Exceptionally versatile respiratory metabolisms drive metabolite production by diverse gut bacteria

Alexander S. Little,[1,2] Isaac T. Younker,[1,2] Matthew S. Schechter,[1,2] Raphaël Méheust,[3] Joshua Stemczynski,[1,2] Kaylie Scorza,[1,2] Paola Nol Bernardino,[1,2] Michael W. Mullowney,[1] Deepti Sharan,[1] Emily Waligurski,[1] Rita Smith,[1] Ramanujam Ramanswamy,[1] William Leiter,[1] David Moran,[1] Mary McMillin,[1] Matthew A. Odenwald,[4] Anthony T. Iavarone,[5] Ashley M. Sidebottom,[1] Anitha Sundararajan,[1] Eric G. Pamer,[1,2,6] A. Murat Eren,[7,8] Samuel H. Light[1,2]*

[1]Duchossois Family Institute, University of Chicago, Chicago, IL, USA
[2]Department of Microbiology, University of Chicago, Chicago, IL, USA
[3]Génomique Métabolique, CEA, Genoscope, Institut François Jacob, Université d'Évry, Université Paris-Saclay, CNRS, Evry, France
[4]Department of Medicine, Section of Gastroenterology, Hepatology, and Nutrition, University of Chicago
[5]QB3/Chemistry Mass Spectrometry Facility, University of California, Berkeley, Berkeley, CA, USA
[6]Department of Medicine, Section of Infectious Diseases & Global Health, University of Chicago Medicine, Chicago, IL, USA.
[7]Helmholtz Institute for Functional Marine Biodiversity, Oldenburg, Germany
[8]Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenbug, Germany
*Address correspondence to Samuel H. Light, samlight@uchicago.edu

Anaerobic respiration encompasses a major class of microbial energy metabolisms that employ reductases to respire different non-oxygen electron acceptors. Respiratory reductases play important roles in multiple geochemical cycles but their significance in other contexts remains unclear. Here we identify three taxonomically distinct families of gut bacteria that encode exceptionally large arsenals of tens to hundreds of respiratory-like reductases per genome. By screening representative species from each family, we discover 22 metabolites used as respiratory electron acceptors in a species-specific manner. Identified respiratory reactions transform multiple classes of dietary- and host-derived fecal metabolites, including bioactive molecules like resveratrol and the immunometabolite itaconate. Reductase substrate-profiling identifies enzyme-substrate pairs and paints a striking picture of reductase evolution, including evidence that reductases with specificities for related cinnamate substrates but distinct mechanisms independently emerged at least four separate times. These studies define an exceptionally versatile form of respiration that directly links microbial central metabolism to the gut metabolome.

1

Interpersonal variability in the composition or activity of the gut microbiome has been linked to numerous aspects of mammalian health, including digestive, metabolic, and autoimmune diseases. One of the most common mechanisms by which gut microbes modulate host physiology is through the production or depletion of bioactive metabolites.[1] Despite the importance of microbially derived metabolites for mammalian biology, the spectrum of metabolites produced by the microbiome (collectively, the gut metabolome) and their relationship to microbial physiology remain only partially defined.

Within microbial ecosystems, different microbes are adapted to metabolize different energy sources.[2] This metabolic specialization enables microbes to non-competitively coexist and is a major factor in shaping the composition of microbial communities.[2] In addition to explaining microbial membership within the gut microbiome, cellular energy demands necessitate high metabolic fluxes. Consequently, the products of microbial energy metabolisms accumulate to relatively high concentrations and can provide a significant source of metabolites within the gut. For example, many members of the gut microbiome ferment carbohydrates into the short chain fatty acids, which have numerous impacts on mammalian metabolism and immunology.[1] Similarly, products from microbial amino acid fermentation within the gut modulate the intestinal barrier and systemic immune responses.[1,3,4] Microbial energy metabolisms are thus critical for explaining microbial adaptions to the gut and the production of key constituents of the gut metabolome.

Cellular respiration is one of the fundamental forms of energy metabolism. In heterotrophic microbes, cellular respiration is defined by the oxidation of an organic electron donor and the passage of resulting electrons through an electron transport chain to a terminal electron acceptor.[5] Electron transfer is coupled to the creation of an ion gradient which, in turn, powers oxidative adenosine triphosphate (ATP) synthesis through ATP synthase.[5] Molecular oxygen serves as the respiratory electron acceptor in aerobic environments and is the hallmark acceptor of cellular respiration. However, some microbes residing in anaerobic environments, such as the lower gastrointestinal tract, possess respiratory metabolisms that utilize non-oxygen, alternative electron acceptors.[5]

The use of different respiratory electron acceptors is well established in environmental microbes, where the activity has received particular attention in the context of biogeochemical cycles (*e.g.*, nitrogen and sulfur cycles).[5,6] In the gut, fermentation represents the dominant type of microbial energy metabolism but respiratory metabolisms also occur, particularly in the context of inflammation. Inflammatory processes generate reactive oxygen and nitrogen species that lead to the accumulation of inorganic electron acceptors tetrathionate and nitrate, which fuel the respiratory growth of enteric pathogens.[7,8] In the non-inflamed gut, specialized sulfate-reducing bacteria use various sulfates as respiratory electron acceptors, while acetogens and methanogens respire carbon dioxide.[9,10] Additional reports have characterized several alternative respiratory mechanisms, but the relevance of organic respiratory electron acceptors for the gut metabolome remains largely unexplored.[11–13]

Here we employ a genome mining-based approach to interrogate respiratory electron acceptor usage in the human gut microbiome. We find that three taxonomically distinct families of gut microbes encode exceptionally large numbers of respiratory-like reductases. We show that representative species from each clade exhibit respiratory activities and identify diverse small molecule electron acceptors used in a strain-dependent manner. We additionally pair specific reductases with their substrates to reveal a convoluted relationship between reductase sequence and substrate specificity. By probing reductases with distinct active site architectures, we show that parallel evolutionary trajectories generated reductases that achieve similar activities through distinct mechanisms. These studies establish a distinctive mode of bacterial respiration, defined by the versatile use of organic electron acceptors, and contextualize the role of energy metabolism in shaping the gut metabolome.

Results

Three distinct groups of gut bacteria possess large reductase arsenals

Previous studies have identified multiple respiratory reductases that use distinct electron acceptors. Characterized respiratory reductases are often evolutionarily related, sharing a level of sequence homology that places them within one of two enzyme superfamilies. “Molybdopterin reductases” (Pfam PF00384) have a catalytic domain that binds the redox-active cofactor molybdopterin and tend to act on inorganic respiratory electron acceptors (**Figure 1A**, **Table S1**).[13,14,23–25,15–22] “Flavin reductases” (Pfam PF00890) have a catalytic domain that binds the redox-active cofactor flavin adenine dinucleotide (FAD) and tend to act on organic respiratory electron acceptors (**Figure 1B**, **Table S1**).[26–31] Structural and biochemical studies of various enzymes within the flavin and molybdopterin superfamilies suggest each contains a broadly conserved electron transfer mechanism but active site distinctions that result in different substrate specificities and activities.[32,33]

Most characterized respiratory reductases are soluble proteins that receive electrons from the electron transport chain in the cytosolic membrane. While respiratory reductases can be associated with either the cytosolic or extracytosolic side of the cytosolic membrane, we observed that most characterized flavin and molybdopterin respiratory reductases contain an N-terminal signal peptide characteristic of extracytosolic localization (**Table S1**). Strikingly, one of the few characterized cytosolic reductases, *Escherichia coli* nitrate reductase, was reported to have lost its signal peptide

in recent evolutionary history, thus providing a seeming "exception that proves the rule" to the general pattern of the extracytosolic localization of flavin and molybdopterin respiratory reductases.[34]

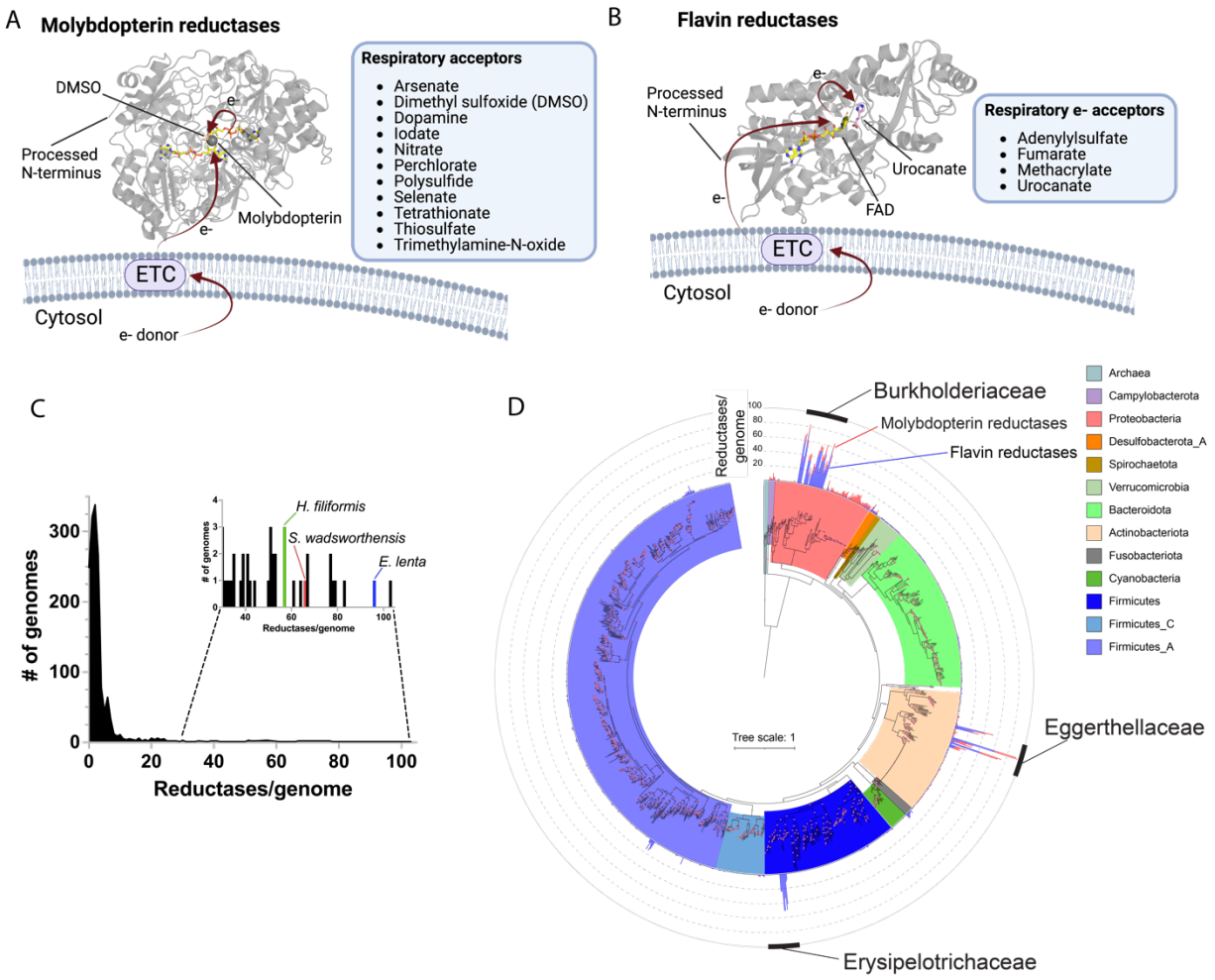


Figure 1. Respiratory reductase orthologs are highly overrepresented in three distinct lineages of gut bacteria. (A) General mechanism and electron acceptors used by different previously characterized respiratory molybdopterin reductases (Pfam PF00384). Arrows highlight electron (e-) transfer path from the electron transport chain (ETC) to dimethylsulfoxide (DMSO) in the DMSO reductase crystal structure (PDB code 4DMR). (B) Electron acceptors and general electron transfer mechanism used by previously characterized respiratory flavin reductases (Pfam PF00890). Arrows highlight electron transfer to substrate in the urocanate reductase co-complex crystal structure (PDB code 6T87). (C) Distribution of the number of flavin/molybdopterin reductases in 1533 representative human gut bacteria genomes and metagenomes genomes. (D) Phylogenetic reconstruction of the evolutionary history of genomes analyzed in (C). The maximum likelihood tree was constructed based on a concatenated alignment of 14 ribosomal proteins under an LG + I + G4 model of evolution (2092 amino acid sites). The number of flavin (blue) and molybdopterin (red) reductases with a computationally predicted signal peptide in each genome are graphed on the outer ring of the tree.

Based on the observed properties of previously characterized respiratory reductases, we reasoned that a survey of genes predicted to encode secreted flavin or molybdopterin reductases could facilitate the identification of respiratory gut microbes. We analyzed 1533 genomes and metagenome-assembled genomes from a recently complied collection of common human gut prokaryotes and observed a striking distribution in the number of reductases per genome.[35] While most genomes encode less than 5 reductases, we identified a small subset of 'high reductase' genomes that encode more than 30 (and as many as 103) reductases (**Figure 1C**, **Table S2**). Consistent with these reductases having an extracellular localization and respiratory functionality, we observed that most containe a predicted N-terminal signal peptide (**Table S2**). Further taxonomic

analyses revealed that the identified 'high reductase genomes form three distinct clades in the: (1) Actinobacteria family Eggerthellaceae, (2) Firmicutes family Erysipelotrichaceae, and (3) Proteobacteria family Burkholderiaceae (**Figure 1D**). Strikingly, species from each clade – including *E. lenta* from Eggerthellaceae, *H. filiformis* from Erysipelotrichaceae, and *S. wadsworthensis* from Burkholderiaceae – are prevalent, moderate-to-low abundance members of the human gut microbiome (**Figure 1C, inset**).[36] These analyses thus reveal three taxonomically distinct groups of gut microbes that encode large numbers of respiratory-like reductases per genome.

High reductase content clades reflect dynamic evolutionary trajectories

To gain insight into the evolutionary history of reductase expansion in Burkholderiaceae, Eggerthellaceae, and Erysipelotrichaceae, we performed a phylogenomic analysis of all classified species in each family (including those that are not associated with the human gut microbiome) and overlaid the reductases per genome data on the resulting tree (**Figure S1A-C, Table S3**). These analyses identified genomes encoding as many as 232 total reductases and established that high reductase genomes form monophyletic clades within the three families. The Erysipelotrichaceae high reductase clade corresponded to the genus *Holdemania*, with the Eggerthellaceae high reductase clade encompassing the genera *Arabia, Adlercreutzia*, *Denitrobacterium*, *D16-34*, *Eggerthella*, *Enteroscipio*, *Gordonibacter*, *Paraeggerthella*, *Parvibacter*, *Raoultibacter*, *Rubneribacter*, *Senegalimassilia*, and *Slackia* and the Burkholderiaceae encompassing the genera *Duodenibacillus*, *Mesosutterella*, *Parasutterella, Sutterella*, and *Turicimonas* (**Figure S1A-C)**. Strikingly, we observed variable but consistently high numbers of reductases per genome among different strains of the same species but found several 'low reductase' species interspersed among the 'high reductase' Eggerthellaceae and Burkholderiaceae clades (**Figure S1A-B)**. These observations are consistent with independent evolutionary trajectories characterized by dramatic expansions in the reductase gene content in ancestral Eggerthellaceae, Erysipelotrichaceae, and Burkholderiaceae species. Furthermore, the variable retention of reductases in descendant lineages suggests that high reductase content genomes confer a context-dependent and probable niche-specific) competitive advantage.

To evaluate reductase evolutionary dynamics on a more contemporary timescale, we performed a 'reductase pangeome' analysis which assessed the variability in the content of paralogous flavin reductases across representative Eggerthellaceae, Erysipelotrichaceae, and Burkholderiaceae family species (*E. lenta*, *H. filiformis*, and *S. wadsworthensis*, respectively). For all three species, we observed that individual strains differed in their reductase content, containing a 'core' set of reductases, which were present in all strains, and an 'accessory' set of reductases present in a subset of strains (**Figure S2**). The *E. lenta* reductase pangenome included the largest number of genomes (84) and revealed the most intricate picture of reductase evolution (23 core and 86 accessory reductases). Prevalence of accessory reductases also varied, ranging from present in a single *E. lenta* strain to absent in a single strain. These results provide evidence of the continued dynamism of reductase evolution and suggest the selective advantage conferred by some reductases is central to the species but strain-dependent for others.

Gut bacteria with large reductase arsenals exhibit respiratory growth properties

To test whether gut microbes with a high number of reductases per genome possessed respiratory capabilities, we selected *E. lenta*, *H. filiformis*, and *S. wadsworthensis* strains for experimental characterization. From each strain's genome we identified a flavin reductase with greater than 50% sequence identity to a previously characterized respiratory urocanate reductase UrdA, suggesting that the small molecule urocanate provided a suitable candidate for preliminary investigations of respiratory growth.[27,29] As use of a respiratory electron acceptor is conditional upon oxidation of an electron donor, we screened common respiratory electron donors for the ability to induce urocanate-dependent growth enhancement. We observed a synergistic formate/urocanate-dependent growth

enhancement of *S. wadsworthensis* and *E. lenta*, establishing formate as a viable electron donor for these strains (**Figure 2A**). *H. filiformis* also exhibited enhanced growth in the presence of urocanate, but requirements for a rich growth medium hindered identification of a respiratory electron donor (**Figure 2A**). Consistent with the observed phenotypes reflecting respiratory activity, growth enhancement of each strain tracked with urocanate reduction to imidazole propionate and depended upon the electron-accepting properties of urocanate (i.e., the reduced reaction product did not enhance growth) (**Figure 2A-2C**). Further supporting the role of respiration in these growth phenotypes, we found that urocanate stimulated ATP synthesis of all three strains (**Figure 2D**). These results thus demonstrate that *E. lenta*, *H. filiformis*, and *S. wadsworthensis* possess respiratory metabolic capabilities and similarly utilize urocanate as a respiratory electron acceptor.

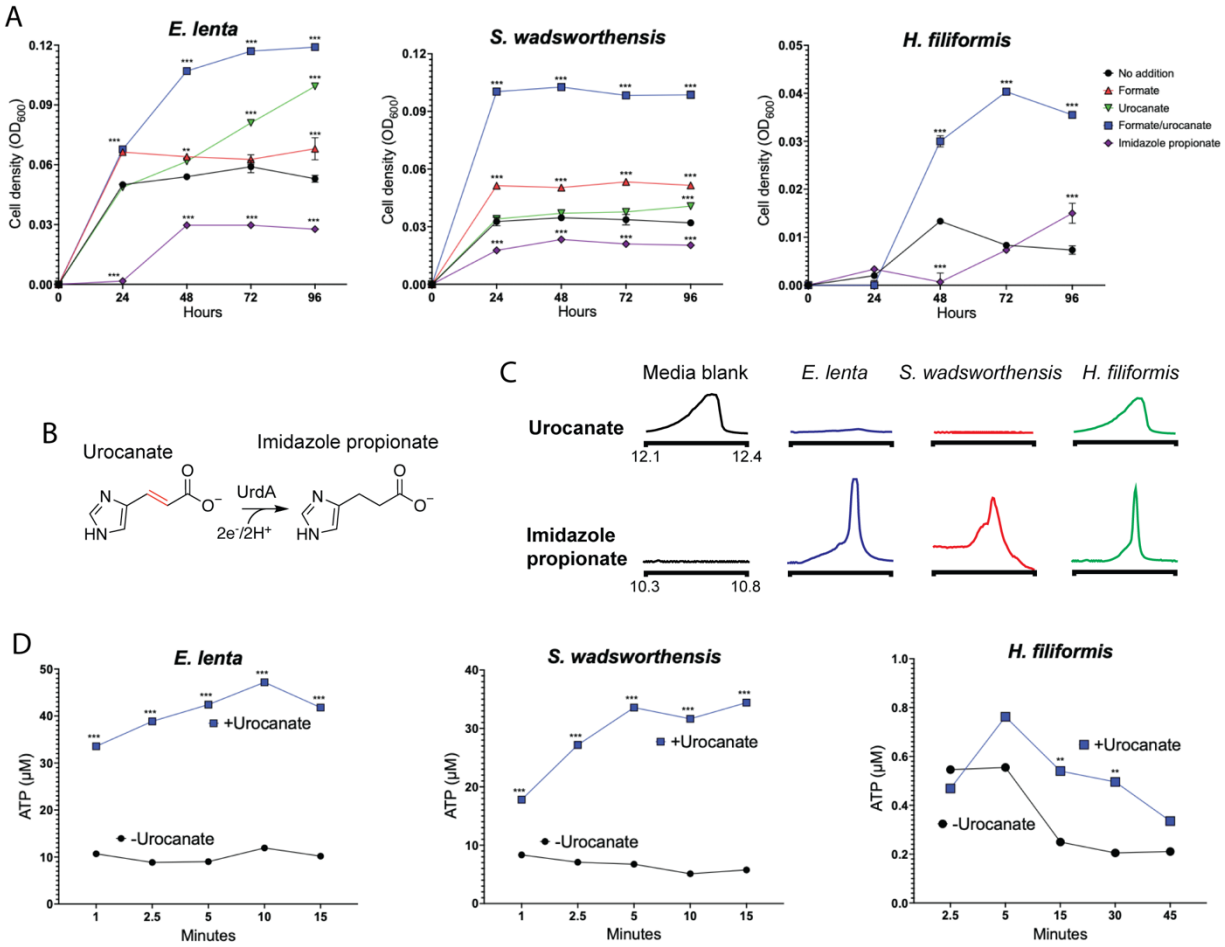


Figure 2. 'High reductase' gut bacteria exhibit respiratory growth properties. (A) Growth of *E. lenta* DSM2243, *S. wadsworthensis* DFI.4.78, and *H. filiformis* DSM12042 strains in the presence of indicated electron donor (formate), electron acceptor (urocanate), and product of urocanate reduction (imidazole propionate). (B) Reaction catalyzed by urocanate reductase UrdA. (C) GC-MS analysis of metabolites in formate/urocanate-supplemented culture supernatants following cultivation of the indicated strains. (D) ATP produced by cells suspended in formate-supplemented buffer. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

Bacteria with large reductase arsenals utilize diverse respiratory electron acceptors

Having established the respiratory capability of our strains, we hypothesized that the high number of reductases per genome reflected an extreme versatility in respiratory electron acceptor utilization capabilities. Since the majority of reductases encoded by these strains are distantly related to functionally characterized enzymes and thus lack obvious substrates, we assembled a panel of small molecules that possessed electron-accepting properties and could be present in the

gastrointestinal tract. We screened this panel for growth enhancement of *E. lenta*, *S. wadsworthensis* and *H. filiformis* and identified nineteen compounds that supported growth of at least one species (**Figure 3A**, **Figures S3-S5**). For most compounds identified in our screen we confirmed that: (1) the presence of an electron donor was required for growth enhancement, (2) substrate depletion coincided with product accumulation, and (3) the reduced product did not impact growth (**Figures S3-S5**). Through complementary enzymatic activity assays, we identified three additional molecules that were used as electron acceptors but, due to poor solubility or other factors, were not associated with enhanced growth under the tested conditions **(Figure 3B & 3C**, **Figure S6**). Collectively, these results define a complex species-dependent pattern of electron acceptor usage further explored below.

Cinnamates

Cinnamates are a class of compounds with a *trans*-phenyl-3-propenoate core that are common polyphenol conjugates in plants, including many foods and beverages.[37] We found that *E. lenta* and *H. filiformis* differentially used a variety of cinnamates as respiratory electron acceptors. While there was some overlap in the derivatives used by the two strains, *H. filiformis* was generally more accommodating to functional group substitutions on the cinnamate phenyl ring, utilizing several derivatives (cinnamate, *m*-coumarate) that were not used by *E. lenta* (**Figure 3A-C**, **Figure S3, S5**). By contrast, *E. lenta* was unique in accommodating carboxyl group modifications, utilizing the cinnamate esters chlorogenate and rosmarinate (**Figure 3A-C**).

In addition to exhibiting distinct substrate specificities, we observed that *E. lenta* and *H. filiformis* modified cinnamates in different ways. While *H. filiformis* exclusively reduced the enoate group, *E. lenta* also reductively dehydroxylated catechol cinnamates – presumably through the action of a previously characterized hydrocinnamate dehydroxylase (**Figure 3B**, **Figure S7**).[13] As dehydroxylation has previously been observed to be catalyzed by molybdopterin reductases, the lack of dehydroxylation is consistent with the observation that *H. filiformis* encodes flavin reductases but no molybdopterin reductases (**Figure 1D, Figure S1C**).[13] *E. lenta* however is capable of both dehydroxylating the catechol caffeate to *m*-coumarate and reducing its enoate to hydrocaffeate. A comparison of the caffeate reaction products and independent experiments with *m*-coumarate and hydrocaffeate revealed that both enoate reduction and catechol dehydroxylation supported respiratory growth (**Figure S7**). Interestingly, *E. lenta* exhibited distinct substrate specificities for dehydroxylation and enoate reduction, as catechol dehydroxylation could occur before or after enoate reduction (i.e., both caffeate and hydrocaffeate were dehydroxylated) but enoate reduction only occurred pre-dehydroxylation (i.e., the enoate on the dehydroxylated product *m*-coumarate was not further reduced) (**Figure S7**). These results thus reveal a complex and variable use of cinnamates as respiratory electron acceptors by *E. lenta* and *H. filiformis*.

(C4)-dicarboxylates

Four-carbon (C4)-dicarboxylates (including intermediates in the tricarboxylic acid cycle, bacterial fermentative products, and amino acids) have been shown to be relevant growth substrates within the gut microbiome.[12] We observed that *S. wadsworthensis* selectively utilized the C4-dicarboxylates aspartate, fumarate, malate, and tartrate as respiratory electron acceptors (**Figure 3A-C, Figure S4**). Each of these C4-dicarboxylates was reduced to succinate, suggesting a metabolic pathway that converges upon fumarate similar to the previously described Enterobacteraceiae mechanism (**Figure 3C, Figure S4**).[12] Among the strains tested, *S. wadsworthensis* is thus uniquely adapted for use of C4-dicarboxylates.

Sulfoxides and catechols

Sulfoxides and catechols are two distinct classes of compounds that have previously been identified as substrates for different molybdopterin reductases. Small molecule sulfoxides are a significant component of cruciferous vegetables and are generated by the nonenzymatic oxidation of sulfides, which may be relevant for gut inflammation when the influx of immune cell-derived reactive oxygen species has been proposed to convert small molecule sulfides into sulfoxides.[38] Dietary plant phenols and conjugated catecholamines (dopamine, epinephrine, norepinephrine) excreted in the bile are significant sources of catechols within the gut.[39] Consistent with the high number of molybdopterin reductases encoded by *E. lenta* and *S. wadsworthensis,* we observed that these microbes reduced multiple sulfoxides (dimethyl sulfoxide, methionine sulfoxide, and methylcysteine sulfoxide) and catechols (catechin and epicatechin) (**Figure 3B & 3C, Figure S3, S4 & S6**). *E. lenta* and *S. wadsworthensis* thus similarly utilize a chemically diverse group of electron acceptors.

Other electron acceptors

Finally, we identified a heterogeneous group of enoate/alkene electron acceptors differentially used by the three tested strains. This group includes the previously discussed product of histidine catabolism, urocanate, which was used by all three strains, and shikimate, a major intermediate in the biosynthesis of aromatic compounds, which was used by *E. lenta* and *S. wadsworthensis* (**Figure 2, Figure 3A-C, Figure S3 & S4**). The alkene resveratrol, a plant phenylpropanoid present in different foods/beverages that has many reported effects on mammalian physiology, was converted by both *E. lenta* and *H. filiformis* (**Figure S6**). Finally, itaconate, a mammalian immune signaling molecule that is produced by the interferon-inducible gene Irg1 in response to infection, was reduced to methylsuccinate by *E. lenta* (**Figure 3C, Figure S3**).[40–42] To our knowledge, itaconate has no previously established links to the gut microbiome

Electron acceptors and reduced products are present within feces and correlate with gut microbiome complexity

To clarify the relevance of identified electron acceptors within the gut microbiome, we measured fecal metabolites in a population of 19 healthy (non-antibiotic-treated) human stool donors and 22 hospitalized antibiotic-treated patients with severely diminished microbiome complexity (**Figure S8A**). As the multiple differences that distinguish these healthy and antibiotic-treated groups precludes the definitive attribution of metabolite differences to the microbiome, we additionally sought a more controlled experimental system. We thus measured electron acceptors and their reduced products in mouse fecal samples before and after treatment with a cocktail of broad-spectrum antibiotics that severely disrupted the gut microbiome (**Figure S8B**).

Distinct compounds were detected in the mice and human samples but in both populations levels of reductase substrates and products tended to be lower in the antibiotic-treated groups (**Figure 3D & 3E, Table S4 & S5**). This pattern was evident for multiple cinnamates, which can be produced from bacterial breakdown of larger dietary polyphenols, urocanate and shikimate, which are intermediates in bacterial metabolic pathways and, perhaps most strikingly, for the immunometabolite itaconate (**Figure 3D, Table S4**). Itaconate and its reduced product methylsuccinate were detected in virtually every healthy human fecal sample but were below the limit of detection for the majority of the antibiotic-treated group (**Figure 3D, Table S4**). These studies establish that the identified respiratory electron acceptors are present within the mammalian gastrointestinal tract and provide evidence that their levels are influenced by the gut microbiome.

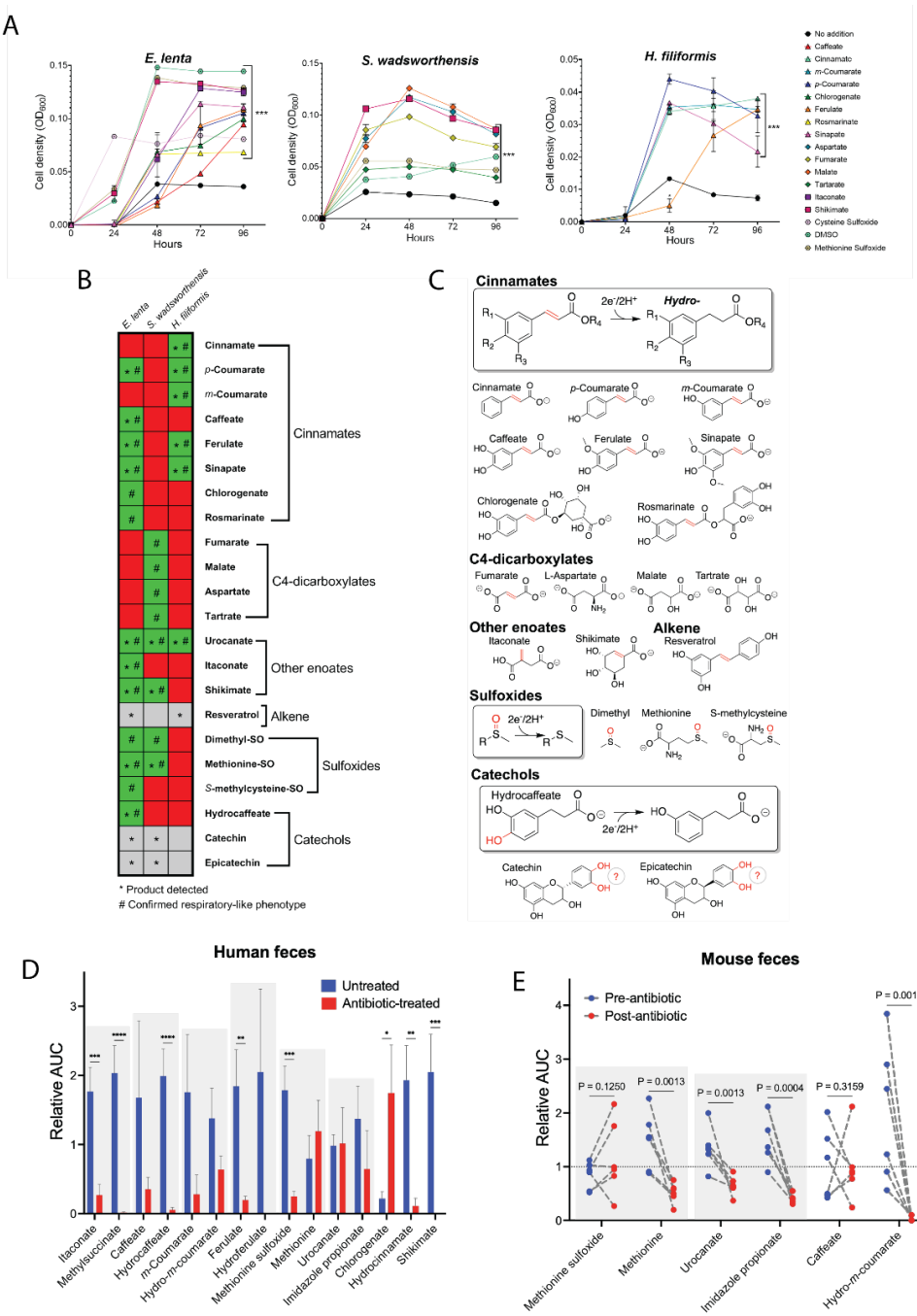


Figure 3. 'High reductase' gut bacteria utilize diverse respiratory electron acceptors. (A) *E. lenta* DSM2243, *S. wadsworthensis* DFI.4.78, and *H. filiformis* DFI.9.20 growth in media supplemented with formate and identified growth-stimulating small molecules. (B) Summary of electron acceptor usage findings. Green indicates growth enhancement, red indicates no enhancement, and grey denotes limited solubility. (C) Summary of identified reductase activities. The electron-accepting bond is highlighted in red. (D) Identified metabolites in fecal samples collected from non-antibiotic versus antibiotic treated subjects as measured by LC-MS. Area under the curve (AUC) values are presented relative to each metabolite's normalized average (normalized to 1) within the dataset as measured by LC-MS. See **Table S4** for original data. (E) Identified metabolites in mouse feces pre- and post-antibiotic treatment. Area under the curve (AUC) values are presented relative to each metabolite's normalized average (normalized to 1) within the dataset. See **Table S5** for original data. *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone (A), multiple unpaired t tests (D, E).

Respiratory electron acceptors selectively induce specific flavin reductases

An examination of the genomic context of *E. lenta*, *S. wadsworthensis*, and *H. filiformis* flavin and molybdopterin reductase genes revealed that they frequently colocalized with putative transcriptional regulators, though notable species-specific differences in the type of regulators were evident. *S. wadsworthensis* and *H. filiformis* reductase genes often colocalized with histidine kinase two-component systems or typical helix-turn-helix cytosolic transcriptional regulators (**Figure 4A & 4B, Table S6**). For *E. lenta*, 74 reductase genes directly neighbor a gene encoding a classical DNA-binding helix-turn-helix domain (**Figure 4A & 4B, Table S6**). Strikingly, 64 of these contain an N-terminal domain with 3-12 predicted transmembrane helices (**Table S6**). Notably, the vast majority of *E. lenta* genes that encode proteins with this transmembrane/helix-turn-helix topology neighbor a reductase gene, suggesting that they may represent an unusual type of reductase-associated transcriptional regulator.

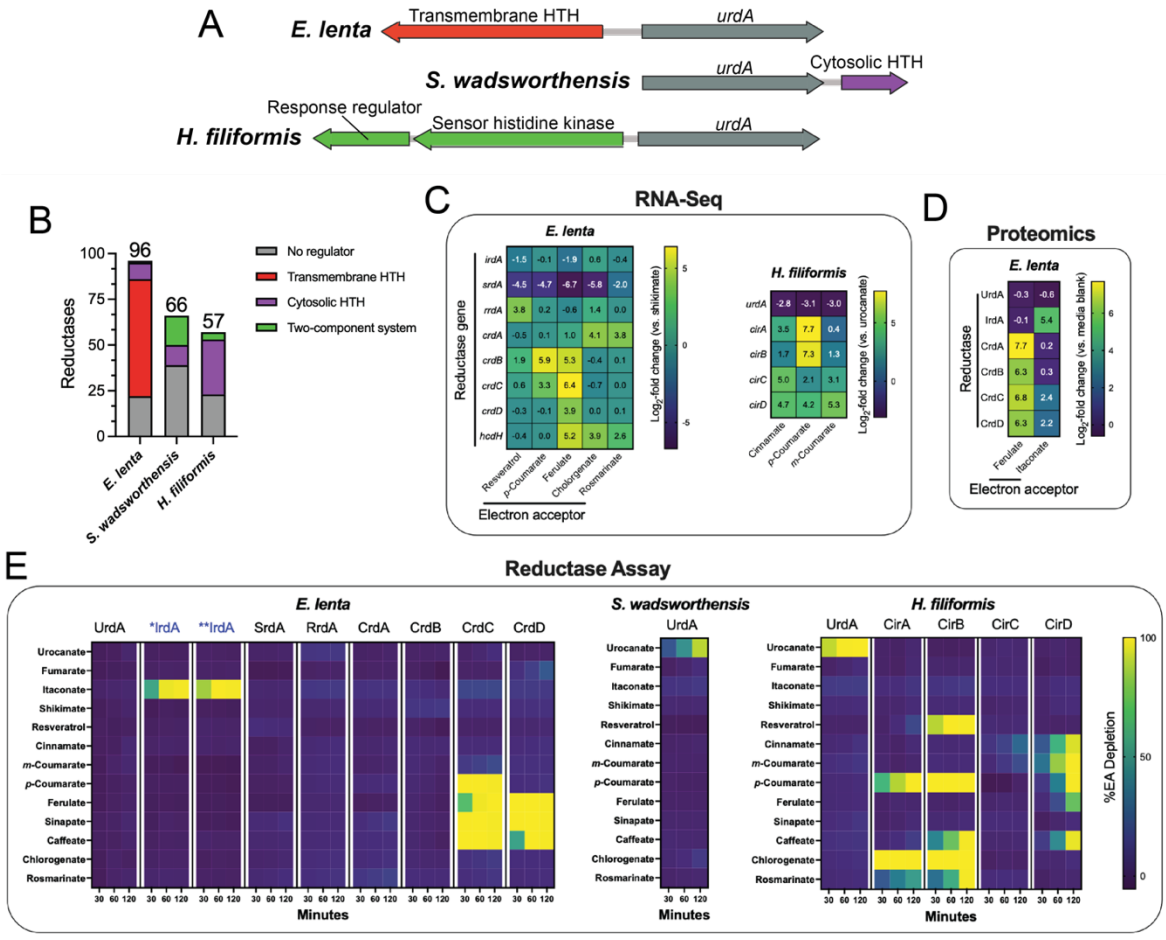


Figure 4. Flavin reductases are induced by their electron acceptors and exhibit narrow substrate specificities. (A) Genomic context of *urdA* urocanate reductases in *E. lenta* DSM2243, *S. wadsworthensis* DFI.4.78, and *H. filiformis* DFI.9.20 genomes. Genes encoding predicted transcriptional regulators are noted. (B) Genomic context of flavin and molybdopterin reductases with respect to adjacent regulatory elements. Either no clear element present (grey), two-component system (green), cytosolic regulator (purple), or transmembrane regulator (red). Cytosolic and transmembrane refers to the cellular localization of the putative signal-receiving domain in predicted transcriptional regulators that contain a DNA-binding helix-turn-helix (HTH) domain. The number of flavin and molybdopterin reductases genes that directly neighbor a predicted transcriptional regulator in *S. wadsworthensis* DFI.4.78, *H. filiformis* DFI.9.20, and *E. lenta* DSM2243 genomes. (C) RNA-Seq results from *E. lenta* DSM2243 and *H. filiformis* DFI.9.20 cells cultivated in media supplemented with indicated electron acceptors. Gene names were assigned based on induction conditions (see **Table S10**). (D) Proteomics results from *E. lenta* DSM2243 cells cultivated in media supplemented with indicated electron acceptors. (E) Activity of recombinant reductases in the presence of indicated electron acceptors. Due to poor recombinant IrdA expression, close homologs from related organisms, *IrdA and **IrdA, are included in these experiments (see text for additional detail).

As bacterial transcriptional regulators often establish autoregulatory circuits that concordantly regulate neighboring genes on the genome, the genomic colocalization of reductase genes with transcriptional regulators suggested an inducible basis of regulation. We reasoned it might be possible to leverage transcriptional responses to identified respiratory electron acceptors to gain insight into reductase substrate specificity. We chose to focus our efforts on the more experimentally tractable flavin reductases and thus excluded sulfoxide substrates, whose growth-promoting properties were inhibited by the established molybdoprotein reductase inhibitor tungstate, from further analyses (**Figure S9**).

After validating our transcriptomic approach by confirming the predicted *E. lenta* urocanate reductase was one of the most strongly induced genes in the presence of urocanate, we performed RNA-seq on *E. lenta* and *H. filiformis* in nine conditions containing a respiratory electron acceptor used by an unknown reductase (**Figure 4C**, **Table S7 & S8**). We further performed proteomic analyses of *E. lenta* grown in the presence of ferulate and itaconate to identify reductase induction at the protein level (**Figure 4D**, **Table S9**). From these experiments we identified at least one flavin reductase as among the most strongly induced genes in each condition. To facilitate downstream analyses, we assigned reductase gene names based on the condition they were induced (*crdA* for the *E. lenta* cinnamate-induced reductase A, *cirB* for the *H. filiformis* cinnamate-induced reductase B, etc.) (**Table S10**).

Transcriptomic and proteomic analyses revealed a complex pattern of reductase regulation. While itaconate, shikimate, and resveratrol each specifically induced a single flavin reductase, distinct cinnamates differentially induced multiple flavin reductases (**Figure 4C & 4D, Table S7 & S8**). For both *E. lenta* and *H. filiformis,* structurally related cinnamates resulted in elevated but variable expression of four distinct flavin reductases (**Figure 4C, Table S7 & S8**). Furthermore, in *E. lenta*, ferulate and chlorogenate, but not *p*-coumarate, induced the previously characterized *E. lenta* hydrocaffeate dehydroxylase *hcdH* (**Figure 4C & S7A**).[43] These findings thus reveal specific but sometimes complex regulatory responses of reductases to related respiratory electron acceptors.

Reductase induction conditions predict substrate specificity

To assess whether enzyme substrate specificity could be predicted based on gene expression patterns, we recombinantly produced thirteen flavin reductases identified in our transcriptomic studies. Due to poor yields of recombinant *E. lenta* itaconate-induced reductase, IrdA, we additionally produced two close IrdA orthologs from *Adlercreutzia muris* and *Berryella wangjianweii* (referred to as *IrdA and **IrdA, respectively). Purified reductases were then assayed on a panel of identified electron acceptors. We identified ten reductases that reduced at least one electron acceptor, with substrate specificity strongly tracking with induction patterns observed in the transcriptomics/proteomics studies (**Figure 4E**). We additionally observed that IrdA was part of the *E. lenta* reductase accessory pangenome (**Figure S2**) and reasoned that this attribute could further validate its functional assignment. We thus tested five *E. lenta* strains (two with *irdA* paralogs and three without) for itaconate reduction and found a one-to-one correlation between the presence of *irdA* and itaconate reduction (**Figure S10**). These studies thus establish the transcriptomic/proteomic induction-based reductase identification method, demonstrate the substrate specificity of a subset of the identified reductases, and show that reductase gene presence predicts strain activities.

Strikingly, these studies revealed marked distinctions in substrate specificity among the cinnamate reductases. While the ability of *E. lenta* and *H. filiformis* to utilize diverse cinnamates could have plausibly reflected the activity of a single promiscuous reductase, we found that both microbes encode multiple reductases with distinct specificities for different subsets of cinnamates (**Figure 4E**). Among *H. filiformis* reductases, CirA utilized *p*-coumarate, chlorogenate, and rosmarinate. CirC weakly utilized cinnamate, while CirD utilized cinnamate, *m*-coumarate, *p*-coumarate, ferulate, and caffeate. Intriguingly, while the cinnamate reductase CirB utilized *p*-coumarate, caffeate,

chlorogenate, and rosmarinate, it also utilized the alkene resveratrol. Among *E. lenta* reductases, CrdC and CrdD both utilized sinapate, ferulate and caffeate, with CrdC uniquely using *p*-coumarate. These results demonstrate that *E. lenta* and *H. filiformis* express multiple cinnamate reductases that discriminate between substrates with relatively subtle structural differences.

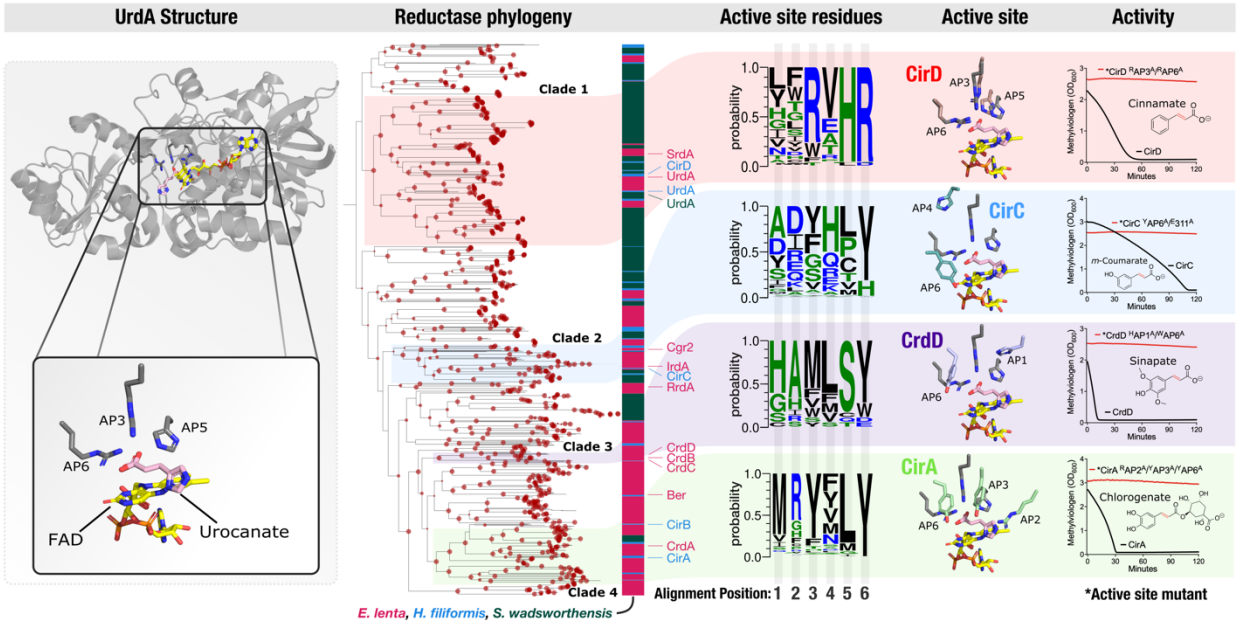


Figure 5. Independent evolutionary trajectories and distinct active sites distinguish flavin reductases with related electron acceptors. UrdA structure – crystal structure of urocanate- and FAD-bound UrdA (PDB code 6T87). Reductase phylogeny – phylogenetic tree of flavin reductases from *E. lenta*, *S. wadsworthensis*, and *H. filiformis* genomes. Bootstrap support values are indicated by the size of red dots at nodes of the tree and range from 70 to 100. Active site residues – representation of sequence identity of active site amino acids in reductase clades 1-4 scaled to frequency within the multiple sequence alignment. Positions within the multiple sequence alignment have been renumbered to active site alignment position (AP). Active site – AlphaFold models of CirD, CirC, CrdD, and CirA cinnamate reductases superimposed to the UrdA crystal structure. Activity -- reductase activity of CirD, CirC, CrdD, and CirA and active site point mutants (indicated by an asterisk). Active site mutations and alignment positions they correspond to: *CrdD H313A (AP1) and W510A (AP6); *CirA R417A (AP2), Y469A (AP3), and Y634A (AP6); *CirC E311A and Y511A (AP6); *CirD R542A (AP3) and R716A (AP6). The y-axis shows the amount of reduced methylviologen in the presence of the indicated electron acceptor.

Complex evolutionary factors relate reductase sequence and substrate specificity

We next explored the relationship between reductase evolution and substrate specificity. We reasoned that new reductase activities could be acquired either by horizontal gene transfer or gene duplication followed by sequence divergence/functional diversification and that these two scenarios would lead to distinct relationships between reductase sequence and substrate specificity. To investigate this, we constructed a phylogenetic tree using flavin reductases from *E. lenta*, *S. wadsworthensis*, and *H. filiformis* genomes (**Figure 5**). Several branches in the resulting tree include reductases from multiple species. For example, urocanate reductases from the three species are monophyletic and thus presumably share a more recent common evolutionary history (**Figure 5**). However, a majority of branches on the tree are highly sub-branched and exclusively contain reductases from a single genus. These patterns suggest that both horizontal gene transfer and gene duplication may have played significant roles in reductase evolution within high reductase per genome bacterial lineages.

The distribution of reductases with different activities in the tree suggests a convoluted evolutionary history, with cinnamate reductases providing a striking example of the complex relationship between amino acid sequence and substrate specificity. Despite catalyzing highly similar reactions, we observed that the eight *E. lenta* and *H. filiformis* reductases induced in the presence of different

cinnamate substrates separated into four phylogenetically distinct reductase clades (**Figure 5, clades 1-4**). Cinnamate reductases typically share greater than 30% amino acid sequence identity to other reductases within their clade – including reductases with distinct substrate specificities. For example, the *H. filiformis* cinnamate reductase CirD and urocanate reductase UrdA are both within reductase clade 1 and share 32% sequence identity. By contrast, cinnamate reductases from each clade share less than 26% sequence identity with cinnamate reductases from other clades, despite utilizing similar substrates (**Figure 5**). These observations provide evidence that flavin reductases with similar cinnamate substrate specificities independently evolved at least four times.

Marked active site distinctions distinguish reductases with related substrates

To clarify how different evolutionary trajectories may have independently generated reductases with similar cinnamate substrate specificities, we turned to previous mechanistic studies of flavin reductases. Fumarate and urocanate reductases were previously shown to contain two conserved active site arginines – one that forms a critical salt bridge with the substrate carboxyl group and a second proposed to facilitate catalytic proton transfer to reduce the substrate enoate group.[44–46] As a similar substrate enoate group is reduced by cinnamate reductases, we analyzed active site amino acid conservation patterns and compared the substrate-bound urocanate reductase crystal structure to AlphaFold structural models of cinnamate reductases from the four reductase clades (**Figure 5**). We found that active site arginines were conserved in reductase clade 1, including in the UrdA urocanate reductases and the CirD cinnamate reductase (**Figure 5,** alignment positions 4 and 6). By contrast, arginines were not conserved at the same positions in reductase clades 2-4. Instead, tyrosine was conserved at alignment position 6 and other active site amino acids exhibited variable, clade-specific patterns of conservation (**Figure 5**).

To test whether the distinct patterns of active site conservation in the four reductase clades reflected important functional distinctions, we generated point mutants that targeted conserved amino acids in a representative cinnamate reductase from each reductase clade (CirA, CirC, CirD, and CrdD). In each case, we found that conserved clade-specific active site amino acids were essential for activity (**Figure 5**). These results thus show that distinct active site architectures functionally distinguish the cinnamate reductase clades and suggest that parallel evolutionary processes generated flavin reductases with similar substrate specificities but significant mechanistic distinctions.

Discussion

The importance of anaerobic respiration is well established in environmental microbes but has received less attention in host-associated communities, such as the gut microbiome. Inorganic compounds tend to be the best-studied non-oxygen respiratory electron acceptors – partially due to their significance in various geochemical cycles.[6] Here we identify taxonomically distinct members of the gut microbiome that possess high numbers of respiratory-like reductases per genome. We find that these bacteria possess respiratory metabolisms and identify a number of organic respiratory electron acceptors that exhibit strain-dependent patterns of usage. These results thus establish a distinct type of respiration, defined by use of organic respiratory electron acceptors, within gut microbial communities.

In addition to the distinctive use of organic small molecule electron acceptors, a striking aspect of the respiratory metabolisms identified in our studies concerns the sheer number of respiratory-like reductases. The identified bacteria with large reductase arsenals routinely encode more than 50 respiratory-like reductases, with some even possessing more than 200 reductases (**Table S3**). This high number of reductases contrasts with previously characterized non-host associated respiratory specialists. For example, the marine and fresh-water-inhabiting bacterium *Shewanella oneidensis* is widely cited as a model organism with exceptionally broad respiratory capabilities and yet possesses “only” 22 molybdopterin and flavin reductases (**Table S2**).[47–49] The evolutionary factors underpinning

the dramatic reductase expansion in gut microbes may reflect the enzymatic diversity required to match the massive chemical diversity of the small molecule electron acceptors present in the gut. Whatever the case, the number of electron acceptors used by these bacteria underscores their distinctive respiratory metabolic strategy defined by the versatile use of organic electron acceptors within the gut.

Another insight from our studies regards the complex relationship between reductase evolution and substrate specificity. We find clear evidence that reductases that use related cinnamate substrates evolved independently at least four separate times (**Figure 5**). Furthermore, these parallel evolutionary trajectories are unlikely to be confined to cinnamate electron acceptors. We observe that *E. lenta* and *S. wadsworthensis* both utilize catechin, sulfoxides, and shikimate as electron acceptors (**Figure 3**). While *E. lenta* reductases responsible for each of these reactions can be inferred from our transcriptomics data or previously published results, *S. wadsworthensis* lacks a close ortholog of all three reductases (**Table S10**). The presently unidentified *S. wadsworthensis* reductases that utilize these electron acceptors presumably emerged from separate evolutionary trajectories. These observations thus provide evidence of considerable dynamism in reductase evolution. The complex relationship between reductase sequence and substrate specificity also highlights the challenge of predicting reductase activity solely based on sequence information. While reductases with high sequence identity tend to utilize the same electron acceptor (e.g., the *E. lenta*, *H. filiformis*, and *S. wadsworthensis* urocanate reductases), the complex relationship between reductase sequence and substrate complicates efforts to predict substrate specificity of reductases with low sequence identity to a characterized reductase.

Outside general considerations of bacterial metabolism, physiology, and evolution, perhaps the most notable aspect of the identified respiratory metabolisms concerns their implication for the gut metabolome. Effects of microbially derived metabolites represent one of the chief mechanisms by which the gut microbiome impacts host physiology. The versatile use of respiratory electron acceptors by bacteria that possess large reductase arsenals thus establishes a link between primary bacterial metabolism and the gut metabolome with direct implications for mammalian biology.

Notably, several respiratory metabolisms identified in our studies have direct links to mammalian biology. Microbiome-derived imidazole propionate, the product of urocanate reduction, impairs insulin signaling through mTORC1 and is elevated in the serum of type 2 diabetes patients.[50] Microbiome-derived hydro-*p*-coumarate (also called 4-hydroxyphenylpropionate or desaminotyrosine) protects against influenza through type I interferon activation.[4,51] Resveratrol impacts multiple enzymes and transcription factors and influences multiple metabolic and immune pathways.[40]

Itaconate is perhaps the most intriguing electron acceptor identified in our studies. In macrophages, itaconate is produced by the interferon-inducible gene Irg1 in response to bacterial and viral infection. In this context, itaconate acts as an immunometabolite that orchestrates direct and indirect mechanisms of pathogen defense.[40–42] Several intracellular bacterial pathogens counteract mammalian itaconate by degrading it to pyruvate and acetyl-CoA.[52,53] The reduction of itaconate to methylsuccinate by the *E. lenta* itaconate reductase represents a second type of microbial transformation of itaconate. The role of methylsuccinate in biology has received little previous research attention and, to our knowledge, neither itaconate nor methylsuccinate have been studied within the gut. It is thus striking that both compounds are dramatically depleted following antibiotic treatment (**Figure 3D**). Future studies will be required to address the effect of these molecules in the gastrointestinal tract.

In addition to the electron acceptors studied here, several *E. lenta* reductases with electron acceptors relevant for mammalian physiology have been previously identified. The *E. lenta* enzyme Cgr2 is a flavin reductase that reduces the cardiac medication digoxin and an unidentified molecule that drives intestinal inflammation through Th17 activation.[35,54,55] The *E. lenta* enzyme DadH is a molybdopterin reductase that dehydroxylates the neurotransmitter dopamine and inactivates the

Parkinson's disease medication L-DOPA.[13,56] Both of these previously studied reductases thus act on bioactive electron acceptors and directly impact host physiology.

Multiple studies show that reductases from bacteria with large reductase arsenals directly impact bioactive metabolites. This connection is particularly striking when considering that electron acceptors used by the vast majority of reductases in these bacteria remain unknown. Identified respiratory metabolisms may thus only scratch the surface of interactions between respiratory reductases and the gut metabolome. As such, bacteria with large reductase arsenals likely represent an attractive focus for future discovery of bioactive microbial metabolites within the gut.

Methods

Detection of molybdopterin and flavin reductases in prokaryotic genomes

Genomes and metagenome-assembled genomes of common prokaryotes in the human gut microbiome were downloaded from the version 2.0 of the Unified Human Gastrointestinal Genome (UHGG) in MGnify database on July 2022.[57] Only species representative genomes identified in more than five distinct samples (1533 total) were included in subsequent analyses. Gene prediction was done with Prodigal (version 2.6.3)[58] in single mode. Molybdopterin and flavin reductases were identified based on the PFAM accessions (PF00384 and PF00890, respectively) using the searching tool Hmmsearch (version 3.3) (E-value cut-off 0.001).[59,60] SIGNALP (version 5.0) was run to predict the putative cellular localization of the proteins using the parameter -org gram.[61]

To extend the genomic reductase content analyses to the identified taxonomic families, selected representative genomes for Burkholderiaceae, Eggerthellaceae and Erysipelotrichaceae (**Figure S1A-C**, **Table S3**) were downloaded from the Genome Taxonomy Database on October 2022 (release 202 of April 27, 2021). The same protocol as described above was performed to detect the flavin and molybdopterin reductases.

Concatenated ribosomal proteins phylogeny

Maximum-likelihood trees were calculated based on the concatenation of 14 ribosomal proteins (L2, L3, L4, L5, L6, L14, L15, L18, L22, L24, S3, S8, S17, and S19) for the UHGG tree and 16 ribosomal proteins for the three family trees (L16 and S10). Homologous protein sequences were aligned using MAFFT (version 7.467) (--auto option), and alignments refined to remove gapped regions using Trimal (version 1.4.1) (--gappyout option).[62,63] The protein alignments were concatenated and manually inspected. Phylogenetic trees were inferred using using IQ-TREE with 1000 ultrafast bootstrap replicates (-bnni -m TEST -st AA -bb 1000 -nt AUTO) (version 1.6.12), using ModelFinder to select the best model of evolution, and with 1000 ultrafast bootstrap.[64–66]

Reductase pangenomics

Pangenomes for *E. lenta*, *S. wadsworthensis*, and *H. filiformis* were calculated anvi'o 7.1 (with parameters anvi-pan-genome --mcl-inflation 10).[67] Briefly, the program (1) performed an all vs. all NCBI BLAST to create a sequence similarity network, (2) used the Markov Cluster algorithm to resolve gene clusters, and (3) was visualized with "anvi-display-pan".[68] Subsequently, the species pangenomes were subsetted for gene clusters annotated with the Pfam PF00890 using "anvi-split".

General bacterial culture conditions

All bacterial strains used in this study are listed in **Table S11**. Anaerobic growth was performed in a Coy vinyl anaerobic chamber under an atmosphere of 2-5% H_2, 2-5% CO_2, and balance nitrogen. Strains used in this study were initially grown on BHI media (BD BACTO Brain Heart Infusion

#237500) plates for 48-72 hours at 37 °C to yield visible colonies. Overnight cultures were grown for 24 hours at 37 °C using liquid BHI for *E. lenta* and *S. wadsworthensis*, or liquid BHI-NoDex (Alpha Bioscience Brain Heart Infusion W/O Dextrose #B02-114) plus 0.01% tween-80 for *H. filiformis*.

ATP assay

Bacterial cultures were grown and conditioned overnight in 10 mM urocanate to induce reductase expression. Cells were washed and resuspended in phosphate assay buffer (40 mM potassium phosphate, 10 mM magnesium sulfate, 2 mM sodium formate, pH 7.0) then allowed to incubate at room temperature for 45 minutes to achieve a resting cell suspension. To initiating the reaction, cells were mixed with buffer of buffer supplemented with 2 mM urocanate in a 96-well plate. At specified time points 10 µL of mixture was taken out and placed into a separate 96-well plate with 90 µL DMSO to quench the reaction and liberate cellular ATP. ATP concentration was measured by using an ATP determination kit (Invitrogen #A22066). Sample luminescence was read using black-well plates (Corning 96-well Black #3915) in a BioTek Cytation 5 plate reader and compared to a standard curve of known ATP concentrations.

Respiratory electron acceptor screen

A collection of potential respiratory electron acceptors were selected for testing based on: 1) their likely presence in the gut, 2) commercial availability, and 3) previous reports of their reduction by species in the Eggerthellaceae family (in the case of some phenylpropanoid derivatives).[13,69,70] Selected potential electron acceptors were directly dissolved into media at 10 mM in either 50% BHI media containing 10 mM sodium formate (*E. lenta* and *S. wadsworthensis*), or 100% BHI-NoDex with 0.01% tween-80 (*H. filiformis*) and filter sterilized. Molecules were not dissolved in solvents first to avoid potential issues of their additional usage by bacteria. Media was arrayed into deep-well (1 mL) 96-well plates (USA Scientific PlateOne #1896-1110) for growth. Strains were inoculated with normalized 1/1000 inoculations from overnight cultures, and plates were sealed with adhesive seals to minimize evaporation. Time points were taken at 24, 48, 72, and 96 hours by mixing with a multichannel pipette and moving 200 µl into a separate 96-well plate (Greiner Bio-One 96-Well Microplates #658162) and reading absorbance at OD_{600} in a BioTek Cytation 5 plate reader normalized to media blanks.

Colorimetric assay of catechol dehydroxylation

Catechol dehydroxylation assays were performed essentially as previously described.[43] Briefly, 3 mL of BHI medium supplemented with L-cysteine-HCl, and sodium formate (10 mM) were aliquoted. Seed cultures were prepared by inoculating wells, in triplicate, with either, *and E. lenta*, *S. wadsworthensis*, or *H. filiformis*. Cultures were incubated at 37°C for 18-24 hours to afford dense cultures. Next, 900 µL of medium containing 500 µM of catechol compound were aliquoted into the wells of a 96 deep-well plate. To these wells, 100 µL of dense seed culture (or sterile control) was added. Plates were then incubated at 37 °C for 72 hr. Every 24 hours, 200 uL were aspirated, sealed and centrifuged at 2000 rpm for 10 min at 4°C, and the supernatant was aspirated and transferred to a fresh 96-well plate. An aliquot (35 µL) of supernatant was then screened via the catechol colorimetric assay. Absorbance was immediately measured at 500 nm using a plate reader. The colorimetric assay for catechol dehydroxylation was based on the Arnow test.[71] Briefly, 50 µL of 0.5 M aqueous HCl was added to 50 µL of culture supernatant. After mixing, 50 µL of an aqueous solution containing both sodium molybdate and sodium nitrite (0.1 g/mL each) was added, which produced a yellow color. Finally, 50 µL of 1 M aqueous NaOH was added followed by pipetting up and down to mix. This allowed the characteristic pink color to develop. Absorbance was measured at 500 nm immediately using a microplate reader.

Metabolite mass spectrometry

Metabolites were extracted from microbial spent media following sample incubation at -80 °C for between one and twelve hours. Four volumes of methanol containing internal standards were added to each sample (one volume). Tubes were then centrifuged (-10 °C, 20,000 x *g* for 15 min) followed by transfer of 100 μL of supernatant to mass spectrometry vials. Metabolites were extracted from fecal samples by adding 80% methanol to 100 mg/mL of material and stored at -80 °C for at least one hour. Samples measured by LC-MS were then homogenized and centrifuged (-10 °C, 20,000 x *g* for 15 min) followed by transfer of 100 μL of supernatant to mass spectrometry vials. Culture supernatants measured by TMS derivatization and GCMS were dried completely under nitrogen stream at 30 °C. To dried samples, 50 μL of freshly prepared 20 mg/mL methoxamine (Sigma, 226904) in pyridine (Sigma, 270970) was added and incubated for 90 min at 30 °C and 1400 rpm. After samples were cooled to room temperature, 80 μL of derivatizing reagent (N,O-Bis(trimethylsilyl)trifluoroacetamide, 1% Trimethylchlorosilane, Sigma, B-023) and 70 μL of ethyl acetate were added and mixed in a at 70 °C for one hour at 1400 rpm. Samples were cooled to room temperature and diluted with 400 μL of ethyl acetate for analysis. Culture supernatants were analyzed using a GC-MS (Agilent 7890A GC system, Agilent 5975C MS detector) with an electron impact ionization source, a HP-5MSUI column (30 m x 0.25 mm, 0.25 μm, Agilent Technologies 19091S-433UI), and 1 μL injection. Culture supernatant data analyses were performed using MassHunter Quantitative Analysis software (version B.10, Agilent Technologies). Normalized peak areas were calculated by dividing raw peak areas of targeted analytes by averaged raw peak areas of internal standards. Aside from dehydroshikimate, all metabolite identifications from culture supernatant were confirmed by matching to authentic standard spectra and retention time and spectra in the NIST Tandem Mass Spectral Library Version 2.3. Thus, these metabolites culture supernatant are identified at confidence Level 1.[72] Shikimate reduction products were not readily available as standards; therefore, LC-MS/MS fragmentation was used to identify putative products of a shikimate reduction reaction in microbial cultures. All LC-MS analyses were performed using a Thermo Scientific Vanquish Flex UHPLC coupled with an IQ-X mass spectrometer (Thermo Fisher Scientific). To detect shikimate and its putative hydroshikimate product in *E. lenta* cultures, the chromatographic method used was an isocratic 100% mobile phase A (5% acetonitrile, 0.1% formic acid in water) for 0.2 min, followed by a gradient of 0 to 97% mobile phase B (99.9% acetonitrile, 0.1% formic acid) for 4.5 min with a wash of 100% mobile phase B for 1 min. Shikimate and its putative dehydroshikimate product were detected by mass spectrometry using negative ionization. A hydroshikimate (3,4,5-trihydroxycyclohexanecarboxylic acid) standard was not readily available, therefore, it is reported at confidence Level 3.[72] To support the assignment of the dehydroshikimate structure, analysis of the MS^2 fragmentation spectrum and a comparison to the shikimate MS^2 fragmentation spectrum has been included with putative fragment ion structures (**Fig. S11**). All metabolites from mouse fecal pellet and human stool were measured by LC-MS analyses. The chromatographic method used was an isocratic 100% mobile phase A for 0.2 min, followed by a gradient of 0 to 40% mobile phase B for 2.8 min, then a gradient from 40% to 100% mobile phase B over 1.5 min, with a final wash of 100% mobile phase B for 1.5 min. Mouse fecal pellet and human stool sample data analysis was performed using FreeStyle software (version 1.8 SP2, Thermo Scientific). Metabolite identification was established by matching accurate mass, retention time, and MS^2 fragmentation pattern from candidate ions in experimental samples to data from authentic standards of each metabolite of interest. Thus, metabolites reported meeting these criteria from mouse fecal pellet samples and human stool samples are identified at confidence Level 1.[72] Normalized peak areas were calculated by dividing raw peak areas of targeted analytes by averaged raw peak areas of internal standards. Plasma samples were processed and analyzed alongside culture supernatant samples as a biological control to evaluate metabolite extraction efficiency and instrument performance. Recovery, retention time, and coefficient of variation (CV) were calculated for internal standards. Additional method details are included in the Supplementary Information.

Collection of mouse fecal samples

All mouse experiments were performed in accordance with and approved by the Institutional Animal Care and Use Committee of the University of Chicago under protocol 72599. Six- to nine-week-old female specific pathogen-free C57BL/6 mice from Jackson Laboratories were used for all experiments. Mice were housed in sterile, autoclaved cages with irradiated food and acidified, autoclaved water upon arriving at on site mouse facility. Mouse handling and cage changes were performed by investigators wearing sterile gowns, masks, and gloves in a sterile biosafety hood. Mice were cohoused with their original shipment group until after antibiotic treatment. Fresh fecal pellets were collected directly from mice.

Following collection of “pre-antibiotic” treatment fecal pellets, mice were administered metronidazole, neomycin, and vancomycin (MNV) each at 0.25 g/L in drinking water for 3 days. The drinking water was then changed to acidified, autoclaved water and the mice were transferred to clean cages. Mice were injected intraperitoneally with clindamycin 0.4 mg total, 2 days after the end of the MNV regimen and received a cage change each day following to prevent additional antibiotic consumption through coprophagy. “Post-antibiotic” treatment fecal pellets were collected 2 days after the clindamycin injection.

Collection of human fecal samples

Fecal samples from healthy, non-antibiotic treated subjects were collected through a prospective fecal collection protocol approved by the institutional review boards at the University of Chicago (UC). All donors provided written and informed consent for IRB-approved biospecimen collection and analysis (UC Protocol IRB20-1384). The study was conducted in accordance with the Declaration of Helsinki.

Fecal samples from patients treated with broad-spectrum antibiotics were selected from samples collected as part of an ongoing prospective cohort study of hospitalized adult hepatology patients at a single institution. Inclusion criteria for the study included: age ≥18 years, ability to provide informed consent (either themselves or by proxy if decisionally incapacitated), and being treated on the hepatology consult service. Subjects who were younger than 18 years, unable to provide consent, had prior solid organ transplant, or a prior colectomy were excluded. Patients were enrolled as soon as possible upon hospital admission, most within 48 hours. All samples were obtained under a protocol that was approved by the University of Chicago IRB (IRB21-0327), and written informed consent was obtained from all participants or their surrogate decision makers.

Upon collection, fecal samples were refrigerated at +4°C for less than 24 hours prior to aliquoting and storing at -80°C until processing for shotgun metagenomic and metabolomic analysis (details below). Demographic and clinical data, including medication administration, were collected through review of the medical record and through the University of Chicago Center for Research Informatics (CRI).

The 22 samples from antibiotic-treated patients included in this study were chosen based on antibiotic administration and metagenomic findings that would be expected after broad-spectrum antibiotic exposure. The most common antibiotic exposure was a combination of intravenous (IV) vancomycin, IV cefepime, and IV metronidazole; however, one patient was also exposed to meropenem, two patients exposed to piperacillin-tazobactam, and one patient exposed to ciprofloxacin and ceftriaxone. The metagenomic changes included 8 samples with >90% relative abundance of Proteobacteria and 14 samples with >90% relative abundance of *Enterococcus*, which indicates that broad spectrum antibiotic exposure had marked effects on the gut microbiome composition in these samples.

Fecal DNA extraction

DNA was extracted using the QIAamp PowerFecal Pro DNA kit (Qiagen). Prior to extraction, samples were subjected to mechanical disruption using a bead beating method. Briefly, samples were suspended in a bead tube (Qiagen) along with lysis buffer and loaded on a bead mill homogenizer (Fisherbrand). Samples were then centrifuged, and supernatant was resuspended in a reagent that effectively removed inhibitors. DNA was then purified routinely using a spin column filter membrane and quantified using Qubit.

Metagenomic analyses of human fecal samples

Human fecal samples underwent shot-gun DNA sequencing. After undergoing mechanical disruptions with a bead beater (BioSpec Product), samples were further purified with QIAamp mini spin columns (Qiagen). Purified DNA was quantified with a Qubit 2.0 fluorometer and sequenced on the Illumina HiSeq platform, producing around 7 to 8 million PE reads per sample with read length of 150 bp. Adapters were trimmed off from the raw reads, and their quality was assessed and controlled using Trimmomatic (v.0.39), then human genome was removed by kneaddata (v0.7.10, https://github.com/biobakery/kneaddata).[73] Taxonomy was profiled using metaphlan4.[74] Alpha-diversity (a reflection of the number of unique bacterial taxa and their relative abundances) of fecal samples was estimated using Inverse Simpson Index. Raw sequencing data generated were deposited to the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under accession number PRJNA912122.

16S amplicon sequencing analyses of mouse fecal samples

The V4-V5 region within 16S rRNA gene was amplified using universal bacterial primers – 563F (5'-nnnnnnnn-NNNNNNNNNNNN-AYTGGGYDTAAA- GNG-3') and 926R (5'-nnnnnnnn-NNNNNNNNNNNN-CCGTCAATTYHT- TTRAGT-3'), where 'N' represents the barcodes, 'n' are additional nucleotides added to offset primer sequencing. Amplicons were then purified using magnetic beads, then quantified, and pooled at equimolar concentrations. Illumina sequencing-compatible Combinatorial Dual Index (CDI) adapters were ligated onto pooled amplicons using the QIAseq 1-step amplicon library kit (Qiagen). Library QC was performed using Qubit and Tapestation and sequenced on Illumina MiSeq platform to generate 2x250bp reads, generating 5000-10000 reads per sample. Raw V4-V5 16S rRNA gene sequence data is demultiplexed and processed through the dada2 pipeline into Amplicon Sequence Variants (ASVs) with minor modifications in R (v4.0.3). Specifically, reads were first trimmed at 210 bp for forward reads and 150 for reverse reads to remove low quality nucleotides. Chimeras were detected and removed using the default consensus method in the dada2 pipeline. Then, ASVs with length between 300 bp and 360 bp were kept and deemed as high quality ASVs. Taxonomy of the resultant ASVs were assigned to the genus level using the RDP classifier (v2.13) with a minimum bootstrap confidence score of 80. Species-level classification can be provided using blastn (v2.13.0) and refseq_rna database (updated 2022-06-10).

RNA extraction, library preparation, sequencing, and transcriptomics analyses

E. lenta and *H. filiformis* were grown in 50 mL of media supplemented with 10 mM formate and 10 mM of the tested electron acceptor. A BHI baseline media was used for *E. lenta* cultures. A no-dextrose BHI media supplemented with 0.01% tween-80 was utilized for *H. filiformis* cultures. Overnight cultures were diluted 1/100 into flasks, and allowed to grow shaking for 24 hours. Whole cultures were pelleted and used for RNA extraction. Total RNA from biological replicates was extracted using the Maxwell RSC instrument (Promega). RNA was quantified using Qubit, a fluorometric method and integrity was measured through a TapeStation, Agilent Technologies.

Ribosomal RNA was removed using NEBNext rRNA Depletion Kit for bacteria. Libraries from ribo-depleted samples were constructed using the NEB Ultra Directional RNA library prep kit for Illumina. Briefly, 10-500ng total RNA was subjected to ribosomal RNA depletion and rRNA-free samples were fragmented based on their RNA integrity number. Post cDNA synthesis, Illumina compatible adapters were ligated onto the inserts and final library quality was evaluated using TapeStation (Agilent technologies). Libraries were normalized using qPCR and then sequenced on Illumina's NovaSeq 6000 platform using 2x150bp read configuration.

High quality sequence reads were mapped to the NCBI reference genomes of *E. lenta* strain DSM2243, (GCF_000024265.1/), *H. filliformis* strain AF24-29 (GCF_003459085.1/), and accordingly. The reads were mapped to the respective reference genomes using Bowtie2 v2.4.5 and sorted with Samtools v1.6 using default parameters.[75,76] All sequence data were deposited to the NCBI SRA database under the bioproject ID. Read counts were generated using featureCounts v2.0.1 with the flags -p, -B, and -C.[77] Gene expression was quantified as the total number of reads uniquely aligning to the respective reference genomes, binned by annotated gene coordinates.

Differential gene expression and related quality control analyses were determined using the Bioconductor package DESEQ2 in R.[78,79] Normalization of raw read counts was performed by a scaling method implemented within DESEQ2 package, which accounts for differences in library size and library composition. Reads were also normalized for batch effects where applicable by ComBat-seq.[80] Log_2-fold change shrinkage was performed using apgelm method in DESeq2.[81] Differential expression of pairwise comparisons (of the different conditions) was assessed using the negative binomial test with a Benjamini–Hochberg false discovery rate (FDR) adjustment applied for multiple testing corrections.

Proteomics mass spectrometry

Samples for proteomic analysis were prepped using a modified version of the filter-aided sample preparation (FASP) procedure. A portion of thawed cell pellet was mixed with 500 µL 1x SDS buffer (40 mL of 1x contained: 4g glycerol, 0.67g Tris HCl, 0.68g Tris base, 0.8g SDS, and 6mg EDTA) and dithiothreitol to 20 mM. Mixture was sonicated at 20% pulsing 2-sec on 2-sec off for 5 minutes. This mixture was then heated at 95 °C for 20 minutes, followed by 37 °C for 30 minutes. Iodoacetamide was added to 60 mM and mixture was incubated in the dark for 60 minutes, before dithiothreitol was again added to 60 mM. This prepared lysate was then mixed 1:8 (25µl:200µl) with exchange buffer (8 M urea, 0.2% (w/v) deoxycholate, 1M ammonium bicarbonate pH 8) and dispensed to a filter unit. Mixture was spun at 14,000g for 10 minutes, discarding filtrate. The filter unit was washed 3 times each with 200 µL exchange buffer and spun at 14,000g for 10 minutes discarding the filtrate. The filter unit was then washed twice with 200 µL digestion buffer (0.2% (w/v) deoxycholate, 50 mM ammonium bicarbonate pH 8) and spun at 14,000g for 10 minutes discarding the filtrate. The filter unit was transferred to a passive collection tube and 100 µL digestion buffer + trypsin (~1:50 trypsin:protein) and incubated at 37 °C overnight. The following day, the tube was centrifuged at 14,000g for 10 minutes, but filtrate was not discarded. 50 µL of peptide recovery buffer (50 mM ammonium bicarbonate pH 8) was added to the filter unit and spun at 14,000g for 10 minutes, and repeated once more. The resulting three filtrates were transfer to a LoBind tube. To the filtrate, 900 µL of ethyl acetate and 2.5 µL of trifluoroacetic acid were added and vortexed. Mixture was sonicated (10 seconds at 10%), and centrifuged at 16,000g for 10 minutes. The upper organic layer was removed and discarded, taking care not to disturb the phase boundary. The addition of ethyl acetate, sonication, centrifugation, and layer removal was repeated twice more, but without the addition of trifluoroacetic acid. Sample tubes were heated uncovered to 60 °C for 5 minutes to evaporate residual ethyl acetate, then frozen at -80 °C until analysis.

Samples of trypsin-digested proteins were analyzed using a Synapt G2-Si ion mobility mass spectrometer that was equipped with a nanoelectrospray ionization source (Waters, Milford, MA). The mass spectrometer was connected in line with an Acquity M-class ultra-performance liquid

chromatography system that was equipped with trapping (Symmetry C18, inner diameter: 180 µm, length: 20 mm, particle size: 5 µm) and analytical (HSS T3, inner diameter: 75 µm, length: 250 mm, particle size: 1.8 µm) columns (Waters). The mobile phase solvents were water and acetonitrile, both of which contained 0.1% (volume/volume) formic acid. Data-independent, ion mobility-enabled, high-definition mass spectra and tandem mass spectra were acquired in the positive ion mode.[82–85]

Data acquisition was controlled using MassLynx software (version 4.1) and tryptic peptide identification and relative quantification using a label-free approach were performed using Progenesis QI for Proteomics software (version 4.0, Waters).[86] Data were searched against the *E. lenta* protein database to identify tryptic peptides.[87]

Recombinant reductase production and purification

Hexa-histidine-tagged reductase constructs were expressed in *E. coli* Rosetta cells using either the pET28a or pMCSG53 expression vectors, as previously described.[88] Briefly, expression was performed using 500 mL cultures of Luria-Bertani broth (BD LB Broth, Miller #244610) + 100 µM riboflavin, grown to an OD of 0.7-1.0 at 37 °C, then induced with 1 mM β-d-1-thiogalactopyranoside overnight at 20 °C. Kanamycin (30 mg/mL, pET28a vector) or carbenicillin (100 mg/mL, pMCSG53 vector) were used for selection. Cultures were pelleted, resuspended in lysis buffer (50 mM Tris-HCl, 300 mM NaCl, 1 mM dithiothreitol, 10 mM imidazole, pH 7.5), and incubated with 500 µg/mL lysozyme on ice for 30 minutes. Cells were ruptured by sonication (Branson Digital Sonifier SFX 250) with 40 15-second pulses between 40-second waits. Insoluble material was removed by centrifugation. Resulting supernatants were mixed with Ni-charged resin beads (Bio-Rad Profinity IMAC) to purify the His-tagged reductases. Beads were washed using lysis buffer, and bound protein was eluted using lysis buffer containing an additional 500 mM imidazole. The eluted protein was dialyzed (Slide-A-Lyzer MINI Dialysis Device, 7K MWCO) into 50 mM HEPES, 100 mM NaCl, pH 7.0 for subsequent assays.

Recombinant reductase assay

Assays were performed in an anaerobic chamber under an atmosphere of 2-5% H_2, 2-5% CO_2, and balance nitrogen. Potential electron acceptors were dried in 96-well plates (Falcon 96-well Polystyrene Microplates #351172) to allow for a 1 mM final concentration at a 200 µL reaction volume. Reaction buffer consisted of a base buffer of 50 mM HEPES and 100 mM NaCl pH 7.0, was autoclaved then immediately sparged with nitrogen and placed into the anaerobic chamber. Powders of methyl viologen (CAS #75365-73-0), sodium dithionite (CAS #7775-14-6), and FAD (CAS #146-14-5) were weighted out and placed into the anaerobic chamber 24 hours before use. All components (plastics, buffers, etc.) were placed in the anaerobic chamber for at least 24 hours prior to assay use. Powders were resuspended in the base buffer immediately before usage. Methyl viologen was resuspended to 50 mM, sodium dithionite to 25 mM, and FAD to 1 mM. The working master mix consisted of both methyl viologen and sodium dithionite added to 100 µM, which yields a deep purple color. FAD was added to 40 µM, the deep purple color should remain. Purified reductases were added to a final concentration of 5 µg/well, and the resulting mixture was multichannel pipetted into the electron acceptor containing plates. Assays of CirA, CirC, CirD, CrdD and their point mutants used reductases at 0.5, 15, 5, 0.5 µg/well, respectively. Resuspension of the dried electron acceptor molecules begins the reaction. The plates were adhesive-sealed and read in a BioTek Epoch 2 at OD_{600} for 2 hours, where a reduction in absorbance indicates a consumption of methyl viologen as an electron donor in the reaction.

Reductase phylogenetics

Genomes for *E. lenta*, *S. wadworthensis*, and *H. filiformis* were downloaded using ncbi-genome-download (https://github.com/kblin/ncbi-genome-download) (see **Table S12** for genome accessions). We then used anvi'o v7.1 to convert genome FASTA files into contigs databases (https://anvio.org/m/contigs-db) using the contigs workflow (https://doi.org/10.1186/s13059-020-02195-w), during which Prodigal v2.6.3 identified open reading frames (contigs workflow).[58,89] Next, we used the EcoPhylo workflow implemented in anvi'o (https://anvio.org/m/ecophylo) in 'tree-mode' to recover reductase genes in genomes using the Pfam model PF00890 and to explain their phylogeny. Briefly, the EcoPhylo workflow (1) used the program hmmsearch in HMMER v3.3.2[90] to identify reductases, (2) removed hits that had less than 80% model coverage to minimize the likelihood of false positives due to partial hits, (3) dereplicated resulting sequences using MMseqs2 13.45111[91] to avoid redundancy, (4) calculated a multiple sequence alignment with MUSCLE v3.8.1551,[92] trimmed the alignment by removing columns of the alignment with trimal v1.4.rev15 (with the parameters '-gappyout')[63] (5) removed sequences that have more than 50% gaps using the anvi'o program anvi-script-reformat-fasta, (6) calculated a phylogenetic tree with the final alignment with IQ-TREE 2.2.0-beta COVID-edition[93] (with the parameters '-nt AUTO -m WAG -B 1000') that resulted in a NEWICK formatted tree file, and finally (7) visualized the tree in the anvi'o interactive interface.

Reductase active site conservation analyses

AlphaFold models of CirA, CirC, CirD, and CrdD were download from Uniprot.[94] Active site amino acids were identified by independently superimposing N- and C-terminal domains of AlphaFold models to the substrate-bound urocanate reductase crystal structure (PDB code 6T87) using PyMOL v2.5.1 (http://www.pymol.org/pymol) 'align'.[46] Next, sequences in the monophyletic clade surrounding the experimentally validated sequences were subsetted from the reductase multiple sequence alignment using the program anvi-script-reformat-fasta and alignment positions were sliced in Jalview v2.11.2.5).[95] Finally, the extent of conservancy among the active site-associated residues were visualized with WebLogo 3 (https://weblogo.threeplusone.com/create.cgi).

References

1. Krautkramer, K. A., Fan, J. & Bäckhed, F. Gut microbial metabolites as multi-kingdom intermediates. *Nat. Rev. Microbiol.* **19**, 77–94 (2021).
2. Johnson, D. R., Goldschmidt, F., Lilja, E. E. & Ackermann, M. Metabolic specialization and the assembly of microbial communities. *ISME J.* **6**, 1985–91 (2012).
3. Liu, Y. *et al.* Clostridium sporogenes uses reductive Stickland metabolism in the gut to generate ATP and produce circulating metabolites. *Nat. Microbiol.* **7**, 695–706 (2022).
4. Dodd, D. *et al.* A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* **551**, 648–652 (2017).
5. Moodie, A. D. & Ingledew, W. J. Microbial anaerobic respiration. *Adv. Microb. Physiol.* **31**, 225–69 (1990).
6. Pfister, C. A. *et al.* Conceptual Exchanges for Understanding Free-Living and Host-Associated Microbiomes. *mSystems* **7**, e0137421 (2022).
7. Winter, S. E. *et al.* Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* **467**, 426–429 (2010).
8. Winter, S. E. *et al.* Host-derived nitrate boosts growth of E. coli in the inflamed gut. *Science (80-.).* **339**, 708–711 (2013).
9. Gibson, G. R., Macfarlane, G. T. & Cummings, J. H. Occurrence of sulphate-reducing bacteria in human faeces and the relationship of dissimilatory sulphate reduction to methanogenesis in the large gut. *J. Appl. Bacteriol.* **65**, 103–111 (1988).
10. Smith, N. W., Shorten, P. R., Altermann, E. H., Roy, N. C. & McNabb, W. C. Hydrogen cross-feeders of the human gastrointestinal tract. *Gut Microbes* **10**, 270–288 (2019).
11. Light, S. H. *et al.* A flavin-based extracellular electron transfer mechanism in diverse Gram-

positive bacteria. *Nature* **562**, 140–144 (2018).
12. Schubert, C. & Unden, G. C4-Dicarboxylates as Growth Substrates and Signaling Molecules for Commensal and Pathogenic Enteric Bacteria in Mammalian Intestine. *J. Bacteriol.* **204**, e0054521 (2022).
13. Rekdal, V. M. *et al.* A widely distributed metalloenzyme class enables gut microbial metabolism of host-and diet-derived catechols. *Elife* **9**, (2020).
14. Bilous, P. T., Cole, S. T., Anderson, W. F. & Weiner, J. H. Nucleotide sequence of the dmsABC operon encoding the anaerobic dimethylsulphoxide reductase of Escherichia coli. *Mol. Microbiol.* **2**, 785–95 (1988).
15. Silvestro, A., Pommier, J., Pascal, M. C. & Giordano, G. The inducible trimethylamine N-oxide reductase of Escherichia coli K12: its localization and inducers. *Biochim. Biophys. Acta* **999**, 208–16 (1989).
16. Heinzinger, N. K., Fujimoto, S. Y., Clark, M. A., Moreno, M. S. & Barrett, E. L. Sequence analysis of the phs operon in Salmonella typhimurium and the contribution of thiosulfate reduction to anaerobic energy metabolism. *J. Bacteriol.* **177**, 2813–20 (1995).
17. Hensel, M., Hinsley, A. P., Nikolaus, T., Sawers, G. & Berks, B. C. The genetic basis of tetrathionate respiration in Salmonella typhimurium. *Mol. Microbiol.* **32**, 275–87 (1999).
18. Cruz-García, C., Murray, A. E., Klappenbach, J. A., Stewart, V. & Tiedje, J. M. Respiratory nitrate ammonification by Shewanella oneidensis MR-1. *J. Bacteriol.* **189**, 656–62 (2007).
19. Krafft, T. *et al.* Cloning and nucleotide sequence of the psrA gene of Wolinella succinogenes polysulphide reductase. *Eur. J. Biochem.* **206**, 503–10 (1992).
20. Saltikov, C. W. & Newman, D. K. Genetic identification of a respiratory arsenate reductase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10983–8 (2003).
21. Krafft, T., Bowen, A., Theis, F. & Macy, J. M. Cloning and sequencing of the genes encoding the periplasmic-cytochrome B-containing selenate reductase of Thauera selenatis. *DNA Seq.* **10**, 365–77 (2000).
22. Bender, K. S. *et al.* Identification, characterization, and classification of genes encoding perchlorate reductase. *J. Bacteriol.* **187**, 5090–6 (2005).
23. McPherson, M. J., Baron, A. J., Pappin, D. J. & Wootton, J. C. Respiratory nitrate reductase of Escherichia coli. Sequence identification of the large subunit gene. *FEBS Lett.* **177**, 260–4 (1984).
24. Lledó, B., Martínez-Espinosa, R. M., Marhuenda-Egea, F. C. & Bonete, M. J. Respiratory nitrate reductase from haloarchaeon Haloferax mediterranei: biochemical and genetic analysis. *Biochim. Biophys. Acta* **1674**, 50–9 (2004).
25. Yamazaki, C. *et al.* A novel dimethylsulfoxide reductase family of molybdenum enzyme, Idr, is involved in iodate respiration by Pseudomonas sp. SCT. *Environ. Microbiol.* **22**, 2196–2212 (2020).
26. Cole, S. T. Nucleotide sequence coding for the flavoprotein subunit of the fumarate reductase of Escherichia coli. *Eur. J. Biochem.* **122**, 479–84 (1982).
27. Light, S. H. *et al.* Extracellular electron transfer powers flavinylated extracellular reductases in Gram-positive bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 26892–26899 (2019).
28. Mikoulinskaia, O., Akimenko, V., Galouchko, A., Thauer, R. K. & Hedderich, R. Cytochrome c-dependent methacrylate reductase from Geobacter sulfurreducens AM-1. *Eur. J. Biochem.* **263**, 346–352 (1999).
29. Bogachev, A. V., Bertsova, Y. V., Bloch, D. A. & Verkhovsky, M. I. Urocanate reductase: Identification of a novel anaerobic respiratory pathway in Shewanella oneidensis MR-1. *Mol. Microbiol.* **86**, 1452–1463 (2012).
30. Speich, N. *et al.* Adenylylsulphate reductase from the sulphate-reducing archaeon Archaeoglobus fulgidus: cloning and characterization of the genes and comparison of the enzyme with other iron-sulphur flavoproteins. *Microbiology* **140 (Pt 6**, 1273–84 (1994).
31. Méheust, R., Huang, S., Rivera-Lugo, R., Banfield, J. F. & Light, S. H. Post-translational flavinylation is associated with diverse extracytosolic redox functionalities throughout bacterial life. *Elife* **10**, 1–22 (2021).
32. Jardim-Messeder, D. *et al.* Fumarate reductase superfamily: A diverse group of enzymes whose evolution is correlated to the establishment of different metabolic pathways. *Mitochondrion* **34**, 56–66 (2017).
33. Le, C. (Chip), 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).
34. Ize, B. *et al.* Remnant signal peptides on non-exported enzymes: Implications for the evolution of prokaryotic respiratory chains. *Microbiology* **155**, 3992–4004 (2009).
35. Alexander, M. *et al.* Human gut bacterial metabolism drives Th17 activation and colitis. *Cell Host Microbe* **30**, 17-30.e9 (2022).
36. Kraal, L., Abubucker, S., Kota, K., Fischbach, M. A. & Mitreva, M. The prevalence of species and strains in the human microbiome: A resource for experimental efforts. *PLoS One* **9**, 1–11 (2014).
37. Clifford, M. N. Chlorogenic acids and other cinnamates - nature, occurrence, dietary burden, absorption and metabolism. *J. Sci. Food Agric.* **80**, 1033–1043 (2000).
38. Rivera-Chávez, F. & Bäumler, A. J. The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. *Annu. Rev. Microbiol.* **69**, 31–48 (2015).
39. Asano, Y. *et al.* Critical role of gut microbiota in the production of biologically active, free catecholamines in the gut lumen of mice. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **303**, 1288–1295 (2012).
40. Diaz-Gerevini, G. T. *et al.* Beneficial action of resveratrol: How and why? *Nutrition* **32**, 174–178 (2016).
41. Bentley, R. & Haslam, E. The Shikimate Pathway — A Metabolic Tree with Many Branche. *Crit. Rev. Biochem. Mol. Biol.* **25**, 307–384 (1990).
42. Michelucci, A. *et al.* Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 7820–7825 (2013).
43. Maini Rekdal, V. *et al.* A widely distributed metalloenzyme class enables gut microbial metabolism of host- and diet-derived catechols. *Elife* **9**, (2020).
44. Leys, D. *et al.* Structure and mechanism of the flavocytochrome c fumarate reductase of Shewanella putrefaciens MR-1. *Nat. Struct. Biol.* **6**, 1113–1117 (1999).
45. Pankhurst, K. L. *et al.* A proton delivery pathway in the soluble fumarate reductase from Shewanella frigidimarina. *J. Biol. Chem.* **281**, 20589–97 (2006).
46. Venskutonytė, R. *et al.* Structural characterization of the microbial enzyme urocanate reductase mediating imidazole propionate production. *Nat. Commun.* **12**, 1347 (2021).
47. Heidelberg, J. F. *et al.* Genome sequence of the dissimilatory metal ion–reducing bacterium Shewanella oneidensis. *Nat. Biotechnol.* **20**, 1118–1123 (2002).
48. Hau, H. H. & Gralnick, J. A. Ecology and Biotechnology of the Genus *Shewanella*. *Annu. Rev. Microbiol.* **61**, 237–258 (2007).
49. Ikeda, S. *et al.* Shewanella oneidensis MR-1 as a bacterial platform for electro-biotechnology. *Essays Biochem.* **65**, 355–364 (2021).
50. Koh, A. *et al.* Microbially Produced Imidazole Propionate Impairs Insulin Signaling through mTORC1. *Cell* **175**, 947-961.e17 (2018).
51. Steed, A. L. *et al.* The microbial metabolite desaminotyrosine protects from influenza through type I interferon. *Science (80-.).* **357**, 498–502 (2017).
52. Sasikaran, J., Ziemski, M., Zadora, P. K., Fleig, A. & Berg, I. A. Bacterial itaconate degradation promotes pathogenicity. *Nat. Chem. Biol.* **10**, 371–7 (2014).
53. Wang, H. *et al.* An essential bifunctional enzyme in Mycobacterium tuberculosis for itaconate dissimilation and leucine catabolism. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15907–15913 (2019).
54. Haiser, H. J. *et al.* Predicting and Manipulating Cardiac Drug Inactivation by the Human Gut Bacterium Eggerthella lenta. *Science (80-.).* **341**, 295–298 (2013).
55. 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**, 1–32 (2018).
56. 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**, (2019).
57. Almeida, A. *et al.* A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).
58. Hyatt, D. *et al.* Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* (2010). doi:10.1186/1471-2105-11-119
59. Mistry, J. *et al.* Pfam: The protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419 (2021).

60. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* (1998). doi:10.1093/bioinformatics/14.9.755
61. Almagro Armenteros, J. J. *et al.* SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**, 420–423 (2019).
62. Katoh, K. & Standley, D. M. A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* (2016). doi:10.1093/bioinformatics/btw108
63. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* (2009). doi:10.1093/bioinformatics/btp348
64. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* (2015). doi:10.1093/molbev/msu300
65. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Von Haeseler, A. & Jermiin, L. S. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* (2017). doi:10.1038/nmeth.4285
66. Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the ultrafast bootstrap approximation. *Mol. Biol. Evol.* (2018). doi:10.1093/molbev/msx281
67. Delmont, T. O. & Eren, A. M. Linking pangenomes and metagenomes: the Prochlorococcus metapangenome. *PeerJ* **6**, e4320 (2018).
68. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
69. 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).
70. García-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).
71. Arnow, L. E. COLORIMETRIC DETERMINATION OF THE COMPONENTS OF 3,4-DIHYDROXYPHENYLALANINETYROSINE MIXTURES. *J. Biol. Chem.* **118**, 531–537 (1937).
72. Sumner, L. W. *et al.* Proposed minimum reporting standards for chemical analysis. *Metabolomics* **3**, 211–221 (2007).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Blanco-Miguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species with MetaPhlAn 4. *bioRxiv* 2022.08.22.504593 (2022). doi:10.1101/2022.08.22.504593
75. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
76. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, (2021).
77. 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).
78. 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).
79. Team, R. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. *http://www.R-project.org/* (2016).
80. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma.* **2**, (2020).
81. Zhu, A., Ibrahim, J. G. & Love, M. I. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics* **35**, 2084–2092 (2019).
82. Plumb, R. S. *et al.* UPLC/MS(E); a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Commun. Mass Spectrom.* **20**, 1989–94 (2006).
83. Shliaha, P. V, Bond, N. J., Gatto, L. & Lilley, K. S. Effects of traveling wave ion mobility separation on data independent acquisition in proteomics studies. *J. Proteome Res.* **12**, 2323–39 (2013).
84. Helm, D. *et al.* Ion mobility tandem mass spectrometry enhances performance of bottom-up proteomics. *Mol. Cell. Proteomics* **13**, 3709–15 (2014).
85. Distler, U. *et al.* Drift time-specific collision energies enable deep-coverage data-independent acquisition proteomics. *Nat. Methods* **11**, 167–70 (2014).

86. Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).
87. Sayers, E. W. *et al.* Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).
88. Rivera-Lugo, R. *et al.* Distinct energy-coupling factor transporter subunits enable flavin acquisition and extracytosolic trafficking for extracellular electron transfer in Listeria monocytogenes *bioRxiv* 2022.11.04.515267 (2022). doi:10.1101/2022.11.04.515267
89. Eren, A. M. *et al.* Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2020).
90. Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).
91. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* (2017). doi:10.1038/nbt.3988
92. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
93. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
94. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
95. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).

Acknowledgments
We thank Huaiying Lin & Nicholas P. Dylla for assistance with data analyses and Laurie Comstock for helpful feedback. Research reported in this publication was supported by funding from the National Institutes of Health (T32DK007074 to M.A.O., 1S10OD020062-01 to A.T.I., and K22AI144031 & R35GM146969 to S.H.L) and the Searle Scholars Program (to S.H.L).

Supplemental material for:

Exceptionally versatile respiratory metabolisms drive metabolite production by diverse gut bacteria

Alexander S. Little,[1,2] Isaac T. Younker,[1,2] Matthew S. Schechter,[1,2] Raphaël Méheust,[3] Joshua Stemczynski,[1,2] Kaylie Scorza,[1,2] Paola Nol Bernardino,[1,2] Michael W. Mullowney,[1] Deepti Sharan,[1] Emily Waligurski,[1] Rita Smith,[1] Ramanujam Ramanswamy,[1] William Leiter,[1] David Moran,[1] Mary McMillin,[1] Matthew A. Odenwald,[4] Anthony T. Iavarone,[5] Ashley M. Sidebottom,[1] Anitha Sundararajan,[1] Eric G. Pamer,[1,2,6] A. Murat Eren,[7,8] Samuel H. Light[1,2]*

[1]Duchossois Family Institute, University of Chicago, Chicago, IL, USA
[2]Department of Microbiology, University of Chicago, Chicago, IL, USA
[3]Génomique Métabolique, CEA, Genoscope, Institut François Jacob, Université d'Évry, Université Paris-Saclay, CNRS, Evry, France
[4]Department of Medicine, Section of Gastroenterology, Hepatology, and Nutrition, University of Chicago
[5]QB3/Chemistry Mass Spectrometry Facility, University of California, Berkeley, Berkeley, CA, USA
[6]Department of Medicine, Section of Infectious Diseases & Global Health, University of Chicago Medicine, Chicago, IL, USA.
[7]Helmholtz Institute for Functional Marine Biodiversity, Oldenburg, Germany
[8]Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenbug, Germany
*Address correspondence to Samuel H. Light, samlight@uchicago.edu

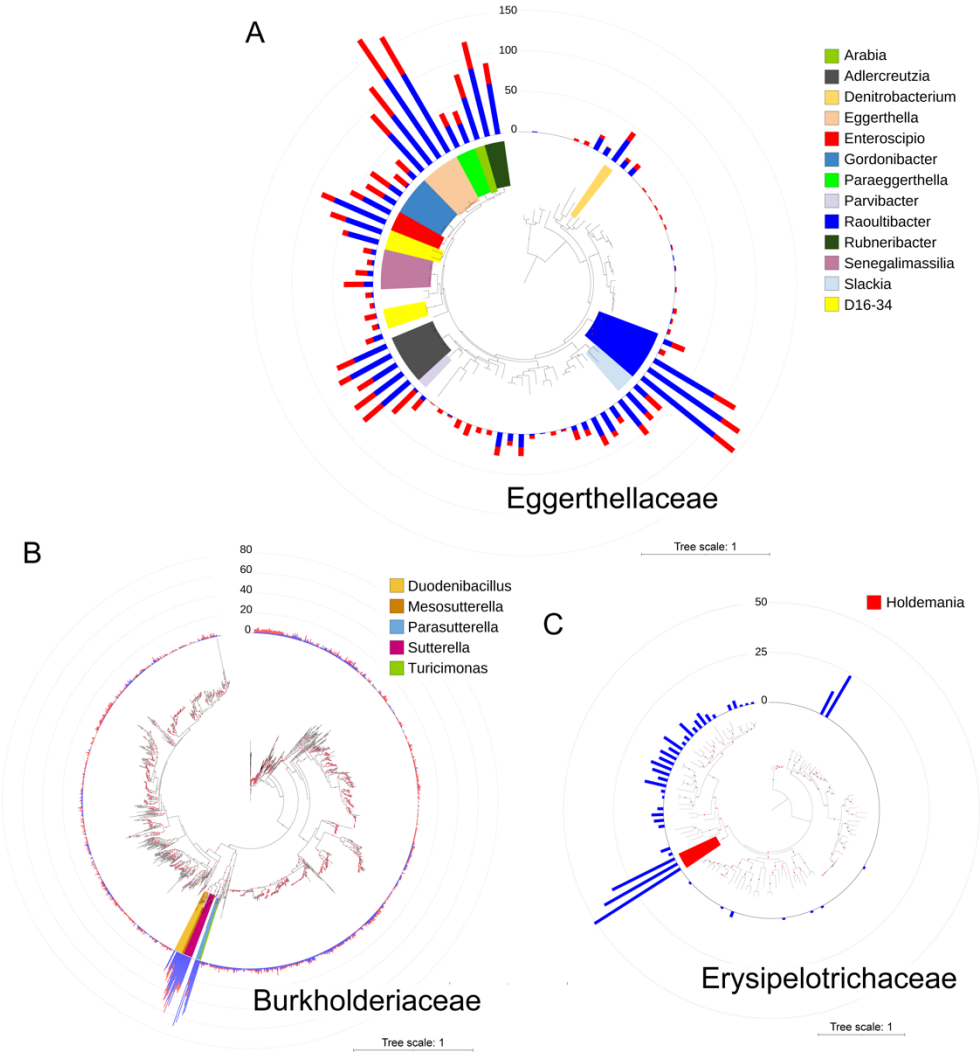


Figure S1. The distribution of reductases per genome in three taxonomic families. Phylogenetic trees constructed with representative genomes from each Genome Taxonomy Database (GTDB) species in (A) Eggerthellaceae (88 genomes, 2387 amino acid sites), (B) Burkholderiaceae (1510 genomes, 2379 amino acid sites), and (C) Erysipelotrichaceae (116 genomes, 2464 amino acid sites)**.** Each maximum likelihood tree was constructed based on a concatenated alignment of 16 ribosomal proteins under an LG + I + G4 model of evolution. The numbers of flavin (blue) and molybdopterin (red) reductases with a computationally predicted signal peptide in each genome are graphed on the outer ring of the trees.

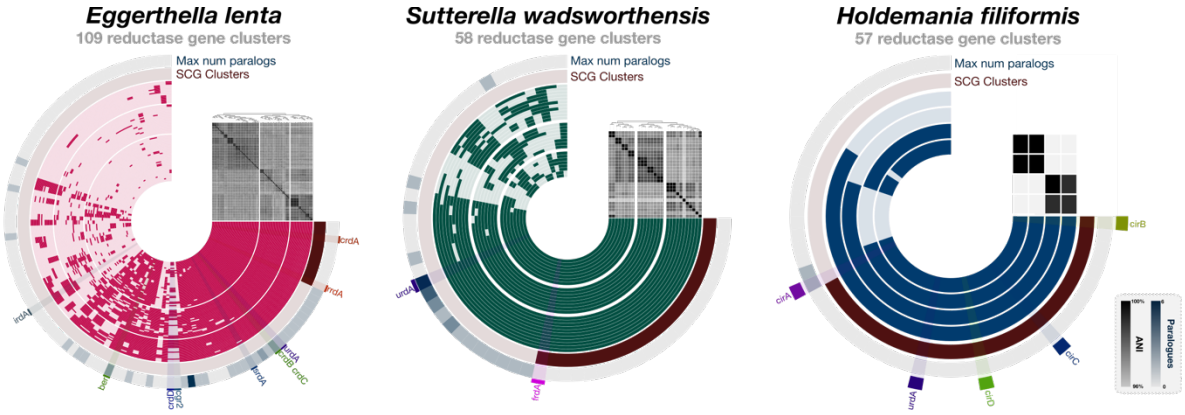


Figure S2. Flavin reductase pangenomes of *Eggerthella lenta*, *Sutterella wadsworthensis*, and *Holdemania filiformis*. The inner rings represent unique genomes. The radial elements of each pangenome represent gene cluster presence (darker color) or absence (lighter color) across the genomes. The outermost concentric circle, "Max num paralogues," indicates the maximum number of paralogues (defined as reductases with ~60% sequence identity) one genome contributes to the gene cluster. The second outermost circle, "SCG clusters," indicates single-copy core reductase – i.e., gene clusters for which every genome contributed exactly one gene.

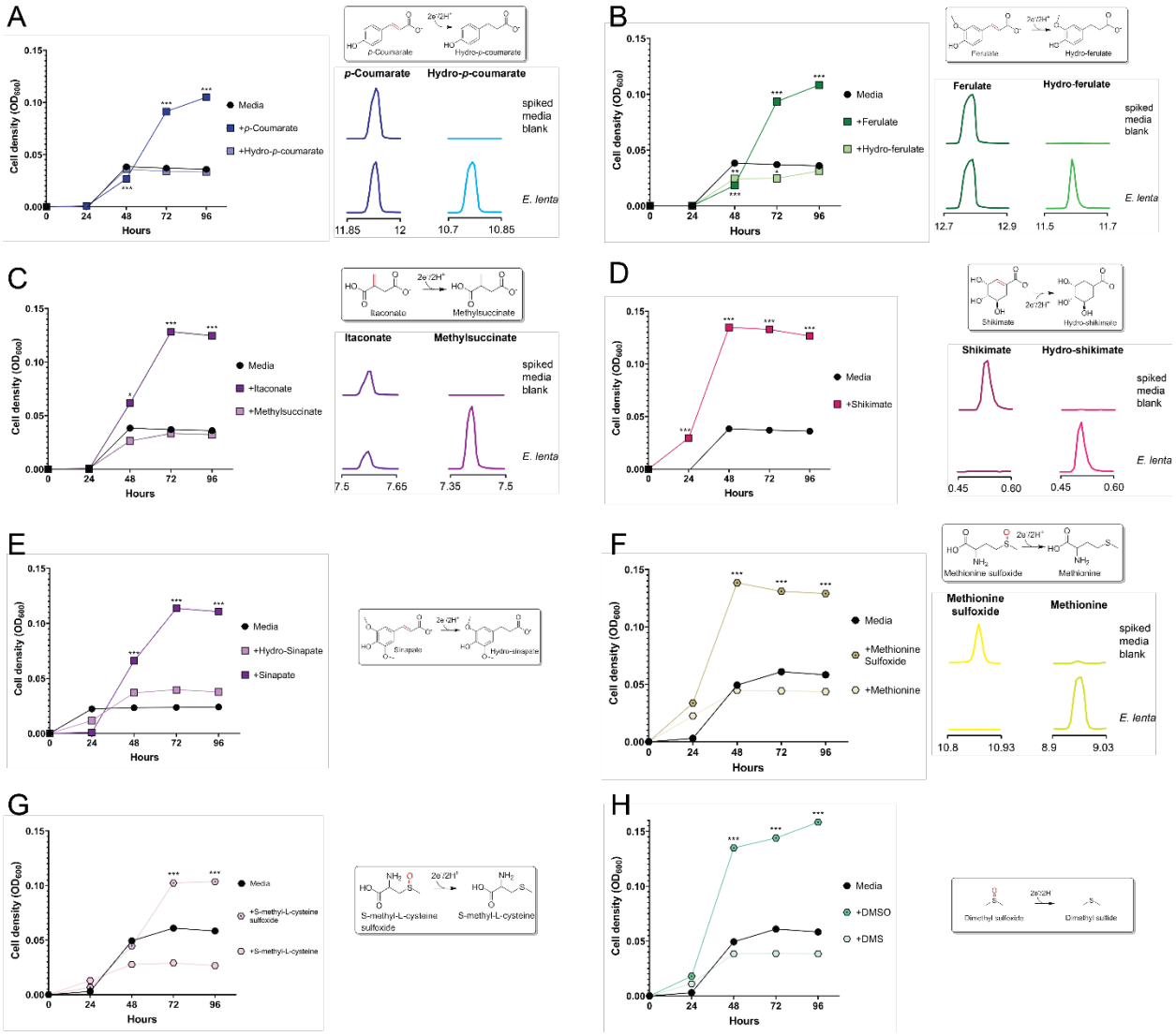


Figure S3. *E. lenta* uses multiple respiratory electron acceptors. *E. lenta* DSM2243 growth in formate-supplemented media provisioned with electron acceptors: (A) *p*-coumarate, (B) ferulate, (C) itaconate, (D) shikimate, (E) sinapate, (F) methionine sulfoxide, (G) S-methyl-L-cysteine sulfoxide, and (H) dimethyl sulfoxide. A 'no electron acceptor' condition and conditions with predicted reduction products are included as controls. GC- or LC-MS analyses show peaks for electron acceptors and their reduced products in uninoculated and inoculated growth media. (A), (B), (C), (E) were measured by GC-MS; (D) was measured by LC-MS; support for identification of hydro-shikimate in (D) is provided in **Figure S11**. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

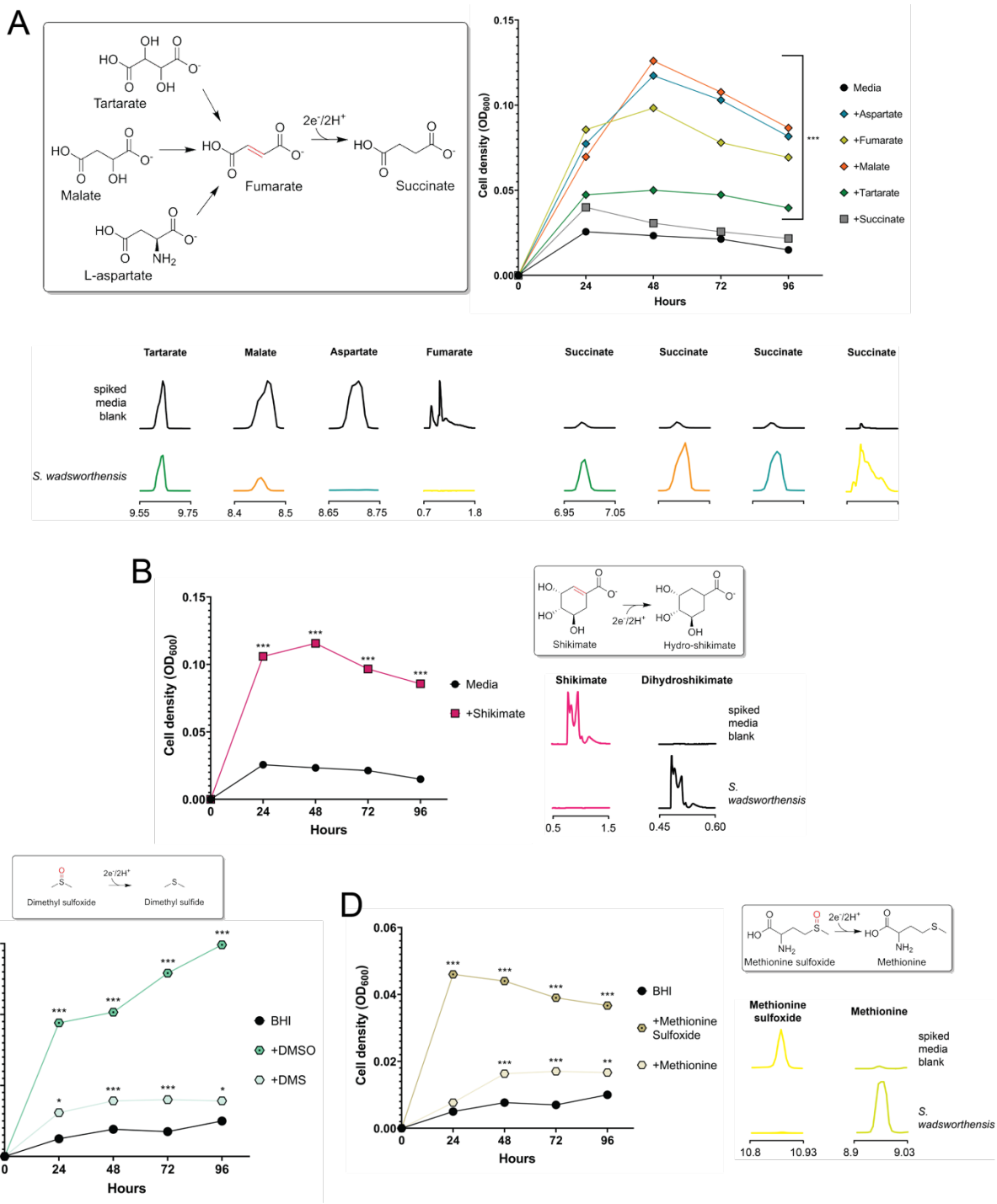


Figure S4. *S. wadsworthensis* growth-stimulating electron acceptors. *S. wadsworthensis* growth in formate-supplemented media provisioned with electron acceptors: (A) labeled C4-dicarboxylates, (B) shikimate, (C) dimethyl sulfoxide, and (D) methionine sulfoxide. A 'no electron acceptor' condition and conditions with predicted products are included as controls. GC- or LC-MS analyses show peaks for electron acceptors and their reduced products in uninoculated and inoculated growth media. (A) fumarate and succinate were measured by LC-MS, all others measured by GC-MS; (B) was measured by LC-MS; support for identification of hydro-shikimate in (D) is provided in **Figure S13**. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

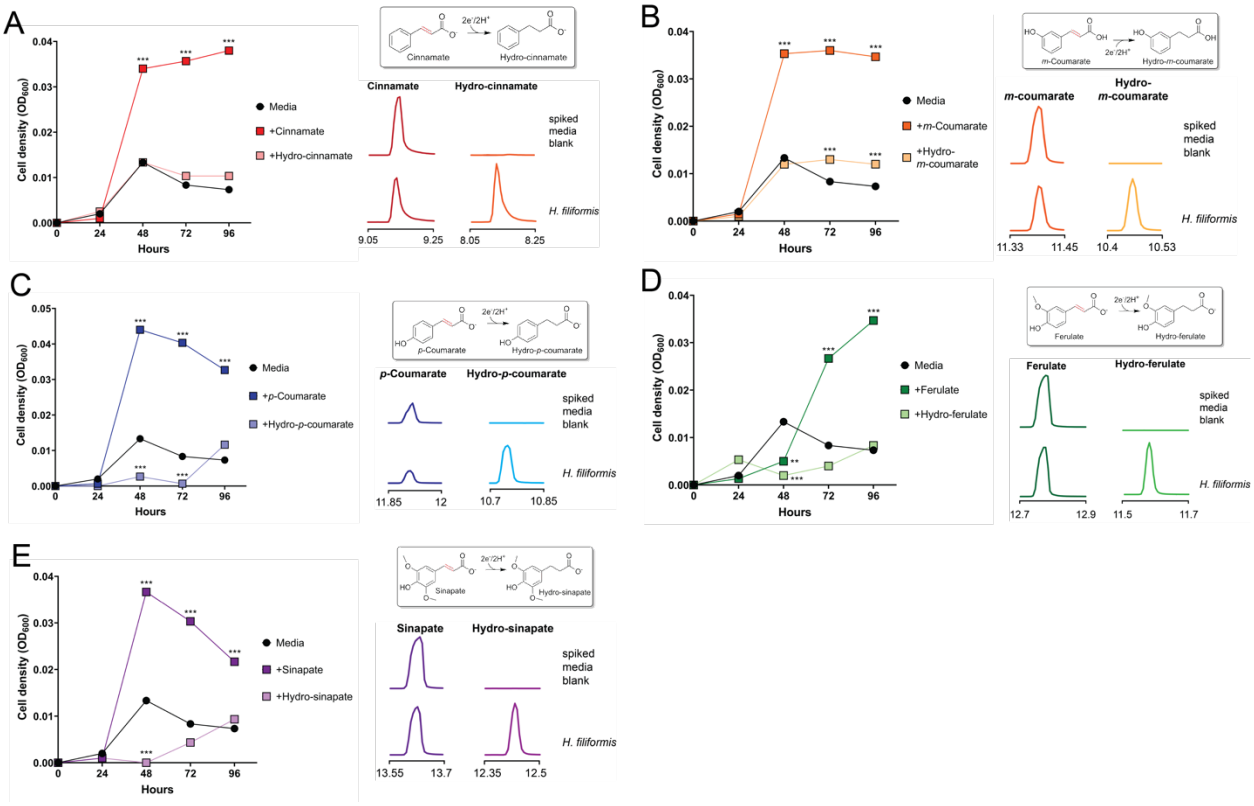


Figure S5. *H. filiformis* growth-stimulating electron acceptors. *H. filiformis* growth in media provisioned with electron acceptors: (A) cinnamate (B) *m*-coumarate, (C) *p*-coumarate (D) ferulate, (E) sinapate. A 'no electron acceptor' condition and conditions with predicted products are included as controls. GC-MS analyses show peaks for electron acceptors and their reduced products in uninoculated and inoculated growth media. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

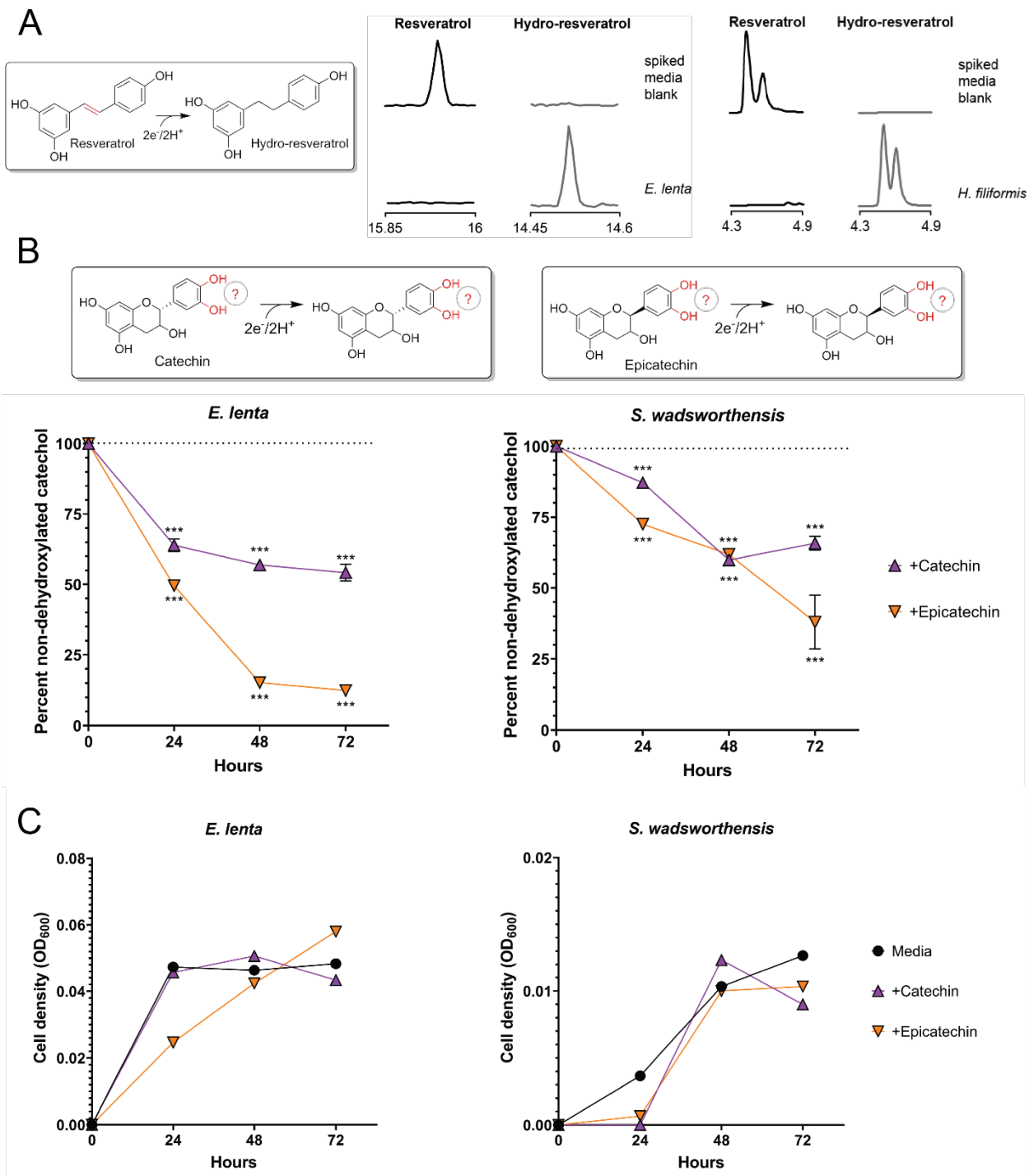


Figure S6. Electron acceptors not observed to stimulate growth. (A) GC-MS analysis of resveratrol-spiked media before and after *E. lenta* DSM2243 growth and LC-MS analysis of resveratrol-spiked media before and after H. filiformis DFI.9.20 growth. The low solubility of resveratrol hindered experiments to assess whether this electron acceptor supported respiratory growth. (B) Enzymatic assay of catechin and epicatechin dehydroxylation by *E. lenta* DSM2243 and *S. wadsworthensis* DFI.4.78 strains. This assay tracks dehydroxylation and doesn't distinguish between the two possible dehydroxylation products. (C) Growth of *E. lenta* DSM2243 and *S. wadsworthensis* DFI.4.78 in the presence of catechin and epicatechin. Data are mean ±SD (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA.

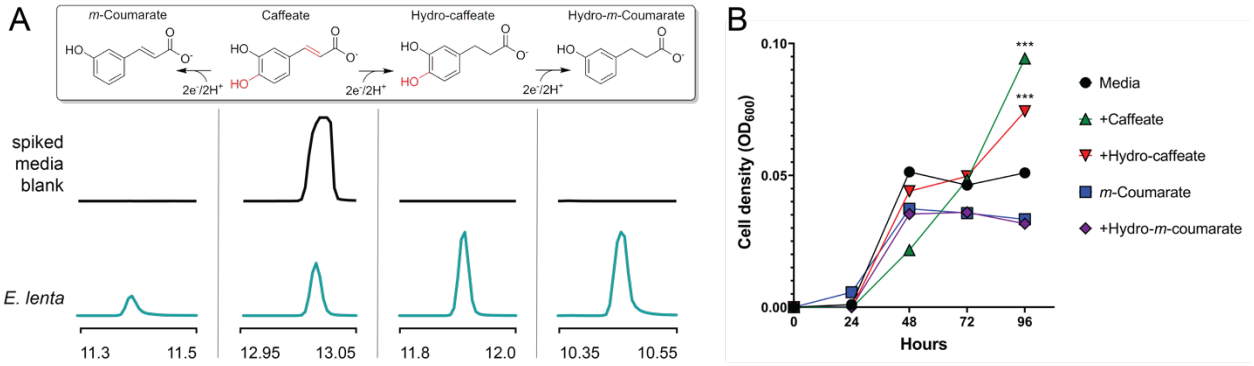


Figure S7. Caffeate utilization by *E. lenta*. (A) GC-MS analysis of supernatant collected from *E. lenta* DSM2243 grown in caffeate-spiked media. Proposed reaction pathways are shown, with peaks for each compound provided beneath its chemical structure. The previously characterized hydrocaffeate dehydroxylase may catalyze the observed dehydroxylation reactions. The observed caffeate reduction to hydrocaffeate provides evidence of a caffeate reductase, while the accumulation of *m*-coumarate suggests that this enzyme may specifically use caffeate. (B) *E. lenta* DSM2243 growth in media supplemented with formate and different cinnamates. The pattern of cinnamate-dependent growth enhancement supports the conclusions that: (1) dehydroxylation can support respiratory growth and (2) *m*-coumarate is a poor electron acceptor for *E. lenta*. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

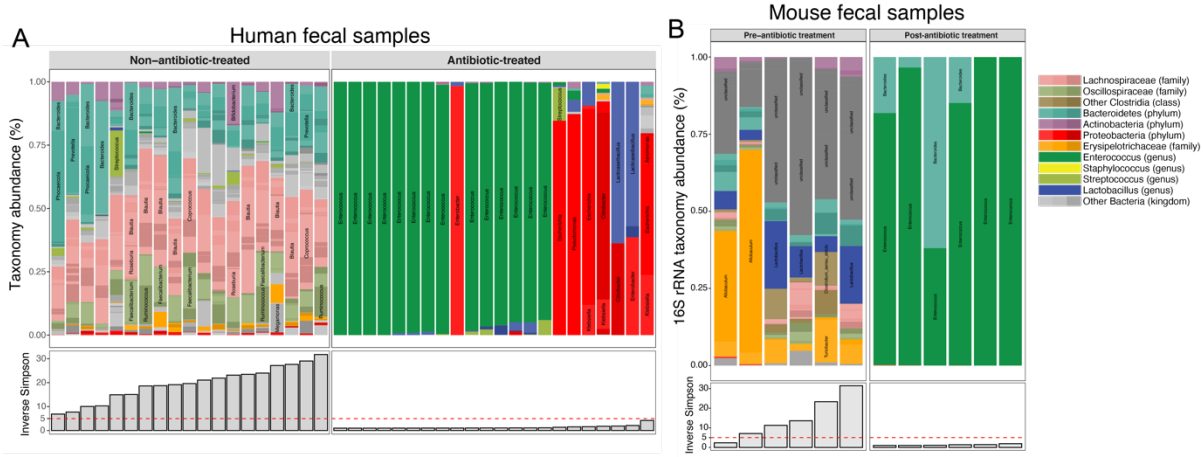


Figure S8. Microbiome composition of fecal samples used for metabolite measurements. (A) Taxonomy abundance in human fecal samples used for metabolomics analyses assessed by shotgun metagenomics. (B) Taxonomy abundance in mouse fecal samples used for metabolomics analyses assessed by 16S rRNA amplicon sequencing. The Inverse Simpson measure of microbiome diversity is also presented.

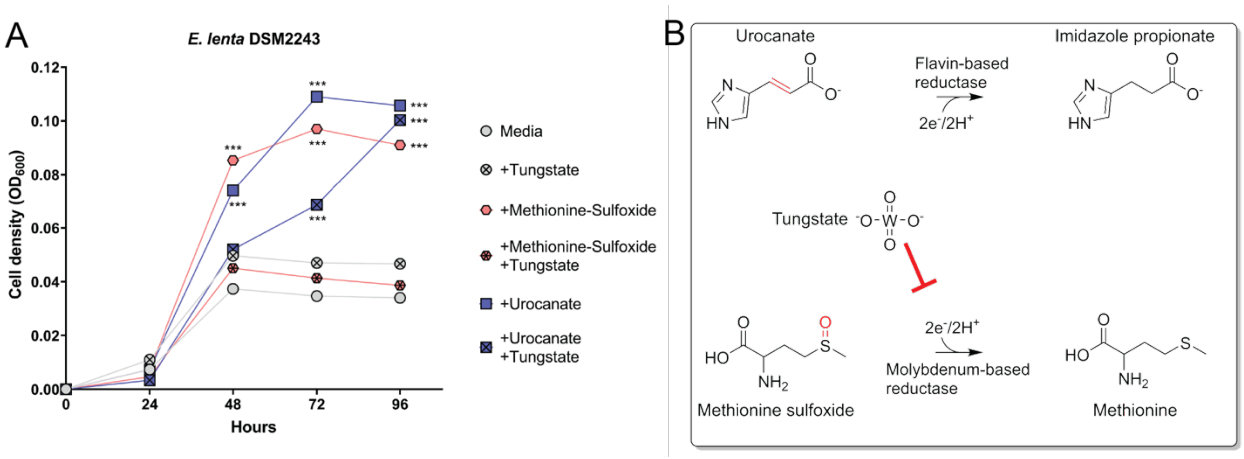


Figure S9. Tungstate inhibits the use of sulfoxide electron acceptors. (A) The effect of the molybdopterin reductase inhibitor, tungstate, on *E. lenta* DSM2243 growth. Media was supplemented with formate and the noted electron acceptor. (B) Reactions catalyzed by urocanate and sulfoxide reductases. Tungstate's selective growth inhibition is consistent with sulfoxide, but not urocanate, reduction being catalyzed by a molybdopterin reductase. Data are mean (n = 3). *p < 0.05, **p < 0.01, *** p < 0.001. Two-way ANOVA, multiple test vs media alone.

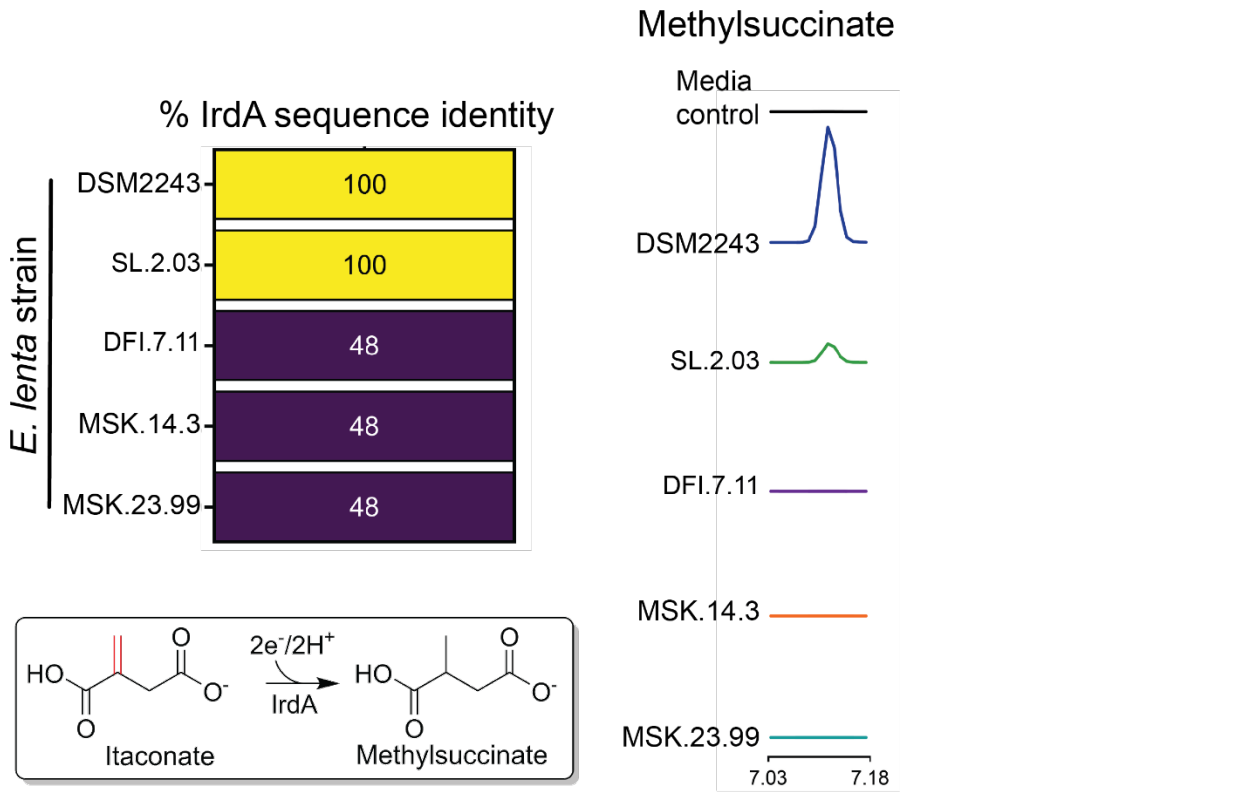
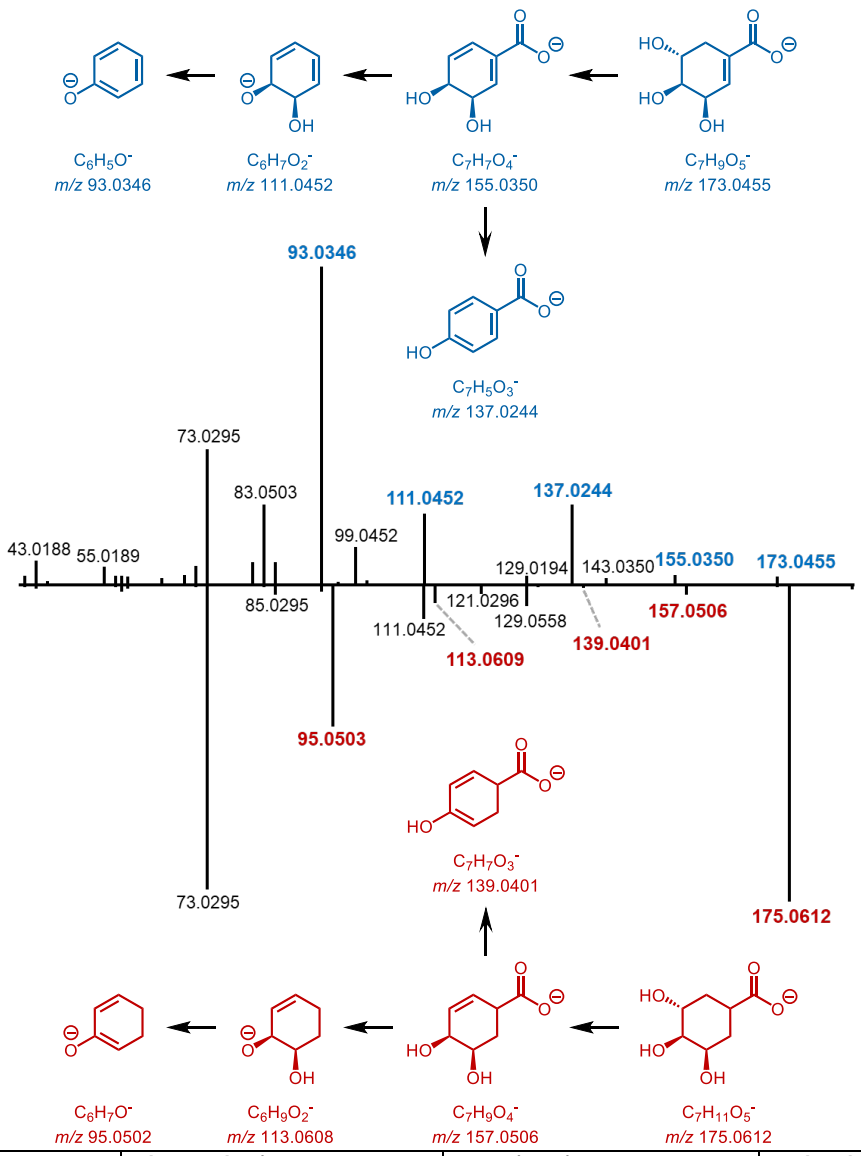


Figure S10. Presence of *irdA* predicts itaconate reductase activity of *E. lenta* strains. The sequence identity of the reductase with greatest similarity to IrdA is shown for indicated *E. lenta* strains strain. The methylsuccinate LC-MS peak of media collected after cultivation with the indicated strain.



theoretical *m/z*	observed *m/z*	error (ppm)	molecular formula
shikimate			
173.0455	173.0455	0.00	C7H9O5-
155.0350	155.0350	0.00	C7H7O4-
137.0244	137.0244	0.00	C7H5O3-
111.0452	111.0452	0.00	C6H7O2-
93.0346	93.0346	0.00	C6H5O-
hydro-shikimate			
175.0612	175.0612	0.00	C7H11O5-
157.0506	157.0506	0.00	C7H9O4-
139.0401	139.0401	0.00	C7H7O3-
113.0608	113.0609	0.88	C6H9O2-
95.0502	95.0503	1.05	C6H7O-

Figure S11. Butterfly plot of LCMS-MS fragmentation spectra for shikimate at top with key putative fragment ions highlighted in blue and hydro-shikimate at bottom with key putative fragment ions highlighted in red. The difference in two Dalton between the key putative fragments in these spectra support the assignment of m/z 175.0612 as hydro-shikimic acid as shown.

Supplementary Methods

Measurement of substrate and reduced products using gas chromatography–mass spectrometry (GC-MS) and liquid chromatography-mass spectrometry (LC-MS) metabolomics.

Metabolite Extraction from Liquid Culture. Samples were incubated at -80 ºC for between one and twelve hours. Four volumes of methanol containing internal standards (d^7,^{15}N-proline, Cambridge Isotopes, DNLM-7562 and U-^{13}C-palmitic acid, Cambridge Isotopes, CLM-409) were added to each culture supernatant (one volume) in a microcentrifuge tube. Tubes were then centrifuged at -10 ºC, 20,000 x *g* for 15 min followed by transfer of 100 μL of supernatant to pre-labeled mass spectrometer autosampler vials (Microliter, 09-1200).

Metabolite Extraction from Fecal Material. Metabolites were extracted from fecal samples by adding 80% methanol (containing $^{13}C_9$-phenylalanine, Cambridge Isotopes, CLM-2250; d_6-succinate, DLM-831; and $^{13}C_{11}$-tryptophan CLM-4290 internal standards) to 100 mg/mL and stored at -80 °C for at least one hour in beadruptor tubes (Fisherbrand; 15-340-154). Samples were then homogenized at 4 °C on a Bead Mill 24 Homogenizer (Fisher; 15-340-163) set at 1.6 m/s with six thirty-second cycles, five seconds off per cycle. Samples were then centrifuged at -10 °C, 20,000 x *g* for 15 min to generate supernatants for subsequent metabolomic analysis.

Derivatization of Culture Supernatants Methods for GC-MS Analyses. Culture supernatants were dried completely under nitrogen stream at 30 ºC (Biotage SPE Dry 96 Dual, 3579M). To dried samples, 50 μL of freshly prepared 20 mg/mL methoxamine (Sigma, 226904) in pyridine (Sigma, 270970) was added and incubated in a ThermoMixer (Eppendorf) for 90 min at 30 °C and 1400 rpm. After samples were cooled to room temperature, 80 μL of derivatizing reagent (N,O-Bis(trimethylsilyl)trifluoroacetamide, 1% Trimethylchlorosilane, Sigma, B-023) and 70 μL of ethyl acetate (Sigma, 439169) were added and mixed in a ThermoMixer at 70 ºC for one hour at 1400 rpm. Samples were cooled to room temperature and diluted with 400 μL of ethyl acetate for analysis.

GC-MS Chromatography and Instrument Parameters. Culture supernatants were analyzed using a GC-MS (Agilent 7890A GC system, Agilent 5975C MS detector) with an electron impact ionization source, a HP-5MSUI column (30 m x 0.25 mm, 0.25 μm, Agilent Technologies 19091S-433UI), and 1 μL injection. Oven ramp parameters: 1 min hold at 60 ºC, 16 ºC per min up to 300 ºC with a 7 min hold at 300 ºC, for a total runtime of 23 minutes. Inlet temperature was 280 ºC, transfer line was 300 ºC, and source temperature was 230 ºC. Ionization was achieved using a 70 eV electron beam. Ions were measured over a 50 to 600 m/z range.

GC-MS Data Analyses. Culture supernatant data analyses were performed using MassHunter Quantitative Analysis software (version B.10, Agilent Technologies). Metabolite identifications were confirmed by matching to authentic standard spectra and retention time and spectra in the NIST Tandem Mass Spectral Library Version 2.3 (see **Table S13** for validation methods of each metabolite). Thus, all metabolites reported from mouse fecal pellet samples and human stool samples are identified at confidence Level 1.[72] Normalized peak areas were calculated by dividing raw peak areas of targeted analytes by averaged raw peak areas of internal standards d7-proline and U-^{13}C-palmitate.

LC-MS Parameters for Shikimate Reduction Analysis. Shikimate reduction products were not readily available as standards, therefore LC-MS/MS fragmentation was used to identify putative products of a shikimate reduction reaction in microbial cultures. All LC-MS analyses were performed using a Thermo Scientific Vanquish Flex UHPLC coupled to an IQ-X mass spectrometer (Thermo Fisher Scientific). To detect shikimate and its putative hydroshikimate product in *Eggerthella lenta* cultures, the chromatographic method used was an isocratic 100% mobile phase A (5% acetonitrile, 0.1% formic acid in water) for 0.2 min, followed by a gradient of 0 to 97% mobile phase B (99.9% acetonitrile, 0.1% formic acid) for 4.5 min with a wash of 100% mobile phase B for 1 min. Shikimate and its putative dehydroshikimate product were detected by mass spectrometry using negative ionization. Flow from the UHPLC was ionized with the heated electrospray ionization (HESI) source set to 2400 V, ion transfer tube temp set to 200 °C, vaporizer temperature set to 400 °C, and sheath, aux and sweep gases set to arbitrary values of 40, 5, and 1, respectively. Data for MS^1 was acquired using a maximum inject time of 50 ms, a normalized AGC target of 25%, a 100–1300 m/z

quadrupole scan range, and a resolution of 60,000. All tandem MS^2 mass spectral data was acquired using a 1.5 m/z quadrupole isolation window, a maximum inject time of 22 ms, a normalized AGC target of 20%, and a resolution of 15,000. Each of the metabolite ions of interest were added to an inclusion list for MS^2 fragmentation by a normalized higher-energy collisional dissociation energy of 30%. A hydroshikimate (3,4,5-trihydroxycyclohexanecarboxylic acid) standard was not readily available, therefore, it is reported at confidence Level 3.[72] To support the assignment of the dehydroshikimate structure, analysis of the MS^2 fragmentation spectrum and a comparison to the shikimate MS^2 fragmentation spectrum has been included with putative fragment ion structures (**Fig. S11**).

LC-MS Chromatography and Instrument Parameters for Fecal Material Analyses. All metabolites from mouse fecal pellet and human stool were measured by LC-MS analyses. Reversed-phase chromatography was performed at a 300 μl min−1 flow rate on a Waters CORTECS T3 C18 RP-UHPLC column (100 × 2.1 mm inner diameter, 1.6 μm particle size, 120 Å pore size (1 Å = 0.1 nm)). Mobile phase A was 5% acetonitrile in water with 0.1% formic acid and mobile phase B was acetonitrile with 0.1% formic acid. For all mouse fecal pellet and human stool analyses, the chromatographic method used was an isocratic 100% mobile phase A for 0.2 min, followed by a gradient of 0 to 40% mobile phase B for 2.8 min, then a gradient from 40% to 100% mobile phase B over 1.5 min, with a final wash of 100% mobile phase B for 1.5 min. To acquire negative ionization mass spectra, flow from the UHPLC was ionized with the heated electrospray ionization (HESI) source set to 3500 V, ion transfer tube temp set to 190 °C, vaporizer temperature set to 120 °C, and sheath, aux and sweep gases set to arbitrary values of 50, 10, and 1, respectively. For positive ionization mass spectra, flow from the UHPLC was ionized with the heated electrospray ionization (HESI) source set to 3500 V, ion transfer tube temp set to 300 °C, vaporizer temperature set to 350 °C, and sheath, aux and sweep gases set to arbitrary values of 50, 10, and 1, respectively. Data for both positive and negative mode MS and MS^2 were acquired using the EASY-IC internal calibration to fluoranthene, automatic maximum inject time, and standard AGC target. Further, data for all MS^1 was acquired using a 100–365 m/z quadrupole scan range and a resolution of 120,000. All tandem MS^2 mass spectral data was acquired using a 0.7 m/z quadrupole isolation window and a resolution of 15,000. Each of the metabolite ions of interest were added to an inclusion list for MS^2 fragmentation by stepped higher-energy collisional dissociation absolute energies of 20, 40, 60, and 80 eV.

LC-MS Data Analysis. Mouse fecal pellet and human stool sample data analysis was performed using FreeStyle software (version 1.8 SP2, Thermo Scientific). Metabolite identification was established by matching accurate mass, retention time, and MS^2 fragmentation pattern from candidate ions in experimental samples to data from authentic standards of each metabolite of interest. Thus, all metabolites reported from mouse fecal pellet samples and human stool samples are identified at confidence Level 1.[72] Normalized peak areas were calculated by dividing raw peak areas of targeted analytes by averaged raw peak areas of internal standards ($^{13}C_9$-phenylalanine and $^{13}C_{11}$-tryptophan for metabolites measured by positive ionization; d_6-succinate and $^{13}C_{11}$-tryptophan for metabolites measured by negative ionization).

GC-MS and LC-MS Quality Control. Biological control plasma samples were processed and analyzed alongside culture supernatant samples to evaluate metabolite extraction efficiency and instrument performance. The samples were extracted with methanol containing known internal standard (IS) concentrations. Recovery, retention time, and coefficient of variation (CV) were calculated for ISs. Method blanks containing no metabolites were included throughout the run.