Article

Diet-derived urolithin A is produced by a dehydroxylase encoded by human gut *Enterocloster* species

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

Urolithin A (uroA) is a polyphenol derived from the multi-step metabolism of dietary ellagitannins by the human gut microbiota. Once absorbed, uroA can trigger mitophagy and aryl hydrocarbon receptor signaling pathways, altering host immune function, mitochondrial health, and intestinal barrier integrity. Most individuals harbor a microbiota capable of uroA production; however, the mechanisms underlying the dehydroxylation of its catechol-containing 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* species. We show that the *ucd* operon encodes a predicted molybdopterin-dependent enzyme complex that dehydroxylates urolithins at a specific position (9-OH). By interrogating publicly available metagenomics datasets, we observed that uroC-metabolizing *Enterocloster* species and *ucd* operon genes are prevalent in human feces. In ex vivo experiments with human fecal samples, only samples actively transcribing *ucd* could produce uroA, possibly explaining differences in urolithin metabolism between individuals. Collectively, this work identifies *Enterocloster* species and the *ucd* operon as important contributors to uroA production and establishes a multi-omics framework to further our mechanistic understanding of polyphenol metabolism by the human gut microbiota.

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

Diet is a strong modulator of the composition and function of the gut microbiota[16–19]. Polyphenols are a diverse class of plant secondary metabolites found in fruits, vegetables, and nuts (reviewed in ref. 20) that are poorly absorbed by the host and reach the large intestine relatively intact[7,21]. Ellagitannins are a large sub-group of polyphenols that belong to the family of hydrolysable tannins and are characterized by central glucose (open-chain or pyranose forms) linked to diverse pyrogallol-like moieties[20]. Camu camu, a berry rich in the ellagitannin castalagin, has been shown to impact anti-cancer immunity via the gut microbiome and is currently in clinical trials in combination with immune checkpoint inhibitors (NCT05303493, NCT06049576)[22,23].

[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. ✉e-mail: bastien.castagner@mcgill.ca

Depending on microbiota composition, ellagitannins like castalagin can be hydrolyzed and reduced by gut bacteria into bioactive metabolites (ellagic acid, urolithins, nasutins) according to different metabolic phenotypes characterized by the terminal metabolites observed in biological fluids[24] (Supplementary Fig. 1).

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

While urolithin metabolites are prevalent in human populations, few gut bacteria have been reported to metabolize urolithins[32–34]. Most known urolithin metabolizers belong to the *Eggerthellaceae* family (*Gordonibacter urolithinfaciens, Gordonibacter pamelaeae, Ellagibacter isourolithinifaciens*) and can perform multiple metabolic steps in the urolithin metabolism pathway, yielding either urolithin C (uroC) or isourolithin A (isouroA) from ellagic acid[35]. Recently, certain members of the *Enterocloster* spp. (*Lachnospiraceae* family) were reported to dehydroxylate uroC to uroA and isouroA to urolithin B (uroB) both in vitro and in vivo[36,37]. These findings shed light on the minimal bacterial community required for the complete metabolism of ellagic acid to uroA; however, the genes and enzymes responsible for these dehydroxylation reactions remain unknown (Fig. 1a).

Here, we use a multi-omics enzyme identification framework to uncover uroC dehydroxylase (*ucd*) genes and enzymes in *Enterocloster* spp. We find that the UcdCFO enzyme complex specifically dehydroxylates 9-hydroxy urolithins and that both metabolizing species and *ucd* genes are prevalent and actively transcribed in human feces during ex vivo metabolism. We further demonstrate that dehydroxylation of urolithins by *Enterocloster* spp. can inactivate growth-delaying catechols. Our study sheds light on the genetic and chemical basis underlying the complex reciprocal interactions between urolithins and the gut microbiota.

Results

A subset of *Enterocloster* species convert uroC to uroA
Members of the *Enterocloster* spp. have previously been shown to dehydroxylate uroC in vitro[36] and in vivo[37] under anaerobic conditions (Fig. 1a, full metabolic pathway in Supplementary Fig. 1). To determine the prevalence of uroC metabolism within this genus, we incubated a subset of *Enterocloster* spp. type strains 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 single-copy marker proteins, genomes, or proteomes (Supplementary Fig. 2a, c). Based on the above results, we chose to perform a more in-depth analysis on *E. asparagiformis* and *E. bolteae* to identify the metabolic gene clusters involved in uroC dehydroxylation.

A dehydroxylase is upregulated upon uroC treatment
To understand when uroC-metabolizing enzymes were being expressed, we first sought to characterize the kinetics of uroC
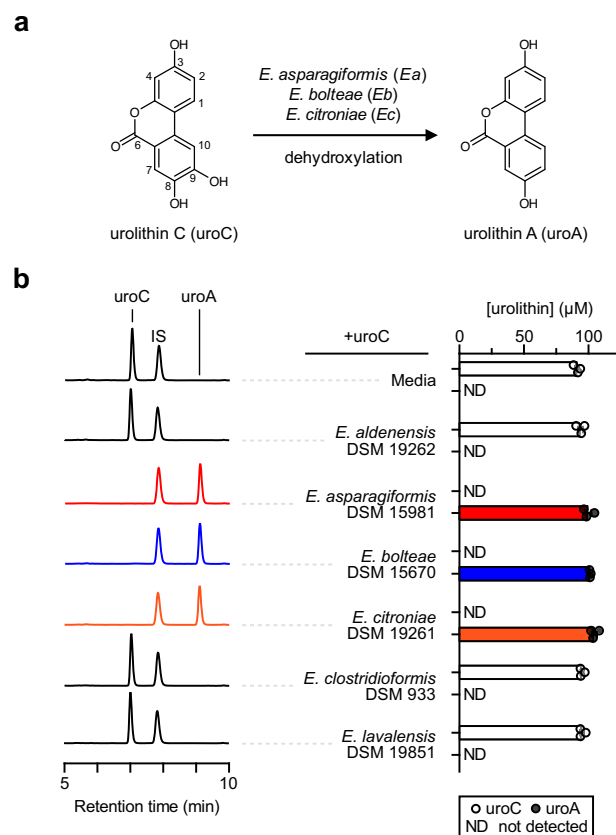

Fig. 1 | Urolithin C metabolism is limited to a subset of *Enterocloster* spp.
a Reaction scheme of uroC dehydroxylation by gut resident *Enterocloster* spp. via unknown enzymes. **b** LC-MS screen of *Enterocloster* spp. type strains for uroC dehydroxylation activity. UroC (100 μM) was added to cultures (in mABB + H media) at the start of growth, and urolithins were extracted after a 24 h anaerobic incubation and then analyzed by LC-MS. Representative chromatograms (λ = 305 nm) for each experimental group (left, from one representative biological replicate). The same scale was used for each chromatogram. Quantification of urolithin peak areas relative to a salicylic acid internal standard (IS) (right, *n* = 3 biological replicates). Coloring for uroC-metabolizing is consistent throughout the manuscript: *E. asparagiformis* (red), *E. bolteae* (blue), *E. citroniae* (orange). Data are represented as mean ± SEM. Source data are provided as a Source data file.

dehydroxylation in rich media (mABB + H). Therefore, a simultaneous growth and metabolism experiment was designed, whereby uroC (or its vehicle, dimethyl sulfoxide (DMSO)) was spiked into *E. asparagiformis* and *E. bolteae* cultures during the exponential phase of growth. Urolithin metabolites were then measured by LC-MS (Fig. 2a). Treatment with uroC during the exponential phase did not affect the 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. Although minor uroA production (3.0 ± 0.4 μM) was observed in DMSO-induced *E. asparagiformis* following uroC treatment, metabolism of uroC to uroA was inducible as cells originating from bacteria grown with uroC were capable of quantitative uroA production (Fig. 2e). Consequently, we performed RNA sequencing to compare gene expression in DMSO and uroC-treated cultures of *E. asparagiformis* and *E. bolteae*. Since uroA was detected in both bacterial cultures as soon as 2 h post spike-in (Fig. 2c), this timepoint was selected to isolate mRNA.
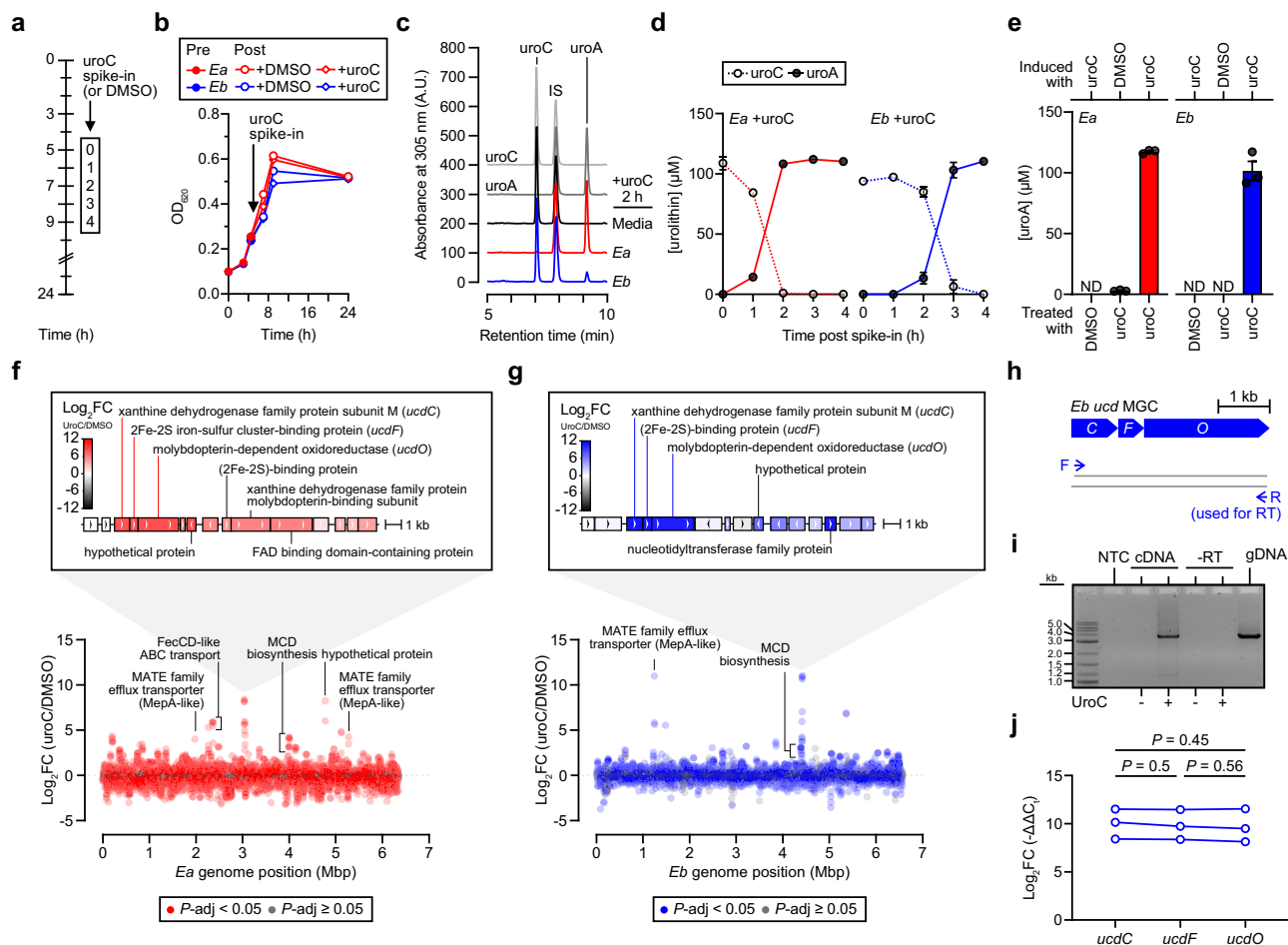
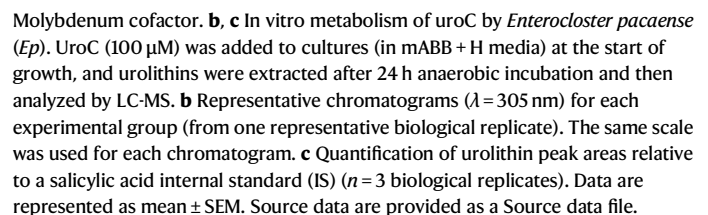**Fig. 2 | Urolithin C treatment upregulates a putative dehydroxylase operon.**
a Experimental design of uroC spike-in experiments during exponential growth. For each biological replicate in this design, (**b**) growth, (**c**, **d**) metabolism, (**f**, **g**) and RNA-seq results are matched. **b** Growth curve (optical density (OD) at 620 nm) of DMSO or uroC-spiked *E. asparagiformis* (*Ea*) and *E. bolteae* (*Eb*) cultures ($n = 4$ biological replicates). **c** Representative chromatograms ($\lambda = 305$ nm) of cultures sampled 2 h post-spike-in (from one representative biological replicate). The same scale was used for each chromatogram. IS, internal standard. **d** Quantification of urolithin concentrations 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. The treated 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 (**f**) *Ea* and (**g**) *Eb* based on DESeq2 (two-sided) analysis ($n = 4$ biological replicates for each bacterium). Data points are colored according

to their adjusted *P*-value (based on the Benjamini-Hochberg-corrected Wald statistic). The genomic organization around the differentially expressed genes is depicted above the Manhattan plots. Genes are colored according to their log₂fold-change (log₂FC) values. **h** Primer (forward (F) and reverse (R)) design for RT-PCR (**i**) experiment targeting the *Eb ucd* metabolic gene cluster (MGC). **i** 1% agarose gel image of RT-PCR amplicons using primers (**h**) that span the full-length (~3.6 kb) *Eb ucd* MGC (from one biological replicate). NTC, no template control. **j** RT-qPCR expression of each gene in the *Eb ucd* operon following treatment with DMSO or uroC (100 μM) for 2 h ($n = 3$ biological replicates). Gene expression profiles of each target gene in the *Eb ucd* gene cluster displayed as log₂FC with lines connecting paired biological replicates; repeated-measures one-way ANOVA with Tukey's multiple comparisons test. Exact *P*-values are noted above the data points. Data are represented as mean ± SEM (behind symbols) in (**b**, **d**, **e**). Source data and statistical details are provided as a Source data file.

RNA sequencing of uroC-induced cultures revealed a distinct gene cluster, which we term uroC dehydroxylase (*ucd*), that was highly and differentially expressed (log₂FC > 8) in both *E. asparagiformis* (Fig. 2f) and E. bolteae (Fig. 2g). In both bacteria, these clusters contained adjacent genes that were expressed to similar log₂FC values: a xanthine dehydrogenase family protein (subunit M), a (2Fe-2S)-binding protein, and a molybdopterin-dependent oxidoreductase (Fig. 2f, g). These genes will hereafter be referred to as *ucdC* (FAD-binding subunit), *ucdF* (2Fe2S-binding subunit), and *ucdO* (molybdopterin cytosine dinucleotide (MCD)-binding subunit), respectively. Interestingly, we also observed an upregulation of genes involved in efflux (MepA-like multidrug and toxin extrusion (MATE) transporters) and iron transport (FecCD-like) (Fig. 2f, g), suggesting a link between uroC metabolism, iron uptake, and efflux.

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

We next sought to characterize the *ucd* metabolic gene cluster in *E. bolteae* since this bacterium is considered a core species of the gut microbiome[8]. Based on the proximity, sense, and expression levels of each of the three genes by RNA-seq (Fig. 2g), we hypothesized that all three genes in the cluster were organized in an operon. We designed a gene-specific RT-PCR assay that would enable the detection of full-length polycistronic *ucdCFO* genes using cDNA from DMSO- or uroC-treated *E. bolteae* as a template (Fig. 2h). An amplicon of the expected size (~3.6 kb) was detected only in cDNA derived from uroC-treated *E. bolteae*, validating the inducibility of these genes (Fig. 2i). Long-read sequencing of the obtained amplicon (GenBank PQ855390.1) yielded a sequence corresponding to the *E. bolteae ucdCFO* metabolic gene cluster with 100% identity. Using an independent set of *E. bolteae* cultures, we then performed RT-qPCR on DMSO- or uroC-treated *E.*

Fig. 3 | Urolithin C metabolism correlates with *ucd* operon prevalence in gut bacteria. a Maximum likelihood phylogenetic tree (left) of *Enterocloster* spp. and catechol-metabolizing *Eggerthellaceae* based on the Genome Taxonomy Database (GTDB). Bacteria that convert uroC to uroA are labeled with gray squares. 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 (right). Domain annotations for each gene are denoted below according to InterPro annotations for corresponding proteins. ND, Not detected; Moco, Molybdenum cofactor. **b**, **c** In vitro metabolism of uroC by *Enterocloster pacaense* (*Ep*). 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 and then analyzed by LC-MS. **b** Representative chromatograms (λ = 305 nm) for each experimental group (from one representative biological replicate). The same scale was used for each chromatogram. **c** Quantification of urolithin peak areas relative to a salicylic acid internal standard (IS) (*n* = 3 biological replicates). Data are represented as mean ± SEM. Source data are provided as a Source data file.

bolteae with all three genes in the putative operon as targets. Similar to our RNA-seq results, all three genes were highly induced (mean $\log_2$FC ≥ 9.7 for all three *ucd* genes) relative to DMSO controls and were expressed at the same level (Fig. 2j). These results indicate that the *ucdCFO* genes are transcribed as a single polycistronic mRNA and therefore form a uroC-inducible operon.

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

Next, we aimed to determine the substrate scope of the *ucd* operon. Multiple urolithins possess pyrogallol, catechol, and phenol structural motifs that are dehydroxylated at various positions by gut bacteria (Supplementary Fig. 1). Interestingly, *E. bolteae* only metabolizes the 9-position hydroxyl group of urolithins (Supplementary Fig. 3a–d) and does not require adjacent hydroxyl groups since isouroA is dehydroxylated to uroB (Supplementary Fig. 3d). Since dehydroxylation of urolithins in *E. bolteae* is position specific, we hypothesized that the *ucd* operon would be induced by other 9-hydroxy urolithins (uroM6 and isouroA). Therefore, we performed RT-qPCR on DMSO-, uroM6-, uroC-, or isouroA-treated *E. bolteae* cultures using the *ucdO* gene as a target. Each urolithin significantly induced the expression of the *ucd* operon to a similar extent (Supplementary Fig. 3e, f). In addition, *E. bolteae* cell suspensions induced with uroC were capable of dehydroxylating uroM6 (Supplementary Fig. 3g) and isouroA (Supplementary Fig. 3h), indicating that the same proteins induced by uroC can metabolize structurally similar 9-hydroxy urolithins. Thus, it is likely that the same metabolic enzymes, encoded by the *E. bolteae ucd* operon, are acting on 9-hydroxy urolithins.

Presence of *ucd* homologs predicts uroC metabolism

We wondered whether novel metabolizers of uroC could be discovered based on *ucd* operon or individual Ucd protein sequence homology. Homology searches using the *E. bolteae ucd* operon sequence revealed additional hits: *Enterocloster pacaense* (many isolates), *Enterocloster lavalensis* (single metagenome-assembled

genome), and uncharacterized *Enterocloster* sp005845215 (two isolates) (Supplementary Fig. 4). Of the three hits, we characterized *E. pacaense*, also named *Lachnoclostridium pacaense* in the literature[38], as its type strain was available from culture collections. *E. pacaense* possessed genomic sequences with high homology (86.6% nucleotide identity) and identical functional annotations to the *E. bolteae ucd* operon sequence (Fig. 3a). When incubated with uroC, *E. pacaense* CCUG 71489 T quantitatively produced uroA (Fig. 3b, c). We searched for homologs of the *E. bolteae ucd* in the genomes of urolithin- and catechol-metabolizing bacteria belonging to the *Eggerthellaceae* but could not identify any hits. Notably, *Eggerthellaceae* lack 9-hydroxy urolithin dehydroxylase activity[33,35], which correlates with an absence of *ucd*-like operons in their genomes (Fig. 3a). At the individual protein level, Ucd proteins from *Enterocloster* spp. formed clades distant from related xanthine dehydrogenase family proteins from other taxa (Supplementary Fig. 5–7).

We next investigated the genomic context surrounding the *ucd* operon in all 4 confirmed uroC metabolizers to identify potential protein coding sequences involved in regulation or transport. In these bacteria, 5 protein-coding sequences with greater than 60% amino acid sequence identity were detected in all genomes (Supplementary Fig. 8a). These genes encode a small hypothetical protein, a putative membrane-associated hypothetical protein, two macrocin-O-methyltransferases (TylF), and a Crp/Fnr transcriptional regulator. We then searched for these genes in our RNA-seq datasets (Fig. 2f, g) and observed that all 5 genes were significantly upregulated upon uroC treatment (Supplementary Fig. 8b, c). While we cannot validate the function of these protein-coding sequences in their hosts, we hypothesize that the hypothetical proteins and Crp/Fnr transcriptional regulator, which can regulate molybdenum cofactor and xanthine dehydrogenase expression under anaerobic conditions[39], are involved in sensing uroC and upregulating *ucd* transcription. Overall, these comparative genomics data indicate that the presence of a *ucd* operon in genomes predicts uroC metabolism by gut bacteria.

Molybdopterin cofactor biosynthesis is upregulated by uroC

In addition to the three genes in the *ucd* operon, we observed a significant increase ($log_2FC \geq 2.6$) in 9 molybdopterin cytosine dinucleotide (MCD) biosynthesis genes upon uroC treatment (Fig. 2f, g and Supplementary Fig. 9a). These 9 genes, which recapitulate the function of 10 genes in *E. coli*, are involved in molybdenum cofactor biosynthesis (*moaAC*, *mogA*, *moeA*), molybdate ion transport (*modABCE*), cytosine addition to the molybdenum cofactor (*mocA*), and MCD cofactor insertion into the active site (*xdhC*) (Supplementary Fig. 9b)[39]. All 9 genes cluster in the genomes of *E. asparagiformis* and *E. bolteae* and are organized into 2 adjacent operons (Supplementary Fig. 9c) that are induced upon uroC treatment. These findings imply that uroC dehydroxylation is MCD-dependent, which differs from the predicted bis-molybdopterin guanine dinucleotide (bis-MGD) requirement of catechol dehydroxylases in *Eggerthellaceae*[10,14].

The UcdCFO complex enables anaerobic electron transport

Since oxidoreductases utilize a variety of cofactors and coenzymes for catalytic activity, we sought to determine the redox coenzymes and conditions necessary for uroC dehydroxylation. Therefore, we performed metabolism assays using crude lysates from uroC-induced *E. bolteae*. As these crude lysates alone did not metabolize uroC, various redox coenzymes (NADP+, NAD+, FAD, NADPH, NADH, and $FADH_2$) were added to lysates to test uroC dehydroxylation (Fig. 4a). We additionally included a reducing agent, sodium dithionite ($Na_2S_2O_4$), and a potential electron donor, sodium formate (HCO_2Na), as controls. NAD(P)H-treated lysates significantly dehydroxylated uroC to uroA compared to the no cofactor control. However, aerobic incubation of crude lysates supplemented with NAD(P)H completely inhibited uroC dehydroxylation (Fig. 4a), indicating that the active enzyme complex requires a strictly anaerobic environment for dehydroxylation.

To confirm that *ucd* operon-encoded proteins were expressed in *E. bolteae* crude lysates, we performed untargeted proteomics and compared protein expression upon DMSO or uroC treatment. All 3 proteins encoded by the *ucd* operon (UcdC, UcdF, and UcdO) were the most differentially expressed proteins in the uroC treatment group (Fig. 4b). In addition, proteins involved in MCD biosynthesis were strongly increased upon uroC treatment (Fig. 4b, c and Supplementary Fig. 9a), pointing to the coordination between MCD biosynthesis and active Ucd oxidoreductase assembly. These multi-omics datasets implicate all three *ucdCFO* and MCD biosynthesis genes in the metabolism of uroC to uroA, as demonstrated by the strong positive correlation between transcript and protein differential expression (Fig. 4c).

To validate the function of the *E. bolteae ucd* operon (Fig. 4d), we attempted heterologous expression of *E. bolteae* UcdCFO in *E. coli*; however, all expression and activity assays were unsuccessful despite the inclusion of *mocA* and *xdhC* genes involved in MCD maturation in our expression plasmids. This lack of activity likely resulted from the choice of heterologous host and from the complex assembly of active molybdoenzymes[40]. We therefore attempted to express UcdCFO in a phenol-degrading soil bacterium, *Rhodococcus erythropolis*, using a thiostrepton-inducible expression system[41] (pTipQC2-*ucdCFO*, Supplementary Fig. 10a, b), previously used to express the anaerobic *E. lenta* Cgr2 protein[42]. Despite the poor yield of soluble Ucd proteins (Supplementary Fig. 10c, d), we observed dehydroxylation of 9-hydroxy urolithins (uroM5, uroM6, uroD, uroC, and isouroA) in NADH-treated crude lysates of *R. erythropolis* transformed with pTipQC2-*ucdCFO*, but not in lysis buffer or no insert controls (pTipQC2) (Fig. 4e and Supplementary Fig. 11a–i).

Contrary to what was observed in *E. bolteae*, NADH supplementation was not strictly required in this heterologous system, though additional NADH significantly improved uroA production (Supplementary Fig. 11j). In an attempt to validate the NADH requirement of UcdCFO, we constructed a UcdO-His$_6$ fusion protein through

site-directed mutagenesis (Supplementary Fig. 12a, b). Although UcdO-His$_6$ and its associated proteins could be enriched by affinity chromatography (Supplementary Fig. 12c), we could not demonstrate a strict NADH requirement in this heterologous system as fractions enriched in UcdO (and associated proteins) were capable of uroC dehydroxylation in the absence of additional NADH (Supplementary Fig. 12d). These heterologous expression experiments thus confirm that the *ucd* operon confers 9-hydroxy urolithin dehydroxylase activity that can be enhanced through NADH supplementation.

To gain an understanding of the structural organization of proteins encoded by the *E. bolteae ucd* operon, we performed structural modeling using AlphaFold3 (AF3)[43,44]. By inputting a single copy of each Ucd protein, AF3 predicted a heterotrimeric UcdCFO enzyme complex (Fig. 4f) with a configuration similar to crystal structures of xanthine dehydrogenase/oxidase family proteins: 4-hydroxybenzoyl-CoA reductase from *Thauera aromatica* (*Ta* 4-HBCR)[45] (Fig. 4g), carbon monoxide dehydrogenase from *Afipia carboxidovorans* OM5 (*Ac* CODH)[46], and bovine milk xanthine dehydrogenase from *Bos taurus* (*Bt* XDH)[47] (Supplementary Fig. 13a–d). Structural superposition of the AF3 model onto *Ta* 4-HBCR, an anaerobic enzyme that dehydroxylates 4-hydroxybenzoyl-CoA to benzoyl-CoA, revealed that these enzymes have minor differences in the FAD-binding subunit (*Eb* UcdC and Ta 4-HBCRβ) as *Ta* 4-HBCRβ has an additional loop that binds a 4Fe-4S cluster (Supplementary Fig. 13e)[45]. However, iron-sulfur- (*Eb* UcdF and *Ta* 4-HBCRγ) and MCD- (*Eb* UcdO and *Ta* 4-HBCRα) binding subunits shared conserved 2Fe-2S and molybdopterin cofactor binding domains, respectively. The predicted UcdCFO enzyme complex with superposed ligands from *Ta* 4-HBCR supported a complete electron transport chain whereby electrons would flow from reduced FAD to two 2Fe-2S clusters, then to the MCD cofactor, and finally to uroC as the terminal electron acceptor (Fig. 4h and Supplementary Fig. 13f, g). Using homology modeling, we further identified the putative uroC binding site in UcdO, which overlaps with the salicylic acid (inhibitor) binding site in the *Bos taurus* xanthine dehydrogenase structure (Supplementary Fig. 13h). The putative uroC binding site in *Eb* UcdO was densely populated with aromatic residues like tyrosine (Y375, Y536, Y624, Y632), tryptophan (W345), and phenylalanine (F458, F464), which could form π-π stacking interactions with uroC, orienting it toward the molybdenum cofactor (Supplementary Fig. 13i). In addition, a conserved glutamate residue (E741), essential for catalytic activity in both mammalian and bacterial xanthine dehydrogenase family proteins[48,49], was placed near the MCD-binding site (Supplementary Fig. 13h–j). Altogether, the predicted AF3 UcdCFO complex shares structural features with xanthine dehydrogenase family proteins, likely following a similar electron transfer mechanism to *Ta* 4-HBCRβ.

Disruption of the uroC catechol rescues lag time extension

To further investigate the physiological role of urolithin dehydroxylation in *Enterocloster* spp., we performed growth experiments in different media conditions. We observed a concentration-dependent increase in lag time for all uroC-metabolizing bacteria when uroC was added prior to growth in rich medium containing hemin (mABB + H) (Fig. 5a, b). As catechols are common structural motifs in iron-binding siderophores[50], we hypothesized that uroC could be altering iron availability (or vice versa) in the growth medium via its catechol moiety. Incubation of *Enterocloster* spp. in a medium lacking added iron (mABB) exacerbated the growth delay by uroC (Fig. 5c and Supplementary Fig. 14a); however, this growth delay was partially rescued upon supplementation of different iron sources (hemin, Fe(II)SO$_4$, or Fe(III) pyrophosphate) (Fig. 5c and Supplementary Fig. 14b, c). To validate that iron chelation could extend the lag time of *Enterocloster* spp., 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)
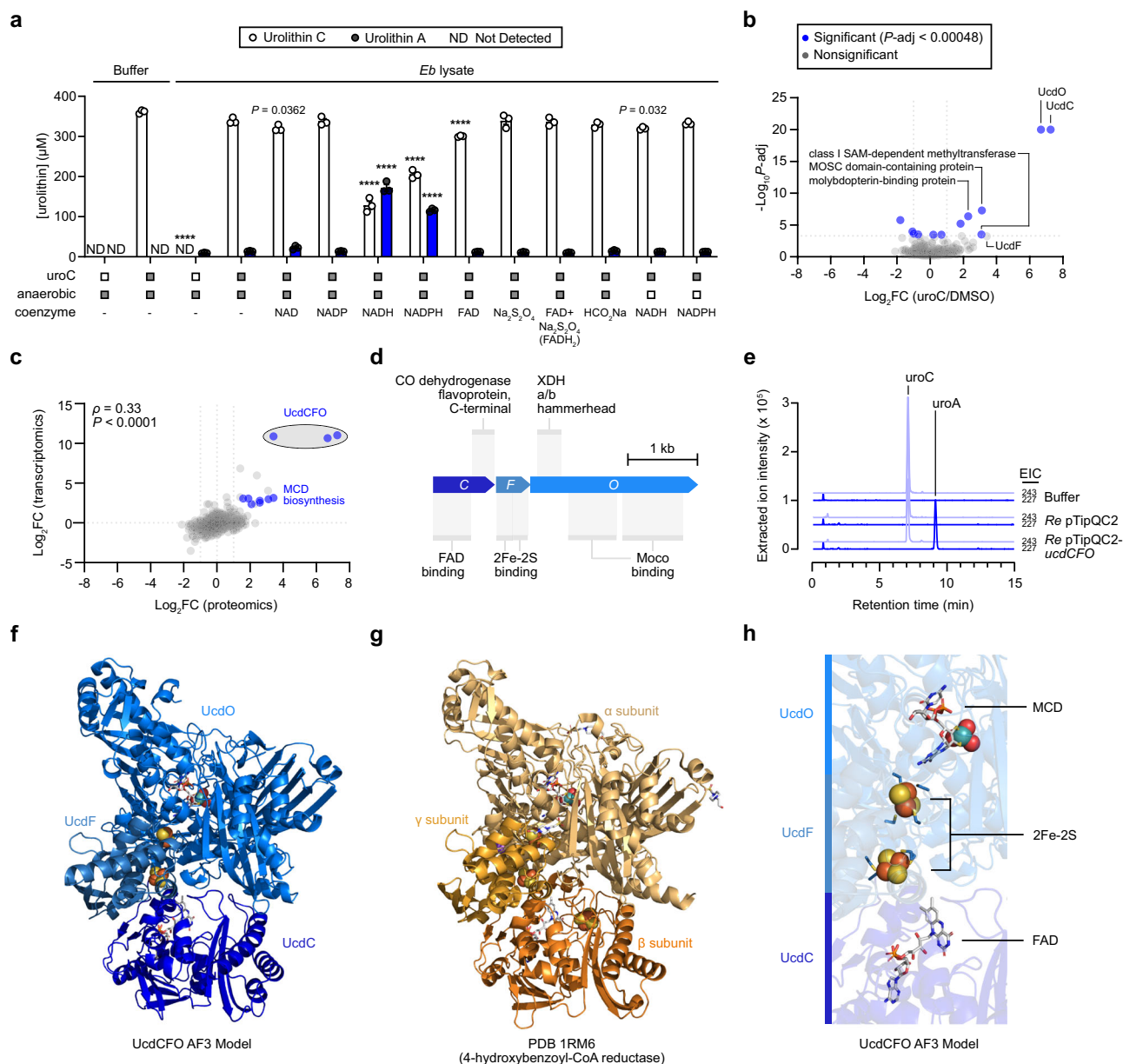
Fig. 4 | The UcdCFO complex enables anaerobic electron transport from NADH to uroC. a Quantification of urolithin concentrations in crude uroC-induced *Eb* lysates (*n* = 3 biological replicates) re-treated with DMSO or uroC (350 µM) and various coenzymes at room temperature under anaerobic or aerobic conditions for 20 h. Treatments are indicated by gray squares. NAD(P), nicotinamide adenine dinucleotide (phosphate); FAD, flavin adenine dinucleotide; Na$_2$S$_2$O$_4$, sodium dithionite (reducing agent); HCO$_2$Na, sodium formate (possible electron donor). Data are represented as mean ± SEM; two-way ANOVA with Dunnett's multiple comparisons test against the *Eb* lysate incubated with uroC anaerobically (no coenzyme). Statistical significance is denoted in the plot with exact *P*-values when 0.05 > *P* ≥ 0.0001 (****). **b** 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 (two-sided) with Benjamini-Hochberg correction for multiple comparisons). Gray, *P*-adj ≥ cutoff *P*-value (0.00048). Blue, *P*-adj < cutoff *P*-value (0.00048). **c** Scatter plot showing the correlation between gene and protein expression (log$_2$FC values) induced in uroC-treated *Eb* using the datasets in Figs. 2g and 4b, respectively. The non-parametric

Spearman rank correlation test (two-sided) was used for statistical analysis, approximate *P* < 0.0001. **d** Functional domains in the *E. bolteae ucd* operon based on InterPro annotations. **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. Crude lysates were supplemented with 2 mM NADH and incubated anaerobically for 72 h at room temperature. **f** Quaternary structure prediction of the proteins encoded by the *Eb ucd* operon using AlphaFold3 (AF3). Small molecule ligands from the X-ray crystal structure of PDB 1RM6 (4-hydroxybenzoyl-CoA reductase from *Thauera aromatica*) were placed into UcdCFO through structural superposition. Subunits are colored in shades of blue. **g** X-ray crystal structure of PDB 1RM6 (4-hydroxybenzoyl-CoA reductase from *Thauera aromatica*). A single αβγ unit (colored in shades of orange) of the (αβγ)$_2$ heterohexamer is shown. **h** Small molecule ligands from PDB 1RM6 in the superposed UcdCFO model. Source data and statistical details are provided as a Source data file.

pyrophosphate could partially rescue growth delay (Supplementary Fig. 14c). Interestingly, uroA, which lacks a catechol moiety, did not impact the growth of the tested bacteria in either mABB or mABB + H media compared to uroC (Fig. 5c and Supplementary Fig. 14a, b). We

therefore screened a panel of catechol- and/or pyrogallol-containing urolithins (uroM5, uroM6, uroD, and uroC) along with some of their dehydroxylated products (uroM7, uroA, isouroA, uroB) (Fig. 5d) to test whether these structural features impacted lag time. Overall, catechol-
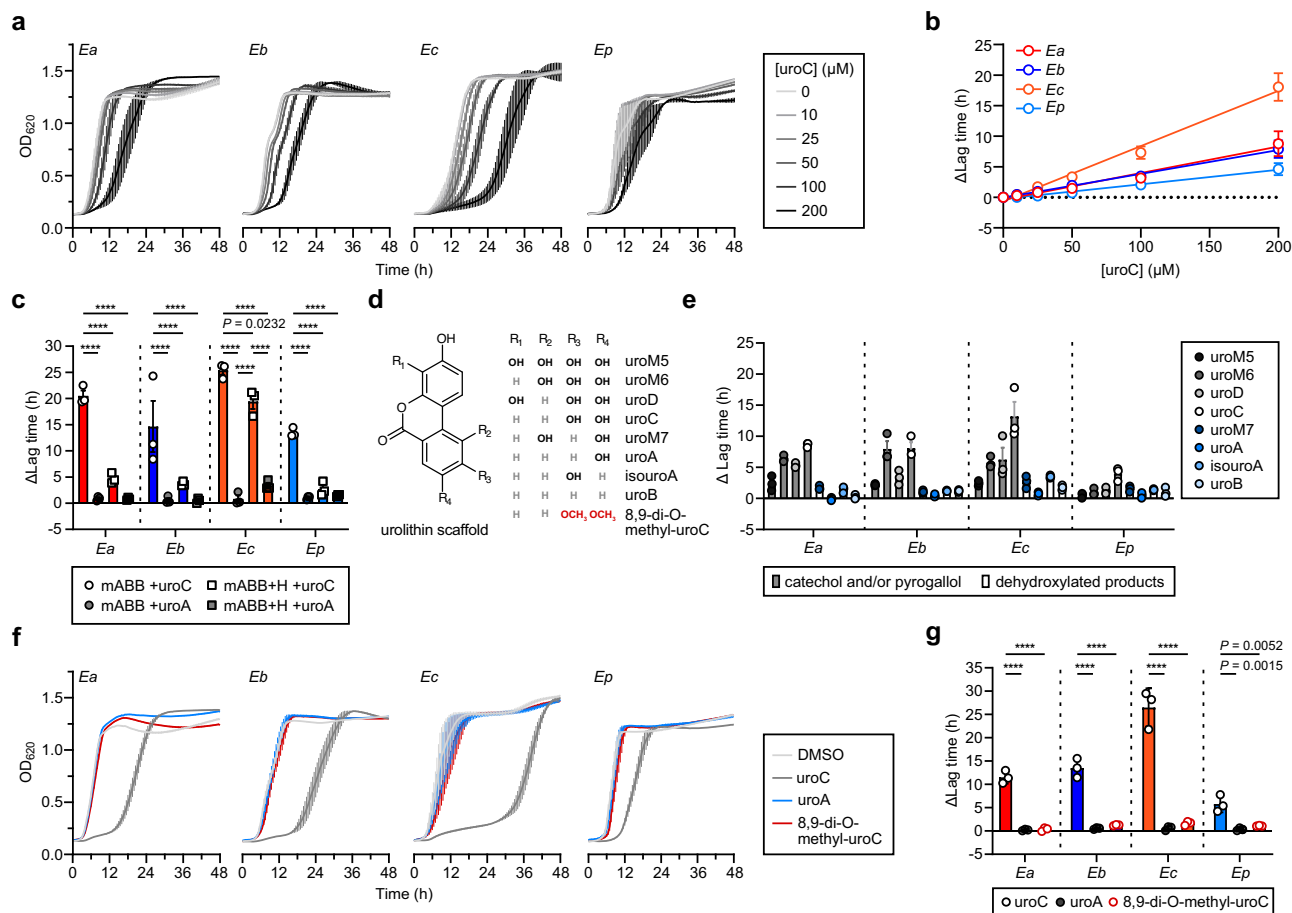
Fig. 5 | The catechol moiety of uroC delays *Enterocloster* spp. growth. a Growth curves (optical density (OD) at 620 nm) of uroC-metabolizing *Enterocloster* spp. treated with increasing concentrations of uroC in rich mABB + H media (7.7 μM hemin) (*n* = 3 biological replicates). Data are represented as mean ± SEM. **b** Quantification of the difference in lag time compared to the DMSO control for growth curves in (**a**). Data are represented as mean ± SEM; lines were fitted using simple linear regression. **c** Quantification of the difference in lag time of *Enterocloster* spp. grown in mABB (no added iron) or mABB + H (7.7 μM hemin) compared to respective DMSO controls for growth curves in Supplementary Fig. 14a, b (*n* = 3 biological replicates). Data are represented as mean ± SEM; repeated measures two-way ANOVA (matching by biological replicate) with Tukey's multiple comparisons test. Statistical significance is denoted in the plot with exact *P*-values when 0.05 > *P* ≥ 0.0001 (****). **d** Structures of tested urolithins. **e** Quantification of the difference in lag time of *Enterocloster* spp. grown in mABB (no added iron)

treated with 100 μM urolithins in (**d**) compared to respective DMSO controls (*n* = 3 biological replicates). Data are represented as mean ± SEM; repeated measures two-way ANOVA (matching by biological replicate) with Tukey's multiple comparisons test (main treatment effect). *P*-values for comparisons are provided in the Source data file. **f** Growth curves (OD at 620 nm) of uroC-metabolizing *Enterocloster* spp. grown in mABB media and treated with DMSO (vehicle), 100 μM of uroC, uroA, or 8,9-di-O-methyl-uroC (*n* = 3 biological replicates). Data are represented as mean ± SEM. **g** Quantification of the difference in lag time compared to the DMSO control for growth curves in (**f**). Data are represented as mean ± SEM; repeated measures two-way ANOVA (matching by biological replicate) with Tukey's multiple comparisons test. Statistical significance is denoted in the plot with exact *P*-values when 0.05 > *P* ≥ 0.0001 (****). Source data and statistical details are provided as a Source data file.

and/or pyrogallol-containing urolithins delayed growth more than their dehydroxylated counterparts (Fig. 5e). This effect was most apparent when comparing Ucd substrates and products (uroM6 to uroM7 and uroC to uroA). Curiously, uroM5 (catechol and pyrogallol moieties) delayed growth less than uroC (catechol alone), suggesting that compound polarity and other physicochemical properties may influence lag time extension. To further confirm that the catechol moiety of uroC was responsible for delaying growth, we synthesized a methylated analog of uroC (8,9-di-O-methyl-uroC, Fig. 5d and Supplementary Fig. 15a–f), and tested its effect on growth in mABB medium. Like uroA, 8,9-di-O-methyl-uroC did not delay the growth of uroC-metabolizing bacteria (Fig. 5f, g). These data demonstrate that both the catechol moiety of uroC and iron availability are determinants of uroC-mediated lag phase extension in *Enterocloster* spp.

Dehydroxylation of catechols by gut bacteria has been observed for diverse classes of compounds like neurotransmitters, therapeutic drugs, and diet-derived polyphenols[14]. Although catechol

dehydroxylation can provide a growth advantage in some species[14], we hypothesized that dehydroxylation could be a mechanism used by gut bacteria to inactivate catechol-containing compounds that affect their fitness. To determine whether diverse catechols can delay growth, uroC-metabolizing bacteria and the dopamine-metabolizing *Eggerthella lenta* A2 were incubated with catechol-containing compounds and their dehydroxylated counterparts: uroC (uroA), entacapone, dopamine (m-tyramine), caffeic acid (m-coumaric acid) (Supplementary Fig. 16a). Surprisingly, neither dopamine nor caffeic acid (and their dehydroxylated counterparts) delayed the growth of the tested bacteria (Supplementary Fig. 16b). On the other hand, both uroC and the nitrocatechol-containing Parkinson's drug entacapone delayed the growth of *Enterocloster* spp. but did not affect *E. lenta* A2 (Supplementary Fig. 16b)[8]. Thus, catechol-containing compounds show differential effects on the growth of gut bacteria, depending on their structure. These results prompted us to investigate the effect of uroC on a more diverse panel of gut bacteria including *E. aldenensis, E.*
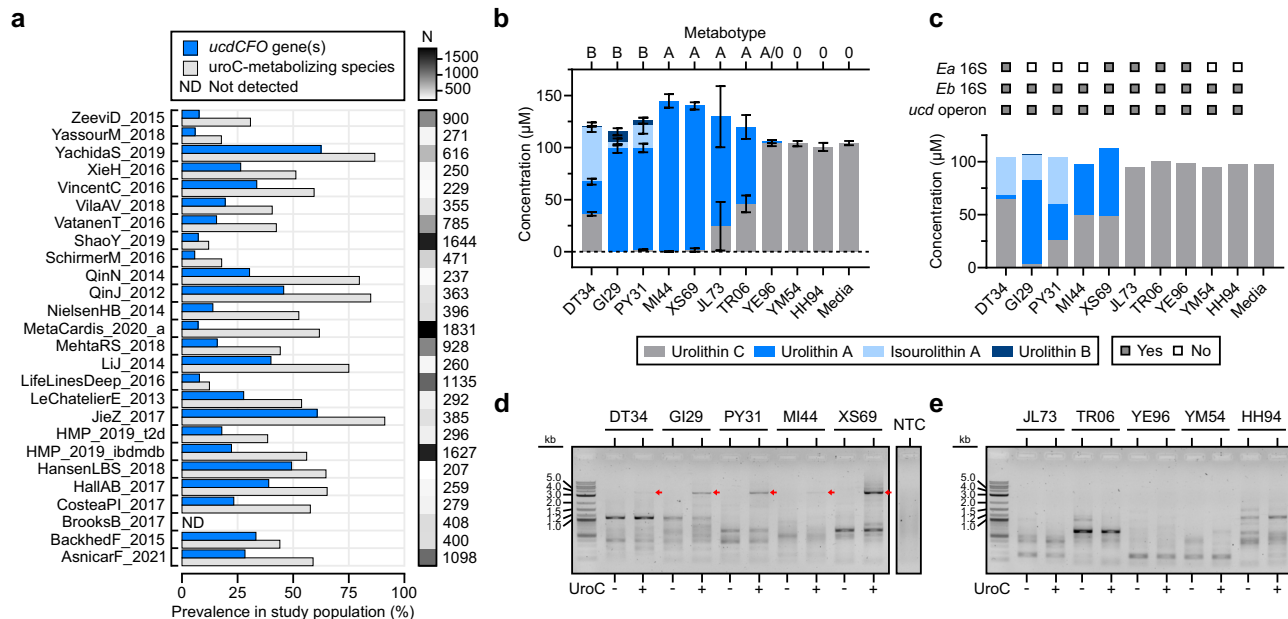
Fig. 6 | The *ucd* operon is prevalent in metagenomes and actively transcribed in urolithin C-metabolizing human fecal samples. a Prevalence of *ucd* operon (at least one gene) and of a uroC-metabolizing species (at least one species) in fecal metagenomes from the CuratedMetagenomicData R package. Only studies with ≥ 200 participants are depicted. All 86 studies are available in Supplementary Fig. 17. **b** Summary of urolithin concentrations in human fecal slurries (*n* = 10 healthy donors) incubated with 100 μM uroC for 48 h. Data are represented as mean ± SEM (*n* = 3 experimental replicates). Participants (5 female, 5 male) were aged 21–43 years with BMI between 18.1–26.3. **c** Summary of urolithin concentrations in fecal slurries (*n* = 10 healthy donors) incubated with 100 μM uroC for 48 h. The presence of uroC-metabolizing species and of the *ucd* operon is denoted above the graph if the bacterium or operon was detected in the uroC-treated fecal slurry

(Supplementary Fig. 18b, f, g). Data are representative of 1 replicate for each donor where DNA and RNA were also extracted from fecal slurries. **d**, **e** *ucd* gene-specific RT-PCR on fecal microbiota communities from 10 healthy donors using the primer set described in Fig. 2h. Samples are matched to the urolithin metabolism data in (**c**) and 16S rRNA sequencing in Supplementary Fig. 18a–e (all from 1 replicate for each donor). 1% agarose gel of amplicons derived from (**d**) uroC-metabolizing and (**e**) non-metabolizing fecal slurries. Bands corresponding to the *E. bolteae ucd* operon (~3.6 kb) are labeled with red arrows. The no template control (NTC) is the same for (**d**, **e**). See Supplementary Fig. 18h, i for the no reverse transcriptase control PCR reactions on the same samples. Source data are provided as a Source data file.

clostridioformis, and *E. lavalensis*, which do not metabolize uroC, along with *Gordonibacter* spp., which produce uroC from dietary ellagic acid[33]. Treatment with uroC delayed the growth of *E. aldenensis*, *E. clostridioformis*, and *E. lavalensis* to varying extents (Supplementary Fig. 16c); however, there was no difference in growth between the DMSO-, uroC-, and uroA-treated cultures of *Enterococcus faecium* and *Gordonibacter* spp. (Supplementary Fig. 16d, e). Thus, *Enterocloster* spp., regardless of their ability to metabolize uroC, are sensitive to uroC-mediated lag time extension, while other bacteria are insensitive to its effects on growth.

Enterocloster spp. and *ucd* are prevalent in human feces

We next wondered whether uroC-metabolizing *Enterocloster* spp. and their *ucd* operons were prevalent and active in human fecal samples. We first utilized uniformly processed metagenomic data from the curatedMetagenomicData R package[51]. After filtering for fecal samples (86 studies, *n* = 21,030 subjects), we counted the prevalence of at least one uroC-metabolizing species and at least one *ucd* gene homolog (Methods). The prevalence of both features was variable across studies (Fig. 6a for studies with > 200 participants, Supplementary Fig. 17a for all studies). Combining all studies, the prevalence of at least one uroC-metabolizing species and at least one *ucdCFO* gene homolog was 9,343/21,030 (44.9%) and 4,356/21,030 (20.7%), respectively. *E. bolteae* was the most prevalent and abundant uroC-metabolizing species detected in gut metagenomes (Supplementary Fig. 17b, c) and correlated strongly with overall *ucd* abundance (Supplementary Fig. 17d). These findings suggest that uroC-metabolizing *Enterocloster* spp. and *ucd* operon genes are prevalent in human fecal metagenomic samples

and reflect the variable urolithin metabolism profiles (metabotypes) in the general population[25,52].

Next, we performed ex vivo metabolism assays to determine whether *Enterocloster* spp. could metabolize uroC in the context of a complex community. Fecal slurries from 10 healthy individuals were first profiled according to their uroC metabotypes (Fig. 6b)[24]. Individuals clustered into metabotypes A (only uroA produced), B (uroA and isouroA/uroB), and 0 (no terminal urolithin metabolites). Stools JL73, TR06, and YE96 displayed variable metabolism patterns and did not metabolize uroC in some experiments, likely reflecting differences in activity between aliquots of feces (Fig. 6b). We then repeated metabolism assays using fecal slurries from all 10 healthy individuals and extracted urolithins, DNA, and RNA from each culture. In this experiment, only 5/10 fecal slurries metabolized uroC to uroA (Fig. 6c). We hypothesized that differences in metabotypes could be explained by microbial composition. Therefore, long-read V1-V9 16S rRNA sequencing was performed on fecal slurries. Both DMSO- and uroC-treated fecal slurries within individuals had similar microbial compositions and diversity metrics (Supplementary Fig. 18a–d) but showed differences in composition between individuals and metabolism status (Supplementary Fig. 18a, e). Surprisingly, all samples contained 16S rRNA sequences mapping to *E. bolteae*, and many non-metabolizing fecal slurries contained *E. asparagiformis* (Fig. 6c and Supplementary Fig. 18b). 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. 18f, g). These data indicate that ex vivo uroC metabolism in fecal samples cannot be reliably predicted

based solely on the presence of uroC-metabolizing *Enterocloster* spp. and *ucd* operon genes.

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. 18h, i). These data demonstrate that *ucd* transcription correlates with uroC metabolism in complex fecal communities.

Discussion

We identified genes and proteins that are sufficient for the metabolism of urolithins by gut resident *Enterocloster* spp. through a combination of transcriptomics, comparative genomics, and untargeted proteomics. Our study reveals a novel multi-subunit molybdoenzyme (urolithin C dehydroxylase, Ucd) that specifically catalyzes the dehydroxylation of 9-hydroxy urolithins. By screening diverse urolithins, we observed that dehydroxylation by Ucd was restricted to the 9-position regardless of the number or position of other hydroxyl groups surrounding it, suggesting that the Ucd active site can specifically orient the 9-position hydroxyl group towards the MCD cofactor. This selectivity is likely mediated though polar contacts with the 3-position hydroxyl group, which remains intact throughout the urolithin metabolic pathway.

Catechol dehydroxylases are widespread in gut resident *Eggerthella lenta* and *Gordonibacter* spp[10,14,53]. These molybdoenzymes, which belong to the DMSO reductase superfamily[54], dehydroxylate substrates like catechol lignan (Cldh), dopamine (Dadh), DOPAC (Dodh), hydrocaffeic acid (Hcdh), and caffeic acid (Cadh). *E. lenta* dehydroxylases are predicted to form a complex composed of a bis-MGD binding enzyme (substrate binding), a 4Fe-4S cluster-containing protein, and a membrane anchor protein. Indeed, most catechol dehydroxylases from *E. lenta* are predicted to be exported from the cytoplasm (via the twin-arginine protein transport pathway) and anchored to the membrane. In contrast, *Gordonibacter* spp. catechol dehydroxylases, composed of a bis-MGD binding enzyme (substrate binding) and a 4Fe-4S cluster-containing protein, are likely restricted to the cytosol, requiring transporters to access their substrates[14]. Despite catalyzing dehydroxylation reactions, dehydroxylases from the xanthine dehydrogenase family, like Ucd and *Thauera aromatic* 4-HBCR[45], differ significantly from DMSO reductase family dehydroxylases in their subunit composition, cofactors, and coenzymes. Xanthine dehydrogenase family enzymes are often found as heterohexamers in an $(\alpha\beta\gamma)_2$ configuration[48]. In each $\alpha\beta\gamma$ heterotrimer, a complete electron transport chain is formed between the FAD-binding β subunit, the 2Fe-2S cluster-containing γ subunit, and the MCD-binding α subunit, enabling substrate reduction. Based on our metabolism studies in *E. bolteae* and *R. erythropolis*, we propose that NAD(P)H can serve as electron donors for urolithin dehydroxylation; however, alternative low-reduction potential electron donors, like ferredoxins, may also be suitable[45].

By identifying the genetic basis for uroC metabolism by *Enterocloster* spp., we found a novel 9-hydroxy urolithin-metabolizing species: *E. pacaense*. Our data suggests that *ucd*-containing *Enterocloster* spp. are some of the main drivers of urolithin A production in the gut microbiota based on their prevalence in metagenomes and transcriptional activity in uroC-metabolizing fecal samples. However, we cannot conclude that *Enterocloster* spp. are solely responsible for this activity. Rare, strain-specific urolithin A production has been reported for *Bifidobacterium pseudocatenulatum* INIA P815[34], *Streptococcus thermophilus* FUA329[55], and *Enterococcus faecium* FUA027[55], which may be

a result of horizontal gene transfer since dehydroxylase activity is not shared by other members of these taxa. In addition, enzymes acting on the same substrate could arise in phylogenetically distant organisms through convergent evolution. This is relevant in the context of urolithin dehydroxylases as *Ellagibacter isourolithinifaciens* can metabolize uroC, albeit at the 8-position hydroxyl group. Thus, further enzyme discovery efforts are necessary to understand urolithin production in these bacteria.

Although multiple urolithin-metabolizing gut bacteria have been identified, the physiological role of urolithin dehydroxylation has remained elusive. In this study, uroC, and related catechol and pyrogallol-containing urolithins, delayed the growth of *Enterocloster* spp. when added prior to growth (lag phase). This growth-delaying activity was most pronounced in media lacking added iron but could be partially rescued through iron supplementation. These findings may suggest that uroC is sequestering iron under low iron conditions through its catechol, leading to a decrease in the concentration of this micronutrient, similar to a recent study on entacapone[56]. However, alternative explanations like oxidative stress though Fenton chemistry[57] or complexation of toxic uroC with increasing iron concentrations could explain the observed results. We validated that the growth-delaying activity of uroC was dependent on its catechol moiety since dehydroxylation to uroA and methylation of the uroC catechol inhibited lag time extension. Other taxa like *E. lenta*, *Gordonibacter* spp., and *E. faecium* were unaffected by uroC, suggesting that *Enterocloster* spp., regardless of Ucd expression, are sensitive to catechol-containing urolithins, though not to all catechols. Thus, we propose that urolithin dehydroxylation by *Enterocloster* spp. may indirectly limit the toxicity of urolithins by removing their catechol structural motifs. In *Eggerthellaceae*, dehydroxylation of catechols can promote bacterial growth and ATP generation through their use as alternative electron acceptors[14,58,59]. Similar mechanisms may apply to *Enterocloster* spp. as uroC is a terminal electron acceptor, though further mechanistic work is needed to support this hypothesis.

While uroA is the most common terminal metabolite following ellagitannin consumption in humans, its production varies widely[34]. In our hands, the ability of a fecal sample to produce uroA from uroC ex vivo did not correlate with the presence nor abundance of widespread uroC-metabolizing *Enterocloster* spp. or a *ucd* operon homolog. This observation could be explained by poor viability, metabolic inactivity, abundance, or transcriptional repression of *Enterocloster* spp. in fecal samples. However, active transcription of the *ucd* operon correlated perfectly with metabotypes. These findings further emphasize the importance of functional assays such as transcriptomics and ex vivo metabolism to understand the metabolism of xenobiotics by the gut microbiota.

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

Methods

Anaerobic bacterial strains and culturing conditions

Bacterial strains and other materials used in this study are listed in Supplementary Table 1. All bacterial stains were validated by sequencing the 16S rRNA gene. The same culture used for validation was used to make 25% glycerol stocks. Anaerobic strains were grown from glycerol stocks on mABB + H or BHIS (ATCC Medium 1293) agar plates (recipe below) for 48–72 h at 37 °C in a vinyl anaerobic chamber, which was maintained with a gas mixture of 3% H_2, 10% CO_2, 87% N_2. To make overnight cultures, a single colony was inoculated into 5 mL of liquid mABB or mABB + H medium and incubated at 37 °C between 16–48 h,

depending on the bacterium (16–24 h for *Enterocloster* spp. and *E. faecium*, and 48 h for *E. lenta* and *Gordonibacter* spp.).

Human fecal sample collection

Human fecal samples were collected under the approval of protocol A04-M27-15B by the McGill Faculty of Medicine Institutional Review Board. Informed written consent was received from the participants for the use of human samples. Participants were compensated with 10 CAD. Eligibility criteria for the healthy participants were as follows: no diagnosed gastrointestinal disease, no ongoing therapeutic treatment, and no antibiotic usage 3 months prior to the start of the study. Subject information was recorded at the time of sampling and is provided in Supplementary Table 2. Participants (5 female, 5 male) were between 21–43 years old with BMI between 18.1-26.3. Fresh fecal samples were collected between 2017-2019 and placed immediately in an anaerobic chamber, aliquoted, and stored at − 70 °C until use.

Modified anaerobe basal broth (mABB and mABB + H)

For 1 L of modified anaerobe basal broth (mABB), the following components were dissolved in MilliQ water, then autoclaved: 16 g peptone, 7 g yeast extract, 5 g sodium chloride, 1 g starch, 1 g D-glucose monohydrate, 1 g sodium pyruvate, 0.5 g sodium succinate, 1 g sodium thioglycolate, 15 g agar (for plates). The autoclaved solution was allowed to cool, then the following filter-sterilized solutions were added aseptically: 10 mL of 100 mg/mL L-arginine-HCl, 10 mL of 50 mg/mL L-cysteine, 8 mL of 50 mg/mL sodium bicarbonate, 50 µL of 10 mg/mL vitamin K1, 20 mL of 50 mg/mL dithiothreitol, and, for mABB + H, 10 mL of 0.5 mg/mL haemin. The media was then placed in the anaerobic chamber and allowed to reduce for at least 24 h prior to its use in experiments.

Genomic DNA extraction of isolates and identity validation

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

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

Treatments with urolithins and other catechols

Treatments used in this study (urolithin M5, urolithin M6, urolithin M7, urolithin D, urolithin C, urolithin A, isourolithin A, urolithin B, 8,9-di-O-methyl-urolithin C, dopamine, m-tyramine, caffeic acid, m-coumaric acid, entacapone, and 2,2′-bipyridyl) were dissolved in DMSO to a concentration of 10 mM.

Treatment prior to growth (metabolism only)

Overnight cultures of bacteria were diluted 1/50 into fresh media. Treatments (10 mM stocks solutions, dissolved in DMSO) were added to the diluted bacterial suspension to a final concentration of 100 µM, and samples were incubated for 24 h at 37 °C in an anaerobic chamber.

Treatment during growth (spike-in)

Overnight cultures of *Enterocloster* spp. were diluted 1/50 into fresh media and incubated at 37 °C in an anaerobic chamber. After 5 h of incubation (~ OD$_{620}$ from a 200 µL sample ~ 0.4), 10 mM urolithins (or an equivalent volume of DMSO) were added to the growing cultures at a final concentration of 50 or 100 µM for protein expression or RNA expression, respectively. For protein expression analyses and inducibility tests, the cultures were incubated for an additional 4 h. For RNA expression analyses, the cultures were incubated for an additional 2 h.

Treatment prior to growth (growth curves)

Overnight cultures of *Enterocloster* spp. were diluted 1/25 into mABB (with or without added 15.4 µM iron source), depending on the experimental design. Separately, treatments (10 mM stocks solutions, dissolved in DMSO) were prepared in mABB (with or without added 15.4 µM iron source) to a final concentration of 200 µM. In each well of a 96-well plate, 100 µL of 1/25 bacteria and 200 µM treatment were combined. These were plated in technical duplicates. The final concentration of treatment was 100 µM (unless otherwise specified in concentration-response experiments), and the final dilution of bacteria was 1/50.

Urolithin extraction from bacterial cultures

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

Extraction method A. This method was used for cultures. Salicylic acid (3 mg/mL in DMSO) was spiked-in as an internal standard at a final concentration of 50 µg/mL. The cultures and standards were then extracted with 3 volumes of ethyl acetate + 1% formic acid (e.g., 600 µL solvent to 200 µL thawed culture). The organic phase (top) was transferred to a new tube and dried in a vacuum concentrator (Heto Lab) connected to a rotary evaporator (Buchi). After solvent removal, samples were redissolved in 0.5 volumes (relative to the starting culture) of 50% MeOH:H$_2$O. Samples were centrifuged at 20,000 × g for 5 min to pellet insoluble material, then transferred to LC-MS vials. Urolithins were then analyzed by LC-MS.

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

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

LC-MS method to quantify urolithins

Samples (10 µL) were injected into a 1260 Infinity II Single Quadrupole LC/MS system (Agilent) fitted with a Poroshell 120 EC-C18 4.6 × 50 mm, 2.7 µm column (Agilent) incubated at 30 °C. The mobile phase was composed of MilliQ water + 0.1% formic acid (solvent A) and acetonitrile + 0.1% formic acid (solvent B). The flow rate was set to 0.7 mL/min. The gradient was as follows: 0–8 min: 10-30 %B, 8–10 min: 30-100 %B, 10–13.5 min: 100 %B isocratic, 13.5-13.6 min: 100-10 %B, then 13.6–15.5 min: 10 %B. The multiple wavelength detector was set to monitor absorbance at 305 nm. The mass spectrometer (API-ES) was run in the negative mode in both selected ion monitoring (SIM) and scan (100–1000 m/z) modes to validate peak identities. The capillary voltage

was set to 3000 V, drying gas to 10.0 L/min, nebulizer pressure to 30 psig, and drying gas temperature to 350 °C. Peaks were validated based on retention times compared to spike-in standards and mass-to-charge ratios. To quantify urolithins, integrated peak areas (derived from Agilent ChemStation software) for the compounds of interest were compared to spike-in standards of known concentrations. When standards were not available or overlapped with other peaks, the extracted ion chromatogram peak area was used: urolithin M5 ([M-H]: 275), urolithin E ([M-H]: 259), urolithin D ([M-H]: 259), urolithin M7 ([M-H]: 243), urolithin M6 ([M-H]: 259), urolithin C ([M-H]: 243), urolithin G ([M-H]: 243), urolithin A ([M-H]: 227), isourolithin A ([M-H]: 227), urolithin B ([M-H]: 211). Blank runs of 50% MeOH were included at the beginning and end of LC-MS sequences to ensure proper column washing.

Cell suspension assay to test inducibility

E. asparagiformis and *E. bolteae* (10 mL growing cultures in mABB + H media) were grown with 50 µM uroC (or an equivalent volume of DMSO) as detailed in the *Treatment during growth (spike-in)* section above and incubated for 4 h at 37 °C. Cultures were then pelleted at 6500 × g for 3 min, and the supernatants were discarded. The cells were washed with 10 mL of pre-reduced PBS (placed in the anaerobic chamber 24 h before), re-pelleted, and resuspended in 2 mL of pre-reduced PBS. For each condition tested, a 200 µL aliquot of the cell suspension was transferred into a sterile 1.5 mL tube, and 10 mM urolithins (uroM6, uroC, isouroA, or DMSO) were added at a final concentration of 100 µM. Cell suspensions were briefly vortexed and incubated at room temperature in the anaerobic chamber for 16 h prior to freezing and urolithin extraction using *Extraction Method B*.

RNA extraction from isolates

A volume of 1.5 mL of treated (100 µM urolithin C for 2 h) *Enterocloster* spp. culture (see *Enterocloster* spp. urolithin C treatments) was pelleted (6500 g for 3 min), and the supernatant was removed for later LC-MS analysis. The pellet (suspended in 200 µL of media) was then mixed with 800 µL TRI reagent (Zymo Research) and transferred to a ZR BashingBead lysis tube (Zymo Research). Samples were lysed in a Mini Beadbeater 16 (Biospec) according to the following sequence: 1 min ON, 5 min OFF. For RNA sequencing, bead beating was done for a total of 5 min ON. For RT-(q)PCR, bead beating was done for a total of 2 min ON to preserve longer transcripts. RNA isolation was then performed using the Direct-zol RNA Miniprep Kit (Zymo Research) according to the manufacturer's instructions (including an on-column DNase digestion). To ensure complete DNA removal, an additional DNA digestion step was performed on the isolated RNA using the Ambion *DNA-free* DNA Removal Kit (Invitrogen) according to the manufacturer's instructions. The DNA-free RNA was then cleaned up using the *OneStep* PCR Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were initially verified by NanoDrop and 1 % agarose gel electrophoresis. For RNA sequencing, RNA integrity was assessed by Génome Québec using a Bioanalyzer 2100 (Agilent). RNA integrity (RIN) values ranged between 7.5-7.8 for *E. asparagiformis* DSM 15981 and 7.0-7.3 for *E. bolteae* DSM 15670.

RNA-seq of uroC-treated *Enterocloster* spp

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

Analysis of RNA-seq reads was done using Galaxy bioinformatics cloud computing (https://usegalaxy.org/) hosted by Compute Canada

Genetics and Genomics Analysis Platform (GenAP) (https://www.genap.ca/). Genomes and annotations were fetched from the NCBI genome browser: *E. bolteae* (ASM223457v2) (accessed 2022/05/11) and *E. asparagiformis* (ASM2514912v1) (accessed 2023/09/19). Raw reads were first verified for quality using FastQC (v0.73, https://github.com/s-andrews/FastQC) with default parameters. FastQC reports were aggregated into MultiQC (v1.11, https://github.com/MultiQC/MultiQC)[60]. The mean sequence quality scored were above 35 for all samples. Raw reads were then trimmed using Cutadapt (v3.7, https://github.com/marcelm/cutadapt)[61] to trim adapter sequences (R1 sequence: AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC, and R2 sequence: AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT) that were not removed after sequencing using default parameters for paired-end reads. Trimmed reads were then aligned to reference genomes for each bacterium using HISAT2 (v2.2.1, https://github.com/DaehwanKimLab/hisat2)[62] with paired-end parameters and reverse strandedness (RF). Aligned read counts were assigned to features in annotation files (.gtf) using featureCounts (v2.0.1)[63] with the following parameters: reverse strandedness, count fragments instead of reads, GFF feature type filter = "gene", multi-mapping and multi-overlapping features included (-M -O), minimum mapping quality per read of 0, and the rest of the parameters were kept as default. Differential gene expression analysis was then performed using DESeq2 (v2.11.40.7)[64] using default parameters. Differential expression tables were annotated with the Annotate DESeq2/DEXSeq output tables tool (v1.1.0) in Galaxy to include the following: GFF feature type = "CDS", GFF feature identifier = "gene_id", GFF transcript identifier = "transcript_id", GFF attribute to include = "protein_id, product". The "protein_id" was used to query the NCBI database and the NCBI Sequence Viewer was used to investigate the genomic context surrounding genes of interest.

Comparative genomics

The nucleotide sequence for the *Enterocloster bolteae* DSM 15670 *ucd* operon (NCBI NZ_CP022464.2 REGION: complement (4417875.4421605)) was used as a query for BLASTn (megablast) searches using the refseq_genomes database limited to Bacteria (taxid:2). The NCBI multiple sequence alignment (MSA) viewer was used to download alignment figures.

Comparisons between *ucd* operon genomic contexts were performed using CAGECAT (v1.0, https://cagecat.bioinformatics.nl/) using the clinker tool for visualization. GenBank files of *ucd* operon genomic contexts (*E. asparagiformis*: NZ_CP102272.1 [3012635.3088318], *E. bolteae*: NZ_CP022464.2 [4368517.4444988], *E. citroniae*: NZ_KQ235877.1 [350511.426194], *E. pacaense*: NZ_LS999946.1 [186864.253810]) were downloaded from the NCBI Sequence Viewer, and partial genes were trimmed from.gbk files using a text editor to prevent execution errors in the clinker. The identity threshold was set to 0.6, and proteins were manually recolored to highlight conserved features and potential regulators.

Phylogenetic tree construction

Phylogenetic trees based on single copy marker proteins. Trees for *Enterocloster* spp. and *Eggerthellaceae* family members were constructed using PhyloT (v2, https://phylot.biobyte.de/index.cgi) using a pruned tree from the Genome Taxonomy Database (GTDB) for Bacteria release 214[65]. Trees were visualized in iTOL (v7, https://itol.embl.de/)[66].

Phylogenetic trees based on whole genomes and proteomes. Trees were constructed using the Type (Strain) Genome Server (TYGS, https://tygs.dsmz.de)[67,68] with the following GenBank genome accessions: *E. aldenensis* (GCA_003467385.1), *E. asparagiformis* (GCA_025149125), *E. bolteae* (GCA_000154365), *E. clostridioformis* (GCA_900113155), *E. lavalensis* (GCA_003024655), *E. pacaense* (GCA_900566185). For *E. citroniae*, the Integrated Microbial Genomes

ObjectID was used: *E. citroniae* (2928404274). Further information on nomenclature and taxonomy was obtained from the List of Prokaryotic names with Standing in Nomenclature (LPSN, available at https://lpsn.dsmz.de). Taxa names in figures were updated according to the GTDB.

Ucd protein phylogenetic trees. The NCBI RefSeq accessions for UcdC: WP_002569575.1, UcdF: WP_002569574.1, and UcdO: WP_002569573.1 of *Enterocloster bolteae* DSM 15670 were used as queries for BLASTp searches against the refseq_protein database (max target sequences = 250) limited to Bacteria (taxid:2). The FASTA (complete sequence) files for all 3 queries were downloaded and used as inputs for multiple sequence alignment using the MAFFT L-INS-i (v7.525, https://mafft.cbrc.jp/alignment/software/source.html) tool with the following settings: −maxiterate 1000 −localpair −treeout −amino. Outputs were then trimmed using trimAl (v1.4, https://github.com/inab/trimal) using the -gappyout setting. IQ-TREE (v2.0.7, http://www.iqtree.org/) was used to construct a maximum likelihood phylogenetic tree with the following settings: -st AA -m TEST -bb 1000 -alrt 1000. Trees were visualized in iTOL.

RT-PCR analysis to determine *ucd* structure
Isolated RNA samples (500 ng) were reverse transcribed using the LunaScript® RT Master Mix Kit (Primer-free) (NEB) in a reaction volume of 10 μL containing the ucdCFO_RT-PCR_r primer (Supplementary Table 3) at a final concentration of 1 μM. The No-RT Control included in the kit was used as a no-enzyme control for reverse transcription. The reaction mixtures were incubated in a thermal cycler: 10 min at 55 °C, 1 min at 95 °C. PCR reactions were conducted using the OneTaq 2X Master Mix with Standard Buffer (NEB). The ucdCFO_RT-PCR primer pair (Supplementary Table 3) was added to the master mix (to a final concentration of 0.2 μM), and 1 μL of template (cDNA, -RT, no template, or gDNA) was added for a total reaction volume of 20 μL. PCR tubes were placed in a thermal cycler, and targets were amplified according to the following conditions: 20 s at 94 °C, 31 cycles (20 s at 94 °C, 30 s at 62 °C, 3 min at 68 °C), 5 min at 68 °C. A volume of 5 μL of reaction was directly loaded onto a 1% agarose gel (made with 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were compared to the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB). The rest of the PCR product was then run on a 1% agarose gel, and bands corresponding to the desired products were cut out and purified using the Monarch DNA Gel Extraction Kit (NEB). DNA was quantified using the Qubit dsDNA HS assay kit (Invitrogen) and submitted to Plasmidsaurus for long-read sequencing using Oxford Nanopore Technologies (GenBank PQ855390.1).

RT-qPCR analysis of *E. bolteae* ucd genes
Isolated RNA samples (500 ng) were reverse transcribed using the iScript Reverse Transcription Supermix (Bio-Rad) in a reaction volume of 10 μL. The iScript No-RT Control Supermix was used as a no enzyme control for reverse transcription (-RT). The reaction mixtures were incubated in a thermal cycler: 5 min at 25 °C, 20 min at 48 °C, and 1 min at 95 °C. Both cDNA and -RT controls were diluted 1/20 in nuclease-free water before use. qPCR reactions were conducted using the Luna Universal qPCR Master Mix kit (NEB). The Eb_ucdO_qPCR, Eb_ucdF_qPCR, Eb_ucdC_qPCR, and Eb_dnaK_Ref_qPCR primer pairs (Supplementary Table 3) were added to their respective master mixes (final primer concentration of 250 nM) and 6.6 or 4.4 μL of diluted template (cDNA, -RT, no template) were added to 26.4 or 17.6 μL of master mix for triplicates or duplicates, respectively. All cDNA samples were run in technical triplicates, while other sample types were run in technical duplicates. Replicate mixes were pipetted (10 μL/well) into a MicroAmp Fast 96-Well Reaction Plate (Applied Biosystems), and the plates were sealed, then spun down for 2 min to eliminate air bubbles. The qPCR detection parameters were as follows: SYBR Green detection, ROX reference dye, 10 μL reaction volume. The thermal cycling

conditions were: 1 min at 95 °C, 40 cycles (15 s at 95 °C, 30 s at 60 °C), then melt analysis (60–95 °C). Data were analyzed according to the $2^{-\Delta\Delta Ct}$ method[69] with the dnaK gene serving as the reference gene (*E. bolteae* dnaK RNA-seq $\log_2$FC = 0.122).

Protein extraction from *Enterocloster* spp
All steps other than sonication were carried out under anaerobic conditions. To extract proteins, 10 mL of treated (50 μM urolithin C for 4 h) *Enterocloster* spp. culture (see *Treatment during growth (spike-in)*) were pelleted (6500 g for 3 min), and the supernatant was discarded. The pellet was washed with 10 mL of pre-reduced PBS, pelleted again, and resuspended in 0.6 mL of pre-reduced lysis buffer. For urolithin metabolism assays, the lysis buffer was 1 X PBS, 1 mM DTT, 1% Triton X-100, and 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free). For proteomics, the lysis buffer was 20 mM Tris, pH 7.5, 500 mM NaCl, 10 mM $MgSO_4$, 10 mM $CaCl_2$, and 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free). The resuspended pellet was then sonicated on ice using a Misonix Sonicator 3000 set to power level 2/10 according to the following sequence (aerobically, in a cold room): 20 s ON, 40 s OFF, for a total of 2 min ON. Tubes were centrifuged at $20,000 \times g$ for 2 min to pellet insoluble particles, and 0.4 mL of lysate was transferred to a new 1.5 mL tube (kept on ice). Lysates used in metabolism assays were transported to the anaerobic chamber in a sealed plastic bag containing an anaerobic gas generating system to minimize loss in activity.

Urolithin metabolism assays in *E. bolteae* lysates
Protein lysates (described above) were aliquoted (50 μL aliquots) into 1.5 mL tubes, then treated with DMSO or urolithin C (10 mM stock) at a final concentration of 350 μM. Cofactors (NADPH, NADH, NADP, NAD, and FAD), each dissolved to a final concentration of 30 mM in lysis buffer (immediately before the assay was run) were added individually to the lysates at a final concentration of 1 mM. In select samples, sodium formate (HCO_2Na) was added to a final concentration of 2 mM (from a 40 mM stock solution in lysis buffer) and sodium dithionite ($Na_2S_2O_4$) was added to a final concentration of 5 mg/mL (from a 100 mg/mL stock solution in lysis buffer). The lysates were incubated at room temperature in an anaerobic chamber for 20 h prior to freezing at − 70 °C. Samples were then extracted using *Extraction Method C*.

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

Proteomics analysis of uroC-treated *E. bolteae*
Extracted proteins were submitted for proteomic analysis at the RI-MUHC. For each sample, protein lysates were loaded onto a single stacking gel band to remove lipids, detergents, and salts. The gel band was reduced with DTT, alkylated with iodoacetic acid, and digested with trypsin. Extracted peptides were re-solubilized in 0.1% aqueous formic acid and loaded onto a Thermo Acclaim Pepmap (Thermo, 75 μm ID X 2 cm C18 3 μm beads) precolumn and then onto an Acclaim Pepmap Easyspray (Thermo, 75 μm ID X 15 cm with 2 um C18 beads) analytical column separation using a Dionex Ultimate 3000 uHPLC at 250 nL/min with a gradient of 2–35% organic (0.1% formic acid in acetonitrile) over 3 h. Peptides were analyzed using a Thermo Orbitrap Fusion mass spectrometer operating at 120,000 resolution (FWHM in MS1) with HCD sequencing (15,000 resolution) at top speed for all peptides with a charge of 2 + or greater. The raw data were converted into *.mgf format (Mascot generic format) for searching using the Mascot 2.6.2 search engine (Matrix Science) against *Enterocloster bolteae* DSM 15670 proteins (NCBI assembly GCF_002234575.2) and a

database of common contaminant proteins. Mascot was searched with a fragment ion mass tolerance of 0.100 Da and a parent ion tolerance of 5.0 ppm. O-63 of pyrrolysine, carboxymethyl of cysteine and j + 66 of leucine/isoleucine indecision were specified in Mascot as fixed modifications. Deamidation of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable modifications.

The database search results were loaded into Scaffold Q + Scaffold_5.0.1 (Proteome Sciences) for statistical treatment and data visualization. Scaffold (v5.3.0) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm[70] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[71]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Protein quantification and differential expression were determined in Scaffold using the following parameters: Quantitative method was set to total spectra, the minimum value was set to 0.5 in case proteins were not detected in one condition, and statistical tests were performed using Fisher's exact test (two-sided) with the Benjamini-Hochberg multiple test correction at a significance level set to 0.05.

Protein structures and homology modeling

The amino acid sequences (single copies) for the *E. bolteae* UcdC (WP_002569575.1), UcdF (WP_002569574.1), and UcdO (WP_002569573.1) were used as inputs for AlphaFold3 (AF3) modeling using the AlphaFold server (https://alphafoldserver.com/). In parallel, the UniProt accession for UcdO (A8RZR2) was queried in the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/) and the integrated FoldSeek tool (https://github.com/steineggerlab/foldseek) was used to generate a list of proteins with similar structures from solved crystal structures in the Protein Data Bank (PDB)[72]. Foldseek hits of published X-ray crystal structures PDB 1RM6 (4-hydroxybenzoyl-CoA reductase from *Thauera aromatica*), PDB 1ZXI (Carbon monoxide dehydrogenase from *Afipia carboxidovorans* OM5), and PDB 3UNI (Xanthine dehydrogenase from *Bos taurus*) were fetched from the PDB and imported into PyMOL (v2.4.1). The UcdO subunit (chain C) in the predicted AF3 UcdCFO complex was superposed onto the following chains in the published PDBs using the "super" command: PDB 1RM6: chain A; PDB 1ZXI: chain B; PDB 3UNI: chain A.

Active site modeling of *Bt* XDH (PDB 3UNI) was performed by highlighting (showing as sticks) amino acid residues surrounding the salicylic acid ligand (active site inhibitor) and key residues identified in the literature (e.g., E1261). Active site modeling of the AF3 UcdCFO complex was performed by first superposing the structure onto PDB 1RM6, as described above. Then, amino acids (from the AF3 UcdCFO model) surrounding the Mo metal (from 1RM6) were selected and shown as sticks.

Multiple sequence alignments based on protein sequences of *Eb* UcdO (UniProt A8RZR2), *Ta* 4-HBCDα (UniProt O33819), *Ac* CoxL (UniProt P19919), and *Bt* XDH (UniProt P80457 – amino acids 1-523 removed) were generated in Benchling using the Benchling Clustal Omega sequence alignment tool (https://benchling.com).

Plasmid construction and transformation

Plasmid construction in E. coli NEB10β. Primers flanking the *E. bolteae ucd* operon (NCBI NZ_CP022464 REGION: complement(4417875.4421605)) were designed in Benchling using the Primer3 tool. Tails including 6 bases, followed by restriction sites for NdeI and XhoI were included on the forward and reverse primers, respectively (*ucdCFO*: Eb_ucdCFO_NdeI_f and Eb_ucdCFO_XhoI_r; *ucdCFO*-His:

Eb_ucdCFO_NdeI_f and Eb_ucdCFO_His_XhoI_r) (Supplementary Table 3). In the Eb_ucdCFO-His_XhoI_r primer, the stop codon of *ucdO* in the *ucd* operon was mutated to an alanine, yielding a C-terminal His$_6$-tagged protein.

PCR was performed using the Q5 High-Fidelity polymerase (NEB) with *E. bolteae* DSM 15670 genomic DNA as a template. The target was amplified according to the following cycling conditions: 30 s at 98 °C, 30 cycles (10 s at 98 °C, 20 s at 60 °C, 80 s at 72 °C), 2 min at 72 °C. The *ucdCFO* PCR product was purified using the Monarch PCR & DNA Cleanup Kit (NEB) according to the manufacturer's instructions for products ≥ 2 kb. The resulting purified PCR product and the pTipQC2 plasmid (Hokkaido Systems Science Co.) were digested overnight (16 h) with NdeI and XhoI (both from NEB) in rCutSmart buffer according to the manufacturer's instructions (~ 600−1000 ng DNA per 50 μL reaction). Double-digested DNA was migrated on a 0.6% agarose gel and bands corresponding to the desired products were cut out and purified using the Monarch DNA Gel Extraction kit (NEB). The purified products were ligated using the Hi-T4 DNA Ligase (NEB): a ~ 3:1 insert:plasmid molar ratio ligation reaction was set up on ice, then incubated at room temperature for 2 h. The ligation mixture (2 μL) was electroporated (1.8 kV, 25 μF, 200 Ω) into 40 μL electrocompetent *E. coli* NEB10β cells (according to the Quick-n'-Dirty Electrocompetent *E. coli* protocol (dx.doi.org/10.17504/protocols.io.bjpykmpw)) using 0.1 cm gap cuvettes (Bio-Rad). The cuvette was immediately filled with 1 mL pre-warmed LB post-shock and cells were allowed to recover at 37 °C for 30 min before plating on LB + 100 μg/mL ampicillin. After an overnight incubation at 37 °C, colonies were picked and grown in selective LB + 100 μg/mL ampicillin. Plasmids were purified using the Plasmid DNA Miniprep Kit (BioBasic), and size was confirmed with a diagnostic restriction digest (10 μL reactions). The final plasmid constructs (pTipQC2-*ucdCFO* and pTipQC2-*ucdCFO*-His) was submitted to Plasmidsaurus for long-read sequencing using Oxford Nanopore Technologies (v14 library preparation chemistry, R10.4.1 flow cells) (GenBank PQ855388.1 and PQ855389.1, respectively).

pTipQC2-*ucdCFO* transformation into Rhodococcus erythropolis DSM 43066. Electrocompetent *R. erythropolis* DSM 43066 were prepared according to a modified protocol from P. Lessard 2002 (Electrocompetent *Rhodococcus erythropolis* AN12). Briefly, 50 mL LB were inoculated with 1 mL of a stationary phase (48−72 h growth from a single colony) *R. erythropolis* DSM 43066 culture and grown aerobically for 16 h at 30 °C with shaking at 200 RPM. The next day, cells were pelleted at 5000 × *g* for 10 min at 4 °C and washed according to the following sequence: 2 washes of (10 mL of ice-cold sterile MilliQ water), 10 mL of ice-cold sterile 10% glycerol. The final pellet was then resuspended in 5 mL of ice-cold sterile 10% glycerol. The resuspended electrocompetent *R. erythropolis* DSM 43066 were aliquoted (50 μL/aliquot), then 3 μL (~ 0.5−1 μg) of pTipQC2-*ucdCFO* plasmid was added to appropriate tubes and incubated for 30 min on ice. Cells with plasmid were transferred to 0.1 cm gap cuvettes (Bio-Rad) and electroporated (1.8 kV, 25 μF, 200 Ω). Time constants were between 4.3−4.6 ms. The cuvette was immediately filled with 1 mL LB post-shock and cells were allowed to recover at 30 °C for 2.5 h before plating 100 μL dilutions (1/10 dilution, undiluted, and concentrated recovery culture) on LB + 30 μg/mL chloramphenicol at 30 °C. After 2-3 days of incubation, colonies were picked and grown in selective liquid LB + 30 μg/mL chloramphenicol at 30 °C with shaking at 200 RPM. Plasmid-positive colonies were identified by colony PCR using the pTipQC2-*ucdCFO*_cPCR primer set (Supplementary Table 3) and validated by diagnostic restriction digests and whole-plasmid sequencing.

Heterologous expression of UcdCFO

All growth steps below were performed in selective media (LB + 30 μg/mL chloramphenicol) in aerobic conditions at 30 °C with shaking at 220 RPM. Single colonies of *R. erythropolis* DSM 43066 harboring the

pTipQC2 (empty plasmid), pTipQC2-*ucdCFO*, or pTipQC2-*ucdCFO*-His were inoculated into 15 mL selective media and grown for 72 h to produce overnight cultures. Overnight cultures were then thoroughly resuspended and diluted 1:10 into 20 mL fresh selective media and grown for ~ 8 h until OD_{600} values reached ~ 0.6. Thiostrepton (5 mg/mL in DMSO) was added to a final concentration of 1 μg/mL and cultures were incubated aerobically for 16 h at 30 °C to induce protein expression. The next morning, cultures were pelleted and resuspended in 0.2 volumes of lysis buffer (1 X PBS, 1 mM DTT, 1% Triton X-100, 2 mg/mL lysozyme, and 1 tablet/100 mL SIGMAFAST protease inhibitor (EDTA-free)). If lysates were destined for further purification by immobilized metal affinity chromatography, lysis buffer was supplemented with 15 mM imidazole. The resuspended cells in lysis buffer (4 mL) were incubated on ice for 1 h with shaking, then sonicated on ice using a Misonix Sonicator 3000 set to power level 3/10 according to the following sequence (aerobically, in a cold room): 20 s ON, 40 s OFF, for a total of 4 min ON. Crude lysates were transported to the anaerobic chamber in a sealed plastic bag containing an anaerobic gas generating system to minimize loss in activity. To evaluate dehydroxylase activity, crude lysates or protein fractions were treated in the same manner detailed in *Urolithin metabolism assays in E. bolteae lysates* with the following modifications: lysates were treated with PBS or NADH (~ 2 mM final concentration in PBS), and incubations lasted 72 h.

Enrichment of UcdCFO-His proteins
Immobilized metal chelate affinity columns were freshly prepared by loading 1 mL of nickel-nitrilotriacetic acid (Ni-NTA) agarose (QIAGEN) slurry into Bio-Spin chromatography columns (Bio-Rad). The Ni-NTA agarose bed (0.5 mL) was then washed with 5 mL wash buffer (1X PBS, 1 mM DTT, 15 mM imidazole). Crude lysates described above (*Heterologous expression of UcdCFO*) were centrifuged for 2 min at 20,000 × g to pellet insoluble materials, and the entire supernatant (4 mL) was loaded onto the Ni-NTA column. The flow-through was collected and re-loaded onto the column to maximize the binding of UcdO-His and associated proteins. The column was washed with 14 mL of wash buffer, then bound proteins were eluted with 4 mL of elution buffer (1X PBS, 1 mM DTT, 200 mM imidazole) and 1 mL fractions were collected. The first fraction, which was the most active, was carried forward for further analyses (SDS-PAGE and activity assays).

SDS-PAGE analysis of UcdCFO proteins
Crude lysates described above (*Heterologous expression of UcdCFO*) were centrifuged for 2 min at 20,000 × g. The insoluble pellet was separated from the soluble supernatant. The insoluble pellet (from 100 μL of crude lysate) was resuspended in 100 μL of 1X reducing loading dye (62.5 mM Tris-HCl (pH 6.8), 2% (w/v) SDS, 10% glycerol, 0.01% (w/v) bromophenol blue, 38 mM DTT). The soluble fraction was diluted with 3X reducing loading dye to a final concentration of 1X. All samples were heated at 95 °C for 5 min, then 10 μL were loaded onto a 10% bis-tris polyacrylamide protein gel. Gels were fixed and stained with GelCode Blue Stain Reagent (Fisher) according to the manufacturer's instructions or using homemade 10X colloidal Coomassie solution (0.4% (w/v) Coomassie Brilliant Blue R250, 10% (w/v) citric acid, 8% (w/v) ammonium sulfate, and 20% (v/v) methanol).

Growth curves
Overnight cultures of *Enterocloster* spp. were treated as described in the *Treatment prior to growth (growth curves)* sub-section of *Treatments with urolithins and other catechols*. Once plated, 96-well plates were sealed with a Breathe-Easy membrane and placed in a pre-warmed plate reader inside the anaerobic chamber (BioTek Epoch 2). The optical density at 620 nm was recorded every 30 min for 48 h. Kinetic analysis was performed in BioTek Gen6 Software (v1.03.01) using the built-in kinetic analysis.

curatedMetagenomicData meta-analysis
All 93 metagenomic studies (22,588 samples and their metadata) available in the curatedMetagenomicData R package[51] (v3.8.0, 10.18129/B9.bioc.curatedMetagenomicData) were downloaded locally (ExperimentHub snapshotDate(): 2023-04-24, accessed on 2023-06-06) and transferred to the Narval cluster hosted by the Digital Research Alliance of Canada. Metagenomic data for urolithin C-metabolizing *Enterocloster* spp. were obtained by querying the "relative_abundance" (pre-processed using MetaPhlAn3) and "gene_families" (pre-processed using HUMAnN3) entries in individual study datasets[73]. For individual taxa (containing partial strings "bolteae", "citroniae", "asparagiformis", "asparagiforme", "Enterocloster", or "47FAA" (corresponding to *E. pacaense*)), relative abundance (%) was extracted from the rows of the "relative_abundance" datasets using the stringr R package (v1.5.0, https://github.com/tidyverse/stringr). Prevalence (relative abundance in sample > 0) was then calculated for each sample.

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

Fecal slurry preparation and treatment
Frozen (− 70 °C) fecal samples were brought into the anaerobic chamber and allowed to thaw. The samples were suspended in 1 mL mABB medium per 0.1 g feces and homogenized by breaking apart large pieces with a sterile loop and by vortexing. Large particles were pelleted by centrifuging the tubes at 700 × g for 3 min. The supernatants (containing bacteria) were transferred to new tubes and centrifuged at 6500 × g for 5 min to pellet the cells. The supernatants were discarded, and the cell pellets were washed with 5 mL of fresh media. The cell suspensions were once again centrifuged at 6500 × g for 5 min, and the resulting cell pellets were resuspended in 600 μL media per 0.1 g feces. Resuspended cells were treated with either 100 μM urolithin C or an equivalent volume of DMSO and incubated at 37 °C anaerobically for 48 h. 200–300 μL volumes were removed from the batch cultures and immediately frozen at − 70 °C for later extraction of urolithins (using *Extraction Method A*), DNA, and RNA.

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

Long read 16S rRNA sequencing in fecal slurries
Long-read 16S PCR reactions were conducted using the Platinum SuperFi II Green PCR Master Mix (Invitrogen). The ONT_16S_27F_GGK and ONT_16S_1492R_GGK primer pairs (Supplementary Table 3) were added to their respective master mixes (final primer concentration of 0.2 μM), and 1 μL of template (~ 10 ng) was added (for a total reaction volume of 25 μL). PCR tubes were placed in a thermal cycler, and targets were amplified according to the following cycling conditions: 30 s at 98 °C, 30 cycles (10 s at 98 °C, 10 s at 60 °C, 30 s at 72 °C), 5 min at

72 °C, and hold at 4 °C. Amplicons were quantified using the Qubit dsDNA HS assay kit (Invitrogen) to verify that amplicon concentrations were reasonably balanced (range = 18.36–24.00 ng/µL). Barcoding of amplicons was performed with 2 µL of PCR reaction according to the manufacturer's instructions (for ONT kit SQK-AMB111-24). Barcoding reactions were incubated in a thermal cycler for 10 min at 65 °C, then for 2 min at 80 °C. 10 µL of each barcoding reaction were pooled, and proteins were digested using heat-labile proteinase K (NEB) by incubating the pooled library for 15 min at 37 °C, followed by heat inactivation for 10 min at 55 °C. Amplicons were purified using Agencourt AMPure XP beads (Beckman Colter) using 0.7 volumes of beads-to-library. Following 70% EtOH washes and drying steps, the library was eluted using 15 µL of the provided elution buffer (EB), yielding a library with a concentration of 30 ng/µL using the Qubit dsDNA HS assay kit (Invitrogen). 11 µL of the eluted DNA library were transferred to a new tube and combined with 1 µL of Rapid Adapter T (RAP T). This mixture was incubated at room temperature for 10 min. Since the library was concentrated, it was diluted 1:2 in EB before combining with SB II and LB II, then loaded into a primed Flongle Flow Cell (R9.4.1) in a MinION device following the manufacturer's instructions. Sequencing was allowed to proceed for ~ 20 h until pore exhaustion or enough reads were obtained. Base calling & demultiplexing was performed using Guppy (v6.4.6) using the "SUP" super high accuracy model for R9.4.1 flow cells. The raw reads were filtered for a length between 1500 ± 200 bp. Filtered reads were assigned to taxa using Emu[74] (v3.4.4, https://github.com/treangenlab/emu) by mapping 16S rRNA sequences to the emu_database database (based on the NCBI 16S RefSeq with the entry for *E. asparagiformis* changed to the sequence obtained by ONT sequencing (GenBank accession PP280819) since the RefSeq sequence for this bacterium contained multiple N nucleotides that biased the assignment of *E. asparagiformis* to *E. lavalensis*). Data were not rarefied or scaled. Count tables were then used to create a phyloseq (v1.44.0, https://github.com/joey711/phyloseq) object in R[75]. Stacked bar plots were generated using ggnested (v0.1.0, https://github.com/gmteunisse/ggnested) and fantactic (v0.2.0, https://github.com/gmteunisse/Fantactic). Diversity analyses were performed using Microbiome Analyst (v2.0, https://www.microbiomeanalyst.ca/)[76].

Total RNA extraction from fecal slurries

A 300 µL fecal slurry aliquot (treated with either 100 µM urolithin C or an equivalent volume DMSO for 48 h) was thawed and pelleted (6500 g for 3 min). The pellet (in 200 µL media) was then mixed with 800 µL TRI reagent (Zymo Research). Samples were lysed in a Mini Beadbeater 16 (Biospec) according to the following sequence: 1 min ON, 5 min OFF for a total of 5 min ON. RNA isolation was then performed using the Directzol RNA Miniprep Kit (Zymo Research) according to the manufacturer's instructions (including an on-column DNase digestion). To ensure complete DNA removal, an additional DNA digestion step was performed on the isolated RNA using the Ambion DNA-free DNA Removal Kit (Invitrogen) according to the manufacturer's instructions. The DNA-free RNA was then cleaned up using the OneStep PCR Inhibitor Removal Kit (Zymo Research). RNA concentration and quality were verified by Qubit RNA BR assay kit (Invitrogen) and 1% agarose gel electrophoresis.

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

Total RNA was extracted from frozen fecal slurries as previously described (see *Total RNA extraction from fecal slurries*), and subsequently reverse transcribed as described above (see *RT-PCR analysis to determine ucd structure*) in a reaction volume of 5 µL. PCR reactions were conducted using the OneTaq 2X Master Mix with Standard Buffer (NEB). The ucdCFO_RT-PCR primer pair (Supplementary Table 3) was added to the master mix (to a final concentration of 0.2 µM), and 1 µL of the template (cDNA, -RT, or no template) was added for a total reaction

volume of 20 µL. PCR tubes were placed in a thermal cycler, and targets were amplified according to the following conditions: 30 s at 94 °C, 45 cycles (30 s at 94 °C, 1 min at 61 °C, 4 min at 68 °C), 5 min at 68 °C. A volume of 10 µL of reaction was directly loaded onto a 1% agarose gel (made with 1X TAE buffer) containing SafeView Classic (Abm). PCR product sizes were compared to the Quick-Load® Purple 1 kb Plus DNA Ladder (NEB).

PCR analysis of *ucd* operon prevalence in fecal slurries

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

Synthesis of di-O-methyl-urolithin C

Di-O-methyl-urolithin C (3-Hydroxy-8,9-dimethoxy-6H-dibenzo[b,d] pyran-6-one, CAS 126438-35-5) was synthesized based on previously reported Ullmann-type coupling conditions for urolithin derivatives[77]. Resorcinol (213 mg, 2 mmol) and 2-bromo-4,5-dimethoxybenzoic acid (261 mg, 1 mmol) were dissolved in 1 mL of 8% w/v NaOH (in MilliQ H_2O) and heated in a thermo-shaker set to 100 °C for 20 min (in 1.7 mL tube). Then, 200 µL of a 10% w/v Cu(II)SO$_4$ pentahydrate solution was added, and the reaction was heated at 100 °C for 1 h. The reaction solution (pink-red coloration) contained an insoluble precipitate which was collected by centrifugation ($20,000 \times g$ for 30 s). The insoluble pellet was washed 7 times with 1 mL of MilliQ H_2O until the pH of the wash solution was equal to the pH of MilliQ H_2O (~ pH 6). The pellet was dried by lyophilization for 16 h (0.0010 mbar, − 90 °C) and the product was recovered as a pale pink solid (94 mg, 35% yield).

^{1}H NMR (600 MHz, (CD$_3$)$_2$SO)): δ = 10.22 (s, 1H), 8.21 (d, J = 8.76 Hz, 1H), 7.69 (s, 1H), 7.54 (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); HRMS: m/z [M + Na]$^+$ calculated for C$_{15}$H$_{12}$NaO$_5$: 295.0577, found: 295.0585.

Statistical analyses and graphing

Statistical methods were not used to determine sample sizes, experiments were not randomized, and the investigators were not blinded. Data points for in vitro and ex vivo assays were assumed to be normally distributed, though this was not formally tested. Correlation analyses were performed using the non-parametric Spearman rank correlation (ρ). Statistical tests on bacterial relative abundances were performed using the Kruskal-Wallis test on untransformed relative abundance values, which are skewed toward 0. Statistical analyses for large datasets are detailed in the relevant methods sections. Details related to each test performed are supplied in the Figure legends, and statistical test results are provided in the Source data file for each figure panel, where applicable. In all cases, α = 0.05 and tests were two-tailed. Data were plotted in GraphPad Prism (v10.0.0) or using the ggplot2 (v3.4.2, https://ggplot2.tidyverse.org/) R package. Figures were assembled in Affinity Designer (v1.10.6.1665).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

RNA sequencing reads were deposited in the NCBI SRA BioProject ID PRJNA996126 under BioSample accession codes SAMN36514640 (*Enterocloster bolteae* DSM 15670) [https://www.ncbi.nlm.nih.gov/biosample/SAMN36514640/] and SAMN36514641 (*Enterocloster asparagiformis* DSM 15981) [https://www.ncbi.nlm.nih.gov/biosample/?term=SAMN36514641]. Untargeted proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner

repository with the dataset identifier PXD048514. Oxford Nanopore 16S rRNA sequencing reads of healthy human fecal slurries were deposited in the NCBI SRA BioProject ID PRJNA1073957. The 16S rRNA sequence for *E. asparagiformis* DSM 15981 used in the Emu database search was deposited in GenBank under accession PP280819.1 [https://www.ncbi.nlm.nih.gov/nuccore/PP280819]. The partial *ucd* operon cDNA sequence related to Fig. 2i was deposited in GenBank under accession PQ855390.1 [https://www.ncbi.nlm.nih.gov/nuccore/PQ855390]. The plasmid sequences for pTipQC2-*ucdCFO* and pTipQC2-*ucdCFO*-His were deposited in GenBank under accessions PQ855388.1 [https://www.ncbi.nlm.nih.gov/nuccore/PQ855388] and PQ855389.1 [https://www.ncbi.nlm.nih.gov/nuccore/PQ855389], respectively. Previously published X-ray crystal structures are available in the Protein Data Bank (PDB) under accessions 1RM6, 1ZXI, and 3UNI. Source data for Figs. 1–6 and Supplementary Figs. 3, 10–12, 14–18 are provided as Source data files. Source data are provided in this paper.

Code availability

All code, tables, and RData files obtained from the analysis of curatedMetagenomicData were deposited in Zenodo [https://doi.org/10.5281/zenodo.8302320].

References

1. Knight, R. et al. The microbiome and human biology. *Annu. Rev. Genom. Hum. Genet.* **18**, 65–86 (2017).
2. Taguer, M. & Maurice, C. F. The complex interplay of diet, xenobiotics, and microbial metabolism in the gut: Implications for clinical outcomes. *Clin. Pharmacol. Ther.* **99**, 588–599 (2016).
3. Almeida, A. et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).
4. Spanogiannopoulos, P., Bess, E. N., Carmody, R. N. & Turnbaugh, P. J. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat. Rev. Microbiol.* **14**, 273–287 (2016).
5. Ervin, S. M. et al. Structural insights into endobiotic reactivation by human gut microbiome-encoded sulfatases. *Biochemistry* **59**, 3939–3950 (2020).
6. Yao, L, et al. A biosynthetic pathway for the selective sulfonation of steroidal metabolites by human gut bacteria. *Nat. Microbiol.* **7**, 1404–1418 (2022).
7. Tomas-Barberan, F. A. & Espin, J. C. Effect of food structure and processing on (Poly)phenol-Gut microbiota interactions and the effects on human health. *Annu. Rev. Food Sci. Technol.* **10**, 221–238 (2019).
8. Maier, L. et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–628 (2018).
9. Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R. & Goodman, A. L. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature* **570**, 462–467 (2019).
10. Maini Rekdal, V., Bess, E. N., Bisanz, J. E., Turnbaugh, P. J. & Balskus, E. P. Discovery and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism. *Science* **364**, eaau6323 (2019).
11. Campbell, C. et al. Bacterial metabolism of bile acids promotes generation of peripheral regulatory T cells. *Nature* **581**, 475–479 (2020).
12. Paik, D. et al. Human gut bacteria produce Tau(Eta)17-modulating bile acid metabolites. *Nature* **603**, 907–912 (2022).
13. Cerqueira, F. M., Photenhauer, A. L., Pollet, R. M., Brown, H. A. & Koropatkin, N. M. Starch digestion by Gut Bacteria: Crowdsourcing for carbs. *Trends Microbiol.* **28**, 95–108 (2020).
14. Maini Rekdal, V. et al. A widely distributed metalloenzyme class enables gut microbial metabolism of host- and diet-derived catechols. *Elife* **9**, https://doi.org/10.7554/eLife.50845 (2020).
15. Garcia-Villalba, R., Selma, M. V., Espin, J. C. & Tomas-Barberan, F. A. Identification of novel urolithin metabolites in human feces and urine after the intake of a pomegranate extract. *J. Agric. Food Chem.* **67**, 11099–11107 (2019).
16. Carmody, R. N. et al. Diet dominates host genotype in shaping the murine gut microbiota. *Cell Host Microbe* **17**, 72–84 (2015).
17. Carmody, R. N. et al. Cooking shapes the structure and function of the gut microbiome. *Nat. Microbiol.* **4**, 2052–2063 (2019).
18. Sonnenburg, E. D. et al. Diet-induced extinctions in the gut microbiota compound over generations. *Nature* **529**, 212–215 (2016).
19. Rothschild, D. et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature* **555**, 210–215 (2018).
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).
21. Van Duynhoven, J. et al. Metabolic fate of polyphenols in the human superorganism. *Proc. Natl. Acad. Sci. USA* **108**, 4531–4538 (2011).
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).
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).
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).
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).
26. Ghosh, S., Banerjee, M., Haribabu, B. & Jala, V. R. Urolithin A attenuates arsenic-induced gut barrier dysfunction. *Arch. Toxicol.* **96**, 987–1007 (2022).
27. Singh, R. et al. Enhancement of the gut barrier integrity by a microbial metabolite through the Nrf2 pathway. *Nat. Commun.* **10**, 89 (2019).
28. Denk, D. et al. Expansion of T memory stem cells with superior antitumor immunity by Urolithin A-induced mitophagy. *Immunity* **55**, 2059–2073 (2022).
29. 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).
30. 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).
31. Luan, P. et al. Urolithin A improves muscle function by inducing mitophagy in muscular dystrophy. *Sci. Transl. Med.* **13**, eabb0319 (2021).
32. Selma, M. V. et al. Isolation of human intestinal bacteria capable of producing the bioactive metabolite isourolithin A from ellagic acid. *Front. Microbiol.* **8**, 1521 (2017).
33. Selma, M. V., Beltran, D., Garcia-Villalba, R., Espin, J. C. & Tomas-Barberan, F. A. Description of urolithin production capacity from ellagic acid of two human intestinal Gordonibacter species. *Food Funct.* **5**, 1779–1784 (2014).
34. Gaya, P., Peirotén, A., Medina, M., Alvarez, I. & Landete, J. M. Bifidobacterium pseudocatenulatum INIA P815: The first bacterium able to produce urolithins A and B from ellagic acid. *J. Funct. Foods* **45**, 95–99 (2018).
35. Garcia-Villalba, R. et al. Metabolism of different dietary phenolic compounds by the urolithin-producing human-gut bacteria Gordonibacter urolithinfaciens and Ellagibacter isourolithinifaciens. *Food Funct.* **11**, 7012–7022 (2020).
36. Iglesias-Aguirre, C. E. et al. Gut bacteria involved in ellagic acid metabolism To yield human Urolithin metabotypes revealed. *J. Agric. Food Chem.* **71**, 4029–4035 (2023).

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

38. Hosny, M. et al. Clostridium pacaense: a new species within the genus Clostridium. *New Microbes New Infect.* **28**, 6–10 (2019).

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

40. Leimkuhler, S. & Iobbi-Nivol, C. Bacterial molybdoenzymes: old enzymes for new purposes. *FEMS Microbiol. Rev.* **40**, 1–18 (2016).

41. Nakashima, N. & Tamura, T. Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. *Appl. Environ. Microbiol.* **70**, 5557–5568 (2004).

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

43. Varadi, M. et al. AlphaFold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).

44. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).

45. Unciuleac, M., Warkentin, E., Page, C. C., Boll, M. & Ermler, U. Structure of a xanthine oxidase-related 4-hydroxybenzoyl-CoA reductase with an additional [4Fe-4S] cluster and an inverted electron flow. *Structure* **12**, 2249–2256 (2004).

46. 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).

47. 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).

48. Hille, R., Hall, J. & Basu, P. The mononuclear molybdenum enzymes. *Chem. Rev.* **114**, 3963–4038 (2014).

49. 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).

50. Kramer, J., Ozkaya, O. & Kummerli, R. Bacterial siderophores in community and host interactions. *Nat. Rev. Microbiol.* **18**, 152–163 (2020).

51. Pasolli, E. et al. Accessible, curated metagenomic data through ExperimentHub. *Nat. Methods* **14**, 1023–1024 (2017).

52. 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).

53. 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).

54. 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).

55. 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).

56. Pereira, F. C, et al. The Parkinson's disease drug entacapone disrupts gut microbiome homoeostasis via iron sequestration. *Nat. Microbiol.* 9, 3165–3183 (2024).

57. Salgado, P., Melin, V., Duran, Y., Mansilla, H. & Contreras, D. The reactivity and reaction pathway of fenton reactions driven by substituted 1,2-dihydroxybenzenes. *Environ. Sci. Technol.* **51**, 3687–3693 (2017).

58. 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).

59. Bae, M. et al. Metatranscriptomics-guided discovery and characterization of a polyphenol-metabolizing gut microbial enzyme. *Cell Host Microbe* **32**, 1887–1896 (2024).

60. 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).

61. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).

62. 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).

63. 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).

64. 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).

65. Parks, D. H. et al. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* **38**, 1079–1086 (2020).

66. Letunic, I. & Bork, P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* **52**, W78–W82 (2024).

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

68. 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).

69. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.* **3**, 1101–1108 (2008).

70. 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).

71. 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).

72. Van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* **42**, 243–246 (2024).

73. Beghini, F. et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* **10**, e65088 (2021).

74. 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).

75. McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

76. Lu, Y. et al. MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data. *Nucleic Acids Res.* **51**, W310–W318 (2023).

77. 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).

Acknowledgements

Author contributions

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

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-56266-2.

Correspondence and requests for materials should be addressed to Bastien Castagner.

Peer review information *Nature Communications* thanks Yaowei Fang, Duncan Kountz, and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.