The dietary ellagitannin metabolite urolithin A is produced by a molybdenum-dependent dehydroxylase encoded by prevalent human gut *Enterocloster* spp.

Reilly Pidgeon[1], Sacha Mitchell[1], Michael Shamash[2], Layan Suleiman[1], Lharbi Dridi[1], Corinne F. Maurice[2, 3], Bastien Castagner[1*]

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

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

[3] McGill Centre for Microbiome Research, Montreal, Quebec, Canada

Abstract

Urolithin A (uroA) is a polyphenol derived from the multi-step metabolism of dietary ellagitannins by the human gut microbiota that can affect host health by stimulating mitophagy. Most individuals harbor a microbiota capable of uroA production; however, the mechanisms underlying the dehydroxylation of its catechol-containing dietary precursor (uroC) are unknown. Here, we use a combination of untargeted bacterial transcriptomics, proteomics, and comparative genomics to uncover an inducible uroC dehydroxylase (*ucd*) operon in *Enterocloster* spp. We show that *Enterocloster* spp. are sensitive to iron chelation by uroC, and dehydroxylation to uroA rescues growth by disrupting the iron-binding catechol. Importantly, only microbiota samples actively transcribing *ucd* could produce uroA, establishing *ucd*-containing *Enterocloster* spp. as keystone urolithin metabolizers. Overall, this work identifies *Enterocloster* spp. and the *ucd* operon as main contributors to uroA production and establishes a multi-omics framework to further our mechanistic understanding of polyphenol metabolism by the human gut microbiota.

26 **Main**

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

36 Diet is a strong modulator of the composition and function of the gut microbiota [16-19]. Diet-
37 derived polyphenols are a diverse class of plant secondary metabolites found in fruits,
38 vegetables, and nuts (reviewed in [20]) that are poorly absorbed by the host and reach the large
39 intestine relatively intact [7,21]. Ellagitannins are a large sub-group of polyphenols that belong to
40 the family of hydrolysable tannins and are characterized by a central glucose (open-chain or
41 pyranose forms) linked to diverse pyrogallol-like moieties [20]. Camu camu, a berry rich in the
42 ellagitannin castalagin, has been shown to impact anti-cancer immunity via the gut microbiome,
43 and is currently in clinical trials (NCT05303493, NCT06049576) in combination with immune
44 checkpoint inhibitors [22,23]. Depending on microbiota composition, ellagitannins can be
45 hydrolyzed and reduced by gut bacteria into bioactive metabolites (ellagic acid, urolithins,
46 nasutins) according to different metabolic phenotypes characterized by the terminal metabolites
47 observed in biological fluids [24] (Supplementary Fig. 1).

48 Urolithin A (uroA) is the most common terminal metabolite of ellagitannin metabolism and has
49 reported pharmacological activities both within the gut environment and systemically following
50 absorption [25]. In the gut, uroA can attenuate colitis by increasing the expression of epithelial
51 tight junction proteins [26-28] *via* the activation of aryl hydrocarbon receptor (AhR)-Nrf2 pathways
52 [29,30]. Additionally, uroA can enhance immunotherapy in colorectal cancer models by activating
53 Pink1-dependent mitophagy pathways in T cells, improving anti-tumor CD8+ T cell immunity [31].
54 Clinical trials in healthy individuals have demonstrated that uroA is safe, bioavailable, and can
55 be detected in its aglycone, glucuronidated, and sulfated forms in plasma [25,27]. Once absorbed
56 by the host, uroA can trigger mitophagy in muscle cells, improving muscle function in animal
57 models of ageing and Duchenne muscular dystrophy [26,29,32,33]. Overall, uroA can enhance gut
58 barrier integrity, modulate the immune system, and promote mitochondrial health in the host,
59 thus showing promise as a postbiotic to treat age-related conditions [34-36].

60 While urolithin metabolism is prevalent in human populations, few gut bacteria have been
61 reported to metabolize urolithins [34-36]. Most known urolithin metabolizers belong to the
62 *Eggerthellaceae* family (*Gordonibacter urolithinfaciens*, *Gordonibacter pamelaeae*, *Ellagibacter*
63 *isourolithinifaciens*) and can perform multiple metabolic steps in the urolithin metabolism
64 pathway, yielding either urolithin C (uroC) or isourolithin A (isouroA) from ellagic acid [37].
65 Recently, certain members of the *Enterocloster* spp. (*Lachnospiraceae* family) were reported
66 to dehydroxylate uroC to uroA and isouroA to urolithin B (uroB) both *in vitro* and *in vivo* [38,39].
67 These findings shed light on the minimal bacterial community required for the complete
68 metabolism of ellagic acid to uroA; however, the genes and enzymes responsible for these
69 dehydroxylation reactions remain unknown (Fig. 1A).

70 Here, we use a multi-omics enzyme identification framework to uncover uroC dehydroxylase
71 (*ucd*) genes and enzymes in *Enterocloster* spp. and their relative, *Lachnoclostridium pacaense*.

72 We find that the UcdCFO enzyme complex specifically dehydroxylates 9-hydroxy urolithins and
73 that both metabolizing species and *ucd* genes are prevalent and actively transcribed in human
74 feces during *ex vivo* metabolism. We further demonstrate that *Enterocloster* spp. growth is
75 delayed by uroC and that dehydroxylation may be a mechanism to inactivate its iron-binding
76 catechol moiety. Our study sheds light on the genetic and chemical basis underlying the
77 complex reciprocal interactions between urolithins and the gut microbiota.

Results

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

Members of the *Enterocloster* spp. have previously been shown to dehydroxylate uroC *in vitro* [38] and *in vivo* [39] under anaerobic conditions (Fig. 1A, full metabolic pathway in Supplementary Fig. 1). To determine the prevalence of uroC metabolism within this genus, we incubated all available *Enterocloster* spp. type strains (Methods) with uroC and quantified urolithin concentrations by liquid chromatography-mass spectrometry (LC-MS). Of the tested bacteria, only *E. asparagiformis* , *E. bolteae* , and *E. citroniae* dehydroxylated uroC to produce uroA (Fig. 1B). Interestingly, uroC metabolism was not predicted by phylogeny, as uroC-metabolizing species did not cluster based on 16S rRNA genes, genomes, or proteomes (Fig. 1C, Supplementary Fig. 2A,B, respectively), suggesting gain or loss of metabolic gene clusters throughout the evolution of *Enterocloster* spp. Based on these results, we chose to perform more in-depth analysis on *E. asparagiformis* and *E. bolteae* to identify the metabolic gene clusters involved in uroC dehydroxylation.

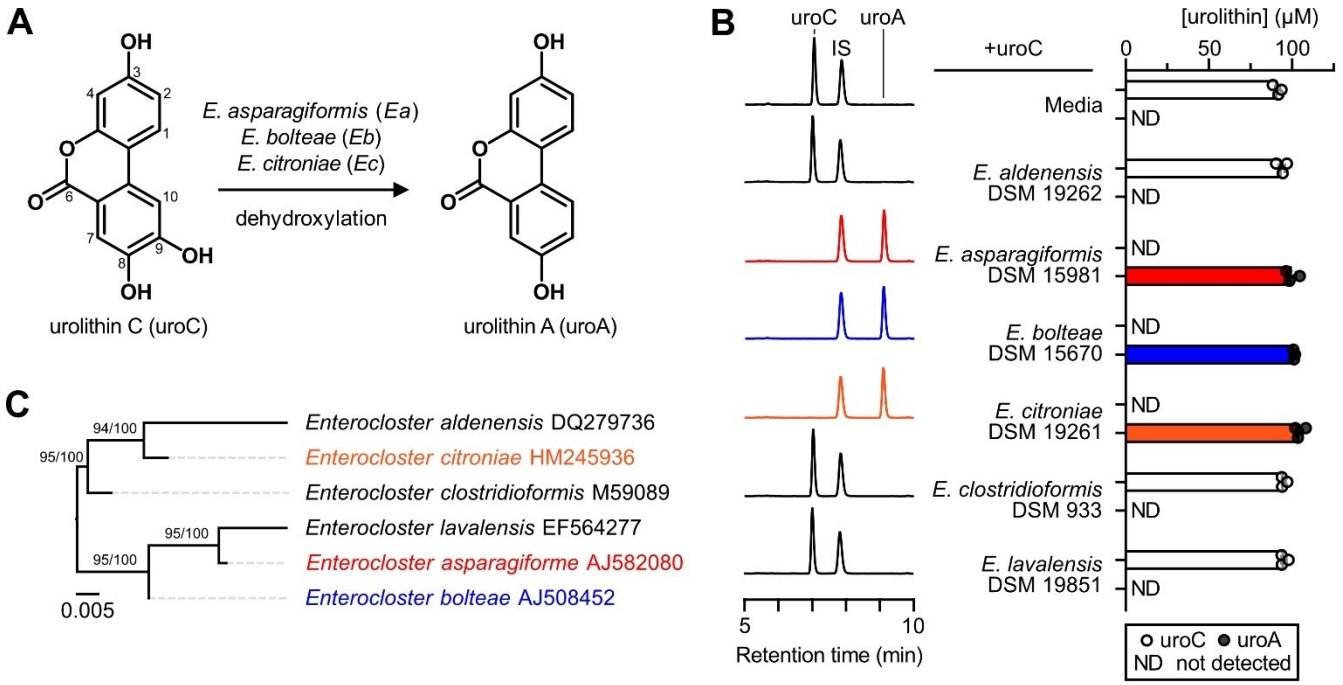


Figure 1. Urolithin C metabolism by *Enterocloster* spp. is not predicted by phylogeny.

A) Reaction scheme of uroC dehydroxylation by gut resident *Enterocloster* spp. via unknown enzymes. **B)** LC-MS screen of *Enterocloster* spp. type strains for dehydroxylation activity. 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. Left: Representative chromatograms (λ = 305 nm) for each experimental group (from one representative biological replicate). The same scale was used for each chromatogram. Right: Quantification of urolithin peak areas relative to a salicylic acid internal standard (IS) (n = 3 biological replicates). Data are represented as mean ± SEM. **C)** Phylogenetic tree of tested *Enterocloster* spp. type strain 16S rRNA sequences constructed using the Genome-to-Genome Distance Calculator (GGDC) Phylogeny Server [40]. 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. The GenBank accession numbers are provided to the right of each taxon. Source data and statistical details are provided as a Source data file.

A putative urolithin C dehydroxylase metabolic gene cluster is upregulated upon urolithin C treatment.

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

We next sought to determine whether uroC metabolism is inducible or constitutive. To test for inducibility, both bacteria were treated with DMSO or uroC during exponential growth, then washed and resuspended in PBS, yielding cell suspensions unable to synthesize new proteins. Metabolism of uroC to uroA was inducible as only cells originating from bacteria grown with

124 uroC were capable of uroA production (Fig. 2E). Consequently, we performed RNA-sequencing
125 to compare gene expression in DMSO and uroC-treated cultures of *E. asparagiformis* and *E.*
126 *bolteae*. Since uroA was detected in both bacterial cultures as soon as 2 h post spike-in (Fig.
127 2C), this timepoint was selected to isolate mRNA.

128 RNA sequencing of uroC-induced cultures revealed a distinct gene cluster, which we term uroC
129 dehydroxylase (*ucd*), that was highly and differentially expressed ($\log_2$FC > 8) in both *E.*
130 *asparagiformis* (Fig. 2F) and *E. bolteae* (Fig. 2G). In both bacteria, these clusters contained
131 adjacent genes that were expressed to similar $\log_2$FC values: a xanthine dehydrogenase family
132 protein subunit M, a (2Fe-2S)-binding protein, and a molybdopterin-dependent oxidoreductase
133 (Fig. 2F,G). These genes will hereafter be referred to as *ucdC* (for coenzyme), *ucdF* (for
134 ferredoxin), and *ucdO* (for oxidoreductase), respectively. Interestingly, we also observed an
135 upregulation of genes involved in efflux (MepA-like multidrug and toxin extrusion (MATE)
136 transporters) and iron transport (FecCD-like) (Fig. 2F,G), suggesting a link between uroC
137 metabolism, iron uptake, and efflux.

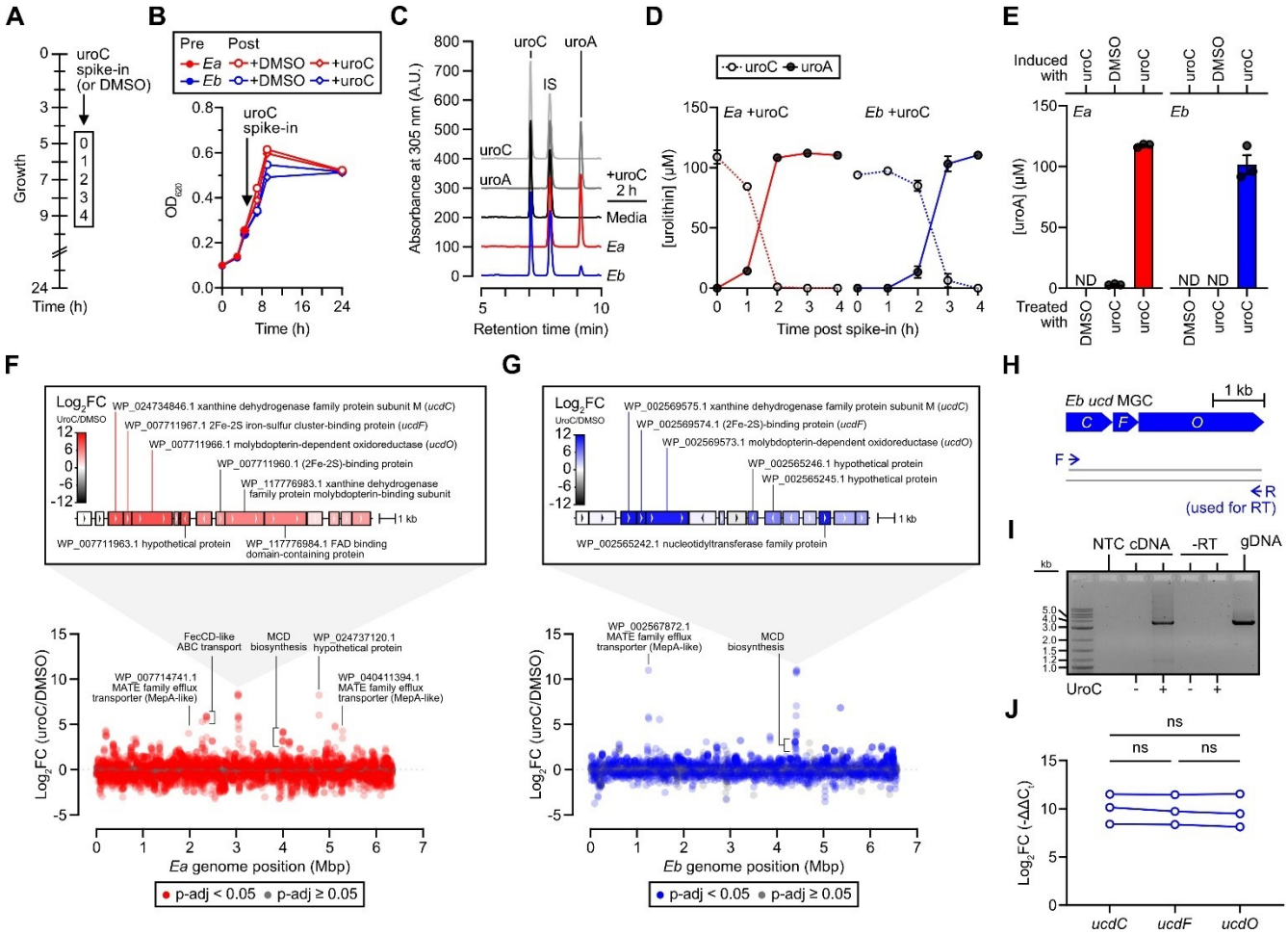


Figure 2. Urolithin C treatment upregulates a putative dehydroxylase operon.

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-seq (F-G) results are matched. **B)** Growth curve (optical density (OD) at 620 nm) of DMSO or uroC-spiked *E. asparagiformis* (*Ea*) and *E. bolteae* (*Eb*) type strain cultures according to the design in (A). 200 μL of culture were sampled at each timepoint and OD_{620} was measured in a 96-well plate (n = 4 biological replicates). The same sampled culture was then frozen and extracted for analysis by LC-MS (C,D). **C)** Representative chromatograms (λ = 305 nm) of cultures sampled 2 h post-spike-in (from one representative biological replicate). The same scale was used for each chromatogram. **D)** Quantification of urolithin concentrations from peak areas relative to a salicylic acid internal standard (IS) over 4 h in uroC-spiked *Ea* and *Eb* type strain cultures (n = 4 biological replicates). **E)** Quantification of urolithin A concentrations in DMSO- or uroC-treated *Ea* and *Eb* cell suspensions. Cell suspensions were prepared from *Ea* and *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 DMSO or 100 μM uroC (n = 3 biological replicates). **F,G)** Manhattan plots of genes altered by uroC treatment in *Ea* (F) and *Eb* (G) based on DESeq2 analysis (n = 4 biological replicates). Data points are colored according to their adjusted p-value (based on the Benjamini-Hochberg-corrected Wald statistic). Grey, p-adj ≥ 0.05. Red or blue, p-adj < 0.05 for *Ea* and *Eb*, respectively. The genomic organization around the differentially expressed genes (generated from the NCBI Sequence

159 Viewer) is depicted above Manhattan plots, which show the most highly and differentially
160 expressed genes by RNA-seq. Genes are colored according to their log_2FC values. NCBI
161 accessions for select proteins encoded by highlighted genes are provided. **H)** Primer design for
162 RT-PCR (I) experiment targeting the *Eb ucd* gene cluster. The same reverse primer was used
163 for both the reverse transcription step and the subsequent PCR reaction. **I)** 1% agarose gel
164 image of RT-PCR amplicons using primers (H) that span the full-length *Eb ucd* gene cluster
165 (from one biological replicate). NTC, no template control. **J)** RT-qPCR expression of each gene
166 in the *Eb ucd* operon. Growing *Eb* cultures were treated with DMSO or uroC (100 μM) for 2 h
167 before RNA isolation and reverse transcription (n = 3 biological replicates). Gene expression
168 profiles of each target gene in the *Eb ucd* gene cluster displayed as log_2FC (equivalent to -
169 $\Delta\Delta C_t$, where $\Delta\Delta C_t = \Delta C_{t\ uroC} - \Delta C_{t\ DMSO}$) with lines connecting paired biological replicates;
170 repeated-measures one-way ANOVA with Tukey's multiple comparisons test; ns, not
171 significant. Data are represented as mean ± SEM (behind symbols) in (B,D,E). FC, fold change
172 (uroC/DMSO); Source data and statistical details are provided as a Source data file.

173

174 **The *ucd* metabolic gene cluster is organized in an operon.**

175 We next sought to characterize the *ucd* metabolic gene cluster in *E. bolteae* since this
176 bacterium is considered a core species of the gut microbiome [8]. Based on the proximity, sense,
177 and expression levels of each of the three genes by RNA-seq (Fig. 2G), we hypothesized that
178 all three genes in the cluster were organized in an operon. We designed a gene-specific RT-
179 PCR assay that would enable the detection of full-length polycistronic *ucdCFO* genes using
180 cDNA from DMSO- or uroC-treated *E. bolteae* as a template (Fig. 2H). An amplicon of the
181 expected size (~3.6 kb) was detected only in cDNA derived from uroC-treated *E. bolteae*,
182 validating the inducibility of these genes (Fig. 2I). Long-read sequencing of the obtained
183 amplicon yielded a sequence corresponding to the *E. bolteae ucdCFO* metabolic gene cluster
184 with 100% identity (Supplementary Sequence 1). Using an independent set of *E. bolteae*
185 cultures, we then performed RT-qPCR on DMSO- or uroC-treated *E. bolteae* with all three
186 genes in the putative operon as targets. Similar to our RNA-seq results, all three genes were
187 highly induced (mean log_2FC ≥ 9.7 for all three *ucd* genes) relative to DMSO controls and were
188 expressed at the same level (Fig. 2J). These results indicate that the *ucdCFO* genes are
189 transcribed as a single polycistronic mRNA and therefore form a uroC-inducible operon.

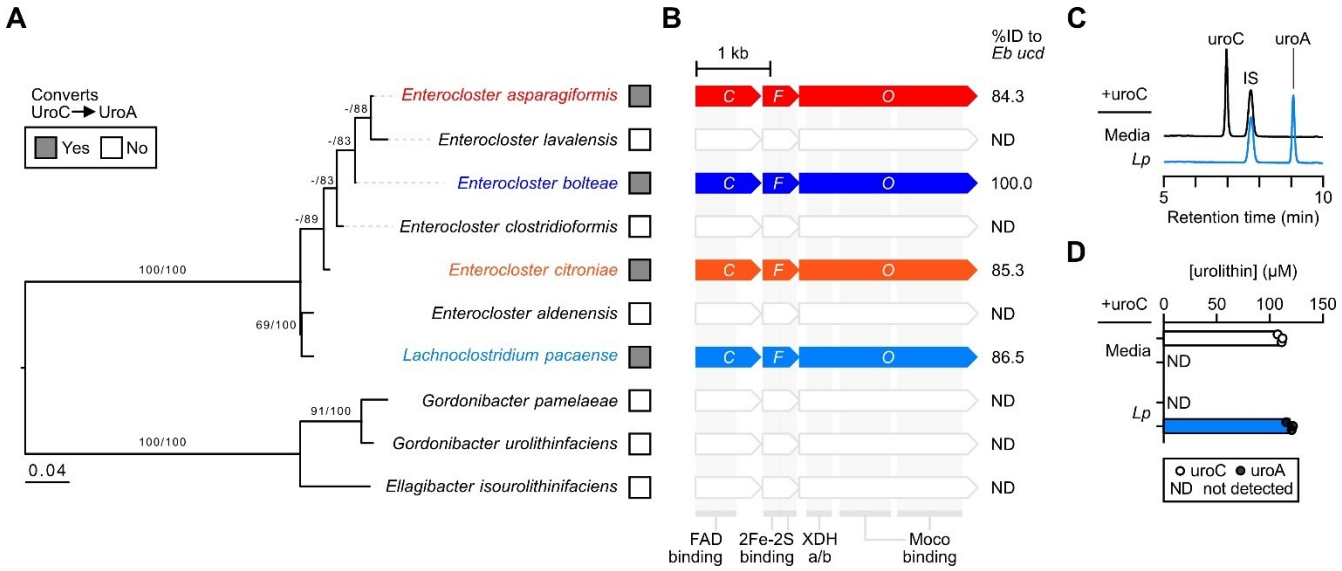
190 **The *ucd* operon is induced by 9-hydroxy urolithins.**

191 Next, we aimed to determine the substrate scope of the *ucd* operon. Multiple urolithins
192 possess pyrogallol, catechol, and phenol structural motifs that are dehydroxylated at various
193 positions by gut bacteria (Supplementary Fig. 1). Interestingly, *E. bolteae* only metabolizes the
194 9-position hydroxyl group of urolithins (Supplementary Fig. 3A-D) and does not require adjacent
195 hydroxyl groups since isouroA is dehydroxylated to uroB (Supplementary Fig. 3D) [38]. Since
196 dehydroxylation of urolithins in *E. bolteae* is position specific, we hypothesized that the *ucd*
197 operon would be induced by other 9-hydroxy urolithins (uroM6 and isouroA). Therefore, we
198 performed RT-qPCR on DMSO-, uroM6-, uroC-, or isouroA-treated *E. bolteae* cultures using
199 the *ucdO* gene as a target. Each urolithin significantly induced the expression of the *ucd* operon
200 to a similar extent (Supplementary Fig. 3E-F). In addition, *E. bolteae* cell suspensions induced
201 with uroC were capable of dehydroxylating uroM6 (Supplementary Fig. 3G) and isouroA
202 (Supplementary Fig. 3H), indicating that the same proteins induced by uroC can metabolize
203 structurally similar 9-hydroxy urolithins. Thus, it is likely that the same metabolic enzymes,
204 encoded by the *E. bolteae ucd* operon, are acting on 9-hydroxy urolithins.

205 **Presence of *ucd* operon homologs in genomes predicts urolithin C metabolism by gut**
206 **bacteria.**

207 We wondered whether novel metabolizers of uroC could be discovered based on
208 nucleotide sequence homology to the *ucd* operon. Homology searches using the *E. bolteae*
209 *ucd* operon sequence confirmed that only uroC-metabolizing *Enterocloster* spp. (*E.*
210 *asparagiformis*, *E. bolteae*, and *E. citroniae*) possessed homologs of the *ucdCFO* genes with a
211 similar organization (Supplementary Fig. 4). In addition, the gut bacterium *Lachnoclostridium*
212 *pacaense* [41] was identified as another hit (Supplementary Fig. 4). The type strain of this
213 bacterium (CCUG 71489T = Marseille-P3100) was closely related to *Enterocloster* spp. based

214 on 16S rRNA, whole genome, and whole proteome phylogenies (Fig. 3A, Supplementary Fig.
215 2A,B, respectively). *L. pacaense* possessed genomic sequences with high homology (86.5%
216 nucleotide identity) and identical functional annotations to the *E. bolteae ucd* operon sequence
217 (Fig. 3B). When incubated with uroC, *L. pacaense* CCUG 71489T quantitatively produced uroA
218 (Fig. 3C,D). We searched for homologs of the *E. bolteae ucd* in the genomes of urolithin- and
219 catechol-metabolizing bacteria belonging to the *Eggerthellaceae* but could not identify any hits.
220 Notably, *Eggerthellaceae* lack 9-hydroxy urolithin dehydroxylase activity [35,37], which correlates
221 with an absence of *ucd*-like operons in their genomes (Fig. 3B). These comparative genomics
222 data indicate that the presence of a *ucd* operon in genomes predicts uroC metabolism by gut
223 bacteria.



224

225 **Figure 3. Urolithin C metabolism correlates with *ucd* operon prevalence in gut bacteria.**

226 **A)** Phylogenetic tree of *Enterocloster* spp., *Lachnoclostridium pacaense* (*Lp*), and catechol-
227 metabolizing *Eggerthellaceae* type strain 16S rRNA sequences constructed using the Genome-
228 to-Genome Distance Calculator (GGDC) Phylogeny Server [40]. Maximum likelihood (ML) tree
229 inferred under the GTR+GAMMA model and rooted by midpoint-rooting. The branches are
230 scaled in terms of the expected number of substitutions per site. The numbers above the
231 branches are support values when larger than 60% from ML (left) and maximum parsimony
232 (right) bootstrapping. Bacteria that convert uroC to uroA are labeled with grey squares. **B)** NBCI
233 Multiple Sequence Aligner viewer hits for BLASTn searches using the *E. bolteae* DSM 15670
234 *ucd* operon nucleotide sequence as a query against the NCBI refseq_genomes database. Only

235 hits with ≥ 90 % query coverage and species-level taxonomic resolution are displayed with %
236 identity to the query sequence. Domain annotations for each gene are denoted below according
237 to InterPro annotations for corresponding proteins. ND, Not detected; Moco, Molybdenum
238 cofactor. **C,D)** *In vitro* metabolism of uroC by *Lp*. UroC (100 μM) was added to cultures (in
239 mABB+H media) at the start of growth and urolithins were extracted after 24 h anaerobic
240 incubation, then analyzed by LC-MS. **C)** Representative chromatograms (λ = 305 nm) for each
241 experimental group (from one representative biological replicate). The same scale was used
242 for each chromatogram. **D)** Quantification of urolithin peak areas relative to a salicylic acid
243 internal standard (IS) (n = 3 biological replicates). Data are represented as mean ± SEM.
244 Source data and statistical details are provided as a Source data file.

245

246 **A molybdopterin cofactor biosynthetic gene cluster is upregulated upon urolithin C**
247 **treatment.**

248 In addition to the three genes in the *ucd* operon, we observed a significant increase
249 ($log_2FC \geq 2.6$) in 9 molybdopterin cytosine dinucleotide (MCD) biosynthesis genes upon uroC
250 treatment (Fig. 2F,G, Supplementary Fig. 5A). These 9 genes, which recapitulate the function
251 of 10 genes in *E. coli* [42], are involved in molybdenum cofactor biosynthesis (*moaAC*, *mogA*,
252 *moeA*), molybdate ion transport (*modABCE*), cytosine addition to the molybdenum cofactor
253 (*mocA*), and MCD cofactor insertion into the active site (*xdhC*) (Supplementary Fig. 5B) [42]. All
254 9 genes cluster in the genomes of *E. asparagiformis* and *E. bolteae* and are organized into 2
255 adjacent operons (Supplementary Fig. 5C) that are induced upon uroC treatment. Based on
256 sequence homology to *E. coli* oxidoreductases and MCD biosynthetic machinery, proteins
257 encoded by the *ucd* operon belong to the xanthine dehydrogenase family [43]. These findings
258 imply that uroC dehydroxylation is MCD-dependent, which differs from the bis-molybdopterin
259 guanine dinucleotide requirement of catechol dehydroxylases in *Eggerthellaceae* [10,14].

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

261 Since oxidoreductases utilise a variety of cofactors and coenzymes for catalytic activity,
262 we sought to determine the redox coenzymes and conditions necessary for uroC

263 dehydroxylation. Therefore, we performed metabolism assays using crude lysates from uroC-
264 induced *E. bolteae*. As crude lysates alone did not metabolize uroC, various redox coenzymes
265 (NADPH, NADH, and FAD) were added to lysates to promote uroC dehydroxylation (Fig. 4A).
266 Only NADH-treated lysates yielded quantitative dehydroxylation of uroC to uroA compared to
267 the no cofactor control (Fig. 4A). Interestingly, the addition of free FAD partially inhibited uroC
268 dehydroxylation in NADH-treated lysates (Fig. 4A), likely by decreasing the free NADH pool.
269 NADPH, which differs from NADH by a phosphate group on the 2'-OH group of the adenosine
270 moiety, was unable to promote uroC dehydroxylation, indicating some specificity in the redox
271 cofactors necessary for dehydroxylation. Aerobic incubation of crude lysates supplemented
272 with NADH completely inhibited uroC dehydroxylation (Fig. 4B), indicating that the active
273 enzyme complex requires a strictly anaerobic environment for dehydroxylation, as has been
274 demonstrated for various metalloenzymes [44].

275 To confirm that *ucd* operon-encoded proteins were expressed in *E. bolteae* crude lysates, we
276 performed untargeted proteomics and compared protein expression upon DMSO or uroC
277 treatment. Indeed, all 3 proteins encoded by the *ucd* operon (UcdC, UcdF, and UcdO) were the
278 most differentially expressed proteins in the uroC treatment group (Fig. 4C). In addition,
279 proteins involved in MCD biosynthesis were also strongly increased upon uroC treatment (Fig.
280 4C,D, Supplementary Fig. 5A), pointing to the coordination between MCD biosynthesis and
281 active UcdCFO oxidoreductase assembly. These multi-omics datasets implicate all three
282 *ucdCFO* genes and MCD biosynthesis genes in the metabolism of uroC to uroA, as
283 demonstrated by the strong positive correlation between transcript and protein differential
284 expression (Fig. 4D).

285 To validate the function of the *E. bolteae ucd* operon, we attempted heterologous expression
286 of *E. bolteae* UcdCFO in *E. coli*; however, all expression and activity assays were unsuccessful

287 despite the inclusion of *mocA* and *xdhC* genes involved in MCD maturation in our expression
288 plasmids. This lack of activity likely resulted from the choice of heterologous host and from the
289 complex assembly of active molybdoenzymes [43]. We therefore attempted to express UcdCFO
290 in the phenol-degrading soil bacterium *Rhodococcus erythropolis* using a thiostrepton-inducible
291 expression system [45] (pTipQC2-*ucdCFO*, Supplementary Fig. 6A,B), previously used to
292 express the anaerobic *E. lenta* Cgr2 protein [44]. Despite the poor yield of soluble protein
293 (Supplementary Fig. 6C,D), we were able to observe uroM6 and uroC dehydroxylation at the
294 9-position in crude lysates of *R. erythropolis* transformed with pTipQC2-*ucdCFO*, but not in the
295 no insert control (pTipQC2) (Fig. 4E, Supplementary Fig. 6E-G), thus confirming that the *ucd*
296 operon confers 9-hydroxy urolithin dehydroxylase activity.

297 To gain an understanding of the structural organization of proteins encoded by the *ucd* operon,
298 we performed modeling using AlphaFold2 [46,47]. Structures of each protein encoded by the *E.*
299 *bolteae ucd* operon (Fig. 4F) were superposed onto published X-ray crystal structures of
300 xanthine dehydrogenase family enzymes with similar folds [48], yet from different taxonomic
301 domains: *Afipia carboxidovorans* carbon monoxide dehydrogenase [49] and *Bos taurus* xanthine
302 dehydrogenase [50]. The 3 proteins encoded by the *ucd* operon formed subunits in an
303 oxidoreductase complex with a similar quaternary structure to the published crystal structures
304 (Fig. 4G, Supplementary Fig. 7A,B). The predicted quaternary structure of the UcdCFO enzyme
305 complex supported a complete electron transport chain whereby electrons would flow from
306 reduced FAD to two 2Fe-2S clusters, then to the MCD cofactor, and finally to uroC as the
307 terminal electron acceptor (Fig. 4H, Supplementary Fig. 7C,D). This model supports our
308 findings in crude lysates whereby NADH serves as an electron donor to reduce UcdC-bound
309 FAD (Fig. 4A). Using homology modeling, we further identified the putative uroC binding site in
310 UcdO, which overlaps with the salicylic acid ligand in the *Bos taurus* xanthine dehydrogenase

structure (Supplementary Fig. 7E). 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 (Supplementary Fig. 7F), orienting it towards the molybdenum cofactor.

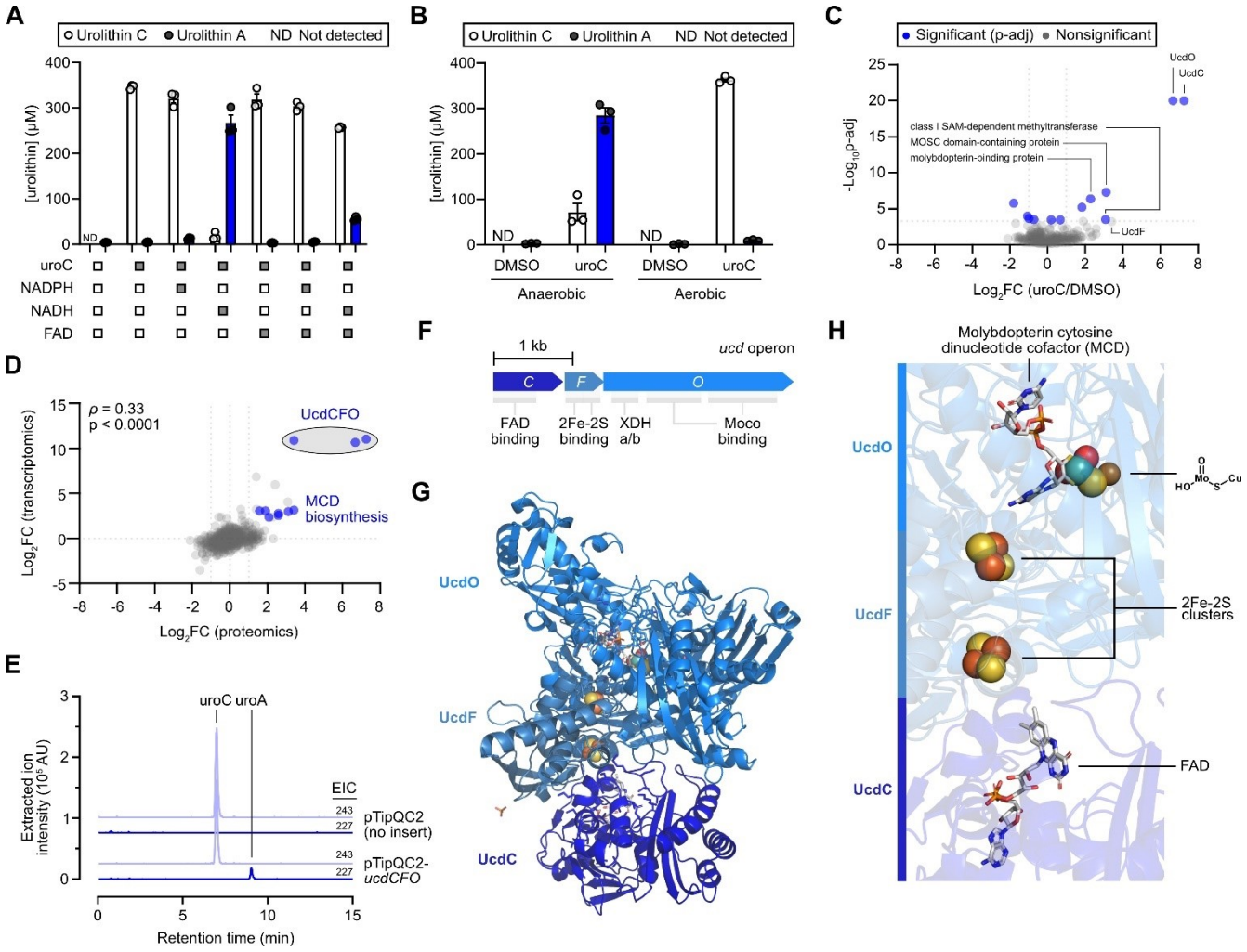


Figure 4. The UcdCFO complex enables anaerobic electron transport from NADH to uroC.

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). 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). Data are represented as mean ± SEM for (A,B). **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)** LC-MS extracted ion chromatograms (EIC) of uroC ($[M-H]^-$ = 243) and uroA ($[M-H]^-$ = 227) from a representative anaerobic uroC dehydroxylation assay using crude lysates of *R. erythropolis* harboring either pTipQC2 (no insert) or pTipQC2-*ucdCFO* plasmids. **F)** Domains of genes in the *ucd* operon based on InterPro annotations. **G)** 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). **H)** 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.

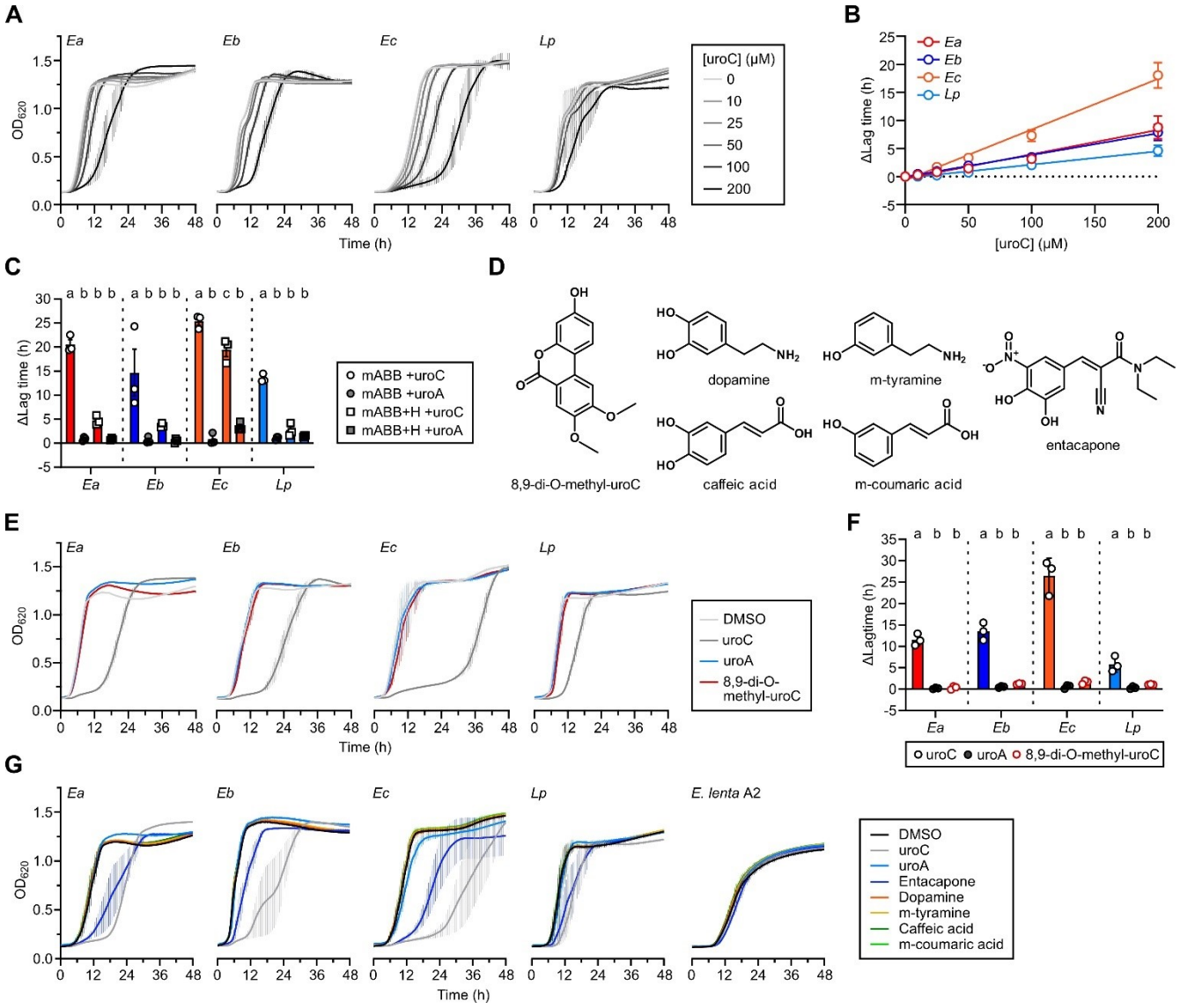
Disruption of the urolithin C catechol moiety rescues growth delay in iron-limited media

To gain an understanding of why *Enterocloster* spp. and *L. pacaense* metabolize 9-hydroxy urolithins, we performed growth experiments in different media conditions. When uroC was added prior to growth in rich medium containing hemin (mABB+H), a concentration-dependent increase in lag time was observed for all uroC-metabolizing bacteria (Fig. 5A,B). As catechols are common structural motifs in iron-binding siderophores [51], we hypothesized that uroC was delaying growth by altering iron availability in the growth medium *via* its catechol moiety. Incubation of *Enterocloster* spp. and *L. pacaense* in medium lacking added iron (mABB) exacerbated the growth delay by uroC (Fig. 5C, Supplementary Fig. 8A,B); however, this growth delay was partially rescued upon supplementation of different iron sources (hemin, $Fe(II)SO_4$, or Fe(III) pyrophosphate) (Fig. 5C, Supplementary Fig. 8C). To validate that iron chelation could extend the lag time of *Enterocloster* spp. and *L. pacaense*, we incubated all four uroC-metabolizers with 2,2'-bipyridyl (biP) in mABB media. As observed with uroC, biP delayed the growth of all tested bacteria, but supplementation of $Fe(II)SO_4$, or Fe(III) pyrophosphate could partially rescue growth delay (Supplementary Fig. 8C). Interestingly,

357 uroA, which lacks a catechol moiety, did not impact the growth of the tested bacteria in either
358 mABB or mABB+H media compared to uroC (Fig. 5C, Supplementary Fig. 8A,B). To confirm
359 that the catechol moiety of uroC was responsible for delaying growth, we synthesized a
360 methylated analogue of uroC (8,9-di-O-methyl-uroC, Fig. 5D, Supplementary Fig. 9A-F), which
361 is unable to bind iron [52], and tested its effect on growth in mABB. Like uroA, 8,9-di-O-methyl-
362 uroC did not delay growth of uroC-metabolizing bacteria (Fig. 5E,F). These data demonstrate
363 that both the catechol moiety of uroC and iron availability are essential to uroC-mediated lag
364 phase extension.

365 Dehydroxylation of catechols by gut bacteria has been observed for diverse classes of
366 compounds like neurotransmitters, therapeutic drugs, and diet-derived polyphenols [14]. Although
367 catechol dehydroxylation can promote growth in some species [14], we hypothesized that
368 dehydroxylation could be a mechanism used by gut bacteria to inactivate catechol-containing
369 compounds that affect their fitness. To determine whether diverse catechols can delay growth,
370 uroC-metabolizing bacteria and the dopamine-metabolizing *Eggerthella lenta* A2 were
371 incubated with catechol-containing compounds and their dehydroxylated counterparts: uroC
372 (uroA), entacapone, dopamine (m-tyramine), caffeic acid (m-coumaric acid) (Fig. 5D).
373 Surprisingly, neither dopamine nor caffeic acid (and their dehydroxylated counterparts) delayed
374 the growth of the tested bacteria (Fig. 5G). On the other hand, both uroC and the nitrocatechol-
375 containing Parkinson's drug entacapone delayed the growth of *Enterocloster spp*. and *L.*
376 *pacaense* but did not affect *E. lenta* A2 (Fig. 5G) [8]. Thus, catechol-containing compounds show
377 differential effects on the growth of gut bacteria, depending on their structure. These results
378 prompted us to investigate the effect of uroC on a more diverse panel of gut bacteria including
379 *E. aldenensis, E. clostridioformis*, and *E. lavalensis*, which do not metabolize uroC, along with
380 *Gordonibacter* spp., which produce uroC from dietary ellagic acid [35]. Treatment with uroC

delayed the growth of *E. aldenensis* , *E. clostridioformis,* and *E. lavalensis* to varying extents (Supplementary Fig. 10A); however, there was no difference in growth between the DMSO-, uroC-, and uroA-treated cultures of *Enterococcus faecium* and *Gordonibacter* spp. (Supplementary Fig. 10B,C). Thus, all *Enterocloster* spp. tested showed sensitivity to uroC-mediated lag time extension, while other bacteria were insensitive to its effects on growth.



387 **Figure 5. The catechol moiety of uroC delays *Enterocloster* spp. growth in an iron-**
388 **dependent manner.**

389 **A)** Growth curves (optical density (OD) at 620 nm) of uroC-metabolizing *Enterocloster* spp. and
390 *L. pacaense* treated with increasing concentrations of uroC in rich mABB+H media (7.7 μM
391 hemin) (n = 3 biological replicates). Data are represented as mean ± SEM. **B)** Quantification of
392 the difference in lag time compared to the DMSO control for growth curves in (A). Data are
393 represented as mean ± SEM; lines were fitted using simple linear regression. **C)** Quantification
394 of the difference in lag time of *Enterocloster* spp. and *L. pacaense* grown in mABB (no added
395 iron) or mABB+H (7.7 μM hemin) compared to respective DMSO controls for growth curves in
396 Supplementary Fig. 8A,B. Data are represented as mean ± SEM; repeated measures two-way
397 ANOVA (matching by biological replicate) with Tukey's multiple comparisons test. Significant
398 differences between treatments for individual bacteria are denoted by a different lowercase
399 letter above each plot. **D)** Structures of tested catechols and their derivatives. **E)** Growth curves
400 (OD at 620 nm) of uroC-metabolizing *Enterocloster* spp. and *L. pacaense* grown in mABB
401 media and treated with DMSO (vehicle), 100 μM of uroC, uroA, or 8,9-di-O-methyl-uroC (n = 3
402 biological replicates). Data are represented as mean ± SEM. **F)** Quantification of the difference
403 in lag time compared to the DMSO control for growth curves in (E). Data are represented as
404 mean ± SEM; repeated measures two-way ANOVA (matching by biological replicate) with
405 Tukey's multiple comparisons test. Significant differences between treatments for individual
406 bacteria are denoted by a different lowercase letter above each plot. **G)** Growth curves (OD at
407 620 nm) of uroC-metabolizing *Enterocloster* spp., *L. pacaense*, and *E. lenta* A2 grown in mABB
408 media and treated with DMSO (vehicle), 100 μM of uroC, uroA, entacapone, dopamine, m-
409 tyramine, caffeic acid or m-coumaric acid. Data are represented as mean ± SEM.

410

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

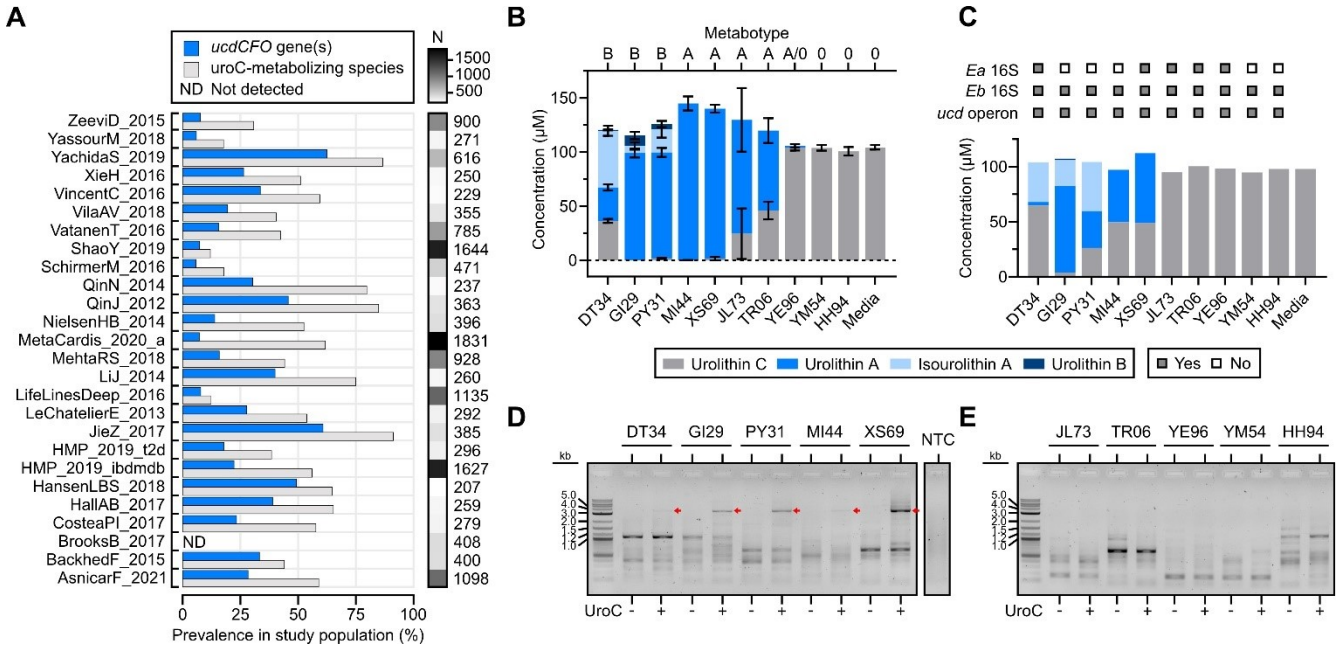
413 We next wondered whether uroC-metabolizing *Enterocloster* spp. and their *ucd* operons
414 were prevalent and active in human fecal samples. We first utilized uniformly processed
415 metagenomic data from the curatedMetagenomicData R package [53]. After filtering for fecal
416 samples (86 studies, n = 21,030 subjects), we counted the prevalence of at least one uroC-
417 metabolizing species and at least one *ucd* gene homolog (Methods). The prevalence of both
418 features was variable across studies (Fig. 6A for studies with >200 participants, Supplementary
419 Fig. 11A for all studies). Combining all studies, the prevalence of at least one uroC-metabolizing
420 species and at least one *ucdCFO* gene homolog was 9,343/21,030 (44.9%) and 4,356/21,030

421 (20.7%), respectively. *E. bolteae* was the most prevalent and abundant uroC-metabolizing
422 species detected in gut metagenomes (Supplementary Fig. 11B,C) and correlated strongly with
423 ucd abundance (Supplementary Fig. 11D). These findings suggest that uroC-metabolizing
424 *Enterocloster* spp. and *ucd* operon genes are prevalent in human fecal metagenomic samples
425 and reflect the variable urolithin metabolism profiles (metabotypes) in the general population
426 [25,54].

427 Next, we performed *ex vivo* metabolism assays to determine whether *Enterocloster* spp. could
428 metabolize uroC in the context of a complex community. Fecal slurries from 10 healthy
429 individuals were first profiled according to their uroC metabotypes (Fig. 6B) [24]. Individuals
430 clustered into metabotypes A (only uroA produced), B (uroA and isouroA/uroB), and 0 (no
431 terminal urolithin metabolites). While we observed metabotypes A and B in uroC-metabolizing
432 fecal slurries, all slurries produced some amount of uroA from uroC. Stools JL73, TR06, and
433 YE96 displayed variable metabolism patterns and did not metabolize uroC in some
434 experiments, likely reflecting differences in activity between aliquots of feces (Fig. 6B). We then
435 repeated metabolism assays using fecal slurries from all 10 healthy individuals and extracted
436 urolithins, DNA, and RNA from each culture. In this experiment, only 5/10 fecal slurries
437 metabolized uroC to uroA (Fig. 6C). We hypothesized that differences in metabotypes could be
438 explained by microbial composition. Therefore, long-read V1-V9 16S rRNA sequencing was
439 performed on fecal slurries. Both DMSO- and uroC-treated fecal slurries within individuals had
440 similar microbial compositions and diversity metrics (Supplementary Fig. 12A-D) but showed
441 differences in composition between individuals and metabolism status (Supplementary Fig.
442 12A,E). Surprisingly, all samples contained 16S rRNA sequences mapping to *E. bolteae*, and
443 many non-metabolizing fecal slurries contained *E. asparagiformis* (Fig. 6C, Supplementary Fig.
444 12B). We then assayed genomic DNA from treated fecal slurries for the presence of the *ucd*

operon by PCR and found that 10/10 individuals (19/20 conditions) yielded a detectable amplicon of the expected size (~3.6 kb) (Supplementary Fig. 12F,G). These data indicate that the prevalence of uroC-metabolizing *Enterocloster* spp. 16S rRNA and *ucd* operon genes does not predict metabolism in fecal samples.

We then surmised that the *ucd* operon would be transcribed only in fecal slurries actively metabolizing uroC. Using a gene-specific reverse primer that binds to *ucdO* (Fig. 2H), the full-length *ucd* operon was reverse transcribed and amplified in RNA extracted from DMSO- and uroC-treated fecal slurries. An amplicon (~3.6 kb) corresponding to the *ucd* operon was only detected in uroC-metabolizing fecal slurries (Fig. 6D) when treated with uroC and entirely absent from non-metabolizing slurries (Fig. 6E). This amplicon was absent in no reverse transcriptase controls, indicating no gDNA contamination (Supplementary Fig. 12H,I). These data demonstrate that *ucd* transcription correlates with uroC metabolism in complex fecal communities and that *E. bolteae* is keystone species involved in urolithin A production.



459 **Figure 6. The *ucd* operon is prevalent in metagenomes and actively transcribed in**

460 **urolithin C-metabolizing human fecal samples.**

461 **A)** Prevalence of *ucd* operon (at least one gene) and of a uroC-metabolizing species (at least
462 one species) in fecal metagenomes from the CuratedMetagenomicData R package. Only
463 studies with ≥200 participants are depicted. All 86 studies are available in Supplementary Fig.
464 11. **B)** Summary of urolithin concentrations in fecal slurries (n = 10 healthy donors) incubated
465 with 100 μM uroC for 48 h. Data are represented as mean ± SEM (n = 3-6 experimental
466 replicates). **C)** Summary of urolithin concentrations in fecal slurries (n = 10 healthy donors)
467 incubated with 100 μM uroC for 48 h. Presence of uroC-metabolizing species and of the *ucd*
468 operon is denoted above the graph if the bacterium or operon was detected in the uroC-treated
469 fecal slurry (Supplementary Fig. 12B,F,G). Data are representative of 1 replicate where DNA
470 and RNA was also extracted from fecal slurries. **D,E)** *ucd* gene-specific RT-PCR on fecal
471 microbiota communities from 10 healthy donors using the primer set described in Fig. 2H.
472 Samples are matched to the urolithin metabolism data in C) and 16S rRNA sequencing in
473 Supplementary Fig. 12A-E. 1% agarose gel of amplicons derived from **D)** uroC-metabolizing
474 and **E)** non-metabolizing fecal slurries. Bands corresponding to the *E. bolteae ucd* operon (~3.6
475 kb) are labeled with red arrows. The no template control (NTC) is the same for (D,E). See
476 Supplementary Fig. 12H,I for the no reverse transcriptase control PCR reactions on the same
477 samples. Source data are provided as a Source data file.

478

479 **Discussion**

480 We identified genes and proteins that are essential for the metabolism of urolithins by
481 gut resident *Lachnospiraceae* through a combination of transcriptomics, comparative
482 genomics, and untargeted proteomics. Our study reveals a novel multi-subunit molybdoenzyme
483 (urolithin C dehydroxylase, Ucd) that catalyzes the dehydroxylation of 9-hydroxy urolithins
484 including uroM6, uroC, and isouroA. Importantly, prevalence analysis in published data and *ex*
485 *vivo* transcriptomics established *E. bolteae* as a keystone urolithin-metabolizing member of the
486 gut microbiota.

487 Catechol dehydroxylases are widespread in gut resident *Eggerthella lenta* and
488 *Gordonibacter* spp. [10,14,55]. These molybdoenzymes, which belong to the DMSO reductase
489 superfamily [56], dehydroxylate substrates like catechol lignan (Cldh), dopamine (Dadh), DOPAC

490 (Dodh), hydrocaffeic acid (Hcdh), and caffeic acid (Cadh), which can promote growth by using
491 these substrates as alternative electron acceptors [14]. A recent survey of reductases in gut
492 bacteria established that most respiratory reductases contain N-terminal signal sequences and
493 are translocated across the cytoplasmic membrane, while non-respiratory reductases, which
494 lack signal sequences, remain in the cytoplasm [57]. The UcdCFO enzyme complex we found in
495 *Enterocloster* spp. differs from catechol dehydroxylases in *Eggerthellaceae* in important ways
496 as it does not require a catechol structural motif for activity, belongs to the xanthine oxidase
497 superfamily, and is composed of 3 subunits that each lack signal sequences. Based on the
498 absence of signal sequences and the cytoplasmic localization of xanthine dehydrogenases, the
499 *ucd* operon likely encodes for a non-respiratory reductase serving a different role than
500 previously characterized catechol dehydroxylases [57].

501 In rich media conditions, uroC, but not uroA, extended the lag phase of growth in both
502 uroC-metabolizing and non-metabolizing *Enterocloster* spp. This growth delay was not
503 observed for other taxa, suggesting that *Enterocloster* spp. are especially sensitive to uroC-
504 mediated iron chelation. In addition to the *ucd* operon, uroC-treated *E. bolteae* and *E.*
505 *asparagiformis* upregulated gene clusters related to efflux (MepA-like MATE family efflux
506 transporters) and iron/siderophore transport (FecCD-like ABC transporter). These responses
507 are analogous to antimicrobial resistance mechanisms raised against entacapone and other
508 non-antibiotic drugs [8]. This suggests that non-respiratory uroC dehydroxylation could serve as
509 an additional strategy that evolved in *E. asparagiformis*, *E. bolteae*, *E. citroniae*, and *L.*
510 *pacaense* to overcome catechol-mediated iron chelation.

511 While uroA is the most common terminal metabolite following ellagitannin consumption
512 in humans, its production varies widely [36]. Interestingly, the ability of a fecal sample to produce
513 uroA from uroC did not correlate with the presence of widespread uroC-metabolizing

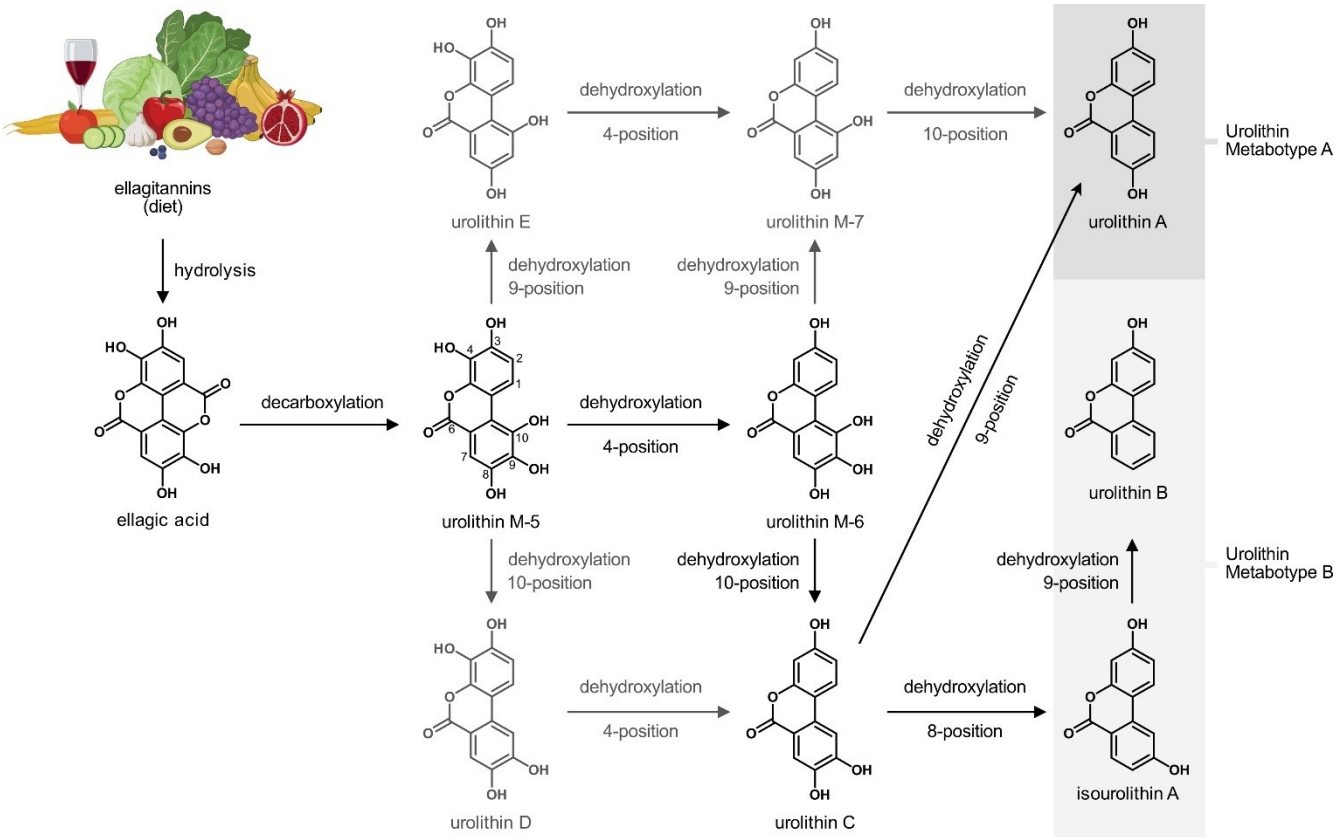
514 *Enterocloster* spp. or a *ucd* gene homologue, likely owing to poor viability or dead bacteria in
515 fecal samples. However, the active transcription of the *ucd* operon correlated perfectly with
516 metabotypes. These findings further emphasize the importance of functional assays such as
517 transcriptomics and *ex vivo* metabolism to understand the metabolism of xenobiotics by the gut
518 microbiota.

519 By identifying the genetic basis for metabolism of uroC, we found a novel metabolizing
520 species that could not have been predicted based on phylogeny alone. Our data suggests that
521 *ucd*-containing *Enterocloster* spp., and the closely related *L. pacaense*, are the main drivers of
522 urolithin A production in the gut microbiota based on their prevalence in metagenomes and
523 activity in fecal samples. However, we cannot conclude that they are solely responsible for this
524 activity. Rare, strain-specific urolithin A production has been reported for *Bifidobacterium*
525 *pseudocatenulatum* INIA P815 [36], *Streptococcus thermophilus* FUA329 [58], and *Enterococcus*
526 *faecium* FUA027 [58], which may be a result of horizontal gene transfer since it is not shared by
527 other members of the taxa. Thus, further enzyme discovery efforts are necessary to understand
528 urolithin production in these bacteria.

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

534

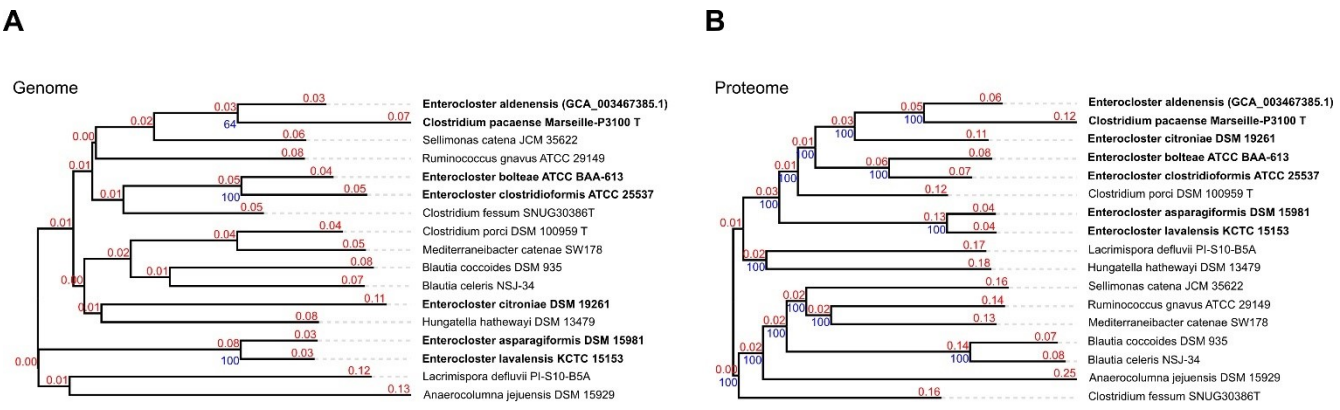
535 **Supplementary Figures**



536

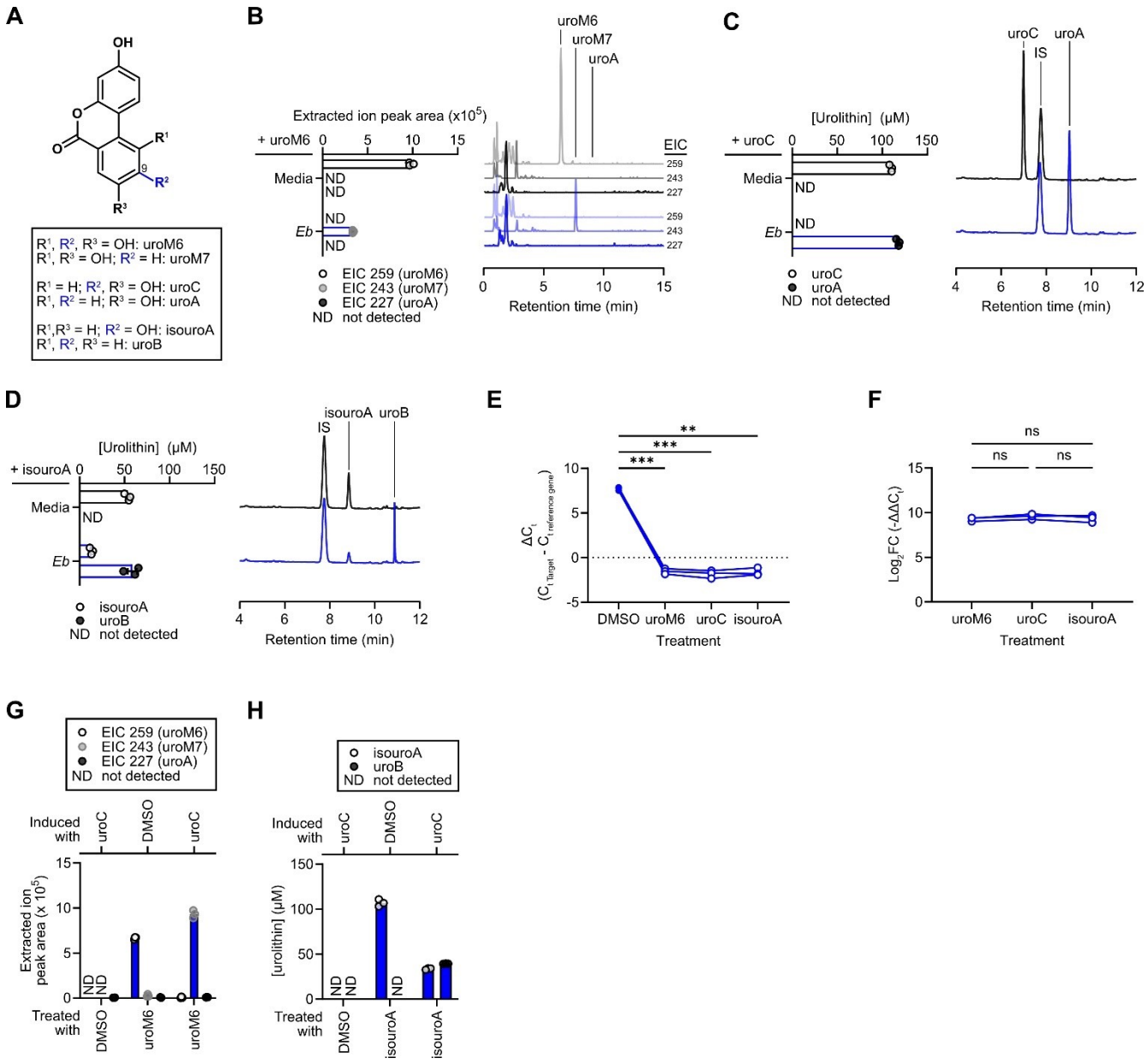
537 **Supplementary Figure 1. Ellagitannins are metabolized by gut bacteria.**

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



Supplementary Figure 2. *Enterocloster* spp. whole genome and proteome phylogenetic trees.

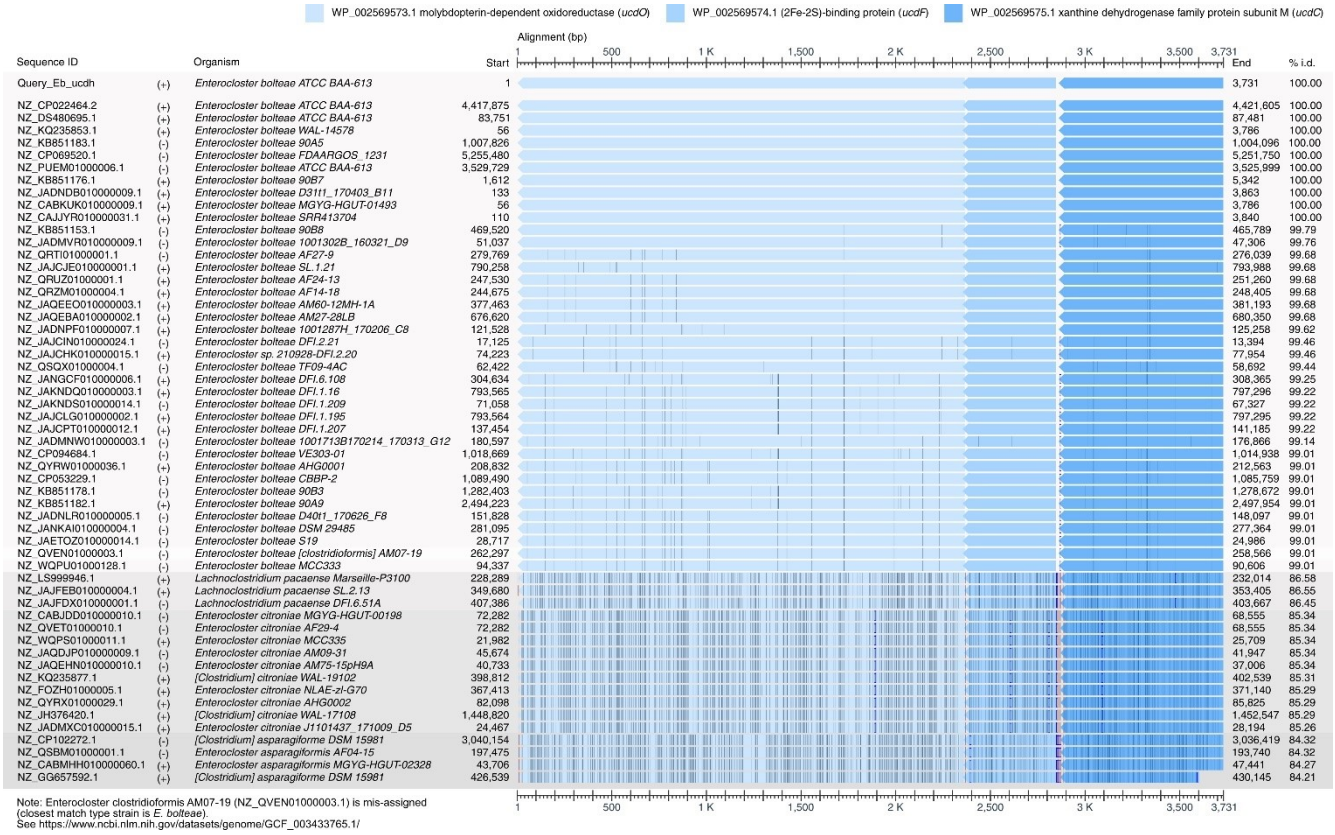
A) Whole genome phylogenetic tree of *Enterocloster* spp. The tree was inferred with FastME 2.1.6.1 from Genome BLAST Distance Phylogeny (GBDP) distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap 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. Source data are provided as a Source data file.



Supplementary Figure 3. *E. bolteae* dehydroxylates urolithins at the 9-position.

A) Chemical structures of urolithin M-6, urolithin C, and isourolithin A along with their dehydroxylated counterparts urolithin M7, urolithin A, and urolithin B. **B)** Quantification of extracted ion chromatogram (EIC) peak areas from *Eb* cultures sampled after 24 h of growth with 100 μM uroM6 (n = 3 biological replicates) with representative extracted ion chromatograms (EIC) to the right (from one representative biological replicate). The same scale was used for each chromatogram. **C)** Quantification of urolithin peak areas from *Eb* cultures sampled after 24 h of growth with 100 μM uroC (n = 3 biological replicates) with representative chromatograms (λ = 305 nm) to the right (from one biological replicate). The same scale was used for each chromatogram. **D)** Quantification of urolithin peak areas from *Eb* cultures sampled after 24 h of growth with 100 μM isouroA (n = 3 biological replicates) with representative chromatograms (λ = 305 nm) to the right (from one biological replicate). The

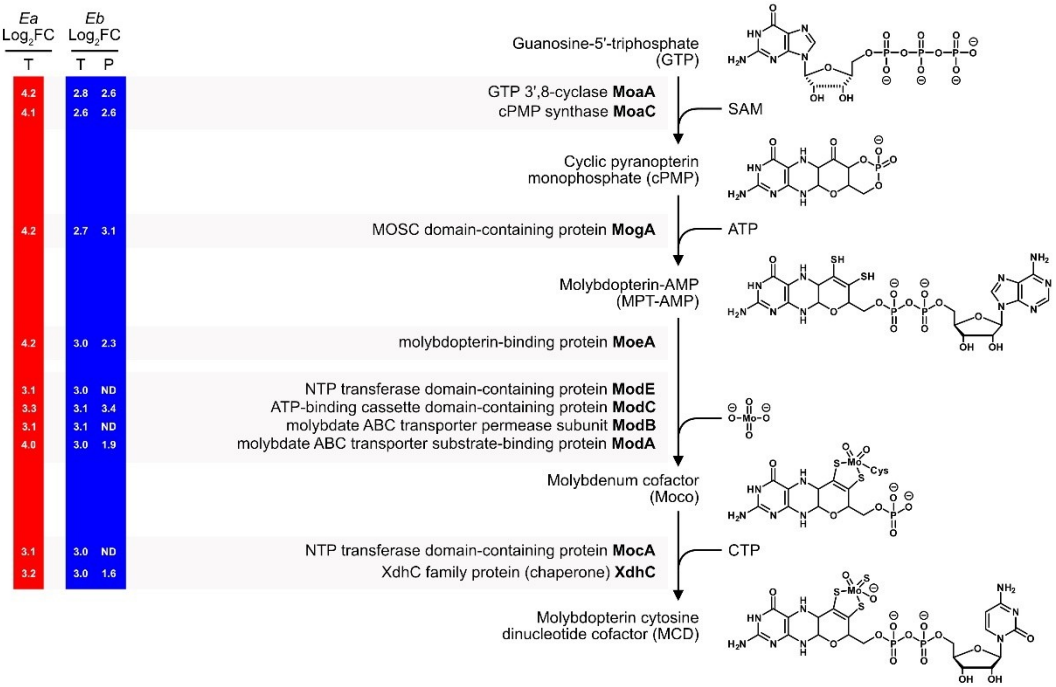
same scale was used for each chromatogram. **E,F)** RT-qPCR expression of the *Eb ucdO* gene. Growing *Eb* cultures were treated with DMSO, uroM6, uroC, or isouroA (100 μM) for 2 h before RNA isolation and reverse transcription (n = 3 biological replicates). **E)** Differential *Eb ucdO* gene expression comparing DMSO, uroM6, uroC, and isouroA is displayed as target-specific ΔCt ($C_{t\ MoO\ Gene}$ - $C_{t\ dnaK\ Reference\ Gene}$) values. Data are presented as individual ΔCt values with lines connecting paired biological replicates (from the same pre-spike culture); repeated-measures one-way ANOVA with Dunnett's multiple comparisons test; **, $p < 0.01$; ***, $p < 0.001$. **F)** Gene expression profile of the *Eb ucdO* gene in different urolithin treatment groups displayed as $\log_2$FC (equivalent to $-\Delta\Delta C_t$, where $\Delta\Delta C_t = \Delta C_{t\ urolithin} - \Delta C_{t\ DMSO}$). Data are presented as individual $\log_2$FC values with lines connecting paired biological replicates; repeated-measures one-way ANOVA with Tukey's multiple comparisons test. **G)** Quantification of extracted ion chromatogram (EIC) peak areas in DMSO- or uroC-treated *Eb* cell suspensions. 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). **H)** Quantification of urolithin concentrations in DMSO- or uroC-treated *Eb* cell suspensions (n = 3 biological replicates). For B-D and G-H, data are represented as mean ± SEM. ND, not detected; ns, not significant; FC, fold change; FDR, false discovery rate. Source data and statistical details are provided as a Source data file.



Supplementary Figure 4. BLASTn searches using the *E. bolteae ucd* operon genomic sequence identifies homologues in gut bacteria.

NBCI Multiple Sequence Aligner viewer hits for BLASTn searches using the *E. bolteae* DSM 15670 *ucd* operon 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. Source data are provided as a Source data file.

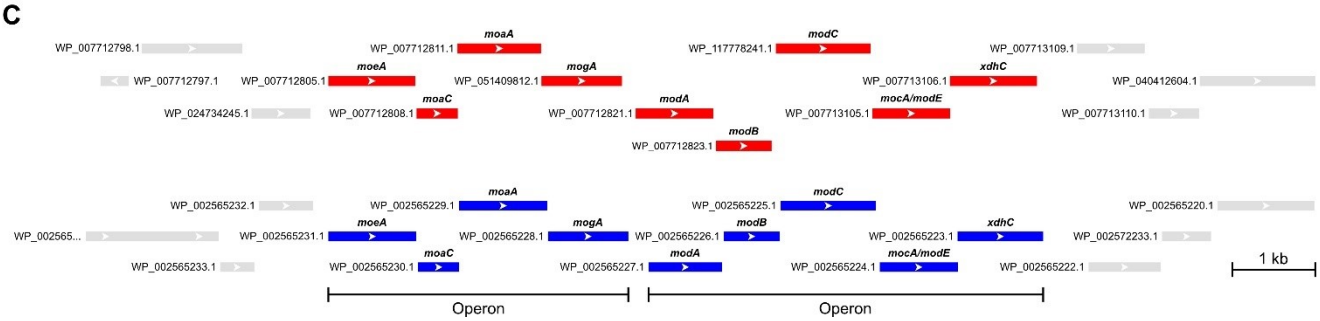
A



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



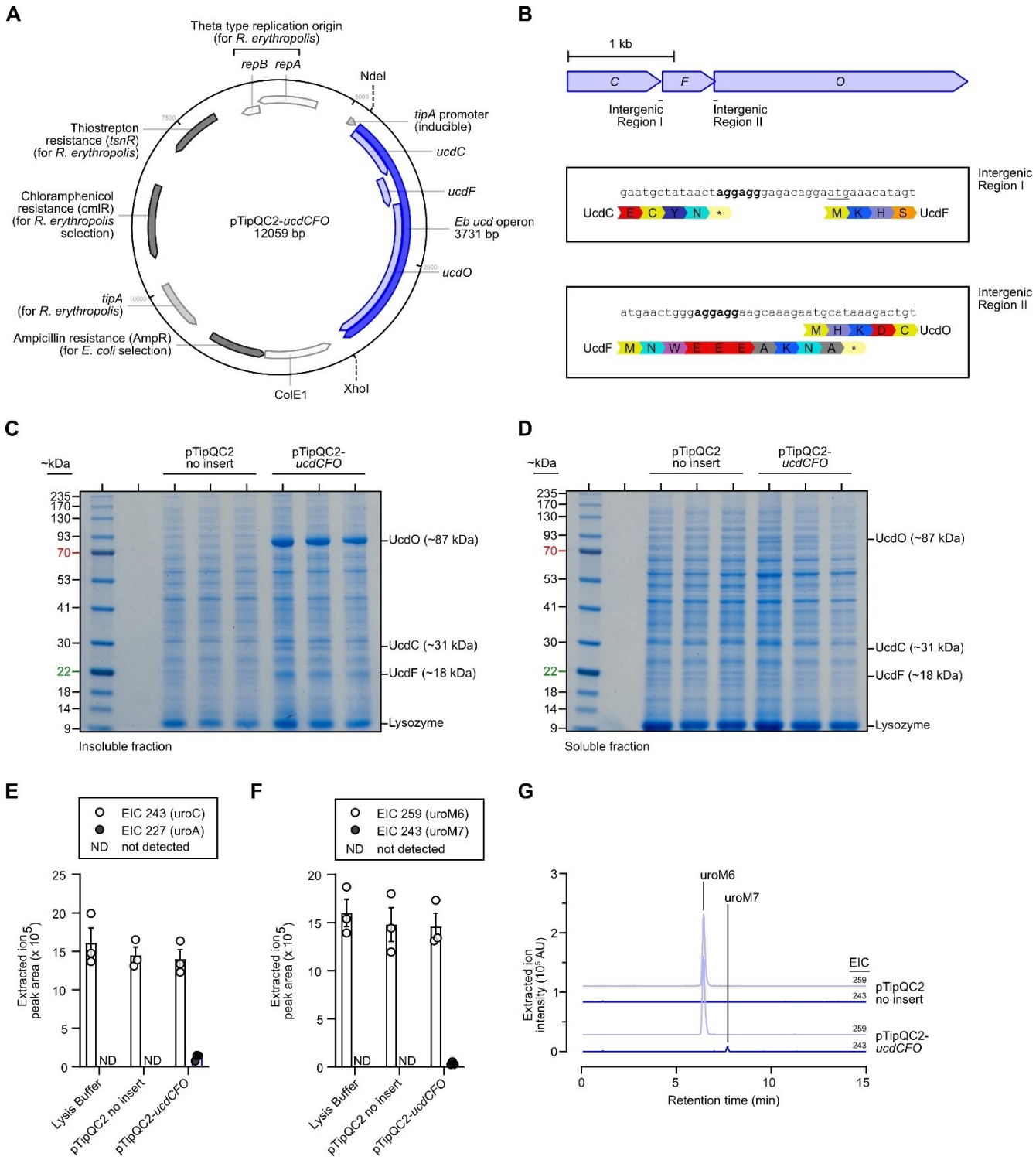
604

605 **Supplementary Figure 5. Urolithin C treatment upregulates molybdopterin cytosine**

606 **dinucleotide cofactor biosynthetic gene clusters.**

607 **A)** Molybdopterin cytosine dinucleotide (MCD) cofactor biosynthetic pathway. The log_2FC
608 values of MCD cofactor biosynthetic genes upregulated by uroC in both transcriptomics (T) and
609 proteomics (P) datasets (when available) are provided to the left of the figure for both *Ea* and
610 *Eb*. **B)** Table of molybdenum cofactor biosynthetic genes found in the genomes of *Ea* and *Eb*.
611 Annotations are based on NCBI (GTF files) and UniProt gene/protein names. Roles are
612 assigned based on required proteins for molybdenum cofactors in the xanthine
613 oxidase/dehydrogenase family of enzymes. Accessions, primary sequence length, and
614 annotated domains (with positions within the primary sequence) are also provided. **C)** Genomic
615 organization of the MCD cofactor biosynthetic genes for *Ea* and *Eb* (generated from the NCBI
616 Sequence Viewer). Source data are provided as a Source data file.

617



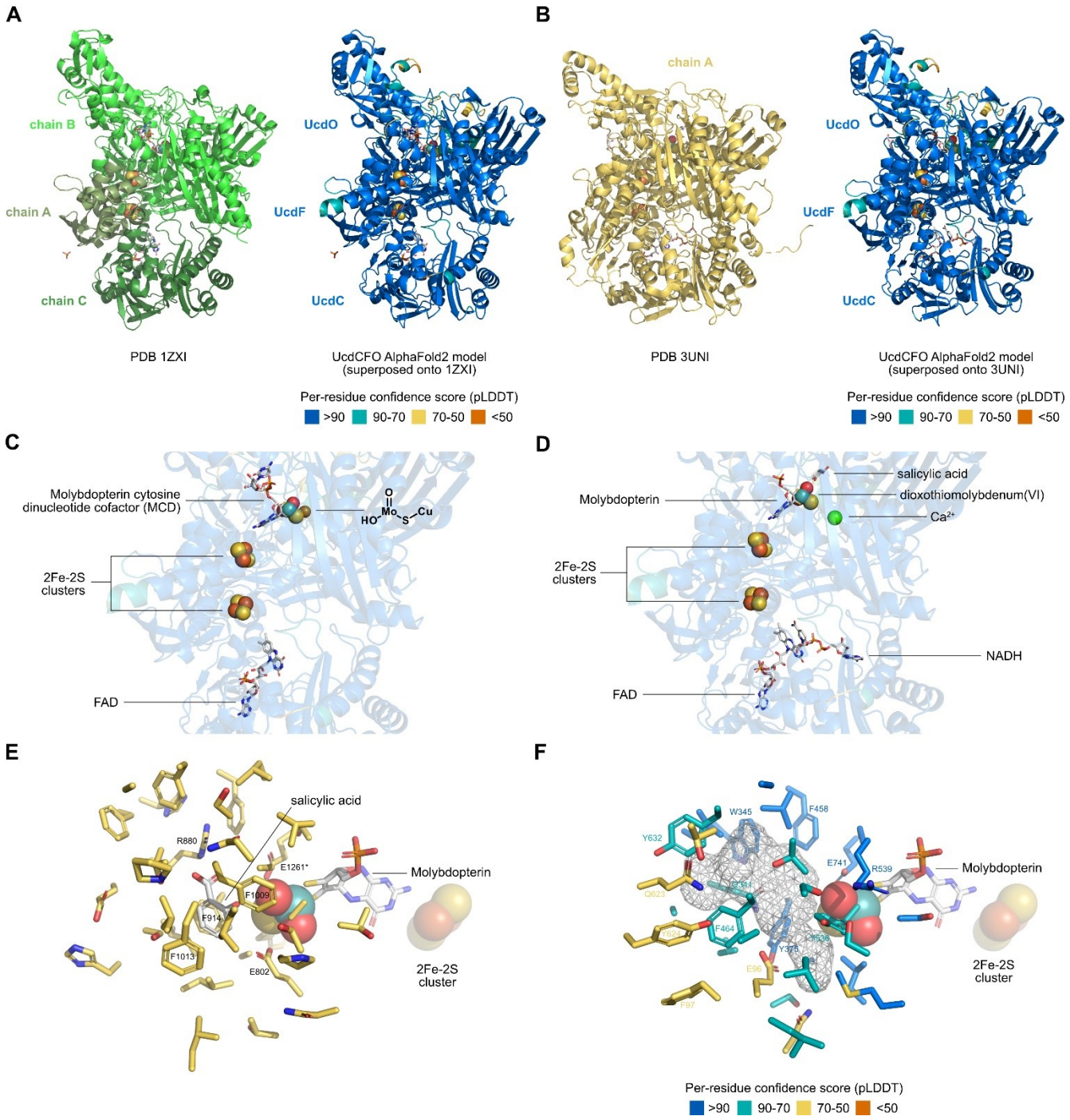
618

619 **Supplementary Figure 6. Heterologous expression of *E. bolteae ucdCFO* genes in *R.***

620 ***erythropolis*.**

621 **A)** Map of the *E. coli* – *R. erythropolis* pTipQC2-*ucdCFO* shuttle plasmid generated in
622 Benchling. **B)** Genomic organization of the wild-type *E. bolteae ucd* operon with intergenic

623 regions highlighted. Shine-Dalgarno consensus sequences (bold) are denoted between
624 translational stop and start (underlined) sites for *ucdC* and *ucdF* (Intergenic Region I) and for
625 *ucdF* and *ucdO* (Intergenic Region II). **C,D)** SDS-PAGE gels (10% bis-tris) stained with colloidal
626 Coomassie dye of the insoluble (C) and soluble (D) fractions from thiostrepton-induced (1
627 µg/mL) *R. erythropolis* harboring pTipQC2 no insert or pTipQC2-*ucdCFO* plasmids (n = 3
628 biological replicates). UcdCFO complex proteins are labeled on the right side of each gel image.
629 **E,F)** Quantification of extracted ion chromatogram (EIC) peak areas from crude lysates of
630 thiostrepton-induced *R. erythropolis* harboring pTipQC2 no insert or pTipQC2-*ucdCFO*
631 plasmids (n = 3 biological replicates). Crude lysates were incubated anaerobically with 2 mM
632 NADH and 357 µM uroC (E) or 357 µM uroM6 (F) for 72 h before extraction and analysis by
633 LC-MS. Data are represented as mean ± SEM. **G)** LC-MS extracted ion chromatograms (EIC)
634 of uroM6 ($[M-H]^-$ = 259) and uroM7 ($[M-H]^-$ = 243) from a representative anaerobic uroC
635 dehydroxylation assay using crude lysates of *R. erythropolis* harboring either pTipQC2 (no
636 insert) or pTipQC2-*ucdCFO* plasmids. Source data are provided as a Source data file.



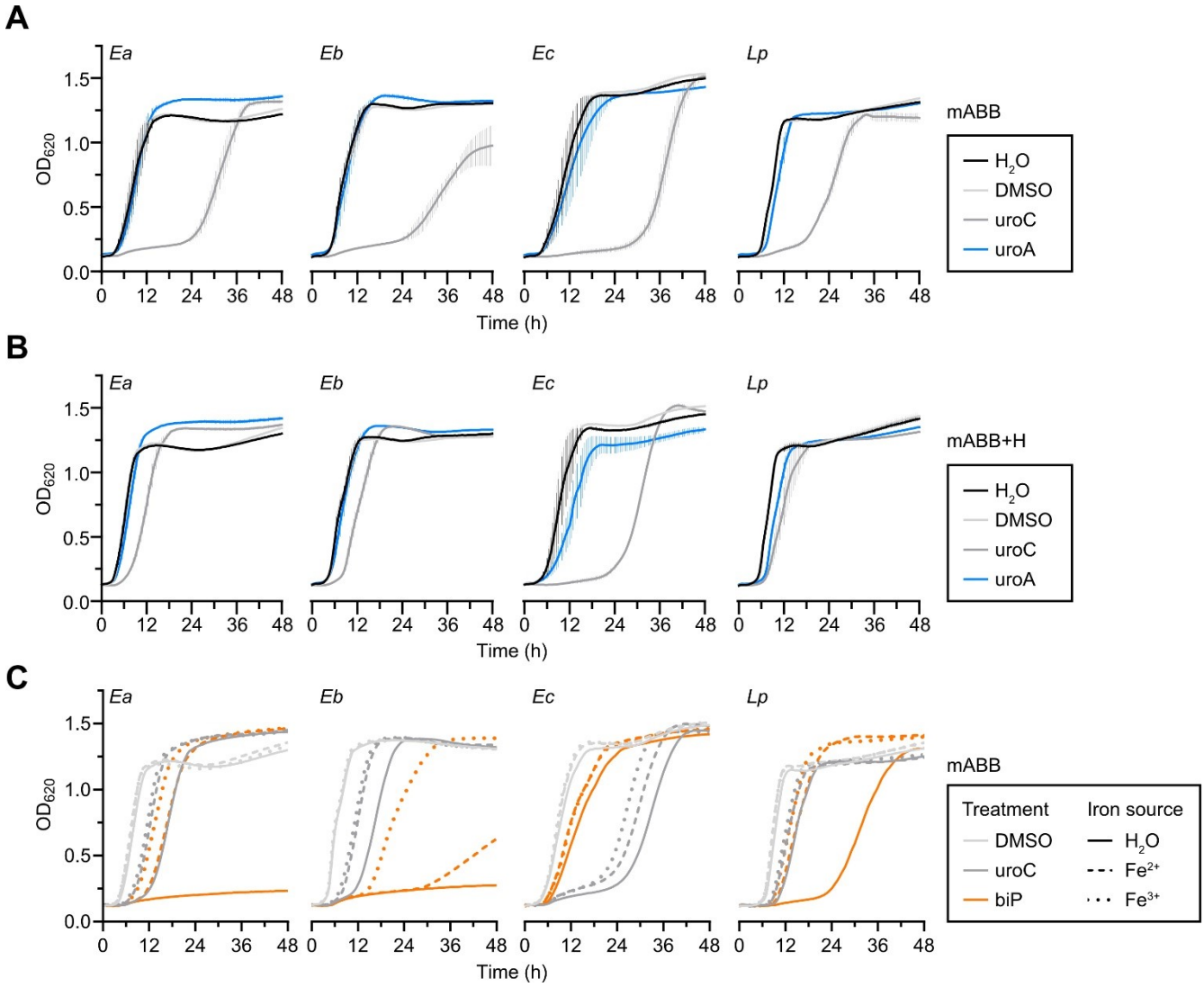
Supplementary Figure 7. The AlphaFold2 model of the *E. bolteae* UcdCFO enzyme complex has a similar quaternary structure to xanthine dehydrogenase superfamily crystal structures.

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

644 dehydrogenase with NADH bound). PDB 1ZXI is colored in shades of green (according to chain
645 ID), PDB 3UNI is colored in gold (single chain A shown), and the *Eb* Ucd enzyme complex is
646 colored according to its per-residue confidence score as indicated in the legend. The various
647 ligands (cofactors, coenzymes, ions, small molecules) of PDB 1ZXI and PDB 3UNI are included
648 in the respective *Eb* Ucd enzyme models. **C,D)** Cofactors from X-ray crystal structures **C)** PDB
649 1ZXI and **D)** PDB 3UNI modeled into the AlphaFold2 *Eb* Ucd enzyme complex, showing a
650 complete electron transport chain from a bound FAD molecule to a molybdopterin cofactor *via*
651 two 2Fe-2S clusters. **E)** Xanthine dehydrogenase active site from PDB 3UNI. Side chains within
652 8 Å of the salicylic acid ligand are colored in gold. Residues important for substrate (purine)
653 binding and catalysis are labeled with their one letter amino acid code and sequence position.
654 Position E1261* is catalytically important (acts as a general base) and is conserved in
655 XDH/xanthine oxidase enzymes [59]. **F)** *Eb* Ucd enzyme complex active site modeling. Side
656 chains within 8 Å of the salicylic acid ligand from PDB 3UNI are colored on the superposed *Eb*
657 Ucd enzyme complex according to their per-residue confidence score as indicated in the
658 legend. Residues surrounding the predicted active site are labeled with their one letter amino
659 acid code and sequence position (in the *Eb* MoO protein). The predicted urolithin binding site
660 is depicted by the surface (mesh) created by the active site residues. The surface was rendered
661 using the cavities and pockets only (culled) setting with a cavity detection cutoff of 5 solvent
662 radii in PyMOL. Source data are provided as a Source data file.

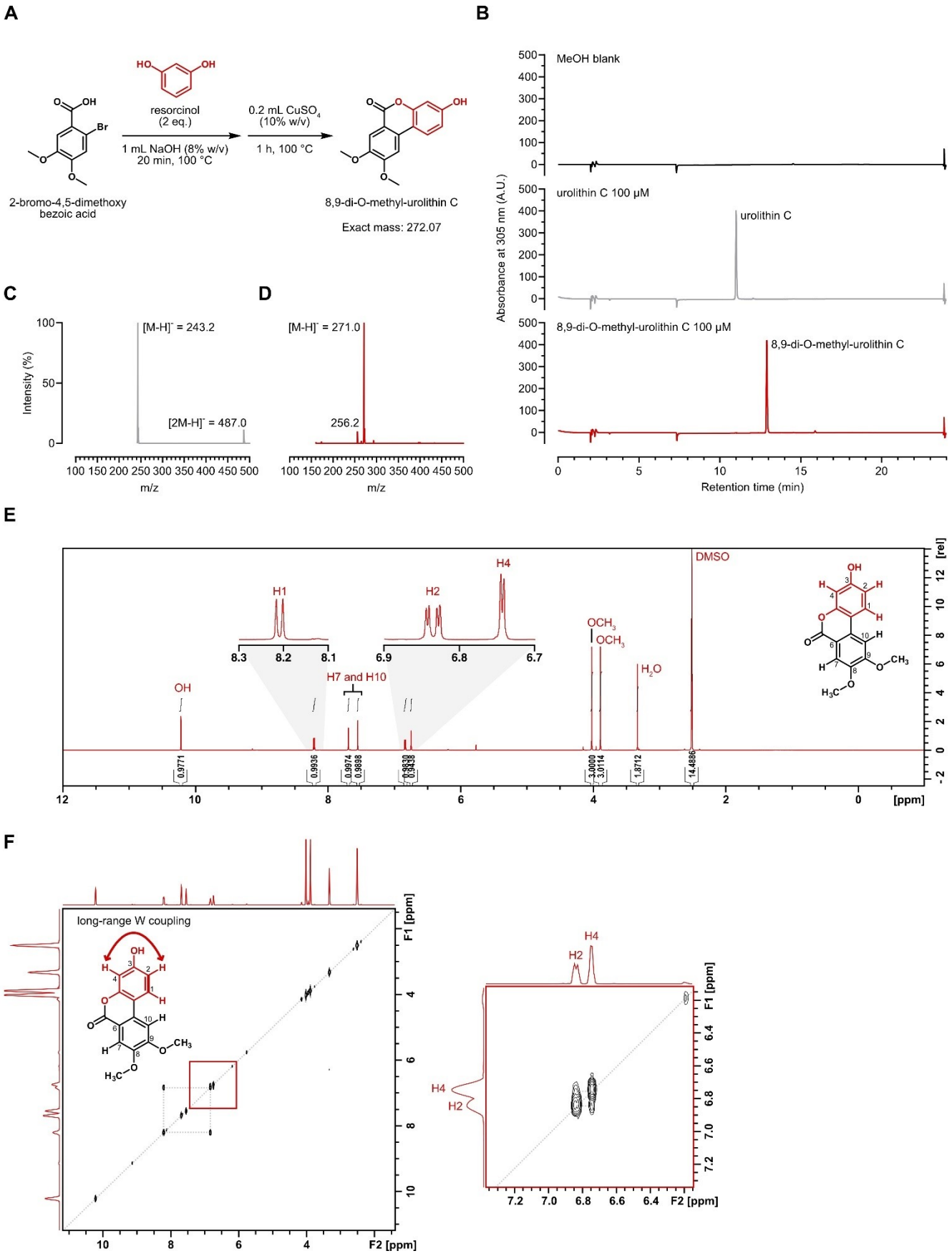
663

664



Supplementary Figure 8. Iron supplementation rescues lag time extension by uroC and 2,2'-bipyridyl.

A,B) Growth curves (optical density (OD) at 620 nm) of uroC-metabolizing *Enterocloster* spp. and *L. pacaense* treated with H_2O, DMSO (vehicle), or 100 μM of uroC or uroA in mABB medium (lacking added iron) (A) or mABB+H (B) medium (containing 7.7 μM hemin) (n = 3 biological replicates). Data are represented as mean ± SEM. **C)** Growth curves (optical density (OD) at 620 nm) of uroC-metabolizing *Enterocloster* spp. and *L. pacaense* treated with DMSO (vehicle), 100 μM of uroC, or 2,2'-bipyridyl (biP) in mABB media (lacking added iron) supplemented with solutions containing no iron (H_2O), 7.7 μM Fe^{2+} ($Fe(II)SO_4$), or 7.7 μM Fe^{3+} (Fe(III) pyrophosphate) (n = 3 biological replicates). Data are represented as means without error bars for clarity. Source data are provided as a Source data file.


A
resorcinol
(2 eq.)
1 mL NaOH (8% w/v)
20 min, 100 °C
0.2 mL CuSO4
(10% w/v)
1 h, 100 °C
2-bromo-4,5-dimethoxy
bezoic acid
8,9-di-O-methyl-urolithin C
Exact mass: 272.07
B
MeOH blank
urolithin C 100 µM
urolithin C
8,9-di-O-methyl-urolithin C 100 µM
8,9-di-O-methyl-urolithin C
Absorbance at 305 nm (A.U.)
Retention time (min)
C
[M-H]⁻ = 243.2
[2M-H]⁻ = 487.0
Intensity (%)
m/z
D
[M-H]⁻ = 271.0
256.2
m/z
E
DMSO
H1
H2
H4
H7 and H10
OH
OCH3
OCH3
H2O
[ppm]
F
long-range W coupling
F1 [ppm]
F2 [ppm]
H4
H2

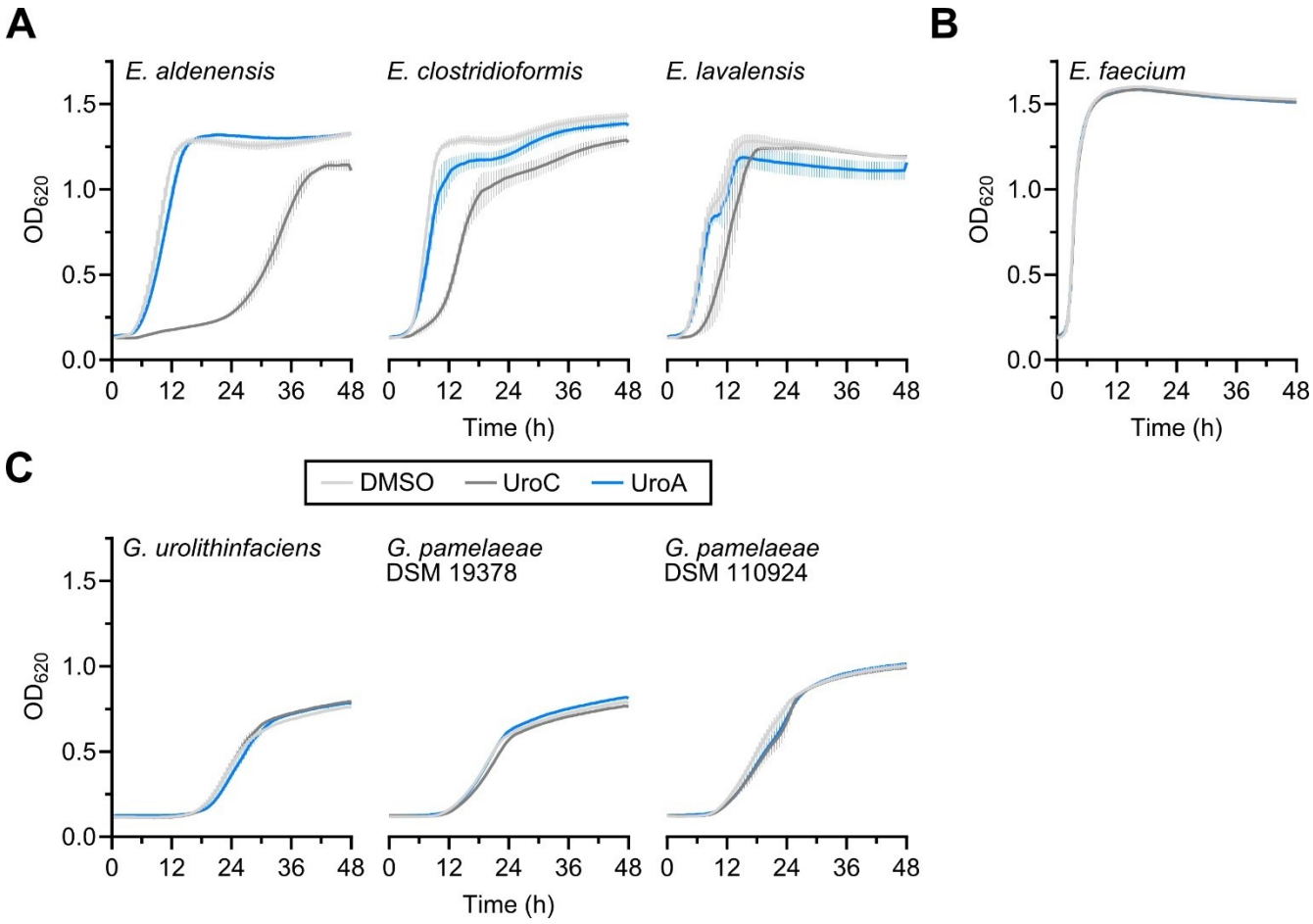
678

679 **Supplementary Figure 9. Synthesis and characterization of 8,9-di-O-methyl-urolithin C**

680 **A)** Reaction scheme for the synthesis of 8,9-di-O-methy-urolithin C. **B)** Aligned reversed-phase
681 (C18) HPLC chromatograms (λ = 305 nm) of 10 µL injections of the following solutions: MeOH
682 blank, urolithin C 100 µM, 8,9-di-O-methy-urolithin C 100 µM. **C,D)** Negative ESI-MS spectra
683 of urolithin C (C) and 8,9-di-O-methy-urolithin C (D) following chromatographic separation (B).
684 **E)** ^{1}H NMR spectrum (600 MHz, $(CD_3)_2SO$)) of 8,9-di-O-methy-urolithin C. **F)** COSY NMR
685 spectrum (600 MHz, $(CD_3)_2SO$)) of 8,9-di-O-methy-urolithin C. The grey diagonal line denotes
686 self correlation between protons. The right panel corresponds to the area in the red box.
687 Coupling between protons is shown on the 8,9-di-O-methy-urolithin C using bold lines and bold
688 arrows (for long range coupling).

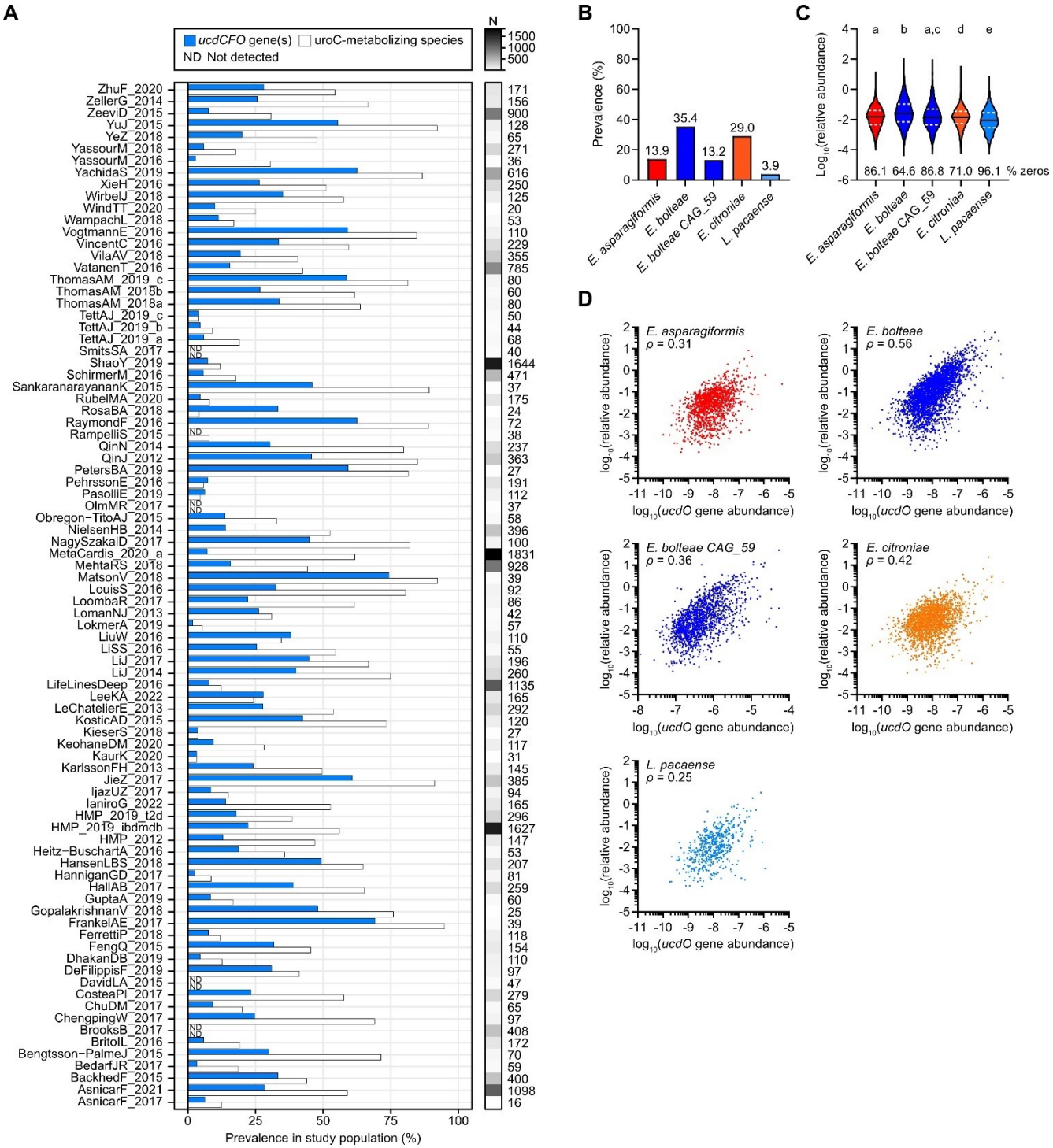
689

690



Supplementary Figure 10. Urolithin C differentially affects the growth of gut bacteria *in vitro*.

A,B,C) Growth curves (optical density (OD) at 620 nm) of non-uroC metabolizing *Enterocloster* spp. (A), *Enterococcus faecium* (B) and *Gordonibacter* spp. (C) treated with DMSO (vehicle), or 100 µM of uroC or uroA in mABB medium (n = 3 biological replicates for *Enterocloster* spp. and *E. faecium*, and n = 4 biological replicates for *Gordonibacter* spp.). Data are represented as mean ± SEM.

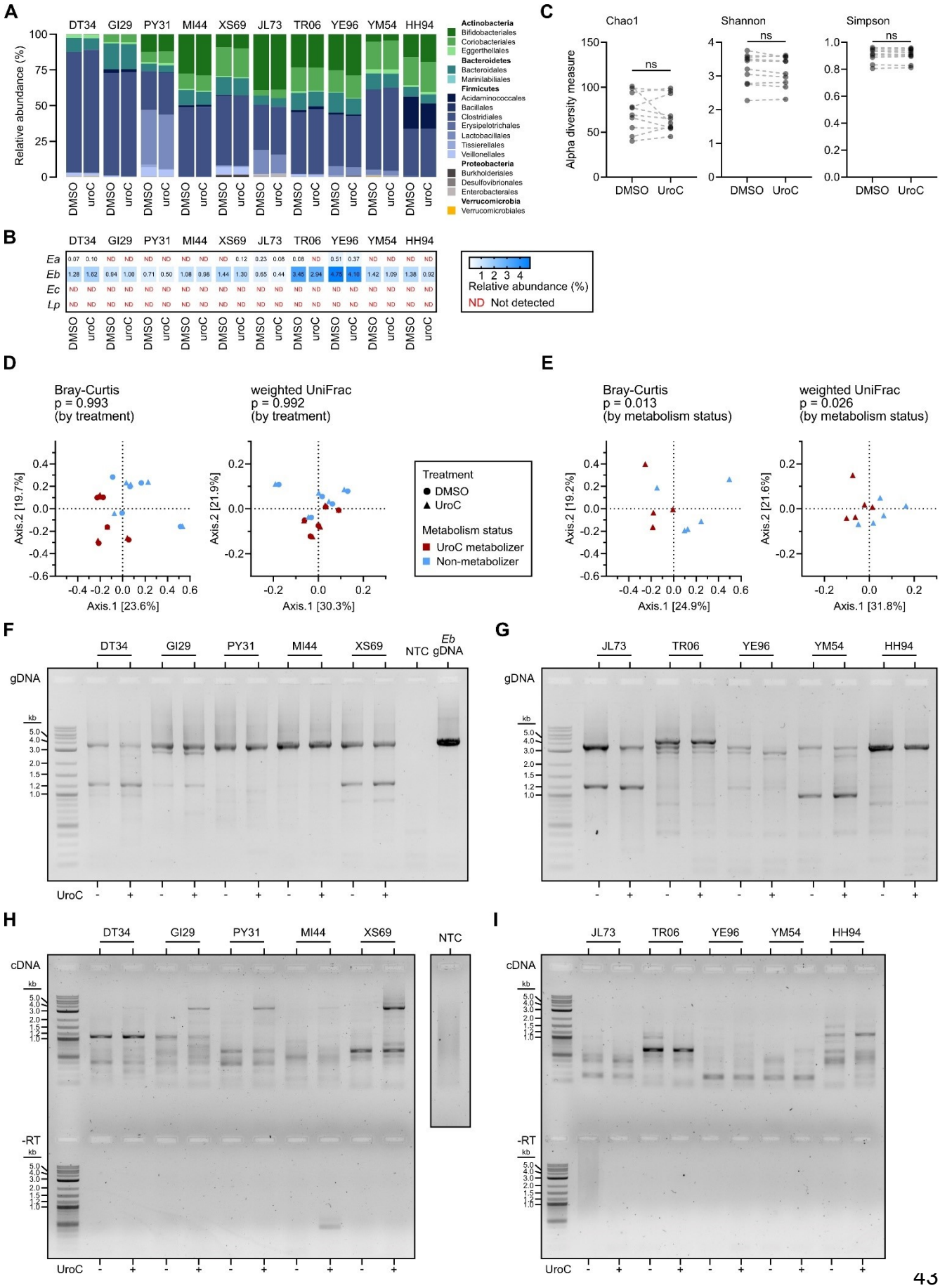


Supplementary Figure 11. Urolithin C-metabolizing species and genes are prevalent and correlate with *ucdO* gene abundance in human gut metagenomes.

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

706 prevalence plot as a heatmap. Details on the study populations (from the
707 curatedMetagenomicData R package) can be found in the Source Data file. **B)** Prevalence of
708 uroC-metabolizing species in human fecal metagenomes from the curatedMetagenomicData R
709 package. Data are reported for 86 studies (N=21,030 individuals) and are colored according to
710 the species. **C)** Violin plot of the $\log_{10}$(relative abundance) of uroC-metabolizing species in
711 human fecal metagenomes. The solid horizontal line corresponds to the median and the dashed
712 white lines correspond to the first and third quartiles. The percentage of zeroes are denoted
713 below the plotted distributions. Differences between groups were determined using the Kruskal-
714 Wallis test on untransformed relative abundance values. Significant differences between
715 groups are denoted by a different lowercase letter above each plot. **D)** Correlation between the
716 *ucdO* gene abundance in reads per kilobase per million mapped reads (RPKM) and the relative
717 abundance of each uroC-metabolizing species in fecal metagenomes. Both values are
718 illustrated on a $\log_{10}$ scale. Spearman rho (ρ) values are denoted above the scatter plots. All
719 correlations were significant, $P < 0.0001$. Source data are provided as a Source data file.

720


A
DT34 GI29 PY31 MI44 XS69 JL73 TR06 YE96 YM54 HH94
Relative abundance (%)
DMSO uroC
Actinobacteria
Bifidobacteriales
Coriobacteriales
Eggerthellales
Bacteroidetes
Bacteroidales
Marinilabiliales
Firmicutes
Acidaminococcales
Bacillales
Clostridiales
Erysipelotrichales
Lactobacillales
Tissierellales
Veillonellales
Proteobacteria
Burkholderiales
Desulfovibrionales
Enterobacterales
Verrucomicrobia
Verrucomicrobiales
B
Ea
Eb
Ec
Lp
Relative abundance (%)
ND Not detected
C
Chao1
Shannon
Simpson
Alpha diversity measure
ns
DMSO UroC
D
Bray-Curtis
p = 0.993
(by treatment)
weighted UniFrac
p = 0.992
(by treatment)
Axis.1 [23.6%]
Axis.2 [19.7%]
Axis.1 [30.3%]
Axis.2 [21.9%]
Treatment
DMSO
UroC
Metabolism status
UroC metabolizer
Non-metabolizer
E
Bray-Curtis
p = 0.013
(by metabolism status)
weighted UniFrac
p = 0.026
(by metabolism status)
Axis.1 [24.9%]
Axis.2 [19.2%]
Axis.1 [31.8%]
Axis.2 [21.6%]
F
DT34 GI29 PY31 MI44 XS69 NTC Eb gDNA
gDNA
kb
UroC
G
JL73 TR06 YE96 YM54 HH94
H
DT34 GI29 PY31 MI44 XS69 NTC
cDNA
-RT
I
JL73 TR06 YE96 YM54 HH94

721

722 **Supplementary Figure 12. Urolithin C-metabolizing species and the *ucd* operon are**
723 **prevalent in fecal slurries, but only *ucd* transcription correlates with urolithin C**
724 **metabolism.**

725 **A)** Stacked bar plot of bacterial percent relative abundance (based on V1-V9 16S rRNA gene
726 sequencing) in DMSO- or uroC-treated fecal slurries from 10 healthy donors (from one
727 experimental replicate). Bars are colored according to the phylum (bold) and order. **B)** Heatmap
728 of the percent relative abundance of uroC-metabolizing *Enterocloster* spp. in A). **C)** Alpha
729 diversity plots between DMSO- and uroC-treated fecal slurries according to Chao1, Shannon,
730 and Simpson diversity metrics. Lines between data points connect paired biological replicates;
731 Wilcoxon test; ns, not significant. **D,E)** Principal coordinate analyses of dissimilarities between
732 16S rRNA compositions based on the Bray-Curtis and weighted UniFrac distance methods.
733 Data points are colored according to the treatment used and the metabolism status of the fecal
734 slurry; PERMANOVA test according to the treatment (D) for all fecal slurries or uroC metabolism
735 status (E) for uroC-treated slurries. **F,G)** PCR with *ucd*-specific primers on gDNA extracted from
736 an *E. bolteae* isolate and fecal microbiota communities from 10 healthy donors (from one
737 experimental replicate). 1% agarose gel of amplicons derived from uroC-metabolizing (F) and
738 non-metabolizing (G) fecal slurries. Samples are matched to the urolithin metabolism data in
739 Fig. 6C and 16S rRNA sequencing in A) and B). The no template control (NTC) and *E. bolteae*
740 positive control are the same for both gels. **H,I)** 1% agarose gel of amplicons from Fig. 6D,E
741 including the no reverse transcriptase (-RT) in uroC-metabolizing (H) and non-metabolizing (I)
742 fecal slurries. The NTC is derived from the final lane of the gel in H). Source data are provided
743 as a Source data file.

744

745 **Supplementary Sequence 1. cDNA sequence of *E. bolteae ucdCFO* transcript, 3605 bp**

746 **(band in Fig. 2I)**

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

748 AAGCACAGCCAGGGTTATGGCAGGCGCAACGGATCTGATTCCGCCAATGAAAGACAAGGTTATATCACCAGAGTATATCATTG
749 ACCTTAAGAAAATTCCAGGTTTGGATTATCTGGAATACGATGACAGGGAAGGTTTAAAAATAGGGGCGCTGACAACACTGCGT
750 ACCATAGAGACATCTCCTCTGGTTAAAGAAAAGAATCCGGCAGTGGCTCATGCCGCCAAGGTAGTGGCATCCACACAAATCAG
751 GACAAAGGGCACCATGGCAGGCAATATCTGCAATGCGTCCCCATCCTGCGATACAGCCCCCAACCTACTGGCTCAGGGCGCTA
752 AGATATTGGTACAGGGTCCCAACAAGGACCGGGTCATTCAAATCGAAGACTTTTTCCTGGGTGTCAAAAAGACTTCCCTGGAG
753 CCGGGAGAAATCGTGACAGGCATTGTGATTCCGCCGCTGGCTGAGAATGAGAGGGCTGCATACATAAAACACGCGGTCAGAAA
754 GGCGATGGACCTGGCCATCATCGGCGTCGCAGTCAAGATTAAGGTGGAAGACGGCGTCTGCACAGACGCCAGGATTGCCCTGG
755 GCGCGGTTGCGGCCACGCCGGTACGGGCACCCGGGGCGGAGGAAGCCCTGATTGGCAAAGAACTTACCGATGAAGTGATTGTG
756 AAGGCATCGGAGGAAGCAATGAATTCCTGTCATCCAATCTCTGATATCCGCGCATCGGCCGAATACAGGAAGGATATGATCCG
757 CGTGTTTACAAAACGTGCCATTAAGCAGGCCATGGAATGCTATAACTAGGAGGGAGACAGGAATGAAACATAGTATAAGTTTA
758 ATCATCAACGGAGATCCGGTTGATGCTATTGTGAAGGATAACCTGACGCTTCTGGATTTTCTTCGTGACCAGTTGTTTCTGAC
759 AGGCACGAAAAAGGGGTGTGAGGAAGGGGAGTGTGGCGCCTGTACCGTGATGCTTGACGGTAAGCCGGTTAACTCCTGCTGTA
760 CCCTGGCCGTGGAGTGCGACGGCCATGAGATCATAACCGTGGAGGGAATTGCCAGGGAGGGGATGCTTCATCCCATTCAGAAG
761 CAGTTTATAGAGAAGTGGGCCATGCAGTGCGGCTACTGTACTCCGGGTATGATTATGTCCGCCAAAGCTCTTCTGGATGTGAA
762 TAAACATCCTACGGAACTGGAAATCCGGGAAGCCATTGAGGGCAATCTGTGCCGCTGCACCGGTTATGCAAAGATTGTGGAGG
763 CAATCCAGGCAGCCGCTGCGCAGATGAACTGGGAGGAGGAAGCAAAGAATGCATAAAGACTGTGACAAACATTATTTTAAAAA
764 ACCGGAATTTTACCGTCTGACCGGCGAGAACAATTATGTCAGGATTGACGCTGAAGACAAGGTAACAGGACACGGCCAGTATG
765 TGGGTGACATCATGTTTCCGGATATGCTTACCGGAAAAATGGTCAGAAGCCCTTATGCATCTGCAAAAATCCTGTCCATTGAC
766 ACCAGCGAGGCAGAGAAGCTTCCAGGTGTAAAGTGCATCCTCACTGCCAGGGACTTTGAGTGGAAGTCCCTGGTGGGAAATGG
767 AGAATTTGCAGCTGAGTTCGCGGACAAGGAAGTATTGTGTTCTGAGAAGGTGCGCCAGGTTGGTGACGATGTGGCAGCCGTGG
768 CAGCCGTGGATGAGGAAACCGCCCAGCGGGCCGCGGATCTGATTAAGGTAGAGTACCAGGTGCTTCCGGGGGTGTTTGACCCC
769 TTTGAGGCAATGGAGGAGAATGCCCCCGAGGTAAACTGGGAGGGCAAAGGCATACACAATATCGGCATGCAGTCCGTGATGAA
770 GGCGGGCACGGATATTGATGAGGAATTTGACCGCGCATCCTATGTGCAGCACAGGGATTACAAGACCCACCGTATGGTACATG
771 CGGCCATGGAGCCCCACGGTGCGGTGGCTACCTACAGGAATGGGACCTACACCATCTGGATGTCCACCCAGATGTCCTTTGTG
772 GACCAGTTCTGGTATGCCCGCTGCCTTGGCGTTGGAGAGAACCAGGTGCGGGTAATCAAGCCTCTTGTGGGCGGCGGTTTCGG
773 CGGCAAGCTGGATTCCTATTCCTTTGGCCTTTGCGCTGCCAAGATGGCGGAGATGACCGGACGCCCGGTACGCATGATTCTGT
774 CCCGCGAAGAAGTATTCCAGACCACGCGCAACCGCCACCCCATTTACATGCATATTGACACTGCCTTTGGGACAGACGGTAAG
775 CTGCTGGCAAAGAAATGCTACCATGTGCTGGACGGCGGGCCTTACGGCGGGTCCGGTGTTGCGGCCTGCGCCCAGTCCACATT
776 GTGGGCCAACTTCCCTTACAAGATGAATTCCGTGGATTTCCTGGCAAGGCGTGTGTATACCAATAACCCTTCTGCAGGAGCCA
777 TGAGAGGATATACGGCATGCCAGGTGCATTTTGCCCATGACCTGAACATGCAGTTCGCGGCTGACCAGATGGGCATTGACCCG
778 GTTGAGTTCCGCAAGATAAGCGCGGCAGACCCCGGATACGTGGCTCCGGCCGGACTGGCAATCACCAGCTGCGCATACAAGGA
779 GACACTGGATACGGCTGCAAAGGAAATTGGCTGGTATGAGAAGAAGGATAAGCTGAAAAAGGGAGAGGGCATTGGTTTTGCGG
780 GCACCGGCTTTGTATCGGGCACCGGATTCGCGGTGCTGGAAGCGCCCAACCAGAGCTCTGCGTGCGTGACCCTGCGCATGAAC
781 AAGCGGGGCATGGCCACCCTGTACATCGGCTCCCACGACATAGGACAGGGTTCCGATACGGTTATGACGGCCATTGTGGCCGA
782 GGAACTGGGGCTTCCCATGGATATGGTCAAGACCTTTATGTCCGACACCTTCCTGACGCCCTGGGACTCCGGCTCTTACGGCA
783 GCCGTGTTACCTTCCTGGCCGGCAATGCGGCCCGCCGCGCTGCCGTGGATGCAAAACGCCAGCTGTTTGAGGTCATTGCGCCT
784 ATGTGGGGGGTGATGCCGGAGACATTAGAGTGTCTGGACGGAAAGGTAATCAGCAAGGAGAAGGCAGAGTATCAGATGTCCAT
785 TGGAGACGCCATGTTCAAGTATATGACGGTCAAGGGCGGCGATGAACTGATTGGCGTGGGATCCTATTACCACCGTACCGACA
786 ATTCACAGTATAACGGCAACAATACCACCAACTACGCGCCTGCATACAGCTTCTCCACCGGAGCAGCCCATCTGACTGTGGAC
787 GAGGAGACCGGCGTCCTGGATATTGATGAATTTGTATTTGCCCATGACTGCGGCCGCGCACTGAATAAGAGGGCCGTGGAAGG
788 CCAGCTGGAGGGATCCATCGGCATGGGCCTGGGCTATGCTGTCTATGAGCACAATGTAACCAGGGAAGGAAAGATTCTCAATC
789 CCAACTTCCGCGATTACCGGCTGCCCACTGCCCTTGATATGCCGAAGATGCGCACCTTTTACGACTTTACACCGGACGAGGAA
790 GGACCTCTGGGCGCAAAGGAAGCTGGCGAGGGCTCCGCAGCACCTGTGGCGCCTGCCATTGCCAATGCGGTCAACATGGCAAC
791 CGGCGTGTACTTCACAGAGCTGCCTCTGGACCCGGA

792

Supplementary Sequence 2. pTipQC2-*ucdCFO*

>pTipQC2-ucdCFO coverage: 717x

AAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCA
GAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATAC
CTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTT
ACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGA
GATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTC
GGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACT
TGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGA
GCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACC
GCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACTAGAGTCCCGCTGAGGCGGCGTAGCAGGTCAGCCGCC
CCAGCGGTGGTCACCAACCGGGGTGGAACGGCGCCGGTATCGGGTGTGTCCGTGGCGCTCATTCCAACCTCCGTGTGTTTGTG
CAGGTTTCGCGTGTTGCAGTCCCTCGCACCGGCACCCGCAGCGAGGGGCTCACGGGTGCCGGTGGGTCGACTAGTTCAGTGAT
GGTGATGGTGATGctcgagCTGCCT**TTATTTTGAATTCCTGTCATCCTTCATGCCGTGCAGCGCTCTCCAGATATGCTCCGGG**
TCCAGAGGCAGCTCTGTGAAGTACACGCCGGTTGCCATGTTGACCGCATTGGCAATGGCAGGCGCCACAGGTGCTGCGGAGCC
CTCGCCAGCTTCCTTTGCGCCCAGAGGTCCTTCCTCGTCCGGTGTAAAGTCGTAAAAGGTGCGCATCTTCGGCATATCAAGGG
CAGTGGGCAGCCGGTAATCGCGGAAGTTGGGATTGAGAATCTTTCCTTCCCTGGTTACATTGTGCTCATAGACAGCATAGCCC
AGGCCCATGCCGATGGATCCCTCCAGCTGGCCTTCCACGGCCCTCTTATTCAGTGCGCGGCCGCAGTCATGGGCAAATACAAA
TTCATCAATATCCAGGACGCCGGTCTCCTCGTCCACAGTCAGATGGGCTGCTCCGGTGGAGAAGCTGTATGCAGGCGCGTAGT
TGGTGGTATTGTTGCCGTTATACTGTGAATTGTCGGTACGGTGGTAATAGGATCCCACGCCAATCAGTTCATCGCCGCCCTTG
ACCGTCATATACTTGAACATGGCGTCTCCAATGGACATCTGATACTCTGCCTTCTCCTTGCTGATTACCTTTCCGTCCAGACA
CTCTAATGTCTCCGGCATCACCCCCCACATAGGCGCAATGACCTCAAACAGCTGGCGTTTTGCATCCACGGCAGCGCGGCGGG
CCGCATTGCCGGCCAGGAAGGTAACACGGCTGCCGTAAGAGCCGGAGTCCCAGGGCGTCAGGAAGGTGTCGGACATAAAGGTC
TTGACCATATCCATGGGAAGCCCCAGTTCCTCGGCCACAATGGCCGTCATAACCGTATCGGAACCCTGTCCTATGTCGTGGGA
GCCGATGTACAGGGTGGCCATGCCCCGCTTGTTCATGCGCAGGGTCACGCACGCAGAGCTCTGGTTGGGCGCTTCCAGCACCG
CGAATCCGGTGCCCGATACAAAGCCGGTGCCCGCAAAACCAATGCCCTCTCCCTTTTTCAGCTTATCCTTCTTCTCATACCAG
CCAATTTCCTTTGCAGCCGTATCCAGTGTCTCCTTGTATGCGCAGCTGGTGATTGCCAGTCCGGCCGGAGCCACGTATCCGGG
GTCTGCCGCGCTTATCTTGCGGAACTCAACCGGGTCAATGCCCATCTGGTCAGCCGCGAACTGCATGTTCAGGTCATGGGCAA
AATGCACCTGGCATGCCGTATATCCTCTCATGGCTCCTGCAGAAGGGTTATTGGTATACACACGCCTTGCCAGGAAATCCACG
GAATTCATCTTGTAAGGGAAGTTGGCCCACAATGTGGACTGGGCGCAGGCCGCAACACCGGACCCGCCGTAAGGCCCGCCGTC
CAGCACATGGTAGCATTTCTTTGCCAGCAGCTTACCGTCTGTCCCAAAGGCAGTGTCAATATGCATGTAAATGGGGTGGCGGT
TGCGCGTGGTCTGGAATACTTCTTCGCGGGACAGAATCATGCGTACCGGGCGTCCGGTCATCTCCGCCATCTTGGCAGCGCAA
AGGCCAAAGGAATAGGAATCCAGCTTGCCGCCGAAACCGCCGCCCACAAGAGGCTTGATTACCCGCACCTGGTTCTCTCCAAC
GCCAAGGCAGCGGGCATACCAGAACTGGTCCACAAAGGACATCTGGGTGGACATCCAGATGGTGTAGGTCCCATTCCTGTAGG
TAGCCACCGCACCGTGGGGCTCCATGGCCGCATGTACCATACGGTGGGTCTTGTAATCCCTGTGCTGCACATAGGATGCGCGG
TCAAATTCCTCATCAATATCCGTGCCCGCCTTCATCACGGACTGCATGCCGATATTGTGTATGCCTTTGCCCTCCCAGTTTAC
CTCGGGGGCATTCTCCTCCATTGCCTCAAAGGGGTCAAACACCCCCGGAAGCACCTGGTACTCTACCTTAATCAGATCCGCGG
CCCGCTGGGCGGTTTCCTCATCCACGGCTGCCACGGCTGCCACATCGTCACCAACCTGGCGCACCTTCTCAGAACACAATACT
TCCTTGTCCGCGAACTCAGCTGCAAATTCTCCATTTCCCACCAGGGACTTCCACTCAAAGTCCCTGGCAGTGAGGATGCACTT
TACACCTGGAAGCTTCTCTGCCTCGCTGGTGTCAATGGACAGGATTTTTGCAGATGCATAAGGGCTTCTGACCATTTTTCCGG
TAAGCATATCCGGAAACATGATGTCACCCACATACTGGCCGTGTCCTGTTACCTTGTCTTCAGCGTCAATCCTGACATAATTG
TTCTCGCCGGTCAGACGGTAAAATTCCGGTTTTTTAAAATAATGTTTGTCACAGTCTTTATGCATTCTTTGCTTCCTCCTCCC
AGTTCATCTGCGCAGCGGCTGCCTGGATTGCCTCCACAATCTTTGCATAACCGGTGCAGCGGCACAGATTGCCCTCAATGGCT
TCCCGGATTTCCAGTTCCGTAGGATGTTTATTCACATCCAGAAGAGCTTTGGCGGACATAATCATACCCGGAGTACAGTAGCC
GCACTGCATGGCCCACTTCTCTATAAACTGCTTCTGAATGGGATGAAGCATCCCCTCCCTGGCAATTCCCTCCACGGTTATGA
TCTCATGGCCGTCGCACTCCACGGCCAGGGTACAGCAGGAGTTAACCGGCTTACCGTCAAGCATCACGGTACAGGCGCCACAC
TCCCCTTCCTCACACCCCTTTTTCGTGCCTGTCAGAAACAACTGGTCACGAAGAAAATCCAGAAGCGTCAGGTTATCCTTCAC
AATAGCATCAACCGGATCTCCGTTGATGATTAAACTTATACTATGTTTCATTCCTGTCTCCCTCCTAGTTATAGCATTCCATG
GCCTGCTTAATGGCACGTTTTGTAAACACGCGGATCATATCCTTCCTGTATTCGGCCGATGCGCGGATATCAGAGATTGGATG
ACAGGAATTCATTGCTTCCTCCGATGCCTTCACAATCACTTCATCGGTAAGTTCTTTGCCAATCAGGGCTTCCTCCGCCCCGG
GTGCCCGTACCGGCGTGGCCGCAACCGCGCCCAGGGCAATCCTGGCGTCTGTGCAGACGCCGTCTTCCACCTTAATCTTGACT
GCGACGCCGATGATGGCCAGGTCCATCGCCTTTCTGACCGCGTGTTTTATGTATGCAGCCCTCTCATTCTCAGCCAGCGGCGG
AATCACAATGCCTGTCACGATTTCTCCCGGCTCCAGGGAAGTCTTTTTGACACCCAGGAAAAAGTCTTCGATTTGAATGACCC
GGTCCTTGTTGGGACCCTGTACCAATATCTTAGCGCCCTGAGCCAGTAGGTTGGGGGCTGTATCGCAGGATGGGGACGCATTG
CAGATATTGCCTGCCATGGTGCCCTTTGTCCTGATTTGTGTGGATGCCACTACCTTGGCGGCATGAGCCACTGCCGGATTCTT

849 **TTCTTTAACCAGAGGAGATGTCTCTATGGTACGCAGTGTTGTCAGCGCCCCTATTTTTAAACCTTCCCTGTCATCGTATTCCA**
850 **GATAATCCAAACCTGGAATTTTCTTAAGGTCAATGATATACTCTGGTGATATAACCTTGTCTTTCATTGGCGGAATCAGATCC**
851 **GTTGCGCCTGCCATAACCCTGGCTGTGCTTCCAAGCTCAAGGAACAGATTGCACGCCTCCCCGATGGTCTTAGGAGCAAGATA**
852 **CTCGAATTGAGGTAATACCATcat**atgTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGACGCCGTCCACGCTGCCT
853 CCTCACGTGACGTGAGGTGCAAGCCCGGACGTTCCGCGTGCCACGCCGTGAGCCGCCGCGTGCCGTCGGCTCCCTCAGCCCGG
854 GCGGCCGTGGGAGCCCGCCTCGATATGTACACCCGAGAAGCTCCCAGCGTCCTCCTGGGCCGCGATACTCGACCACCACGCAC
855 GCACACCGCACTAACGATTCGGCCGGCGCTCGATTCGGCCGGCGCTCGATTCGGCCGGCGCTCGATTCGGCCGGCGCTCGATT
856 CGGCCGGCGCTCGATTCGGCCGAGCAGAAGAGTGAACAACCACCGACCACGCTTCCGCTCTGCGCGCCGTACCCGACCTACCT
857 CCCGCAGCTCGAAGCAGCTCCCGGGAGTACCGCCGTACTCACCCGCCTGTGCTCACCATCCACCGACGCAAAGCCCAACCCGA
858 GCACACCTCTTGCACCAAGGTGCCGACCGTGGCTTTCCGCTCGCAGGGTTCCAGAAGAAATCGAACGATCCAGCGCGGCAAGG
859 TTCAAAAAGCAGGGGTTGGTGGGGAGGAGGTTTTGGGGGGTGTCGCCGGGATACCTGATATGGCTTTGTTTTGCGTAGTCGAA
860 TAATTTTCCATATAGCCTCGGCGCGTCGGACTCGAATAGTTGATGTGGGCGGGCACAGTTGCCCCATGAAATCCGCAACGGGG
861 GGCGTGCTGAGCGATCGGCAATGGGCGGATGCGGTGTTGCTTCCGCACCGGCCGTTCGCGACGAACAACCTCCAACGAGGTCA
862 GTACCGGATGAGCCGCGACGACGCATTGGCAATGCGGTACGTCGAGCATTCACCGCACGCGTTGCTCGGATCTATCGTCATCG
863 ACTGCGATCACGTTGACGCCGCGATGCGCGCATTCGAGCAACCATCCGACCATCCGGCGCCGAACTGGGTCGCACAATCGCCG
864 TCCGGCCGCGCACACATCGGATGGTGGCTCGGCCCCAACCACGTGTGCCGCACCGACAGCGCCCGACTGACGCCACTGCGCTA
865 CGCCCACCGCATCGAAACCGGCCTCAAGATCAGCGTCGGCGGCGATTTCGCGTATGGCGGGCAACTGACCAAAAACCCGATTC
866 ACCCCGATTGGGAGACGATCTACGGCCCGGCCACCCCGTACACATTGCGGCAGCTGGCCACCATCCACACACCCCGGCAGATG
867 CCGCGTCGGCCCGATCGGGCCGTGGGCCTGGGCCGCAACGTCACCATGTTCGACGCCACCCGGCGATGGGCATACCCGCAGTG
868 GTGGCAACACCGAAACGGAACCGGCCGCGACTGGGACCATCTCGTCCTGCAGCACTGCCACGCCGTCAACACCGAGTTCACGA
869 CACCACTGCCGTTCACCGAAGTACGCGCCACCGCGCAATCCATCTCCAAATGGATCTGGCGCAATTTCACCGAAGAACAGTAC
870 CGAGCCCGACAAGCGCATCTCGGTCAAAAAGGCGGCAAGGCAACGACACTCGCCAAACAAGAAGCCGTCCGAAACAATGCAAG
871 AAAGTACGACGAACATACGATGCGAGAGGCGATTATCTGATGGGCGGAGCCAAAAATCCGGTGCGCCGAAAGATGACGGCAGC
872 AGCAGCAGCCGAAAAATTCGGTGCCTCCACTCGCACAATCCAACGCTTGTTTGCTGAGCCGCGTGACGATTACCTCGGCCGTG
873 CGAAAGCTCGCCGTGACAAAGCTGTCGAGCTGCGGAAGCAGGGGTTGAAGTACCGGGAAATCGCCGAAGCGATGGAACTCTCG
874 ACCGGGATCGTCGGCCGATTACTGCACGACGCCCGCAGGCACGGCGAGATTTCAGCGGAGGATCTGTCGGCGTAACCAAGTCA
875 GCGGGTTGTCGGGTTCCGGCCGGCGCTCGGCACTCGGACCGGCCGGCGGATGGTGTTCTGCCTCTGGCGCAGCGTCAGCTACC
876 GCCGAAGGCCTGTCATCGACCGGCTTCGACTGAAGTATGAGCAACGTCACAGCCTGTGATTGGATGATCCGCTCACGCTCGAC
877 CGCTACCTGTTCAGCTGCCGCCCGCTGGGCATGAGCAACGGCCAACTCTCGTTCAAGAGCTCGACCGCGCGGGTCCCGGACGG
878 GGAAGAGCGGGGAGCTTTGCCAGAGAGCGACGACTTCCCCTTGCGTTGGTGATTGCCGGTCAGGGCAGCCATCCGCCATCGTC
879 GCGTAGGGTGTCACACCCCAGGAATCGCGTCACTGAACACAGCAGCCGGTAGGACGACCATGACTGAGTTGGACACCATCGCA
880 AATCCGTCCGATCCCGCGGTGCAGCGGATCATCGATGTCACCAAGCCGTCACGATCCAACATAAAGACAACGTTGATCGAGGA
881 CGTCGAGCCCCTCATGCACAGCATCGCGGCCGGGGTGGAGTTCATCGAGGTCTACGGCAGCGACAGCAGTCCTTTTCCATCTG
882 AGTTGCTGGATCTGTGCGGGCGGCAGAACATACCGGTCCGCCTCATCGACTCCTCGATCGTCAACCAGTTGTTCAAGGGGGAG
883 CGGAAGGCCAAGACATTCGGCATCGCCCGCGTCCCTCGCCCGGCCAGGTTCGGCGATATCGCGAGCCGGCGTGGGGACGTCGT
884 CGTTCTCGACGGGGTGAAGATCGTCGGGAACATCGGCGCGATAGTACGCACGTCGCTCGCGCTCGGAGCGTCGGGGATCATCC
885 TGGTGGACAGTGACATCACCAGCATCGCGGACCGGCGTCTCCAAAGGGCCAGCCGAGGTTACGTCTTCTCCCTTCCCGTCGTT
886 CTCTCCGGTCGCGAGGAGGCCATCGCCTTCATTCGGGACAGCGGTATGCAGCTGATGACGCTCAAGGCGGATGGCGACATTTC
887 CGTGAAGGAACTCGGGGACAATCCGGATCGGCTGGCCTTGCTGTTCGGCAGCGAAAAGGGTGGGCCTTCCGACCTGTTCGAGG
888 AGGCGTCTTCCGCCTCGGTTTCCATCCCCATGATGAGCCAGACCGAGTCTCTCAACGTTTCCGTTTCCCTCGGAATCGCGCTG
889 CACGAGAGGATCGACAGGAATCTCGCGGCCAACCGATAAGCGCCTCTGTTCCTCGGACGCTCGGTTCCTCGACCTCGATTCGT
890 CAGTGATGATCACCTCACACGGCAGCGATCACCACTGACATATCGAGGTCAACGGTCGTGGTCCGGGCGGGCACTCCTCGAAG
891 GCGCGGCCGACGCCCTTGAACGACTCGATGACTCTAGACGCATCCGAAACCTCCACCCCACTCACCTAGTCCGACATCCGTAC
892 CTTGGAAACCGACCTGTATTGGCATTTCAGTTGGACATCGACCAGTGGCGTTGCTAGGTTCAAGACCATGTCCAGCCCGAAGG
893 CGTCCAGACTCTAGCCACCGGAGGTAGTCCGGTGGCCACATCCCGTCGCGCCCGAACGTCACGCTCTTGTGTGGCCTTCCCTT
894 GTTGTTTGCGATCAGTGGCACACCTCTACCGTCTGAATTTCGAGTCTGGCCTCGGCTGCGCACATCTCGCACTGTGACGCTGT
895 CAGGTCACCCGCTTCGCGGCTACCAGTTCCTTTCATCGAATCGAGCTTCCGGTGCCGCCGCGCAGCCTCCCTGACCATCCTCA
896 GATTTTATGGAGTCTCGCAGTGCCTTTCGCTATCTACGTCCTCGGGCTTGCTGTCTTCGCCCAGGGCACATCCGAGTTCATGT
897 TGTCCGGACTCATACCGGACATGGCCCGTGACCTCGGGGTTTCGGTCCCCGCCGCCGGACTCCTCACCTCCGCCTTCGCGGTC
898 GGGATGATCATCGGCGCTCCGTTGATGGCTATCGCCAGCATGCGGTGGCCCCGGCGACGCGCCCTTCTGACATTCCTCATCAC
899 GTTCATGCTGGTCCACGTCATCGGCGCGCTCACCAGCAGCTTCGAGGTCTTGCTGGTCACACGCATCGTGGGAGCCCTCGCCA
900 ATGCCGGATTCTTGGCAGTGGCCCTGGGGGCGGCGATGGCGATGGTGCCCGCCGACATGAAAGGGCGCGCCACGTCCGTCCTC
901 CTCGGCGGTGTCACGATCGCATGTGTAGCCGGTGTTCCCGGGGGCGCCTTCCTGGGTGAAATGTGGGGCTGGCGTGCAGCGTT
902 CTGGGCTGTCGTCGTCATCTCCGCCCCTGCAGTGGTGGCGATTATGTTCGCCACCCCGGCCGAGCCGCTTGCAGAGTCCACAC
903 CGAATGCCAAGCGTGAACTGTCCTCACTGCGCTCACGCAAGCTCCAGCTCATGCTTGTCCTCGGGGCGCTGATCAACGGCGCA
904 ACGTTCTGTTCGTTCACGTACATGGCGCCCACGCTCACCGACATCTCCGGTTTCGACTCCCGTTGGATTCCGTTGCTGCTGGG
905 GCTGTTCGGGCTCGGATCGTTCATCGGTGTCAGCGTCGGAGGCAGGCTCGCCGACACCCGGCCGTTCCAACTGCTCGCTGTCG
906 GGTCCGCAGCACTGTTGACGGGATGGATCGTCTTCGCTCTCACGGCATCCCACCCCGCGGTGACATTGGTGATGCTGTTCGTG
907 CAGGGCGCTTTGTCCTTCGCGGTCGGCTCGACTTTGATCTCCCAGGTGCTCTACGCCGCCGACGCGGCACCGACCTTGGGTGG

908 ATCGTTCGCGACGGCCGCGTTCAACGTCGGTGCTGCACTGGGACCGGCGCTCGGCGGGTTGGCGATCGGCATGGGTCTGAGCT
909 ACCGCGCCCCGCTCTGGACGAGCGCCGCGCTGGTGACACTCGCGATCGTCATCGGCGCAGCCACCTTGTCTCTGTGGCGGCGA
910 CCAGCGTCTGTCCACGAATCTGTCCCCGCCTGACCAGAAACCAGGATCTGTGAGTGTGGTGACTGATCTGTGCACGCTCAGCA
911 GTCACCGCGCGCTCGCGTCGTACCGAGGGCCAGCGCCAACAGGTGTGTGGAGCTCTGCCCCTGCCTCTTTCACGCGAACTCAC
912 TGTTCAGTGCGGCGATACGTGCTCGGTGAGTTCCACTACAGCGACCATGACTAGAATTGATCTCCTCGACCGCCAATTGGGCA
913 TCTGAGAATCATCTGCGTTTCTCGCACGCAACGTACTTGCAACGTTGCAACTCCTAGTGTTGTGAATCACACCCCACCGGGGG
914 GTGGGATTGCAGTCACCGATTTGGTGGGTGCGCCCAGGAAGATCACGTTTACATAGGAGCTTGCAATGAGCTACTCCGTGGGA
915 CAGGTGGCCGGCTTCGCCGGAGTGACGGTGCGCACGCTGCACCACTACGACGACATCGGCCTGCTCGTACCGAGCGAGCGCAG
916 CCACGCGGGCCACCGGCGCTACAGCGACGCCGACCTCGACCGGCTGCAGCAGATCCTGTTCTACCGGGAGCTGGGCTTCCCGC
917 TCGACGAGGTCGCCGCCCTGCTCGACGACCCGGCCGCGGACCCGCGCGCGCACCTGCGCCGCCAGCACGAGCTGCTGTCCGCC
918 CGGATCGGGAAACTGCAGAAGATGGCGGCGGCCGTGGAGCAGGCGATGGAGGCACGCAGCATGGGAATCAACCTCACCCCGGA
919 GGAGAAGTTCGAGGTCTTCGGCGACTTCGACCCCGACCAGTACGAGGAGGAGGTCCGGGAACGCTGGGGGAACACCGACGCCT
920 ACCGCCAGTCCAAGGAGAAGACCGCCTCGTACACCAAGGAGGACTGGCAGCGCATCCAGGACGAGGCCGACGAGCTCACCCGG
921 CGCTTCGTCGCCCTGATGGACGCGGGTGAGCCCGCCGACTCCGAGGGGGCGATGGACGCCGCCGAGGACCACCGGCAGGGCAT
922 CGCCCGCAACCACTACGACTGCGGGTACGAGATGCACACCTGCCTGGGCGAGATGTACGTGTCCGACGAACGTTTCACGCGAA
923 ACATCGACGCCGCCAAGCCGGGCCTCGCCGCCTACATGCGCGACGCGATCCTCGCCAACGCCGTCCGGCACACCCCCTGAGCG
924 GTGGTCGTGGCCCGGGTCTCCCGCCCGGTCTCACCCCACGGCTCACTCCCGGGCCACGACCACCGCCGTCCCGTACGCGCACA
925 CCTCGGTGCCCACGTCCGCCGCCTCCGTCACGTCGAAACGGAAGATCCCCGGGTACCGAGCTCGTCAGGTGGCACTTTTCGGG
926 GAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAA
927 ATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTG
928 CCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCG
929 AACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTG
930 CTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGT
931 TGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTG
932 ATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCAT
933 GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAAT
934 GGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGG
935 ATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGG
936 TCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAAC
937 TATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCAT
938 ATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAA
939 ATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCT
940 GCGCGTAATCTGCTGCTTGCAAAC

941

942 NdeI and XhoI restriction sites are lowercase and the *ucdCFO* insert is in bold.

943 **Online Methods**

944 **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
Strain, strain background	*Eggerthella lenta* A2	DSMZ	DSM 110911	
Strain, strain background	*Enterococcus faecium* DSM 20477	DSMZ	DSM 20477	Type strain
Strain, strain background	*Gordonibacter pamelaeae* DSM 19378	DSMZ	DSM 19378	Type strain
Strain, strain background	*Gordonibacter pamelaeae* DSM 110924	DSMZ	DSM 110924	
Strain, strain background	*Gordonibacter urolithinfaciens* DSM 27213	DSMZ	DSM 27213	Type strain
Strain, strain background	*Rhodococcus erythropolis* DSM 43066	DSMZ	DSM 43066	Type strain
Strain, strain background	*Escherichia coli* NEB10β	NEB	C3019	Cloning strain
Strain, strain background	*Escherichia coli* NEB10β pTipQC2	This study		
Strain, strain background	*Escherichia coli* NEB10β pTipQC2-*ucdCFO*	This study		
Strain, strain background	*Rhodococcus erythropolis* DSM 43066 pTipQC2	This study		
Strain, strain background	*Rhodococcus erythropolis* DSM 43066 pTipQC2-*ucdCFO*	This study		

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Commercial assay or kit	One-4-All Genomic DNA Minipreps Kit	BioBasic	BS88503	
Commercial assay or kit	EZ-10 Spin Column Plasmid DNA Miniprep Kit	BioBasic	BS414	
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	Monarch DNA Gel Extraction Kit	NEB	T1020	
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	GelCode Blue Stain Reagent	Fisher	24590	
Commercial assay or kit	XhoI Restriction Enzyme	NEB	R0146	
Commercial assay or kit	NdeI Restriction Enzyme	NEB	R0111	
Commercial assay or kit	Hi-T4 DNA Ligase	NEB	M2622	

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Plasmids and primers	pTipQC2	Hokkaido System Science Co.	RE-0006	GenBank AB127591.1
Plasmids and primers	pTipQC2-*ucdCFO*	This work	This work	See Supplementary Sequence 2
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
Commercial assay or kit	BLUelf Prestained Protein Ladder	FroggaBio	PM008-0500	5-235 kDa
Chemical compound, drug	SIGMAFAST Protease Inhibitor Cocktail, EDTA-free	Sigma	S8830-2TAB	
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

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
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	Ferrous sulfate heptahydrate	BioShop	FER005	CAS 7782-63-0
Chemical compound, drug	Iron(III) pyrophosphate	Sigma	P6526-100G	CAS 10058-44-3
Chemical compound, drug	LB Broth (Miller)	BioShop	LBL407.1	
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

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Chemical compound, drug	Lysozyme	BioShop	LYS702.10	CAS 12650-88-3
Chemical compound, drug	Triton X-100	Sigma	X100-100ML	CAS 9036-19-5
Chemical compound, drug	Glycerol	Sigma	G2025	CAS 56-81-5
Chemical compound, drug	Sodium dodecyl sulfate (SDS)	BioShop	SDS001	CAS 151-21-3
Chemical compound, drug	Bromophenol blue	BioShop	BRO777	CAS 6262-5-28-9
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	Caffeic acid	Sigma	C0625-2G	CAS 331-39-5
Chemical compound, drug	m-coumaric acid	MedChemExpress	HY-113357	CAS 588-30-7
Chemical compound, drug	Dopamine hydrochloride	Sigma	H8502-5G	CAS 62-31-7
Chemical compound, drug	m-tyramine	MedChemExpress	HY-128975	CAS 38449-59-1
Chemical compound, drug	Entacapone	Sigma	SML0654-10MG	CAS 130929-57-6

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Chemical compound, drug	2,2'-bipyridyl	Sigma	D216305-10G	CAS 366-18-7
Chemical compound, drug	Resorcinol	Sigma	398047-100G	CAS 108-46-3
Chemical compound, drug	2-bromo-4,5-dimethoxybenzoic acid	Sigma	441074-5G	CAS 6286-46-0
Chemical compound, drug	Sodium hydroxide	Sigma	221465-500G	CAS 1310-73-2
Chemical compound, drug	Copper (II) sulfate pentahydrate	Sigma	209198-100G	CAS 7758-99-8
Chemical compound, drug	TRI reagent	Zymo Research	R2050-1-50	
Chemical compound, drug	NADH disodium salt, trihydrate, reduced	BioBasic	NB0642.SIZE.1g	CAS 606-68-8
Chemical compound, drug	NADPH	BioBasic	Y4433000.SIZE.100mg	CAS 2646-71-1
Chemical compound, drug	FAD disodium salt hydrate	Sigma	F6625-100MG	CAS 84366-81-4
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	MilliQ water system	Millipore	IQ7000	
Lab equipment, instruments	Gene Pulser Cuvette (0.1 cm gap)	Bio-Rad	1652089	
Lab equipment, instruments	Gene Pulser Xcell Total Electroporation System	Bio-Rad	1652660	
Lab equipment, instruments	AnaeroPack System	Mitsubishi Gas Chemical Company	10-01	

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
Lab equipment, instruments	Vinyl anaerobic chambers	Coy Laboratory Products	Type A	
Lab equipment, instruments	Multiscan FC Microplate Photometer	Thermo Fisher	51119100	
Lab equipment, instruments	BioTek Epoch 2 Microplate Spectrophotometer	Agilent	EPOCH2NS-SN	
Lab equipment, instruments	Costar 96-Well Clear Round Bottom	Corning	3788	
Lab equipment, instruments	Breathe-Easy Sealing Membranes	Sigma	Z380059-1PAK	
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	

Reagent type or resource	Designation	Source or reference	Identifiers	Additional information
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	BioTek Gen6 v1.03.01	Agilent	N/A	
Software	R Studio 2023.06.0+421	Posit Software	N/A	
Software	Affinity Designer	Serif Ltd	N/A	

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

947 **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
Eb_ucdCFO_NdeI_f	Cloning	This work	CCCTGGcatatgATGGTATTACCTCAATTCGAG
Eb_ucdCFO_XhoI_r	Cloning	This work	CTGTCTctcgagCTGCCTTTATTTTGAATTCCTG
pTipQC2-ucdCFO_cPCR_f	cPCR	This work	CGAGGGAGCTTCCAGGGGGAAA
pTipQC2-ucdCFO_cPCR_r	cPCR	This work	GAAGGCCAGCTGGAGGGATCCA

948 **Adapter sites** are bolded. Restriction sites are lowercase.

949 **Anaerobic bacterial strains and culturing conditions**

950 Bacterial strains used in this study are listed in the *Resources table*. All bacterial stains
951 were validated by sequencing the 16S rRNA gene (see Genomic DNA extraction and 16S rRNA
952 sequencing of bacterial isolates). The same culture used for validation was used to make 25%
953 glycerol stocks. Anaerobic strains were grown from glycerol stocks on mABB+H agar plates
954 (recipe below) for 48-72 h at 37 °C in a vinyl anaerobic chamber, which was maintained with a
955 gas mixture of 3% H_2, 10% CO_2, 87% N_2. To make overnight cultures, a single colony was
956 inoculated into 5 mL of liquid mABB or mABB+H medium and incubated at 37 °C between 16-
957 48h, depending on the bacterium (16-24 h for *Enterocloster* spp. and *E. faecium*, and 48 h for
958 *E. lenta* and *Gordonibacter* spp.).

959 **Human fecal sample collection**

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

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

969 For 1 L of modified anaerobe basal broth (mABB), the following components were
970 dissolved in MilliQ water, then autoclaved: 16 g peptone, 7 g yeast extract, 5 g sodium chloride,
971 1 g starch, 1 g D-glucose monohydrate, 1 g sodium pyruvate, 0.5 g sodium succinate, 1 g
972 sodium thioglycolate, 15 g agar (for plates). The autoclaved solution was allowed to cool, then

973 the following filter-sterilized solutions were added aseptically: 10 mL of 100 mg/mL L-arginine-
974 HCl, 10 mL of 50 mg/mL L-cysteine, 8 mL of 50 mg/mL sodium bicarbonate, 50 µL of 10 mg/mL
975 vitamin K1, 20 mL of 50 mg/mL dithiothreitol, and, for mABB+H, 10 mL of 0.5 mg/mL haemin.
976 The media was then placed in the anaerobic chamber and allowed to reduce for at least 24 h
977 prior to its use in experiments.

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

979 The identities of all bacteria in this study were validated by full-length (V1-V9) 16S rRNA
980 sequencing. DNA was first extracted from 0.5-1 mL of overnight culture using the One-4-All
981 Genomic DNA Miniprep Kit (BioBasic) according to the manufacturer's instructions. The purified
982 genomic DNA (2 µL, ~ 20 ng) was used as a template for PCR reactions (25 µL reaction volume)
983 using the Q5 High-Fidelity polymerase (NEB). PCR tubes were placed in a thermal cycler and
984 targets (~1.5 kb) were amplified according to the following cycling conditions: 30 s at 98 °C, 30
985 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
986 the reaction was mixed with 6X loading buffer and loaded onto a 1% agarose gel (made with
987 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were compared to the
988 ExcelBand 100 bp+3K DNA Ladder (SMOBIO).

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

994 **Treatments with urolithins and other catechols**

995 All treatments used in this study (urolithin C, urolithin M6, urolithin A, isourolithin A, 8,9-
996 di-O-methyl-urolithin C, dopamine, m-tyramine, caffeic acid, m-coumaric acid, entacapone, and
997 2,2'-bipyridyl) were dissolved in DMSO to a concentration of 10 mM.

998 *Treatment prior to growth (metabolism only)*: Overnight cultures of bacteria were diluted 1/50
999 into fresh mABB+H. Treatments (10 mM stocks solutions, dissolved in DMSO) were added to
1000 the diluted bacterial suspension to a final concentration of 100 μM and samples were incubated
1001 for 24 h at 37 °C in an anaerobic chamber.

1002 *Treatment during growth (spike-in)*: Overnight cultures of *Enterocloster* spp. were diluted 1/50
1003 into fresh mABB+H and incubated at 37 °C in an anaerobic chamber. After 5 hours of incubation
1004 (~OD_{620} from a 200 μL sample ~ 0.4), 10 mM urolithins (or an equivalent volume of DMSO)
1005 were added to the growing cultures at a final concentration of 50 or 100 μM for protein
1006 expression or RNA expression, respectively. For protein expression analyses and inducibility
1007 tests, the cultures were incubated for an additional 4 h. For RNA expression analyses, the
1008 cultures were incubated for an additional 2 h.

1009 *Treatment prior to growth (growth curves)*: Overnight cultures of *Enterocloster* spp. were diluted
1010 1/25 into mABB (with or without added 15.4 μM iron source), depending on the experimental
1011 design. Separately, treatments (10 mM stocks solutions, dissolved in DMSO) were prepared in
1012 mABB (with or without added 15.4 μM iron source) to a final concentration of 200 μM. In each
1013 well of a 96-well plate, 100 μL of 1/25 bacteria and 200 μM treatment were combined. These
1014 were plated in technical duplicates. The final concentration of treatment was 100 μM (unless
1015 otherwise specified in concentration-response experiments) and the final dilution of bacteria
1016 was 1/100.

1017 **Urolithin extraction from fecal slurries or bacterial cultures**

1018 Frozen (-70 °C) fecal slurries or bacterial cultures were thawed at room temperature. For
1019 quantification of urolithin concentrations, urolithin standards (stock 10 mM in DMSO) were
1020 spiked into separate media aliquots immediately before extraction.

1021 *Extraction Method A*: This method was used for cultures. Salicylic acid (3 mg/mL in DMSO)
1022 was spiked-in as an internal standard at a final concentration of 50 µg/mL. The cultures and
1023 standards were then extracted with 3 volumes of ethyl acetate + 1% formic acid (e.g., 600 µL
1024 solvent to 200 µL thawed culture). The organic phase (top) was transferred to a new tube and
1025 dried in a vacuum concentrator (Heto Lab) connected to a rotary evaporator (Buchi). After
1026 solvent removal, samples were redissolved in 0.5 volumes (relative to the starting culture) of
1027 50% MeOH:H_2O. Samples were centrifuged at 20,000 x g for 5 min to pellet insoluble material,
1028 then transferred to LC-MS vials. Urolithins were then analyzed by LC-MS.

1029 *Extraction Method B*: This method was used for cultures and pre-induced cell suspensions.
1030 Samples were diluted with an equal volume of MeOH, vortexed briefly, and incubated at room
1031 temperature for 10 min. Samples were centrifuged at 20,000 x g for 5 min to pellet insoluble
1032 material, then transferred to LC-MS vials. Urolithins were then analyzed by LC-MS.

1033 *Extraction Method C*: This method was used to extract urolithins from crude bacterial lysates.
1034 Lysates were diluted with 3 volumes of MeOH, vortexed briefly, and incubated at room
1035 temperature for 10 min. Samples were centrifuged at 20,000 x g for 5 min to pellet insoluble
1036 material, then transferred to LC-MS vials. Urolithins were then analyzed by LC-MS.

1037 **LC-MS method to quantify urolithins**

1038 Samples (10 µL) were injected into a 1260 Infinity II Single Quadrupole LC/MS system
1039 (Agilent) fitted with a Poroshell 120 EC-C18 4.6x50 mm, 2.7 µm column (Agilent). The mobile

1040 phase was composed of MilliQ water + 0.1% formic acid (solvent A) and acetonitrile + 0.1%
1041 formic acid (solvent B). The flow rate was set to 0.7 mL/min. The gradient was as follows: 0-8
1042 min: 10-30 %B, 8-10 min: 30-100 %B, 10-13.5 min: 100 %B isocratic, 13.5-13.6 min: 100-10
1043 %B, then 13.6-15.5 min: 10 %B. The multiple wavelength detector was set to monitor
1044 absorbance at 305 nm. The mass spectrometer was run in negative mode in both selected ion
1045 monitoring (SIM) and scan (100-1000 m/z) modes to validate peak identities. Peaks were
1046 validated based on retention times compared to spike-in standards and mass-to-charge ratios.
1047 To quantify urolithins, peak areas for the compounds of interest were compared to spike-in
1048 standards of known concentration(s). When standards were not available (urolithin M7), the
1049 extracted ion chromatogram was used ($[M-H]^-$ = 243).

1050 **Synthesis of di-O-methyl-urolithin C**

1051 Di-O-methyl-urolithin C (3-Hydroxy-8,9-dimethoxy-6H-dibenzo[b,d]pyran-6-one, CAS
1052 126438-35-5) was synthesized based on previously described Ullmann-type coupling
1053 conditions for urolithin derivatives [60]. Resorcinol (213 mg, 2 mmol) and 2-bromo-4,5-
1054 dimethoxybenzoic acid (261 mg, 1 mmol) were dissolved in 1 mL of 8% w/v NaOH (in MilliQ
1055 H_2O) and heated in a thermo-shaker set to 100 °C for 20 min (in 1.7 mL tube). Then, 200 µL of
1056 a 10% w/v Cu(II)SO_4 pentahydrate solution was added and the reaction was heated at 100 °C
1057 for 1 h. The reaction solution (pink-red coloration) contained an insoluble precipitate which was
1058 collected by centrifugation (20,000 x g for 30 s). The insoluble pellet was washed 7 times with
1059 1 mL of MilliQ H_2O until the pH of the wash solution was equal to the pH of MilliQ H_2O (~pH 6).
1060 The pellet was dried by lyophilization for 16 h (0.0010 mbar, -90 °C) and the product was
1061 recovered as a pale pink solid (94 mg, 35% yield).

1062 ^{1}H NMR (600 MHz, $(CD_3)_2SO$)): δ = 10.22 (s, 1H), 8.21 (d, J = 8.76 Hz, 1H), 7.69 (s, 1H), 7.54
1063 (s, 1H), 6.83 (dd, J = 8.67, 2.37 Hz, 1H), 6.74 (d, J = 2.34 Hz, 1H), 4.02 (s, 3H), 3.89 (s, 3H);
1064 HRMS: m/z $[M+Na]^+$ calculated for $C_{15}H_{12}NaO_5$: 295.0577, found: 295.0585.

1065 **Phylogenetic tree construction**

1066 Phylogenetic trees based on the 16S rRNA gene were constructed using the DSMZ
1067 single-gene phylogeny server (https://ggdc.dsmz.de/phylogeny-service.php#)[61] with GenBank
1068 16S rRNA sequence accessions: *E. aldenensis* (DQ279736), *E. asparagiformis* (AJ582080), *E.*
1069 *bolteae* (AJ508452), *E. citroniae* (HM245936), *E. clostridioformis* (M59089), *E. lavalensis*
1070 (EF564277), *L. pacaense* (LT799004), *G. pamelaeae* (AM886059), *G. urolithinfaciens*
1071 (HG000667), *E. isourolithinifaciens* (MF322780).

1072 Phylogenetic trees based on whole genomes and proteomes were constructed using the
1073 Type (Strain) Genome Server (TYGS, https://tygs.dsmz.de)[40,62] with the following GenBank
1074 genome accessions: *E. aldenensis* (GCA_003467385.1), *E. asparagiformis*
1075 (GCA_025149125), *E. bolteae* (GCA_000154365), *E. clostridioformis* (GCA_900113155), *E.*
1076 *lavalensis* (GCA_003024655), *L. pacaense* (GCA_900566185). For *E. citroniae*, the Integrated
1077 Microbial Genomes ObjectID was used: *E. citroniae* (2928404274). Further information on
1078 nomenclature and taxonomy was obtained from the List of Prokaryotic names with Standing in
1079 Nomenclature (LPSN, available at https://lpsn.dsmz.de).

1080 **Cell suspension assay to test inducibility**

1081 Bacteria (10 mL growing cultures in mABB+H media) were grown with 50 μM uroC (or
1082 an equivalent volume of DMSO) as detailed in the *Treatment during growth (spike-in)* section
1083 above and incubated for 4 h at 37 °C. Cultures were then pelleted at 6,500 x g for 3 min and
1084 the supernatants were discarded. The cells were washed with 10 mL of pre-reduced PBS

1085 (placed in the anaerobic chamber 24 h before), re-pelleted, and resuspended in 2 mL of pre-
1086 reduced PBS. For each condition tested, a 200 µL aliquot of the cell suspension was transferred
1087 into a sterile 1.5 mL tube, and 10 mM urolithins (uroM6, uroC, isouroA, or DMSO) were added
1088 at a final concentration of 100 µM. Cell suspensions were briefly vortexed and incubated at
1089 room temperature in the anaerobic chamber for 16h prior to freezing and urolithin extraction
1090 using *Extraction Method B*.

1091 **RNA extraction from isolates**

1092 A volume of 1.5 mL of treated (100 µM urolithin C for 2 h) *Enterocloster* spp. culture (see
1093 *Enterocloster* spp. urolithin C treatments) was pelleted (6,500 g for 3 min) and the supernatant
1094 was removed for later LC-MS analysis. The pellet (suspended in 200 µL of media) was then
1095 mixed with 800 µL TRI reagent (Zymo Research) and transferred to a ZR BashingBead lysis
1096 tube (Zymo Research). Samples were lysed in a Mini Beadbeater 16 (Biospec) according to
1097 the following sequence: 1 min ON, 5 min OFF. For RNA-sequencing, bead beating was done
1098 for a total of 5 min ON. For RT-(q)PCR, bead beating was done for a total of 2 min ON to
1099 preserve longer transcripts. RNA isolation was then performed using the Direct-zol RNA
1100 Miniprep Kit (Zymo Research) according to the manufacturer's instructions (including an on-
1101 column DNase digestion). To ensure complete DNA removal, an additional DNA digestion step
1102 was performed on the isolated RNA using the Ambion *DNA-free* DNA Removal Kit (Invitrogen)
1103 according to the manufacturer's instructions. The DNA-free RNA was then cleaned up using
1104 the *OneStep* PCR Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were
1105 initially verified by NanoDrop and 1 % agarose gel electrophoresis. For RNA-sequencing, RNA
1106 integrity was assessed by Génome Québec using a Bioanalyzer 2100 (Agilent). RNA integrity
1107 (RIN) values ranged between 7.5-7.8 for *E. asparagiformis* DSM 15981 and 7.0-7.3 for *E.*
1108 *bolteae* DSM 15670.

1109 **RNA-sequencing of urolithin C-treated *E. bolteae* and *E. asparagiformis* isolates**

1110 Total RNA was sent to Génome Québec for library preparation and RNA-sequencing.
1111 Briefly, total RNA was prepared for Illumina sequencing using the NEBNext rRNA Depletion Kit
1112 (Bacteria) (NEB) kit to remove rRNA and using the NEBNext Multiplex Oligos for Illumina (NEB)
1113 kit (stranded/directional). Prepared libraries were quality checked with a Bioanalyzer 2100
1114 (Agilent) prior to sequencing. Sequencing was performed on a NovaSeq 6000 (Illumina) with
1115 the following flow cell/settings: S4 flow cell, 100 bp, 25 M reads, paired end.

1116 Analysis of RNA-seq reads was done using Galaxy bioinformatics cloud computing
1117 (https://usegalaxy.org/) hosted by Compute Canada Genetics and Genomics Analysis Platform
1118 (GenAP) (https://www.genap.ca/). Genomes and annotations were fetched from the NCBI
1119 genome browser: *E. bolteae* (ASM223457v2) (accessed 2022/05/11) and *E. asparagiformis*
1120 (ASM2514912v1) (accessed 2023/09/19). Raw reads were first verified for quality using
1121 FastQC (v0.73, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) with default
1122 parameters. FastQC reports were aggregated into MultiQC (v1.11) [63]. The mean sequence
1123 quality scored were above 35 for all samples. Raw reads were then trimmed using Cutadapt
1124 (v3.7) [64] to trim adapter sequences (R1 sequence:
1125 AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC, and R2 sequence:
1126 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT) that were not removed after sequencing
1127 using default parameters for paired end reads. Trimmed reads were then aligned to reference
1128 genomes for each bacterium using HISAT2 (v2.2.1) [65] with paired-end parameters and reverse
1129 strandedness (RF). Aligned read counts were assigned to features in annotation files (.gtf)
1130 using featureCounts (v2.0.1) [66] with the following parameters: reverse strandedness, count
1131 fragments instead of reads, GFF feature type filter = “gene”, multi-mapping and multi-
1132 overlapping features included (-M -O), minimum mapping quality per read of 0, and the rest of

1133 the parameters were kept as default. Differential gene expression analysis was then performed
1134 using DESeq2 (v2.11.40.7) [67] using default parameters. Differential expression tables were
1135 annotated with the Annotate DESeq2/DEXSeq output tables tool (v1.1.0) in Galaxy to include
1136 the following: GFF feature type = “CDS”, GFF feature identifier = “gene_id”, GFF transcript
1137 identifier = “transcript_id”, GFF attribute to include = “protein_id, product”. The “protein_id” was
1138 used to query the NCBI database and the NCBI Sequence Viewer was used to investigate the
1139 genomic context surrounding genes of interest.

1140 **Comparative genomics**

1141 The nucleotide sequence for the *Enterocloster bolteae* DSM 15670 *ucd* operon (NCBI
1142 NZ_CP022464 REGION: complement(4417875..4421605)) was used as a query for BLASTn
1143 (megablast) searches using the refseq_genomes database limited to Bacteria (taxid:2). The
1144 NCBI multiple sequence alignment (MSA) viewer was used to download alignment figures.

1145 **RT-PCR analysis of *E. bolteae* to determine *ucd* operon structure**

1146 Isolated RNA samples (500 ng) were reverse transcribed using the LunaScript® RT
1147 Master Mix Kit (Primer-free) (NEB) in a reaction volume of 10 µL containing the ucdCFO_RT-
1148 PCR_r primer at a final concentration of 1 µM. The No-RT Control included in the kit was used
1149 as a no-enzyme control for reverse transcription. The reaction mixtures were incubated in a
1150 thermal cycler: 10 min at 55˚C, 1 min at 95˚C. PCR reactions were conducted using the OneTaq
1151 2X Master Mix with Standard Buffer (NEB). The ucdCFO_RT-PCR primer pair was added to
1152 the master mix (to a final concentration of 0.2 µM) and 1 µL of template (cDNA, -RT, no
1153 template, or gDNA) was added for a total reaction volume of 20 µL. PCR tubes were placed in
1154 a thermal cycler and targets were amplified according to the following conditions: 20s at 94˚C,
1155 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
1156 was directly loaded onto a 1% agarose gel (made with 1X TAE buffer) containing SafeView

1157 Classic (Abm). PCR product sizes were compared to the Quick-Load® Purple 1 kb Plus DNA
1158 Ladder (NEB). The rest of the PCR product was then run on a 1% agarose gel and bands
1159 corresponding to the desired products were cut out and purified using the Monarch DNA Gel
1160 Extraction kit (NEB). DNA was quantified using the Qubit dsDNA HS assay kit (Invitrogen) and
1161 submitted to Plasmidsaurus for long-read sequencing using Oxford Nanopore Technologies
1162 (Supplementary Sequence 1).

1163 **RT-qPCR analysis of *E. bolteae ucd* operon genes**

1164 Isolated RNA samples (500 ng) were reverse transcribed using the iScript Reverse
1165 Transcription Supermix (Bio-Rad) in a reaction volume of 10 µL. The iScript No-RT Control
1166 Supermix was used as a no enzyme control for reverse transcription (-RT). The reaction
1167 mixtures were incubated in a thermal cycler: 5 min at 25 °C, 20 min at 48 °C, and 1 min at 95
1168 °C. Both cDNA and -RT controls were diluted 1/20 in nuclease-free water before use. qPCR
1169 reactions were conducted using the Luna Universal qPCR Master Mix kit (NEB). The
1170 Eb_ucdO_qPCR, Eb_ucdF_qPCR, Eb_ucdC_qPCR, and Eb_dnaK_Ref_qPCR primer pairs
1171 were added to their respective master mixes (final primer concentration of 250 nM) and 6.6 or
1172 4.4 µL of diluted template (cDNA, -RT, no template) were added to 26.4 or 17.6 µL of master
1173 mix for triplicates or duplicates, respectively. All cDNA samples were run in technical triplicates,
1174 while other sample types were run in technical duplicates. Replicate mixes were pipetted (10
1175 µL/well) into a MicroAmp Fast 96-Well Reaction Plate (Applied Biosystems) and the plates were
1176 sealed, then spun down for 2 min to eliminate air bubbles. The qPCR detection parameters
1177 were as follows: SYBR Green detection, ROX reference dye, 10 µL reaction volume. The
1178 thermal cycling conditions were: 1 min at 95 °C, 40 cycles (15 s at 95 °C, 30 s at 60 °C), then
1179 melt analysis (60-95 °C). Data were analyzed according to the $2^{-\Delta\Delta Ct}$ method [68] with the dnaK
1180 gene serving as the reference gene (*E. bolteae* dnaK RNA-seq $\log_2$FC = 0.122).

1181 **Protein extraction from *Enterocloster* spp.**

1182 All steps other than sonication were carried out under anaerobic conditions. To extract
1183 proteins, 10 mL of treated (50 μM urolithin C for 4 h) *Enterocloster* spp. culture (see
1184 *Enterocloster* spp. urolithin C treatments) were pelleted (6,500 g for 3 min) and the supernatant
1185 was discarded. The pellet was washed with 10 mL of pre-reduced PBS, pelleted again, and
1186 resuspended in 0.5 mL of pre-reduced lysis buffer (20 mM Tris, pH 7.5, 500 mM NaCl, 10 mM
1187 $MgSO_4$, 10 mM $CaCl_2$, and 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free)). The
1188 resuspended pellet was then sonicated on ice using a Misonix Sonicator 3000 set to power
1189 level 2/10 according to the following sequence (aerobically, in a cold room): 20 s ON, 40 s OFF,
1190 for a total of 2 min ON. Tubes were centrifuged at 20,000 x g for 2 min to pellet insoluble
1191 particles and 0.4 mL of lysate was transferred to a new 1.5 mL tube (kept on ice). Lysates used
1192 in metabolism assays were transported to the anaerobic chamber in a sealed plastic bag
1193 containing an anaerobic gas generating system to minimize loss in activity.

1194 **Urolithin metabolism assays using crude lysates from uroC-induced *E. bolteae***

1195 Protein lysates (described above) were aliquoted (50 μL aliquots) into 1.5 mL tubes, then
1196 treated with DMSO or urolithin C (10 mM stock) at a final concentration of 350 μM. Cofactors
1197 (NADPH, NADH, and FAD, each dissolved to a final concentration of 30 mM (in lysis buffer
1198 immediately before the assay was run) were added individually to the lysates at a final
1199 concentration of 1 mM. The lysates were incubated at room temperature in an anaerobic
1200 chamber for 20 h prior to freezing at -70 °C. Samples were then extracted using *Extraction*
1201 *Method C*.

1202 To assess the oxygen sensitivity of crude lysates from uroC-induced *E. bolteae*, samples were
1203 prepared as described above. After adding DMSO or uroC and NADH (under anaerobic
1204 conditions), tubes were either incubated at room temperature inside the anaerobic chamber or

1205 just outside of the chamber for 20 h. Afterwards, samples were frozen at -70 °C and then
1206 extracted using *Extraction Method C*.

1207 **Proteomics analysis of uroC-treated *E. bolteae***

1208 Extracted proteins were submitted for proteomic analysis at the RI-MUHC. For each
1209 sample, protein lysates were loaded onto a single stacking gel band to remove lipids,
1210 detergents, and salts. The gel band was reduced with DTT, alkylated with iodoacetic acid, and
1211 digested with trypsin. Extracted peptides were re-solubilized in 0.1% aqueous formic acid and
1212 loaded onto a Thermo Acclaim Pepmap (Thermo, 75 um ID X 2 cm C18 3 um beads) precolumn
1213 and then onto an Acclaim Pepmap Easyspray (Thermo, 75 um ID X 15 cm with 2 um C18
1214 beads) analytical column separation using a Dionex Ultimate 3000 uHPLC at 250 nL/min with
1215 a gradient of 2-35% organic (0.1% formic acid in acetonitrile) over 3 hours. Peptides were
1216 analyzed using a Thermo Orbitrap Fusion mass spectrometer operating at 120,000 resolution
1217 (FWHM in MS1) with HCD sequencing (15,000 resolution) at top speed for all peptides with a
1218 charge of 2+ or greater. The raw data were converted into *.mgf format (Mascot generic format)
1219 for searching using the Mascot 2.6.2 search engine (Matrix Science) against *Enterocloster*
1220 *bolteae* DSM 15670 proteins (NCBI assembly GCF_002234575.2) and a database of common
1221 contaminant proteins. Mascot was searched with a fragment ion mass tolerance of 0.100 Da
1222 and a parent ion tolerance of 5.0 ppm. O-63 of pyrrolysine, carboxymethyl of cysteine and j+66
1223 of leucine/isoleucine indecision were specified in Mascot as fixed modifications. Deamidation
1224 of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable
1225 modifications.

1226 The database search results were loaded into Scaffold Q+ Scaffold_5.0.1 (Proteome Sciences)
1227 for statistical treatment and data visualization. Scaffold (v5.3.0) was used to validate MS/MS
1228 based peptide and protein identifications. Peptide identifications were accepted if they could be

1229 established at greater than 95.0% probability by the Peptide Prophet algorithm [69] with Scaffold
1230 delta-mass correction. Protein identifications were accepted if they could be established at
1231 greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities
1232 were assigned by the Protein Prophet algorithm [70]. Proteins that contained similar peptides and
1233 could not be differentiated based on MS/MS analysis alone were grouped to satisfy the
1234 principles of parsimony. Proteins sharing significant peptide evidence were grouped into
1235 clusters. Protein quantification and differential expression were determined in Scaffold using
1236 the following parameters: Quantitative method was set to total spectra, the minimum value was
1237 set to 0.5 in case proteins were not detected in one condition, and statistical tests were
1238 performed using Fisher's exact test with the Benjamini-Hochberg multiple test correction at a
1239 significance level set to 0.05.

1240 **Protein structures**

1241 The protein FASTA sequences (NCBI RefSeq accessions for UcdC: WP_002569575.1,
1242 UcdF: WP_002569574.1, UcdO: WP_002569573.1) for the *Enterocloster bolteae* DSM 15670
1243 *ucd* operon were used as a query for BLASTp searches against the UniProtKB reference
1244 proteomes + Swiss-Prot databases. The AlphaFold2 protein structures for matches (UniProt
1245 accessions for UcdC: A8RZR5, UcdF: A8RZR3, UcdO: A8RZR2) of each protein were
1246 downloaded and imported into PyMOL (v2.4.1). Foldseek (in 3Di/AA mode) was used to
1247 generate a list of proteins with similar structures from solved crystal structures in the Protein
1248 Data Bank (PDB) [48].

1249 Hits of published X-ray crystal structures (PDB: 1ZXI (from *Afipia carboxidovorans* OM5) and
1250 PDB: 3UNI (from *Bos taurus*)) were fetched from the PDB and imported into PyMOL. The
1251 AlphaFold2 structures for each Ucd protein were aligned to the following chains in the published

1252 PDBs using the “super” command: PDB 1ZXI: UcdC to chain C, UcdF to chain A, UcdO to chain
1253 B; PDB 3UNI: UcdC, UcdF, and UcdO to chain A.

1254 **pTipQC2-*ucdCFO* Plasmid Construction, Purification, and Transformation**

1255 *Plasmid construction in E. coli* NEB10β: Primers flanking the *E. bolteae ucd* operon
1256 (NCBI NZ_CP022464 REGION: complement(4417875..4421605)) were designed in Benchling
1257 using the Primer3 tool. Tails including 6 random bases, followed by restriction sites for NdeI
1258 and XhoI were included on the forward and reverse primers, respectively (Eb_ucdCFO_NdeI_f
1259 and Eb_ucdCFO_XhoI_r). PCR was performed using the Q5 High-Fidelity polymerase (NEB)
1260 with *E. bolteae* DSM 15670 genomic DNA as a template. The target was amplified according
1261 to the following cycling conditions: 30 s at 98 °C, 30 cycles (10 s at 98 °C, 20 s at 60 °C, 80 s
1262 at 72 °C), 2 min at 72 °C. The *ucdCFO* PCR product was purified using the Monarch PCR &
1263 DNA Cleanup Kit (NEB) according to the manufacturer’s instructions for products ≥ 2 kb. The
1264 resulting purified PCR product and the pTipQC2 plasmid (Hokkaido Systems Science Co.) were
1265 digested overnight (16 h) with NdeI and XhoI (both from NEB) in rCutSmart buffer according to
1266 the manufacturer’s instructions (~600-1000 ng DNA per 50 µL reaction). Double digested DNA
1267 was migrated on a 0.6% agarose gel and bands corresponding to the desired products were
1268 cut out and purified using the Monarch DNA Gel Extraction kit (NEB). The purified products
1269 were ligated using the Hi-T4 DNA Ligase (NEB): a ~3:1 insert:plasmid molar ratio ligation
1270 reaction was set up on ice, then incubated at room temperature for 2 h. The ligation mixture (2
1271 µL) was electroporated (1.8 kV, 25 µF, 200 Ω) into 40 µL electrocompetent *E. coli* NEB10β cells
1272 (according to the Quick-n’-Dirty Electrocompetent *E. coli* protocol
1273 (dx.doi.org/10.17504/protocols.io.bjpykmpw) using 0.1 cm gap cuvettes (Bio-Rad). The cuvette
1274 was immediately filled with 1 mL pre-warmed LB post-shock and cells were allowed to recover
1275 at 37 °C for 30 min before plating on LB + 100 µg/mL ampicillin. After an overnight incubation

1276 at 37 °C, colonies were picked and grown in selective LB + 100 µg/mL ampicillin. Plasmids
1277 were purified using the Plasmid DNA Miniprep Kit (BioBasic) and size was confirmed with a
1278 diagnostic restriction digest (10 µL reactions). The final plasmid construct (pTipQC2-*ucdCFO*)
1279 was submitted to Plasmidsaurus for long-read sequencing using Oxford Nanopore
1280 Technologies (v14 library preparation chemistry, R10.4.1 flow cells) (Supplementary Sequence
1281 2).

1282 *pTipQC2-ucdh transformation into Rhodococcus erythropolis DSM 43066*:
1283 Electrocompetent *R. erythropolis* DSM 43066 were prepared according to a modified protocol
1284 from P. Lessard 2002. Briefly, 50 mL LB were inoculated with 1 mL of a stationary phase (48-
1285 72 h growth from a single colony) *R. erythropolis* DSM 43066 culture and grown aerobically for
1286 16 h at 30 °C with shaking at 200 RPM. The next day, cells were pelleted at 5,000 x g for 10
1287 min at 4 °C and washed according to the following sequence: 2 washes of (10 mL of ice cold
1288 sterile MilliQ water), 10 mL of ice cold sterile 10% glycerol. The final pellet was then
1289 resuspended in 5 mL of ice cold sterile 10% glycerol. The resuspended electrocompetent *R.*
1290 *erythropolis* DSM 43066 were aliquoted (50 µL/aliquot), then 3 µL (~0.5-1 µg) of pTipQC2-
1291 *ucdCFO* plasmid was added to appropriate tubes and incubated for 30 min on ice. Cells with
1292 plasmid were transferred to 0.1 cm gap cuvettes (Bio-Rad) and electroporated (1.8 kV, 25 µF,
1293 200 Ω). Time constants were between 4.3-4.6 ms. The cuvette was immediately filled with 1
1294 mL LB post shock and cells were allowed to recover at 30 °C for 2.5 h before plating 100 µL
1295 dilutions (1/10 dilution, undiluted, and concentrated recovery culture) on LB + 30 µg/mL
1296 chloramphenicol at 30 °C. After 2-3 days of incubation, colonies were picked and grown in
1297 selective liquid LB + 30 µg/mL chloramphenicol at 30 °C with shaking at 200 RPM. Plasmid-
1298 positive colonies were identified by colony PCR using the pTipQC2-*ucdCFO*_cPCR primer set
1299 and validated by diagnostic restriction digests and whole-plasmid sequencing.

1300 **Heterologous expression of UcdCFO enzymes in *Rhodococcus erythropolis*:**

1301 All growth steps below were performed in selective media (LB + 30 µg/mL
1302 chloramphenicol) in aerobic conditions at 30 °C with shaking at 200 RPM, unless otherwise
1303 specified. Single colonies of *R. erythropolis* DSM 43066 harboring the pTipQC2 (empty
1304 plasmid) or pTipQC2-*ucdCFO* were inoculated into 5 mL selective media and grown for 72 h to
1305 produce overnight cultures. Overnight cultures were then thoroughly resuspended and diluted
1306 1:10 into 25 mL fresh selective media and grown for ~8 h until OD_{600} values reached ~0.6.
1307 Thiostrepton (5 mg/mL in DMSO) was added to a final concentration of 1 µg/mL and cultures
1308 were incubated aerobically for 16 h at 25 °C to induce protein expression. The next morning,
1309 cultures were pelleted and resuspended in 0.2 volumes of lysis buffer (20 mM Tris, pH 7.5, 500
1310 mM NaCl, 10 mM $MgSO_4$, 10 mM $CaCl_2$, 2 mM DTT, 1% Triton X-100, 2 mg/mL lysozyme, and
1311 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free)). The resuspended cells in lysis
1312 buffer were incubated on ice for 1 h with shaking, then sonicated on ice using a Misonix
1313 Sonicator 3000 set to power level 2/10 according to the following sequence (aerobically, in a
1314 cold room): 20 s ON, 40 s OFF, for a total of 4 min ON. Crude lysates were transported to the
1315 anaerobic chamber in a sealed plastic bag containing an anaerobic gas generating system to
1316 minimize loss in activity and treated in the same manner detailed in *Urolithin metabolism assays*
1317 *using crude lysates from uroC-induced E. bolteae*.

1318 **SDS-PAGE analysis of UcdCFO proteins in crude lysates**

1319 Crude lysates described above (*Heterologous expression of UcdCFO enzymes in*
1320 *Rhodococcus erythropolis*) were centrifuged for 2 min at 20,000 x g. The insoluble pellet was
1321 separated from the soluble supernatant. The insoluble pellet (from 100 µL of crude lysate) was
1322 resuspended in 100 µL of 1X reducing loading dye (62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS,
1323 10% glycerol, 0.01% (w/v) bromophenol blue, 38 mM DTT). The soluble fraction was diluted

1324 with 3X reducing loading dye to a final concentration of 1X. All samples were heated at 95 °C
1325 for 5 min, then 10 µL were loaded onto a 10% bis-tris polyacrylamide protein gel. Gels were
1326 fixed and stained with GelCode Blue Stain Reagent (Fisher) according to the manufacturer's
1327 instructions.

1328 **Growth curves with catechols in different iron-containing media conditions**

1329 Overnight cultures of *Enterocloster* spp. were treated as described in the *Treatment prior*
1330 *to growth (growth curves)* sub-section of *Treatments with urolithins and other catechols*.

1331 Once plated, 96-well plates were sealed with a Breathe-Easy membrane and placed in a pre-
1332 warmed plate reader inside the anaerobic chamber (BioTek Epoch 2). The optical density at
1333 620 nm was recorded every 30 min for 48 h. Kinetic analysis was performed in BioTek Gen6
1334 Software using the built-in kinetic analysis.

1335 **curatedMetagenomicData meta-analysis of *Enterocloster ucd* operon in human fecal**
1336 **metagenome datasets**

1337 All 93 metagenomic studies (22,588 samples and their metadata) available in the
1338 curatedMetagenomicData R package [53] (v3.8.0) were downloaded locally (ExperimentHub
1339 snapshotDate(): 2023-04-24, accessed on 2023-06-06) and transferred to the Narval cluster
1340 hosted by the Digital Research Alliance of Canada. Metagenomic data for urolithin C-
1341 metabolizing *Enterocloster* spp. were obtained by querying the "relative_abundance" (pre-
1342 processed using MetaPhlAn3 and "gene_families" (pre-processed using HUMAnN3) entries in
1343 individual study datasets [71]. For individual taxa (containing partial strings "bolteae", "citroniae",
1344 "asparagiformis", "asparagiforme", "Enterocloster", or "47FAA" (corresponding to *L.*
1345 *pacaense*)), relative abundance (%) was extracted from the rows of the "relative_abundance"

1346 datasets using the stringr R package (v1.5.0, https://github.com/tidyverse/stringr). Prevalence
1347 (relative abundance in sample > 0) was then calculated for each sample.

1348 For specific genes, the NCBI protein accessions for each gene of the *ucd* operon (*ucdO*, *ucdF*,
1349 *ucdC*) was used to search the UniProt database. UniRef90 accession numbers corresponding
1350 to hits (C5EGQ4, G5HFF3, A8RZR5, respectively) were then extracted from the rows of the
1351 "gene_families" datasets using the stringr R package. Prevalence (abundance in sample > 0)
1352 was then calculated for each sample. R scripts and RData files are available in Zenodo (see
1353 Data Availability).

1354 **Fecal slurry preparation and treatment**

1355 Frozen (-70 °C) fecal samples were brought into the anaerobic chamber and allowed to
1356 thaw. The samples were suspended in 1 mL mABB medium per 0.1 g feces and homogenized
1357 by breaking apart large pieces with a sterile loop and by vortexing. Large particles were pelleted
1358 by centrifuging the tubes at 700 x g for 3 min. The supernatants (containing bacteria) were
1359 transferred to new tubes and centrifuged at 6,500 x g for 5 min to pellet the cells. The
1360 supernatants were discarded, and the cell pellets were washed with 5 mL of fresh media. The
1361 cell suspensions were once again centrifuged at 6,500 x g for 5 min and the resulting cell pellets
1362 were resuspended in 600 µL media per 0.1 g feces. Resuspended cells were treated with either
1363 100 µM urolithin C or an equivalent volume of DMSO and incubated at 37˚C anaerobically for
1364 48 h. 200-300 µL volumes were removed from the batch cultures and immediately frozen at -
1365 70˚C for later extraction of urolithins (using *Extraction Method A*), DNA, and RNA.

1366 **Genomic DNA extraction from fecal slurries**

1367 A 300 µL fecal slurry aliquot (between 50-100 mg wet weight) was pelleted (10,000 g for
1368 5 min) and the supernatant was removed for later LC-MS analysis. The pellet was then mixed

1369 with 750 µL of ZymoBIOMICS lysis solution (Zymo Research) and transferred to a ZR
1370 BashingBead lysis tube (Zymo Research). Samples were lysed in a Mini Beadbeater 16
1371 (Biospec) according to the following sequence: 1 min ON, 5 min OFF for a total of 5 min ON.
1372 DNA was then purified using the ZymoBIOMICS DNA Miniprep Kit (Zymo Research) according
1373 to the manufacturer's instructions (including the *OneStep* PCR Inhibitor Removal step). Purified
1374 DNA samples were quantified using the Qubit dsDNA HS assay kit (Invitrogen).

1375 **Long read 16S rRNA sequencing of microbial communities in fecal slurries**

1376 Long read 16S PCR reactions were conducted using the Platinum SuperFi II Green PCR
1377 Master Mix (Invitrogen). The ONT_16S_27F_GGK and ONT_16S_1492R_GGK primer pairs
1378 (see *Primer sequences table*) were added to their respective master mixes (final primer
1379 concentration of 0.2 µM) and 1 µL of template (~10 ng) was added (for a total reaction volume
1380 of 25 µL). PCR tubes were placed in a thermal cycler and targets were amplified according to
1381 the following cycling conditions: 30 s at 98 °C, 30 cycles (10 s at 98 °C, 10 s at 60 °C, 30 s at
1382 72 °C), 5 min at 72 °C, and hold at 4 °C. Amplicons were quantified using the Qubit dsDNA HS
1383 assay kit (Invitrogen) to verify that amplicon concentrations were reasonably balanced (range
1384 = 18.36-24.00 ng/µL). Barcoding of amplicons was performed with 2 µL of PCR reaction
1385 according to the manufacturer's instructions (for ONT kit SQK-AMB111-24). Barcoding
1386 reactions were incubated in a thermal cycler for 10 min at 65 °C, then for 2 min at 80 °C. 10 µL
1387 of each barcoding reaction were pooled and proteins were digested using heat-labile proteinase
1388 K (NEB) by incubating the pooled library for 15 min at 37 °C, followed by heat inactivation for
1389 10 min at 55 °C. Amplicons were purified using Agencourt AMPure XP beads (Beckman
1390 Coulter) using 0.7 volumes of beads-to-library. Following 70% EtOH washes and drying steps,
1391 the library was eluted using 15 µL of the provided elution buffer (EB), yielding a library with a
1392 concentration of 30 ng/µL using the Qubit dsDNA HS assay kit (Invitrogen). 11 µL of the eluted

1393 DNA library were transferred to a new tube and combined with 1 µL of Rapid Adapter T (RAP
1394 T). This mixture was incubated at room temperature for 10 min. Since the library was
1395 concentrated, it was diluted 1:2 in EB before combining with SB II and LB II, then loaded into a
1396 primed Flongle Flow Cell (R9.4.1) in a MinION device following the manufacturer's instructions.
1397 Sequencing was allowed to proceed for ~20 h until pore exhaustion or enough reads were
1398 obtained. Base calling & demultiplexing was performed using Guppy (v6.4.6) using the "SUP"
1399 super high accuracy model for R9.4.1 flow cells. The raw reads were filtered for a length
1400 between 1500 ± 200 bp. Filtered reads were assigned to taxa using Emu [72] (v3.4.4, GitLab
1401 Project ID: 19618062) by mapping 16S rRNA sequences to the emu_database database
1402 (based on the NCBI 16S RefSeq with the entry for *E. asparagiformis* changed to the sequence
1403 obtained by ONT sequencing (GenBank accession PP280819) since the RefSeq sequence for
1404 this bacterium contained multiple N nucleotides that biased the assignment of *E. asparagiformis*
1405 to *E. lavalensis*). Data were not rarefied or scaled. Count tables were then used to create a
1406 phyloseq (v1.44.0, https://github.com/joey711/phyloseq) object in R [73]. Stacked bar plots were
1407 generated using ggnested (v0.1.0, https://github.com/gmteunisse/ggnested) and fantaxtic
1408 (v0.2.0, https://github.com/gmteunisse/Fantaxtic). Diversity analyses were performed using
1409 Microbiome Analyst (https://www.microbiomeanalyst.ca/) [74].

1410 **Total RNA extraction from fecal slurries**

1411 A 300 µL fecal slurry aliquot (treated with either 100 µM urolithin C or an equivalent
1412 volume DMSO for 48 h) was thawed and pelleted (6,500 g for 3 min). The pellet (in 200 µL
1413 media) was then mixed with 800 µL TRI reagent (Zymo Research). Samples were lysed in a
1414 Mini Beadbeater 16 (Biospec) according to the following sequence: 1 min ON, 5 min OFF for a
1415 total of 5 min ON. RNA isolation was then performed using the Direct-zol RNA Miniprep Kit
1416 (Zymo Research) according to the manufacturer's instructions (including an on-column DNase

1417 digestion). To ensure complete DNA removal, an additional DNA digestion step was performed
1418 on the isolated RNA using the Ambion DNA-free DNA Removal Kit (Invitrogen) according to the
1419 manufacturer's instructions. The DNA-free RNA was then cleaned up using the OneStep PCR
1420 Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were verified by Qubit
1421 RNA BR assay kit (Invitrogen) and 1% agarose gel electrophoresis.

1422 **RT-PCR analysis of the *ucd* operon in fecal slurries**

1423 Total RNA was extracted from frozen fecal slurries as previously described (see Total
1424 RNA extraction from microbial communities), and subsequently reverse transcribed as
1425 described above (see *RT-PCR analysis of E. bolteae ucd operon structure*) in a reaction volume
1426 of 5 µL. PCR reactions were conducted using the OneTaq 2X Master Mix with Standard Buffer
1427 (NEB). The ucdCFO_RT-PCR primer pair was added to the master mix (to a final concentration
1428 of 0.2 µM) and 1 µL of template (cDNA, -RT, or no template) was added for a total reaction
1429 volume of 20 µL. PCR tubes were placed in a thermal cycler and targets were amplified
1430 according to the following conditions: 30s at 94˚C, 45 cycles (30s at 94˚C, 1 min at 61˚C, 4 min
1431 at 68˚C), 5 min at 68˚C. A volume of 10 µL of reaction was directly loaded onto a 1% agarose
1432 gel (made with 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were
1433 compared to the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB).

1434 **PCR analysis of the *ucd* operon prevalence in fecal slurries**

1435 Genomic DNA (gDNA) was extracted from frozen fecal slurries as previously described
1436 (see Genomic DNA extraction from microbial communities). PCR reactions and product
1437 visualization was conducted on the gDNA as described above (see RT-PCR analysis of the *ucd*
1438 operon in microbial communities). In this case, 5 µL of PCR product was loaded onto the gels
1439 instead of 10 µL.

1440 **Statistical analyses and graphing**

1441 Statistical methods were not used to determine sample sizes, experiments were not
1442 randomized, and the investigators were not blinded. Data points related to uroC metabolism
1443 and RT-qPCR were assumed to be normally distributed, though this was not formally tested.
1444 Correlation analyses were performed using the non-parametric Spearman rank correlation (ρ).
1445 Statistical tests on bacterial relative abundances were performed using the Kruskal-Wallis test
1446 on untransformed relative abundance values, which are skewed towards 0. Statistical analyses
1447 for large datasets are detailed in the relevant methods sections. Details related to each test
1448 performed are supplied in the figure legends. In all cases, $\alpha = 0.05$ and tests were two-tailed.
1449 Data were plotted in GraphPad Prism (v10.0.0) or using the ggplot2 (v3.4.2) R package. Figures
1450 were assembled in Affinity Designer (v1.10.6.1665).

1451 **Data Availability**

1452 RNA-seq reads were deposited in the NCBI SRA BioProject ID PRJNA996126 under
1453 BioSample accession codes SAMN36514640 (*Enterocloster bolteae* DSM 15670) and
1454 SAMN36514641 (*Enterocloster asparagiformis* DSM 15981). Reviewer link:
1455 https://dataview.ncbi.nlm.nih.gov/object/PRJNA996126?reviewer=fpbuj6eeuebv6mij3pme3mp
1456 8ep. Untargeted proteomics data have been deposited to the ProteomeXchange Consortium
1457 via the PRIDE partner repository with the dataset identifier PXD048514 and
1458 10.6019/PXD048514 [75]. Oxford Nanopore 16S rRNA sequencing reads of healthy human fecal
1459 slurries were deposited in the NCBI SRA BioProject ID PRJNA1073957. Reviewer link:
1460 https://dataview.ncbi.nlm.nih.gov/object/PRJNA1073957?reviewer=nh1a04lg59enlti318qs23s
1461 6ot. The 16S rRNA sequence for *E. asparagiformis* DSM 15670 used in the Emu database
1462 search was deposited in GenBank under accession PP280819. All original code, tables, and
1463 RData files obtained from the analysis of curatedMetagenomicData were deposited in Zenodo
1464 (https://doi.org/10.5281/zenodo.8302320).

1465 **References**

1466 1 Knight, R. *et al.* The Microbiome and Human Biology. *Annu Rev Genomics Hum Genet*
1467 **18**, 65-86 (2017). https://doi.org/10.1146/annurev-genom-083115-022438
1468 2 Taguer, M. & Maurice, C. F. The complex interplay of diet, xenobiotics, and microbial
1469 metabolism in the gut: Implications for clinical outcomes. *Clin Pharmacol Ther* **99**, 588-
1470 599 (2016). https://doi.org/10.1002/cpt.366
1471 3 Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut
1472 microbiome. *Nat Biotechnol* **39**, 105-114 (2021). https://doi.org/10.1038/s41587-020-
1473 0603-3
1474 4 Spanogiannopoulos, P., Bess, E. N., Carmody, R. N. & Turnbaugh, P. J. The microbial
1475 pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol*
1476 **14**, 273-287 (2016). https://doi.org/10.1038/nrmicro.2016.17
1477 5 Ervin, S. M. *et al.* Structural Insights into Endobiotic Reactivation by Human Gut
1478 Microbiome-Encoded Sulfatases. *Biochemistry* **59**, 3939-3950 (2020).
1479 https://doi.org/10.1021/acs.biochem.0c00711
1480 6 Yao, L. *et al.* A biosynthetic pathway for the selective sulfonation of steroidal metabolites
1481 by human gut bacteria. *Nat Microbiol* **7**, 1404-1418 (2022).
1482 https://doi.org/10.1038/s41564-022-01176-y
1483 7 Tomas-Barberan, F. A. & Espin, J. C. Effect of Food Structure and Processing on
1484 (Poly)phenol-Gut Microbiota Interactions and the Effects on Human Health. *Annu Rev*
1485 *Food Sci Technol* **10**, 221-238 (2019). https://doi.org/10.1146/annurev-food-032818-
1486 121615
1487 8 Maier, L. *et al.* Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature*
1488 **555**, 623-628 (2018). https://doi.org/10.1038/nature25979
1489 9 Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R. & Goodman, A. L.
1490 Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature*
1491 **570**, 462-467 (2019). https://doi.org/10.1038/s41586-019-1291-3
1492 10 Maini Rekdal, V., Bess, E. N., Bisanz, J. E., Turnbaugh, P. J. & Balskus, E. P. Discovery
1493 and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism.
1494 *Science* **364**, eaau6323 (2019). https://doi.org/10.1126/science.aau6323
1495 11 Campbell, C. *et al.* Bacterial metabolism of bile acids promotes generation of peripheral
1496 regulatory T cells. *Nature* **581**, 475-479 (2020). https://doi.org/10.1038/s41586-020-
1497 2193-0
1498 12 Paik, D. *et al.* Human gut bacteria produce Tau(Eta)17-modulating bile acid metabolites.
1499 *Nature* **603**, 907-912 (2022). https://doi.org/10.1038/s41586-022-04480-z
1500 13 Cerqueira, F. M., Photenhauer, A. L., Pollet, R. M., Brown, H. A. & Koropatkin, N. M.
1501 Starch Digestion by Gut Bacteria: Crowdsourcing for Carbs. *Trends Microbiol* **28**, 95-
1502 108 (2020). https://doi.org/10.1016/j.tim.2019.09.004
1503 14 Maini Rekdal, V. *et al.* A widely distributed metalloenzyme class enables gut microbial
1504 metabolism of host- and diet-derived catechols. *Elife* **9** (2020).
1505 https://doi.org/10.7554/eLife.50845
1506 15 Garcia-Villalba, R., Selma, M. V., Espin, J. C. & Tomas-Barberan, F. A. Identification of
1507 Novel Urolithin Metabolites in Human Feces and Urine after the Intake of a Pomegranate
1508 Extract. *J Agric Food Chem* **67**, 11099-11107 (2019).
1509 https://doi.org/10.1021/acs.jafc.9b04435
1510 16 Carmody, R. N. *et al.* Diet dominates host genotype in shaping the murine gut microbiota.
1511 *Cell Host Microbe* **17**, 72-84 (2015). https://doi.org/10.1016/j.chom.2014.11.010

17 Carmody, R. N. *et al.* Cooking shapes the structure and function of the gut microbiome. *Nat Microbiol* **4**, 2052-2063 (2019). https://doi.org/10.1038/s41564-019-0569-4
18 Sonnenburg, E. D. *et al.* Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**, 212-215 (2016). https://doi.org/10.1038/nature16504
19 Rothschild, D. *et al.* Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210-215 (2018). https://doi.org/10.1038/nature25973
20 Quideau, S., Deffieux, D., Douat-Casassus, C. & Pouysegu, L. Plant polyphenols: chemical properties, biological activities, and synthesis. *Angew Chem Int Ed Engl* **50**, 586-621 (2011). https://doi.org/10.1002/anie.201000044
21 van Duynhoven, J. *et al.* Metabolic fate of polyphenols in the human superorganism. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4531-4538 (2011). https://doi.org/10.1073/pnas.1000098107
22 Messaoudene, M. *et al.* A Natural Polyphenol Exerts Antitumor Activity and Circumvents Anti-PD-1 Resistance through Effects on the Gut Microbiota. *Cancer Discov* **12**, 1070-1087 (2022). https://doi.org/10.1158/2159-8290.CD-21-0808
23 Pang, S. A., Elkrief, A., Capella, M. P. & Miller, W. H., Jr. Two Cases of Durable and Deep Responses to Immune Checkpoint Inhibition-Refractory Metastatic Melanoma after Addition of Camu Camu Prebiotic. *Curr Oncol* **30**, 7852-7859 (2023). https://doi.org/10.3390/curroncol30090570
24 Garcia-Villalba, R. *et al.* Urolithins: a Comprehensive Update on their Metabolism, Bioactivity, and Associated Gut Microbiota. *Mol Nutr Food Res* **66**, e2101019 (2022). https://doi.org/10.1002/mnfr.202101019
25 Singh, A. *et al.* Direct supplementation with Urolithin A overcomes limitations of dietary exposure and gut microbiome variability in healthy adults to achieve consistent levels across the population. *Eur J Clin Nutr* **76**, 297-308 (2022). https://doi.org/10.1038/s41430-021-00950-1
26 Denk, D. *et al.* Expansion of T memory stem cells with superior anti-tumor immunity by Urolithin A-induced mitophagy. *Immunity* **55**, 2059-2073 e2058 (2022). https://doi.org/10.1016/j.immuni.2022.09.014
27 Andreux, P. A. *et al.* The mitophagy activator urolithin A is safe and induces a molecular signature of improved mitochondrial and cellular health in humans. *Nat Metab* **1**, 595-603 (2019). https://doi.org/10.1038/s42255-019-0073-4
28 Ryu, D. *et al.* Urolithin A induces mitophagy and prolongs lifespan in C. elegans and increases muscle function in rodents. *Nat Med* **22**, 879-888 (2016). https://doi.org/10.1038/nm.4132
29 Ghosh, S., Banerjee, M., Haribabu, B. & Jala, V. R. Urolithin A attenuates arsenic-induced gut barrier dysfunction. *Arch Toxicol* **96**, 987-1007 (2022). https://doi.org/10.1007/s00204-022-03232-2
30 Singh, R. *et al.* Enhancement of the gut barrier integrity by a microbial metabolite through the Nrf2 pathway. *Nat Commun* **10**, 89 (2019). https://doi.org/10.1038/s41467-018-07859-7
31 Shen, P. X. *et al.* Urolithin A ameliorates experimental autoimmune encephalomyelitis by targeting aryl hydrocarbon receptor. *EBioMedicine* **64**, 103227 (2021). https://doi.org/10.1016/j.ebiom.2021.103227
32 Ghosh, S., Haribabu, B. & Jala, V. R. Microbial metabolite Urolithin A induces expansion of Regulatory T Cells in aryl hydrocarbon receptor (AhR)-dependent manner. *Journal of Immunology* **206**, 113.103-113.103 (2021).

1559 33 Xia, B. *et al.* Urolithin A exerts antiobesity effects through enhancing adipose tissue
1560 thermogenesis in mice. *PLoS Biol* **18**, e3000688 (2020).
1561 https://doi.org/10.1371/journal.pbio.3000688
1562 34 Selma, M. V. *et al.* Isolation of Human Intestinal Bacteria Capable of Producing the
1563 Bioactive Metabolite Isourolithin A from Ellagic Acid. *Front Microbiol* **8**, 1521 (2017).
1564 https://doi.org/10.3389/fmicb.2017.01521
1565 35 Selma, M. V., Beltran, D., Garcia-Villalba, R., Espin, J. C. & Tomas-Barberan, F. A.
1566 Description of urolithin production capacity from ellagic acid of two human intestinal
1567 Gordonibacter species. *Food Funct* **5**, 1779-1784 (2014).
1568 https://doi.org/10.1039/c4fo00092g
1569 36 Gaya, P., Peirotén, A., Medina, M., Alvarez, I. & Landete, J. M. Bifidobacterium
1570 pseudocatenulatum INIA P815: The first bacterium able to produce urolithins A and B
1571 from ellagic acid. *Journal of Functional Foods* **45**, 95-99 (2018).
1572 https://doi.org/10.1016/j.jff.2018.03.040
1573 37 Garcia-Villalba, R. *et al.* Metabolism of different dietary phenolic compounds by the
1574 urolithin-producing human-gut bacteria Gordonibacter urolithinfaciens and Ellagibacter
1575 isourolithinifaciens. *Food Funct* **11**, 7012-7022 (2020).
1576 https://doi.org/10.1039/d0fo01649g
1577 38 Iglesias-Aguirre, C. E. *et al.* Gut Bacteria Involved in Ellagic Acid Metabolism To Yield
1578 Human Urolithin Metabotypes Revealed. *J Agric Food Chem* **71**, 4029-4035 (2023).
1579 https://doi.org/10.1021/acs.jafc.2c08889
1580 39 Iglesias-Aguirre, C. E. *et al.* In vivo administration of gut bacterial consortia replicates
1581 urolithin metabotypes A and B in a non-urolithin-producing rat model. *Food Funct* **14**,
1582 2657-2667 (2023). https://doi.org/10.1039/d2fo03957e
1583 40 Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Goker, M. TYGS and
1584 LPSN: a database tandem for fast and reliable genome-based classification and
1585 nomenclature of prokaryotes. *Nucleic Acids Res* **50**, D801-D807 (2022).
1586 https://doi.org/10.1093/nar/gkab902
1587 41 Hosny, M. *et al.* Clostridium pacaense: a new species within the genus Clostridium. *New*
1588 *Microbes New Infect* **28**, 6-10 (2019). https://doi.org/10.1016/j.nmni.2018.12.003
1589 42 Zupok, A., Iobbi-Nivol, C., Mejean, V. & Leimkuhler, S. The regulation of Moco
1590 biosynthesis and molybdoenzyme gene expression by molybdenum and iron in bacteria.
1591 *Metallomics* **11**, 1602-1624 (2019). https://doi.org/10.1039/c9mt00186g
1592 43 Leimkuhler, S. & Iobbi-Nivol, C. Bacterial molybdoenzymes: old enzymes for new
1593 purposes. *FEMS Microbiol Rev* **40**, 1-18 (2016). https://doi.org/10.1093/femsre/fuv043
1594 44 Koppel, N., Bisanz, J. E., Pandelia, M. E., Turnbaugh, P. J. & Balskus, E. P. Discovery
1595 and characterization of a prevalent human gut bacterial enzyme sufficient for the
1596 inactivation of a family of plant toxins. *Elife* **7** (2018). https://doi.org/10.7554/eLife.33953
1597 45 Nakashima, N. & Tamura, T. Isolation and characterization of a rolling-circle-type
1598 plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-
1599 recombinant-protein expression. *Appl Environ Microbiol* **70**, 5557-5568 (2004).
1600 https://doi.org/10.1128/AEM.70.9.5557-5568.2004
1601 46 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**,
1602 583-589 (2021). https://doi.org/10.1038/s41586-021-03819-2
1603 47 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the
1604 structural coverage of protein-sequence space with high-accuracy models. *Nucleic*
1605 *Acids Res* **50**, D439-D444 (2022). https://doi.org/10.1093/nar/gkab1061

48 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

49 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

50 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

51 Kramer, J., Özkaya, Ö. & Kümmerli, R. Bacterial siderophores in community and host interactions. *Nature Reviews Microbiology* **18**, 152-163 (2020). https://doi.org/10.1038/s41579-019-0284-4

52 Pals, M. J., Wijnberg, L., Yildiz, C. & Velema, W. A. Catechol-Siderophore Mimics Convey Nucleic Acid Therapeutics into Bacteria. *Angew Chem Int Ed Engl* **n/a**, e202402405 (2024). https://doi.org/10.1002/anie.202402405

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

54 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

55 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

56 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

57 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

58 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

59 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

60 Cozza, G. *et al.* Urolithin as a Converging Scaffold Linking Ellagic acid and Coumarin Analogues: Design of Potent Protein Kinase CK2 Inhibitors. *ChemMedChem* **6**, 2273-2286 (2011). https://doi.org/10.1002/cmdc.201100338

61 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

62 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

63 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

64 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

65 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

66 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

67 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

68 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

69 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

70 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

71 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

72 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

73 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

74 Lu, Y. *et al.* MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data. *Nucleic Acids Research* **51**, W310-W318 (2023). https://doi.org/10.1093/nar/gkad407

75 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

1689 **Acknowledgements**

1690 This research was funded by the Canadian Institutes of Health Research (CIHR) grant PJT-
1691 437944 to B.C. and C.F.M. and a Weston Family Microbiome Initiative's 2021 Transformational
1692 Research program to B.C. B.C. holds a tier II Canada Research Chair (CRC) in Therapeutic
1693 Chemistry. C.F.M. holds a tier II CRC in Gut Microbial Physiology and is a Azrieli Global Scholar
1694 in the Humans & Microbiome program. R.P. is supported by the CIHR Canada Graduate
1695 Scholarship-Doctoral (#493808 for 2023-2026) and by the Fonds de Recherche du Québec-
1696 Santé : Bourse de formation au doctorat (#316063 for 2022-2023). M.S. is supported by the
1697 CIHR Canada Graduate Scholarship to Honor Nelson Mandela (#DF2-187718 for 2023-2026),
1698 the Fonds de Recherche du Québec-Santé: Bourse de formation au doctorat (#311071 for
1699 2022-2023).

1700 RNA sample preparation and sequencing was performed by Génome Québec. Untargeted
1701 proteomics sample preparation and analysis was performed by the Research Institute of the
1702 McGill University Health Centre (RI-MUHC) Proteomics and Molecular Analysis Platform.
1703 Preliminary small-molecule MALDI-TOF mass spectrometry was performed by Mark Hancock
1704 at the McGill SPR-MS Facility. High resolution mass spectrometry was performed by the McGill
1705 Chemistry Characterization Facility. Molecular biology experiments were enabled by the McGill
1706 University Imaging and Molecular Biology Platform (IMBP). High-performance cloud computing
1707 was enabled by Calcul Québec (https://www.calculquebec.ca/) and the Digital Research
1708 Alliance of Canada (https://alliancecan.ca/). RNA-seq data analysis was performed using two
1709 instances of Galaxy: (https://usegalaxy.org/) and Compute Canada Genetics and Genomics
1710 Analysis Platform (GenAP) (https://www.genap.ca/).

1711 **Author information**

1712 **Authors and Affiliations**

1713 *Department of Pharmacology & Therapeutics, McGill University, 3655 Prom. Sir-William-Osler,*

1714 *Montreal, Quebec, H3G 1Y6, Canada*

1715 Reilly Pidgeon, Sacha Mitchell, Layan Suleiman, Lharbi Dridi & Bastien Castagner

1716 *Department of Microbiology & Immunology, McGill University, 3775 University Street, Montreal,*

1717 *Quebec, H3A 2B4, Canada*

1718 Michael Shamash & Corinne F. Maurice

1719 **Contributions**

1720 **R.P.** designed the study, performed microbiology experiments, performed bioinformatics
1721 analyses, analyzed data, created figures, and wrote the initial manuscript with B.C. **S.M.**
1722 performed experiments on microbial communities, analyzed data, and created figures. **M.S.**
1723 performed Nanopore sequencing and processed raw sequencing reads. **L.S.** performed
1724 microbiology experiments. **L.D.** consulted on experimental design and methodology. **C.M.**
1725 supervised research, obtained ethical approval for the use of human fecal samples, and
1726 obtained research funding. **B.C.** designed the study, supervised the research, obtained
1727 research funding, and wrote the initial manuscript with R.P. All authors reviewed and edited the
1728 manuscript.

1729 **Corresponding authors**

1730 Correspondence to Bastien Castagner (bastien.castagner@mcgill.ca).

1731 **Ethics declarations**

1732 The authors declare no competing interests.