGGCGGCGGTTTCGG
621 CGGCAAGCTGGATTCCTATTCCTTTGGCCTTTGCGCTGCCAAGATGGCGGAGATGACCGGACGCCCGGTACGCATGATTCTGT
622 CCCGCGAAGAAGTATTCCAGACCACGCGCAACCGCCACCCCATTTACATGCATATTGACACTGCCTTTGGGACAGACGGTAAG
623 CTGCTGGCAAAGAAATGCTACCATGTGCTGGACGGCGGGCCTTACGGCGGGTCCGGTGTTGCGGCCTGCGCCCAGTCCACATT
624 GTGGGCCAACTTCCCTTACAAGATGAATTCCGTGGATTTCCTGGCAAGGCGTGTGTATACCAATAACCCTTCTGCAGGAGCCA
625 TGAGAGGATATACGGCATGCCAGGTGCATTTTGCCCATGACCTGAACATGCAGTTCGCGGCTGACCAGATGGGCATTGACCCG
626 GTTGAGTTCCGCAAGATAAGCGCGGCAGACCCCGGATACGTGGCTCCGGCCGGACTGGCAATCACCAGCTGCGCATACAAGGA
627 GACACTGGATACGGCTGCAAAGGAAATTGGCTGGTATGAGAAGAAGGATAAGCTGAAAAAGGGAGAGGGCATTGGTTTTGCGG
628 GCACCGGCTTTGTATCGGGCACCGGATTCGCGGTGCTGGAAGCGCCCAACCAGAGCTCTGCGTGCGTGACCCTGCGCATGAAC
629 AAGCGGGGCATGGCCACCCTGTACATCGGCTCCCACGACATAGGACAGGGTTCCGATACGGTTATGACGGCCATTGTGGCCGA
630 GGAACTGGGGCTTCCCATGGATATGGTCAAGACCTTTATGTCCGACACCTTCCTGACGCCCTGGGACTCCGGCTCTTACGGCA
631 GCCGTGTTACCTTCCTGGCCGGCAATGCGGCCCGCCGCGCTGCCGTGGATGCAAAACGCCAGCTGTTTGAGGTCATTGCGCCT
632 ATGTGGGGGGTGATGCCGGAGACATTAGAGTGTCTGGACGGAAAGGTAATCAGCAAGGAGAAGGCAGAGTATCAGATGTCCAT
633 TGGAGACGCCATGTTCAAGTATATGACGGTCAAGGGCGGCGATGAACTGATTGGCGTGGGATCCTATTACCACCGTACCGACA
634 ATTCACAGTATAACGGCAACAATACCACCAACTACGCGCCTGCATACAGCTTCTCCACCGGAGCAGCCCATCTGACTGTGGAC
635 GAGGAGACCGGCGTCCTGGATATTGATGAATTTGTATTTGCCCATGACTGCGGCCGCGCACTGAATAAGAGGGCCGTGGAAGG
636 CCAGCTGGAGGGATCCATCGGCATGGGCCTGGGCTATGCTGTCTATGAGCACAATGTAACCAGGGAAGGAAAGATTCTCAATC
637 CCAACTTCCGCGATTACCGGCTGCCCACTGCCCTTGATATGCCGAAGATGCGCACCTTTTACGACTTTACACCGGACGAGGAA
638 GGACCTCTGGGCGCAAAGGAAGCTGGCGAGGGCTCCGCAGCACCTGTGGCGCCTGCCATTGCCAATGCGGTCAACATGGCAAC
639 CGGCGTGTACTTCACAGAGCTGCCTCTGGACCCGGA

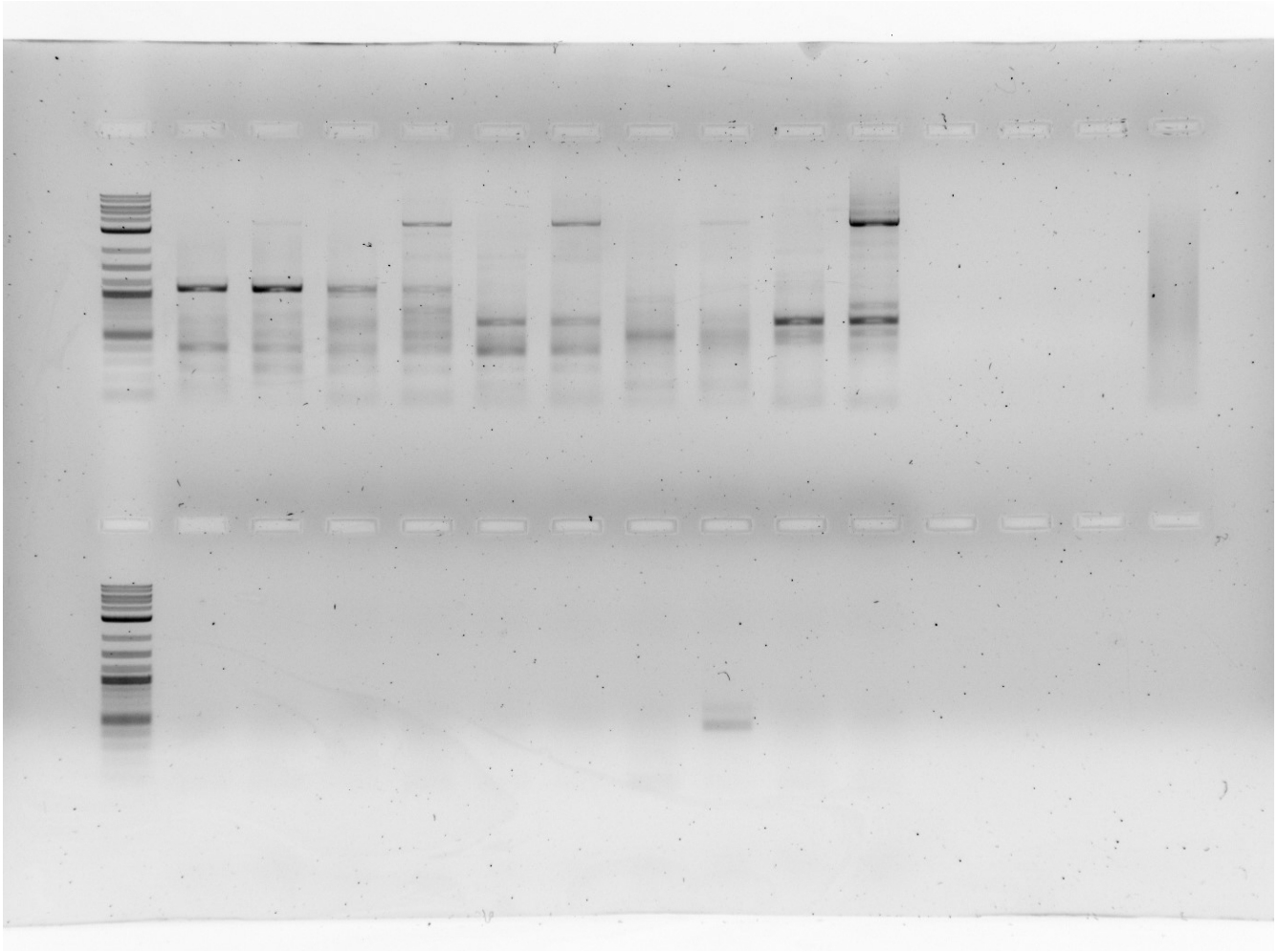
640 Full uncropped gel image from Fig. 4B:

641 

642 The ladder used here is the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550).

643

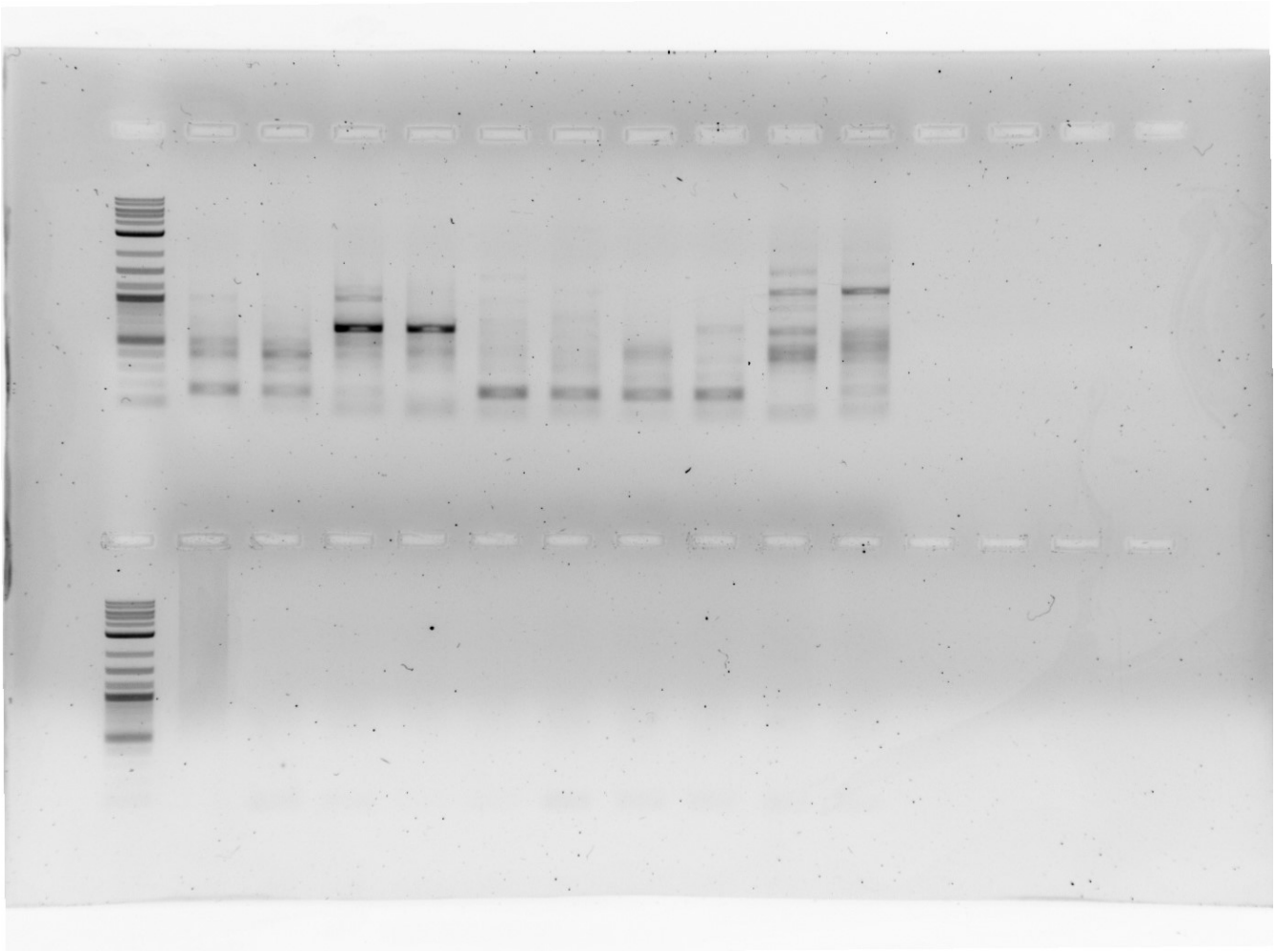
644 Full uncropped gel image from Fig. 6D and Extended Data Fig. 8F:



645

646 The ladder used here is the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550).

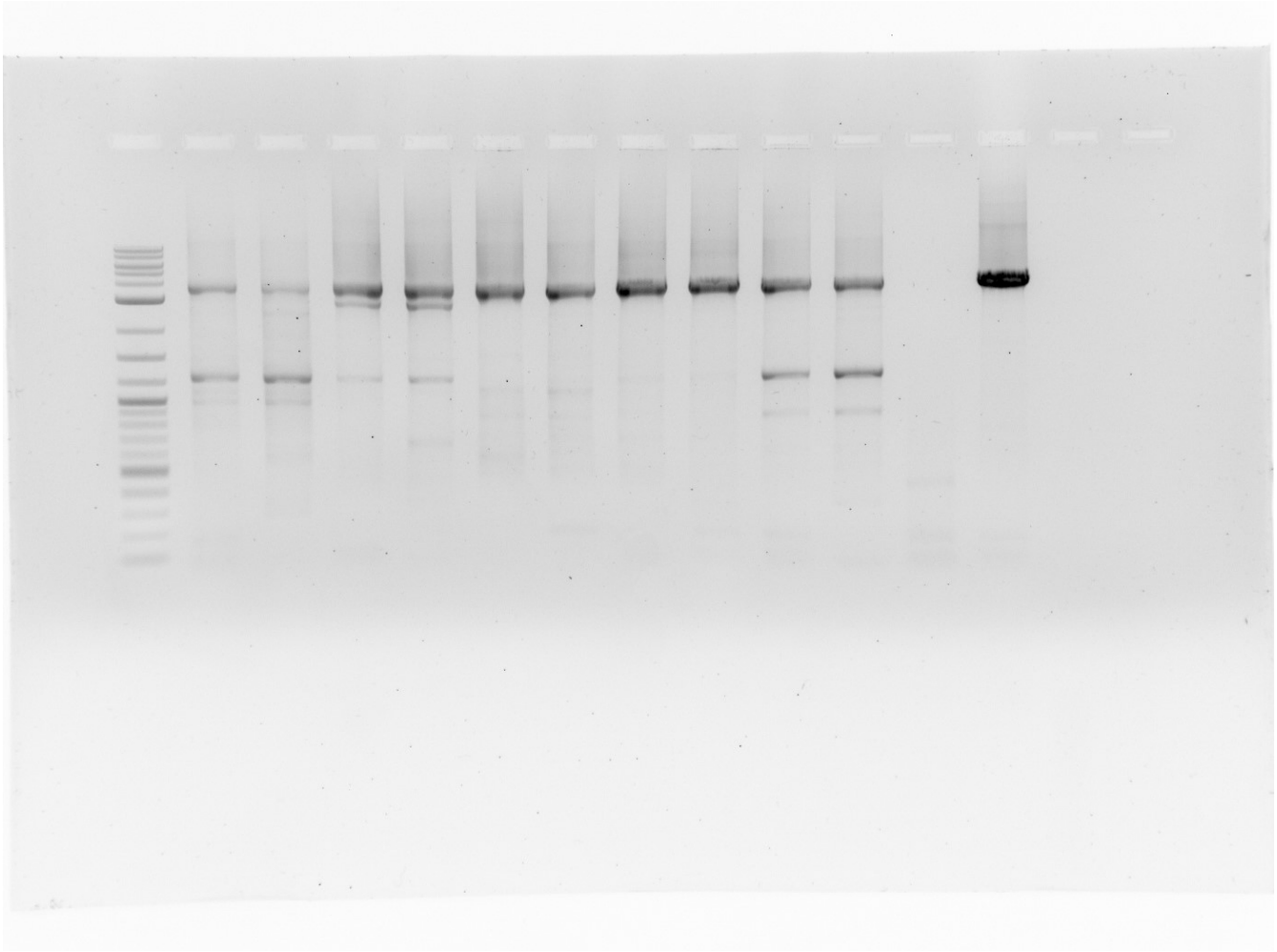
647 Full uncropped gel image from Fig. 6E and Extended Data Fig. 8G:



648

649 The ladder used here is the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550).

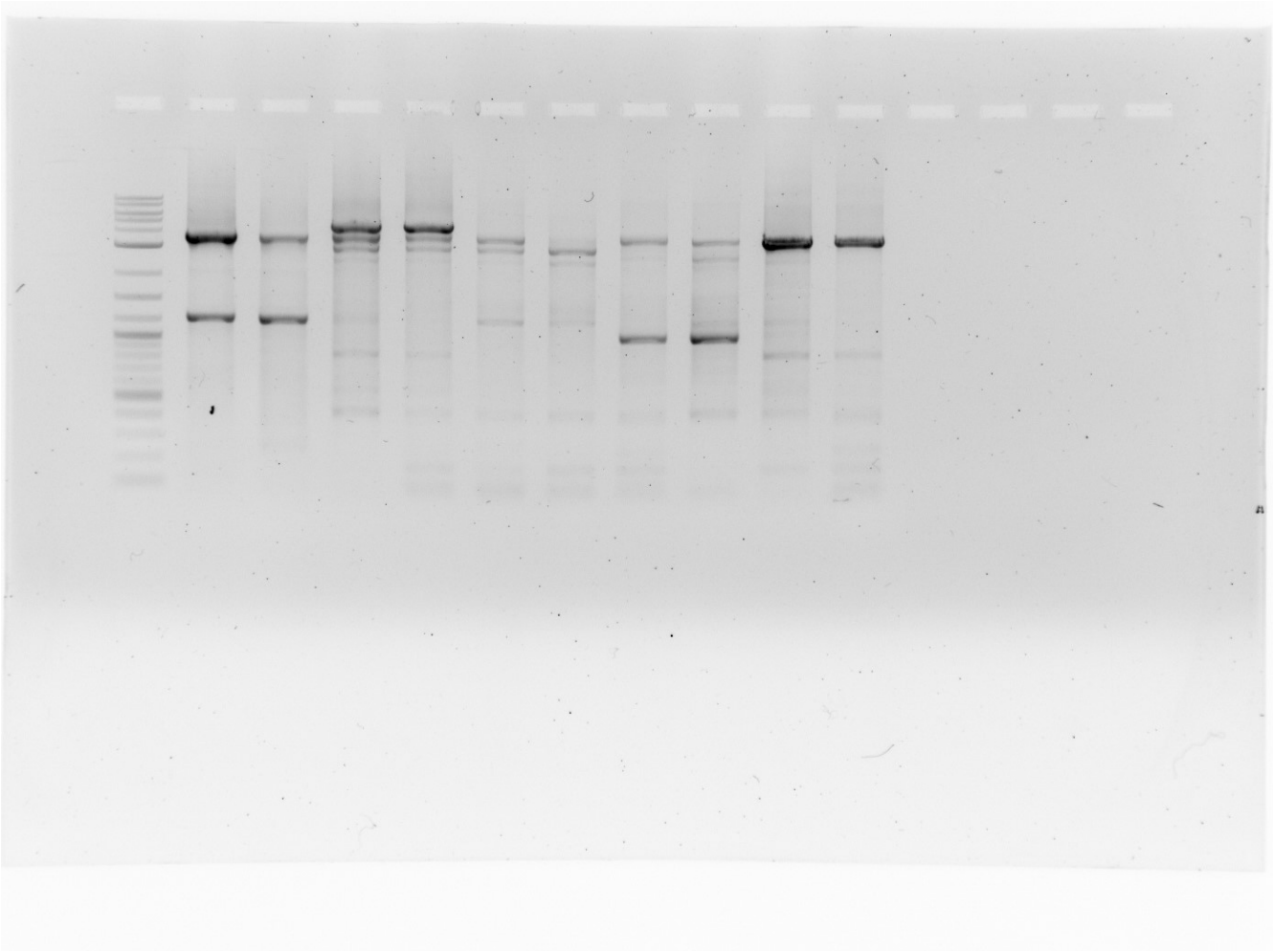
650 Full uncropped gel image from Extended Data Fig. 8D:



651

652 The ladder used here is the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550).

653 Full uncropped gel image from Extended Data Fig. 8E:



654

655 The ladder used here is the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB, N0550).

656 **Online Methods**

657 **Resources table**

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background	*Enterocloster aldenensis* DSM 19262	DSMZ	DSM 19262	Type strain
Strain, strain background	*Enterocloster asparagiformis* DSM 15981	DSMZ	DSM 15981	Type strain
Strain, strain background	*Enterocloster bolteae* DSM 15670	DSMZ	DSM 15670	Type strain
Strain, strain background	*Enterocloster citroniae* DSM 19261	DSMZ	DSM 19261	Type strain
Strain, strain background	*Enterocloster clostridioformis* DSM 933	DSMZ	DSM 933	Type strain
Strain, strain background	*Enterocloster lavalensis* DSM 19851	DSMZ	DSM 19851	Type strain
Strain, strain background	*Lachnoclostridium pacaense* CCUG 71489T	CCUG	CCUG 71489T	Type strain
Commercial assay or kit	One-4-All Genomic DNA Minipreps Kit	BioBasic	BS88503	
Commercial assay or kit	Direct-zol RNA Miniprep Kit	Zymo Research	R2051	
Commercial assay or kit	ZR BashingBead Lysis Tubes (0.1 & 0.5 mm)	Zymo Research	S6012-50	
Commercial assay or kit	Ambion *DNA-free* DNA Removal Kit	Invitrogen	AM1906	
Commercial assay or kit	*OneStep* PCR Inhibitor Removal Kit	Zymo Research	D6030	
Commercial assay or kit	iScript Reverse Transcription Supermix	Bio-Rad	1708840	RT-PCR with random hexamers
Commercial assay or kit	LunaScript RT Master Mix Kit (Primer-free)	NEB	E3025	For gene-specific RT-PCR
Commercial assay or kit	Luna Universal qPCR Master Mix	NEB	M3003	
Commercial assay or kit	Monarch PCR & DNA Cleanup Kit	NEB	T1030	
Commercial assay or kit	Q5 High-Fidelity Polymerase	NEB	M0491	For 16S PCR and cloning
Commercial assay or kit	One*Taq* Quick-Load 2X Master Mix with Standard Buffer	NEB	M0486	For colony PCR
Commercial assay or kit	One*Taq* 2X Master Mix with Standard Buffer	NEB	M0482	For *ucd*-specific PCR
Commercial assay or kit	Platinum SuperFi II Green PCR Master Mix	Invitrogen	12369010	For ONT 16S PCR on fecal slurries
Commercial assay or kit	dNTP Mixture (10 mM)	BioBasic	DD0056	
Commercial assay or kit	Qubit dsDNA High Sensitivity (HS) Assay Kit	Invitrogen	Q32851	

Commercial assay or kit	Qubit RNA Broad Range (BR) Assay Kit	Invitrogen	Q10210	
Commercial assay or kit	Amplicon barcoding kit AB01-24	ONT	SQK-AMB111.24	
Commercial assay or kit	Thermolabile proteinase K	NEB	P8111	
Commercial assay or kit	Agencourt AMPure XP	Beckman Coulter	A63880	
Commercial assay or kit	Flongle Flow Cell (R9.4.1)	ONT	FLO-FLG001	
Plasmids and primers	Custom oligonucleotides (Desalted, Dried Form)	Invitrogen	10336022	See Primer sequences table
Commercial assay or kit	ExcelBand 100 bp+3K DNA Ladder	SMOBIO	DM2300	0.1-3 kb
Commercial assay or kit	Quick-Load Purple 1 kb Plus DNA Ladder	NEB	N0550	0.1-10 kb
Chemical compound, drug	Acetonitrile	Sigma	34998	CAS 75-05-8
Chemical compound, drug	Ethyl acetate	Sigma	34858	CAS 141-78-6
Chemical compound, drug	Methanol	Sigma	34860	CAS 67-56-1
Chemical compound, drug	Ethanol	Commercial Alcohols	P016EAAN	CAS 64-17-5
Chemical compound, drug	Formic acid	Sigma	F0507	CAS 64-18-6
Chemical compound, drug	Peptone	BioShop	PEP403.1	
Chemical compound, drug	Yeast extract	BioShop	YEX401.500	CAS 8013-01-2
Chemical compound, drug	Sodium chloride	BioShop	SOD004.5	CAS 7647-14-5
Chemical compound, drug	Starch (from potato)	Sigma	S2004-500G	CAS 9005-25-8
Chemical compound, drug	D-glucose monohydrate	BioShop	GLU601.1	CAS 14431-43-7
Chemical compound, drug	Sodium pyruvate	BioShop	PYR302.100	CAS 113-24-6
Chemical compound, drug	Sodium succinate dibasic	Sigma	14160-100G	CAS 150-90-3
Chemical compound, drug	Sodium thioglycolate	TCI	M0053	CAS 367-51-1
Chemical compound, drug	Agar	Fisher	BP1423-500	CAS 9002-18-0
Chemical compound, drug	L-arginine HCl	BioShop	ARG006.100	CAS 1119-34-2
Chemical compound, drug	L-cysteine	BioShop	CYS555.100	CAS 52-90-4
Chemical compound, drug	Sodium bicarbonate	BioShop	SOB308.500	CAS 144-55-8
Chemical compound, drug	Haemin	Sigma	H9039-1G	CAS 16009-13-5
Chemical compound, drug	Vitamin K1	Sigma	95271-250MG	CAS 84-80-0
Chemical compound, drug	Dithiothreitol (DTT)	Fisher	BP172-25	CAS 3483-12-3

Chemical compound, drug	Tris base	Sigma	10708976001	CAS 77-86-1
Chemical compound, drug	Magnesium sulfate heptahydrate	BioShop	MAG511.1	CAS 10034-99-8
Chemical compound, drug	Calcium chloride dihydrate	Fisher	BP510-500	CAS 10035-04-8
Chemical compound, drug	SIGMAFAST Protease Inhibitor Cocktail, EDTA-free	Sigma	S8830-2TAB	
Chemical compound, drug	Urolithin C	TRC	U847015	CAS 165393-06-6
Chemical compound, drug	Urolithin A	TRC	U847000	CAS 1143-70-0
Chemical compound, drug	Urolithin A	Sigma	SML1791-5MG	CAS 1143-70-0
Chemical compound, drug	Urolithin M-6	TRC	U847040	CAS 1006683-97-1
Chemical compound, drug	Isourolithin A	TRC	I917520	CAS 174023-48-4
Chemical compound, drug	Urolithin B	TRC	U847005	CAS 1139-83-9
Chemical compound, drug	Salicylic acid	Sigma	84210-100G	CAS 69-72-7
Chemical compound, drug	TRI reagent	Zymo Research	R2050-1-50	
Chemical compound, drug	Agarose	Fisher	BP160-500	CAS 9012-36-6
Chemical compound, drug	SafeView Classic	Abm	G108	
Chemical compound, drug	PBS pH 7.4 (10X)	Gibco	70011-044	
Chemical compound, drug	Nuclease-free water	Invitrogen	10977-015	
Lab equipment, instruments	AnaeroPack System	Mitsubishi Gas Chemical Company	10-01	
Lab equipment, instruments	Vinyl anaerobic chambers	Coy Laboratory Products	Type A	
Lab equipment, instruments	Multiscan FC Microplate Photometer	Thermo Fisher	51119100	
Lab equipment, instruments	Costar 96-Well Clear Round Bottom	Corning	3788	
Lab equipment, instruments	LifeECO Thermal Cycler	BIOER	TC-96	
Lab equipment, instruments	Nanodrop 2000c	Thermo Fisher	ND-2000C	
Lab equipment, instruments	Mini Beadbeater 16	Biospec	607	
Lab equipment, instruments	Viia7 qPCR machine	Applied Biosystems	4453535	
Lab equipment, instruments	MicroAmp Fast 96-Well Reaction Plates (0.1 mL)	Applied Biosystems	4346907	
Lab equipment, instruments	PlateSeal Film, Clear Polyester, X-Large, Non-Sterile	PlateSeal	PS-PETSTXL-100	
Lab equipment, instruments	LC-MS (Agilent 1260 Infinity II LC connected to a 6120 Quadrupole MS)	Agilent		

Lab equipment, instruments	Poroshell 120 EC-C18 column (4.6x50 mm, 2.7 µm)	Agilent	699975-902	
Lab equipment, instruments	Poroshell 120 EC-C18 column (4.6x5 mm, 2.7 µm) guard column	Agilent	820750-911	
Lab equipment, instruments	Vacuum concentrator	Heto Lab	VR-1	
Lab equipment, instruments	Sonicator	Misonix	S3000	
Lab equipment, instruments	Rotary evaporator	Buchi	R-300	
Software	Agilent OpenLab CDS	Agilent	N/A	
Software	GraphPad Prism 10	GraphPad by Dotmatics	N/A	
Software	Mascot 2.6.2	Matrix Science	N/A	
Software	Scaffold 5	Proteome Software Inc	N/A	
Software	R Studio 2023.06.0+421	Posit Software	N/A	
Software	Affinity Designer	Serif Ltd	N/A	

NEB = New England Biolabs, ONT = Oxford Nanopore Technologies, TRC = Toronto Research Chemicals

Primer sequences table

Designation	Purpose	Source	5'-Sequence-3'
16S_V1_27_f	16S PCR		AGAGTTTGATCMTGGCTAG
16S_V9_1492_r	16S PCR		TACGGYTACCTTGTTAYGACTT
ONT_16S_27_f_GGK	16S PCR		**GGK**AGRGTTYGATYMTGGCTCAG
ONT_16S_1492_r_GGK	16S PCR		**GGK**CCGGYTACCTTGTTACGACTT
Eb_ucdO_qPCR_f	qPCR	This work	GTGGATGAGGAAACCGCCCAGC
Eb_ucdO_qPCR_r	qPCR	This work	TGCCCTCCCAGTTTACCTCGGG
Eb_ucdF_qPCR_f	qPCR	This work	GCGCCTGTACCGTGATGCTTGA
Eb_ucdF_qPCR_r	qPCR	This work	CCCTCCCTGGCAATTCCCTCCA
Eb_ucdC_qPCR_f	qPCR	This work	AGCCCCCAACCTACTGGCTCAG
Eb_ucdC_qPCR_r	qPCR	This work	GCCTGTCACGATTTCTCCCGGC
Eb_dnaK_Ref_qPCR_f	qPCR	This work	GGTGCTGTTGGTAGGCGGTTCC
Eb_dnaK_Ref_qPCR_r	qPCR	This work	GGGTCTTGCTGGGTTCCTTGCC
ucdCFO_RT-PCR_f	RT-PCR	This work	GCAATCTGTTCCTYGAGCTTGG
ucdCFO_RT-PCR_r	RT-PCR	This work	CGTGSAGCGCTCTCCAGATATG

Adapter sites are bolded.

662 **Bacterial strains and culturing conditions**

663 Bacterial strains used in this study are listed in the *Resources table*. All bacterial stains
664 were validated by sequencing the 16S rRNA gene (see Genomic DNA extraction and 16S rRNA
665 sequencing of bacterial isolates). The same culture used for validation was used to make 25%
666 glycerol stocks. Strains were grown from glycerol stocks on mABB+H agar plates (recipe below)
667 for 48-72 h at 37 °C in a vinyl anaerobic chamber, which was maintained with a gas mixture of
668 3% H_2, 10% CO_2, 87% N_2. To make overnight cultures, a single colony was inoculated into 5
669 mL of liquid mABB+H medium and incubated at 37 °C between 16-24 h.

670 **Human fecal sample collection**

671 Human fecal samples were collected under the approval of protocol A04-M27-15B by
672 the McGill Faculty of Medicine Institutional Review Board. Informed written consent was
673 received from the participants for the use of human samples. Eligibility criteria for the healthy
674 participants were as follows: body mass index between 18.5–30, no diagnosed gastrointestinal
675 disease, no ongoing therapeutic treatment, and no antibiotic usage 3 months prior to the start
676 of the study. Subject information was recorded at the time of sampling. The age of donors
677 ranged from 21–40 years. Fresh fecal samples were collected and placed immediately in an
678 anaerobic chamber, aliquoted, and stored at −70 °C until use.

679 **Modified anaerobe basal broth (mABB and mABB+H)**

680 For 1 L of modified anaerobe basal broth (mABB), the following components were
681 dissolved in MilliQ water, then autoclaved: 16 g peptone, 7 g yeast extract, 5 g sodium chloride,
682 1 g starch, 1 g D-glucose monohydrate, 1 g sodium pyruvate, 0.5 g sodium succinate, 1 g
683 sodium thioglycolate, 15 g agar (for plates). The autoclaved solution was allowed to cool, then
684 the following filter-sterilized solutions were added aseptically: 10 mL of 100 mg/mL L-arginine-
685 HCl, 10 mL of 50 mg/mL L-cysteine, 8 mL of 50 mg/mL sodium bicarbonate, 50 µL of 10 mg/mL

686 vitamin K1, 20 mL of 50 mg/mL dithiothreitol, and, for mABB+H, 10 mL of 0.5 mg/mL haemin.
687 The media was then placed in the anaerobic chamber and allowed to reduce for at least 24 h
688 prior to its use in experiments.

689 **Genomic DNA extraction of isolates and identity validation**

690 The identities of all bacteria in this study were validated by full-length (V1-V9) 16S rRNA
691 sequencing. DNA was first extracted from 0.5-1 mL of overnight culture using the One-4-All
692 Genomic DNA Miniprep Kit (BioBasic) according to the manufacturer's instructions. The purified
693 genomic DNA (2 µL, ~ 20 ng) was used as a template for PCR reactions (25 µL reaction volume)
694 using the Q5 High-Fidelity polymerase (NEB). PCR tubes were placed in a thermal cycler and
695 targets (~1.5 kb) were amplified according to the following cycling conditions: 30 s at 98 °C, 30
696 cycles (10 s at 98 °C, 20 s at 60 °C, 45 s at 72 °C), 2 min at 72 °C, and hold at 10 °C. 5 µL of
697 the reaction was mixed with 6X loading buffer and loaded onto a 1% agarose gel (made with
698 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were compared to the
699 ExcelBand 100 bp+3K DNA Ladder (SMOBIO).

700 PCR products (~1.5 kb) were purified using the Monarch PCR & DNA Cleanup Kit (NEB)
701 according to the manufacturer's instructions for products < 2 kb. Purified 16S PCR products
702 were eluted in nuclease-free water, quantified using the Qubit dsDNA HS assay kit (Invitrogen),
703 and adjusted to 30 ng/µL. Samples were submitted to Plasmidsaurus for long-read sequencing
704 using Oxford Nanopore Technologies (v14 library preparation chemistry, R10.4.1 flow cells).

705 **Urolithin extraction from fecal slurries or bacterial cultures**

706 Frozen (-70 °C) fecal slurries or bacterial cultures were thawed at room temperature. For
707 quantification of urolithin concentrations, urolithin standards (stock 10 mM in DMSO) were
708 spiked into media immediately before extraction.

709 *Extraction Method A*: Salicylic acid (3 mg/mL in DMSO) was spiked-in as an internal standard
710 at a final concentration of 50 µg/mL. The cultures and standards were then extracted with 3
711 volumes of ethyl acetate + 1% formic acid (e.g., 600 µL solvent to 200 µL thawed culture). The
712 organic phase (top) was transferred to a new tube and dried in a vacuum concentrator (Heto
713 Lab) connected to a rotary evaporator (Buchi). After solvent removal, samples were redissolved
714 in 0.5 volumes (relative to the starting culture) of 50% MeOH:H_2O. Samples were centrifuged
715 at 20,000 x g for 5 min to pellet insoluble material, then transferred to LC-MS vials. Urolithins
716 were then analyzed by LC-MS.

717 *Extraction Method B*: The cultures, suspensions, or supernatants were diluted with an equal
718 volume of MeOH, vortexed briefly, and incubated at room temperature for 10 min. Samples
719 were centrifuged at 20,000 x g for 5 min to pellet insoluble material, then transferred to LC-MS
720 vials. Urolithins were then analyzed by LC-MS.

721 **LC-MS method to quantify urolithins**

722 Samples (10 µL) were injected into a 1260 Infinity II Single Quadrupole LC/MS system
723 (Agilent) fitted with a Poroshell 120 EC-C18 4.6x50 mm, 2.7 µm column (Agilent). The mobile
724 phase was composed of MilliQ water + 0.1% formic acid (solvent A) and acetonitrile + 0.1%
725 formic acid (solvent B). The flow rate was set to 0.7 mL/min. The gradient was as follows: 0-8
726 min: 10-30 %B, 8-10 min: 30-100 %B, 10-13.5 min: 100 %B isocratic, 13.5-13.6 min: 100-10
727 %B, then 13.6-15.5 min: 10 %B. The multiple wavelength detector was set to monitor
728 absorbance at 305 nm. The mass spectrometer was run in negative mode in both selected ion
729 monitoring (SIM) and scan (100-1000 m/z) modes to validate peak identities. Peaks were
730 validated based on retention times compared to spike-in standards and mass-to-charge ratios.
731 To quantify urolithins, peak areas for the compounds of interest were compared to spike-in

732 standards of known concentration(s). When standards were not available (urolithin M7), the
733 extracted ion chromatogram was used ($[M-H]^-$ = 243).

734 **Isolate treatments with urolithins**

735 *Treatment during growth (spike-in)*: Overnight cultures of *Enterocloster* spp. were diluted
736 1/50 into fresh mABB+H and incubated at 37 °C. After 5 hours of incubation (~OD_{620} from a
737 200 µL sample ~ 0.4), 10 mM urolithins (or an equivalent volume of DMSO) were added to the
738 growing cultures at a final concentration of 50 or 100 µM for protein expression or RNA
739 expression, respectively. For protein expression analyses and inducibility tests, the cultures
740 were incubated for an additional 4 h. For RNA expression analyses, the cultures were incubated
741 for an additional 2 h.

742 *Treatment prior to growth*: Overnight cultures of *Enterocloster* spp. were diluted 1/50 into fresh
743 mABB+H. Separately, treatments (10 mM stocks solutions, dissolved in DMSO) were prepared
744 in fresh mABB+H medium to a final concentration of 111 µM. In a 96-well plate, 180 µL of
745 treated media were combined with 20 µL of 1/5 culture. These were plated in technical
746 triplicates. The final concentration of treatment was 100 µM (unless otherwise specified in
747 concentration-response experiments) and the final dilution of bacteria was 1/100.

748 In both cases, urolithins were extracted from cultures using *Extraction Method A*.

749 **Phylogenetic tree construction**

750 Phylogenetic trees based on the 16S rRNA gene were constructed using the DSMZ
751 single-gene phylogeny server (https://ggdc.dsmz.de/phylogeny-service.php#)[57] with GenBank
752 16S rRNA sequence accessions: *E. aldenensis* (DQ279736), *E. asparagiformis* (AJ582080), *E.*
753 *bolteae* (AJ508452), *E. citroniae* (HM245936), *E. clostridioformis* (M59089), *E. lavalensis*

754 (EF564277), *L. pacaense* (LT799004), *G. pamelaeae* (AM886059), *G. urolithinfaciens*
755 (HG000667), *E. isourolithinifaciens* (MF322780).

756 Phylogenetic trees based on whole genomes and proteomes were constructed using the
757 Type (Strain) Genome Server (TYGS, https://tygs.dsmz.de)[55,58] with the following GenBank
758 genome accessions: *E. aldenensis* (GCA_003467385.1), *E. asparagiformis*
759 (GCA_025149125), *E. bolteae* (GCA_000154365), *E. clostridioformis* (GCA_900113155), *E.*
760 *lavalensis* (GCA_003024655), *L. pacaense* (GCA_900566185). For *E. citroniae*, the Integrated
761 Microbial Genomes ObjectID was used: *E. citroniae* (2928404274). Further information on
762 nomenclature and taxonomy was obtained from the List of Prokaryotic names with Standing in
763 Nomenclature (LPSN, available at https://lpsn.dsmz.de).

764 **Growth curves**

765 Bacterial cultures were prepared as described above (see *Treatment during growth*
766 *(spike-in)* in the *Isolate treatments with urolithins section*). At each timepoint, 200 µL of culture
767 were transferred to a Costar 96-Well Clear Round Bottom (Corning) and the OD_{620} was
768 measured using a Multiscan FC Microplate Photometer (Thermo Fisher).

769 **Resting cell suspension assay to test inducibility**

770 Bacteria (10 mL growing cultures in mABB+H media) were grown with 50 µM uroC (or
771 an equivalent volume of DMSO) as detailed in the *Treatment during growth (spike-in)* section
772 above and incubated for 4 h at 37 °C. Cultures were then pelleted at 6,500 x g for 3 min and
773 the supernatants were discarded. The cells were washed with 10 mL of pre-reduced PBS
774 (placed in the anaerobic chamber 24 h before), re-pelleted, and resuspended in 2 mL of pre-
775 reduced PBS, yielding resting cell suspensions. For each condition tested, a 200 µL aliquot of
776 the resting cell suspension was transferred into a sterile 1.5 mL tube, and 10 mM urolithins

777 (uroM6, uroC, isouroA, or DMSO) were added at a final concentration of 100 μM. Resting cell
778 suspensions were briefly vortexed and incubated at room temperature in the anaerobic
779 chamber for 16h prior to freezing and urolithin extraction using *Extraction Method B*.

780 **RNA extraction from isolates**

781 A volume of 1.5 mL of treated (100 μM urolithin C for 2 h) *Enterocloster* spp. culture (see
782 *Enterocloster* spp. urolithin C treatments) was pelleted (6,500 g for 3 min) and the supernatant
783 was removed for later LC-MS analysis. The pellet (suspended in 200 μL of media) was then
784 mixed with 800 μL TRI reagent (Zymo Research) and transferred to a ZR BashingBead lysis
785 tube (Zymo Research). Samples were lysed in a Mini Beadbeater 16 (Biospec) according to
786 the following sequence: 1 min ON, 5 min OFF. For RNA-sequencing, bead beating was done
787 for a total of 5 min ON. For RT-(q)PCR, bead beating was done for a total of 2 min ON to
788 preserve longer transcripts. RNA isolation was then performed using the Direct-zol RNA
789 Miniprep Kit (Zymo Research) according to the manufacturer's instructions (including an on-
790 column DNase digestion). To ensure complete DNA removal, an additional DNA digestion step
791 was performed on the isolated RNA using the Ambion *DNA-free* DNA Removal Kit (Invitrogen)
792 according to the manufacturer's instructions. The DNA-free RNA was then cleaned up using
793 the *OneStep* PCR Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were
794 initially verified by NanoDrop and 1 % agarose gel electrophoresis. For RNA-sequencing, RNA
795 integrity was assessed by Génome Québec using a Bioanalyzer 2100 (Agilent). RNA integrity
796 (RIN) values ranged between 7.5-7.8 for *E. asparagiformis* DSM 15981 and 7.0-7.3 for *E.*
797 *bolteae* DSM 15670.

798 **RNA-sequencing of urolithin C-treated *E. bolteae* and *E. asparagiformis* isolates**

799 Total RNA was sent to Génome Québec for library preparation and RNA-sequencing.
800 Briefly, total RNA was prepared for Illumina sequencing using the NEBNext rRNA Depletion Kit

801 (Bacteria) (NEB) kit to remove rRNA and using the NEBNext Multiplex Oligos for Illumina (NEB)
802 kit (stranded/directional). Prepared libraries were quality checked with a Bioanalyzer 2100
803 (Agilent) prior to sequencing. Sequencing was performed on a NovaSeq 6000 (Illumina) with
804 the following flow cell/settings: S4 flow cell, 100 bp, 25 M reads, paired end.

805 Analysis of RNA-seq reads was done using Galaxy bioinformatics cloud computing
806 (https://usegalaxy.org/) hosted by Compute Canada Genetics and Genomics Analysis Platform
807 (GenAP) (https://www.genap.ca/). Genomes and annotations were fetched from the NCBI
808 genome browser: *E. bolteae* (ASM223457v2) (accessed 2022/05/11) and *E. asparagiformis*
809 (ASM2514912v1) (accessed 2023/09/19). Raw reads were first verified for quality using
810 FastQC (v0.73, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) with default
811 parameters. FastQC reports were aggregated into MultiQC (v1.11) [59]. The mean sequence
812 quality scored were above 35 for all samples. Raw reads were then trimmed using Cutadapt
813 (v3.7) [60] to trim adapter sequences (R1 sequence:
814 AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC, and R2 sequence:
815 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT) that were not removed after sequencing
816 using default parameters for paired end reads. Trimmed reads were then aligned to reference
817 genomes for each bacterium using HISAT2 (v2.2.1) [61] with paired-end parameters and reverse
818 strandedness (RF). Aligned read counts were assigned to features in annotation files (.gtf)
819 using featureCounts (v2.0.1) [62] with the following parameters: reverse strandedness, count
820 fragments instead of reads, GFF feature type filter = “gene”, multi-mapping and multi-
821 overlapping features included (-M -O), minimum mapping quality per read of 0, and the rest of
822 the parameters were kept as default. Differential gene expression analysis was then performed
823 using DESeq2 (v2.11.40.7) [63] using default parameters. Differential expression tables were
824 annotated with the Annotate DESeq2/DEXSeq output tables tool (v1.1.0) in Galaxy to include

825 the following: GFF feature type = “CDS”, GFF feature identifier = “gene_id”, GFF transcript
826 identifier = “transcript_id”, GFF attribute to include = “protein_id, product”. The “protein_id” was
827 used to query the NCBI database and the NCBI Sequence Viewer was used to investigate the
828 genomic context surrounding genes of interest.

829 **Comparative genomics**

830 The nucleotide sequence for the *Enterocloster bolteae* DSM 15670 *ucd* operon (NCBI
831 NZ_CP022464 REGION: complement(4417875..4421605)) was used as a query for BLASTn
832 (megablast) searches using the refseq_genomes database limited to Bacteria (taxid:2). The
833 NCBI multiple sequence alignment (MSA) viewer was used to download alignment figures.

834 **RT-PCR analysis of *E. bolteae* to determine *ucd* operon structure**

835 Isolated RNA samples (500 ng) were reverse transcribed using the LunaScript® RT
836 Master Mix Kit (Primer-free) (NEB) in a reaction volume of 10 µL containing the ucdCFO_RT-
837 PCR_r primer at a final concentration of 1 µM. The No-RT Control included in the kit was used
838 as a no-enzyme control for reverse transcription. The reaction mixtures were incubated in a
839 thermal cycler: 10 min at 55˚C, 1 min at 95˚C. PCR reactions were conducted using the OneTaq
840 2X Master Mix with Standard Buffer (NEB). The ucdCFO_RT-PCR primer pair was added to
841 the master mix (to a final concentration of 0.2 µM) and 1 µL of template (cDNA, -RT, no
842 template, or gDNA) was added for a total reaction volume of 20 µL. PCR tubes were placed in
843 a thermal cycler and targets were amplified according to the following conditions: 20s at 94˚C,
844 31 cycles (20s at 94˚C, 30s at 62˚C, 3 min at 68˚C), 5 min at 68˚C. A volume of 5 µL of reaction
845 was directly loaded onto a 1% agarose gel (made with 1X TAE buffer) containing SafeView
846 Classic (Abm). PCR product sizes were compared to the Quick-Load® Purple 1 kb Plus DNA
847 Ladder (NEB). The rest of the PCR product was then run on a 1% agarose gel and bands
848 corresponding to the desired products were cut out and purified using the Monarch DNA Gel

849 Extraction kit (NEB). DNA was quantified using the Qubit dsDNA HS assay kit (Invitrogen) and
850 submitted to Plasmidsaurus for long-read sequencing using Oxford Nanopore Technologies
851 (Supplementary Information).

852 **RT-qPCR analysis of *E. bolteae ucd* operon genes**

853 Isolated RNA samples (500 ng) were reverse transcribed using the iScript Reverse
854 Transcription Supermix (Bio-Rad) in a reaction volume of 10 µL. The iScript No-RT Control
855 Supermix was used as a no enzyme control for reverse transcription (-RT). The reaction
856 mixtures were incubated in a thermal cycler: 5 min at 25 °C, 20 min at 48 °C, and 1 min at 95
857 °C. Both cDNA and -RT controls were diluted 1/20 in nuclease-free water before use. qPCR
858 reactions were conducted using the Luna Universal qPCR Master Mix kit (NEB). The
859 Eb_ucdO_qPCR, Eb_ucdF_qPCR, Eb_ucdC_qPCR, and Eb_dnaK_Ref_qPCR primer pairs
860 were added to their respective master mixes (final primer concentration of 250 nM) and 6.6 or
861 4.4 µL of diluted template (cDNA, -RT, no template) were added to 26.4 or 17.6 µL of master
862 mix for triplicates or duplicates, respectively. All cDNA samples were run in technical triplicates,
863 while other sample types were run in technical duplicates. Replicate mixes were pipetted (10
864 µL/well) into a MicroAmp Fast 96-Well Reaction Plate (Applied Biosystems) and the plates were
865 sealed, then spun down for 2 min to eliminate air bubbles. The qPCR detection parameters
866 were as follows: SYBR Green detection, ROX reference dye, 10 µL reaction volume. The
867 thermal cycling conditions were: 1 min at 95 °C, 40 cycles (15 s at 95 °C, 30 s at 60 °C), then
868 melt analysis (60-95 °C). Data were analyzed according to the $2^{-\Delta\Delta Ct}$ method [64] with the dnaK
869 gene serving as the reference gene (*E. bolteae* dnaK RNA-seq $\log_2$FC = 0.122).

870 **Protein extraction**

871 All steps other than sonication were carried out under anaerobic conditions. To extract
872 proteins, 10 mL of treated (50 µM urolithin C for 4 h) *Enterocloster* spp. culture (see

873 *Enterocloster* spp. urolithin C treatments) were pelleted (6,500 g for 3 min) and the supernatant
874 was discarded. The pellet was washed with 10 mL of pre-reduced PBS, pelleted again, and
875 resuspended in 0.5 mL of pre-reduced lysis buffer (20 mM Tris, pH 7.5, 500 mM NaCl, 10 mM
876 $MgSO_4$, 10 mM $CaCl_2$, and 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free)). The
877 resuspended pellet was then sonicated on ice using a Misonix Sonicator 3000 set to power
878 level 2/10 according to the following sequence (aerobically, in a cold room): 20 s ON, 40 s OFF,
879 for a total of 2 min ON. Tubes were centrifuged at 20,000 x g for 2 min to pellet insoluble
880 particles and 0.4 mL of lysate was transferred to a new 1.5 mL tube (kept on ice). Lysates used
881 in metabolism assays were transported to the anaerobic chamber in a sealed plastic bag
882 containing an anaerobic gas generating system to minimize loss in activity.

883 **Urolithin metabolism assays using crude lysates from uroC-induced *E. bolteae***

884 Protein lysates (described above) were aliquoted (50 µL aliquots) into 1.5 mL tubes, then
885 treated with DMSO or urolithin C (10 mM stock) at a final concentration of 350 µM. Cofactors
886 (NADPH, NADP, NADH, and FAD, each dissolved to a final concentration of 30 mM (in lysis
887 buffer immediately before the assay was run) were added individually to the lysates at a final
888 concentration of 1 mM. The lysates were incubated at room temperature in an anaerobic
889 chamber for 20 h prior to freezing at -70 °C. Samples were then extracted using *Extraction*
890 *Method B*.

891 To assess the oxygen sensitivity of crude lysates from uroC-induced *E. bolteae*, samples were
892 prepared as described above. After adding DMSO or uroC and NADH (under anaerobic
893 conditions), tubes were either incubated at room temperature inside the anaerobic chamber or
894 just outside of the chamber for 20 h. Afterwards, samples were frozen at -70 °C and then
895 extracted using *Extraction Method B*.

896 **Proteomics analysis of uroC-treated *E. bolteae***

897 Extracted proteins were submitted for proteomic analysis at the RI-MUHC. For each
898 sample, protein lysates were loaded onto a single stacking gel band to remove lipids,
899 detergents, and salts. The gel band was reduced with DTT, alkylated with iodoacetic acid, and
900 digested with trypsin. Extracted peptides were re-solubilized in 0.1% aqueous formic acid and
901 loaded onto a Thermo Acclaim Pepmap (Thermo, 75 um ID X 2 cm C18 3 um beads) precolumn
902 and then onto an Acclaim Pepmap Easyspray (Thermo, 75 um ID X 15 cm with 2 um C18
903 beads) analytical column separation using a Dionex Ultimate 3000 uHPLC at 250 nL/min with
904 a gradient of 2-35% organic (0.1% formic acid in acetonitrile) over 3 hours. Peptides were
905 analyzed using a Thermo Orbitrap Fusion mass spectrometer operating at 120,000 resolution
906 (FWHM in MS1) with HCD sequencing (15,000 resolution) at top speed for all peptides with a
907 charge of 2+ or greater. The raw data were converted into *.mgf format (Mascot generic format)
908 for searching using the Mascot 2.6.2 search engine (Matrix Science) against *Enterocloster*
909 *bolteae* DSM 15670 proteins (NCBI assembly GCF_002234575.2) and a database of common
910 contaminant proteins. Mascot was searched with a fragment ion mass tolerance of 0.100 Da
911 and a parent ion tolerance of 5.0 ppm. O-63 of pyrrolysine, carboxymethyl of cysteine and j+66
912 of leucine/isoleucine indecision were specified in Mascot as fixed modifications. Deamidation
913 of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable
914 modifications.

915 The database search results were loaded into Scaffold Q+ Scaffold_5.0.1 (Proteome Sciences)
916 for statistical treatment and data visualization. Scaffold (v5.3.0) was used to validate MS/MS
917 based peptide and protein identifications. Peptide identifications were accepted if they could be
918 established at greater than 95.0% probability by the Peptide Prophet algorithm [65] with Scaffold
919 delta-mass correction. Protein identifications were accepted if they could be established at

920 greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities
921 were assigned by the Protein Prophet algorithm [66]. Proteins that contained similar peptides and
922 could not be differentiated based on MS/MS analysis alone were grouped to satisfy the
923 principles of parsimony. Proteins sharing significant peptide evidence were grouped into
924 clusters. Protein quantification and differential expression were determined in Scaffold using
925 the following parameters: Quantitative method was set to total spectra, the minimum value was
926 set to 0.5 in case proteins were not detected in one condition, and statistical tests were
927 performed using Fisher's exact test with the Benjamini-Hochberg multiple test correction at a
928 significance level set to 0.05.

929 **Protein structures**

930 The protein FASTA sequences (NCBI RefSeq accessions for UcdC: WP_002569575.1,
931 UcdF: WP_002569574.1, UcdO: WP_002569573.1) for the *Enterocloster bolteae* DSM 15670
932 *ucd* operon were used as a query for BLASTp searches against the UniProtKB reference
933 proteomes + Swiss-Prot databases. The AlphaFold2 protein structures for matches (UniProt
934 accessions for UcdC: A8RZR5, UcdF: A8RZR3, UcdO: A8RZR2) of each protein were
935 downloaded and imported into PyMOL (v2.4.1). Foldseek (in 3Di/AA mode) was used to
936 generate a list of proteins with similar structures from solved crystal structures in the Protein
937 Data Bank (PDB) [44].

938 Hits of published X-ray crystal structures (PDB: 1ZXI (from *Afipia carboxidovorans* OM5) and
939 PDB: 3UNI (from *Bos taurus*)) were fetched from the PDB and imported into PyMOL. The
940 AlphaFold2 structures for each Ucd protein were aligned to the following chains in the published
941 PDBs using the "super" command: PDB 1ZXI: UcdC to chain C, UcdF to chain A, UcdO to chain
942 B; PDB 3UNI: UcdC, UcdF, and UcdO to chain A.

943 **curatedMetagenomicData meta-analysis of *Enterocloster ucd* operon in human fecal**
944 **metagenome datasets**

945 All 93 metagenomic studies (22,588 samples and their metadata) available in the
946 curatedMetagenomicData R package [47] (v3.8.0) were downloaded locally (ExperimentHub
947 snapshotDate(): 2023-04-24, accessed on 2023-06-06) and transferred to the Narval cluster
948 hosted by the Digital Research Alliance of Canada. Metagenomic data for urolithin C-
949 metabolizing *Enterocloster* spp. were obtained by querying the "relative_abundance" (pre-
950 processed using MetaPhlAn3 and "gene_families" (pre-processed using HUMAnN3) entries in
951 individual study datasets [67]. For individual taxa (containing partial strings “bolteae”, “citroniae”,
952 “asparagiformis”, “asparagiforme”, “Enterocloster”, or “47FAA” (corresponding to *L.*
953 *pacaense*)), relative abundance (%) was extracted from the rows of the "relative_abundance"
954 datasets using the stringr R package (v1.5.0, https://github.com/tidyverse/stringr). Prevalence
955 (relative abundance in sample > 0) was then calculated for each sample.

956 For specific genes, the NCBI protein accessions for each gene of the *ucd* operon (*ucdO*, *ucdF*,
957 *ucdC*) was used to search the UniProt database. UniRef90 accession numbers corresponding
958 to hits (C5EGQ4, G5HFF3, A8RZR5, respectively) were then extracted from the rows of the
959 "gene_families" datasets using the stringr R package. Prevalence (abundance in sample > 0)
960 was then calculated for each sample. R scripts and RData files are available in Zenodo (see
961 Data Availability).

962 **Fecal slurry preparation and treatment**

963 Frozen (-70 °C) fecal samples were brought into the anaerobic chamber and allowed to
964 thaw. The samples were suspended in 1 mL mABB medium per 0.1 g feces and homogenized
965 by breaking apart large pieces with a sterile loop and by vortexing. Large particles were pelleted
966 by centrifuging the tubes at 700 x g for 3 min. The supernatants (containing bacteria) were

967 transferred to new tubes and centrifuged at 6,500 x g for 5 min to pellet the cells. The
968 supernatants were discarded, and the cell pellets were washed with 5 mL of fresh media. The
969 cell suspensions were once again centrifuged at 6,500 x g for 5 min and the resulting cell pellets
970 were resuspended in 600 µL per 0.1 g feces. Resuspended cells were treated with either 100
971 µM urolithin C or an equivalent volume DMSO and incubated at 37˚C anaerobically for 48 h.
972 200-300 µL volumes were removed from the batch cultures and immediately frozen at -70˚C
973 for later extraction of urolithins (using *Extraction Method A*), DNA, and RNA.

974 **Genomic DNA extraction from microbial communities**

975 A 300 µL fecal slurry aliquot (between 50-100 mg wet weight) was pelleted (10,000 g for
976 5 min) and the supernatant was removed for later LC-MS analysis. The pellet was then mixed
977 with 750 µL of ZymoBIOMICS lysis solution (Zymo Research) and transferred to a ZR
978 BashingBead lysis tube (Zymo Research). Samples were lysed in a Mini Beadbeater 16
979 (Biospec) according to the following sequence: 1 min ON, 5 min OFF for a total of 5 min ON.
980 DNA was then purified using the ZymoBIOMICS DNA Miniprep Kit (Zymo Research) according
981 to the manufacturer's instructions (including the *OneStep* PCR Inhibitor Removal step). Purified
982 DNA samples were quantified using the Qubit dsDNA HS assay kit (Invitrogen).

983 **Long read 16S rRNA sequencing of microbial communities**

984 Long read 16S PCR reactions were conducted using the Platinum SuperFi II Green PCR
985 Master Mix (Invitrogen). The ONT_16S_27F_GGK and ONT_16S_1492R_GGK primer pairs
986 (see *Primer sequences table*) were added to their respective master mixes (final primer
987 concentration of 0.2 µM) and 1 µL of template (~10 ng) was added (for a total reaction volume
988 of 25 µL). PCR tubes were placed in a thermal cycler and targets were amplified according to
989 the following cycling conditions: 30 s at 98 °C, 30 cycles (10 s at 98 °C, 10 s at 60 °C, 30 s at
990 72 °C), 5 min at 72 °C, and hold at 4 °C. Amplicons were quantified using the Qubit dsDNA HS

991 assay kit (Invitrogen) to verify that amplicon concentrations were reasonably balanced (range
992 = 18.36-24.00 ng/μL). Barcoding of amplicons was performed with 2 μL of PCR reaction
993 according to the manufacturer's instructions (for ONT kit SQK-AMB111-24). Barcoding
994 reactions were incubated in a thermal cycler for 10 min at 65 °C, then for 2 min at 80 °C. 10 μL
995 of each barcoding reaction were pooled and proteins were digested using heat-labile proteinase
996 K (NEB) by incubating the pooled library for 15 min at 37 °C, followed by heat inactivation for
997 10 min at 55 °C. Amplicons were purified using Agencourt AMPure XP beads (Beckman
998 Coulter) using 0.7 volumes of beads-to-library. Following 70% EtOH washes and drying steps,
999 the library was eluted using 15 μL of the provided elution buffer (EB), yielding a library with a
1000 concentration of 30 ng/μL using the Qubit dsDNA HS assay kit (Invitrogen). 11 μL of the eluted
1001 DNA library were transferred to a new tube and combined with 1 μL of Rapid Adapter T (RAP
1002 T). This mixture was incubated at room temperature for 10 min. Since the library was
1003 concentrated, it was diluted 1:2 in EB before combining with SB II and LB II, then loaded into a
1004 primed Flongle Flow Cell (R9.4.1) in a MinION device following the manufacturer's instructions.
1005 Sequencing was allowed to proceed for ~20 h until pore exhaustion or enough reads were
1006 obtained. Base calling & demultiplexing was performed using Guppy (v6.4.6) using the "SUP"
1007 super high accuracy model for R9.4.1 flow cells. The raw reads were filtered for a length
1008 between 1500 ± 200 bp. Filtered reads were assigned to taxa using Emu [68] (v3.4.4, GitLab
1009 Project ID: 19618062) by mapping 16S rRNA sequences to the emu_database database
1010 (based on the NCBI 16S RefSeq with the entry for *E. asparagiformis* changed to the sequence
1011 obtained by ONT sequencing (GenBank accession PP280819) since the RefSeq sequence for
1012 this bacterium contained multiple N nucleotides that biased the assignment of *E. asparagiformis*
1013 to *E. lavalensis*). Count tables were then used to create a phyloseq (v1.44.0,
1014 https://github.com/joey711/phyloseq) object in R [69]. Stacked bar plots were generated using

1015 ggnested (v0.1.0, https://github.com/gmteunisse/ggnested) and fantaxtic (v0.2.0,
1016 https://github.com/gmteunisse/Fantaxtic), and β-diversity analysis was performed using the
1017 phyloseq ordinate() and plot_ordination() functions.

1018 **Total RNA extraction from microbial communities**

1019 A 300 µL fecal slurry aliquot (treated with either 100 µM urolithin C or an equivalent
1020 volume DMSO for 48 h) was thawed and pelleted (6,500 g for 3 min). The pellet (in 200 µL
1021 media) was then mixed with 800 µL TRI reagent (Zymo Research). Samples were lysed in a
1022 Mini Beadbeater 16 (Biospec) according to the following sequence: 1 min ON, 5 min OFF for a
1023 total of 5 min ON. RNA isolation was then performed using the Direct-zol RNA Miniprep Kit
1024 (Zymo Research) according to the manufacturer's instructions (including an on-column DNase
1025 digestion). To ensure complete DNA removal, an additional DNA digestion step was performed
1026 on the isolated RNA using the Ambion DNA-free DNA Removal Kit (Invitrogen) according to the
1027 manufacturer's instructions. The DNA-free RNA was then cleaned up using the OneStep PCR
1028 Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were verified by Qubit
1029 RNA BR assay kit (Invitrogen) and 1% agarose gel electrophoresis.

1030 **RT-PCR analysis of the *ucd* operon in microbial communities**

1031 Total RNA was extracted from frozen fecal slurries as previously described (see Total
1032 RNA extraction from microbial communities), and subsequently reverse transcribed as
1033 described above (see *RT-PCR analysis of E. bolteae ucd operon structure*) in a reaction volume
1034 of 5 µL. PCR reactions were conducted using the OneTaq 2X Master Mix with Standard Buffer
1035 (NEB). The ucdCFO_RT-PCR primer pair was added to the master mix (to a final concentration
1036 of 0.2 µM) and 1 µL of template (cDNA, -RT, or no template) was added for a total reaction
1037 volume of 20 µL. PCR tubes were placed in a thermal cycler and targets were amplified
1038 according to the following conditions: 30s at 94˚C, 45 cycles (30s at 94˚C, 1 min at 61˚C, 4 min

1039 at 68˚C), 5 min at 68˚C. A volume of 10 µL of reaction was directly loaded onto a 1% agarose
1040 gel (made with 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were
1041 compared to the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB).

1042 **PCR analysis of the *ucd* operon in microbial communities**

1043 Genomic DNA (gDNA) was extracted from frozen fecal slurries as previously described
1044 (see Genomic DNA extraction from microbial communities). PCR reactions and product
1045 visualization was conducted on the gDNA as described above (see RT-PCR analysis of the *ucd*
1046 operon in microbial communities). In this case, 5 µL of PCR product was loaded onto the gels
1047 instead of 10 µL.

1048 **Statistical analyses and graphing**

1049 Statistical methods were not used to determine sample sizes, experiments were not
1050 randomized, and the investigators were not blinded. Data points related to uroC metabolism
1051 and RT-qPCR were assumed to be normally distributed, though this was not formally tested.
1052 Correlation analyses were performed using the non-parametric Spearman rank correlation (ρ).
1053 Statistical tests on bacterial relative abundances were performed using the Kruskal-Wallis test
1054 on untransformed relative abundance values, which are skewed towards 0. Statistical analyses
1055 for large datasets are detailed in the relevant methods sections. Details related to each test
1056 performed are supplied in the figure legends. In all cases, $\alpha = 0.05$ and tests were two-tailed.
1057 Data were plotted in GraphPad Prism (v10.0.0) or using the ggplot2 (v3.4.2) R package. Figures
1058 were assembled in Affinity Designer (v1.10.6.1665).

1059 **Data Availability**

1060 RNA-seq reads were deposited in the NCBI SRA BioProject ID PRJNA996126 under
1061 BioSample accession codes SAMN36514640 (*Enterocloster bolteae* DSM 15670) and
1062 SAMN36514641 (*Enterocloster asparagiformis* DSM 15981). Reviewer link:
1063 https://dataview.ncbi.nlm.nih.gov/object/PRJNA996126?reviewer=fpbuj6eeuebv6mij3pme3mp
1064 8ep.

1065 Untargeted proteomics data have been deposited to the ProteomeXchange Consortium via the
1066 PRIDE partner repository with the dataset identifier PXD048514 and 10.6019/PXD048514. [70]

1067 Oxford Nanopore 16S rRNA sequencing reads of healthy human fecal slurries were deposited
1068 in the NCBI SRA BioProject ID PRJNA1073957. Reviewer link:
1069 https://dataview.ncbi.nlm.nih.gov/object/PRJNA1073957?reviewer=nh1a04lg59enlti318qs23s
1070 6ot.

1071 The 16S rRNA sequence for *E. asparagiformis* DSM 15670 used in the Emu database search
1072 was deposited in GenBank under accession PP280819.

1073 All original code, tables, and RData files obtained from the analysis of
1074 curatedMetagenomicData were deposited in Zenodo
1075 (https://doi.org/10.5281/zenodo.8302320).

1076 **References**

1077 1 Knight, R. *et al.* The Microbiome and Human Biology. *Annu Rev Genomics Hum Genet*
1078 **18**, 65-86 (2017). https://doi.org/10.1146/annurev-genom-083115-022438
1079 2 Taguer, M. & Maurice, C. F. The complex interplay of diet, xenobiotics, and microbial
1080 metabolism in the gut: Implications for clinical outcomes. *Clin Pharmacol Ther* **99**, 588-
1081 599 (2016). https://doi.org/10.1002/cpt.366
1082 3 Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut
1083 microbiome. *Nat Biotechnol* **39**, 105-114 (2021). https://doi.org/10.1038/s41587-020-
1084 0603-3
1085 4 Spanogiannopoulos, P., Bess, E. N., Carmody, R. N. & Turnbaugh, P. J. The microbial
1086 pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol*
1087 **14**, 273-287 (2016). https://doi.org/10.1038/nrmicro.2016.17
1088 5 Ervin, S. M. *et al.* Structural Insights into Endobiotic Reactivation by Human Gut
1089 Microbiome-Encoded Sulfatases. *Biochemistry* **59**, 3939-3950 (2020).
1090 https://doi.org/10.1021/acs.biochem.0c00711
1091 6 Yao, L. *et al.* A biosynthetic pathway for the selective sulfonation of steroidal metabolites
1092 by human gut bacteria. *Nat Microbiol* **7**, 1404-1418 (2022).
1093 https://doi.org/10.1038/s41564-022-01176-y
1094 7 Tomas-Barberan, F. A. & Espin, J. C. Effect of Food Structure and Processing on
1095 (Poly)phenol-Gut Microbiota Interactions and the Effects on Human Health. *Annu Rev*
1096 *Food Sci Technol* **10**, 221-238 (2019). https://doi.org/10.1146/annurev-food-032818-
1097 121615
1098 8 Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*
1099 **555**, 623-628 (2018). https://doi.org/10.1038/nature25979
1100 9 Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R. & Goodman, A. L.
1101 Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature*
1102 **570**, 462-467 (2019). https://doi.org/10.1038/s41586-019-1291-3
1103 10 Maini Rekdal, V., Bess, E. N., Bisanz, J. E., Turnbaugh, P. J. & Balskus, E. P. Discovery
1104 and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism.
1105 *Science* **364**, eaau6323 (2019). https://doi.org/10.1126/science.aau6323
1106 11 Campbell, C. *et al.* Bacterial metabolism of bile acids promotes generation of peripheral
1107 regulatory T cells. *Nature* **581**, 475-479 (2020). https://doi.org/10.1038/s41586-020-
1108 2193-0
1109 12 Paik, D. *et al.* Human gut bacteria produce Tau(Eta)17-modulating bile acid metabolites.
1110 *Nature* **603**, 907-912 (2022). https://doi.org/10.1038/s41586-022-04480-z
1111 13 Cerqueira, F. M., Photenhauer, A. L., Pollet, R. M., Brown, H. A. & Koropatkin, N. M.
1112 Starch Digestion by Gut Bacteria: Crowdsourcing for Carbs. *Trends Microbiol* **28**, 95-
1113 108 (2020). https://doi.org/10.1016/j.tim.2019.09.004
1114 14 Maini Rekdal, V. *et al.* A widely distributed metalloenzyme class enables gut microbial
1115 metabolism of host- and diet-derived catechols. *Elife* **9** (2020).
1116 https://doi.org/10.7554/eLife.50845
1117 15 Garcia-Villalba, R., Selma, M. V., Espin, J. C. & Tomas-Barberan, F. A. Identification of
1118 Novel Urolithin Metabolites in Human Feces and Urine after the Intake of a Pomegranate
1119 Extract. *J Agric Food Chem* **67**, 11099-11107 (2019).
1120 https://doi.org/10.1021/acs.jafc.9b04435
1121 16 Carmody, R. N. *et al.* Diet dominates host genotype in shaping the murine gut microbiota.
1122 *Cell Host Microbe* **17**, 72-84 (2015). https://doi.org/10.1016/j.chom.2014.11.010

1123 17 Carmody, R. N. *et al.* Cooking shapes the structure and function of the gut microbiome.
1124 *Nat Microbiol* **4**, 2052-2063 (2019). https://doi.org/10.1038/s41564-019-0569-4
1125 18 Sonnenburg, E. D. *et al.* Diet-induced extinctions in the gut microbiota compound over
1126 generations. *Nature* **529**, 212-215 (2016). https://doi.org/10.1038/nature16504
1127 19 Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut
1128 microbiota. *Nature* **555**, 210-215 (2018). https://doi.org/10.1038/nature25973
1129 20 Quideau, S., Deffieux, D., Douat-Casassus, C. & Pouysegu, L. Plant polyphenols:
1130 chemical properties, biological activities, and synthesis. *Angew Chem Int Ed Engl* **50**,
1131 586-621 (2011). https://doi.org/10.1002/anie.201000044
1132 21 van Duynhoven, J. *et al.* Metabolic fate of polyphenols in the human superorganism.
1133 *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4531-4538 (2011).
1134 https://doi.org/10.1073/pnas.1000098107
1135 22 Garcia-Villalba, R. *et al.* Urolithins: a Comprehensive Update on their Metabolism,
1136 Bioactivity, and Associated Gut Microbiota. *Mol Nutr Food Res* **66**, e2101019 (2022).
1137 https://doi.org/10.1002/mnfr.202101019
1138 23 Singh, A. *et al.* Direct supplementation with Urolithin A overcomes limitations of dietary
1139 exposure and gut microbiome variability in healthy adults to achieve consistent levels
1140 across the population. *Eur J Clin Nutr* **76**, 297-308 (2022).
1141 https://doi.org/10.1038/s41430-021-00950-1
1142 24 Andreux, P. A. *et al.* The mitophagy activator urolithin A is safe and induces a molecular
1143 signature of improved mitochondrial and cellular health in humans. *Nat Metab* **1**, 595-
1144 603 (2019). https://doi.org/10.1038/s42255-019-0073-4
1145 25 Denk, D. *et al.* Expansion of T memory stem cells with superior anti-tumor immunity by
1146 Urolithin A-induced mitophagy. *Immunity* **55**, 2059-2073 e2058 (2022).
1147 https://doi.org/10.1016/j.immuni.2022.09.014
1148 26 Ryu, D. *et al.* Urolithin A induces mitophagy and prolongs lifespan in C. elegans and
1149 increases muscle function in rodents. *Nat Med* **22**, 879-888 (2016).
1150 https://doi.org/10.1038/nm.4132
1151 27 Ghosh, S., Banerjee, M., Haribabu, B. & Jala, V. R. Urolithin A attenuates arsenic-
1152 induced gut barrier dysfunction. *Arch Toxicol* **96**, 987-1007 (2022).
1153 https://doi.org/10.1007/s00204-022-03232-2
1154 28 Singh, R. *et al.* Enhancement of the gut barrier integrity by a microbial metabolite through
1155 the Nrf2 pathway. *Nat Commun* **10**, 89 (2019). https://doi.org/10.1038/s41467-018-
1156 07859-7
1157 29 Shen, P. X. *et al.* Urolithin A ameliorates experimental autoimmune encephalomyelitis
1158 by targeting aryl hydrocarbon receptor. *EBioMedicine* **64**, 103227 (2021).
1159 https://doi.org/10.1016/j.ebiom.2021.103227
1160 30 Ghosh, S., Haribabu, B. & Jala, V. R. Microbial metabolite Urolithin A induces expansion
1161 of Regulatory T Cells in aryl hydrocarbon receptor (AhR)-dependent manner. *Journal of*
1162 *Immunology* **206**, 113.103-113.103 (2021).
1163 31 Xia, B. *et al.* Urolithin A exerts antiobesity effects through enhancing adipose tissue
1164 thermogenesis in mice. *PLoS Biol* **18**, e3000688 (2020).
1165 https://doi.org/10.1371/journal.pbio.3000688
1166 32 Selma, M. V. *et al.* Isolation of Human Intestinal Bacteria Capable of Producing the
1167 Bioactive Metabolite Isourolithin A from Ellagic Acid. *Front Microbiol* **8**, 1521 (2017).
1168 https://doi.org/10.3389/fmicb.2017.01521
1169 33 Selma, M. V., Beltran, D., Garcia-Villalba, R., Espin, J. C. & Tomas-Barberan, F. A.
1170 Description of urolithin production capacity from ellagic acid of two human intestinal

Gordonibacter species. *Food Funct* **5**, 1779-1784 (2014). https://doi.org/10.1039/c4fo00092g

34 Gaya, P., Peirotén, A., Medina, M., Alvarez, I. & Landete, J. M. Bifidobacterium pseudocatenulatum INIA P815: The first bacterium able to produce urolithins A and B from ellagic acid. *Journal of Functional Foods* **45**, 95-99 (2018). https://doi.org/10.1016/j.jff.2018.03.040

35 Garcia-Villalba, R. *et al.* Metabolism of different dietary phenolic compounds by the urolithin-producing human-gut bacteria Gordonibacter urolithinfaciens and Ellagibacter isourolithinifaciens. *Food Funct* **11**, 7012-7022 (2020). https://doi.org/10.1039/d0fo01649g

36 Iglesias-Aguirre, C. E. *et al.* Gut Bacteria Involved in Ellagic Acid Metabolism To Yield Human Urolithin Metabotypes Revealed. *J Agric Food Chem* **71**, 4029-4035 (2023). https://doi.org/10.1021/acs.jafc.2c08889

37 Iglesias-Aguirre, C. E. *et al.* In vivo administration of gut bacterial consortia replicates urolithin metabotypes A and B in a non-urolithin-producing rat model. *Food Funct* **14**, 2657-2667 (2023). https://doi.org/10.1039/d2fo03957e

38 Hosny, M. *et al.* Clostridium pacaense: a new species within the genus Clostridium. *New Microbes New Infect* **28**, 6-10 (2019). https://doi.org/10.1016/j.nmni.2018.12.003

39 Zupok, A., Iobbi-Nivol, C., Mejean, V. & Leimkuhler, S. The regulation of Moco biosynthesis and molybdoenzyme gene expression by molybdenum and iron in bacteria. *Metallomics* **11**, 1602-1624 (2019). https://doi.org/10.1039/c9mt00186g

40 Leimkuhler, S. & Iobbi-Nivol, C. Bacterial molybdoenzymes: old enzymes for new purposes. *FEMS Microbiol Rev* **40**, 1-18 (2016). https://doi.org/10.1093/femsre/fuv043

41 Koppel, N., Bisanz, J. E., Pandelia, M. E., Turnbaugh, P. J. & Balskus, E. P. Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *Elife* **7** (2018). https://doi.org/10.7554/eLife.33953

42 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org/10.1038/s41586-021-03819-2

43 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439-D444 (2022). https://doi.org/10.1093/nar/gkab1061

44 van Kempen, M. *et al.* Fast and accurate protein structure search with Foldseek. *Nat Biotechnol* (2023). https://doi.org/10.1038/s41587-023-01773-0

45 Resch, M., Dobbek, H. & Meyer, O. Structural and functional reconstruction in situ of the [CuSMoO2] active site of carbon monoxide dehydrogenase from the carbon monoxide oxidizing eubacterium Oligotropha carboxidovorans. *J Biol Inorg Chem* **10**, 518-528 (2005). https://doi.org/10.1007/s00775-005-0006-4

46 Ishikita, H., Eger, B. T., Okamoto, K., Nishino, T. & Pai, E. F. Protein conformational gating of enzymatic activity in xanthine oxidoreductase. *J Am Chem Soc* **134**, 999-1009 (2012). https://doi.org/10.1021/ja207173p

47 Pasolli, E. *et al.* Accessible, curated metagenomic data through ExperimentHub. *Nat Methods* **14**, 1023-1024 (2017). https://doi.org/10.1038/nmeth.4468

48 Cortes-Martin, A. *et al.* The gut microbiota urolithin metabotypes revisited: the human metabolism of ellagic acid is mainly determined by aging. *Food Funct* **9**, 4100-4106 (2018). https://doi.org/10.1039/c8fo00956b

49 Bess, E. N. *et al.* Genetic basis for the cooperative bioactivation of plant lignans by Eggerthella lenta and other human gut bacteria. *Nat Microbiol* **5**, 56-66 (2020). https://doi.org/10.1038/s41564-019-0596-1

50 Le, C. C., Bae, M., Kiamehr, S. & Balskus, E. P. Emerging Chemical Diversity and Potential Applications of Enzymes in the DMSO Reductase Superfamily. *Annu Rev Biochem* **91**, 475-504 (2022). https://doi.org/10.1146/annurev-biochem-032620-110804

51 Little, A. S. *et al.* Dietary- and host-derived metabolites are used by diverse gut bacteria for anaerobic respiration. *Nat Microbiol* **9**, 55-69 (2024). https://doi.org/10.1038/s41564-023-01560-2

52 Pinske, C., Bonn, M., Kruger, S., Lindenstrauss, U. & Sawers, R. G. Metabolic deficiences revealed in the biotechnologically important model bacterium Escherichia coli BL21(DE3). *PLoS One* **6**, e22830 (2011). https://doi.org/10.1371/journal.pone.0022830

53 Xia, M. *et al.* Improvement of Urolithin A Yield by In Vitro Cofermentation of Streptococcus thermophilus FUA329 with Human Gut Microbiota from Different Urolithin Metabotypes. *J Agric Food Chem* (2024). https://doi.org/10.1021/acs.jafc.3c09734

54 Zhang, X. *et al.* Isolation and characterization of a novel human intestinal Enterococcus faecium FUA027 capable of producing urolithin A from ellagic acid. *Front Nutr* **9**, 1039697 (2022). https://doi.org/10.3389/fnut.2022.1039697

55 Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Goker, M. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* **50**, D801-D807 (2022). https://doi.org/10.1093/nar/gkab902

56 Pauff, J. M., Cao, H. & Hille, R. Substrate Orientation and Catalysis at the Molybdenum Site in Xanthine Oxidase: CRYSTAL STRUCTURES IN COMPLEX WITH XANTHINE AND LUMAZINE. *J Biol Chem* **284**, 8760-8767 (2009). https://doi.org/10.1074/jbc.M804517200

57 Meier-Kolthoff, J. P., Goker, M., Sproer, C. & Klenk, H. P. When should a DDH experiment be mandatory in microbial taxonomy? *Arch Microbiol* **195**, 413-418 (2013). https://doi.org/10.1007/s00203-013-0888-4

58 Meier-Kolthoff, J. P. & Goker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* **10**, 2182 (2019). https://doi.org/10.1038/s41467-019-10210-3

59 Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016). https://doi.org/10.1093/bioinformatics/btw354

60 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011). https://doi.org/10.14806/ej.17.1.200

61 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-915 (2019). https://doi.org/10.1038/s41587-019-0201-4

62 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014). https://doi.org/10.1093/bioinformatics/btt656

63 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8

64 Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101-1108 (2008). https://doi.org/10.1038/nprot.2008.73

65 Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* **74**, 5383-5392 (2002). https://doi.org/10.1021/ac025747h

66 Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* **75**, 4646-4658 (2003). https://doi.org/10.1021/ac0341261

67 Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* **10**, e65088 (2021). https://doi.org/10.7554/eLife.65088

68 Curry, K. D. *et al.* Emu: species-level microbial community profiling of full-length 16S rRNA Oxford Nanopore sequencing data. *Nat Methods* **19**, 845-853 (2022). https://doi.org/10.1038/s41592-022-01520-4

69 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217 (2013). https://doi.org/10.1371/journal.pone.0061217

70 Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-D552 (2022). https://doi.org/10.1093/nar/gkab1038

1284 **Acknowledgements**

1285 This research was funded by the Canadian Institutes of Health Research (CIHR) grant PJT-
1286 437944 to B.C. and C.F.M. and a Weston Family Microbiome Initiative's 2021 Transformational
1287 Research program to B.C. B.C. holds a tier II Canada Research Chair (CRC) in Therapeutic
1288 Chemistry. C.F.M. holds a tier II CRC in Gut Microbial Physiology and is a Azrieli Global Scholar
1289 in the Humans & Microbiome program. R.P. is supported by the CIHR Canada Graduate
1290 Scholarship-Doctoral (#493808 for 2023-2026) and by the Fonds de Recherche du Québec-
1291 Santé : Bourse de formation au doctorat (#316063 for 2022-2023). M.S. is supported by the
1292 CIHR Canada Graduate Scholarship to Honor Nelson Mandela (#DF2-187718 for 2023-2026),
1293 the Fonds de Recherche du Québec-Santé: Bourse de formation au doctorat (#311071 for
1294 2022-2023).

1295 RNA sample preparation and sequencing was performed by Génome Québec. Untargeted
1296 proteomics sample preparation and analysis was performed by the Research Institute of the
1297 McGill University Health Centre (RI-MUHC) Proteomics and Molecular Analysis Platform.
1298 Small-molecule MALDI-TOF mass spectrometry was performed by Mark Hancock at the McGill
1299 SPR-MS Facility. Molecular biology experiments were enabled by the McGill University Imaging
1300 and Molecular Biology Platform (IMBP). High-performance cloud computing was enabled by
1301 Calcul Québec (https://www.calculquebec.ca/) and the Digital Research Alliance of Canada
1302 (https://alliancecan.ca/). RNA-seq data analysis was performed using two instances of Galaxy:
1303 (https://usegalaxy.org/) and Compute Canada Genetics and Genomics Analysis Platform
1304 (GenAP) (https://www.genap.ca/).

1305 **Author information**

1306 **Authors and Affiliations**

1307 *Department of Pharmacology & Therapeutics, McGill University, 3655 Prom. Sir-William-Osler,*

1308 *Montreal, Quebec, H3G 1Y6, Canada*

1309 Reilly Pidgeon, Sacha Mitchell, Layan Suleiman, Lharbi Dridi & Bastien Castagner

1310 *Department of Microbiology & Immunology, McGill University, 3775 University Street, Montreal,*

1311 *Quebec, H3A 2B4, Canada*

1312 Michael Shamash & Corinne F. Maurice

1313 **Contributions**

1314 **R.P.** designed the study, performed microbiology experiments, performed bioinformatics
1315 analyses, analyzed data, created figures, and wrote the initial manuscript with B.C. **S.M.**
1316 performed experiments on microbial communities, analyzed data, and created figures. **M.S.**
1317 performed Nanopore sequencing and processed raw sequencing reads. **L.S.** performed
1318 microbiology experiments. **L.D.** consulted on experimental design and methodology. **C.M.**
1319 supervised research, obtained ethical approval for the use of human fecal samples, and
1320 obtained research funding. **B.C.** designed the study, supervised the research, obtained
1321 research funding, and wrote the initial manuscript with R.P. All authors reviewed and edited the
1322 manuscript.

1323 **Corresponding authors**

1324 Correspondence to Bastien Castagner (bastien.castagner@mcgill.ca).

1325 **Ethics declarations**

1326 The authors declare no competing interests.